









A Genetic Algorithm and Neural Network Stacking Ensemble Approach to Improve NO₂ Level Estimations

Javier González-Enrique¹ , Juan Jesús Ruiz-Aguilar² ,
José Antonio Moscoso-López² , Steffanie Van Rooode¹ ,
Daniel Urda¹ , and Ignacio J. Turias¹ 

¹ Department of Computer Science Engineering,
Polytechnic School of Engineering, University of Cádiz,
Avda. Ramón Puyol, s/n, 11202 Algecira, Cádiz, Spain
{javier.gonzalezzenrique, steffanie.vanroode,
daniel.urda, ignacio.turias}@uca.es

² Department of Industrial and Civil Engineering,
Polytechnic School of Engineering, University of Cádiz,
Avda. Ramón Puyol, s/n, 11202 Algeciras, Cádiz, Spain
{juanjesus.ruiz, joseantonio.moscoso}@uca.es

Abstract. This work investigates the possible improvements that a stacked ensemble can provide to NO₂ estimations in a monitoring network located in the Bay of Algeciras (Spain). In the proposed ensemble, ANNs, linear and nonlinear genetic algorithms models have been used as the individual learners in the first stage. The non-linear GA models produce better results than linear GA models as they are able to detect useful relationships between variables that are ignored in the linear case. The outputs of the individual learners have been employed as the inputs of the ANN models of the second stage. The most accurate of these models produced the final NO₂ estimation. The obtained results are promising as this final stage-2 model is able to outperform all the other estimation models considered in this work. This can be explained due to its ability to exploit the advantages offered by each individual model from stage-1 and then find an optimal combination of their outputs in order to increase the global estimation performance. The improvement of these NO₂ estimations can be very useful to improve the autonomous capacities for monitoring networks.

Keywords: Artificial neural networks · Genetic algorithms · Air pollution · NO₂ · Ensembles

1 Introduction

Ensemble methods are machine learning algorithms where the performance or classification accuracy is improved as a result of the combination of individual models. Different variants and approaches can be found in the scientific literature to create these ensembles. A first approach employs the same learner but changes the training datasets. Between this type of ensembles Boosting [1], Bagging [2], Random forest [3] and

AdaBoost [4] can be cited. Another possible approach relies on the use of different learning methods. In this case, majority voting, weighted voting and averaging are the most common techniques. Finally, stacking ensembles [5] are based on the use of the outputs of individual models as inputs of a second stage algorithm as a way to improve the performance of the models.

Air pollution is one of the most important environmental problems that must be faced in order to preserve the quality of living of the population. Nitrogen dioxide (NO_2) is one of the main pollutants. Its origins are manifold, but it is very related to combustion processes [6] and the reactions between nitrogen oxides and ozone [7]. It has harmful effects on human health [8] and is considered to be the main reason for air quality loss in urban areas [9].

The main objective of this paper is to improve NO_2 estimations in a monitoring network located in the Bay of Algeciras (Spain). To achieve this goal, a stacking ensemble is proposed. Artificial neural networks (ANNs), linear and nonlinear genetic algorithms (GAs) are employed as individual learners. Besides, ANNs are used as the second stage algorithm. This ensemble produces promising results outperforming all the individual models and other stacking ensembles that are also calculated. The importance of improving the NO_2 estimations is related to their ability to give monitoring networks autonomous capabilities, such as missing data imputation or detection of decalibration situations.

The rest of this paper is organized as follows. Section 2 describes the area of study and the database. Section 3 presents the methods used in this work. Section 4 describes the experimental design. Results are discussed in Sect. 5. Finally, the conclusions are shown in Sect. 6.

2 Data and Area Description

The Bay of Algeciras area is a heavily industrialized region which is located in the south of Spain and includes a population of nearly 300,000 inhabitants. The sources of NO_2 are numerous including not only the mentioned industries but very heavy traffic in the urban areas. Additionally, the Port of Algeciras Bay is one of the most prominent ship-trading ports in Europe. Thus, vessels constitute another important source of gaseous air pollution in this area.

