

Article

Characterization of Surface Ozone Behavior at Different Regimes

Nádia F. Afonso and José C. M. Pires * 

Laboratório de Engenharia de Processos, Ambiente Biotecnologia e Energia (LEPABE), Departamento de Engenharia Química, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal; nadia.afonso@iol.pt

* Correspondence: jcpires@fe.up.pt; Tel.: +351-225-082-262

Received: 25 July 2017; Accepted: 12 September 2017; Published: 14 September 2017

Abstract: Previous studies showed that the influence of meteorological variables and concentrations of other air pollutants on O₃ concentrations changes at different O₃ concentration levels. In this study, threshold models with artificial neural networks (ANNs) were applied to characterize the O₃ behavior at an urban site (Porto, Portugal), describing the effect of environmental and meteorological variables on O₃ concentrations. ANN characteristics, and the threshold variable and value, were defined by genetic algorithms (GAs). The considered predictors were hourly average concentrations of NO, NO₂, and O₃, and meteorological variables (temperature, relative humidity, and wind speed) measured from January 2012 to December 2013. Seven simulations were performed and the achieved models considered wind speed (at 4.9 m·s⁻¹), temperature (at 17.5 °C) and NO₂ (at 26.6 µg·m⁻³) as the variables that determine the change of O₃ behavior. All the achieved models presented a similar fitting performance: $R^2 = 0.71\text{--}0.72$, $RMSE = 14.5\text{--}14.7 \mu\text{g}\cdot\text{m}^{-3}$, and the index of agreement of the second order of 0.91. The combined effect of these variables on O₃ concentration was also analyzed. This statistical model was shown to be a powerful tool for interpreting O₃ behavior, which is useful for defining policy strategies for human health protection concerning this air pollutant.

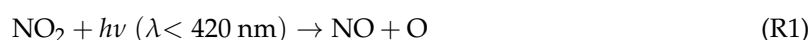
Keywords: air pollution; artificial neural network; genetic algorithms; surface ozone; threshold models

1. Introduction

Surface ozone (O₃) is considered one of the most concerning air pollutants in Europe. It is a secondary pollutant (it is not directly emitted), generated by chemical reactions that occur in the atmosphere between primary air pollutants (nitrogen oxides—NO_x—and volatile organic compounds—VOCs) catalyzed by sunlight [1]. The impact of this air pollutant has been studied in different areas [2–5]. Concerning human health, O₃ can cause injuries to airway epithelial cells (and lung diseases such as asthma), hyperplasia, headaches, and nausea, particularly in sensitive people, such as children and elderly [6–10]. Regarding vegetation, O₃ can damage plant leaves (decreases in both leaf photosynthesis and leaf area), reducing crop yields associated with a high negative economic impact [11–13]. Moreover, as a strong oxidant, it is responsible for the degradation of material via corrosion [14].

As already mentioned, O₃ is produced by chemical reactions between primary pollutants present in the atmosphere. It can also be transported from other locations by the wind (horizontal transport) and from the stratosphere (vertical transport) [15,16]. Thus, the atmosphere works as an open chemical reactor, in which reaction kinetics depend on the concentrations of reactants (primary air pollutants), mixture (wind speed and direction), temperature (influencing exponentially kinetic reaction constants—Arrhenius equation), and solar radiation. Thus, the O₃ behavior is highly dependent on the environment (urban, rural, or background) and meteorological variables.

In urban areas, O₃ concentrations are usually lower than the values observed in rural areas [17–20]. This phenomenon occurs mainly due to high NO_x concentrations. Photochemical equilibrium is defined by the following equations [21–23]:



The photochemical reaction of NO₂ decomposition (chemical reaction R1) leads to the production of NO molecules and oxygen atoms that combine with molecular oxygen to produce ozone (chemical reaction R2). O₃ can also react with NO, forming NO₂ (chemical reaction R3). The complexity of phenomena associated with O₃ formation makes it hard to understand and predict its concentration in ambient air. Moreover, there are insufficient data (e.g., inventories of air pollutant emissions) to develop consistent phenomenological models to describe O₃ behavior.

Alternatively, statistical models have the capability of characterizing the relationship between variables using collected data and they involve mere pattern-recognition using mathematical operations. One of the most applied statistical models is the artificial neural network (ANN). ANNs are nonlinear models, which are inspired in the biological neural processing system [24,25]. These models are composed by artificial neurons (grouped in layers; three layers—input, hidden, and output—are often applied) that receive an input value and converts to an output through a selected function (activation function). Additionally, ANNs are characterized by a high fitting performance. The rapid development of computer hardware has increased the processing capabilities, which have led to achievement of ANN models with less computation time [26]. Therefore, these models have been used in a wide range of applications, including classification, regression, and mapping [27–29]. However, there are too many variables that need to be defined before the model parameters can be determined, including (i) the number of processing neurons in the hidden layer, and (ii) the activation function for each neuron. In recent years, genetic algorithms (GAs) have been applied to help in the definition of these variables. GAs are commonly applied to generate high-quality solutions for optimization and search problems, based on bio-inspired operators, such as mutation, crossover, and selection [30,31]. In GAs, a set of candidate solutions (called population—a group of individuals) are iteratively modified through the mentioned genetic operators in order to find a group of better solutions for the next generation (new iteration). GAs present the following advantages: (i) continuous or discrete variables can be optimized; (ii) a derivative function is not required; (iii) multivariable problems can be optimized; (iv) extremely complex cost surfaces can be dealt with; and (v) a list of optimal solutions (and not just a single one) is provided.

In Porto (selected area in this study—Portugal), an increasing trend of O₃ concentrations (147% higher) has been observed since the 19th century due to the photochemical production of this pollutant, associated with the increase in anthropogenic emissions mainly due to traffic [32]. Additionally, in the north of Portugal, Lamas d'Olo is a rural site where the highest O₃ concentrations are usually measured. Consequently, this site is often selected to evaluate O₃ behavior. Russo et al. [33] mentioned that high-ozone episodes can be explained by several factors: (i) atmospheric stagnation; (ii) horizontal transport by the wind of ozone-rich air masses; (iii) high solar radiation and temperature; and (iv) the influence of local winds (sea breezes and valley winds). Carvalho et al. [19] observed a positive correlation between O₃ concentrations with temperature and a negative correlation with relative humidity. Regarding the effect of wind field, the northeast flow from Spain (Galicia and Asturias) was observed, and this can be associated with the long-range transport of atmospheric pollutants to Portugal. Fernández-Guisuraga et al. [34] compared O₃ trends at urban and rural sites. At the rural site, O₃ concentrations were mainly influenced by the wind (transport), showing low variability with the concentrations of other pollutants. On the other hand, at the urban site, most of the variance was explained by the NO₂/NO_x ratio. Several research studies can be found where

