

Document downloaded from:

<http://hdl.handle.net/10251/63854>

This paper must be cited as:

Capilla Roma, CA. (2015). Artificial neural network approach for forecasting nitrogen oxides concentrations. *Environmental Engineering Science*. 32(9):781-788.
doi:10.1089/ees.2014.0556.



The final publication is available at

[http://dx.doi.org/ 10.1089/ees.2014.0556](http://dx.doi.org/10.1089/ees.2014.0556)

Copyright Mary Ann Liebert

Additional Information

Artificial neural network approach for forecasting nitrogen oxides concentrations

Carmen Capilla*

Department of Applied Statistics and Operations Research and Quality, Polytechnic University of Valencia, Spain

Abstract

This paper presents the application of feed-forward multilayer perceptron networks to forecast hourly nitrogen oxides levels 24 hours in advance. Input data were meteorological variables, average hourly traffic and nitrogen oxides hourly levels. The introduction of four periodic components (sine and cosine terms for the daily and weekly cycles) was analyzed in order to improve the models' prediction power. The data were measured during three years at monitoring stations in Valencia (Spain) in two locations with high traffic density. The models' evaluation criteria were the mean absolute error, the root mean square error, the mean absolute percentage error, and the correlation coefficient between observations and predictions. Comparisons of multilayer perceptron-based models proved that the insertion of the four additional seasonal input variables improved the ability of obtaining more accurate predictions, which emphasizes the importance of taking into account the seasonal character of nitrogen oxides. When using seasonal components as predictors, the root mean square error (RMSE) improves from 20.29 to 19.35 when predicting nitrogen dioxide, and from 45.07 to 42.37 when forecasting nitric oxides if the model includes seasonal components At one study location. At the other location the RMSE changes from 23.76 to 23.05 when predicting nitrogen dioxide and from 33.94 to 33.10 for the other pollutant's forecasts. Neural networks did not require very exhaustive information about air pollutants, reaction mechanisms, meteorological parameters or traffic characteristics, and they had the ability of allowing nonlinear and complex relationships between very different predictor variables in an urban environment.

Key words: air quality; nitrogen oxides concentration; urban atmospheric pollution; multilayer perceptron; seasonal variability.

Introduction

Air pollution in urban areas is a worldwide growing problem. Administrations introduce plans and regulations to reduce pollutants emissions. Emission directives establish limit values for concentrations of pollutants with the aim of avoiding, preventing and reducing the harmful effects on human health and the environment as a whole (Senger, 2000). Modeling temporal variations in the concentration of pollutants is useful when evaluating the effectiveness of these plans. These models allow long and short term forecasting of pollution levels. They have to take into account the link between climate and pollutants, which plays an important role in the variability.

*Corresponding author: Department of Applied Statistics and Operations Research and Quality, Polytechnic University of Valencia, Camino de Vera s/n, 46022 Valencia, Spain. Phone: +34 963877490, fax: +34 963877499. E-mail:ccapilla@eio.upv.es

Tools to predict pollution levels can be used in different ways. Deterministic and statistical models have been developed for the purpose of forecasting. Deterministic models are not appropriate as prognostic models in coastal zones (Rye, 1995). They are more suitable for extensive areas such as whole regions and large cities. One of the causes of the uncertainty of deterministic models is the lack of sufficient data as they require precise data from the emission and transportation of pollutants (deriving from traffic) and meteorological conditions. As the complexity of a problem increases, the theoretical understanding decreases due to ill-defined interactions between systems, and statistical approaches are required.

Statistical models are able to establish a relationship between input variables (predictors) and output variables, without detailing the causes and effects of the formation of pollutants. Autoregressive statistical models are often used to analyze the seasonality, trend and autocorrelation of pollutants variability (Box et al., 2008; Prada-Sanchez et al., 2000), but they are limited by their weakness when modeling non-linear temporal variations. Classification and regression trees have also been applied to study the pollutant variability (Ryan, 1995; Gardner and Dorling, 2000). The possible presence of chaotic dynamics in pollutant concentrations allows modeling non-linear time series (Kocak et al., 2000). Donnelly et al. (2015) applied a non-parametric kernel regression model to forecast nitrogen dioxide concentrations 48h in advance, using temporal variations and correlations with meteorology. The model had low computational resources and gave index of agreement values between 0.74 and 0.94.

During the last two decades the use of artificial neural networks, and, in particular, the application of multilayer perceptron (MP) models, has been developed to forecast pollutant concentrations (Gardner and Dorling, 1998). Neural networks have been shown to be effective alternatives to more traditional statistical techniques (Shalkoff, 1992). The neural network models can be trained to approximate virtually any smooth, measurable function (Hornik et al., 1989). Unlike other statistical techniques, they make no prior assumptions concerning the data distribution. They can model highly non-linear functions and can be trained to accurately generalize when presented with new, unseen data (Bishop, 1995). These features of neural networks make them an attractive alternative to developing numerical models, and also when choosing between statistical approaches. Ibarra-Berastegi et al. (2008) focused on the prediction of hourly levels up to 8 h ahead for five pollutants at six locations in the area of Bilbao (Spain) using MP. The best performance of these models at the different sensors in the area, was obtained for the prediction of nitrogen dioxide (NO₂) 1 hour ahead (correlation coefficient between observations and predictions=0.9),

and the worse results were observed for the prediction of ozone 8 hour ahead. Caselli et al. (2009) compared the MP and multivariate regression models to predict critical pollution events. The regression models gave less accurate results mostly for 1 day forecasting and failed when fitting spiked high values of pollutant concentrations. Arhami et al. (2013) investigated the combination of artificial neural networks and Monte Carlo simulations, to quantify model uncertainty when predicting several pollutants at two urban sampling locations of a developing country. They used meteorological parameters and seasonal components as predictors, and validated the models with simulated data. They concluded that this methodology allowed selecting input variables and models' architecture in their study area. The best mean square errors obtained were 18.7 for nitrogen dioxide predictions, and 27.5 for nitric oxides. Elangasinghe et al. (2014) developed an artificial neural network model for predicting nitrogen dioxide concentrations at a site near a major highway in Auckland, New Zealand. They compared models with different inputs (meteorological parameters, hour, day and month). Their study revealed that carefully choosing of inputs can give more reliable nitrogen dioxide forecasts, but the authors indicated that the inclusion of emission rates might improved the methodology. The artificial neural network model outperformed a linear regression model based on the same input parameters.

The objective of this study is to investigate, for the first time, the capability of the MP method to forecast NO₂ and nitric oxide (NO) concentrations in the Valencia urban area (Spain). The primary goal of the work is to predict concentrations 24 hours ahead at two different locations. This forecasting period of 24 h has been selected for practical regulatory reasons; shorter time forecasts are of minimal value for air quality management purposes. NO₂ and NO are relevant air pollutants in Valencia (European Communities 2007). They are mainly a consequence of motor vehicle emissions, and industrial pollution plays a smaller part (Ballester et al., 2005). Tenias et al. (1998) reported a significant connection between a 10 µg/m³ increase in NO₂ level and the relative risk of asthma emergency visits in Valencia. Daily levels of NO₂ in this city are also associated with cardiovascular admissions (Ballester et al., 2001; Ballester et al., 2006). This pollutant is precursor of other secondary pollutants that are related to photochemical smog and acid rain. Ambient air NO₂ is in large part originated by the oxidation of NO. The link between climate and these pollutants plays an important role on their variability and has to be taken into account when selecting optimal pollutant reduction strategies to avoid exceeding emission directives. In this paper, several MP models are designed and compared to establish the most efficient

performer as a forecasting tool using meteorological and traffic variables, pollutants concentrations, and seasonal components as predictors.

Materials and methods

Study area and database

The study area is located in Valencia (Spain). This city has around one million inhabitants, and Mediterranean climatology and structure. An air pollution network managed by the local government since 1995, measures pollution variables in its urban area. The traffic network of the local municipality measures the number of vehicles (NV) circulating every hour at locations close to the pollution monitoring sites. The data used in this work were hourly observations from the air pollution and traffic networks. The study considers two monitoring stations: Pista Silla and Viveros. Fig. 1 presents the map of the study monitoring sites' location. These two stations were selected because several high pollution episodes were registered to them during the period 2002-2005. The limit value of NO₂ for the protection of human health in a calendar year (averaging period) was exceeded at Pista Silla, in 2003, 2004 and 2005. At this site, the highest annual NO₂ mean was observed in 2003. Mass concentrations of nitrogen oxides are determined using the chemiluminescence method. Pollutants concentrations are expressed in µg/m³. The volumes are standardized at a temperature of 293°K and a pressure of 101.3 kPa. The air pollution monitoring site at Pista Silla also measures wind speed (WS, m/s), wind direction (WD, degrees), temperature (T, °C), solar radiation (SR, W/m²), relative humidity (RH, %) and pressure (P, mbar). At Viveros location, WS, WD, T and SR observations were provided by the National Institute of Meteorology, which manages a meteorological station close to the air pollution station. Pista Silla station is in a roadside site located a few meters from a motorway, and Viveros is in an avenue close to the city center. The distance between them is 2,6 km. Traffic density is high at both sites. The matrix of data (hourly measurements) had 18339 entries for Pista Silla (years 2003-2005) and 16221 for Viveros (years 2002-2004), after eliminating rows with missing values. Table 1 shows averages, coefficients of variation and maximum values of pollutants, meteorological and traffic variables for both stations, during the study period.

