Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 32 (2007) 1218–1225
Expert Systems with Applications
www.elsevier.com/locate/eswa

Forecasting airborne pollen concentration time series with neural and neuro-fuzzy models

José Luis Aznarte M. [a], José Manuel Benítez Sánchez [a,*], Diego Nieto Lugilde [b], Concepción de Linares Fernández [b], Consuelo Díaz de la Guardia [b], Francisca Alba Sánchez [b]

[a] Department of Computer Science and Artificial Intelligence, University of Granada, ETSI Informatica, 18071 Granada, Granada, Spain
[b] Department of Botany, University of Granada, Spain

Abstract

Forecasting airborne pollen concentrations is one of the most studied topics in aerobiology, due to its crucial application to allergology. The most used tools for this problem are single lineal regressions and autoregressive models (ARIMA). Notwithstanding, few works have used more sophisticated tools based in Artificial Intelligence, as are neural or neuro-fuzzy models. In this work, we applied some of these models to forecast olive pollen concentrations in the atmosphere of Granada (Spain). We first studied the overall performance of the selected models, then considering the data segmented into intervals (low, medium and high concentration), to test how they behave on each interval. Experimental results show an advantage of the neuro-fuzzy models against classical statistical methods, although there is still room for improvement.[1]

Keywords: Aerobiology; Airborne pollen; Neuro-fuzzy; Time series; Forecasting

1. Introduction and problem description

Forecasting future airborne pollen concentrations is undeniably of a high importance because of its medical, environmental and biological effects. The presence and amount of airborne pollen depends on a wide range of factors including meteorological (temperature, rain, humidity, wind, etc.), biological (phenological and physiological state of plants, plants distribution, etc.) and geological (topography) issues. Actually, this is a highly chaotic and thus a hard to model problem.

Application of classic statistical methods to this problem has yielded results not entirely satisfactory (Díaz de la Guardia et al., 2003; Galán, Cariñanos, García-Mozo, Alcázar, & Domínguez-Vílchez, 2001). Models based on Soft Computing techniques have proved successful in a number of hard time series problems, including electric load forecasting (Dash, Liew, Rahman, & Dash, 1995; Kalaitzakis, Stavrakakis, & Anagnostakis, 2002; Kim, Yu, & Song, 2002; Tamimi & Egbert, 2000), financial forecasting (Kuo, 2001; Kuo & Xue, 1998; Vázquez Abad, Fdez-Riverola, & Corchado, 2000), etc.

Regarding Aerobiology, some works have applied Neural Networks to pollen forecasting, reporting encouraging results (Castellano-Méndez, Aira, Iglesias, Jato, & González-Manteiga, 2005; Ranzi, Lauriola, Marletto, & Zinoni, 2003; Sánchez-Mesa, Galán, Martínez Heras, & Hervás-Martínez, 2002). In this work, we have selected models which combine Fuzzy Systems and Neural Networks to model the airborne pollen concentration and compare their performance with classical methods which have been used before.

* Corresponding author. Tel.: +34 958 246143; fax: +34 958 243317.
E-mail address: J.M.Benitez@decsai.ugr.es (J.M. Benítez Sánchez).

[1] This research has been partially supported by Ministerio de Ciencia y Tecnología under project TIC2003-04650.

0957-4174/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2006.02.011

2. Objectives

Two were the main objectives of this work:

- To apply neural models and neuro-fuzzy models to the airborne pollen concentration data, evaluate their performance and compare them against the most used statistical methods.
- To study if forecasting errors vary in different parts of the state-space depending on the absolute value of the series.

3. Data and methodology

The study was carried out on daily aerobiological data obtained over 11 years, from 1992 to 2003 inclusive, in the city of Granada (Southern Spain). Hence, around 4000 data points were available. The data were obtained following the standard methodology of the Spanish Aerobiological Network (Dominguez, Galán, Villamandos, & Infante, 1991), and are measured in grains per cubic meter (grains/m^3) of air.

Only *Olea europaea* L. pollen values were considered, because this species is one of the most allergenic in the Iberian Peninsula, and has a very strong seasonality. Fortunately, it also has a very specific pollen morphology (it is monoespecific) which allows us to indentify it perfectly on the species level. This is important in order to reduce the study to just one type of pollen hence producing a less noisy dataset with a consistent phenological behaviour. In addition, there exist other statistical studies about this pollen series, so more information was available for modelling (Alba & Díaz de la Guardia, 1998; Alba, Díaz de la Guardia, Ocaña, & Valderrama, 2002; Díaz de la Guardia et al., 2003).