ANN models were applied to determine O₃ trends and to predict its concentration (to provide early warning to the population when high O₃ concentration episodes occur). Comrie [35] compared the performance of an ANN model with a multiple linear regression (MLR) model to predict daily average O₃ concentrations in different cities with distinct climate and O₃ regimes. ANN models presented slightly better performance than MLR. Abdul-Wahab and Al-Alawi [36] developed ANN models to predict O₃ concentrations through meteorological and environmental data. The contribution of the meteorological data was defined between 33% and 41%, while the remaining variation was attributed to chemical pollutants. NO, SO₂, relative humidity (the highest contribution), non-methane hydrocarbon, and NO₂ were the variables that most influenced the O₃ concentrations. Additionally, temperature also presents an important role, while solar radiation had a lower effect than expected. Pires et al. [37] compared threshold autoregressive (TAR) models, autoregressive (AR) models, and ANN in the prediction of the next day hourly average O₃ concentrations. In the training period, ANN presented a higher performance. However, in the test period, TAR models presented more accurate results and the distinction became greater when the evaluation was performed for the prediction of extreme values.

In recent studies, O₃ concentrations have shown different behaviors regarding certain explanatory variables [25,37], which can be classified as O₃ regimes. This observation can be justified by the chemical reactions associated with O₃ formation/destruction that are influenced by certain variables, such as temperature, solar radiation, and wind speed [32]. To take these regimes account, threshold regression models were considered in this study [38]. Thus, GAs were used to define the threshold variable and value (the value of the explanatory variable corresponding to the change of the regime; two regimes were selected), the number of hidden neurons, and the activation function in the hidden and output layers. In this study, hourly average O₃ concentrations were modeled using threshold models with an ANN, whose structure was iteratively optimized by GAs. The achieved models enable the characterization of O₃ variability with selected meteorological and environmental variables in different regimes.

2. Materials and Methods

2.1. Data

Air quality data were obtained from an urban background site (*Sobreiras—Lordelo do Ouro*, see Figure 1) of the Air Quality Monitoring Network (AQMN) of Porto, Portugal. The AQMN is managed by the Regional Commission of Coordination and Development of Northern Portugal (*Comissão de Coordenação e Desenvolvimento Regional do Norte*), under the responsibility of the Ministry of Environment. Hourly average concentrations of NO, NO₂, and O₃ from the period from January 2012 to December 2013 (8760 hourly average values in 2012 and 7481 in 2013) were used to develop the proposed models. NO and NO₂ were obtained through the chemiluminescence method according to European Union (EU) Directive 1999/30/EC (European Community). According to EU Directive 2002/3/EC, O₃ measurements were performed through UV-absorption photometry using the equipment 41 M UV Photometric Ozone Analyzer (Environment S.A., Poissy, France). This monitoring equipment was subject to a rigid maintenance program, calibrated every 4 weeks. Measurements were continuously registered, and hourly average concentrations (in $\mu\text{g}\cdot\text{m}^{-3}$) were recorded.

The meteorological data were collected in a meteorological station located at *Pedras Rubras*, which is managed by *Instituto Português do Mar e da Atmosfera* (IPMA, I.P.); these values are considered representative for the entire Metropolitan Area of Porto. In this study, hourly averages of temperature (T , °C), relative humidity (RH , %), and wind speed (WS , $\text{m}\cdot\text{s}^{-1}$) were used to analyze the influence of meteorological conditions on O₃ concentrations.



Figure 1. Monitoring site of *Sobreiras—Lordelo do Ouro* (from <http://qualar.apambiente.pt/>).

2.2. Statistical Model

In this study, threshold models with ANNs were defined with GAs, aiming to evaluate the effect of environmental and meteorological variables in O_3 concentrations. The applied model is defined as the following:

$$y = \begin{cases} \text{net}_1(x_i), & \text{if } x_d \leq r \\ \text{net}_2(x_i), & \text{if } x_d > r \end{cases} \quad (1)$$

where y is the output variable, net_1 and net_2 are ANN models, x_i are the exploratory variables, x_d is the threshold variable, and r is the threshold value. Applied feedforward ANN models had three layers (input, hidden, and output) and considered eight input variables (hourly average data): NO concentration, NO_2 concentration (due to the chemical reactions R1 and R3), the ratio NO_2/NO (due to the equilibrium constant of the chemical reaction R3), T , RH , $1/RH$ (as RH usually shows a negative effect on O_3 levels), WS , and $1/WS$ (the same as the RH effect). The output variable was the hourly average O_3 concentrations measured at the same time of the input data to infer the direct influence of these variables on O_3 chemistry. Regarding the activation functions, the linear function was considered for the output neuron and four functions were selected by GAs: sigmoid, hyperbolic tangent, inverse, and radial basis. The data were divided in training (75%) and validation (25%) sets and the early stopping method (ANN training procedure is stopped when an increase in validation error is observed) was applied to avoid overfitting. The division of the data was performed by time: 75% for training (January 2012 to 25 May 2013); 25% for validation (25 May 2013 to 19 December 2013). In the training set, O_3 concentrations ranged from 0 to $161 \mu\text{g}\cdot\text{m}^{-3}$, while O_3 concentrations ranged from 0 to $170 \mu\text{g}\cdot\text{m}^{-3}$ in the validation set.

GAs are a search and optimization technique based on Darwin principles of evolution and natural genetics [30,31]. This procedure begins with a set of individuals (population) that is randomly generated. Each individual (also called chromosome) is a binary code string and contains information about a set of parameters, which is a potential solution to a given problem. To evaluate the quality of the proposed solution (to rank the individuals in the population), a fitness function should be defined. To create new chromosomes for the next generation, the fittest chromosomes are submitted to the genetic operations [30]: (i) selection; (ii) crossover; (iii) mutation. These new chromosomes are then evaluated according to the fitness function, and the ones with the highest performance were selected. The repetition of this procedure generates a sequence of populations containing better solutions. The termination criteria can be (i) to stop after a previously defined maximum number of generations is achieved, or (ii) to stop when a desired fitness value is achieved. In this study, GAs were used to define the threshold variable and value, the number of hidden neurons, and the activation function in the

hidden layer, and to select the explanatory variables to be used in each ANN model. The determination of the models was coded by the authors with MATLAB[®] software (R2014a, MathWorks, Natick, MA, USA, 2014) using the following specifications:

- a population size of 100;
- a selection probability of 0.20 (proportion of the individuals of the new generation obtained by selection operator);
- a selection criterion based on elitism (a small proportion of the fittest candidates is copied unchanged into the next generation);
- a crossover probability of 0.70 (proportion of the individuals of the new generation obtained by crossover operator);
- a mutation probability of 0.1 (proportion of the individuals of the new generation obtained by mutation operator);
- an evaluation of root mean squared error (RMSE) in training and validation sets;
- a stopping criterion based on the maximum number of generations.