Factors mainly contributing to NO₂ and NO concentrations are connected with the source activity (e.g. traffic) and periodic variations in nature (e.g. photochemical reactions in the atmosphere). Periodic components are expected in the time series at the week level and in the form of daily variations. Fig. 2 represents the average diurnal cycle of NO₂ and NO levels at Pista Silla during the study period. The average weekly cycle at the same station for the two pollutants can be seen in Fig. 3. The average traffic strength is also plotted in these diagrams with another scale on the right side. There are differences between hours, with concentrations' peaks associated with high traffic density. The daily and weekly variations of the two pollutants depend on average daily and weekly traffic variations. The same cyclical patterns were observed at Viveros station.

Multilayer perceptrons

The MP models used in this work were composed of three layers of neurons: the input layer, the hidden layer, and the output layer. The models were compared with different number I of predictors X_i or neurons in the input layer. The predictors were pollutants concentrations, meteorological parameters, traffic variable or seasonal components (sine and cosine terms for the daily and weekly cycles). The models' output Y was the prediction of NO₂ or NO concentrations 24 hours in advance; therefore the number of neurons in the output layer was equal to 1. Table 2 shows the models that were analyzed. The number of neurons H in the hidden layer was determined by experimentation, training the neural networks with values of H from 5 to 30. Greater values of H did not give a better performance. The MP networks were trained with two backpropagation algorithms: the scaled conjugate gradient algorithm (SCG) and the Levenberg-Marquard algorithm (LM). It has been proved that both algorithms converge faster and perform better than other backpropagation algorithms (Moller, 1993). The output Y^o can be expressed as follows:

$$Y^o = f^o(b^o + \sum_{j=1}^H w_j^o f^h(b_j^h + \sum_{i=1}^I w_{ij}^h X_i)) \quad (1)$$

where o denotes the elements of the output layer and h indicates the elements of the hidden layer. w_j^o is the weight that connects the neuron j of the hidden layer with the neuron of the output layer, and w_{ij}^h is the weight that connects the neuron i of the input layer with the neuron j of the hidden layer. b^o is the bias of the neuron of the output layer, and b_j^h is the bias of neuron j of the hidden layer. f^o is the transfer function of the neuron of the output layer. In this work, the linear transfer function has been applied for

f^o . f^h is the transfer function of neuron j of the hidden layer. The most widely used f^h are the hyperbolic transfer function (tansig) and the logarithmic sigmoid function (logsig):

$$tansig(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

$$logsig(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

Fig. 4 shows an MP model with I neurons in the input layer, H neurons in the hidden layer, and one neuron in the output layer.

Overtraining does occur when the MP memorizes the patterns introduced to it, and it is not capable of identifying new situations. The early stopping technique can be used to avoid this problem (Sarle, 1995). In this technique, the available database is separated into three subsets: the training set, the validation set, and the test set. The training set is used to update the network weights and biases. During the training, the validation set is used to guarantee the generalization capability of the model, and training should stop before the error on the validation set begins to rise. The test set is a new set used to check the generalization of the MP. In this work, the models were trained on data from the first year. Data from the second year were used as the validation set, and observations from the third year are the test data set. The computations were performed with the *Neural Network Toolbox* of MATLAB.

Evaluation criteria

Four statistical parameters were obtained to compare the performance of the MP models with the test dataset. They are the most used indices to assess the quality of an estimator (Willmott et al., 1985; Gomez-Sanchis et al., 2006). The correlation coefficient r between the forecasted values Y_f and the observations Y quantifies the global description of the model. We computed the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - Y_{fi})^2}{n}} \quad (4)$$

where n is the number of observations in the test data set. The mean absolute error (MAE) also measures how close forecasts are to observations:

$$MAE = \frac{\sum_{i=1}^n |Y_i - Y_{fi}|}{n} \quad (5)$$

An expression of accuracy of predictions as a percentage can be computed with the mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - Y_{fi}}{Y_i} \right| \quad (6)$$

Results and discussion

The best performance indices results at Pista Silla station are shown in Table 3. It contains the best results obtained with the models MP1 to MP4 (output $(NO_2)_{t+24}$), and the models MP5 to MP8 (output NO_{t+24}). The table indicates the backpropagation algorithm, the transfer function, and the number of neurons in the hidden layer. The comparison of the results indicates that the most accurate predictions of $(NO_2)_{t+24}$ at Pista Silla, were obtained with the Levenberg-Marquard (LM) algorithm. The MP2 model using the tansig transfer function and this algorithm gives good values of the four evaluation criteria when the number of neurons in the hidden layer is $n_h=14$. The values of these indices show that the obtained model is a good estimator. The predictors are meteorological parameters, traffic, seasonal components and $(NO_2)_t$ concentrations. Fig. 5 represents predictions with this model corresponding to the first 100 hours of the test data set. The period is from 00:00h on January 1st 2003 to 3:00h (am) on January 5th 2003. Valencia shows clear seasonal variations, with increasing activity during the coldest months. Nitrogen oxides emissions are higher during times of lower temperatures and also increase with reduced traffic speed, which in Valencia also occurs during winter months when urban locations have a greater density of traffic. The optimal neural network matches actual observations very appropriately, and captures the concentrations' peaks and troughs. Fig. 6 gives the scatterplot of the data set test predictions versus observations. The evaluation of NO_{t+24} predictions at Pista Silla is also in Table 3. In this case, the LM algorithm also gives most accurate forecasts. The MP6 model with $n_h=10$, this algorithm, and the logsig transfer function, provided the best value of the RMSE. In this case, the predictors are meteorology, traffic, seasonality and NO_t concentration. The model MP8 gives better prediction results in terms of MAE and MAPE. This model includes all the potential predictors (see Table 2). The Fig. 7 shows actual and predicted values of NO_{t+24} with this model for the first 100 hours of the test data set (from 00:00h on January 1st 2003 to 3:00h am on January 5th 2003). The neural model fits the observed values correctly. The accuracy of predictions of these models for the two pollutants can be compared using the

two adimensional parameters MAPE and r. MAPE, which computes this accuracy as a percentage, is better for $(\text{NO}_2)_{t+24}$ (MAPE=0.45) than for NO_{t+24} (MAPE=1.13).

Table 4 shows the best predictions results of the test data set at Viveros station. Forecasts evaluations of $(\text{NO}_2)_{t+24}$ can be seen in the first four rows (models MP1 to MP4). Very similar results were obtained with the two transfer functions. The MP2 model gives the best RMSE result as can be seen in the table, when using the SCG algorithm, the tansig transfer function and with a number of neurons in the hidden layer $n_h=10$. The inputs to the neural network are meteorological variables, traffic, seasonal components, and $(\text{NO}_2)_t$ levels. Predictions of this model in the first 100 hours of the test data set are plotted in Fig. 8. The period is from 17:00h on February 25th 2002 (the first complete record), to 21:00h February 28th 2002. The value of the MAPE=0.98 indicates that the fit is less accurate than at Pista Silla, where the MAPE results are smaller with the four models (see Table 3). The correlation coefficients between $(\text{NO}_2)_{t+24}$ observations and predictions are very small in Viveros (see Table 4). This indicates that the linear agreement between this two variables is very poor, and then the performance of the model MP1 to MP4 worse than at Pista Silla. Fig. 5 also shows a closer agreement between observations and forecasts than Fig. 8. The model MP4, that includes NO_t as predictor, has a smaller MAPE value.

The evaluation indices for the predictions of NO_{t+24} at Viveros are also in Table 4. The model that provides the best RMSE of NO_{t+24} forecasts is the MP8 model at Viveros station. The transfer function is the tansig, with the Levenberg-Marquard backpropagation algorithm, and the number of neurons in the hidden layer is $n_h=10$. Fig. 9 shows the time series of NO_{t+24} predictions and actual values in the first 100 hours with complete data of the test data set at Viveros (the first 20 hours had missing records). Predictions are less accurate than at Pista Silla, where the correlation coefficient resulted for this pollutant $r=0.66$ and the MAPE is equal to 1.13 with this model. At Viveros, these parameters are $r=0.52$ and MAPE=3.32. The correlation coefficient is better when predicting NO_{t+24} than $(\text{NO}_2)_{t+24}$ at Viveros, but the agreement between actual and predicted values, expressed as a percentage, is better for $(\text{NO}_2)_{t+24}$. The MP models might perform worse at Viveros because the nitrogen oxides' temporal variations might depend on other environmental parameters not considered here, such as other meteorology parameters. At Viveros location, wind speed, wind direction, temperature and solar radiation observations were used as inputs. At Pista Silla, pressure and relative humidity were also available.

The MP models have the advantages of making no prior assumptions concerning the data distribution. They have also modeled the non-linear relationships existing between inputs and outputs, and have been

trained to accurately generalize when presented with new, unseen data. They are easy to use and understand compared to other statistical methods. These models can also be applied in other areas with similar air quality problems and meteorological and traffic influences. Future work will involve their application to forecast other pollutants. However, the MP is a “black box” learning approach, cannot interpret relationships between inputs and output, and cannot deal with uncertainties. There are other linear statistical methods where a greater understanding of the cause and effect can be obtained, but they are not useful in this study given the highly non-linear relationships between inputs and outputs. Other approaches such as non-parametric regression models have recently (Donnelly et al. 2015) performed well for air quality prediction in other places. A comparison of MP models with these approaches in the study locations is a promising research area. These models may be a good alternative to the methods used in this work.