All the aforementioned facts highlight the importance that an adequate pollution control strategy has to preserve the wellbeing of the population. With this purpose, a pollution monitoring network is located in this area. It is composed of 14 stations and records hourly measures of NO_2 . Figure 1 shows the location of the Bay of Algeciras and the situation of the monitoring stations (depicted using their codes). Table 1 shows the correspondence between stations and their codes.



Fig. 1. Area of study

Table 1. Location of the NO₂ monitoring stations

Monitoring station	Code	Lat.	Long.
EPSA Algeciras	1	36°8'11.70" N	5°27'11.44" W
Campamento	2	36°10'45.96" N	5°22'37.9" W
Los Cortillijos	3	36°11'25.74" N	5°26'8.73" W
Esc. Hostelería	4	36°12'13.97" N	5°23'1.33" W
Col. Los Barrios	5	36°11'5.154" N	5°29'18.362" W
Col. Carteya	6	36°12'34.798" N	5°23'29.803" W
El Rinconcillo	7	36°09'42.95" N	5°26'31.765" W
Palmones	8	36°10'35.274" N	5°26'24.648" W
Est. San Roque	9	36°12'47.218" N	5°25'55.945" W
El Zabal	10	36°10'21.47" N	5°20'30.192" W
Economato	11	36°11'23.032" N	5°22'51.85" W
Guadarranque	12	36°10'55.55" N	5°24'41.6" W
La Línea	13	36°09'22.252" N	5°20'22.823" W
Madrevieja	14	36°12'6.66" N	5°24'19.8" W

The database used in this work contains hourly NO₂ concentration measures that were obtained by the aforementioned monitoring stations during a period of 6 years (2010–2015). This database was normalized as a previous step. Then, it was split into two different datasets. A first one including records from 2010 to 2014, which was used to select the best parameters of the models and train them. The second one includes only measures taken in 2015 and was used as the test set. The results are provided using only the test set in order to determine the performance of the models with unseen data.

3 Methods

This section presents a brief description of the methods and techniques used in this work.

3.1 Artificial Neural Networks

Backpropagation feedforward multilayer perceptron [10], which includes at least one hidden layer different from the input and output layers, is the most widely used design for ANNs. According to [11], ANNs with enough neurons and a single hidden layer can be considered as universal approximators of any nonlinear function.

In this work, backpropagation neural networks (BPNNs) with a single hidden layer have been used to create hourly NO₂ estimation models. The Levenberg–Marquardt algorithm [12] has been employed for optimization purposes. Additionally, the early stopping technique [13] has been applied to the training process with the aim of avoiding overfitting and ensuring good generalization capabilities in the models.

So as to determine the optimal number of hidden neurons, authors have used a 5-fold cross validation resampling procedure, which has been used previously with good results [14–18].

3.2 Genetic Algorithms

Genetic algorithms [19] are search methods inspired by the natural selection processes. The decision variables for a particular problem are encoded into strings of a certain alphabet. These strings are known as chromosomes and act as candidate solutions of the problem (which are known as individuals). The set of all the individuals is known as the population. In order to determine the goodness of each possible solution, a fitness value is calculated for each individual of the population.

The general process starts with the generation of a random initial population. This population evolves from one generation to another through the application of genetic operators. It moves towards a global optimum solution of the problem according to the fitness values obtained. Selection, crossover and mutation can be found among the genetic operators. The process continues and new generations are created until the stopping criteria are met. The interested reader can find a more detailed explication of this process in the work of [20].