Figure 2 shows an example of chromosome (37 bits). It is divided in 8 sets of bits (SB_i). SB₁ (3 bits) defines the threshold variable (from the explanatory variables; the maximum number of 8) through the conversion from binary to decimal numbers (MATLAB function *bin2dec*). SB₂ (8 bits) defines the threshold value. With the threshold variable already defined, the maximum (x_{max}) and minimum (x_{min}) values of this variable are determined. Threshold value is calculated based on Equation (2).

$$r = \frac{bin2dec(SB_2)}{255} \times (x_{max} - x_{min}) + x_{min}. \tag{2}$$

SB₃ and SB₆ (2 bits) define the activation function for the hidden layer of each ANN: 00—log-sigmoid (*logsig*); 01—hyperbolic tangent sigmoid (*tansig*); 10—inverse (*netinv*); 11—radial basis (*radbas*). SB₄ and SB₇ (3 bits) define the number of neurons in the hidden layer through the conversion from binary to decimal number (1 to 8). SB₅ and SB₈ (8 bits) define the explanatory variables that are used in each ANN (1 bit for each explanatory variable): 0—not selected; 1—selected.

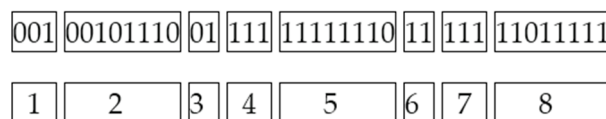


Figure 2. Example of a chromosome.

3. Results and Discussion

3.1. Air Quality and Meteorological Data Characterization

During the analyzed period, the hourly average O₃ concentrations were between 0 and 170 μg·m⁻³ (not exceeding the information neither the alert threshold—180 and 240 μg·m⁻³, respectively). Regarding O₃ exceedances to EU limits for the protection of human health, the 8 h average O₃ concentrations were higher than 120 μg·m⁻³ twice in September 2012, twice in July 2013, and once in August 2013. Figure 3 shows the average daily profile of O₃ concentrations. As a photochemical pollutant, its concentration increases during the daylight period, presenting a maximum between 14 and 15 h and a minimum at night time. The observed profile is characteristic of an urban site, as it does not present a high amplitude of concentrations (due to the presence of high NO_x concentrations).

Figure 4 shows the monthly average values of NO, NO₂, and O₃ concentrations, as well as the analyzed meteorological variables (temperature, relative humidity, and wind speed). High O₃ concentrations were observed in April 2012 (63.7 μg·m⁻³) and from March to July 2013 (59.9–71.3 μg·m⁻³). In this period, low concentrations of NO and NO₂ were also measured.

High temperatures and low relative humidity were also observed. On the other hand, lower O₃ concentrations were measured for periods with high NO and NO₂ concentrations and lower temperatures. These observations are in agreement with other research studies in which the behavior of O₃ was analyzed [18,39–41]. Pires et al. [40] compared several linear models to predict O₃ concentrations at an urban site in Porto. The correlation analysis performed between O₃ concentrations and meteorological variables showed also negative correlations with NO, NO₂, and RH and a positive correlation with *T*. In another study focusing on the same region [41], O₃ concentrations were negatively correlated with NO, NO₂, and *RH* and positively correlated with *T* and *WS*. Zhang, Wang, Park, and Deng [18] analyzed high O₃ concentration episodes and related them with meteorological variables. O₃ concentrations were highly correlated with maximum temperature and minimum relative humidity. The effect of minimum *WS* was also analyzed at urban, suburban, and rural sites. O₃ concentrations were positively (negatively) correlated with minimum *WS* at urban (suburban and rural) sites. Shan, Yin, Zhang, Ji, and Deng [39] analyzed the effect of meteorological variables on O₃ concentrations at an urban site in China. Daily average O₃ concentrations were negatively correlated with pressure and *RH*, and positively correlated with temperature, solar radiation, sunshine duration, and wind speed.

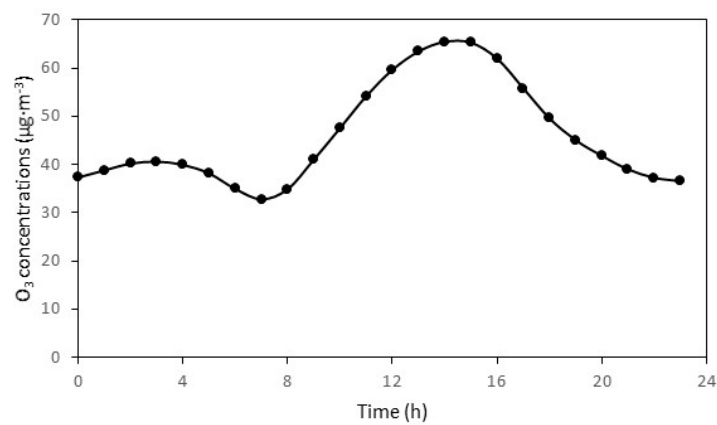


Figure 3. Daily average profile of O₃ concentrations at the monitoring site.