Summary

In this work, neural network models, and more particularly, multilayer perceptron models have been developed in order to forecast nitrogen oxides levels 24 hours in advance. This forecasting period was selected because the shorter time forecasts are of minimal value for air quality management purposes. Management of control and public warning strategies for nitrogen oxides levels, require accurate forecasts of the concentration of these pollutants, and their dependence on meteorological and traffic variables. The multilayer perceptron performed better when predicting nitric oxide at the study locations, if the models included daily and weekly seasonal components. These components were introduced with four additional input variables $\sin(2\pi h/24)$, $\cos(2\pi h/24)$, $\sin(2\pi d/7)$ and $\cos(2\pi d/7)$, $h=1, 2, \dots, 24$, $d=1, 2, \dots, 7$. In one of the monitoring stations (Pista Silla) the best forecast of nitrogen dioxide were obtained when including as input variables meteorology, traffic, seasonality and $(\text{NO}_2)_t$ level. At the other site (Viveros) the introduction of seasonality improve the performance of the multilayer perceptron when predicting nitrogen dioxide in terms of the root mean square error.

The relative importance of meteorological and vehicle emission variables on surface nitrogen oxides predictions is of great interest to establish the legislative measures that permit to reduce their levels. Models with different architectures have been considered. They have allowed predicting nitrogen oxides concentrations with accuracy. Mechanisms involved in nitrogen oxides concentration are complex and

non-linear. Neural networks do not require very exhaustive information about air pollutants, reaction mechanisms, meteorological parameters or traffic flow. They have also the advantage of allowing nonlinear relationships between very different inputs or predictor variables.

The application of the models could be extended to forecast other pollutants hourly levels in the study area. The predictor variables of this work represent an excellent starting point, but models that consider other inputs should be taken into account for future work. The results support other studies done in other parts of the world using artificial neural networks (ANN) techniques (Ibarra-Berastegi et al. 2008, Elangasinghe et al. 2014). It is the first time that this methodology is applied to Valencia urban area to predict nitrogen oxides concentrations. If the models were to be used as an operational air quality forecast models, forecast traffic data would also be required. These forecast traffic could be obtained by applying ANN using as inputs seasonal components. Forecast traffic values can be used as models' input and might improve their performance. A comparison of the models using representative values and forecast traffic data would be useful to evaluate their effects. Future research will also focus on the development of other neural network models for atmospheric pollutants prediction. Finally, the methodology developed in this study could also be applied in different areas with similar air quality problems and meteorological and traffic influences.

Acknowledgements

The author is indebted to the referees for their constructive comments, which have contributed to improving this article.

References

- Arhami, M., Kamali, N., and Rajabi, M.M. (2013). Predicting hourly air pollutant levels using artificial networks coupled with uncertainty analysis by Monte Carlo simulations. *Environ Sci Pollut Res.* 5, 4777.
- Ballester, F., Tenias, J.M., and Pérez-Hoyos, S. (2001). Air pollution and emergency hospital admissions for cardiovascular diseases in Valencia, Spain. *J Epidemiol Commun Health.* 55, 57.

- Ballester, F., Iñíguez, C., and García, F. (2005). *ENHIS-1 Project: WP5 Health Impact Assessment. Local City Report Valencia*. Available at <http://www.apheis.org/CityReports2005/Valencia.pdf>. Date of access April 2006.
- Ballester, F., Rodríguez, P., Iñíguez, C., Sáez, M., Daponte, A. et al. (2006). Air pollution and cardiovascular admissions association in Spain: result within the EMECAS project. *J Epidemiol Commun Health*. 60, 328.
- Bishop, C.M. (1995). *Neural Networks for Pattern Recognition*. Oxford, UK: Clarendon Press.
- Box, G.E.P., Jenkins, G.M., and Reinsel, G.G. (2008). *Time Series Analysis. Forecasting and Control*, 4th edition. New Jersey, USA: John Wiley & Sons.
- Caselli, M., Trizio, L., De Gennaro, G., and Ielpo, P. (2009). A simple feedforward neural network for the PM₁₀ forecasting: Comparison with a radial basis function networks and a multivariate linear regression model. *Water, Air Soil Poll*. 201, 365.
- Donnelly, A., Misstear, B., and Broderick, B. (2015). Real time air quality forecasting using integrated parametric and non-parametric regression techniques. *Atmos Environ*. 103, 53.
- Elangasinghe, M.A., Singhal, N., Dirks, K.N. and Salmond, J.A. (2014). Development of an ANN-based air pollution forecasting system with explicit knowledge through sensitivity analysis. *Atmos Pol Res*. 5, 696.
- European Communities (2007). *Europe's Environment: the Fourth Assessment*. European Environment Agency, Office for Official Publications of the European Community, Luxembourg
- Gardner M.W., and Dorling S.R. (1998). Artificial neural networks (the multilayer perceptron) - a review of applications in the atmospheric sciences. *Atmos Environ*. 32(14-15), 2627-2636.
- Gardner, M.W., and Dorling, S.R. (2000). Statistical surface ozone models: an improved methodology to account for non-linear behaviour. *Atmos Environ*. 34, 21.
- Gomez-Sanchis, J., Martin-Guerrero, J.D., Soria-Olivas, E., Vila-Francés, J., Carrasco, J.L., and del Valle-Tascón, S. (2006). Neural networks for analysing the relevance of input variables in the prediction of tropospheric ozone concentration. *Atmos Environ*. 40, 6173.
- Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*. 2, 359.

- Ibarra-Berastegi, G., Elias, A., Barona, A., Saenz, J., Ezcurra, A., and Diaz de Argandoña, J. (2008). From diagnosis to prognosis for forecasting air pollution using neural networks: Air pollution monitoring in Bilbao. *Environ Modell Softw.* 23, 622.
- Kocak, K., Saylan, L., and Sen, L. (2000). Nonlinear times series prediction of ozone concentration in Istanbul. *Atmos Environ.* 34, 1267.
- Moller, M.F., (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks.* 6, 525.
- Prada-Sanchez, J.M., Febrero-Bande, M., Cotos-Yanez, T., Gonzalez-Manteiga, W., Bermudez-Cela, L., and Lucas-Dominguez, T. (2000). Prediction of SO₂ pollution incident near a power station using partially linear models and an historical matrix of predictor-response vectors. *Environmetrics.* 11, 209.
- Ryan, W.F. (1995). Forecasting ozone episodes in the Baltimore metropolitan area. *Atmos Environ.* 29, 2387.
- Rye, P.J. (1995). Modelling photochemical smog in the Perth region. *Math Comput Model.* 21,111.
- Sarle, W.S. (1995). Stopped training and other remedies for overfitting. *Proceedings of the 27th Symposium on the Interface*, Fairfax, VA.
- Senger, S.U. (2000). Estimation of the reduction of the concentration of air pollutants until the year 2010 in relation to the UN ECEE Göteborg protocol dated 1. December 1999. *Umweltwiss Schadst Forsch.* 12, 152.
- Shalkoff, R. (1992). *Pattern Recognition: Statistical Structural and Neural Approaches*. New York: John Wiley & Sons.
- Tenias, J.M., Ballester, F., and Rivera, M.L. (1998). Association between hospital medical emergency visits for asthma and air pollution in Valencia, Spain. *Occup Environ Med.* 55, 541.
- Willmott, C., Ackleson, S., Davis, R., Feddema, J., Klink, K., Legates, D., O'Donnell, J., and Rowe, C., (1985). Statistics for the evaluation and comparison of models. *J Geophys Res.* 90, 8995.

Table 1 Mean, coefficient of variation (CV) and maximum values of pollution, meteorological and traffic data for the two monitoring stations. Nitrogen dioxide (NO₂) and nitric oxide (NO) are expressed in µg/m³. Wind speed: WS in m/s. Wind direction: WD in degrees. Temperature: T in °C. Solar radiation: SR in W/m². Relative humidity: RH in %. Pressure: P in mbar. Number of vehicles circulating every hour: NV.

Station	Variable	Average	CV	Minimum	Maximum
P.Silla	NO ₂	58.8	0.51		249
	NO	52.0	1.14		624
	WS	1.1	0.82		8.6
	WD	187.8	0.57		360
	T	18.7	0.36		38.2
	RH	60.8	0.25		92
	P	1022.2	0.01		1044.7
	SR	153.3	1.60		947
Viveros	NV	2945.7	0.57		38712
	NO ₂	36.72	0.67		238
	NO	19.6	1.93		596
	WS	1.8	0.72		11.9
	WD	168.3	0.69		360
	T	18.9	0.34		38.2
	SR	170	1.53		1033.3
	NV	1088.7	0.66		13456

Table 2 Models analyzed in the paper

Model	Output variable	Input variables
MP1	(NO ₂) _{t+24}	Meteorology _t , traffic _t , (NO ₂) _t
MP2	(NO ₂) _{t+24}	Meteorology _t , traffic _t , Seasonality _{t+24} , (NO ₂) _t
MP3	(NO ₂) _{t+24}	Meteorology _t , traffic _t , (NO ₂) _t , NO _t
MP4	(NO ₂) _{t+24}	Meteorology _t , traffic _t , Seasonality _{t+24} , (NO ₂) _t , NO _t
MP5	NO _{t+24}	Meteorology _t , traffic _t , NO _t
MP6	NO _{t+24}	Meteorology _t , traffic _t , Seasonality _{t+24} , NO _t
MP7	NO _{t+24}	Meteorology _t , traffic _t , NO _t , (NO ₂) _t
MP8	NO _{t+24}	Meteorology _t , traffic _t , Seasonality _{t+24} , NO _t , (NO ₂) _t

Table 3 Best performance criteria results at Pista Silla for the models MP1, MP2, MP3 and MP4 (output (NO₂)_{t+24}) and models MP5, MP6, MP7 and MP8 (output NO_{t+24})

Model	Transfer function	n _h	Learning algorithms	RMSE	MAE	MAPE	r
MP1	tansig	16	LM	20.29	16.39	0.51	0.59
MP2	tansig	14	LM	19.35	15.38	0.45	0.63
MP3	logsig	30	LM	20.48	16.32	0.48	0.56
MP4	logsig	10	LM	20.49	16.55	0.50	0.65
MP5	logsig	7	LM	45.07	27.85	1.31	0.61
MP6	logsig	10	LM	42.37	26.48	1.19	0.66
MP7	logsig	10	LM	45.26	28.39	1.42	0.59
MP8	logsig	7	LM	42.57	25.57	1.13	0.66

Table 4 Best performance criteria results at Viveros for the models MP1, MP2, MP3 and MP4 (output $(NO_2)_{t+24}$) and models MP5, MP6, MP7 and MP8 (output NO_{t+24})

Model	Transfer function	n_h	Learning algorithms	RMSE	MAE	MAPE	r
MP1	logsig	5	SCG	23.76	18.16	0.93	0.03
MP2	tansig	10	SCG	23.05	18.31	0.98	0.07
MP3	logsig	16	SCG	23.28	18.41	0.93	0.15
MP4	logsig	12	SCG	23.32	18.31	0.91	0.05
MP5	logsig	5	LM	33.78	19.86	3.52	0.49
MP6	logsig	16	LM	33.19	19.03	3.16	0.52
MP7	tansig	10	LM	33.94	19.26	3.20	0.49
MP8	tansig	10	LM	33.10	19.10	3.32	0.52

FIG. 1 Map of the study monitoring sites' locations.



