Effects of Weather and Anthropogenic Precursors on Ground-Level Ozone Concentrations in Malaysian Cities

Syaza Ezzati Baidrulhisham 1, Norazian Mohamed Noor 1,2,* , Zulkarnain Hassan 1, Andrei Victor Sandu 3,4,* , Petrica Vizureanu 3,5, Ahmad Zia Ul-Saufie 2,6, Mohd Remy Rozainy Mohd Arif Zainol 2,7,8,* , Aeslina Abdul Kadir 2,9 and György Deák 10

Abstract: Ground-level ozone ($O_3$) is a significant source of air pollution, mainly in most urban areas across the globe. Ground-level $O_3$ is not emitted directly into the atmosphere. It results from photochemical reactions between precursors and is influenced by weather factors such as temperature. This study investigated the spatial and temporal analysis of ground-level ozone and analyzed the significant anthropogenic precursors and the weather parameters associated with ground-level ozone during daytime and nighttime at three cities in peninsular Malaysia, namely, Kuala Terengganu, Perai, and Seremban from 2016 to 2020. Secondary data were acquired from the Department of Environment (DOE), Malaysia, including hourly data of $O_3$ with trace gases and weather parameters. The secondary data were analyzed using temporal analysis such as descriptive statistics, box plot, and diurnal plot as well as spatial analysis such as contour plot and wind rose diagram. Spearman correlation was used to identify the association of $O_3$ with its precursors and weather parameters. The results show that a higher concentration of $O_3$ during the weekend due to “ozone weekend effects” was pronounced, however, a slightly significant effect was observed in Perai. The two monsoonal seasons in Malaysia had a minimal effect on the study areas except for Kuala Terengganu due to its geographical location. The diurnal pattern of $O_3$ concentration indicates bimodal peaks of $O_3$ precursors during the peak traffic hours in the morning and evening with the highest intensity of $O_3$ precursors detected in Perai. Spearman correlation analysis determined that the variations in $O_3$ concentrations during day and nighttime generally coincide with the influence of nitrogen oxides (NO) and temperature. Lower NO concentration will increase the amount of $O_3$ concentration and an increasing amount of $O_3$ concentration is influenced by the higher temperature of its surroundings. Two predictive models, i.e., linear (multiple linear regression) and nonlinear models (artificial neural network) were developed and evaluated to predict the next day and nighttime $O_3$ levels. ANN resulted in better prediction for all areas with better prediction identified for daytime $O_3$ levels.

Keywords: air quality modeling; ozone; anthropogenic precursors; Spearman correlation; multiple linear regression; ANN
1. Introduction

Ground-level ozone ($O_3$) is one of the most common pollutants found in the Earth’s troposphere layers. It is therefore commonly referred to as a ‘bad’ ozone. The increased $O_3$ content in the Earth’s surface has consequences for human health, natural ecosystems, and vegetation development [1]. $O_3$ is primarily created by complex photochemical processes caused by anthropogenic precursors. A large percentage of ground-level ozone formation occurs in the presence of sunlight, typically the ultraviolet (UV) spectrum, when nitrogen oxides (NOx), carbon monoxide (CO), non-methane hydrocarbons (NmHCs), methane (CH4), and volatile organic compounds (VOCs) react in the atmosphere. These are referred to as ozone precursors [2]. Ozone absorbs radiation and thus works as a powerful greenhouse gas. Despite rising temperatures, $O_3$ affects evaporation rates, cloud formation, precipitation amounts, and atmospheric circulation. These effects are most noticeable in areas where ozone precursors are emitted [3].

Ozone is an uncommon pollutant since it is not emitted directly into the environment. The transformation of land use from forest to agronomic, residential, and town creates a change, causing the $O_3$ precursor to rise, which can be seen clearly between urban and rural areas [4]. This process usually results in the development of large concentrations of $O_3$ downwind and in suburban regions resulting in higher anthropogenic precursors causing a high concentration of $O_3$ [5]. A greater source of ground-level ozone can be found in urban areas due to greater emissions from vehicles, power stations, industrial boilers, refineries, and chemical factories. The increasing population, industrial activities, and automobiles are the concrete reason for significantly increased emissions of VOCs and NOx and, consequently, higher $O_3$ production can be expected in more populous areas.

In urban areas in particular, the concentration of NO released from traffic affects the formation of NO2. During the day, sunlight plays an important role in starting reactions among $O_3$ precursors [4,6]. A high concentration of NO produced by motor vehicles may result in reduced $O_3$ concentration in busy areas, but it will increase the concentration of NO2. High NO concentration will effective the titrate of atom (O) in urban areas due to emissions from the large amount of transportation activity. Through titration reactions with NO, $O_3$ is depleted, resulting in nitrogen dioxide (NO2) and oxygen (O2). The NO2 produced can react along with $O_3$ to produce nitrate (NO3) and O2 [7]. On the contrary, at night there is no sunlight to react with precursors forming ozone. Thus, the presence of NOx at night should have a significant decreasing impact on the concentrations of $O_3$ in an area. Many studies focus on the effect of ground-level ozone in the daytime rather than nighttime due to the significant relationship between $O_3$ with UVB. However, nighttime ozone is related to daytime ozone as the high preference may be carried over to the daytime. Likewise, the nighttime dispersion or accumulation of the precursors has a large impact on the ozone growth after sunrise. In addition, the overnight aloft ozone residual layer can play a role in morning time ozone rise [8].

In addition to $O_3$ precursors, meteorology plays a vigorous part in the atmospheric fate of ground-level $O_3$ in the lower boundary layer, including accumulation, dilution, and transport. Therefore, changes in meteorological conditions such as wind speed/direction, temperature, and relative humidity, can impact temporal variations of $O_3$ and anthropogenic precursors [9]. Correlation analysis is usually used to measure the relationship between the meteorological influences and parameters on tropospheric ozone [10,11]. Its positive or negative orientation provides a reliable reference for determining how specific climatic factors affect ozone concentrations. Temperature directly influences $O_3$ formation pathways by enhancing reaction rates of ozone precursors, ozone formation rates, and mechanism pathways [12]. On the one hand, high temperature is usually associated with sunshine duration, high solar irradiation, and low relative humidity that enhances ground-level ozone formation [13]. On the other hand, wind can significantly influence ground-level $O_3$ through dispersion and transportation [14]. The upwind and downwind
geographical features affect O\textsubscript{3} levels at a site, hence, favorable wind conditions will help transport O\textsubscript{3} and its precursors downwind [10]. In general, wind speed negatively correlates with O\textsubscript{3} levels. Tong et al. [11] revealed that low wind speed (<4 ms\textsuperscript{-1}) enhanced O\textsubscript{3} accumulation. The boundary layer is stable at low wind, thus stabilizing O\textsubscript{3} levels. As for Malaysia, which has tropical weather with two distinct monsoonal seasons—the wet monsoon, i.e., northeast monsoon (NEM) from November to March and the dry monsoon, i.e., southwest monsoon (SWM) from June to September, with inter-monsoon seasons in between—the influence of O\textsubscript{3} levels on the seasonal monsoon are yet to be discovered [15].