To better model the series some preprocessing of data is necessary. Besides of rescaling the dataset into the interval $[0,1]$, special characteristics of the data suggests that further transformations could be used. In particular, the presence of a high variance is normally tackled using a logarithmic transformation. Notwithstanding, in this case, a linear log-like transformation was used instead, considering three different intervals. The first interval, which we shall call *low* interval, concerns all data below 50 grains/m^3 and was selected to try and separate the error of the broad regions of the series with values zero or close to zero, which are known to produce numerical instability in some of the models. The second or *medium* interval was fixed to contain values between 51 and 200 grains/m^3. This second threshold was proposed by the SEAIC (Sociedad Española de Alergología e Inmunología Clínica) for Olea pollen (SEAIC, 2005) as a general turning point between acceptable and risky concentrations, considering the alergological effects on the sensitive population. Fig. 1 shows the transformation applied. The third interval (*high*) includes data from 201 grains/m^3 and above.

Preprocessed data series was divided into two groups: dataset A comprised the years 1992–2002, both included, and was used for training/building the models; and dataset B, which comprised only 2003 data, was used to test the models' performance.

To apply the Soft Computing models, we had to reshape the dataset into a compatible structure: a set of input–output vectors $[x_{t-k}, x_{t-k+1}, \ldots, x_{t-2}, x_{t-1}; x_t]$. To set this structure, i.e. to select the variables used as inputs to the models, we studied the autocorrelation function (acf) and the partial autocorrelation function (pacf) for the transformed dataset (Fig. 2) to get an insight of the inner relations amongst the lagged variables.

These diagrams indicate that present values are strongly influenced by previous days values, decreasing its influence as the time lag increases. Concretely, there is positive partial autocorrelation in the previous 6 days, while the most recent 2 days are those showing a stronger ascendancy over the actual value. For this reason, and taking into account computational efficiency considerations, only two autocorrelation steps were considered here as inputs for the models.

For each of the models, the values of the parameters were selected according to their corresponding authors' indications and after a little tuning through a short trial-and-error stage.

A twofold analysis of the forecast errors was carried out. On the one hand, we aimed at establishing the overall performance of the selected models in one-step-ahead

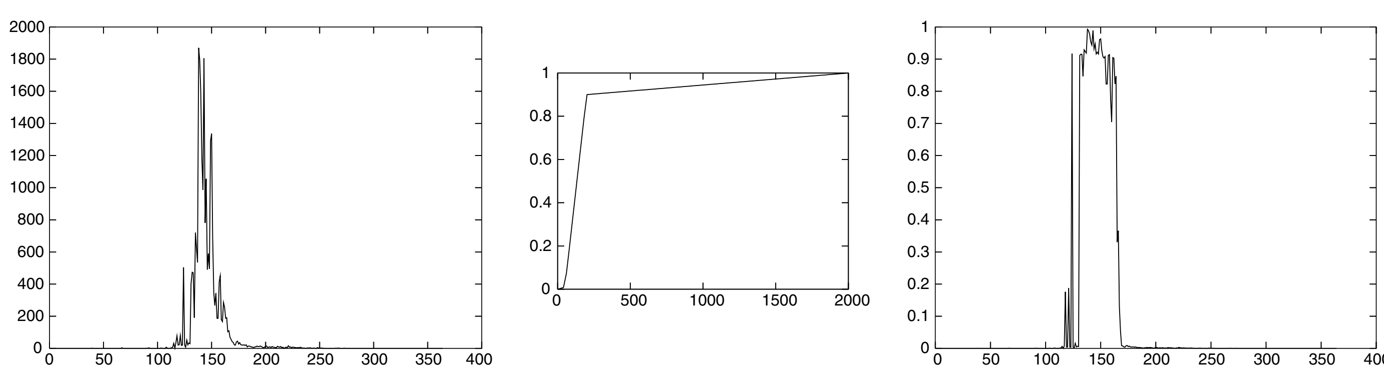

Fig. 1. Sampled 2003 data, linear log-like transformation applied to it and resulting 2003 data.