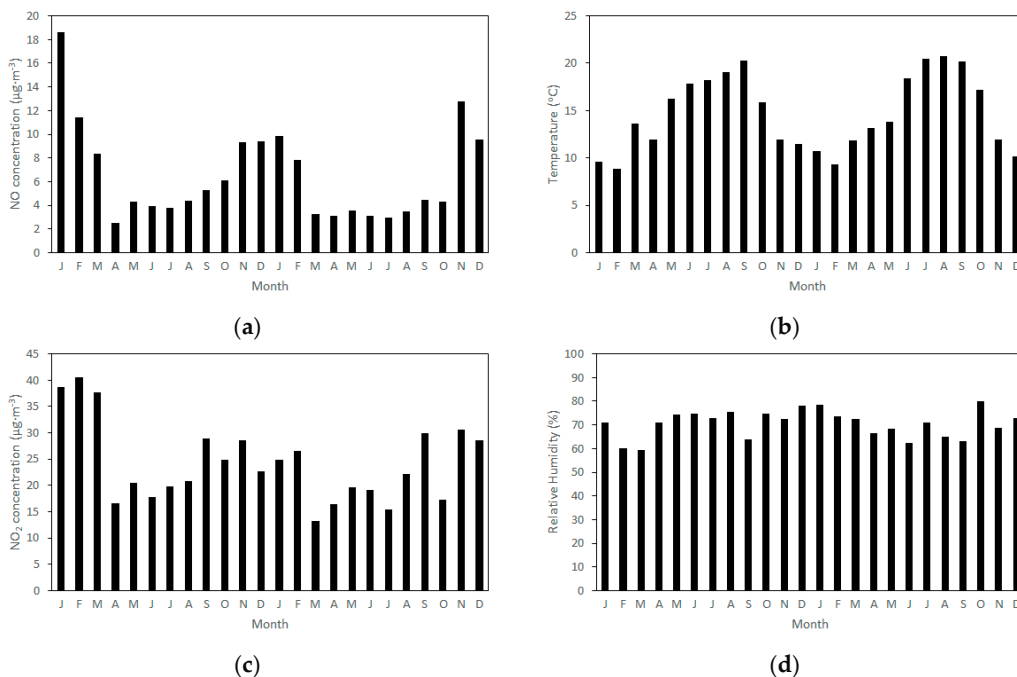


Figure 4. Cont.

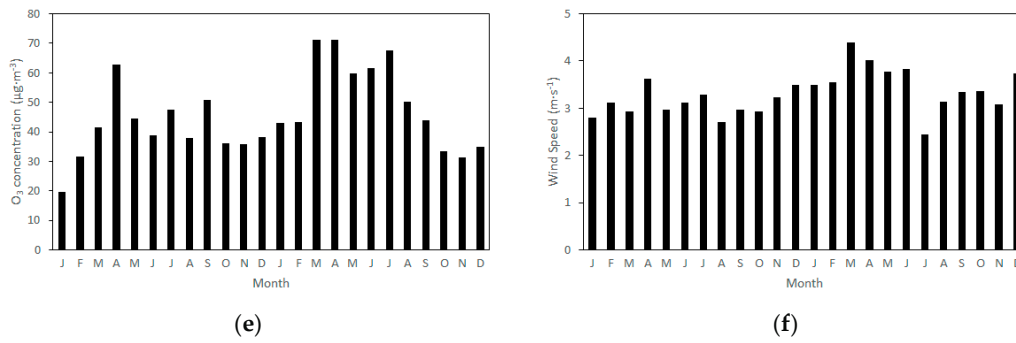


Figure 4. Monthly average values of (a) NO concentrations, (b) temperature, (c) NO₂ concentrations, (d) relative humidity, (e) O₃ concentrations, and (f) wind speed.

3.2. Linear Correlation Analysis

Figure 5 shows the variation in linear correlation between O₃ and meteorological parameters on a monthly basis. Negative correlations were observed for NO (−0.547 to −0.296) and NO₂ (−0.807 to −0.276) concentrations. The effect of these air pollutants was more significant in winter periods than in summer periods. Regarding the effect of meteorological variables, temperature was usually positively correlated with O₃, which was in agreement with what was expected. The highest value ($R = 0.661$) was determined in September 2013 and an unusual negative correlation ($R = -0.376$) was determined in July 2013. *RH* was negatively correlated in almost all periods. The highest impact was also observed in September 2013 ($R = -0.685$) and an unusual positive correlation was determined in July 2013 ($R = 0.419$). Chen et al. [42] demonstrated that *RH* favors O₃ decomposition, justifying the associated negative effect. Regarding *WS*, this variable can have two different effects on O₃ concentrations. Low *WS* can promote the accumulation of O₃ produced in the region (increasing its concentration), while high values reduce the levels of other air pollutants (such as NO_x) that influence the O₃ chemistry (in the case of NO_x, its concentration decrease leads to the increase in O₃ levels). Thus, the effect of *WS* on O₃ concentrations depends on the studied environment: urban or rural. In this study (urban environment), *WS* was positively correlated with O₃ concentrations, with the highest value ($R = 0.593$) in February 2013.

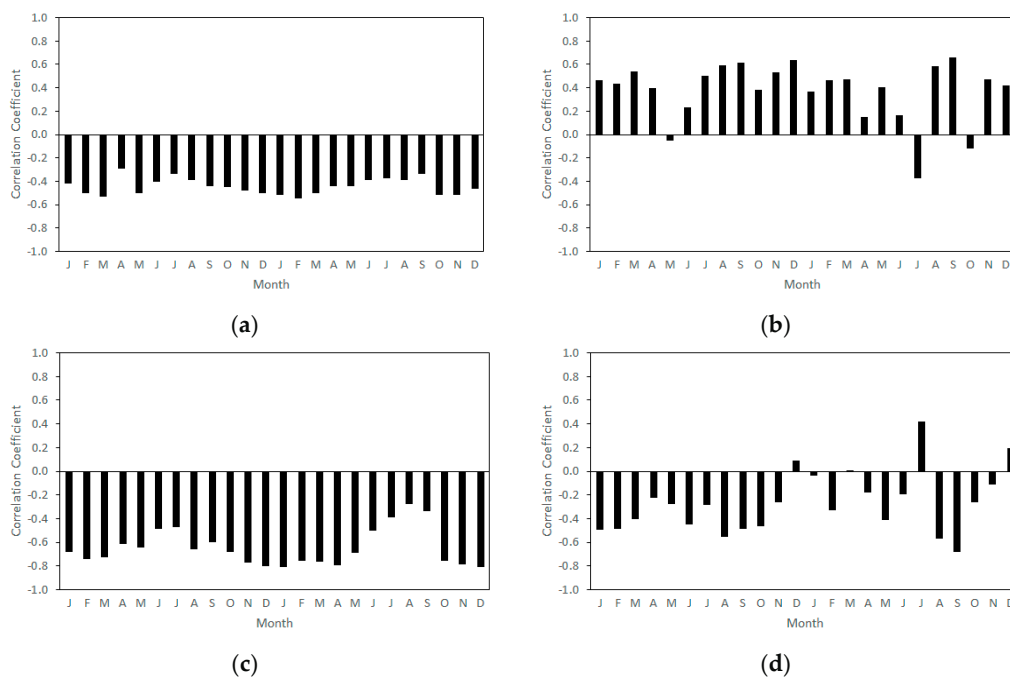
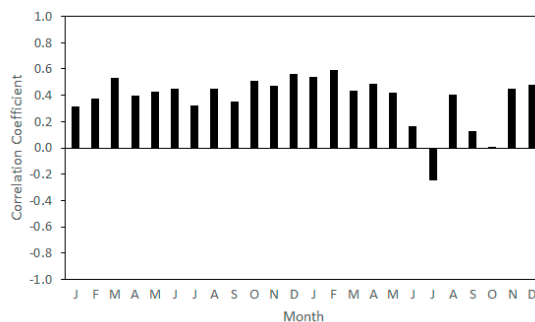


Figure 5. Cont.