In this work, four different genetic algorithms models have been developed in order to estimate the hourly NO₂ concentrations at the EPSA monitoring station (see Table 1). In all these cases, the fitness function that must be minimized is the mean squared error (MSE) between the dependent variable and the estimation produced as the output of a function which is specific for each case. This is shown in Eq. 1.

$$err = MSE(y, \hat{y}) \quad (1)$$

where y is the dependent variable and \hat{y} is the estimation produced by the GA model. The main differences between these models lie on the specific function that is used to

produce the estimations. Equations 2, 3, 4 and 5 show the estimation functions corresponding to GA model 1 (GA-1), GA model 2 (GA-2), GA model 3 (GA-3) and GA model 4 (GA-4) respectively.

$$\hat{y} = \sum_{i=1}^n (w_{1_i} \cdot (S(w_{2_i} \cdot x_i) + w_{3_i} \cdot x_i + w_{4_i} \cdot x_i^{w_{5_i}} + e^{w_{6_i} \cdot x_i + w_{7_i}})) + k \quad (2)$$

$$\hat{y} = \sum_{i=1}^n (S(w_{1_i}) \cdot (S(w_{2_i} \cdot x_i) + w_{3_i} \cdot x_i + w_{4_i} \cdot x_i^{w_{5_i}} + e^{w_{6_i} \cdot x_i + w_{7_i}})) + k \quad (3)$$

$$\hat{y} = \sum_{i=1}^n (w_{1_i} \cdot x_i + S(w_{2_i} \cdot x_i) + w_{3_i} \cdot x_i^{w_{4_i}} + e^{(w_{5_i} \cdot x_i)}) + k \quad (4)$$

$$\hat{y} = \sum_{i=1}^n (w_i \cdot x_i) + k \quad (5)$$

where y is the dependent variable, x_i are the independent variables (predictors), n is the total number of predictors, w_{1_i} , w_{2_i} , w_{3_i} , w_{4_i} , w_{5_i} , w_{6_i} , w_{7_i} , w_i , and k are the weights determined by the GA and S is the sigmoid function, which is expressed in Eq. 6.

$$S(n) = \frac{1}{1 + e^{(-n)}} \quad (6)$$

It is important to note that w_{6_i} in Eqs. 2, 3, and 4 has been constrained within the $[10^{-12}, +\infty)$ interval. The genetic algorithm function provided by MATLAB R2016b has been used to develop the GA models. In this software, the codification of the variables in chromosomes is done internally without any intervention by the user. As can be seen in Eqs. 2–5, GA-1, GA-2, GA-3 present a non-linear behaviour whereas GA-4 fitness function is linear.

3.3 Stacked Ensembles

Stacked ensembles [21] are techniques which are intended to supply an overall prediction or estimation value based on the combination of the outputs of individual models. This type of ensemble can be benefited from the different perspectives offered by the individual models and usually improve their results. A brief description of the ensembles used in this work is presented next:

- Average (*avg*): The final estimation is calculated as the average of the individual models' estimations.
- Weighted average (*wavg*): In this case, each individual model has a different contribution to the final estimation according to the goodness of its estimation power, as is shown in Eq. 7.

$$E_{final} = \sum_{i=1}^j (w_i \cdot E_i) \quad (7)$$

where $w_i \in [0, 1]$ for each $i \in [1, \dots, j]$, $\sum_{i=1}^j w_i = 1$ and E_i represent an individual estimation model.

- ANN weighted ensemble (*ANNwe*): Inspired in the *wavg* ensemble, this work proposes a type of ensemble which uses the individual models as inputs of a BPNN. The obtained model represents the best possible combination of the inputs in order to produce the aggregated output.

4 Experimental Procedure

The objective of this study is to determine the possible improvements in NO₂ estimation models performance when a proposed stacked ensemble is applied. In this approach, GAs are used in conjunction with ANNs. The proposed fitness functions (see Eqs. 2–6) let the GA models capture linear and nonlinear relations between variables and increase their estimation performance.

For estimation purposes, the hourly NO₂ values measured at the EPSA monitoring station (see Table 1) was considered as the dependent variable. In contrast, the hourly NO₂ values corresponding to the rest of the monitoring stations were used as predictor variables. As an initial step, the original database was normalized and divided into two disjoint groups. The first one included hourly NO₂ records going from 2010 to 2014 and was used as the training set. The second one included records belonging to 2015 and acted as the test set.