FIG. 2 Average diurnal cycles at Pista Silla station

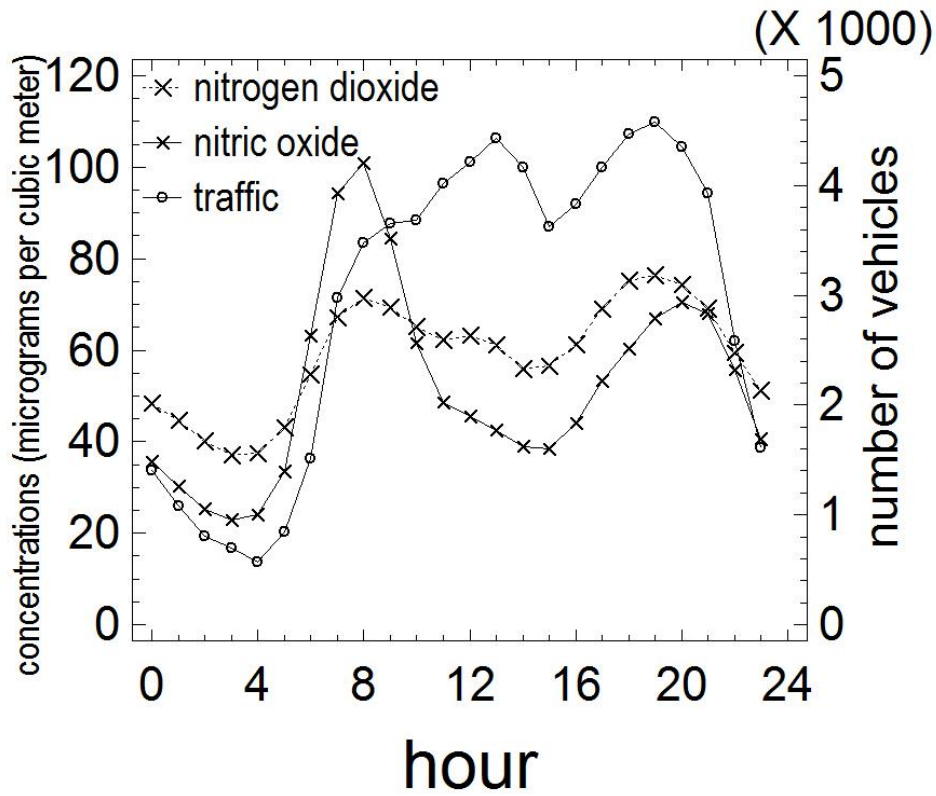


FIG. 3 Average weekly cycles at Pista Silla station

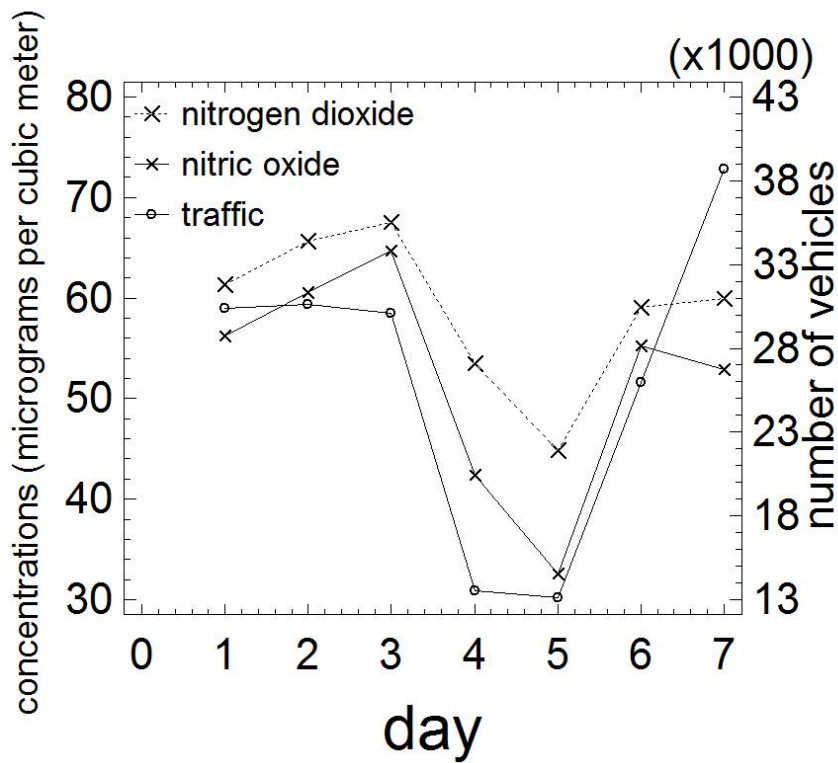


FIG. 4 A multilayer perceptron model

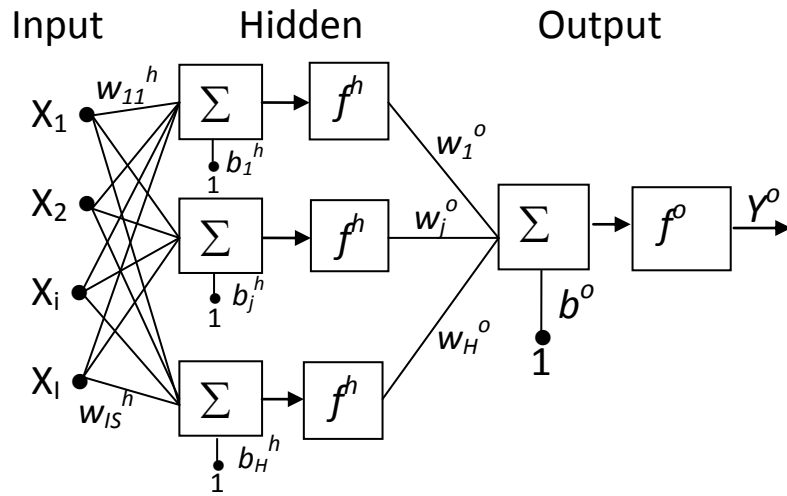


FIG. 5 Nitrogen dioxide prediction 24 hours in advance in the first 100 hours of the test data set at Pista Silla. The model is the multilayer perceptron including $(NO_2)_t$, meteorology, traffic and seasonality as predictors, with 14 neurons in the hidden layer, the tansig transfer function and the Levenberg-Marquard backpropagation algorithm. Correlation coefficient $r=0.63$ and mean absolute error $MAE=15.38$

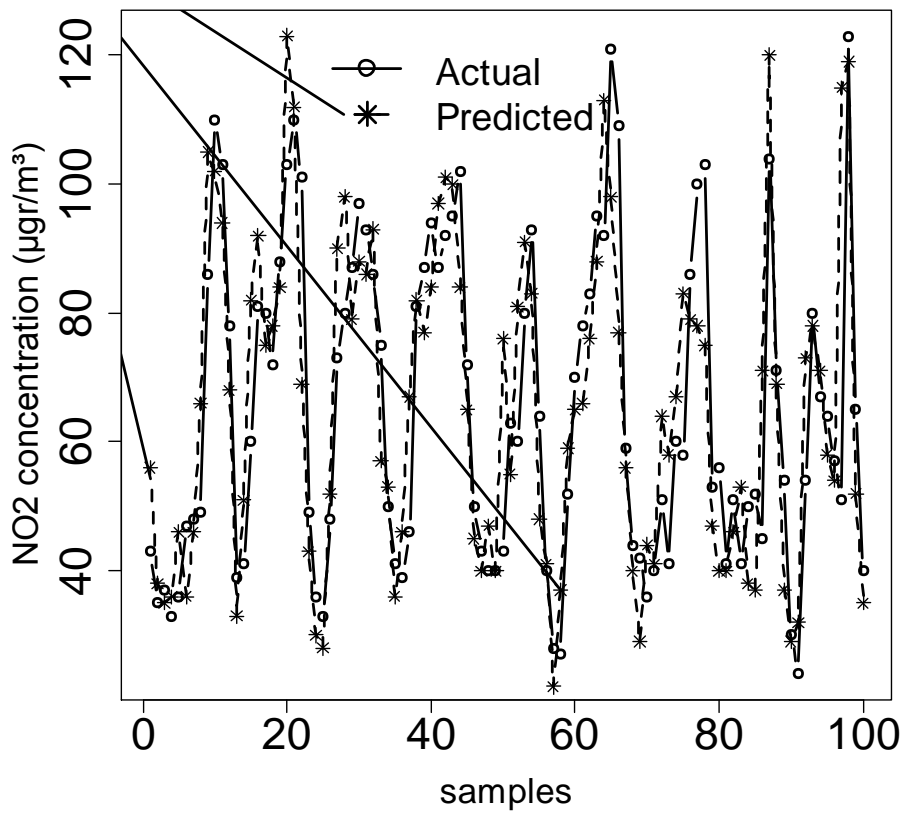


FIG 6 Scatterplot of nitrogen dioxide predictions 24 hours in advance versus observations at Pista Silla. The model is the multilayer perceptron including $(NO_2)_t$, meteorology, traffic and seasonality as predictors, with 14 neurons in the hidden layer, the tansig transfer function and the Levenberg-Marquard backpropagation algorithm