(e)

Figure 5. Temporal variation of linear correlation between O₃ concentrations and the following: (a) NO concentrations, (b) temperature, (c) NO₂ concentrations, (d) relative humidity, and (e) wind speed.

3.3. ANN Models and Interpretation

Seven simulations were performed to determine the models able to describe the relationship between O₃ concentrations with NO, NO₂, *T*, *RH*, and *WS* (measured at same time). These models are threshold models, considering two O₃ regimes where the relationship between output and input variables are different. The change from one regime to another depends on the value (threshold value) of a specific input variable (threshold variable). GAs were used to optimize the ANN characteristics (the number of hidden neurons, the activation function in the hidden layer, and the input variables), and the threshold variable and value. The models were evaluated according their fitting performance in training and validation sets. Table 1 shows the best models for the seven simulations. All of them presented similar fitting performance: (i) $R^2 = 0.71$ – 0.72 ; (ii) *RMSE* between 14.5 and 14.7 $\mu\text{g}\cdot\text{m}^{-3}$; and (iii) the index of agreement of the second order of 0.91. Three explanatory variables were selected, each one with a specific threshold value: (i) *WS* with 4.9 $\text{m}\cdot\text{s}^{-1}$; (ii) *T* with 17.5 °C; and (iii) NO₂ with 26.6 $\mu\text{g}\cdot\text{m}^{-3}$. Generally, hyperbolic tangent and radial basis were the functions selected for the hidden layer, composed by 7 or 8 neurons. In almost all models, all input variables were selected as ANN inputs in both O₃ regimes.

The analysis of the combined effect of input variables (two variables) was performed for the three threshold variables, considering the two regimes determined by the best models in Simulations I, II, and III. Figure 6 shows the combined effect of NO₂, *T*, and *WS* on O₃ concentrations for $WS \leq 4.9 \text{ m}\cdot\text{s}^{-1}$ and for $WS > 4.9 \text{ m}\cdot\text{s}^{-1}$. For $WS \leq 4.9 \text{ m}\cdot\text{s}^{-1}$, O₃ concentrations (i) decreased with NO₂ except when $T > 17 \text{ }^\circ\text{C}$ (without significant variation), (ii) increased with *T* except when $WS > 2.8 \text{ m}\cdot\text{s}^{-1}$ (O₃ presented a maximum between 17 and 20 °C), and (iii) did not change significantly with *WS* except when $T > 26 \text{ }^\circ\text{C}$ (presenting a decreasing tendency). For $WS > 4.9 \text{ m}\cdot\text{s}^{-1}$, O₃ concentrations (i) decreased with NO₂ except when $T > 17 \text{ }^\circ\text{C}$ (presenting a slight increase), (ii) presented high values for high *T* with all tested ranges of NO₂ concentrations (presenting a local maximum— $\approx 87.4 \mu\text{g}\cdot\text{m}^{-3}$ —for $T \approx 11 \text{ }^\circ\text{C}$ and $\text{NO}_2 \approx 6 \mu\text{g}\cdot\text{m}^{-3}$), (iii) increased with *T* for the tested range of *WS*, (iv) did not change significantly with *WS*, and (v) presented higher values than those where $WS \leq 4.9 \text{ m}\cdot\text{s}^{-1}$. The combined effect of *T*-NO₂ is in agreement with what was concluded in linear correlation analysis. The effect of NO₂ is more significant in the winter period, in which temperatures are usually low and NO₂ concentrations are high (see Figure 4). With high NO₂ concentrations, the chemical equilibrium given by Equation (R3) limits the increase in O₃ concentrations. In addition, based on a comparison of the two regimes defined by *WS*, the combined effect of these two variables presented similar behavior; however, O₃ concentrations were higher when $WS > 4.9 \text{ m}\cdot\text{s}^{-1}$ (49–104 $\mu\text{g}\cdot\text{m}^{-3}$) than they were when $WS \leq 4.9 \text{ m}\cdot\text{s}^{-1}$ (11–41 $\mu\text{g}\cdot\text{m}^{-3}$). High values of *WS* are associated with the dispersion of air pollutants, reducing their concentration. As NO₂ concentrations decrease, O₃ concentrations can achieve higher values [17,19,20,22]. The combined effect of *T*-*WS* showed the highest variability of O₃ with *T*, showing a positive correlation between these variables. The O₃ variability with *WS* is almost insignificant.

Regarding the combined effect NO₂-WS, similar conclusions were drawn: O₃ concentrations presented a decreasing tendency with NO₂ concentrations, and their variability with WS is almost insignificant. Figures S1 and S2 present the effect of the same combinations of variables on O₃ concentrations considering T (simulation II) and NO₂ (simulation III) as threshold variables, respectively. Similar analysis can be performed through these figures.

Table 1. ANN models: their input variables, activation functions (AF), number of hidden neurons (HN), and performance indexes (R², RMSE (root mean squared error) and d₂) for each performed simulation (Sim).

Sim	Model	AF	HN	R ² /RMSE/d ₂
I	$O_{3 t} = \begin{cases} \text{net}_1\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } WS \leq 4.9 \\ \text{net}_2\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } WS > 4.9 \end{cases}$	tansig radbas	8 8	0.71/14.7/0.91
II	$O_{3 t} = \begin{cases} \text{net}_1\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } T \leq 17.5 \\ \text{net}_2\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } T > 17.5 \end{cases}$	tansig tansig	7 7	0.72/14.5/0.91
III	$O_{3 t} = \begin{cases} \text{net}_1\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS\right), & \text{if } \text{NO}_2 \leq 26.6 \\ \text{net}_2\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } \text{NO}_2 > 26.6 \end{cases}$	tansig radbas	8 7	0.71/14.7/0.91
IV	$O_{3 t} = \begin{cases} \text{net}_1\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } WS \leq 4.9 \\ \text{net}_2\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } WS > 4.9 \end{cases}$	tansig radbas	8 8	0.71/14.7/0.91
V	$O_{3 t} = \begin{cases} \text{net}_1\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } T \leq 17.5 \\ \text{net}_2\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } T > 17.5 \end{cases}$	tansig tansig	7 7	0.72/14.5/0.91
VI	$O_{3 t} = \begin{cases} \text{net}_1\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS\right), & \text{if } \text{NO}_2 \leq 26.6 \\ \text{net}_2\left(\text{NO}, \text{NO}_2, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } \text{NO}_2 > 26.6 \end{cases}$	tansig radbas	8 8	0.71/14.7/0.91
VII	$O_{3 t} = \begin{cases} \text{net}_1\left(\text{NO}, \text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, WS\right), & \text{if } T \leq 17.5 \\ \text{net}_2\left(\text{NO}_2, \frac{\text{NO}_2}{\text{NO}}, T, RH, \frac{1}{RH}, WS, \frac{1}{WS}\right), & \text{if } T > 17.5 \end{cases}$	tansig tansig	7 7	0.72/14.5/0.91

The application of this statistical methodology allows for the determination of the influence of environmental and meteorological variables on O₃ concentration. Consequently, it is possible to develop more accurate predictive models for this secondary pollutant, which is important for the definition of policy measures for human health protection.