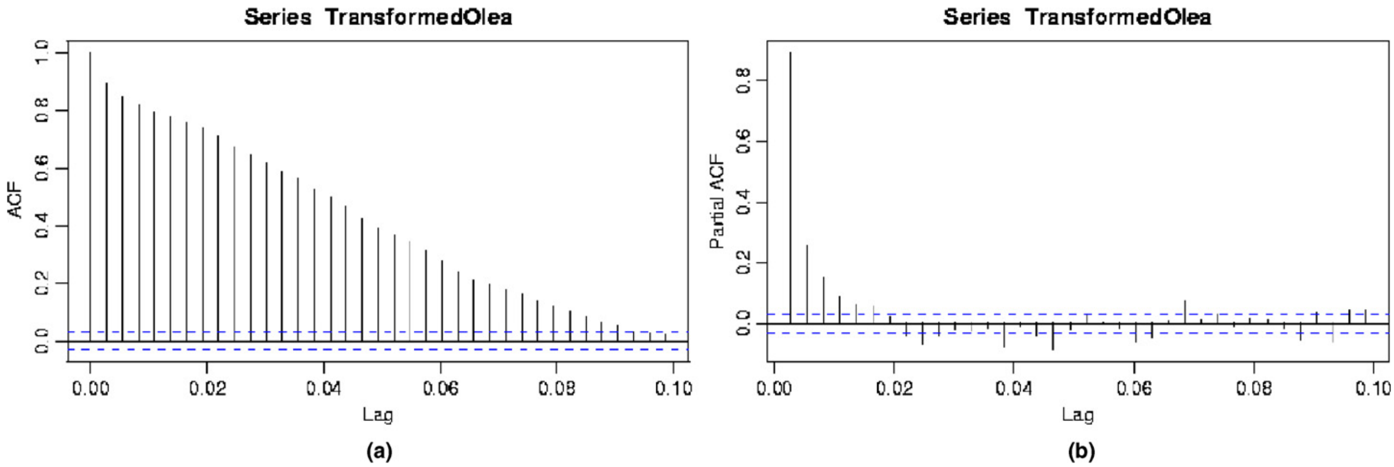

Fig. 2. Autocorrelation (a) and partial autocorrelation (b) diagrams of the transformed data series.

forecasts. Thus we built a model of each type using dataset A to tune its parameters. Then its forecasting accuracy was evaluated by measuring the root mean squared error (RMSE) of the predictions on dataset B.

RMSE is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{1}$$

where N is the number of data available, y_i is the expected output and $\hat{y}_i$ is the output produced by the system for datum i.

Aside from RMSE, we calculated the Theil's U statistic as well, in order to have a measure which is scale-free. Theil's U statistic is defined as

$$U = \frac{\sqrt{\sum_i(\Delta y_i - \Delta\hat{y}_i)^2}}{\sqrt{\sum_i \Delta y_i^2}} \tag{2}$$

where Δy_i and $\Delta\hat{y}_i$ are differences over successive values of the expected and produced outputs, respectively. This measure is not bounded by zero and one. It can be interpreted as the RMSE of the changes for the proposed forecasting model divided by the RMSE of a no-change model. It has the no-change model as the benchmark. Values lower than 1.0 show an improvement over the simple no-change forecast.

On the other hand, we intended to show the accuracy of the one-step-ahead forecasts in terms of the prediction error, but this time we considered it segmented into the intervals that we used to transform the data. We wanted to know if the forecasts were equally good (or bad) on each interval, which could allow us to give the final user an estimation of the accuracy of the forecasts depending on the interval.

4. Considered models

To test the applicability and performance of neural and neuro-fuzzy methods on this problem we have selected the following models: multilayered perceptron (MLP), adaptive neuro-fuzzy inference system (ANFIS), hybrid neuro-fuzzy inference system (HyFIS); generalized regression neural network (GRNN), and NEFPROX. To compare them, we have also selected two of the most popular classic statistical methods applied to time series modelling: ARMA models and Holt–Winters Exponential Smoothing. As a benchmark model, we used a naive predictor, consisting in a model that forecasts the value that the series will take at instant $t+1$ as the value that it took at instant t.

4.1. Classic statistical models

4.1.1. Box–Jenkins models: ARMA

The most popular class of linear time series models consists of autoregressive moving average (ARMA) models, including purely autoregressive (AR) models and purely moving-average (MA) models as special cases (Box & Jenkins, 1970). ARMA models are frequently used to model linear dynamic structures, to depict linear relationships among lagged variables, and to serve as vehicles for linear forecasting.

Box–Jenkins models are one of the most frequently used families of parametric models in time series analysis. This is due to their flexibility in approximating many stationary processes and to their computational efficiency. In Belmonte and Canela (2002) and Alba et al. (2002), there are applications of this approach to the airborne pollen series. In our case, the study of 4the autocorrelation functions of the series suggested that ARMA$(2,0)$ was the most suitable model.

