

XI CONGRESO INTERNACIONAL DE INVESTIGACIÓN EN CIENCIAS ADMINISTRATIVAS  
(ACACIA)

**MODELIZACIÓN DE LA DEMANDA TURÍSTICA EN LA REGIÓN NORTE DE PORTUGAL: APLICACIÓN DE  
LAS REDES NEURONALES**

**MESA DE TRABAJO: FINANZAS Y ECONOMÍA**

Paula Odete Fernandes (responsable); [pof@ipb.pt](mailto:pof@ipb.pt); ESTiG-IPB

Francisco José Terroso Cepeda ([frantece@gmail.com](mailto:frantece@gmail.com))

(Telf.: +351.273.303103; Fax: +351.273.313051)

Departamento de Economia e Gestão

Escola Superior de Tecnologia e de Gestão (ESTiG)

Instituto Politécnico de Bragança (IPB)

Campus de Sta. Apolónia, Apartado 134

5301-857 Bragança, Portugal

Luís Moisés Borge González; [borge@eco.uva.es](mailto:borge@eco.uva.es); Universidad de Valladolid (UVA)

(Telf.: +34.983.423905)

Departamento de Estadística y Econometría

Facultad de Ciências Económicas y Empresariales

Avda. Valle Esgueva, n.º 6

47011 Valladolid, España

ITESO, Universidad Jesuita de Guadalajara.

22-25 Mayo de 2007

## **MODELIZACIÓN DE LA DEMANDA TURÍSTICA EN LA REGIÓN NORTE DE PORTUGAL: APLICACIÓN DE LAS REDES NEURONALES**

### **RESUMEN**

Las Redes Neuronales Artificiales vienen siendo utilizadas en estos últimos años en áreas muy diversas, siendo una de estas áreas la predicción de series temporales. En este trabajo se aplica la metodología de las Redes Neuronales Artificiales con el objetivo de desarrollar un modelo que pueden ser susceptible de ser utilizado como instrumento de predicción de la demanda turística.

En el análisis, se constató que la red que ha presentado más estabilidad, mayor consistencia a lo largo del análisis y ha evidenciado ser adecuada para la modelización y predicción de la serie temporal en estudio, fue entrenada con el algoritmo de *Levenberg-Marquardt*, la función de activación tangente hiperbólica, tiene diez nodos en la capa oculta y cada valor de la serie depende directamente de los doce valores precedentes.

Las "*Pernoctaciones Mensuales en la Región Norte de Portugal*", registradas en el período de Enero de 1987 a Diciembre de 2003, fue la serie elegida para hacer todo el estudio.

**Palabras Clave:** Redes Neuronales Artificiales, Arquitectura, Entrenamiento y Algoritmo *Backpropagation*.

## 1. INTRODUCCIÓN

Las redes neuronales constituyen una metodología con la que se busca emular una de las características propias de los humanos: la capacidad de memorizar y asociar hechos. El hombre es capaz de resolver ciertas situaciones acudiendo a la experiencia acumulada. Una forma de aproximarse a cierto tipo de problemas consiste en la construcción de sistemas que sean capaces de reproducir esta característica humana. Las redes neuronales no son más que un modelo artificial y simplificado del cerebro humano, que es el ejemplo más perfecto del que disponemos, capaz de adquirir conocimiento a través de la experiencia.

Desde el punto de vista de las aplicaciones prácticas de las Redes Neuronales Artificiales (RNA), su principal ventaja frente a otras técnicas reside en el procesado paralelo, adaptativo y no lineal. Actualmente, las redes neuronales son un campo muy consolidado en el que continuamente se desarrollan nuevos métodos y aplicaciones.

Las redes neuronales han demostrado su capacidad para resolver problemas complejos en diversas áreas. Se han desarrollado aplicaciones de RNA para fines tan variados como visión artificial, tratamiento de señales e imágenes, reconocimiento de caracteres manuscritos y de voz, sistemas expertos, análisis de imágenes médicas, control remoto, control de robots, inspección industrial, predicciones económicas y problemas de gestión, exploración científica, entre otras (Zafra *et al.*, 1993; Sohl & Venkatachalam, 1995; Schalkoff, 1997; Haykin, 1999; Shin & Park, 1999; Alon *et al.*, 2001).

Así, en este trabajo, hacemos una somera descripción de la metodología aplicada, RNA, y a continuación se realizará un análisis temporal de referencia, recurriendo al conocimiento de sus valores en el pasado. Este análisis tiene como objetivo principal obtener un modelo que presente más estabilidad y mayor consistencia a lo largo del análisis y que evidencie ser adecuado para la modelización y predicción de la serie temporal en estudio.

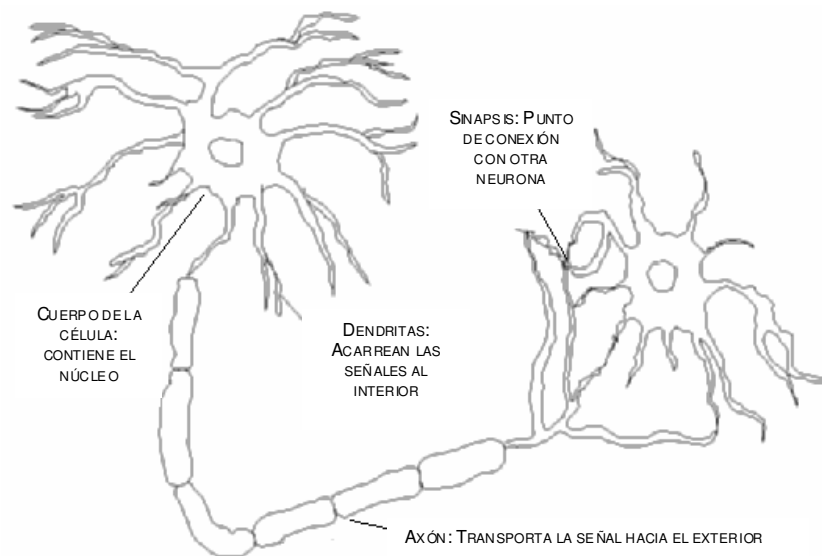
## 2. INSPIRACIÓN BIOLÓGICA: EL CEREBRO HUMANO

La modelización y la teoría de las redes neuronales están inspirados en la estructura y funcionamiento del sistema nervioso humano, dónde la neurona es el elemento fundamental.

Se calcula que el cerebro humano contiene aproximadamente 12 billones de células nerviosas o **neuronas**. Una de las características de las neuronas es su capacidad de comunicarse. Cada neurona tiene de 5.600 a 60.000 **conexiones dendríticas** provenientes de otras neuronas

(véase Figura 1). Estas conexiones transportan los impulsos enviados desde otras neuronas y están conectadas a la membrana de la neurona. Cada neurona tiene una salida denominada **axón**. El contacto de cada axón con una dendrita se realiza a través de la **sinapsis**. Tanto el axón como las dendritas transmiten **la señal** en una única dirección. Se puede decir que el cerebro humano se compone de decenas de billones de neuronas interconectadas entre sí formando circuitos o redes que desarrollan funciones específicas (Zhang *et al.*, 1998; Haykin, 1999; Marques, 1999; Basheer & Hajmeer, 2000; Rodrigues, 2000).

