1. Introduction

The growth in urban traffic congestion has become a serious problem in all large metropolitan areas; it significantly affects the economy and travel behaviour. In addition, it is a cause of discomfort for millions of motorists [1, 2], and has harmful impacts on human health and metropolitan air quality [3]. For instance, in the US, congestion is responsible for wasting 101 billion USD annually. This figure accounts for an extra 2 billion gallons of fuel and 4.8 billion hours of wasted time [4]. Traffic congestion involves major contributing factors such as human, vehicle, and traffic composition [5]. Traffic composition is the distribution of vehicle type; this includes passenger car, mini bus and/or heavy vehicles.

Congestion causes chronic environmental problems such as air and noise pollution. Researchers have recently developed various models to estimate air pollutants resulting from transportation sectors. For instance, multivariate linear regression was implemented to link traffic density and the presence of diesel vehicles emitting air pollutants [6]. Real time information on traffic density, diesel vehicles, and traffic voice (noise) was used to model indoor and outdoor air pollutants [7]. In addition, Adaptive Neural Fuzzy Inference System (ANFIS) was implemented to predict the traffic flow over a short period using 104 changeable parameters [8]. ANFIS has been implemented to determine the level of safety on roads in relation with traffic density, speed and road plane visibility [9]. The results have shown ANFIS ability to enhance safety levels on roads; ANFIS was used to improve the work of Traffic Controllers in decreasing traffic congestion.

The performance of ANFIS controllers in terms of delay, average waiting time and queue length, was analysed and compared with the performances of traditional controllers and normal fuzzy controllers [10]. The results showed that ANFIS performance was superior. ANFIS has become popular for modelling environmental systems due to its accuracy, efficiency and capacity to handle large amount of stochastic (linear, nonlinear) data. For example, ANFIS was used to model driver behaviour using their reaction time and delay [11]. Furthermore, to help developing countries properly estimate solar radiation and benefit from it, ANFIS was implemented to model solar radiation based on metrological variables [12].

Despite its advantages, the complexity of ANFIS model topology, especially at a large number of input variables, is considered the main limitation of its wide implementation. This is because...
ANFIS generates and tries all possible combinations of premises, which are a function of the number of variables. For instance, if a system has (n) inputs and (P) premises, then the number of available rules equals \( N = P^n \). Thus, ANFIS implementation may not be feasible for several variable reasons. Furthermore, human expertise is necessary to optimize the ANFIS structure; however, this is solution is not always viable. For instance, if five inputs and three membership are used in a system, then the rule number becomes \( 3^5 = 243 \) rules, which significantly increases the total number of parameters and computing time [13]. However, a modified ANFIS (MANFIS) is recommended to overcome this limitation.

This study aims to propose a model to estimate CO\(_2\) and NO\(_x\) at the signalized roundabouts. It links the traffic conditions, including delays and percentages of heavy trucks, with the vehicle emissions to estimate air pollution. It applies a systematic search algorithm to choose the best available representative input variables. After that, it modifies the model to minimize the modelling complexity and error as well as providing effective tools to simulate such environmental applications. Such approach could assist decision makers in properly establishing sustainable traffic plans to reduce the impact of traffic congestion on air quality.

2. Methodology

2.1. Study Area and Data Collection

Corum city is the centre of the Corum governorate that is located at the north of Turkey. To analyse the traffic situation in the city, the main intersections were categorized into signalized and non-signalized intersections. This study analysed signalized roundabouts resulting in a total of eight intersections during the summer of 2015. Cameras were also placed on each intersection (numbers according to the size of the intersection) with the aim of defining traffic volume at rush hours. Rush hour periods were divided into three periods, these included mornings, (7:30-9:30) to cover business time in Turkey that starts at 9:00 AM. The second period was between 12:00-14:00 (end of schools), and the third period was between 16:30-18:30 (end of business day). Thus, for each of these intersections six hours of traffic video were recorded.