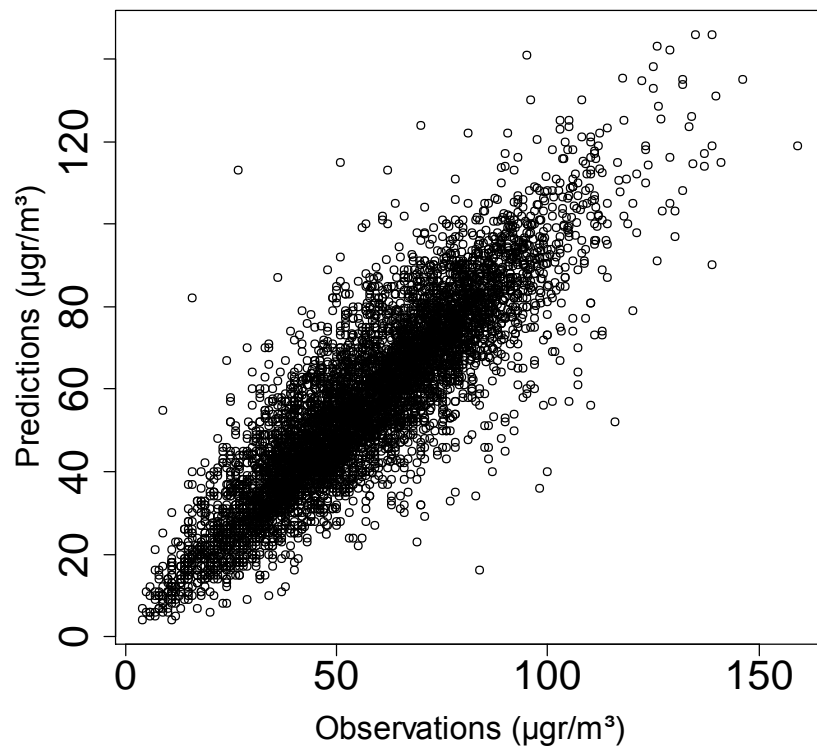


FIG. 7 Nitric oxide prediction 24 hours in advance in the first 100 hours of the test data set at Pista Silla. The model is the multilayer perceptron including NO_t , $(\text{NO}_2)_t$, meteorology, traffic and seasonality as predictors, with 7 neurons in the hidden layer, the logsig transfer function and the Levenberg-Marquard backpropagation algorithm. Correlation coefficient $r=0.66$ and mean absolute error $\text{MAE}=25.57$

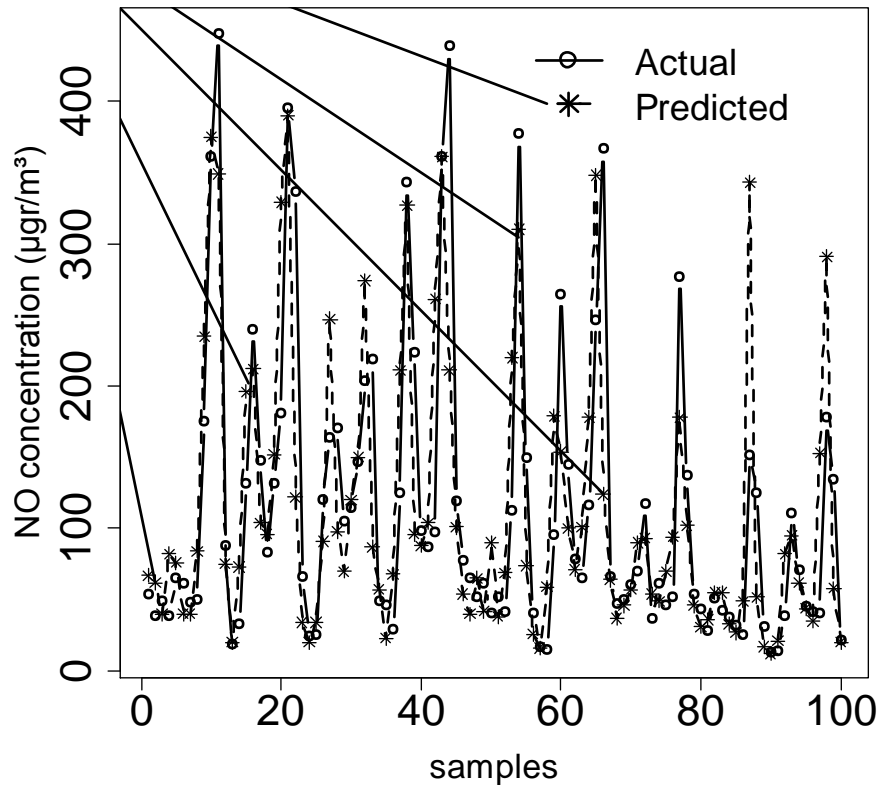


FIG. 8 Nitrogen dioxide prediction 24 hours in advance in the first 100 hours of the test data set at Viveros. The model is the multilayer perceptron including $(NO_2)_t$, meteorology, seasonal components, and traffic as predictors, with 10 neurons in the hidden layer, the tansig transfer function and the Scaled Conjugate Gradient backpropagation algorithm. Correlation coefficient $r=0.07$ and mean absolute error $MAE=18.31$

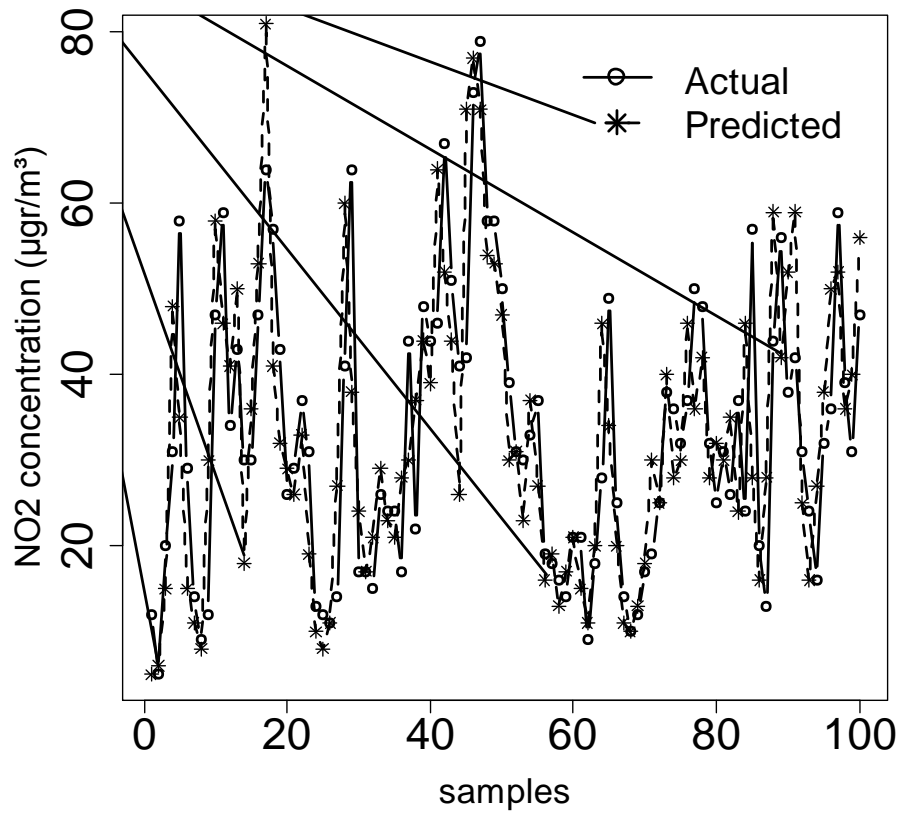
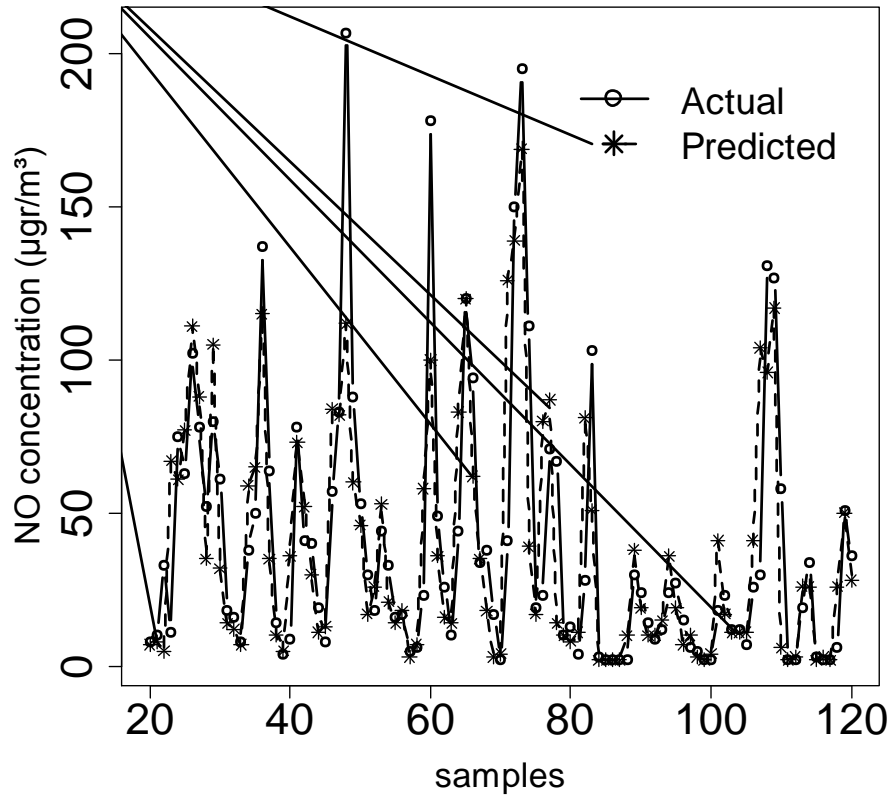