4.1.2. Holt–Winters exponential smoothing

The development of time series models begun with a modelling strategy called classical time series decomposition (Winters, 1960). This approach consists of describing the behaviour of the time series through its non-observable components: trend (T_t), seasonality (S_t), cycle (C_t) and random perturbation (ϵ_t). This is shown in Fig. 3.

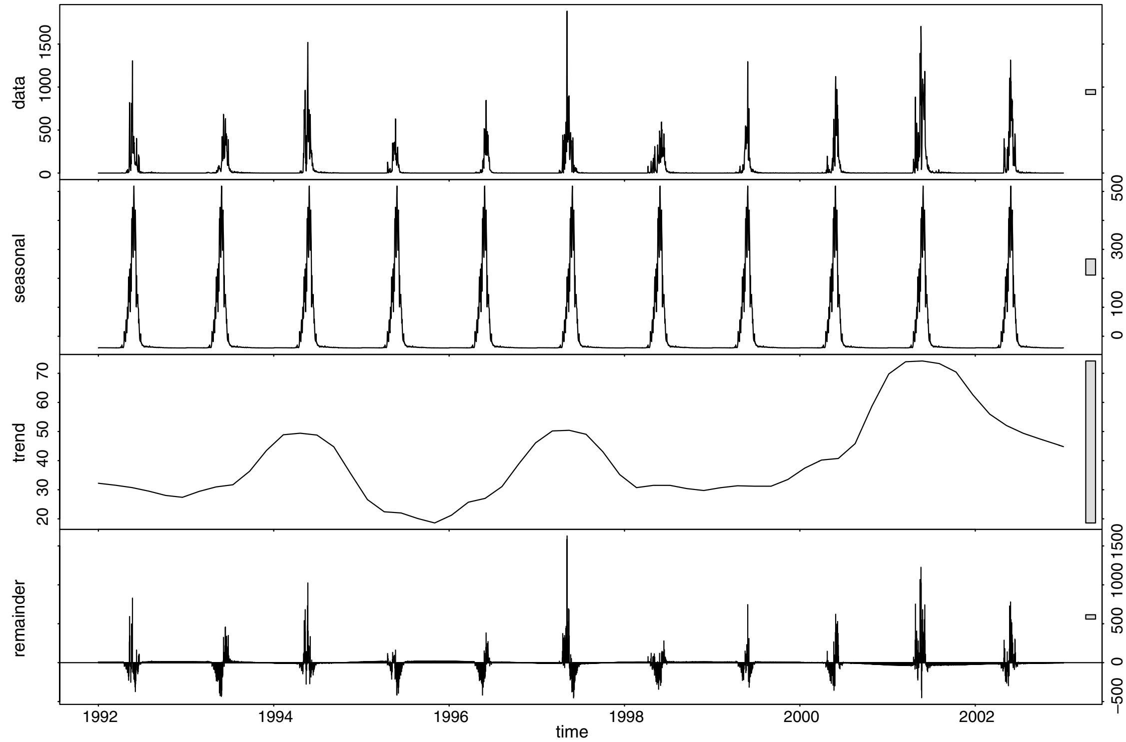

Fig. 3. Decomposition of the series into seasonal and trend components.

Holt–Winters method tries to express the time series as an additive or multiplicative combination of its components:

$$X_t = \mu_t + T_t + S_t + \epsilon_{t-q} \tag{3}$$

where μ_t is the exponentially weighted average of the past values of the series,

$$\mu_t = \alpha X_t + \alpha(1-\alpha)\mu_{t-1} \tag{4}$$

and T_t, S_t are estimated from the data.

4.2. Multilayered perceptron (MLP)

The MLP, trained by the standard backpropagation algorithm is the most widely used neural network approach for complex mappings between input and output. Its mathematical properties for non-linear function approximation are well-documented (Rumelhart, Hinton, & Williams, 1986). In our case, a multilayered perceptron with two inputs and 10 hidden units was used.

4.3. General regression neural networks

The basic GRNN was proposed in 1991 by Specht (1991) as an extension of his probabilistic neural network (PNN). It takes advantage of the fact that given a known joint continuous probability density function $f(\mathbf{x}, y)$ of a vector input $\mathbf{x}$ and a scalar output y, the expected value of y given $\mathbf{x}$ can be computed by estimating the joint pdf using the Parzen estimator.