The prediction of ozone levels using mathematical tools is very useful in providing early warnings to the population. Thus, statistical approaches have been widely used by researchers to study variations in O\textsubscript{3} concentrations with its precursors and weather parameters. Multiple linear regression (MLR) is one of the most common techniques used in the prediction of ground-level O\textsubscript{3} levels. MLR models a linear relationship between the independent and the dependent variables, thus the relationship of the O\textsubscript{3} level and other variables (including other air pollutants, its precursors, and meteorology parameters) can be determined [16]. Therefore, due to its simplicity and reliability, MLR has been successfully used as a predictive tool worldwide [17–20]. Besides providing a simple linear relationship of ozone concentration with its precursors and weather parameters, MLR may not deliver accurate forecasts in some complex situations such as non-linear data and extreme values data. Artificial neural networks (ANN), one of the most common machine learning techniques, can be a useful tool to extract information from non-linear data such as air quality and meteorology. Currently, the application of neural networks has become more popular for predicting ground-level O\textsubscript{3} concentrations. A lot of researchers have effectively adopted ANN as a predictive tool to model O\textsubscript{3} concentration [21–25] and its precursor, NO\textsubscript{2} [26]. Even though many studies have been carried out in the world investigating the association between the weather and ozone concentration using MLR and ANN, there is still a shortage of work, especially in Malaysia.

In this research, the main aim was to analyze the spatial and temporal variability of O\textsubscript{3} concentrations with its precursors and weather parameters during daytime and nighttime in selected urban areas in peninsular Malaysia. Studying the characteristics of O\textsubscript{3} concentrations especially in the rarely pronounced urban areas in peninsular Malaysia are helpful toward a better understanding of the variation in O\textsubscript{3}. At the local level, the photochemical characteristics of O\textsubscript{3} vary with geographical location because the generation of O\textsubscript{3} is significantly affected by local environmental conditions, especially temperature, solar radiation, and atmospheric stagnation. In addition, two predictive models (i.e., linear and nonlinear models) were developed and attempted to forecast daytime and nighttime O\textsubscript{3} levels. Most of the O\textsubscript{3} level predictive models focused on daytime prediction, however, it is important to model the accumulation of O\textsubscript{3} at nighttime as some of the places are showing an increasing trend of O\textsubscript{3} during nighttime.

2. Materials and Methods

2.1. Air Pollutant Dataset

Continuous hourly air quality data from 2016 to 2020 were obtained from the Air Quality Division of the Department of Environment (DOE) Malaysia. This study used 12 variables that were divided into two groups. One group comprised seven air pollutants, such as O\textsubscript{3} (ppm), NO\textsubscript{x} (ppm), NO (ppm), SO\textsubscript{2} (ppm), NO\textsubscript{2} (ppm), CO (ppm), and PM\textsubscript{10} (\textmu g/m\textsuperscript{3}). While weather parameters are wind speed (km/h), temperature (\textdegree C), humidity (%), UVB (W/m\textsuperscript{2}), and wind direction (\textdegree). These data were sorted in Microsoft Excel according to the diurnal cycle since the O\textsubscript{3} level significantly varies during daytime and nighttime. The diurnal cycle of daytime O\textsubscript{3} was from 7:00 am to 7:00 pm and the nighttime cycle was from 8:00 pm to 6:00 am. Malaysia experiences a slightly longer daytime period due to its location near the equator.

In this study, urban areas located in peninsular Malaysia, namely, Kuala Terengganu, Perai, and Seremban were chosen. According to Wan Mahiyuddin et al. [27], ground-level
ozone exposure has reached levels that pose severe health hazards inside the larger Klang Valley sprawling city, which is Malaysia’s most densely populated area. However, there are very limited studies of O$_3$ variation reported in Kuala Terengganu, Seremban, and Perai. The details of the study areas are tabulated in Table 1.

Table 1. Details of Study Areas.

<table>
<thead>
<tr>
<th>Location</th>
<th>Station</th>
<th>Population [28]</th>
<th>Background of Study Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kuala Terengganu</td>
<td>Sekolah Kebangsaan Pusat Chabang Tiga</td>
<td>261,300</td>
<td>Urban Residential areas</td>
</tr>
<tr>
<td></td>
<td>(5.3066, 103.1217)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seremban</td>
<td>Kolej Vokasional Ampangan</td>
<td>467,800</td>
<td>Urban Residential areas</td>
</tr>
<tr>
<td></td>
<td>(2.7250, 101.9708)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sekolah Kebangsaan</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Seberang Jaya 2</td>
<td>630,999</td>
<td>Industrial areas Residential areas</td>
</tr>
<tr>
<td></td>
<td>(5.3979, 100.4043)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.2. Spatial Analysis

In this study, two spatial analyses were conducted: contour plot and wind rose diagram. Contour plot was used to examine the monthly distribution of O$_3$ levels at a regional spatial scale. The spatial interpolation method can effectively estimate the pollutant concentration at any point within the regional ambient air range. Applying the first law of geography, the ordinary kriging method (OKM) was used to analyze the continuous distribution of the concentration of O$_3$ among the stations. A monthly wind rose diagram was generated using Microsoft Excel.

2.3. Measure of Association Using Spearman Correlation

In this research, Spearman Correlation was used to analyze the significant precursors and weather parameters toward O$_3$ concentration and the results are presented in the form of a heat matrix. A nonparametric measure of the relationship between the two sets of ordinal-ranked values is Spearman rank correlation ($\rho$) [29]. A Spearman coefficient is essentially a Pearson correlation coefficient calculated using the ranks of the values of the two variables rather than their actual values. The formula is shown below [30]:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

(1)

where $\rho$ = Spearman’s Correlation coefficient; $d_i$ = difference between two ranks of each observation; and $n$ = number of observations.