FIG. 9 Nitric oxide prediction 24 hours in advance in the first 100 hours with complete records of the test dataset at Viveros. The model is the multilayer perceptron including NO_t , $(\text{NO}_2)_t$, meteorology, traffic and seasonality as predictors, with 10 neurons in the hidden layer, the tansig transfer function and the Levenberg-Marquard backpropagation algorithm. Correlation coefficient $r=0.52$ and mean absolute error $\text{MAE}=19.1$



Supplementary Table S1 Performance criteria results at Pista Silla for models MP1, MP2 and MP3 (output NO₂)

Model	Transfer Function	n _h	SCG algorithm				LM algorithm			
			RMSE	MAE	MAPE	r	RMSE	MAE	MAPE	r
MP1	Tansig	5	22.68	18.58	0.60	0.49	20.96	16.51	0.53	0.58
		7	22.15	18.22	0.59	0.51	21.95	17.43	0.56	0.55
		10	20.83	16.96	0.54	0.59	21.23	17.26	0.53	0.57
		12	23.24	19.02	0.61	0.41	23.25	19.12	0.63	0.55
		14	21.67	17.57	0.54	0.55	20.30	16.49	0.51	0.59
		16	22.52	18.34	0.59	0.45	20.29	16.39	0.51	0.59
		18	21.44	17.59	0.56	0.55	21.42	17.52	0.56	0.57
		20	23.82	19.28	0.61	0.49	21.92	17.65	0.54	0.56
		30	22.83	18.59	0.58	0.47	21.22	17.13	0.53	0.55
	logsig	5	22.88	18.86	0.62	0.52	21.28	17.29	0.54	0.59
		7	23.46	19.27	0.62	0.48	20.84	17.14	0.55	0.58
		10	21.93	17.98	0.56	0.51	21.66	17.62	0.56	0.58
		12	24.06	20.02	0.67	0.44	20.64	16.71	0.52	0.59
		14	23.52	19.14	0.59	0.54	20.58	16.68	0.53	0.58
		16	22.12	18.04	0.58	0.51	21.88	17.74	0.56	0.58
		18	23.04	18.82	0.59	0.42	22.52	18.21	0.56	0.52
		20	22.31	18.40	0.59	0.55	21.94	17.89	0.57	0.57
		30	26.23	21.35	0.67	0.39	21.63	17.66	0.56	0.58
MP2	tansig	5	21.91	17.94	0.57	0.64	20.45	16.52	0.51	0.56
		7	24.45	20.32	0.68	0.50	20.31	16.43	0.51	0.65
		10	21.35	17.39	0.55	0.62	21.19	17.12	0.52	0.66
		12	20.62	16.77	0.52	0.63	20.75	16.69	0.51	0.64
		14	22.23	17.95	0.57	0.57	19.35	15.38	0.45	0.63
		16	21.11	16.99	0.52	0.63	20.89	16.87	0.51	0.63
		18	21.58	17.55	0.54	0.62	21.29	17.18	0.52	0.64
		20	21.38	17.28	0.53	0.63	20.52	16.52	0.51	0.64
		30	22.07	17.94	0.56	0.65	23.43	28.86	0.58	0.59
	logsig	5	26.94	22.42	0.76	0.24	20.69	16.66	0.52	0.65
		7	21.99	17.92	0.56	0.61	20.70	16.66	0.49	0.66
		10	24.84	20.74	0.69	0.42	21.38	17.39	0.54	0.59
		12	21.94	18.73	0.58	0.59	20.71	16.64	0.51	0.66
		14	22.25	18.01	0.56	0.58	22.20	18.03	0.57	0.63
		16	22.51	18.49	0.58	0.57	21.28	17.13	0.51	0.63
		18	24.48	20.29	0.66	0.53	20.61	16.70	0.51	0.59
		20	21.89	17.91	0.55	0.63	21.24	17.10	0.52	0.63
		30	23.34	19.03	0.59	0.54	23.22	28.79	0.56	0.62
MP3	tansig	5	21.44	17.51	0.54	0.56	21.72	17.84	0.58	0.57
		7	21.95	18.03	0.59	0.59	20.44	16.39	0.49	0.53
		10	21.54	17.47	0.54	0.52	21.94	17.39	0.56	0.58
		12	21.60	17.69	0.57	0.55	22.08	17.92	0.56	0.55
		14	21.61	17.66	0.56	0.57	20.81	16.82	0.51	0.54
		16	21.46	17.51	0.55	0.54	22.09	17.93	0.56	0.55
		18	21.77	17.86	0.57	0.53	22.39	18.52	0.61	0.58
		20	21.53	17.46	0.55	0.56	23.33	19.15	0.62	0.54
		30	27.09	21.81	0.64	0.34	22.99	18.56	0.57	0.55
	logsig	5	23.27	18.15	0.59	0.54	21.19	17.27	0.55	0.59
		7	23.79	19.61	0.63	0.51	21.03	17.03	0.53	0.56
		10	22.76	18.66	0.61	0.52	22.01	17.82	0.57	0.57
		12	20.90	16.89	0.52	0.59	20.51	16.51	0.49	0.57
		14	23.01	18.71	0.59	0.55	21.84	18.03	0.59	0.58
		16	23.89	19.40	0.61	0.50	22.96	18.96	0.61	0.56
		18	24.31	20.02	0.67	0.52	22.49	18.29	0.57	0.53
		20	26.70	21.57	0.64	0.49	21.15	17.14	0.53	0.56
		30	22.65	18.09	0.56	0.52	20.48	16.32	0.48	0.56

Supplementary Table S2 Performance criteria results at Pista Silla for models MP4 (output NO₂), MP5 and MP6 (output NO)

Model	Transfer function	n _h	SCG algorithm				LM algorithm				
			RMSE	MAE	MAPE	r	RMSE	MAE	MAPE	r	
MP4	tansig	5	21.08	17.13	0.54	0.59	21.42	17.28	0.54	0.59	
		7	21.65	17.42	0.52	0.59	20.83	16.79	0.52	0.66	
		10	21.56	17.43	0.53	0.63	20.58	16.54	0.51	0.66	
		12	22.36	18.06	0.55	0.59	21.28	17.28	0.51	0.62	
		14	20.95	17.03	0.54	0.65	21.41	17.27	0.53	0.65	
		16	21.55	17.29	0.52	0.65	22.86	18.70	0.58	0.58	
		18	21.22	17.06	0.52	0.64	21.83	17.64	0.55	0.59	
		20	23.37	19.19	0.61	0.60	25.66	20.32	0.54	0.55	
		30	21.63	27.36	0.53	0.63	21.75	17.69	0.56	0.61	
		logsig	5	23.27	18.15	0.59	0.54	20.67	26.68	0.52	0.63
	7		23.79	19.61	0.63	0.51	20.77	16.75	0.51	0.65	
	10		22.76	18.66	0.61	0.52	20.49	16.55	0.50	0.65	
	12		20.91	16.89	0.52	0.59	21.88	17.78	0.55	0.63	
	14		23.01	18.71	0.59	0.55	21.45	17.35	0.54	0.62	
	16		23.89	19.40	0.61	0.50	22.63	18.29	0.58	0.61	
	18		24.31	20.02	0.67	0.52	21.75	17.53	0.56	0.63	
	20		26.70	21.57	0.64	0.49	20.93	16.77	0.51	0.63	
	30		22.65	18.09	0.56	0.52	22.25	17.99	0.57	0.61	
	MP5		tansig	5	46.48	29.59	1.56	0.57	45.46	28.43	1.41
		7		52.73	36.55	2.32	0.38	45.21	27.87	1.34	0.60
10		47.95		30.57	1.52	0.53	46.59	30.66	1.73	0.57	
12		46.58		30.25	1.56	0.57	46.05	29.54	1.52	0.59	
14		49.32		33.39	1.82	0.51	45.49	29.19	1.49	0.59	
16		48.87		31.86	1.71	0.51	45.67	30.27	1.64	0.59	
18		47.75		31.05	1.67	0.54	45.52	29.28	1.52	0.59	
20		46.67		29.16	1.52	0.57	46.31	30.78	1.66	0.59	
30		50.08		34.22	1.92	0.51	46.08	29.19	1.38	0.58	
logsig		5		51.42	33.89	2.02	0.42	46.78	32.54	1.91	0.59
		7	50.02	34.31	2.08	0.48	45.07	27.85	1.31	0.61	
		10	49.09	33.32	1.89	0.51	45.12	27.80	1.31	0.61	
		12	46.34	29.45	1.58	0.57	45.94	27.67	1.27	0.59	
		14	48.10	31.69	1.72	0.53	48.78	30.37	1.64	0.53	
		16	49.82	34.05	1.92	0.48	46.07	30.02	1.63	0.59	
		18	48.33	31.58	1.76	0.53	45.81	29.69	1.47	0.59	
		20	51.34	36.48	1.77	0.49	45.29	29.30	1.44	0.60	
		30	55.27	37.74	1.89	0.36	46.12	29.92	1.54	0.59	
		MP6	tansig	5	45.28	28.36	1.43	0.60	43.87	27.39	1.28
7				45.80	29.52	1.58	0.59	43.55	27.18	1.23	0.64
10	46.35			30.02	1.59	0.58	44.16	28.42	1.54	0.63	
12	45.75			28.84	1.42	0.59	43.27	26.98	1.19	0.65	
14	54.33			37.55	2.02	0.34	44.51	28.15	1.36	0.62	
16	44.51			27.49	1.40	0.62	43.63	27.78	1.29	0.64	
18	46.07			25.47	1.53	0.58	45.07	30.13	1.59	0.63	
20	63.14			44.51	2.21	0.30	45.49	29.40	1.54	0.62	
30	46.32			30.47	1.59	0.58	45.82	29.88	1.45	0.59	
logsig	5			55.93	36.19	2.08	0.17	44.66	27.05	1.29	0.62
	7		46.01	29.00	1.45	0.58	44.88	27.92	1.36	0.61	
	10		47.19	30.18	1.55	0.56	42.37	26.48	1.19	0.66	
	12		47.19	30.18	1.55	0.56	43.28	27.01	1.24	0.64	
	14		45.96	30.28	1.52	0.59	43.28	27.60	1.32	0.65	
	16		48.94	31.51	1.73	0.50	43.36	27.21	1.28	0.64	
	18		47.63	30.79	1.56	0.54	45.13	30.98	1.75	0.64	
	20		47.05	31.03	1.57	0.56	44.18	27.45	1.31	0.62	
	30		46.30	29.49	1.49	0.57	45.14	29.60	1.50	0.61	