The recorded data for each intersection yielded vehicle counts and speeds of 6-types of vehicles: automobile, taxi, minibus, van, bus and heavy trucks. The analysis time of the recorded data was 54 h. Fig. 1 shows the 6th signalized roundabout as a satellite photo. It shows the location of the leg recorder. Table 1 shows a sample of traffic recorded data along the morning rush hours (7:30-9:30). Moreover, it represents an example of how data was extracted. For instance, 1-4 represent the total number of each vehicle type that is leaving leg one into leg number four. Furthermore, the total column at the table represents the total number of vehicles that are leaving from each leg at the roundabout to other legs. Air quality was also monitored during the same period, in a single site at each intersection that is located downwind [14], using Genesis portable air monitoring made by Thermo Fisher at around 15-30 m from the intersection. The measurement was taken every fifteen minutes then the hourly average was used to represent the concentration at peak hour. Finally, modelling and data analysis were performed using MATLAB (7.14).

![Abide signalized roundabout satellite photo (6th roundabout)](image.png)

2.2. Selection of Optimum Inputs and MANFIS Model Development

In order to minimize the data input and simplify the model structure, input selection was implemented. Initially, five inputs were used against each output; these were NO\(_x\), and CO\(_2\), respectively. The inputs represent vehicle average speed, traffic delay per second, percentage of mini trucks (%MT) (2.5 PCU), heavy truck percentage (%HT), and the total of heavy and mini trucks (%H&M). The optimum pair combination (input-output pair) was then determined by searching for the combination with lowest Root Mean Square Error (RMSE). The selection of optimum input number was based on a single iteration using the general bell shape fuzzy function and two membership functions by using a hybrid of least-squares and back-propagation gradient descent methods. In addition, the model performance was evaluated using the following statistical equation [15, 16].

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \hat{x}_i)^2}{n}} \tag{1}
\]

Where \( X_i \) is the actual output and \( \hat{X}_i \) is the predicted output, \( n \) is the number of the outputs.

ANFIS is a multilayer feed-forward network; it performs a fuzzy logic function on incoming signals. To build the fuzzy logic structure, it is essential to (i) select the model inputs (ii) determine the membership functions, and (iii) generate the fuzzy rules. Meanwhile, minimizing the model error needs an optimizing epoch’s number, membership type and number. During the training phase, the shape of the membership function was modified in order to define the relation between input and output. This stage was repeated on several occasions (epochs) until the desired convergence was acquired (usually until the specified minimum square error between the ANFIS output and the actual one is achieved). However, for a first-order Sugeno fuzzy model, a common set of two fuzzy rules and a set of if-then rules are described as follows:

**Rule 1:** \( If \ x \ is \ A_1 \ and \ y \ is \ B_1 \ \text{THEN} \ f_1 = p_1 x + q_1 y + r_1 \tag{2} \)

**Rule 2:** \( If \ x \ is \ A_2 \ and \ y \ is \ B_2 \ \text{THEN} \ f_2 = p_2 x + q_2 y + r_2 \tag{3} \)
where $A_i$ or $B_j$ is a linguistic label (grade), such as “low” or “less”, and $\mu_1$, $\mu_2$, $\mu_3$, $\mu_4$ are the design parameters that are determined by the system developer [17]. Fig. 2 presents the ANFIS model structures, where the circular nodes are fixed and the square nodes have parameters to be learned. The shown five layers are characterized by training and testing phases. The model developer has the capability to choose among the available membership function types in accordance with system demand, simplicity, speed and convenience. The membership function is a parameterized function in which any changes in the corresponding parameters produce a change in the function shape. However, the selection of membership function should fall between 0 and 1 [18].

In this study, to develop MANFIS structure, several steps were implemented. First, optimizing the number of inputs and then determining the best input-output combinations by searching for the lowest error RMSE for training and checking. After that, altering the type of membership functions to determine the best one. Finally, determining the number of membership functions that reduce the RMSE [19, 20].

Fig. 2. ANFIS architecture Sugeno system.

### 3. Results and Discussion

#### 3.1. Inputs Selection for CO$_2$ Modelling

The collected input variables were divided into training (the odd readings) and testing (the even readings). The available variables that may affect vehicles (gasoline and diesel) emissions were selected. It was imperative to choose the factors (inputs) that are relevant to the simulated system. The summary of the optimum twelve inputs and their combinations are shown in Fig. 3. To obtain these results, twenty-five combinations were tested. These combinations were single input-output, double input-output and tribal inputs-output. For levels of single input, the total heavy and mini trucks, and car speed, represent the best results with training RMSE are equal to 494 and 731 and checking RMSE are equal to 1,357 and 1,214, respectively.