The experimental process was divided into two different stages. In the first stage, five estimation models were developed, one using ANNs and the rest using different genetic algorithms approaches. In the case of the ANN models, the BPNNs used a single hidden layer and a different number of hidden neurons (hns) (1 to 50). The Levenberg–Marquardt was selected as the optimization algorithm and the early stopping technique was employed to improve the generalization capabilities of the models. Starting with the training set, a random resampling procedure using 5-fold cross-validation was used for each number of hns and the average performance measures were calculated. This process was repeated 20 times to avoid the effect of randomness in the ANN weights initialization, and the average results were also calculated. Additionally, the individual results for each repetition were also stored so that a multicomparison procedure could determine meaningful differences within the models in a later step. Regarding performance measures, the Pearson correlation coefficient (R), the mean squared error (MSE), the index of agreement (d) and the mean absolute error (MAE) [22] were calculated. These performance indexes are defined in Eqs. (8–11).

$$R = \frac{\sum_{i=1}^N (O_i - \bar{O}) \cdot (P_i - \bar{P})}{\sqrt{\sum_{i=1}^N (O_i - \bar{O})^2 \cdot \sum_{i=1}^N (P_i - \bar{P})^2}} \tag{8}$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (P_i - O_i)^2 \tag{9}$$

$$d = 1 - \frac{\sum_{i=1}^N (P_i - O_i)^2}{\sum_{i=1}^N (|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \tag{10}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |P_i - O_i| \tag{11}$$

where P indicates predicted values and O indicates observed values.

Finally, the best model was obtained using the Friedman test [23] and the Bonferroni method [24], jointly with the mentioned performance measures. The Friedman test let us determine if meaningful differences were present between the models. The Bonferroni method evaluated which of the models were not statistically equivalent. Following the Occam’s razor principle, the model with fewer hns was selected among those showing no significant differences with the model that produced best performance indexes.

After the model selection, a new BPNN model was trained using the whole training dataset and the number of hns of the most accurate model. Then, this model was fed with the inputs of the test set in order to obtain the final NO₂ estimation for the year 2015. Finally, performance measures were calculated through the comparison of observed vs. estimated values.

In the case of the GA models, the fitness functions presented in Sect. 3.2 (Eqs. 1 to 6) were minimized using the training data set. Regarding the parameters that control the genetic algorithms, different tests were carried out in order to select the best possible parameter combination and each combination was repeated 20 times. Table 2 shows the possible values that were tested for each parameter. A detailed description of each parameter can be found in Matlab’s Genetic Algorithm Options web page [25]. Table 3 shows the final combination selected per each GA model.

Table 2. Parameters tested in the GA models

Population size	Crossover function	Crossover fraction (%)	Selection function	Mutation rate (%)	Max. generations
200, 250, 350, 450, 550, 650	Scattered Single point Two point Intermediate Heuristic Arithmetic	75, 80, 85, 90	Stochastic uniform Remainder Uniform Roulette Tournament	2, 4, 6, 8, 10	2000

Table 3. Selected parameters for each AG model

Model	Population size	Crossover function	Crossover fraction (%)	Selection function	Mutation rate (%)
GA-1	450	Heuristic	80	Roulette	4
GA-2	450	Heuristic	80	Roulette	8
GA-3	650	Arithmetic	80	Tournament	6
GA-4	650	Arithmetic	80	Stochastic uniform	8

Once the stopping criteria were met, the corresponding weights were stored for each specific GA model. As the last step, the final NO₂ estimations for the year 2015 were obtained using Eqs. 2, 3, 4, 6 and the corresponding weights and test sets for each case. Finally, values for R, MSE, d and MAE were calculated after comparing the observed NO₂ values against the estimated ones obtained with each model.

In the second stage, *avg*, *wavg* and *ANNwe* ensembles (see Sect. 3.3) were calculated. In the case of the *avg* ensemble, the calculation is straightforward as it only averages the obtained estimations obtained with the individual models. For the *wavg* ensemble (see Eq. 7), each of the estimations from stage-1 was weighted according to its MSE value following an inversely proportional distribution. To calculate the *ANNwe* ensemble, ANN models were trained using the outputs from stage-1 as their inputs and the 2015 NO₂ measured values as their targets. In this case, the same network configuration as stage-1 ANN models was applied. The final output was obtained as the one which produced a lesser MSE value after 20 repetitions.