The analysis was divided into two different time cycles: daytime and nighttime. Table 2 shows the description of the correlation using the following guide for the absolute value of "$\rho$":

Table 2. Description of the correlation related to the value of $\rho$ [31].

<table>
<thead>
<tr>
<th>Value of $\rho$</th>
<th>Station</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00–0.19</td>
<td>Very weak</td>
</tr>
<tr>
<td>0.20–0.39</td>
<td>Weak</td>
</tr>
<tr>
<td>0.40–0.59</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.60–0.79</td>
<td>Strong</td>
</tr>
<tr>
<td>0.80–1.00</td>
<td>Very strong</td>
</tr>
</tbody>
</table>

2.4. Prediction Model

In this study, the daytime and nighttime O$_3$ levels for the next day (the next 24 h) are predicted. As shown in Figure 1, data acquisition, exploration, cleaning, and partitioning
are part of the data preparation. The data acquisition describes the information of data and parameters included in this study (as presented in Section 2.1). Secondly, a descriptive analysis including central tendency (mean and median) and dispersion (standard deviation) analyses were measured in data exploration. Next, data cleaning explained the technique involved in imputing the missing observation. In this study, expectation maximization (EM) was used to fill in the missing data as this method was reported as the most reliable method in estimating missing air pollutant observation data [32]. Before developing the model, the original dataset was partitioned into two datasets for training and validation. Out of the total data, 80% were used to develop the model, whereas the rest of the data were used to validate the model. Parameters that had moderate to strong correlation ($\rho \geq 0.4$) with the $O_3$ level from Spearman correlation analysis were used as the inputs for the prediction models. The details of predictive models are discussed in Sections 2.4.1 and 2.4.2 and the performance evaluation for comparing the performances of the model is explained in Section 2.4.3.

Figure 1. Modeling framework.

2.4.1. Multiple Linear Regression (MLR)

MLR is a conventional method used to model air pollutants. MLR models the connection between two or more independent variables and a dependent variable by fitting a linear equation to the observed data. Multiple linear regression analysis is a statistical technique that can be used to examine the relationship between a single dependent (criterion) variable, $Y$, and several independent (predictor) variables, $X_s$. A random response $Y$ relates to a set of independent variables $X_1, X_2, \ldots, X_i$ based on the multiple regression model. The equation can be read as [33]:

$$Y_i = \beta_0 + \beta_1 \cdot X_1 + \ldots + \beta_k \cdot X_i + \epsilon_i \quad (2)$$

where, $i$ is equal to $n$ observations; $Y_i$ = dependent variable (daytime and nighttime $O_3$ levels); $X_i$ is the explanatory variables (ozone precursors and weather parameters); $\beta_0$ is the y-intercept (constant term); $\beta_k$ is the slope coefficients for each explanatory variable; and $\epsilon$ = the model’s error term (also known as the residuals).
2.4.2. Artificial Neural Network (ANN)

A feed-forward backpropagation neural network (FFBP) was used in this study. The structure of FFBP was composed of three layers of neurons called the input, hidden, and output layers. The first layer of neurons consisted of an input layer, representing independent variables. The second layer was the hidden layer, which is responsible for processing the input weight from the input layer and transferring it to the output layer. The third layer was the output layer, which represents the next-day prediction of daytime and nighttime O$_3$ levels. The operation was conducted using the usual sigmoid function as the activation function.

2.4.3. Performance Indicators

Performance indicators are used to evaluate the predictive model by MLR and ANN. It was used to assess the model’s applicability in predicting daytime and nighttime O$_3$ concentrations. The performance measures applied in this study are mean absolute error (MAE), root mean squared error (RMSE), and correlation ($r$) Table 3 [34].

<table>
<thead>
<tr>
<th>Performance Indicators</th>
<th>Equation</th>
<th>Better Predictability if</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Absolute Error</td>
<td>$MAE = \frac{\sum_{i=1}^{n}</td>
<td>P_i - O_i</td>
</tr>
<tr>
<td>Root Mean Squared Error</td>
<td>$RMSE = \frac{1}{n-1} \sum_{i=1}^{n} (P_i - O_i)^2$</td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>$r = \frac{\sum_{i=1}^{n}(P_i - \overline{P})(O_i - \overline{O})}{(n-1)S_P S_O}$</td>
<td>Closer to $-1$ or $+1$</td>
</tr>
</tbody>
</table>

where;
- $n$ = total number of hourly measurements at a particular site;
- $P_i$ = predicted values of a single set of hourly monitoring data;
- $O_i$ = the observed values from a single set of hourly monitoring records;
- $\overline{P}$ = the mean of the predicted values from a single set of hourly monitoring data;
- $\overline{O}$ = the mean of the observed values from a single set of hourly monitoring data;
- $S_P$ = the standard deviation of the predicted values from a single set of hourly monitoring data;
- $S_O$ = the standard deviation of the observed values from a single set of hourly monitoring data.

3. Results and Discussion

3.1. Spatiotemporal Variation of Ground-Level Ozone

Table 4 represents a data summary for ground-level ozone measurement in the three study areas for the period of 5 years. Overall, the mean concentration of ozone levels for Perai, Kuala Terengganu, and Seremban is below the limit (0.1 ppm) as postulated in the Malaysia Ambient Air Quality Standard (MAAQS). The maximum O$_3$ level in 2018 was recorded as the highest value compared with all study areas during the five year period. The highest ozone concentration reading in Perai was 0.095 ppm in 2018 with a mean of 0.0175 ppm while the lowest maximum value was 0.080 ppm in 2020 with a mean value of 0.01718 ppm. From the mean concentration, Perai and Seremban showed an incremental trend from 2015 to 2019 whereas Kuala Terengganu showed a decremental trend of O$_3$ levels. As a result, the smallest range of data was observed, i.e., 0 ppm to 0.009 ppm. Seremban recorded the highest concentration of O$_3$ starting from 2018 with a mean value of 0.022 ppm. Lower concentrations were recorded in Kuala Terengganu and the lowest was in Perai. The standard deviation indicates that the O$_3$ measurement records showed less variables and a smaller range of O$_3$ levels was observed in Kuala Terengganu.
Table 4. Descriptive statistic of ground-level ozone concentration in Perai, Seremban, and Kuala Terengganu from 2016 to 2020.