For double inputs, the best combination is between the percentage of heavy trucks and average vehicle speed, its training RMSE equalled 0.005 and checking RMSE equalled 1,300. However other double inputs combination has relatively similar results, which combination is between the percentage of heavy trucks and the percentage of mini trucks which has RMSE for training equals to 303 and RMSE for checking equals to 861. For three inputs and single output, the best combination is between the average car’s speed, percentage of heavy trucks and percentage of mini trucks as RMSE$_{training} = 0.002$ and RMSE$_{checking} = 973$. However, using three variables as inputs for the model did not significantly enhance performance. Thus, the optimum number of inputs is assumed to be two, and the best combination is between the heavy truck percentage and vehicle speed.

#### 3.2. Inputs Selection for NO$_x$ Modelling

Fig. 4 shows a summary of the best input combinations for modeling NO$_x$ by testing all twenty-five possible combinations for single, double and tribal inputs. For a single input and one output that is NO$_x$, the percentage of heavy truck represents the best input with RMSE$_{training} = 59.2$ and RMSE$_{checking} = 489.4$, respectively. While for the combination of two inputs, the optimum one is between the delay and percentage of heavy trucks, its results
are RMSE\text{training} = 59.02 and RMSE\text{testing} = 167.6, respectively. However, for NOx modelling the combination of two input variables is considered as the optimum combination, since increasing the input number does not enhance the process significantly as shown in Fig. 4.

### 3.3. Final Structure of MANFIS Model

The performance of two input-output combinations for CO2 and NOx modelling is shown in Table 2. For carbon dioxide, the best representing pair combination is number five that is between the vehicle speed and the percentage of heavy trucks with RMSE\text{training} equals 0.06 and RMSE\text{testing} equals 1,300.7. However, the high testing error may be due to limited training readings which could have been avoided by increasing the number of readings. In addition, the first combination generates a smaller error for the training phase, but it is not considered due to its high testing phase error. On the other hand, for NOx modelling, the traffic delay and the percentage of heavy trucks are considered as the optimum input pair. It produces 0.02 and 253.0 for RMSE\text{training} and RMSE\text{testing}, respectively.

After selecting the best input combination for both CO2 and NOx, the best membership functions were determined. This was done by choosing from eight types of membership functions as shown in Table 3. In this phase, the hybrid training algorithm has been used. Furthermore, three membership functions and three epochs were implemented during the search process. The results show that the triangular membership function best represents CO2 emission with RMSE\text{training} = 0.05 and RMSE\text{testing} = 1,034.2. While for NOx modelling the best membership function is two-sided Gaussian with 0.007 and 241.7 for RMSE\text{training} and RMSE\text{testing}, respectively.

To accomplish the MANFIS structure, the optimum number of membership functions and the best combination for CO2 modelling, b best combination for NOx modelling

| No | CO2 Inputs | RMSE\text{training} | RMSE\text{testing} | NOx Inputs | RMSE\text{training} | RMSE\text{testing} |
|----|------------|----------------------|--------------------|------------|----------------------|--------------------|
| 1  | Delay; Speed | 0.03                 | 3,594.8            | 1          | Delay; Speed         | 0.02               | 699.4             |
| 2  | Delay; % HT  | 1.48                 | 5,889.0            | 2b         | Delay; % HT          | 0.02               | 253.0             |
| 3  | Delay; % MT  | 0.22                 | 5,918.4            | 3          | Delay; % MT          | 0.44               | 1,151.5           |
| 4  | Delay; (H&M) | 2.30                 | 10,964.4           | 4          | Delay; (H&M)         | 0.50               | 2,133.3           |
| 5a | Speed; % HT | 0.06                 | 1,300.7            | 5          | Speed; % HT          | 0.29               | 1,145.8           |
| 6  | Speed; % MT  | 0.48                 | 8,932.7            | 6          | Speed; % MT          | 0.09               | 1,738.0           |
| 7  | Speed; % (H&M)| 0.19               | 1,944.0            | 7          | Speed; % (H&M)       | 0.04               | 378.2             |
| 8  | % HT; % MT   | 304.3                | 861.6              | 8          | % HT; % MT           | 59.05              | 167.6             |
| 9  | % HT; (H&M) | 303.3                | 1,784.4            | 9          | % HT; (H&M)          | 59.0               | 347.2             |
| 10 | % MT; (H&M) | 303.4                | 572.0              | 10         | % MT; (H&M)          | 59.1               | 111.3             |