5 Results and Discussion

The results of the experimental procedure are presented in this section. In the first stage, different models have been developed to estimate the hourly NO₂ concentration values at the EPSA monitoring station (station 1). NO₂ values measured at the other stations have been used as inputs of the models (see Table 1 and Fig. 1). The initial data set has been split into two disjoint datasets and the results are obtained through the comparison of observed vs. estimated values for the test set (2015). This lets us evaluate the performance of the models with unseen data. For comparative purposes, a Lasso model using the same datasets and 5-fold cross validation has also been included. Table 4 shows the performance measures corresponding to stage-1 models.

Table 4. Performance indexes for stage-1 estimation models

Model	nh	\bar{R}	\overline{MSE}	\overline{MAE}	\bar{d}
ANN	6	0.761	238.769	10.988	0.845
GA-1	-	0.716	276.931	11.904	0.814
GA-2	-	0.713	277.556	11.956	0.811
GA-3	-	0.716	276.093	11.980	0.815
GA-4	-	0.650	326.148	13.512	0.757
Lasso	-	0.650	326.155	13.520	0.760

As can be expected, the best ANN model outperforms all the GA models. However, it can be noted that non-linear GA models (GA-1, GA-2 and GA3) beat easily the performance offered by the GA-4 linear model. This indicates that Eqs. 2, 3 and 4 are able to capture linear and also an important amount of non-linear relations between input and output variables. However, the proposed fitness functions cannot compete with ANNs' ability to act as universal approximator of any nonlinear function, as was mentioned in Sect. 3.1.

Table 5. Performance indexes for stage 2 ensembles

Model	nh	\bar{R}	\overline{MSE}	\overline{MAE}	\bar{d}
<i>avg</i>	-	0.726	268.741	11.755	0.815
<i>wavg</i>	-	0.729	266.139	11.679	0.8175
<i>ANNwe</i>	40	0.784	215.608	10.503	0.8690

Table 5 shows the results obtained by the proposed ensembles in the second stage. These methods combine the outputs of stage-1 models with the aim of improving the estimation results.

Results show how *avg* and *wavg* ensembles constitute an improvement over GA models but do not reach the estimation goodness offered by the stage-1 ANN model. This can be explained by the fact that the average and the weighted average operations (to a lesser extent) are highly influenced by extreme values that are far from the mean of the individual learners considering a specific instant of time. In our case, this influence comes primarily from GA4 output. As an example, if AG4 output is removed from the ensemble, the value of MSE for *avg* drops to 260.837 and its R-value rises to 0.735.

In the case of the *ANNwe* ensemble, its performance indexes are far superior when compared to those belonging to all the proposed models in the first and second stages. As can be seen, the second stage ANN can take advantage of the different linear and non-linear relations captured by the GA and ANN of the first stage. Some of them are already present in the ANN model of the first stage, but other ones are provided by the GA models. Considering the results, the proposed two-stage approach guarantees a better estimation performance of the NO₂ concentration values at the EPSA monitoring station.

A comparison between the best models of the first and second stages is presented in Figs. 2 and 3 where estimated versus measured NO₂ hourly values are depicted for January 2015. As can be seen, the fit and adjustment to the observed values are superior in the case of *ANNwe* when compared to the ANN model of the first stage. This confirms what was stated before about the improvement of the estimation models provided by the proposed approach (Fig. 3).

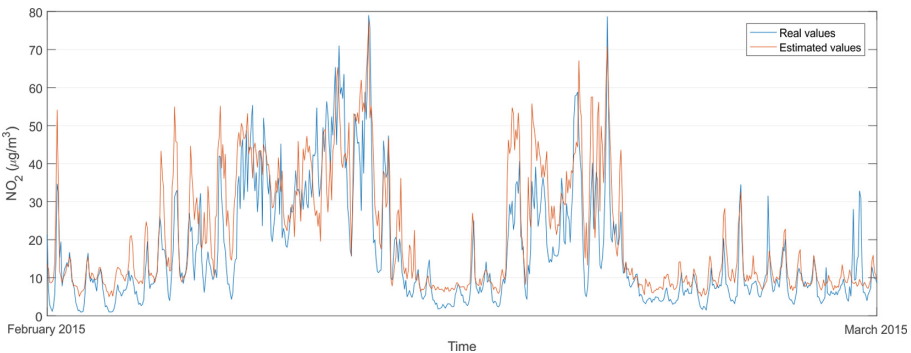


Fig. 2. Estimated vs. real values for January 2015 using the stage-1 ANN most accurate model