**Figura 1:** Neurona y Conexiones Sinápticas.



**Fuente:** Adaptado de SCHALKOFF; (1997:64).

En resumen, la información se transmite de unas neuronas a otras y va siendo procesada a través de las conexiones sinápticas y las propias neuronas. El aprendizaje de las redes neuronales se produce mediante la variación de la efectividad de las sinapsis, de esta manera cambia la influencia que unas neuronas ejercen sobre otras. De esto se deduce que la arquitectura, el tipo y la efectividad de las conexiones en un momento dado, representan en cierto modo la memoria o estado de conocimiento de la red (Vicente, 1993; Church & Curram, 1996; Balkin & Ord, 2000).

### 3. O QUE SON REDES NEURONALES?

Aunque aún es escaso el conocimiento que se tiene sobre la forma en que el cerebro aprende a procesar la información, los científicos han comenzado a desarrollar modelos matemáticos que tratan de mimetizar tales habilidades y que se denominan *redes neuronales artificiales*. La

elaboración de estos modelos supone, en primer lugar, la deducción de los rasgos o características esenciales de las neuronas (o nodos) y sus conexiones y, en segundo lugar, la implementación del modelo en un ordenador de forma que se pueda simular.

Las RNA, son sistemas paralelos para el procesamiento de la información. Están formadas por un conjunto de unidades de procesamiento conectadas entre sí y, por analogía con el cerebro humano, se denomina «neurona» a cada una de estas unidades de procesamiento. Cada neurona se modeliza mediante unidades de proceso y se compone de un núcleo encargado de los cálculos, una red o vector de conexiones de entrada y una salida (Figura 2) (Haykin, 1999; Draghici, 2002). Para establecer una similitud directa entre la actividad sináptica y las RNA se puede considerar como sus elementos principales (Basheer & Hajmeer, 2000; Rodrigues, 2000):

- 1 Función de red. Calcula el valor de base o entrada total a la unidad, generalmente, como una simple suma ponderada de todas las entradas recibidas. Esto es, las entradas multiplicadas por un peso o valor de las conexiones. Equivale a la combinación de las señales excitatorias e inhibitorias de las neuronas biológicas;
- 2 Función de activación. Se encarga de calcular el nivel o estado de activación de la neurona en función de la entrada total. Es la característica principal o determinante de las neuronas, la que mejor define el comportamiento de la misma. Se utilizan diferentes tipos de funciones, desde funciones lineal<sup>1</sup>, umbral<sup>2</sup> a funciones no lineales (la más común es la función sigmoideal<sup>3</sup>: función logística; tangente hiperbólica);
- 3 Conexiones ponderadas. Son equivalente al papel de las conexiones sinápticas. El peso de la conexión equivale a la fuerza o efectividad de la sinapsis. La existencia de conexiones establece si es posible que una unidad influya sobre otra. El valor de los pesos y el signo de los mismos definen la intensidad de la influencia y el tipo (excitatorio/inhibitorio);
- 4 Salida. El valor de salida cumple la función de la tasa de disparo en las neuronas biológicas. Se calcula en función de la activación de la misma, aunque habitualmente no se aplica más que la función identidad, y se toma como salida el valor de activación.

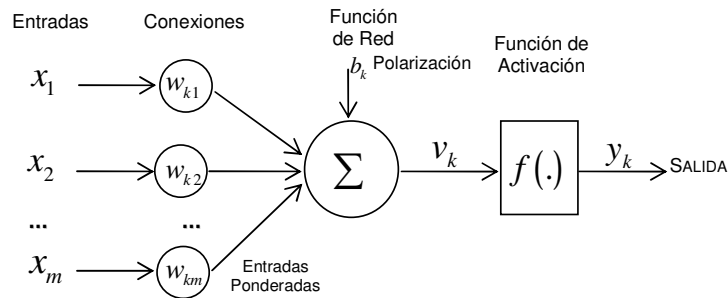
---

<sup>1</sup> La actividad de salida es proporcional a la entrada ponderada total (García & Diez; 1995).

<sup>2</sup> La salida queda fija a uno de dos niveles, dependiendo si la entrada ponderada total es mayor o menor que cierto valor crítico denominado "umbral" (García & Diez; 1995).

<sup>3</sup> La salida varía de forma continua dependiendo de la entrada ponderada total, pero esta dependencia no es lineal (García & Diez; 1995).

**Figura 2: Modelo de Una Neurona.**



**Fuente:** Adaptado de HAYKIN; (1999:11).

El modelo presentado en la Figura 2, incluye también un valor constante denominado polarización (*bias*) aplicado externamente, y representado por  $b_k$ . Este elemento tiene el efecto de aumentar o disminuir la entrada neta de la función de activación dependiendo si es positivo o negativo, respectivamente. En este modelo, cada neurona consta de un conjunto de entradas,  $x_j$ , y una sola salida  $y_k$ . Cada entrada  $j$  está afectada por un coeficiente que se denomina peso y que se representa por  $w_{kj}$ . El primero subíndice,  $k$ , se refiere a la neurona en cuestión y el segundo,  $j$ , se refiere al terminal de entrada de la sinapsis. El peso sináptico de una neurona artificial, a diferencia de una sinapsis del cerebro, varía en un intervalo que incluye tanto valores negativos como positivos. La salida,  $y_k$ , de la neurona es el resultado de la activación de ésta (Haykin, 1999).

La descripción de una neurona  $k$  viene dada por el siguiente par de ecuaciones (Haykin, 1999):

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad [1]$$

y

$$y_k = f(u_k + b_k) \quad [2]$$

siendo  $x_1, x_2, \dots, x_m$  las señales de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  los pesos sinápticos de la neurona  $k$ ,  $u_k$  es la salida de la combinación lineal de las señales de entrada,  $b_k$  es el elemento de polarización,  $f(\cdot)$  es la función de activación e  $y_k$  es la señal de salida de la neurona. La utilización del elemento de polarización permite ajustar el valor producido por las señales de entrada al valor de transición de la función de activación, por lo que la activación de la neurona viene dada por la ecuación:

$$v_k = u_k + b_k \quad [3]$$

Por simplificación matemática y mediante la combinación de las ecuaciones [1] a la [3], se tiene:

$$y_k = f(v_k) \quad [4]$$

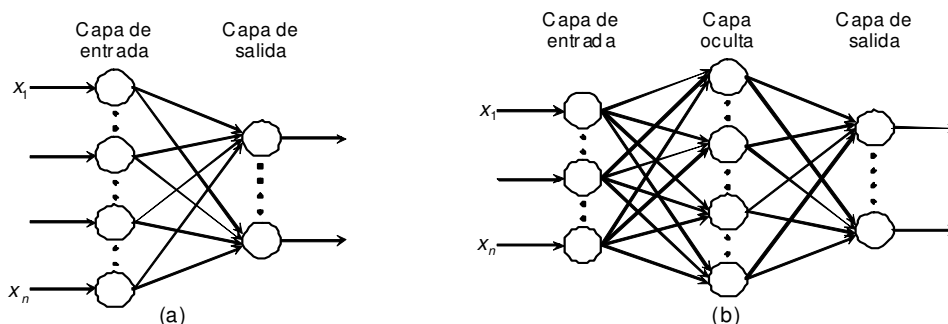
equivalente a:

$$y_k = f\left(\sum_{j=1}^m w_{kj}x_j + b_k\right) \quad [5]$$

El proceso de aplicación de las RNA a un problema en concreto, consta de tres etapas fundamentales:

**1ª Arquitectura de la red:** definir la organización jerárquica de los elementos de procesamiento (neuronas), generalmente organizados en grupos denominados «capas» o «niveles». Para diseñar una red hay que establecer como estarán conectadas unas unidades con otras. Lo habitual es disponer las unidades en forma de capas, pudiéndose hablar de redes de una, de dos o más de dos capas. Se conoce como capa o nivel a un conjunto de neuronas cuyas entradas provienen de la misma fuente y cuyas salidas se dirigen al mismo destino. En cualquier red neuronal existen dos capas típicas y que tienen conexión con el exterior, son la **capa de entrada** a través de la cual los datos son presentados a la red, almacenando la información suministrada a la red, bruta o tras haber realizado un sencillo preproceso de la misma, y la **capa de salida** en la que está almacenada la respuesta dada por la red a una entrada, y puede ser utilizada para los fines que sean precisos (Figura 3a). El resto de las capas existentes entre estas dos capas, capas intermedias encargadas de extraer, procesar y memorizar la información, se denominan **capas ocultas** (García *et al.*, 1996; Silvério, 2000), ver Figura 3b.

**Figura 3:** Perceptron de Una Capa (a). Perceptron Multicapa (b).



**Fuente:** Adaptado de HAYKIN; (1999:13).

Las redes, dependiendo de cómo están conectadas las diferentes capas, se pueden clasificar en redes recurrentes (*feedback*) y redes no recurrentes (*feedforward*). En las primeras, las conexiones permiten que la información pueda volver a lugares por los que ya había pasado, y se admiten las conexiones intra-capas (laterales), permitiendo conexiones de una unidad consigo misma. En las segundas, la información fluye de forma unidireccional de una capa a otra, desde la capa de entrada a las capas ocultas y de éstas a la capa de salida.

**2ª Entrenamiento de la RNA:** el entrenamiento de una red neuronal consiste en la utilización de un algoritmo que ajusta los pesos sinápticos de las conexiones entre las neuronas. Desde sus comienzos, han sido desarrollados varios algoritmos para entrenar una red neuronal. El más utilizado es el algoritmo de **retropropagación (*backpropagation*)**, no sólo por su tradición, sino, esencialmente, por su facilidad de implementación, versatilidad y relativa eficiencia computacional (Kuan & White, 1994; Nam & Schaefer, 1995; Yao *et al.*, 2000). Se han desarrollado un gran número de variaciones del algoritmo de retropropagación del error, cuyo objetivo es mejorar la velocidad y la precisión de los resultados esperados.

En el caso que nos ocupa y para el método de aprendizaje, entre las diferentes variaciones de los algoritmos de entrenamiento por retropropagación del error (*Backpropagation*), se optó por los algoritmos de *Levenberg-Marquardt* [LM] y de *Resilient Backpropagation* [RP]. La elección se debe al hecho de que estos son aconsejables, por su rapidez, para redes de una dimensión intermedia. El [LM], por comportarse mejor para algunos centenares de pesos y el [RP] por conseguir obtener mejores resultados para una gran variedad de problemas.

El entrenamiento de una red usando el algoritmo retropropagación necesita un vector de entrada, un vector de salida y una función del error que es la que debe de minimizarse. Esta función, generalmente, para medir el error entre las salidas de la red y las salidas deseadas, es el error cuadrático medio (Haykin, 1999):

$$E(t) = \frac{1}{2N_p} \sum_{p=1}^{N_p} \sum_{j=1}^{N_n} (d_j^p - y_j^p(t))^2 \quad [6]$$

donde:

$p$ , representa el patrón de la entrada asociado a un determinado ejemplo de entrenamiento;

$j$ , representa una neurona de la capa de salida;

$d_j^p$ , salida deseada de la neurona  $j$ , cuando el patrón de entrada presentado en la red es  $p$ ;

$y_j^p(t)$ , salida de la neurona  $j$  cuando, en el ciclo  $t$ , con patrón de entrada  $p$ ;

$N_j$ , número de neuronas de la capa de salida;

$N_p$ , número de ejemplos de entrenamiento;

$E(t)$ , error de la red en el momento  $t$ .

**3ª Utilización de la Red Neuronal en Modo Recuerdo:** una vez entrenada, la red se encuentra en condiciones de ser utilizada. Para ello, no hay más que presentarle un caso concreto, por ejemplo las últimas observaciones disponibles de una serie de tiempo, para que, utilizando los pesos sinápticos encontrados durante el proceso de entrenamiento, calcule la salida - predicción del dato siguiente de la serie de tiempo.

En el proceso de análisis de los datos con la metodología de las RNA, y con el fin de medir la capacidad de reproducir nuevos patrones del problema no utilizados en el proceso de aprendizaje, es habitual dividir los datos disponibles en tres partes, una destinada al entrenamiento de la red, otra a la validación y otra para efectuar el test. Este último conjunto no es presentado a la red más que para llevar a cabo una validación cruzada, entre los resultados de una red entrenada y el mencionado conjunto.

Esta forma de operar plantea el problema de elegir el número de observaciones que van a integrar cada uno de los mencionados conjuntos. No existe un criterio concluyente de división de la muestra. Se hace necesario establecer un compromiso entre la complejidad del modelo y el número de datos disponible, de forma que el modelo sea más simple que el conjunto de datos que representa. En caso contrario, el modelo sólo memoriza la información no siendo capaz de generalizar lo aprendido a nuevas situaciones (Marques, 1999; Anders & Korn, 1999).

## **4. MODELIZACIÓN DE LA DEMANDA TURÍSTICA EN LA REGIÓN NORTE DE PORTUGAL**

### **4.1. ESTRUCTURA DE LA RED NEURONAL**

A lo largo del análisis, que a continuación se presenta, la serie temporal "*Pernoctaciones Mensuales de Turistas en la Región Norte de Portugal*" pasará a designarse por SO (Serie Original), obtenida en el Instituto Nacional de Estadística (INE) Portugués y referente al período comprendido entre Enero de 1987 y Diciembre de 2003. Los datos están recogidos en el Anexo A -Tabla A.1.