The core GRNN equation is

$$\hat{y}(\mathbf{x}) = \frac{\sum_{i=1}^{n} y_i \cdot h\left(\frac{\delta(\mathbf{x},\mathbf{x}_i)}{\sigma}\right)}{\sum_{i=1}^{n} h\left(\frac{\delta(\mathbf{x},\mathbf{x}_i)}{\sigma}\right)} \tag{5}$$

where h is a Parzen kernel estimator, usually Gaussian, and δ is a distance measure, Euclidean in our case. The width of each kernel centered on data $\mathbf{x}_i$ is represented by σ, and y_i is the expected output for that data.

4.4. Fuzzy rule based systems for time series analysis

Fuzzy rule based system (FRBS) is a popular computing framework based on the concepts of fuzzy set theory, fuzzy *IF–THEN* rules, and fuzzy reasoning (Klir & Yuan, 1995; Zadeh, 1965). It has found successful applications in a wide variety of fields, such as automatic control, data classification, decision analysis, expert systems, robotics, pattern recognition and forecasting, to name a few.

Each of the fuzzy systems described below used four fuzzy labels for each input, resulting in 16 fuzzy rules for each model.

4.4.1. Adaptive neuro-fuzzy inference system

Adaptive neuro-fuzzy inference systems (Jang, 1993) are a special flavour of fuzzy rule based systems (FRBS) with an adaptation procedure which automatically tunes its parameters. It uses TSK-type rules of the form

If x_1 is A and x_2 is B then $p(x_1, x_2)$ (6)

which are trained by a hybrid learning algorithm. The antecedent parameters are trained by a gradient descent variant while the consequent linear parameters are trained by the least squares method.

4.4.2. Hybrid neuro-fuzzy inference system (HyFIS)

As the ANFIS model described above, HyFIS (Kim & Kasabov, 1999) is a FRBS and a neural network. It employs Mamdani-type fuzzy rules of the form

If x_1 is A and x_2 is B then y is C (7)

which are tuned in a two-stage algorithm. The first stage deals with structure learning and sets the number and configuration of fuzzy rules. The second stage, based on gradient descent as well, fine-tunes the parameters of the system to better model the training data supplied to it.

4.4.3. Neuro-fuzzy function approximation (NEFPROX)

NEFPROX (Nauck & Kruse, 1999) is a Mamdani-type FRBS with a neural structure as the one used by ANFIS or HyFIS. It is closely related to NEFCLASS and NEFCON, and as well it has a two-stage learning algorithm, which fixes the structure of the FRBS on its first stage and then fine-tunes the parameters of the system via a heuristic procedure inspired in the gradient descent method.

5. Results and discussion

Results for the first analysis mentioned in Section 3, which deals with overall performance, are summarized in Table 1. As can be easily seen, the neural and neuro-fuzzy models yielded much better results than the classical approaches as well as than the naive predictor. Concretely, ANFIS and NEFPROX produced the lower RMSEs, being the former the best. The results of the Theil's U statistic are coherent with the ones of the RMSE, showing that all the models are suitable for this task (all of them show values under 1).

Table 1
Experimental results using classical Time Series methods and Soft Computing methods over the transformed testing data

Method	RMSE	Theil's U
Naive	0.0918	0.3233
Holt–Winters	0.1102	0.3881
ARMA(2,0)	0.1029	0.3625
MLP	0.0918	0.3234
GRNN	0.0903	0.3181
HyFIS	0.0912	0.3211
NEFPROX	0.0895	0.3162
ANFIS	0.0882	0.3107

It is worth mentioning that the naive predictor behaved quite better than the Holt and Winters and the ARMA approaches. The complexity of the series, which is clearly composed by several regimes (increasing, decreasing, zero regime,...), makes it very difficult to model it completely by a single linear model, whilst the naive model is totally regime-dependent. This also highlights as well that the Soft Computing-based models have an inherent capability to model multiple-regime processes.

Fig. 4 shows both the original 2003 data and the forecasts obtained by using the trained ANFIS and NEFPROX. It is noticeable how both systems manage to model the general behaviour of the data, including sudden rises and falls. NEFPROX, especially, manages to predict the peaks of the series considerably well. Notwithstanding, a close look at these graphs shows that the predicted values sometimes have a 1–2 days delay with respect to the original data. This may be explained by the fact that no exogenous variables (such as temperature, humidity, etc) have been used, and the highly chaotic behaviour of the series is hard to model without that information. In any case, this phenomenon is consistent with other author's results (Alba et al., 2002; Díaz de la Guardia et al., 2003).