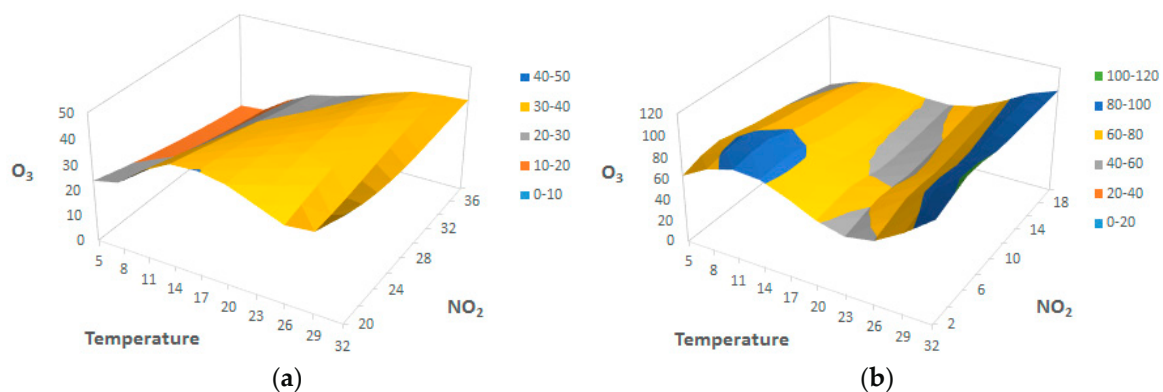


Figure 6. Cont.

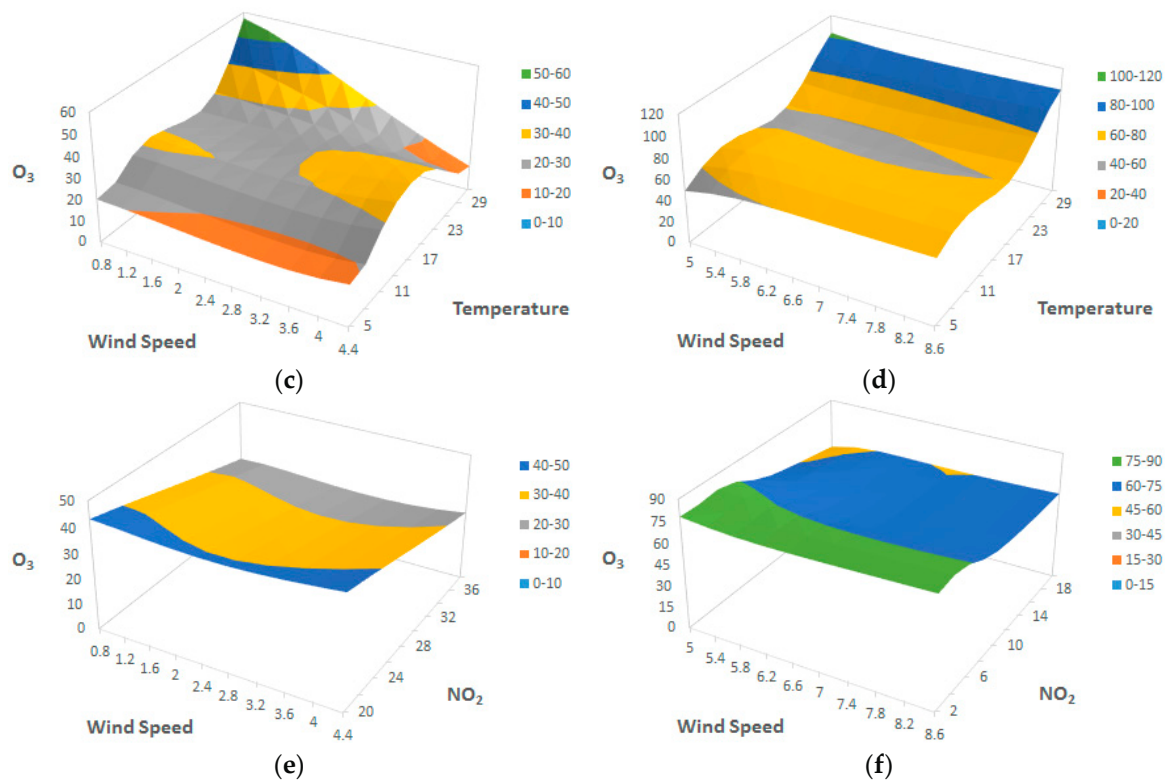


Figure 6. The combined effect of NO₂ concentrations, temperature, and wind speed on O₃ concentrations based on model determined in Simulation I where WS (wind speed) $\leq 4.9 \text{ m}\cdot\text{s}^{-1}$ (a,c,e) and where $WS > 4.9 \text{ m}\cdot\text{s}^{-1}$ (b,d,f).

4. Conclusions

Linear correlation analysis showed a positive relationship between O₃ concentrations with T and WS , while NO, NO₂, and RH showed a negative effect. In the studied period, the highest O₃ concentrations were observed for low NO_x concentrations and high wind speed. Threshold models with ANNs and those defined by genetic algorithms define three important variables that could define different O₃ regimes: (i) a wind speed of $4.9 \text{ m}\cdot\text{s}^{-1}$; (ii) a temperature of $17.5 \text{ }^\circ\text{C}$; and (iii) an NO₂ concentration of $26.6 \mu\text{g}\cdot\text{m}^{-3}$. The achieved models enabled the evaluation of the combined effect of two input variables in different O₃ regimes. This information may be useful for defining policy strategies for human health protection concerning surface ozone.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2076-3417/7/9/944/s1>. Figure S1: The combined effect of NO₂ concentrations, temperature, and wind speed on O₃ concentrations based on the model determined in Simulation II; Figure S2: The combined effect of NO₂ concentrations, temperature, and wind speed on O₃ concentrations based on the model determined in Simulation III.