El análisis se lleva a cabo con dos series identificadas como serie original (SO), serie original en el dominio logarítmico (SO\_Ln). Todas ellas han sido sometidas a todo el análisis, habiendo

sido las estimaciones efectuadas bajo dos procesos distintos. La estimación se realizó mes a mes (*one-step-ahead forecast*), ver Tabla 1.

En el proceso, los datos quedan divididos en tres conjuntos: entrenamiento, validación y conjunto de test. Cuando se habla de valores estimados, éstos hacen referencia a los conjuntos de entrenamiento y validación; por otra parte, cuando se habla de predicciones, éstas se refieren al conjunto de test, puesto que los valores históricos de este conjunto nunca se dan a conocer a la red en las etapas de entrenamiento y validación.

**Tabla 1:** Series sometidas al estudio de las RNA.

DESCRIPCIÓN	DENOMINACIÓN DE LA SERIE
Serie con los valores originales	SO
Aplicación de Logaritmos a la serie SO	SO_Ln

La opción elegida es la de una red de tipo perceptrones multicapa y *feedforward*. Esta elección se fundamenta en los trabajos efectuados por los autores Nam y Schaefer (1995), Faraway y Chatfield (1998) y Thawornwong y Enke (2004). En el presente trabajo se han utilizado RNA de tres capas: una de entrada, una oculta y una de salida. En la capa oculta, se han experimentado alternativamente la función de activación sigmoidea logística [*Logsig*] y la función tangente hiperbólica [*Tansig*]. En la capa de salida se utilizó siempre la función activación lineal [*Lin*], por ser la que mejores resultados proporciona en este tipo de arquitecturas. Dado que siempre se utilizó la misma función de activación para la salida, el análisis a lo largo del estudio nunca recaerá sobre ella sino sobre las anteriores.

En la Tabla 2, se presentan las diferentes arquitecturas estudiadas, observándose organizaciones diferentes, tanto al nivel de la capa de entrada como en la capa oculta. Todas las redes tienen una capa oculta con la variación del número de nodos.

Dada la diversidad de redes propuestas en la literatura, alguna se revela de más utilidad en un estudio que aquellas que se cree, a priori, más apropiadas. Por ello se experimentaron un total de 19 arquitecturas, correspondiendo a la estimación mensual (Tabla 2).

Puesto que se pretende efectuar estimaciones mensuales, se optó por combinar los siguientes supuestos: considerar como retardos de entrada las observaciones anteriores y próximas del

mes que se desea estimar y como la serie presenta un comportamiento estacional, incluir un período de doce meses.

También se presenta el número de parámetros (o de pesos) para cada una de las arquitecturas, calculados mediante la siguiente expresión:

$$N.^{\circ} \text{ Parámetros} = [NR + (NS + 1)] \times NO + NS$$

donde,

$NR$  : N.º de retardos (nodos de entrada);

$NS$  : N.º de nodos en la capa de salida;

$NO$  : N.º de nodos en la capa oculta.

**Tabla 2:** Arquitecturas de las diferentes RNA.

ESTIMACIÓN MENSUAL					
NODOS EN LA CAPA DE ENTRADA	NODOS EN LA CAPA OCULTA	Nº DE PARÁMETROS	NODOS EN LA CAPA DE ENTRADA	NODOS EN LA CAPA OCULTA	Nº DE PARÁMETROS
1,12,13	1	6	1,12,13	4	21
1,12	2	9	1,2,12,13	4	25
1,2,12	2	11	1-12	4	57
1,12,13	2	11	1-13	4	61
1,2,12,13	2	13	1-12	6	85
1-4	2	13	1-13	6	91
1-12	2	29	1,12	10	41
1-13	2	31	1-12	10	141
1,12	4	17	1-13	10	151
1,2,12	4	21			

A lo largo del estudio de las RNA, cuando se esté hablando de la arquitectura de una red, se va a utilizar la siguiente notación: ( $E$ ;  $O$ ;  $S$ ). Donde ' $E$ ' representa los retardos en la capa de entrada, ' $O$ ' el número de nodos de la capa oculta y ' $S$ ' el número de nodos de la capa de salida.

#### 4.2. MÉTODO DE ENTRENAMIENTO DE LAS RNA

El aprendizaje para cada patrón presentado en la entrada de la red es supervisado con la información de cuál es la salida deseada. De este modo, el entrenamiento se realiza a través de un conjunto de datos de entrada que son los que corresponden a los retardos referidos en cada arquitectura de la que se conoce el conjunto de salida.

Los pesos iniciales se han generado de forma aleatoria con valores pequeños. Estos pesos, durante el entrenamiento, han sido ajustados de acuerdo a los algoritmos de aprendizaje, con el objetivo de minimizar el error (diferencia entre el valor deseado y el valor calculado). La actualización de los pesos se hizo una vez presentados todos los elementos del conjunto del entrenamiento, es decir, *batch training*.

Para las diferentes arquitecturas identificadas en la Tabla 2, se han utilizado las funciones de activación [*Logsig*] y [*Tansig*] en la capa oculta. Cada una de las situaciones anteriores ha sido entrenada con los algoritmos *Levenberg-Marquardt* y *Resilient Backpropagation*, en un total de 152 situaciones distintas,

$$\left[ 19(\text{arquitecturas}) \times 2(\text{funciones de activación}) \times 2(\text{algoritmos de entrenamiento}) \times 2(\text{series}) \right].$$

Puesto que un conjunto diferente inicial de pesos conduce normalmente a una solución final también diferente, aunque muy próxima, se han realizado diversas sesiones de entrenamiento para cada una de las situaciones antes identificadas con pesos iniciales diferentes. De este conjunto de sesiones se retuvo la red neuronal con mejores resultados estimados de cada una de las situaciones. En concreto, para las 19 arquitecturas de las redes neuronales mensuales se han realizado 250 sesiones de entrenamiento.

Los criterios que fueran utilizados para elegir el mejor resultado de entre las distintas sesiones, fueran el coeficiente de correlación<sup>4</sup> y Raíz cuadrada del error cuadrático medio (*RMSE*<sup>5</sup>). Se optó por este último, para elegir los mejores modelos/RNA.

A lo largo del proceso de entrenamiento, para cada una de las sesiones, es necesario establecer el número de ciclos y de fijar un objetivo. El error de los resultados obtenidos por las redes en el presente estudio, que se ha fijado como deseable, debe de aproximarse a  $1 \times 10^{-4}$ .