Once the first analysis was done, we considered the error obtained in each of the intervals mentioned in Section 3. To properly understand these results, the distribution of the

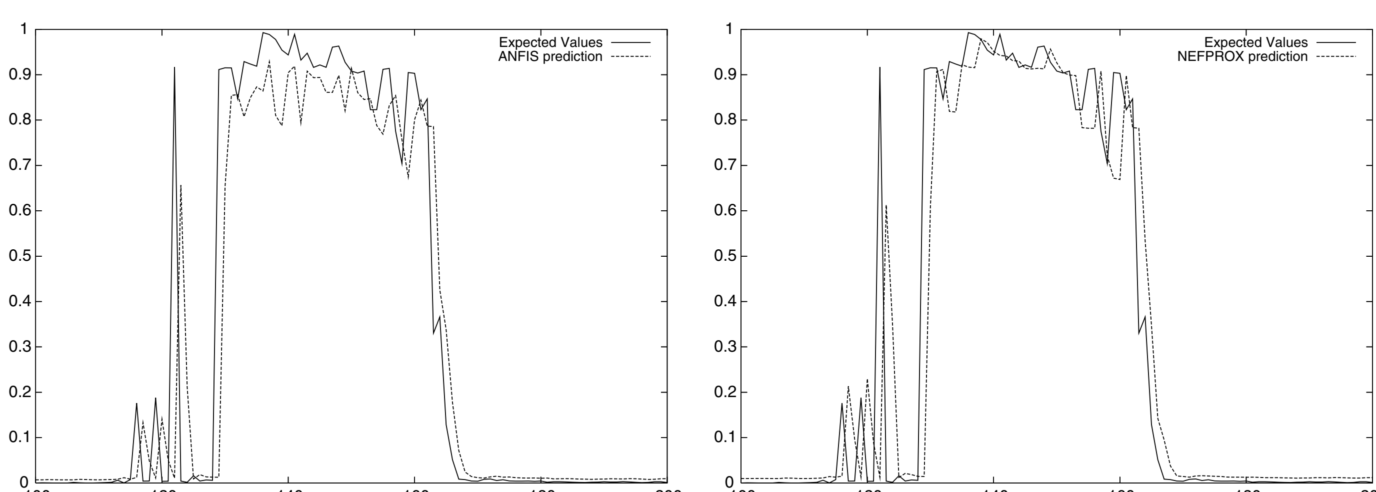

Fig. 4. Values for 2003: expected and forecasted by ANFIS and NEFPROX.

testing data over the intervals must be stated: out of the 363 examples available, 321 were on the *low* interval, 14 on the *medium* interval and 28 on the *high* interval. This is a first hint showing that more attention should be payed to the selection of the thresholds defining the intervals.

The error values for this second experiment are shown in Table 2, where we can see that the lowest values of RMSE are obtained on the *low* interval for all models. The justification for the use of the Theil's U becomes clear now, as well as the need for carefully choosing error measures which are scale free: the values of the Theil's U contradict those of the RMSE, throwing on every case very bad results.

The explanation to this apparent contradiction is given by several joint facts. On the one hand, we know that in the first interval the majority of data are zero, and this leads the RMSE to very low values on every model. On the other hand, in this interval the series under study is considered to be noisy, because of some natural phenomena occurring just before and after the main polinisation period: the long distance pollen transportation from more meridional zones which causes the presence of grains in zones where the flowering has not started (Díaz de la Guardia et al., 2004), the *late deposition* process, which causes the presence of grains weeks after the polinisation period and the so called *resuspension* of grains, which as well affects the pollen counts once the polinisation period has finished.

Table 2
Experimental results obtained by all the models on each transformed interval

Method	Interval	$RMSE_{test}$	Theil's U
Naive	Low	0.0529	30.75
	Medium	0.1764	0.2984
	High	0.2482	0.1881
Holt–Winters	Low	0.0405	23.52
	Medium	0.2911	0.4925
	High	0.3102	0.2351
ARMA(2, 0)	Low	0.0377	21.92
	Medium	0.2224	0.3762
	High	0.3103	0.2352
MLP	Low	0.0454	26.82
	Medium	0.1712	0.2896
	High	0.2665	0.2025
GRNN	Low	0.0427	24.86
	Medium	0.1527	0.2584
	High	0.2705	0.2055
HyFIS	Low	0.0373	21.67
	Medium	0.1670	0.2825
	High	0.2791	0.2115
NEFPROX	Low	0.0445	25.96
	Medium	0.1656	0.2801
	High	0.2607	0.1976
ANFIS	Low	0.0406	23.59
	Medium	0.1567	0.2652
	High	0.2641	0.2001

Another definitive argument to justify the bad results shown for the Theil's U statistic can be obtained by remembering its definition (Section 3). Obviously a no-change model will produce very good overall results when predicting these values composed mostly of zeros, and this will consequently rise the value of the statistic for any other model.