<table>
<thead>
<tr>
<th>Study Area</th>
<th>Year</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Valid</td>
<td>Missing</td>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>8237</td>
<td>5513</td>
<td>8260</td>
<td>8350</td>
<td>8292</td>
</tr>
<tr>
<td></td>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.01543</td>
<td>0.01629</td>
<td>0.01749</td>
<td>0.01757</td>
<td>0.01718</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.01100</td>
<td>0.01050</td>
<td>0.01140</td>
<td>0.01090</td>
<td>0.01185</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>0.014710</td>
<td>0.16441</td>
<td>0.017208</td>
<td>0.017544</td>
<td>0.015890</td>
</tr>
<tr>
<td></td>
<td>Minimum</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.091</td>
<td>0.085</td>
<td>0.095</td>
<td>0.093</td>
<td>0.080</td>
</tr>
<tr>
<td>Seremban</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>6478</td>
<td>5886</td>
<td>8327</td>
<td>8356</td>
<td>8373</td>
</tr>
<tr>
<td></td>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>2306</td>
<td>928</td>
<td>426</td>
<td>406</td>
<td>409</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.01686</td>
<td>0.01748</td>
<td>0.02030</td>
<td>0.02194</td>
<td>0.01943</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.01200</td>
<td>0.01200</td>
<td>0.01600</td>
<td>0.01930</td>
<td>0.01680</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>0.014792</td>
<td>0.016436</td>
<td>0.017167</td>
<td>0.016930</td>
<td>0.014807</td>
</tr>
<tr>
<td></td>
<td>Minimum</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.097</td>
<td>0.103</td>
<td>0.100</td>
<td>0.095</td>
<td>0.094</td>
</tr>
<tr>
<td>Kuala Terengganu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>8335</td>
<td>5975</td>
<td>8215</td>
<td>8320</td>
<td>8337</td>
</tr>
<tr>
<td></td>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>449</td>
<td>836</td>
<td>558</td>
<td>440</td>
<td>447</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.01981</td>
<td>0.01776</td>
<td>0.01739</td>
<td>0.01559</td>
<td>0.01412</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.01800</td>
<td>0.01670</td>
<td>0.01600</td>
<td>0.01520</td>
<td>0.01350</td>
</tr>
<tr>
<td></td>
<td>Std. Deviation</td>
<td>0.013018</td>
<td>0.012303</td>
<td>0.012521</td>
<td>0.011184</td>
<td>0.009916</td>
</tr>
<tr>
<td></td>
<td>Minimum</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.076</td>
<td>0.082</td>
<td>0.069</td>
<td>0.058</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Figure 2 represents the box plot of hourly O\textsubscript{3} concentrations in Seremban, Perai, and Terengganu from 2016 to 2020. The O\textsubscript{3} levels in Seremban and Perai were observed to have higher variability in 2018 and 2019. Seremban had the maximum O\textsubscript{3} level of 0.103 ppm in 2019 and recorded the lowest maximum concentration in 2020 with the value of 0.094 ppm. The highest concentration in Perai was 0.095 ppm in 2018 while the concentration decreased in 2019 and 2020 with maximum values of 0.093 ppm and 0.80 ppm, respectively. The data distribution for all years at Perai were skewed to the right which indicates a positive skew. Conversely, the concentration of O\textsubscript{3} at Kuala Terengganu was the lowest compared with Perai and Seremban. The maximum O\textsubscript{3} concentration was 0.082 ppm in 2017 while the lowest was 0.057 ppm in 2020 as illustrated in Figure 2. Kuala Terengganu experienced a decrease in O\textsubscript{3} concentration values by year starting from 2016 to 2020. It was observed that O\textsubscript{3} levels in Kuala Terengganu were less variable compared with Seremban and Perai. Overall, the distribution of O\textsubscript{3} levels in all study areas were skewed to the right which indicates that extreme concentration occurred. Seremban was observed to have a high exceedance of O\textsubscript{3} levels especially in 2017. The monitoring station observed a decrease in 2020 due to less transportation use and factory shutdowns because of epidemic transmission of COVID-19. Malaysia’s response to the COVID-19 pandemic resulted in unprecedented reductions in economic activity in 2019 to 2021. After accounting for weather changes, lockdown events lowered population-weighted concentrations of NO\textsubscript{2} and particulate matter by around 60% and 31%, respectively, as well as O\textsubscript{3} concentrations [32]. Reductions in transportation sector emissions are largely responsible for the NO\textsubscript{2} variances. The different results in three monitoring stations are believed to be influenced by variations in temperature, humidity, and wind speed.
Figure 2. Hourly box plot of O₃ levels in Seremban, Perai, and Kuala Terengganu. The asterisk (*) on the boxplot indicate the extreme outlier that lie >1.5 times the size of the box.

Monthly variations of tO₃ levels for the three study areas can be analyzed using contour plots and a wind rose diagram as depicted in Figure 3. During the early months of the year, Perai recorded the highest concentrations compared with Seremban and Kuala Terengganu with mean O₃ concentrations of 0.0168 ppm in January, 0.0213 ppm in February, and 0.0219 ppm in March. Perai experienced very light wind (1 km/h to 3 km/h) in January and February that came from the southeast and light wind (3 km/h to 5 km/h) in March from the southeast and south. Latif et al. [6] reported that the northern area of Peninsular Malaysia experienced a hot and dry wind from the northeast, especially from Indochina and the South China Sea to the north of peninsular Malaysia. Nonetheless, the wind rose diagram indicates that Perai did not receive any wind from the northeast direction; hence, the high O₃ levels during this month were most probably due to the low speed of wind. On the other hand, Seremban was observed to experience the highest concentrations of O₃ levels during the inter-monsoon (April) with the value of 0.0224 ppm. During this month, Seremban mostly received light wind (3 km/h to 5 km/h).