Se adoptó una situación de parada anticipada de la fase de entrenamiento, lo que asegura una identificación más rápida de la arquitectura óptima. Posteriormente, se recurre a un segundo conjunto, el de validación que no ha sido utilizado en el entrenamiento, interrumpiendo el entrenamiento cuando el error medido para este conjunto aumenta en varias iteraciones

4

$$r_{A,P} = \frac{\sum_{t=1}^n (A_t - \bar{A})(P_t - \bar{P})}{\sqrt{\sum_{t=1}^n (A_t - \bar{A})^2 (P_t - \bar{P})^2}}; \text{con : } A, \text{ valores reales en el periodo } t; P, \text{ valores pronosticados en el periodo } t; n, \text{ total de observaciones utilizadas.}$$

$$5 \text{ } RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - P_t)^2}{n}}; \text{con : } A, \text{ valores reales en el periodo } t; P, \text{ valores pronosticados en el periodo } t; n, \text{ total de observaciones utilizadas.}$$

sucesivas. Finalmente, se ha usado un tercer conjunto, el conjunto de test, que permite evaluar si los conjuntos de entrenamiento y validación son representativos del universo. Tanto el conjunto de validación como el de test pueden ser usados en la validación de los resultados obtenidos por la red, ya que no entran en el entrenamiento de la misma. Esta operación permite evitar *overfitting*<sup>6</sup>.

La función del error minimizado en el proceso de entrenamiento de las redes neuronales, ha sido la función del error cuadrático medio.

#### **4.3. CONJUNTOS DE ENTRENAMIENTO, VALIDACIÓN Y TEST**

Cada serie ha sido dividida en los conjuntos de entrenamiento, de validación y de test según se refleja a continuación, para el modelo mensual.

La serie original y la serie original en el dominio logarítmico, tienen el mismo número de observaciones, correspondientes al período comprendido entre Enero de 1987 y Diciembre de 2003, en total 204. Estas series se han dividido en el conjunto de entrenamiento, con 180 observaciones correspondientes al periodo que va desde Enero de 1987 hasta Diciembre de 2001. En el conjunto de validación, se han utilizado las 12 observaciones correspondientes al periodo que va desde Enero de 2002 hasta Diciembre de 2002. Finalmente, el conjunto de test vuelve a tener 12 observaciones correspondientes al año 2003. De las 180 observaciones del conjunto de entrenamiento, hay que retirar los meses iniciales correspondientes a los retardos de la entrada de la red neuronal, para cada una de las arquitecturas.

#### **4.4. SELECCIÓN DE LA MEJOR RED**

A continuación vamos examinar el comportamiento de las diferentes redes sobre el conjunto de test. Para ello, vamos a comparar las predicciones que se obtienen en cada análisis con los datos observados del conjunto test que corresponde con las observaciones del año 2003. Puesto que el conjunto de test está siempre separado del conjunto de aprendizaje y del conjunto de validación, los resultados que se obtienen con cada modelo y que se comparan con los datos del test, miden la capacidad de predicción de cada uno de los modelos. Este análisis comparativo entre los distintos modelos seleccionados se hace en todos los casos entre los

---

<sup>6</sup> Puede corresponder a que la red se moldee muy bien para el conjunto utilizado para el entrenamiento, pero al presentar otros vectores de entrada, no es capaz de dar buenas respuestas.

datos de la serie original y los correspondientes valores estimados de cada modelo. El criterio utilizado para tal fin es el de hacer mínimo el *RMSE* para el conjunto de test.

Así, para cada uno de los algoritmos y cada una de las funciones de activación, para las distintas series, se obtuvieron los resultados que se resumen en la Tabla 3 sobre el conjunto del test. Como ya se dijo, aunque en el conjunto del test se obtengan tanto el coeficiente de correlación como el *RMSE*, el criterio utilizado para medir el desempeño del modelo es el del mínimo *RMSE*.

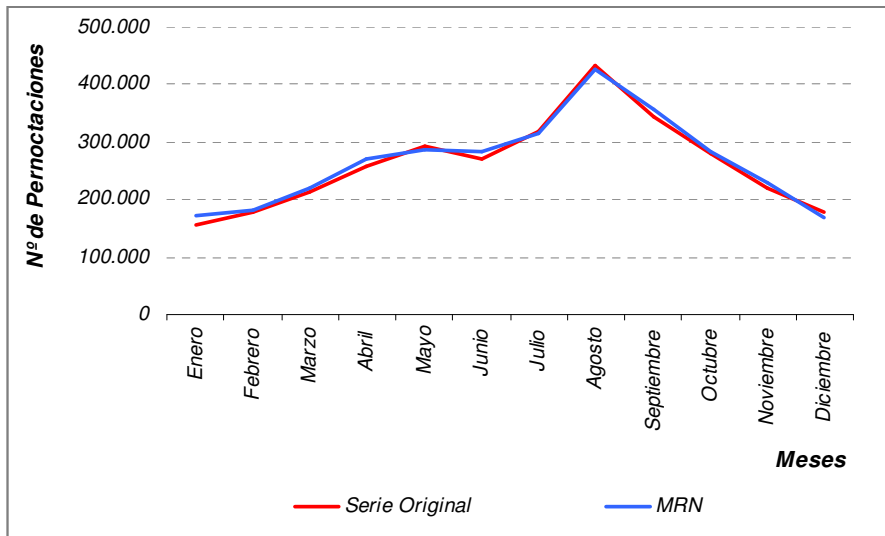
**Tabla 3:** Resumen del comportamiento de las redes para el modelo.

SERIES	ARQUITECTURA	Nº DE PARÁMETROS	ALGORITMO DE ENTRENAMIENTO	FUNCIÓN DE ACTIVACIÓN	INDICADORES DE SELECCIÓN	
					<i>r</i>	<i>RMSE</i>
SO	(1,2,12,13; 4; 1)	25	LM	Logsig	0,990	10.777
	<b>(1-12; 10; 1)</b>	<b>141</b>		<b>Tansig</b>	<b>0,994</b>	<b>9.246</b>
	(1,2,12,13; 4; 1)	25	RP	Logsig	0,989	11.405
	(1,2,12,13; 4; 1)	25		Tansig	0,991	11.298
SO_Ln	(1-12; 10; 1)	141	LM	Logsig	0,992	10.060
	(1-13; 6; 1)	91		Tansig	0,996	9.377
	(1,2,12; 4; 1)	21	RP	Logsig	0,987	12.664
	(1-13; 10; 1)	151		Tansig	0,992	10.913

De entre las dos series, la serie SO es dónde se registró el valor más bajo para el *RMSE* (9.246) y obtuvo el coeficiente de correlación del 99,4%. Esta red ha sido entrenada con el algoritmo [LM] y la función de activación [Tansig] y tiene diez nodos en la capa oculta. Cada valor de la serie depende directamente de los doce valores precedentes. A partir de esta etapa esta red se designará **MRN** (Modelo de la Red Neuronal).