At first sight, it may be surprising that the neural and neuro-fuzzy models obtain worst results than the linear model ARMA on this interval, whose values are mostly zeroes. This fact is explained by considering that those Soft Computing models are based on families of functions that asymptotically tend to zero (logistic, Gaussian,...) and this makes it difficult for them to obtain an absolute zero in the outputs. In fact, many applications of neural networks use to rescale the data to the interval $[0.1, 0.9]$ to avoid this effect.

Notwithstanding, by comparing with the results obtained by the naive predictor, we can conclude that all the models are performing fairly well on this *low* interval. Besides, it should not be forgotten that a high accuracy in the prediction within this interval is only of a relative importance, given the fact that any value included in it would normally not cause any effects on the allergic population.

The values obtained in the *medium* interval confirm what was stated above about the two error measures employed. In this case, the RMSE shows values for all models which are much worst than those of the first interval, while the values of the Theil's U are now a lot better than those obtained for the first interval and show a significative improvement from the no-change model.

It must be recalled that the number of data included in this interval is significantly lower than those of the other intervals. This is probably affecting the predictive capabilities of all the models and can be considered as a justification for the results for this interval being worst than those obtained in the *high* interval.

Regarding each model's results for this interval, we can see that the best models were GRNN, ANFIS and NEFPROX, while all the neural and neuro-fuzzy models showed an improvement over the classical statistical approaches.

The results obtained in the *high* interval are somehow surprising. The naive predictor outperforms any of the other models, a phenomenon which requires an explanation. By studying how each of the models work we can again understand this situation. On the one hand we know that the classical, neural and neuro-fuzzy models are built and/or trained by using the historical information obtained from the training set, and that they try to forecast future values by taking into account that historical information as well as the recent past of the series. Differently, the naive predictor uses *exclusively* the previous day value for predicting, and does not consider in any case the history of the series. This fact, which might be considered as one of the main drawbacks of the naive approach, is in this case an advantage that justifies its good results.

As can be seen in the first line of Fig. 3, the annual peaks of the series show a high volatility (ranging from 400 to 2000 grains/m^3). This together with the fairly small amount of years available (we might say that 11 values do not conform a predictable series), indicates that the models that take into account the whole history of the series tend to predict peak values that are on the average of the available years. Considering all this together with the high value of the peak of the testing year (over 1500 grains/m^3) we can conclude that a naive approach is better in predicting the peaks of this series.

Having said this, the models obtaining better results are NEFPROX, ANFIS and MLP. By looking at the Theil's U statistic results, we see how in general the models performed better in the third interval than in the second one. The number of data available on each of them might explain this result.

6. Conclusions and further work

We have applied five neural and neuro-fuzzy models (MLP, GRNN, HyFIS, NEFPROX and ANFIS) to the modelling and prediction of the airborne pollen series in the city of Granada, Spain. Those models, whose behaviour is mainly non-linear, showed significant improvements over the traditional ARMA and Holt–Winter's linear approaches. The chaotic component of the series under study makes the non-linear behaviour of Soft Computing models more appropriate than the linear one, obtaining better results in all the cases. Concretely, ANFIS and NEFPROX were the models that reached best overall results.

Considering intervals in the domain of the data allowed us to better understand how the models behave, showing that some models have a better performance depending on the area. Accordingly, better results are obtained in general in the *medium* and *high* intervals, while the *low* interval is harder to model for all the models. The best models were HyFIS and ANFIS in the *low* interval, GRNN and ANFIS in the *medium* interval and NEFPROX and ANFIS in the *high* interval.

Results from both versions of the experiment suggest that ANFIS and NEFPROX are the most appropriate models for the present task.

A few remarks can be done concerning some aspects of the experiments. On the one hand, the thresholds of the intervals should be carefully revised attending not only to symptomatological concerns but as well to statistical criteria. Although current criteria are proposed by practitioners and allow them to extract useful information from the predictions, a better choice of the thresholds aiming at a more even distribution of data amongst intervals would lead to a significant improvement of the predictions.

Nonetheless, although our results seem good compared to previous works on the literature (and compared to the results of the classic approaches), the absence of meteorological variables in our study allows us to expect further improvements in the prediction capabilities of future applications.