Perai once again recorded the highest levels of O₃ compared with other areas during the early dry monsoon (June and July), which is the southwest monsoon, while Seremban recorded high O₃ levels at the end of the summer monsoon (September) and during inter-monsoon (October and November). Weak wind (1 km/h to 5 km/h) was detected during the early season of the dry monsoon in Perai which might have influenced the dispersion of pollutants [10]. A gentle breeze (12 km/h to 19 km/h) is at least required to impact the greater dispersion of air pollutants, resulting in lower air pollution concentrations in areas with stronger winds. A similar situation can be observed in Seremban where very light wind (1 km/h to 3 km/h) was detected especially from October to December resulting in almost stagnant air in the inland areas.
Figure 3. Cont.
Figure 3. Monthly contour plot and wind rose diagram of ground-level O$_3$ in Seremban, Kuala Terengganu and Perai.
In the other hand, O$_3$ levels in Kuala Terengganu were detected to be higher at the end of the dry monsoon, i.e., in September and October. Light air (3 km/h to 5 km/h) from the south and southwest was detected in Kuala Terengganu during this time, hence confirming the influence of the southwest monsoon on O$_3$ levels at this area.

Overall, higher O$_3$ levels at all areas can be observed from June to October, i.e., during the southwest monsoon (dry monsoon). However, only Kuala Terengganu shows the influence of the monsoon in contributing to higher O$_3$ levels in that area. Other areas, i.e., Perai and Seremban received light wind (1 km/h to 5 km/h) from the south and southeast. However, during this hot and dry climate, a low amount of precipitation and lower wind speed can be expected, hence it might influence the formation of ozone via complex photochemical reactions with other pollutants in the atmosphere that remain longer in the atmosphere due to poor dispersion [35].

Figure 4 shows weekly variations of O$_3$ levels in Seremban, Kuala Terengganu, and Perai. Generally, the same diurnal trend of O$_3$ variations were observed for the three monitored areas. On a weekly basis, the weekend and weekdays exhibited higher O$_3$ levels starting from 10:00 am until 5:00 pm, whereas the O$_3$ levels started to decrease starting at 6:00 pm until the next-day morning at around 9:00 am. The lowest O$_3$ level was detected in Perai at 8:00 am during the weekday compared with other places, while the highest O$_3$ concentration was found in Seremban at 5:00 pm during the weekdays.

Seremban showed a higher distribution of O$_3$ levels during weekdays compared with the weekend. Extreme concentrations of O$_3$ levels were observed early in the morning (from 1:00 am to 9:00 am) and in the evening (3:00 pm to 6:00 pm). During nighttime, the absence of UV light reduced the O$_3$ photochemical reaction rate and the O$_3$ concentrations because of continuous chemical degradation by NO and titration [36].

In Kuala Terengganu, a higher variability of O$_3$ concentration was monitored during weekdays compared with the weekend. However, during the weekend, a higher concentration of O$_3$ was detected in the morning (from 8:00 am to 12:00 pm) and the concentration was slightly low during the peak wavelength of solar radiation (from 2:00 pm to 4:00 pm) if compared with the variation during weekdays. However, the O$_3$ level increased again from 5:00 pm onwards compared with the weekday variations. Similar O$_3$ variations can be found in Perai with a higher degree of difference in O$_3$ between the weekend and weekdays, especially in Perai because the monitoring areas are located next to the main highway of the South–North and Butterworth–Kulim Express Highway, which results in higher concentrations of O$_3$ during the weekend because many people use the highway to travel to their hometowns on weekends.

Higher O$_3$ levels during the weekend compared with weekdays may be caused by the differences in the number of vehicles. A high level of ozone on the weekend is a well-known phenomenon known as the “ozone weekend effect” (OWE) [37]. The OWE is not pronounced in all areas except for Perai. Low levels of traffic on weekends are thought to be the primary source of OWE, as low levels of NO emission reduce local ozone elimination. The greater amount of ozone compared with other weekdays is most likely due to the residual effect of ozone from the weekend. Ozone concentrations in all cities follow a similar twenty-four-hour cycle, with lower levels during the night and early morning hours and higher levels during the day. The largest concentrations are obtained in the afternoon from 12:00 am to 5:00 pm due to photochemical ozone generation as a result of high UV radiation intensity during these hours. During this time, the high release of vehicle emissions contributes to the high amount of NO that will react with more O$_3$ to produce NO$_2$. Hence, lower O$_3$ concentrations during this period can be expected during weekdays compared with weekends due to traffic emissions. Conversely, the concentration of ozone decreases after 5:00 pm until the next morning due to less intense or no UV radiation [38]. In addition, the population and the background of the area play a major role in influencing the OWE and the amount of O$_3$ produced daily.
Figure 4. Diurnal plot of O_3 level during weekdays and weekends. The asterisk (*) on the boxplot indicate the extreme outlier that lie >1.5 times the size of the box.
3.2. Association of O$_3$ Level with Its Precursors

Figure 5 shows the diurnal pattern of ground-level O$_3$ and its precursors, namely, NO, NO$_x$, NO$_2$, and CO, at all monitoring stations. Generally, a similar pattern of variation of O$_3$ levels with its precursors can be observed for all places except different intensities of pollutants. Higher concentrations of O$_3$ precursors, i.e., CO, NO, NO$_x$, and NO$_2$ were monitored in the morning (7:00 am to 9:00 am) and evening (6:00 pm to 10:00 pm) to coincide with people travelling to perform their daily routines. Photochemical reactions of O$_3$ precursors to form O$_3$ took place more significantly when greater solar radiation was detected during the day starting from 9:00 am. The highest O$_3$ formation was observed in Seremban and Seberang Perai during the maximum intensity solar radiation period (2:00 pm to 4:00 pm). O$_3$ concentration started to drop significantly at the night when no UVB radiation was available.