Supplementary Table S3 Performance criteria results at Pista Silla for models MP7 and MP8 (output NO)

Model	Transfer function	n_h	SCG algorithm				LM algorithm				
			RMSE	MAE	MAPE	r	RMSE	MAE	MAPE	r	
MP7	tansig	5	46.29	28.12	1.52	0.58	45.64	27.89	1.34	0.59	
		7	49.39	32.82	1.80	0.49	46.52	31.00	1.76	0.58	
		10	46.24	29.54	1.46	0.58	45.66	27.93	1.24	0.59	
		12	55.61	35.39	1.88	0.25	45.92	29.22	1.36	0.59	
		14	49.07	32.29	1.68	0.50	45.84	28.45	1.34	0.59	
		16	47.61	29.62	1.48	0.55	47.15	31.67	1.75	0.57	
		18	47.16	30.42	1.59	0.55	49.88	33.69	1.99	0.52	
		20	49.83	31.07	1.55	0.48	46.09	28.23	1.23	0.58	
		30	49.26	32.40	1.66	0.50	46.79	31.23	1.60	0.58	
		logsig	5	48.69	30.59	1.57	0.51	46.76	29.94	1.61	0.59
	7		51.36	32.19	1.61	0.42	45.47	27.63	1.27	0.59	
	10		46.69	28.83	1.34	0.57	45.26	28.39	1.42	0.59	
	12		48.29	29.93	1.50	0.53	47.44	28.74	1.39	0.58	
	14		48.98	32.18	1.73	0.50	48.53	34.46	2.08	0.56	
	16		51.55	34.71	1.90	0.50	45.82	28.46	1.42	0.59	
	18		47.00	29.13	1.36	0.56	46.88	29.29	1.36	0.57	
	20		50.53	32.05	1.35	0.48	48.45	35.38	2.09	0.59	
	30		55.87	36.76	1.83	0.29	46.76	31.05	1.62	0.59	
	MP8		tansig	5	47.47	30.32	1.58	0.54	42.53	26.45	1.21
		7		52.96	34.14	1.96	0.35	43.19	26.70	1.22	0.65
10		48.23		31.07	1.65	0.52	47.27	33.00	1.77	0.58	
12		45.16		28.24	1.33	0.60	43.05	26.56	1.19	0.65	
14		45.48		27.75	1.29	0.59	43.42	27.02	1.18	0.65	
16		43.36		26.76	1.22	0.64	43.54	27.68	1.22	0.65	
18		48.83		32.79	1.64	0.48	46.31	31.31	1.63	0.61	
20		47.37		30.26	1.46	0.55	44.81	28.28	1.32	0.63	
30		46.56		28.73	1.39	0.57	44.56	27.22	1.17	0.62	
logsig		5		48.30	32.62	1.84	0.53	43.38	26.66	1.19	0.64
		7	50.30	32.98	1.83	0.47	42.57	25.57	1.13	0.66	
		10	48.09	29.19	1.45	0.54	42.86	26.58	1.18	0.65	
		12	49.61	32.71	1.78	0.48	45.14	28.28	1.29	0.60	
		14	49.58	31.39	1.59	0.48	45.04	27.59	1.18	0.61	
		16	48.97	32.44	1.81	0.51	43.59	26.37	1.14	0.64	
		18	48.29	29.75	1.48	0.52	43.22	27.35	1.31	0.65	
		20	51.24	34.60	2.01	0.45	43.38	25.97	1.10	0.65	
		30	44.63	27.70	1.33	0.61	44.34	27.78	1.25	0.62	

Supplementary Table S4 Performance criteria results at Viveros for models MP1, MP2 and MP3 (output NO₂)

Model	Transfer function	n _h	SCG				LM				
			RMSE	MAE	MAPE	r	RMSE	MAE	MAPE	r	
MP1	tansig	5	23.76	18.76	0.99	0.12	24.28	19.06	0.99	0.11	
		7	23.78	18.59	0.94	0.10	23.39	18.34	0.93	0.12	
		10	23.25	18.65	0.99	0.03	23.62	18.54	0.96	0.11	
		12	24.29	19.14	1.02	0.09	24.34	19.03	0.96	0.11	
		14	23.72	18.85	0.98	0.09	24.39	19.03	0.94	0.10	
		16	24.51	19.33	1.01	0.06	24.82	19.34	0.94	0.10	
		18	24.54	19.29	0.99	0.08	24.07	18.82	0.94	0.11	
		20	24.68	19.46	0.99	0.09	24.59	19.18	0.97	0.08	
		30	25.01	19.51	0.98	0.07	24.68	19.18	0.97	0.08	
		logsig	5	23.76	18.16	0.93	0.03	24.35	19.14	1.01	0.08
	7		23.29	18.49	0.98	0.06	24.66	19.29	0.99	0.11	
	10		25.33	20.11	1.09	0.02	24.29	19.02	0.97	0.07	
	12		25.66	20.39	1.06	0.02	23.82	18.69	0.95	0.11	
	14		25.18	19.96	1.02	0.01	23.75	18.62	0.95	0.11	
	16		26.19	20.61	1.04	0.03	24.85	19.42	0.98	0.07	
	18		25.51	20.03	1.03	0.03	23.97	18.81	0.95	0.12	
	20		24.92	19.62	0.99	0.09	24.27	19.14	1.06	0.13	
	30		25.51	20.31	1.05	0.08	28.29	20.49	0.96	0.09	
	MP2		tansig	5	23.46	18.54	0.98	0.04	24.74	19.46	1.02
		7		24.12	19.13	0.98	0.03	24.27	19.07	1.03	0.13
10		23.05		18.31	0.98	0.07	24.49	19.22	0.99	0.10	
12		24.25		19.23	0.97	0.06	25.17	19.65	1.01	0.09	
14		24.74		19.44	1.02	0.06	24.72	19.36	0.98	0.14	
16		24.09		18.96	0.98	0.12	24.23	18.86	0.95	0.12	
18		24.08		18.93	0.99	0.07	24.79	19.33	0.98	0.11	
20		24.79		19.59	1.04	0.10	24.16	19.93	0.98	0.11	
30		24.56		19.32	0.99	0.09	25.99	20.36	1.08	0.09	
logsig		5		23.65	18.81	0.99	0.06	23.67	18.49	0.97	0.14
		7	24.01	19.10	0.96	0.03	25.14	19.64	1.01	0.09	
		10	24.48	19.25	0.96	0.05	24.19	18.94	0.98	0.13	
		12	24.74	19.48	0.96	0.06	24.94	19.53	1.00	0.11	
		14	23.98	19.11	0.99	0.05	24.42	19.07	0.96	0.11	
		16	23.69	18.65	0.96	0.12	25.02	19.72	1.06	0.10	
		18	25.46	20.00	1.04	0.02	24.67	19.33	1.03	0.09	
		20	23.66	18.77	0.99	0.08	25.33	19.72	0.98	0.09	
		30	24.99	19.65	0.99	0.09	25.04	19.64	0.97	0.13	
		MP3	tansig	5	23.76	18.78	0.98	0.01	24.22	18.98	0.98
7				23.38	18.57	0.97	0.05	24.04	18.73	0.96	0.11
10	24.63			19.38	1.02	0.08	24.26	19.07	1.00	0.08	
12	24.55			19.22	0.99	0.07	23.64	18.54	0.92	0.12	
14	24.43			19.29	1.01	0.07	24.28	19.01	0.98	0.12	
16	24.39			19.21	0.98	0.08	24.60	19.16	0.92	0.09	
18	23.58			18.58	0.98	0.09	24.30	19.14	1.01	0.09	
20	24.44			19.26	0.93	0.12	24.22	19.14	1.01	0.11	
30	24.59			19.09	0.96	0.10	24.12	18.85	0.87	0.14	
logsig	5			24.83	19.65	1.03	0.03	24.48	19.32	1.03	0.06
	7		26.31	21.06	1.11	0.01	24.58	19.22	0.98	0.08	
	10		25.21	20.02	1.05	0.02	24.18	19.02	0.96	0.08	
	12		24.12	18.98	0.96	0.08	24.27	19.07	0.99	0.09	
	14		23.36	18.25	0.91	0.12	24.17	18.82	0.94	0.11	
	16		23.28	18.41	0.93	0.15	24.12	18.84	0.96	0.10	
	18		23.88	18.88	0.97	0.08	24.48	19.13	0.99	0.08	
	20		26.17	20.49	1.06	0.02	24.34	19.13	0.99	0.08	
	30		24.76	19.48	0.97	0.11	24.86	19.39	0.95	0.08	