References

Alba, F., & Díaz de la Guardia, C. (1998). The effect of air temperature on the starting dates of Ulmus, Platanus and Olea polen seasons in the SE Iberian Peninsula. *Aerobiologia, 14*, 191–194.

Alba, F., Díaz de la Guardia, C., Ocaña, F., & Valderrama, M. (2002). Modelos de regresión lineal dinámica aplicados al polen de Olea y Cupressaceae en la ciudad de Granada. In *XIV Simposio de Palinología de la Asociación de Palinólogos de lengua española (APLE), Salamanca*, Spain.

Belmonte, J., & Canela, M. (2002). Modelling aerobiological time series. Application to Urticaceae. *Aerobiologia, 18*, 287–295.

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.

Castellano-Méndez, M., Aira, M. J., Iglesias, I., Jato, V., & González-Manteiga, W. (2005). Artificial neural network as a useful tool to predict the risk level of *Betula pendula* in the air. *International Journal of Biometeorology, 49*(49), 310–316.

Dash, P. K., Liew, A. C., Rahman, S., & Dash, S. (1995). Fuzzy and neuro-fuzzy computing models for electric load forecasting. *Engineering Applications of Artificial Intelligence*.

Díaz de la Guardia, C., Alba, F., Girón, F., Nieto, D., De Linares, C., Candau, P., et al. (2004). Risk evaluation of exposure to /emph *Olea europaea* L. pollen in andalusia (S Spain). *Polen, 14*, 104–105.

Díaz de la Guardia, C., Alba, F., Trigo, M. M., Galán, C., Ruiz, L., & Sabariego, S. (2003). Aerobiological analysis of *Olea europaea* L. pollen in different localities of Southern Spain asas. *Grana, 42*, 234–243.

Dominguez, E., Galán, C., Villamandos, F., & Infante, F. (1991). Manejo y evaluación de los datos obtenidos en los muestreos aerobiológicos. In *Monografías REA/EAN*, Vol. 1 (pp. 1–18).

Galán, C., Cariñanos, P., García-Mozo, H., Alcázar, P., & Domínguez-Vílchez, E. (2001). Model for forecasting *Olea europaea* L. airborne pollen in South-West Andalusia, Spain. *International Journal of Biometeorology, 45*, 59–63.

Jang, J. -S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*.

Kalaitzakis, K., Stavrakakis, G. S., & Anagnostakis, E. M. (2002). Short-term load forecasting based on artificial neural networks parallels implementation. *Electric Power Systems Research*.

Kim, C., Yu, I., & Song, Y. H. (2002). Kohonen neural network and wavelet transform based approach to short-term load forecasting. *Electric Power Systems Research*.

Kim, J., & Kasabov, N. (1999). HyFIS: adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems. *Neural Networks*.

Klir, G. J., & Yuan, B. (1995). *Fuzzy sets and fuzzy logic*. Prentice Hall.

Kuo, R. J. (2001). A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm. *European Journal of Operational Research*.

Kuo, R. J., & Xue, K. C. (1998). A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights. *Decision Support Systems*.

Nauck, D., & Kruse, R. (1999). Neuro-fuzzy systems for function approximation. *Fuzzy Sets and Systems, 101*(2), 261–271.

Ranzi, A., Lauriola, P., Marletto, V., & Zinoni, F. (2003). Forecasting airborne pollen concentrations: development of local models. *Aerobiologia, 19*, 39–45.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by backpropagation errors. *Nature, 323*, 533–536.

Sánchez-Mesa, J. A., Galán, C., Martínez Heras, J. A., & Hervás-Martínez, C. (2002). The use of a neural network to forecast daily grass pollen concentration in a Mediterranean regions: the southern part of the Iberian Peninsula. *Clinical and Experimental Allergy, 32*, 1606–1612.

Sociedad Española de Alergología e Inmunología Clínica (SEAIC). Grasses, how to interpret the pollen counts. Available from http://polenes.com/en/interpretacion.html, October 2005.

Specht, D. F. (1991). A general regression neural network. *IEEE Transactions on Neural Networks*.

Tamimi, M., & Egbert, R. (2000). Short term electric load forecasting via fuzzy neural collaboration. *Electric Power Systems Research*.

Vázquez Abad, J., Fdez-Riverola, F., & Corchado, J. M. (2000). Forecasting economic cycles with conectionist models. In *I international meeting on economic cycles*.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science, 6*, 324–342.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 3*(8), 338–353.