![Diurnal plot of O$_3$ precursors in (a) Seremban, (b) Kuala Terengganu, and (c) Perai.](image-url)

Figure 5. Diurnal plot of O$_3$ precursors in (a) Seremban, (b) Kuala Terengganu, and (c) Perai.
A noticeable higher concentration of NO, NO\textsubscript{x}, and CO were observed in Perai compared with other areas. There is a lot of housing near the monitoring station and the whole Seberang Perai population was about 428,500 in 2020 [28]. The Butterworth–Kulim Express highway and North–South highway are close to the monitoring area. Furthermore, an industrial area in the Free Industrial Zone (FIZ) is also located within 5 km of the monitoring area. Hence, high air pollutant emissions from vehicles can be expected. High NO concentrations will effective atom titration of (O) in urban areas due to emissions from the large amount of transportation. The negative relationship between O\textsubscript{3} and NO\textsubscript{2} is expected as NO\textsubscript{2} is a precursor of O\textsubscript{3} [39]. This phenomenon is shown in (Equation (3)):

\[
\text{NO} + \text{O}_3 \rightarrow \text{NO}_2 + \text{O}_2 \tag{3}
\]

Hence, it can be seen that the O\textsubscript{3} levels in Perai were significantly less if compared with its precursors, i.e., NO and NO\textsubscript{2}. If compared with Seremban and Kuala Terengganu, the least amount of NO and NO\textsubscript{2} were released and the amount of O\textsubscript{3} formation was greater compared with its precursors. The NO formed inside vehicle engines reacts with O\textsubscript{2} in the air to form NO\textsubscript{2}. Ozone is formed when a combination of nitrogen (N) and O\textsubscript{2} is exposed to UVB by the photolysis of NO\textsubscript{2}. ‘M’ represents an unstable molecule such as nitrogen (N\textsubscript{2}) or O\textsubscript{2}, and O (\textsuperscript{3}P) represents an equilibrium state or low-energy oxygen atom that interacts with molecular O\textsubscript{2} to produce O\textsubscript{3} [40]:

\[
\text{NO}_2 + h\nu \rightarrow \text{NO} + \text{O (}\textsuperscript{3}\text{P}) \tag{4}
\]

\[
\text{O} + \text{O}_2 + \text{M} \rightarrow \text{O}_3 + \text{M} \tag{5}
\]

Since in Seremban and Kuala Terengganu there is less NO compared with Perai, less O\textsubscript{3} was converted to NO\textsubscript{2} and O\textsubscript{2} as in Reaction (3). In addition, Perai has the largest value of CO due to the monitoring station being located at a main highway, and slow traffic during peak hours influenced CO concentrations. The location located near the FIZ also contributed many pollutions to that area. While during midday, the photochemical O\textsubscript{3} formation consumes CO which causes O\textsubscript{3} concentrations to increase and CO concentrations to decrease, which explains the negative association [41].

The oxidation of trace gases in the atmosphere, such as VOCs, probably contributed to O\textsubscript{3} formation in this location, which was largely initiated by the reaction with OH whose concentration is controlled by the concentrations of local VOCs and NO\textsubscript{x} [42]. The phenomenon is shown below [41]:

\[
\text{CO} + \text{OH} \rightarrow \text{H} + \text{CO}_2 \tag{6}
\]

\[
\text{H} + \text{O}_2 + \text{M} \rightarrow \text{HO}_2 + \text{M} \tag{7}
\]

\[
\text{HO}_2 + \text{NO} \rightarrow \text{NO}_2 + \text{OH} \tag{8}
\]

Then, NO\textsubscript{2} undergoes photolysis and O\textsubscript{3} is formed according to Reactions (4) and (5).

3.3. Association of O\textsubscript{3} Level with Other Trace Gases and Weather Parameters

Figure 6 illustrates the Spearman Correlation heat matrix for daytime and nighttime in Perai, Kuala Terengganu, and Seremban. Overall, O\textsubscript{3} concentration was negatively correlated with its precursors, i.e., NO\textsubscript{2}, NO\textsubscript{x}, NO, and CO for all places during day and nighttime. While for weather parameters, the most correlated parameter was temperature followed by wind speed, UVB, and humidity. All weather parameters were positively correlated with O\textsubscript{3} except for humidity, which was negatively correlated. Remarkably, a moderate to strong correlation of CO with NO\textsubscript{2}, NO\textsubscript{x}, and NO were observed in the study areas with a higher correlation value during nighttime.
Figure 6. Spearman correlation matrix of O₃ with trace gases and weather parameters: 
(a) Seremban daytime, (b) Seremban nighttime, (c) Kuala Terengganu daytime, (d) Kuala Terengganu nighttime, 
(e) Perai daytime, and (f) Perai nighttime.
Figure 6a shows the association of daytime and nighttime $O_3$ levels with its precursors and weather parameters at Seremban. A moderate to strong correlation of $O_3$ with weather parameters during daytime was observed with the strongest positive correlation in the temperature ($\rho = 0.605$). A moderate negative correlation was detected with NO and a weak negative correlation of $O_3$ with other precursors (NO$_x$ and CO) were calculated. A similar trend of association of $O_3$ levels with its precursors and weather parameters were observed during nighttime except with lower calculated $\rho$ values. A very strong to strong correlation of NO$_x$ with NO$_2$ and NO was observed during daytime and nighttime in Seremban due to its inter-convertible reaction from NO to NO$_2$ and reactions of NO$_x$ with VOCs.

In Kuala Terengganu, a similar association between $O_3$ levels with the gases and weather parameters during daytime and nighttime was observed if compared with Seremban (Figure 6c,d). Temperature was positively correlated while humidity was negatively correlated with $O_3$ while its precursors also showed a negative correlation with ozone during daytime and nighttime. However, a higher calculated $\rho$-value between $O_3$ level and wind speed was calculated in Kuala Terengganu compared with other places as it is situated near the coastal area, hence land and sea breeze might have influences on air pollutant dispersion.

In Perai, the highest correlation value of CO with $O_3$ concentration was calculated compared with other areas with the value of $-0.468$. In addition, during the daytime, a strong negative correlation was observed with NO ($\rho = -0.704$), temperature, and humidity with $\rho$-values of $-0.742$ and $-0.684$, respectively, while during the nighttime, a strong negative correlation was detected with NO ($\rho = -0.591$). NO$_x$ recorded a moderate correlation with $O_3$ that was ($\rho = -0.467$). CO not only had a high correlation with $O_3$, it also had a significant positive correlation with NO$_x$, NO$_2$, and NO compared with other areas of study with higher values observed during nighttime.