Acknowledgments: This work was financially supported by Project POCI-01-0145-FEDER-006939 (LEPABE) funded by FEDER funds through COMPETE2020—Programa Operacional Competitividade e Internacionalização (POCI)—and by national funds through FCT—Fundação para a Ciência e a Tecnologia. J.C.M.P. acknowledges the FCT Investigator 2015 Programme (IF/01341/2015).

Author Contributions: N.F.A. collected the data, ran the simulations, analyzed the results, and wrote the paper; J.C.M.P. conceived the study, developed the model code, and revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lin, M.Y.; Horowitz, L.W.; Payton, R.; Fiore, A.M.; Tonnesen, G. Us surface ozone trends and extremes from 1980 to 2014: Quantifying the roles of rising asian emissions, domestic controls, wildfires, and climate. *Atmos. Chem. Phys.* **2017**, *17*, 2943–2970. [[CrossRef](#)]
2. Lefohn, A.S.; Malley, C.S.; Simon, H.; Wells, B.; Xu, X.; Zhang, L.; Wang, T. Responses of human health and vegetation exposure metrics to changes in ozone concentration distributions in the European Union, United States, and China. *Atmos. Environ.* **2017**, *152*, 123–145. [[CrossRef](#)]
3. Hollaway, M.J.; Arnold, S.R.; Collins, W.J.; Folberth, G.; Rap, A. Sensitivity of midnineteenth century tropospheric ozone to atmospheric chemistry-vegetation interactions. *J. Geophys. Res.-Atmos.* **2017**, *122*, 2452–2473. [[CrossRef](#)]
4. Proietti, C.; Anav, A.; De Marco, A.; Sicard, P.; Vitale, M. A multi-sites analysis on the ozone effects on gross primary production of European forests. *Sci. Total Environ.* **2016**, *556*, 1–11. [[CrossRef](#)] [[PubMed](#)]
5. Jiang, Z.; Miyazaki, K.; Worden, J.R.; Liu, J.J.; Jones, D.B.A.; Henze, D.K. Impacts of anthropogenic and natural sources on free tropospheric ozone over the Middle East. *Atmos. Chem. Phys.* **2016**, *16*, 6537–6546. [[CrossRef](#)]
6. Jorres, R.A.; Holz, O.; Zachgo, W.; Timm, P.; Koschyk, S.; Muller, B.; Grimminger, F.; Seeger, W.; Kelly, F.J.; Dunster, C.; et al. The effect of repeated ozone exposures on inflammatory markers in bronchoalveolar lavage fluid and mucosal biopsies. *Am. J. Respir. Crit. Care Med.* **2000**, *161*, 1855–1861. [[CrossRef](#)] [[PubMed](#)]
7. Frank, R.; Liu, M.C.; Spannhake, E.W.; Mlynarek, S.; Macri, K.; Weinmann, G.G. Repetitive ozone exposure of young adults—Evidence of persistent small airway dysfunction. *Am. J. Respir. Crit. Care Med.* **2001**, *164*, 1253–1260. [[CrossRef](#)] [[PubMed](#)]
8. Holz, O.; Mucke, M.; Paasch, K.; Bohme, S.; Timm, P.; Richter, K.; Magnussen, H.; Jorres, R.A. Repeated ozone exposures enhance bronchial allergen responses in subjects with rhinitis or asthma. *Clin. Exp. Allergy* **2002**, *32*, 681–689. [[CrossRef](#)] [[PubMed](#)]
9. McConnell, R.; Berhane, K.; Gilliland, F.; London, S.J.; Islam, T.; Gauderman, W.J.; Avol, E.; Margolis, H.G.; Peters, J.M. Asthma in exercising children exposed to ozone: A cohort study. *Lancet* **2002**, *359*, 386–391. [[CrossRef](#)]
10. Goldberg, M.S.; Burnett, R.T.; Brook, J.; Bailar, J.C.; Valois, M.F.; Vincent, R. Associations between daily cause-specific mortality and concentrations of ground-level ozone in Montreal, Quebec. *Am. J. Epidemiol.* **2001**, *154*, 817–826. [[CrossRef](#)] [[PubMed](#)]
11. Carter, C.A.; Cui, X.; Ding, A.; Ghanem, D.; Jiang, F.; Yi, F.; Zhong, F. Stage-specific, nonlinear surface ozone damage to rice production in China. *Sci. Rep.* **2017**, *7*, 44224. [[CrossRef](#)] [[PubMed](#)]
12. Morgan, P.B.; Ainsworth, E.A.; Long, S.P. How does elevated ozone impact soybean? A meta-analysis of photosynthesis, growth and yield. *Plant Cell Environ.* **2003**, *26*, 1317–1328. [[CrossRef](#)]
13. Sadiq, M.; Tai, A.P.K.; Lombardozzi, D.; Val Martin, M. Effects of ozone-vegetation coupling on surface ozone air quality via biogeochemical and meteorological feedbacks. *Atmos. Chem. Phys.* **2017**, *17*, 3055–3066. [[CrossRef](#)]
14. Christodoulakis, J.; Tzanis, C.G.; Varotsos, C.A.; Ferm, M.; Tidblad, J. Impacts of air pollution and climate on materials in Athens, Greece. *Atmos. Chem. Phys.* **2017**, *17*, 439–448. [[CrossRef](#)]
15. Banta, R.M.; Senff, C.J.; White, A.B.; Trainer, M.; McNider, R.T.; Valente, R.J.; Mayor, S.D.; Alvarez, R.J.; Hardesty, R.M.; Parrish, D.; et al. Daytime buildup and nighttime transport of urban ozone in the boundary layer during a stagnation episode. *J. Geophys. Res.-Atmos.* **1998**, *103*, 22519–22544. [[CrossRef](#)]
16. Monier, E.; Weare, B.C. Climatology and trends in the forcing of the stratospheric ozone transport. *Atmos. Chem. Phys.* **2011**, *11*, 6311–6323. [[CrossRef](#)]
17. Tong, L.; Zhang, H.L.; Yu, J.; He, M.M.; Xu, N.B.; Zhang, J.J.; Qian, F.Z.; Feng, J.Y.; Xiao, H. Characteristics of surface ozone and nitrogen oxides at urban, suburban and rural sites in Ningbo, China. *Atmos. Res.* **2017**, *187*, 57–68. [[CrossRef](#)]
18. Zhang, H.; Wang, Y.H.; Park, T.W.; Deng, Y. Quantifying the relationship between extreme air pollution events and extreme weather events. *Atmos. Res.* **2017**, *188*, 64–79. [[CrossRef](#)]
19. Carvalho, A.; Monteiro, A.; Ribeiro, I.; Tchepel, O.; Miranda, A.I.; Borrego, C.; Saavedra, S.; Souto, J.A.; Casares, J.J. High ozone levels in the northeast of Portugal: Analysis and characterization. *Atmo. Environ.* **2010**, *44*, 1020–1031. [[CrossRef](#)]