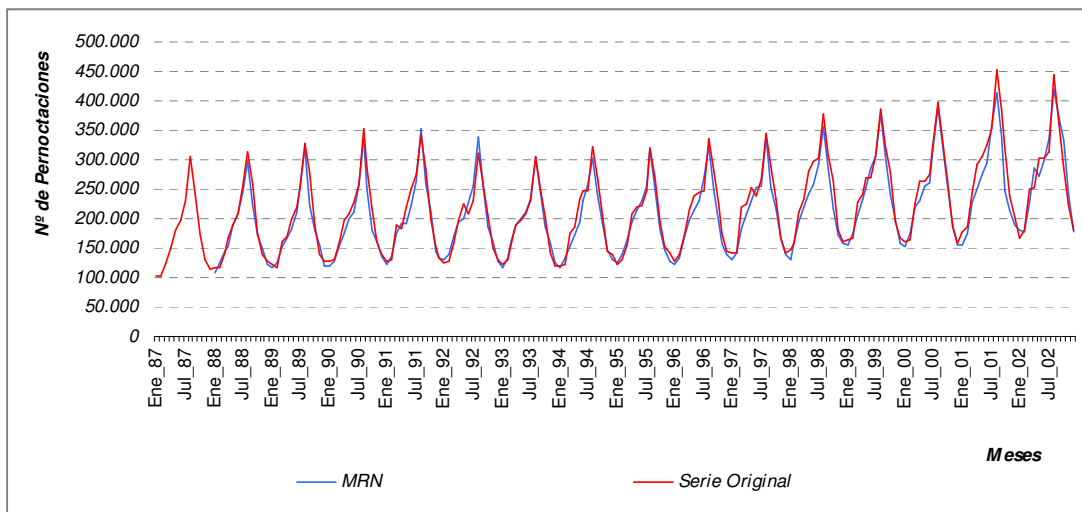
Los valores predichos para el año 2003, para el modelo seleccionado de las pernoctaciones de los turistas en la Región Norte de Portugal, se presentan en la Tabla 4 y su representación gráfica en la Figura 4. Este grafico permite un fácil análisis y comprensión de los resultados producidos. Es notable la excelente aproximación entre los valores pronosticados y los valores observados, pues las líneas casi se sobreponen, lo que significa que la red ha aprendido y hubo memorización de la información.

**Figura 4:** Comparación del MRN con la Serie Original, en el año 2003.



Debido a estos buenos resultados del modelo identificado en el conjunto de test, y con el objeto de resaltar más las diferencias entre los datos obtenidos, se analiza con más detalle su comportamiento, comprobando, en esta ocasión, los resultados del modelo para los conjuntos de entrenamiento y validación junto con la serie original. Los resultados se ven ilustrados en la Figura 5 y vienen recogidos en la Tabla A.2 del Anexo A. Como cabría esperar y dado que la optimización del criterio *RMSE* había sido el utilizado para el conjunto del test y no para el de entrenamiento, el comportamiento del modelo para el conjunto de entrenamiento es inferior al del conjunto de test.

**Figura 5:** Comparación del comportamiento del modelo MRN frente a la Serie Original, en los conjuntos de entrenamiento y validación.



Decir que el modelo MRN es el mejor modelo existente puede ser una afirmación un tanto arriesgada. Sin embargo, sí es razonable admitir que ha sido la arquitectura (1-12;10;1) que ha presentado más estabilidad y mayor consistencia a lo largo del análisis a que ha sido sometida y ha evidenciado ser adecuada para la modelización y predicción de la serie temporal en estudio, Pernoctaciones de Turistas en la Región Norte de Portugal. Así, las predicciones vienen dadas por la ecuación [7].

$$Y_t = b_{2,1} + \sum_{j=1}^n \alpha_j f \left( \sum_{i=1}^m \beta_{ij} y_{t-i} + b_{1,j} \right) \quad [7]$$

donde,

$m$ , número de retardos de entrada;

$n$ , número de nodos en la capa oculta;

$f$ , función de activación sigmoidea *tansig*:  $f(v) = \frac{e^{(v)} - e^{(-v)}}{e^{(v)} + e^{(-v)}}$ ;

$\{\alpha_j, j = 0, 1, \dots, n\}$ , vector de pesos que relacionan los nodos de la capa oculta a los nodos de la capa de salida;

$\{\beta_{ij}, i = 0, 1, \dots, m; j = 1, 2, \dots, n\}$ , pesos que relacionan los nodos de entrada a los nodos de la capa oculta;

$b_{2,1}$  y  $b_{1,j}$ , indican los pesos de los términos independientes (*bias*) asociados a cada nodo de la capa de salida y de la capa oculta, respectivamente. La ecuación indica también la utilización de una función de activación lineal en la capa de salida.

La Tabla 4 presenta las predicciones y los errores relativos absolutos de las predicciones mensuales, en porcentajes, para el año 2003. De acuerdo con la clasificación por Gaw y Law (2002), citando a Lewis, donde señalan que cuando: los errores relativos absolutos de las predicciones sean inferiores a un 10%, se considera una 'precisión elevada'; entre 10 y 20%, se considera 'buena'; entre 20 y 50% se considera 'razonable'; y superiores a 50%, 'imprecisas/poco fiables'; se puede decir que el modelo, para cualquier de los meses, generan predicciones de elevada precisión. Sin embargo, puede decirse que las predicciones pueden considerarse como fiables, dado que se registra una aplastadora mayoría de los errores relativos absolutos inferiores a 10%.

**Tabla 4:** Valores Previstos y Reales de las Pernoctaciones y el Error relativo absoluto de las predicciones, en %, para el período de 2003:01 a 2003:12.

Meses	Serie Original	MRN	Error Relativo Absoluto*
Enero	155.527	171.394	10,20%
Febrero	177.818	180.351	1,42%
Marzo	214.106	220.282	2,88%
Abril	258.519	269.243	4,15%
Mayo	293.531	286.220	2,49%
Junio	271.454	282.712	4,15%
Julio	318.706	314.662	1,27%
Agosto	433.211	427.638	1,29%
Septiembre	343.534	357.190	3,98%
Octubre	281.472	282.544	0,38%
Noviembre	219.463	228.324	4,04%
Diciembre	178.439	167.448	6,16%

Nota:  $*_{ERA} = \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100.$

## 5. CONCLUSIÓN

El modelo, no lineal, obtenido con la metodología empleada permitió evidenciar que las estimaciones/predicciones se ajustan bastante a los datos reales y que en la mayoría de los casos las estimaciones/predicciones detectaron perfectamente el comportamiento que presenta la serie original.

La arquitectura (1-12;10;1) fue la que presentó más estabilidad y mayor consistencia a lo largo del análisis a que ha sido sometida y ha evidenciado ser adecuada para la modelización y predicción de la serie temporal en estudio, Pernoctaciones de Turistas en la Región Norte de Portugal. Esta serie temporal, fue entrenada con el algoritmo de *Levenberg-Marquardt*, la función de activación tangente hiperbólica, tiene diez nodos en la capa oculta y cada valor de la serie depende directamente de los doce valores precedentes.