Overall, high temperature contributes to photochemical ozone chemistry which is significant in the generation of secondary aerosols such as sulfate (SO$_4^{2-}$) and nitrate (NO$_3^-$) that are generated from SO$_2$ and NO$_x$ [43]. This mechanism most likely reduces the negative link between the two species over time. NO is significantly associated with $O_3$ concentration, as increases in its concentration are usually influenced by the complex reaction between NO$_x$ and VOCs in the atmosphere produced by motor vehicles and industrial combustion processes [44]. Humidity is negatively correlated with $O_3$ due to humid days with enhanced cloud cover and thus results in a reduced photochemistry reaction in $O_3$ formation [42]. $O_3$ level has a positive association with temperature due to its formation requiring high UV radiation. Conversely, temperature is negatively correlated with humidity resulting in less humidity in the ambient air with the increase in temperature [43]. Obviously in Perai, SO$_2$ concentration is among the highest associated with $O_3$ compared with other places due to industrial areas being located near the monitoring station. According to Latif et al. [45], the fluctuations of SO$_2$ in the study area may be due to the contributions of SO$_2$ from local and regional sources of biomass burning. As Penang Port is near the monitoring areas, it might be the cause of the high concentrations of SO$_2$ in that area. The distribution of SO$_2$ is influenced by the shipping port because it transfers SO$_2$ to the surrounding areas [46].

In conclusion, parameters that had moderate to strong correlation ($\rho \geq 0.4$) with $O_3$ level were used as input for daytime and nighttime prediction of $O_3$ levels. In the next section, the performances of the linear and nonlinear models were compared and evaluated using several performance measures.

### 3.4. Prediction of Daytime and Nighttime $O_3$ Levels

The linear relationship of $O_3$ levels in daytime and nighttime at all study areas are tabulated in Table 5. The independent parameters for the linear model were the parameters that were strong and moderately correlated ($\rho > 0.4$) with $O_3$ level. From the table, it can be seen that $O_3$ level was correlated with previous $O_3$ concentrations, NO$_x$, and NO. However, no relation was detected with weather parameters indicating that very minimal
to no correlation between weather parameters and O\textsubscript{3} levels was identified using MLR. NO is a significant predictor of variables during the daytime at Seremban. However, NO and temperature are significant during the nighttime. The O\textsubscript{3} concentration decreased 0.027 units when the one unit of NO decreased during the daytime and 0.042 and 0.001 units of NO and temperature during nighttime. In addition, Kuala Terengganu indicated an increase of one unit of NO\textsubscript{x} will result in an increase of 0.331 units of O\textsubscript{3} during the daytime and a decrease of 0.024 of NO during the nighttime. Perai showed that the increase of one unit of NO\textsubscript{x} will result in an increase of 0.197 units of O\textsubscript{3}. The O\textsubscript{3} concentration decreased about 0.002 and 0.325 when the one unit of CO and NO decreased.

Table 5. Model summary for O\textsubscript{3} level predictions.

<table>
<thead>
<tr>
<th>Study Area</th>
<th>Time</th>
<th>Predict Hour</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seremban</td>
<td>Daytime</td>
<td>Next 12 h</td>
<td>(O_{3,t+12} = 0.004 + 0.477 O_3 - 0.027 NO + 0 T + 0 UVB + 0 H)</td>
</tr>
<tr>
<td></td>
<td>Nighttime</td>
<td>Next 12 h</td>
<td>(O_{3,t+12} = 0.047 + 0.276 O_3 - 0.042 NO + 0 WS - 0.001 T + 0 H)</td>
</tr>
<tr>
<td>Kuala Terengganu</td>
<td>Daytime</td>
<td>Next 12 h</td>
<td>(O_{3,t+12} = 0.022 + 0.453 O_3 + 0.311 NO_X - 0.449 NO + 0 WS + 0 T + 0 UVB + 0 H)</td>
</tr>
<tr>
<td></td>
<td>Nighttime</td>
<td>Next 12 h</td>
<td>(O_{3,t+12} = 0.001 + 0.487 O_3 - 0.024 NO + 0 WS + 0 T + 0 H)</td>
</tr>
<tr>
<td>Perai</td>
<td>Daytime</td>
<td>Next 12 h</td>
<td>(O_{3,t+12} = 0.09 + 0.403 O_3 - 0.002 CO + 0.197 NO_X - 0.325 NO + 0 T + 0 UVB + 0 H)</td>
</tr>
<tr>
<td></td>
<td>Nighttime</td>
<td>Next 12 h</td>
<td>(O_{3,t+12} = 0.44 + 0.207 O_3 + 0.002 NO_X - 0.091 NO - 0.001 T + 0 H)</td>
</tr>
</tbody>
</table>

Validation of MLR and ANN predictive models were performed using performance indicators. Table 6 shows the results of the predictive model evaluation of O\textsubscript{3} levels using mean absolute error (MAE), root mean squared error (RMSE), and correlation (\(r\)). Generally, ANN outperformed MLR in predicting the daytime and nighttime O\textsubscript{3} levels in the study areas as a lower error measurement and higher correlation value were calculated. Both of the models estimated better daytime O\textsubscript{3} levels compared with nighttime O\textsubscript{3} in Kuala Terengganu and Perai with average \(r\) differences between daytime and nighttime of 6.66% and 6.97% using MLR and ANN, respectively. Meanwhile, in Seremban, ANN predicted better nighttime O\textsubscript{3} levels compared with daytime and the performances of ANN was observed to be the worst in Seremban (\(r = 0.699\)) compared with other areas (\(r\) value ranging from 0.765 to 0.862). Similarly, MLR predicted better nighttime O\textsubscript{3} concentration (\(r = 0.615\)) and MLR prediction in Seremban was observed to be the worst compared with other places with \(r\) values ranging from 0.683 to 0.747.

Table 6. Predictive Model Performance.

<table>
<thead>
<tr>
<th>Study Area</th>
<th>Time</th>
<th>MAE</th>
<th>RMSE</th>
<th>(r)</th>
<th>MAE</th>
<th>RMSE</th>
<th>(r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seremban</td>
<td>Daytime</td>
<td>0.015</td>
<td>0.019</td>
<td>0.576</td>
<td>0.010</td>
<td>0.012</td>
<td>0.699</td>
</tr>
<tr>
<td></td>
<td>Nighttime</td>
<td>0.014</td>
<td>0.015</td>
<td>0.615</td>
<td>0.005</td>
<td>0.006</td>
<td>0.753</td>
</tr>
<tr>
<td>Kuala Terengganu</td>
<td>Daytime</td>
<td>0.014</td>
<td>0.017</td>
<td>0.739</td>
<td>0.006</td>
<td>0.007</td>
<td>0.805</td>
</tr>
<tr>
<td></td>
<td>Nighttime</td>
<td>0.016</td>
<td>0.021</td>
<td>0.711</td>
<td>0.004</td>
<td>0.006</td>
<td>0.795</td>
</tr>
<tr>
<td>Perai</td>
<td>Daytime</td>
<td>0.061</td>
<td>0.064</td>
<td>0.747</td>
<td>0.006</td>
<td>0.008</td>
<td>0.862</td>
</tr>
<tr>
<td></td>
<td>Nighttime</td>
<td>0.012</td>
<td>0.013</td>
<td>0.683</td>
<td>0.005</td>
<td>0.007</td>
<td>0.765</td>
</tr>
</tbody>
</table>