20. Pires, J.C.M.; Alvim-Ferraz, M.C.M.; Martins, F.G. Surface ozone behaviour at rural sites in Portugal. *Atmos. Res.* **2012**, *104*, 164–171. [[CrossRef](#)]
21. Sanchez, B.; Santiago, J.L.; Martilli, A.; Palacios, M.; Kirchner, F. Cfd modeling of reactive pollutant dispersion in simplified urban configurations with different chemical mechanisms. *Atmos. Chem. Phys.* **2016**, *16*, 12143–12157. [[CrossRef](#)]
22. Pires, J.C.M. Ozone weekend effect analysis in three European urban areas. *Clean-Soil Air Water* **2012**, *40*, 790–797. [[CrossRef](#)]
23. Gressent, A.; Sauvage, B.; Cariolle, D.; Evans, M.; Leriche, M.; Mari, C.; Thouret, V. Modeling lightning-nox chemistry on a sub-grid scale in a global chemical transport model. *Atmos. Chem. Phys.* **2016**, *16*, 5867–5889. [[CrossRef](#)]
24. Agatonovic-Kustrin, S.; Beresford, R. Basic concepts of artificial neural network (ann) modeling and its application in pharmaceutical research. *J. Pharm. Biomed.* **2000**, *22*, 717–727. [[CrossRef](#)]
25. Pires, J.C.M.; Goncalves, B.; Azevedo, F.G.; Carneiro, A.P.; Rego, N.; Assembleia, A.J.B.; Lima, J.F.B.; Silva, P.A.; Alves, C.; Martins, F.G. Optimization of artificial neural network models through genetic algorithms for surface ozone concentration forecasting. *Environ. Sci. Pollut. Res.* **2012**, *19*, 3228–3234. [[CrossRef](#)] [[PubMed](#)]
26. Tawadrous, A.S.; Katsabanis, P.D. Prediction of surface crown pillar stability using artificial neural networks. *Int. J. Numer. Anal. Methods* **2007**, *31*, 917–931. [[CrossRef](#)]
27. Pasini, A. Artificial neural networks for small dataset analysis. *J. Thorac. Dis.* **2015**, *7*, 953–960. [[PubMed](#)]
28. Keshavarzi, A.; Sarmadian, F.; Omran, E.-S.E.; Iqbal, M. A neural network model for estimating soil phosphorus using terrain analysis. *Egypt. J. Remote Sens. Space Sci.* **2015**, *18*, 127–135. [[CrossRef](#)]
29. Zare, M.; Pourghasemi, H.R.; Vafakhah, M.; Pradhan, B. Landslide susceptibility mapping at vaz watershed (iran) using an artificial neural network model: A comparison between multilayer perceptron (mlp) and radial basic function (rbf) algorithms. *Arab. J. Geosci.* **2013**, *6*, 2873–2888. [[CrossRef](#)]
30. Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed.; Addison-Wesley Professional: Boston, MA, USA, 1989.
31. Holland, J.H. *Adaptation in Natural and Artificial Systems*; University of Michigan Press: Ann Arbor, MI, USA, 1975.
32. Alvim-Ferraz, M.C.M.; Sousa, S.I.V.; Pereira, M.C.; Martins, F.G. Contribution of anthropogenic pollutants to the increase of tropospheric ozone levels in the oporto metropolitan area, Portugal since the 19th century. *Environ. Pollut.* **2006**, *140*, 516–524. [[CrossRef](#)] [[PubMed](#)]
33. Russo, A.; Trigo, R.M.; Martins, H.; Mendes, M.T. NO₂, PM10 and O₃ urban concentrations and its association with circulation weather types in Portugal. *Atmos. Environ.* **2014**, *89*, 768–785. [[CrossRef](#)]
34. Fernandez-Guisuraga, J.M.; Castro, A.; Alves, C.; Calvo, A.; Alonso-Blanco, E.; Blanco-Alegre, C.; Rocha, A.; Fraile, R. Nitrogen oxides and ozone in Portugal: Trends and ozone estimation in an urban and a rural site. *Environ. Sci. Pollut. Res.* **2016**, *23*, 17171–17182. [[CrossRef](#)] [[PubMed](#)]
35. Comrie, A.C. Comparing neural networks and regression models for ozone forecasting. *J. Air Waste Manag.* **1997**, *47*, 653–663. [[CrossRef](#)]
36. Abdul-Wahab, S.A.; Al-Alawi, S.M. Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. *Environ. Model. Softw.* **2002**, *17*, 219–228. [[CrossRef](#)]
37. Pires, J.C.M.; Alvim-Ferraz, M.C.M.; Pereira, M.C.; Martins, F.G. Evolutionary procedure based model to predict ground-level ozone concentrations. *Atmos. Pollut. Res.* **2010**, *1*, 215–219. [[CrossRef](#)]
38. Terui, N.; van Dijk, H.K. Combined forecasts from linear and nonlinear time series models. *Int. J. Forecast.* **2002**, *18*, 421–438. [[CrossRef](#)]
39. Shan, W.P.; Yin, Y.Q.; Zhang, J.D.; Ji, X.; Deng, X.Y. Surface ozone and meteorological condition in a single year at an urban site in central-eastern China. *Environ. Monit. Assess.* **2009**, *151*, 127–141. [[CrossRef](#)] [[PubMed](#)]
40. Pires, J.C.M.; Alvim-Ferraz, M.C.M.; Pereira, M.C.; Martins, F.G. Comparison of several linear statistical models to predict tropospheric ozone concentrations. *J. Stat. Comput. Simul.* **2012**, *82*, 183–192. [[CrossRef](#)]

41. Pires, J.C.M.; Martins, F.G.; Sousa, S.I.V.; Alvim-Ferraz, M.C.M.; Pereira, M.C. Selection and validation of parameters in multiple linear and principal component regressions. *Environ. Model. Softw.* **2008**, *23*, 50–55. [[CrossRef](#)]
42. Chen, H.H.; Stanier, C.O.; Young, M.A.; Grassian, V.H. A kinetic study of ozone decomposition on illuminated oxide surfaces. *J. Phys. Chem. A* **2011**, *115*, 11979–11987. [[CrossRef](#)] [[PubMed](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).