Supplementary Table S5 Performance criteria results at Viveros for models MP4 (output NO₂), MP5 and MP6 (output NO)

Model	Transfer function	n _h	SCG algorithm				LM algorithm				
			RMSE	MAE	MAPE	r	RMSE	MAE	MAPE	r	
MP4	tansig	5	24.43	19.66	1.01	0.08	25.28	19.82	1.04	0.06	
		7	23.76	18.97	1.01	0.06	24.91	19.57	1.01	0.09	
		10	25.34	19.83	1.02	0.05	24.79	19.38	0.96	0.09	
		12	23.80	18.84	1.00	0.01	24.54	19.21	0.99	0.06	
		14	24.39	19.25	0.99	0.06	24.38	19.14	0.99	0.13	
		16	24.57	19.43	1.18	0.05	24.31	19.07	0.99	0.10	
		18	24.89	19.84	1.00	0.03	25.15	19.64	0.97	0.10	
		20	24.43	19.27	0.99	0.08	24.83	19.43	0.98	0.10	
		30	25.95	20.33	1.05	0.11	24.45	19.13	0.99	0.12	
		logsig	5	24.56	19.56	1.04	0.03	24.12	18.94	0.99	0.14
	7		23.40	18.77	0.97	0.01	24.34	19.18	1.03	0.25	
	10		26.85	20.77	1.09	0.07	24.94	19.63	1.04	0.09	
	12		23.32	18.31	0.91	0.05	25.32	19.86	1.01	0.06	
	14		25.26	19.95	1.05	0.02	24.72	19.48	1.02	0.07	
	16		24.79	19.74	1.02	0.05	24.75	19.55	1.01	0.09	
	18		24.26	19.06	1.01	0.07	25.66	19.83	0.98	0.08	
	20		23.91	19.02	1.02	0.11	25.16	19.74	1.02	0.09	
	30		24.89	19.74	1.01	0.11	25.19	19.58	0.96	0.11	
	MP5		tansig	5	35.57	22.14	4.33	0.41	34.11	19.99	3.58
		7		34.94	21.90	4.22	0.45	33.81	19.37	3.32	0.49
10		35.10		21.24	4.12	0.44	34.66	21.84	4.29	0.48	
12		35.25		21.74	4.21	0.43	33.86	20.01	3.56	0.49	
14		34.19		21.08	4.05	0.48	33.81	19.83	3.53	0.49	
16		36.00		23.02	4.53	0.41	34.61	21.38	4.05	0.49	
18		35.39		21.32	3.92	0.41	34.51	19.71	3.34	0.48	
20		36.31		22.40	4.22	0.43	35.58	20.34	3.47	0.47	
30		37.59		24.48	4.98	0.36	35.66	22.63	4.49	0.47	
logsig		5		35.71	21.47	4.15	0.39	33.78	19.86	3.52	0.49
		7	35.15	22.10	4.37	0.43	33.96	19.79	3.35	0.49	
		10	35.15	22.56	4.51	0.44	35.87	24.13	3.30	0.44	
		12	38.09	22.79	4.16	0.39	33.99	19.90	3.47	0.49	
		14	36.04	22.11	4.24	0.38	34.42	20.04	3.63	0.47	
		16	43.36	31.67	4.39	0.39	33.88	19.55	3.34	0.49	
		18	39.03	24.60	4.76	0.33	34.46	19.73	3.36	0.48	
		20	36.58	24.16	4.93	0.38	34.36	20.79	2.89	0.48	
		30	35.49	21.91	4.07	0.42	35.68	21.25	3.73	0.47	
		MP6	tansig	5	35.64	21.51	4.06	0.40	33.84	19.09	3.14
7				34.83	21.34	3.93	0.46	33.43	20.39	3.86	0.52
10	34.86			21.43	4.09	0.45	33.47	19.62	3.46	0.51	
12	35.84			23.43	4.75	0.42	33.70	19.56	3.33	0.51	
14	36.49			21.55	3.75	0.34	33.26	18.75	3.03	0.52	
16	35.69			21.52	3.99	0.39	33.65	20.17	3.54	0.51	
18	35.59			21.44	3.90	0.40	33.60	20.66	3.94	0.51	
20	35.07			19.96	3.33	0.43	37.15	20.06	3.35	0.44	
30	33.88			20.64	3.80	0.49	33.99	19.92	3.45	0.49	
logsig	5			37.89	22.43	4.19	0.22	33.94	20.04	3.62	0.49
	7		37.19	21.59	3.75	0.29	33.37	19.90	3.59	0.53	
	10		38.40	23.05	3.89	0.28	33.56	19.61	3.36	0.51	
	12		35.93	21.70	4.04	0.38	33.45	19.75	3.90	0.53	
	14		35.33	22.40	4.28	0.42	33.63	19.25	3.14	0.51	
	16		35.25	21.40	4.01	0.42	33.19	19.03	3.16	0.52	
	18		34.39	21.55	4.12	0.48	33.63	19.91	3.54	0.50	
	20		37.46	24.54	4.91	0.31	34.03	20.75	3.81	0.49	
	30		34.19	20.50	3.68	0.48	34.58	19.73	3.29	0.46	

Supplementary Table S6 Performance criteria results at Viveros for models MP7 and MP8 (output NO)

Model	Transfer function	n_h	SCG algorithm				LM algorithm				
			RMSE	MAE	MAPE	r	RMSE	MAE	MAPE	r	
MP7	tansig	5	35.60	22.67	4.54	0.42	34.57	20.72	3.83	0.48	
		7	35.86	21.62	4.23	0.42	34.09	20.92	3.97	0.49	
		10	35.19	22.69	4.53	0.45	33.94	19.26	3.20	0.49	
		12	37.75	24.36	5.05	0.32	41.61	24.82	5.22	0.40	
		14	36.31	22.36	4.35	0.37	36.67	20.97	3.81	0.44	
		16	35.19	22.26	4.34	0.43	33.85	19.41	3.22	0.49	
		18	34.18	20.99	3.95	0.48	37.70	21.78	3.87	0.47	
		20	35.73	21.82	4.02	0.42	36.72	20.65	3.50	0.46	
		30	40.55	23.79	4.24	0.25	37.62	20.87	3.54	0.44	
		logsig	5	35.82	21.61	4.19	0.42	34.03	21.16	4.09	0.49
	7		37.81	22.11	3.97	0.23	34.16	21.35	4.17	0.49	
	10		35.51	22.12	4.13	0.41	35.95	21.84	4.14	0.48	
	12		37.87	24.54	4.97	0.29	34.45	20.60	3.72	0.49	
	14		36.72	24.09	4.13	0.37	36.92	21.04	3.62	0.47	
	16		34.94	22.05	4.33	0.45	38.84	21.77	3.86	0.46	
	18		34.47	21.96	4.31	0.47	34.89	19.33	3.22	0.47	
	20		34.94	21.60	4.15	0.45	34.43	20.21	3.49	0.49	
	30		34.75	20.92	3.87	0.45	35.72	20.29	3.43	0.48	
	MP8		tansig	5	37.06	21.21	3.80	0.30	33.64	19.67	3.47
		7		34.75	20.08	3.42	0.45	33.51	19.81	3.55	0.51
10		38.55		20.65	3.14	0.37	33.10	19.10	3.32	0.52	
12		36.29		22.88	4.52	0.36	33.28	19.50	3.51	0.52	
14		35.22		21.31	4.04	0.43	34.97	21.35	4.18	0.48	
16		34.04		20.79	3.92	0.49	33.99	20.38	3.68	0.49	
18		34.21		20.72	3.88	0.48	33.64	19.73	3.36	0.51	
20		36.12		22.91	4.56	0.39	33.51	20.70	3.89	0.52	
30		34.74		21.25	4.06	0.45	33.65	20.38	3.57	0.52	
logsig		5		37.35	20.88	3.52	0.30	33.54	19.80	3.52	0.51
		7	35.97	21.63	3.93	0.38	33.63	18.92	3.13	0.50	
		10	35.72	23.39	4.71	0.41	34.24	20.42	3.67	0.49	
		12	35.75	21.20	3.79	0.39	33.64	19.80	3.38	0.52	
		14	35.83	22.43	4.39	0.39	34.44	20.13	2.83	0.44	
		16	37.06	22.97	4.52	0.30	33.78	19.93	3.56	0.50	
		18	35.16	20.68	3.68	0.43	33.56	19.68	3.35	0.51	
		20	35.84	22.91	4.45	0.40	33.91	20.42	3.64	0.49	
		30	34.87	21.33	3.86	0.45	34.13	20.36	3.61	0.51	