Por último, señalar que la precisión de las predicciones no depende únicamente de la óptima elección de un modelo, sino que también es consecuencia de las peculiares características de cada serie, su mayor o menor grado de complicación junto a las posibles distorsiones ocasionales o erráticas que puedan presentarse.

## BIBLIOGRAFÍA

- ALON**, Ilan; **QI**, Min. and **SADOWSKI**, Robert J.; (2001); "Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods"; *Journal of Retailing and Consumer Services*; N.º8, pp.147/156.
- ANDERS**, Ulrich and **KORN**, Olaf; (1999); "Model selection in neural networks"; *Neural Networks*; N.º12, pp.309/323.
- BALKIN**, Sandy D. and **ORD**, J. Keith; (2001); "Automatic neural network modelling univariate time series"; *International Journal of Forecasting*; N.º16, pp.509/515.
- BASHEER**, I.A. and **HAJMEER**, M.; (2000); "Artificial Neural Networks: fundamentals, computing, design and application"; *Journal of Microbiological Methods*; N.º43, pp.3/31.
- CHURCH**, Keith B. and **CURRAM**, Stephen P.; (1996); "Forecasting consumers' expenditure: A comparison between econometric and neural network models"; *International Journal of Forecasting*; N.º12, pp.255/267.
- DRAGHICI**, Sorin; (2002); "On the capabilities of neural networks using limited precision weights"; *Neural Networks*; N.º15, pp.395/414.
- FARAWAY**, Julian and **CHATFIELD**, Chris; (1998); "Time series forecasting with neural networks: a comparative study using the airline data"; *Applied Statistics*; N.º47, pp.231/250.
- GARCÍA**, David y **DÍEZ**, Raúl Pino; (1995); "Análisis comparativo de cálculo de previsiones univariadas y función de transferencia, mediante las metodologías de Box-Jenkins y Redes Neuronales"; *Quesito Journal*; Vol. 19; pp.187/215.
- GARCÍA**, David; **DÍEZ**, Raúl Pino; **RIESTRA**, Cristina S. y **RODRÍGUEZ**, José Luis Mayo; (1996); "Análisis comparativo de los métodos de previsión univariante, Box-Jenkins, redes neuronales artificiales y espacios de estado"; *Estudios de Economía Aplicada*; N.º5, pp.5/33.
- GOH**, Carey and **LAW**, Rob; (2002); "Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention"; *Tourism Management*; N.º23, pp.499/510.
- HAYKIN**, Simon; 1999; "*Neural Networks. A comprehensive foundation*"; New Jersey, Prentice Hall.
- KUAN**, Chung-Ming and **WHITE**, Halbert; (1994); "Artificial Neural Network: An econometric perspective"; *Econometric Reviews*; N.º13, pp.1/91.
- MARQUES**, Jorge Salvador; (1999); "*Reconhecimento de padrões. Métodos Estatísticos e Neurais*"; Lisboa, Instituto Superior Técnico Press.
- NAM**, Kyungdoo and **SCHAEFER**, Thomas; (1995); "Forecasting International Airline Passenger Traffic Using Neural Networks"; *Logistics and Transportation Review*; Volume 31; N.º3, pp.239/251.
- RODRIGUES**, Pedro João S.; (2000); "*Redes neuronais aplicadas à segmentação e classificação de leucócitos em imagens*"; Dissertação de Mestrado em Engenharia Electrónica e Telecomunicações; Universidade de Aveiro.
- SCHALKOFF**, Robert J.; (1997); "*Artificial Neural Networks*"; McGraw-Hill Companies, Inc.

- SHIN**, C.K. and **PARK**, S.C.; (1999); "Memory and neural network based expert system"; *Expert Systems with Applications*; N.º16, pp.145/155.
- SILVÉRIO**, Filipe; (2000); "*Redes Neurais: conceitos e aplicabilidade para Marketing*"; In "*Temas em Métodos Quantitativos*"; Elizabeth Reis & Manuel A. M. Ferreira (Ed.); Edições Sílabo, Lda.
- SOHL**, Jeffrey E. and **VENKATACHALAM**, A.R.; (1995); "A Neural Network approach to forecasting model selection"; *Information and Management*; N.º29, pp.297/303.
- THAWORNWONG**, Suraphan and **ENKE**, David; (2004); "The adaptive selection of financial and economic variables for use with artificial neural networks"; *Neurocomputing*; N.º 56, pp.205/232.
- VICENTE**, Eugenio J. F.; (1993); "*Modelos Conexionistas: conceptos generales, origen y evolución*"; In "*Redes Neuronales Artificiales. Fundamentos y Aplicaciones*"; Olmeda & Barba-Romero (Ed.); Actas de un curso de la Universidad de Alcalá; Universidad de Alcalá de Henares; pp.3/21.
- YAO**, Jingtao; **LI**, Yili and **TAN**, Chem Lim; (2000); "Option price forecasting using neural networks"; *Omega, The International Journal of Management Science*; N.º28, pp.455/466.
- ZAFRA**, José Luis, **CARPINTERO**, Antonio y **OLMEDA**, Ignacio; (1993); "*Algunas Aplicaciones de las Redes Neuronales Artificiales*"; In "*Redes Neuronales Artificiales. Fundamentos y Aplicaciones*"; Olmeda & Barba-Romero (Ed.); Actas de un curso de la Universidad de Alcalá; Universidad de Alcalá de Henares; pp.197/214.
- ZHANG**, Guoqiang; **PATUWO**, B. Eddy and **HU**, Michael Y.; (1998); "Forecasting with artificial neural networks: the state of the art"; *International Journal of Forecasting*; N.º14, pp.35/62.

## ANEXO A

**Tabla A.1:** Valores de la Serie Original, para el período entre 1987:01 y 1995:12.

AÑOS MESES	AÑOS								
	1987	1988	1989	1990	1991	1992	1993	1994	1995
ENERO	102.447	118.011	122.217	126.671	126.826	124.194	121.469	118.606	122.480
FEBRERO	102.123	117.547	116.837	129.802	131.653	127.474	129.284	122.988	130.393
MARZO	125.401	142.687	160.658	158.701	188.999	157.536	154.734	175.261	156.645
ABRIL	150.042	167.118	169.326	197.757	182.290	196.087	189.142	185.525	209.263
MAYO	180.430	189.823	199.158	207.876	219.187	223.918	198.402	232.075	218.666
JUNIO	197.113	207.729	218.595	227.159	251.295	207.907	207.216	248.237	222.720
JULIO	229.293	254.523	252.634	257.633	273.927	231.801	231.453	246.274	247.589
AGOSTO	304.847	315.113	329.014	351.500	341.490	312.026	304.576	322.366	320.750
SEPTIEMBRE	238.542	258.287	278.074	284.867	283.378	259.023	249.583	266.094	269.433
OCTUBRE	173.503	174.359	189.664	216.286	197.241	205.400	202.792	206.256	196.466
NOVIEMBRE	130.187	137.933	138.683	162.062	152.554	149.289	141.976	144.803	152.340
DICIEMBRE	114.229	128.774	127.730	139.683	132.802	130.963	120.748	139.706	140.643
<b>TOTAL</b>	<b>2.048.157</b>	<b>2.211.904</b>	<b>2.302.590</b>	<b>2.459.997</b>	<b>2.481.642</b>	<b>2.325.618</b>	<b>2.251.375</b>	<b>2.408.191</b>	<b>2.387.388</b>

Fuente: Elaboración Propia, basada en los datos del INE.