Ground-level ozone is produced by nonlinear photochemical reactions involving its precursors and is affected by many factors including precursor concentrations, local meteorological conditions, air transportation, deposition, etc. These factors control the production, loss, and transformation of O\textsubscript{3}, and thus its concentrations and spatial distribution. Hence,
ANN that is a nonlinear model can better represent the complex relationship of O$_3$ with its precursors and weather parameters. Feed forward backpropagation neural network (FFBP) is one of the most known ANNs. In several applications, including regression problems, classification problems, or time series predictions, FFBP has been successfully implemented using single autoregressive models [47]. Each input neuron represents one of the input variables (precursors or weather parameters), which are weighted before entering the hidden layer. The output of the hidden layer is based on the output of nonlinear transfer function (sigmoid function), which is applied to modify the sum of the entering weighted values. Thus, multi-variable problems can be managed by ANN as they may describe highly nonlinear connections, such as ozone production, that were overlooked by a conventional statistical technique (e.g., regression model [21]).

4. Conclusions

In this study, three urban areas located in peninsular Malaysia were selected. Hourly air pollutants and a weather parameters dataset from 2016 to 2020 were used to analyze the spatiotemporal variation of O$_3$ levels and its relation to the weather parameters. Seremban recorded the highest annual concentration of O$_3$ levels in 2017, whereas Kuala Terengganu experienced a decrement in annual O$_3$ levels starting from 2016. Slightly lower concentrations were observed in 2020 due to less emission of VOCs and NO$_x$ as a result of the total lockdown in Malaysia caused by COVID-19. The diurnal plot O$_3$ concentration with its precursors (NO$_x$, NO, NO$_2$, and CO) showed similar trends at the three study areas. The peak of O$_3$ precursors was observed during rush traffic hours in the morning and evening and O$_3$ formation was observed to reach its peak hour during the maximum solar intensity of the day (2:00 pm to 5:00 pm) and started to decrease steeply after 6:00 pm. The concentration of ground-level ozone varies according to the population and the location of the monitoring station. VOCs react with ozone precursors during photochemical reaction and increase concentrations of ozone while decreasing concentrations of precursors. The association of ozone and its precursors including weather parameters were tested using non-parametric correlation. Spearman correlation was used to analyze the relation of ozone with one independent variable. The result indicates that there was a strong correlation between ozone with temperature and NO, whereas a moderate correlation between NO$_x$ and CO was observed. A similar trend of association of O$_3$ levels with its precursors and weather parameters was observed during nighttime except with a lower calculated $p$ value. A very strong to strong correlation of NO$_x$ with NO$_2$ and NO during daytime and nighttime was observed in Seremban due to its inter-convertible reaction from NO to NO$_2$ and reactions of NO$_x$ with VOCs. Two predictive models, i.e., a linear model (multiple linear regression) and a nonlinear model (ANN, feed forward back propagation) were attempted to model the daytime and nighttime O$_3$ levels at all study areas. FFBP outperformed the prediction of O$_3$ levels during daytime and nighttime compared with MLR due to a better agreement between the predicted and observed measurement.

In this study, an association between daytime and nighttime O$_3$ levels can only be made with NO$_x$ concentration due to the missing measurement of non-methane hydrocarbon (NmHC) that represents VOC. Excluding VOC as one of the important O$_3$ precursors might cause the loss of important information/findings related to a place. However, this research successfully analyzed daytime and nighttime O$_3$ levels with its precursors at the selected areas. Variation of nighttime O$_3$ levels is usually neglected because O$_3$ cannot be formed photochemically due to the lack of sunlight at night, and NOx will continuously react with O$_3$, which is hypothetically the main period of O$_3$ depletion. Reduced O$_3$ depletion because of the decline in NO$_x$ emission would lead to the accumulation of O$_3$ at night, which would provide higher initial concentrations for photochemical reactions during the daytime, further leading to higher O$_3$ concentrations and exposure hazards at daytime. Therefore, the possible long-term effects on ecosystems caused by nighttime O$_3$ concentrations has important research implications. Prior to this, a long-term study of nighttime O$_3$ variation together with its precursors is required to determine the increment/decrement trend of
this pollutant. Local authorities can strategize mitigation plans accordingly, especially by controlling anthropogenic precursors such as NO\textsubscript{X} and VOC.

**Author Contributions:** Conceptualization, S.E.B. and N.M.N.; methodology, N.M.N.; software, Z.H.; validation and visualization, N.M.N., A.Z.U.-S., A.V.S., and P.V.; formal analysis, N.M.N.; investigation, G.D.; resources, M.R.R.M.A.Z.; data curation, A.A.K.; writing—original draft preparation, S.E.B.; writing—review and editing, N.M.N.; visualization, Z.H.; supervision, N.M.N.; project administration, M.R.R.M.A.Z.; funding acquisition, A.V.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research publication was supported by TUIASI from the University Scientific Research Fund (FCSU).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The author would like to thank to Department of Environment Malaysia for the air pollutant dataset.

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

**References**


2. Sharma, A.; Sharma, S.K.; Mandal, T.K. Ozone sensitivity factor NO\textsubscript{X} or NMHCs. A case study over an urban site in Delhi, India. *Urban Clim.* **2021**, *39*, 100980. [CrossRef]


7. Awang, N.R.; Ramli, N.A.; Yahaya, A.S.; Elbayoumi, M. High night time ground level ozone concentrations in Kemaman: NO\textsubscript{X} and NO\textsubscript{2} concentrations attributions. *Aerosol Air Qual. Res.* **2015**, *15*, 1357–1366. [CrossRef]


23. Pavlak, I.; Jaroslawski, J. Forecasting of surface ozone concentration by using artificial neural networks in rural and urban areas in central Poland. *Atmosphere* 2019, 10, 52. [CrossRef]