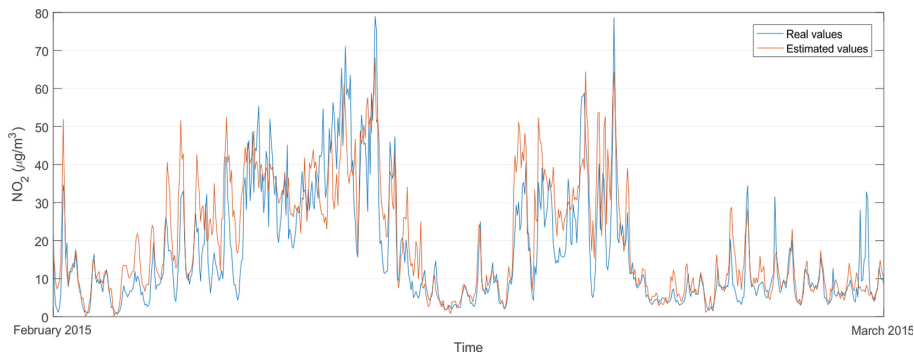


Fig. 3. Estimated vs. real values for January 2015 using the stage-2 *ANNwe* ensemble

6 Conclusions

The aim of this paper is to verify the possible improvements that a stacked ensemble approach can provide to NO₂ estimations over compared to other individual models. This approach uses artificial neural networks, linear and nonlinear genetic algorithms as individual learners. Then, their outputs are used as inputs of the ANN models of the second stage.

Regarding the first stage results, the proposed GA models that use non-linear functions produce much better results if they are compared to GA using a linear fitness function. This indicates that their fitness functions can detect useful relationships between variables that are ignored in the linear approaches. However, ANNs outperform them due to their ability to act as universal approximators (see Sect. 3.1).

The results of both stages show how the *ANNwe* approach outperforms all the other proposed approaches, ensuring a better estimation performance of NO₂ in the monitoring network. The main reason can be found in the fact that all stage-1 models capture different linear and nonlinear relations between the inputs and the targets. Therefore, the *ANNwe* approach is able to exploit the advantages offered by each individual model. Additionally, it is able to find an optimal combination of their outputs in order to increase the global estimation performance.

The use of the proposed model can provide better and more reliable NO₂ estimations if it is compared to the other proposed models. This can be very useful as these estimations can provide robustness and autonomous capabilities to the monitoring network. They also can be helpful in missing values or detection of decalibration situations.

Acknowledgements. This work is part of the coordinated research projects TIN2014-58516-C2-1-R and TIN2014-58516-C2-2-R supported by MICINN (Ministerio de Economía y Competitividad-Spain). Monitoring data have been kindly provided by the Environmental Agency of the Andalusian Government.