**Tabla A.1:** Valores de la Serie Original, para el período entre 1996:01 y 2003:12.

AÑOS MESES	AÑOS							
	1996	1997	1998	1999	2000	2001	2002	2003
ENERO	126.910	140.430	148.218	163.696	162.389	176.690	165.653	155.527
FEBRERO	139.403	141.183	157.415	165.988	162.637	186.586	181.005	177.818
MARZO	172.393	219.465	209.929	228.149	226.010	245.261	249.214	214.106
ABRIL	213.973	224.382	232.767	242.744	262.865	291.395	253.274	258.519
MAYO	239.142	253.833	280.326	269.854	264.497	306.743	302.028	293.531
JUNIO	245.264	238.334	296.612	270.126	273.881	325.568	301.465	271.454
JULIO	248.398	266.993	303.866	306.031	324.962	351.955	314.560	318.706
AGOSTO	336.086	345.672	377.645	385.868	397.405	452.581	444.991	433.211
SEPTIEMBRE	280.769	288.409	309.700	321.248	331.155	383.793	361.181	343.534
OCTUBRE	225.734	232.052	263.522	280.597	263.217	319.417	287.383	281.472
NOVIEMBRE	175.438	166.835	180.796	193.062	186.445	238.925	221.910	219.463
DICIEMBRE	143.163	141.349	161.273	166.990	157.210	202.351	179.766	178.439
<b>TOTAL</b>	<b>2.546.673</b>	<b>2.658.937</b>	<b>2.922.069</b>	<b>2.994.353</b>	<b>3.012.673</b>	<b>3.481.265</b>	<b>3.262.430</b>	<b>3.145.780</b>

Fuente: Elaboración Propia, basada en los datos del INE.

**Tabla A.2:** Valores Estimados de las Pernoctaciones para el período de 1988:01 a 2002:12.

Mes_Años	Valores Estimados	Mes_Años	Valores Estimados	Mes_Años	Valores Estimados	Mes_Años	Valores Estimados	Mes_Años	Valores Estimados
Ene_88	109.463	Ene_91	122.864	Ene_94	117.350	Ene_97	130.725	Ene_00	152.653
Feb_88	126.140	Feb_91	136.378	Feb_94	132.226	Feb_97	141.482	Feb_00	175.446
Mar_88	143.359	Mar_91	175.819	Mar_94	153.435	Mar_97	184.021	Mar_00	220.430
Abr_88	153.410	Abr_91	191.652	Abr_94	171.950	Abr_97	209.496	Abr_00	230.924
May_88	189.725	May_91	190.727	May_94	193.954	May_97	230.266	May_00	254.908
Jun_88	207.066	Jun_91	223.206	Jun_94	230.919	Jun_97	251.968	Jun_00	261.880
Jul_88	243.362	Jul_91	263.494	Jul_94	258.260	Jul_97	254.730	Jul_00	316.616
Ago_88	298.440	Ago_91	353.077	Ago_94	303.053	Ago_97	337.691	Ago_00	388.008
Sep_88	221.132	Sep_91	258.834	Sep_94	238.071	Sep_97	253.396	Sep_00	323.291
Oct_88	171.140	Oct_91	209.897	Oct_94	190.410	Oct_97	217.711	Oct_00	255.086
Nov_88	149.472	Nov_91	144.335	Nov_94	147.974	Nov_97	167.821	Nov_00	189.885
Dic_88	121.025	Dic_91	132.062	Dic_94	129.492	Dic_97	138.751	Dic_00	155.202
Ene_89	117.716	Ene_92	129.332	Ene_95	123.981	Ene_98	131.367	Ene_01	156.134
Feb_89	126.211	Feb_92	139.855	Feb_95	137.824	Feb_98	159.721	Feb_01	175.393
Mar_89	152.963	Mar_92	170.654	Mar_95	163.834	Mar_98	194.919	Mar_01	226.522
Abr_89	166.907	Abr_92	194.438	Abr_95	195.256	Abr_98	220.546	Abr_01	248.776
May_89	183.835	May_92	199.806	May_95	213.517	May_98	242.702	May_01	275.692
Jun_89	210.684	Jun_92	229.090	Jun_95	231.198	Jun_98	258.854	Jun_01	293.160
Jul_89	255.750	Jul_92	254.392	Jul_95	256.199	Jul_98	292.720	Jul_01	359.721
Ago_89	320.174	Ago_92	338.218	Ago_95	320.779	Ago_98	355.829	Ago_01	412.631
Sep_89	222.187	Sep_92	254.973	Sep_95	251.148	Sep_98	285.444	Sep_01	340.132
Oct_89	181.138	Oct_92	186.361	Oct_95	184.276	Oct_98	217.190	Oct_01	247.581
Nov_89	154.424	Nov_92	160.337	Nov_95	148.254	Nov_98	171.542	Nov_01	213.696
Dic_89	120.228	Dic_92	128.558	Dic_95	127.328	Dic_98	159.197	Dic_01	190.044
Ene_90	119.384	Ene_93	116.654	Ene_96	121.768	Ene_99	154.988	Ene_02	179.577
Feb_90	127.209	Feb_93	132.873	Feb_96	134.060	Feb_99	175.119	Feb_02	178.030
Mar_90	153.552	Mar_93	161.931	Mar_96	168.733	Mar_99	204.532	Mar_02	226.711
Abr_90	176.368	Abr_93	188.682	Abr_96	198.332	Abr_99	233.309	Abr_02	285.316
May_90	200.613	May_93	201.364	May_96	213.649	May_99	253.655	May_02	271.130
Jun_90	211.493	Jun_93	211.254	Jun_96	229.949	Jun_99	287.128	Jun_02	301.279
Jul_90	258.690	Jul_93	233.462	Jul_96	269.584	Jul_99	304.644	Jul_02	336.441
Ago_90	334.944	Ago_93	303.360	Ago_96	320.876	Ago_99	378.240	Ago_02	419.899
Sep_90	250.427	Sep_93	245.336	Sep_96	246.136	Sep_99	305.316	Sep_02	372.824
Oct_90	180.014	Oct_93	186.146	Oct_96	193.849	Oct_99	242.797	Oct_02	333.302
Nov_90	161.717	Nov_93	158.309	Nov_96	159.506	Nov_99	197.416	Nov_02	241.988
Dic_90	135.511	Dic_93	125.856	Dic_96	139.495	Dic_99	158.920	Dic_02	176.532

Fuente: Elaboración Propia.