References

1. Drucker, H., Cortes, C., Jackel, L.D., LeCun, Y., Vapnik, V.: Boosting and other ensemble methods. *Neural Comput.* **6**, 1289–1301 (1994). <https://doi.org/10.1162/neco.1994.6.6.1289>
2. Breiman, L.: Bagging predictors. *Mach. Learn.* **24**, 123–140 (1996). <https://doi.org/10.1007/BF00058655>
3. Breiman, L.: Random forests. *Mach. Learn.* **45**, 5–32 (2001). <https://doi.org/10.1023/A:1010933404324>
4. Freund, Y., Schapire, R.E.: Experiments with a new boosting algorithm. In: *Machine Learning International Workshop*, pp. 148–156 (1996)
5. Ting, K.M., Witten, I.H.: Issues in stacked generalization. *J. Artif. Int. Res.* **10**, 271–289 (1999)
6. Rivera, C., et al.: Spatial distribution and transport patterns of NO₂ in the Tijuana - San Diego area. *Atmos. Pollut. Res.* **6**, 230–238 (2015)
7. Finlayson-Pitts, B.J., Pitts, J.N.J.: The atmospheric system. In: Finlayson-Pitts, B.J., Pitts, J. N.J. (eds.) *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*, pp. 15–42. Academic Press, San Diego (2000)
8. Faustini, A., Rapp, R., Forastiere, F.: Nitrogen dioxide and mortality: review and meta-analysis of long-term studies. *Eur. Respir. J.* **44**, 744–753 (2014)
9. Westmoreland, E.J., Carslaw, N., Carslaw, D.C., Gillah, A., Bates, E.: Analysis of air quality within a street canyon using statistical and dispersion modelling techniques. *Atmos. Environ.* **41**, 9195–9205 (2007)
10. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by error propagation. In: Rumelhart, D.E., McClelland, J.L., PDP Research Group (eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Foundations, vol. 1, pp. 318–362. MIT Press, Cambridge (1986)
11. Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**, 359–366 (1989). [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
12. Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **11**, 431–441 (1963)
13. Gardner, M.W., Dorling, S.R.: Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmos. Environ.* **32**, 2627–2636 (1998). [https://doi.org/10.1016/S1352-2310\(97\)00447-0](https://doi.org/10.1016/S1352-2310(97)00447-0)
14. Turias, I.J., González, F.J., Martín, M.L., Galindo, P.L.: Prediction models of CO, SPM and SO₂ concentrations in the Campo de Gibraltar Region, Spain: a multiple comparison strategy. *Environ. Monit. Assess.* **143**, 131–146 (2008). <https://doi.org/10.1007/s10661-007-9963-0>
15. Muñoz, E., Martín, M.L., Turias, I.J., Jimenez-Come, M.J., Trujillo, F.J.: Prediction of PM₁₀ and SO₂ exceedances to control air pollution in the Bay of Algeciras, Spain. *Stoch. Environ. Res. Risk Assess.* **28**, 1409–1420 (2014). <https://doi.org/10.1007/s00477-013-0827-6>
16. Turias, I.J., et al.: Prediction of carbon monoxide (CO) atmospheric pollution concentrations using meteorological variables. *WIT Trans. Ecol. Environ.* **211**, 137–145 (2017). <https://doi.org/10.2495/AIR170141>
17. González-Enrique, J., Turias, I.J., Ruiz-Aguilar, J.J., Moscoso-López, J.A., Jerez-Aragonés, J., Franco, L.: Estimation of NO₂ concentration values in a monitoring sensor network using a fusion approach. *Fresen. Environ. Bull.* **28**, 681–686 (2019)

18. González-Enrique, J., Turias, I.J., Ruiz-Aguilar, J.J., Moscoso-López, J.A., Franco, L.: Spatial and meteorological relevance in NO₂ estimations. A case study in the Bay of Algeciras (Spain). *Stoch. Environ. Res. Risk Assess.* **33**, 801–815 (2019). <https://doi.org/10.1007/s00477-018-01644-0>
19. Holland, J.H.: *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor (1975)
20. Goldberg, D.E.: *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Reading (1989)
21. Polikar, R.: Ensemble based systems in decision making. *Circuits Syst. Mag. IEEE.* **6**, 21–45 (2006). <https://doi.org/10.1109/MCAS.2006.1688199>
22. Willmott, C.J.: Some comments on the evaluation of model performance. *Am. Meteorol. Soc.* **63**, 1309–1313 (1982)
23. Friedman, M.: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.* **32**, 675–701 (1937). <https://doi.org/10.1080/01621459.1937.10503522>
24. Hochberg, Y., Tamhane, A.C.: *Multiple Comparison Procedures*. Wiley, New York (1987)
25. The Mathworks Inc.: *Genetic Algorithm Options*. <https://es.mathworks.com/help/gads/genetic-algorithm-options.html>