*best combination for CO2 modelling, b best combination for NOx modelling

| Code  | Function Description                  | CO2 RMSE\text{train} | CO2 RMSE\text{test} | NOx RMSE\text{train} | NOx RMSE\text{test} |
|-------|--------------------------------------|-----------------------|----------------------|-----------------------|----------------------|
| Trimf | Triangular MF                        | 0.05\(^a\)            | 1,034.2\(^a\)       | 0.013                 | 315.1                |
| trapmf| Trapezoidal MF                       | 0.03                  | 1,449.8              | 0.004                 | 310.4                |
| gbellmf| Generalized bell curve MF             | 0.05                  | 1,139.2              | 0.18                  | 339.7                |
| gaussmf| Gaussian curve MF                     | 0.2                   | 1,106.9              | 0.007                 | 279.0                |
| gauss2mf| Two-sided Gaussian MF                | 0.04                  | 1,393.4              | 0.007\(^b\)          | 241.7\(^b\)         |
| pimf  | Pi-shaped curve MF                   | 0.03                  | 1,392.0              | 0.004                 | 310.5                |
| dsigmf| Composed of the difference between two sigmoidal MF | 0.04 | 1,321.6 | 0.005 | 309.4 |
| psigmf| Product of two sigmoid MF             | 0.36                  | 1,325.6              | 0.005                 | 309.4                |

*The best membership function performance for CO2, b The best performance of the membership function for NOx
Table 4. Performance Summary of Various Membership Functions for CO$_2$ and NO$_x$

| No of function | CO$_2$ $\text{RMSE}_{\text{train}}$ | CO$_2$ $\text{RMSE}_{\text{test}}$ | NO$_x$ $\text{RMSE}_{\text{train}}$ | NO$_x$ $\text{RMSE}_{\text{test}}$ |
|----------------|---------------------------------|---------------------------------|----------------------------------|---------------------------------|
| 2              | 355.1                           | 967.8                           | 67.1                             | 322.7                           |
| 3$^a$          | 0.05                            | 1,034.2                         | 0.007                            | 241.7                           |
| 4              | 0.04                            | 1,460                           | 0.006                            | 403.1                           |
| 5              | 0.05                            | 1,365                           | 0.07                             | 416.7                           |
| 6              | 0.04                            | 1,731                           | 0.08                             | 421.9                           |
| 7              | 2,068                           | 2,578                           | 0.07                             | 423.1                           |

$^a$The best function number

of membership function was determined by keeping the epoch number constant (3 epochs) and altering the number of membership functions from 2 to 7. However, the selected optimum input combinations for both CO$_2$ and NO$_x$ were used in this search. The optimum number of membership function was selected based on the generated smallest RMSE for training and testing phases. Table 4 shows the performance of the model with different function numbers. The best performance for both CO$_2$ and NO$_x$ was achieved with three functions. The training RMSE of CO$_2$ and NO$_x$ were 0.05 and 0.007, respectively, while for the testing phase it was 1,034.2 and 241.7 for CO$_2$ and NO$_x$, respectively. However, as shown in Table 4 increasing the numbers of membership functions does not enhance the model performance.

Developing the model structure by applying MANFIS enhances the overall modelling performance. For instance, it reduces the training RMSE for CO$_2$ by 16% and for NO$_x$ by 65%, respectively, while it reduces the testing RMSE for CO$_2$ by 71% and for NO$_x$ by 4%, respectively. Finally, the final MANFIS structure shown in Fig. 5 illustrates structures of MANFIS model for both CO$_2$ and NO$_x$, respectively. It shows the two inputs, membership functions, the three membership functions, the nine fuzzy rules and the desired output.

Fig. 6 shows the effect of the vehicles’ speed and percentage of heavy truck on CO$_2$ emission. For instance, for a speed value of 10 km/h, as the percentage of heavy trucks decreases, CO$_2$ emission increases. This is related to the increase the number of lighter, or gasoline trucks which have more CO$_2$ emission than the heavy truck [6]. On the other hand, increasing the speed will reduce CO$_2$ emissions. This goes back to different densities of diesel and gasoline. Thus, the consumption varies, and diesel consumption is less than gasoline.
Fig. 7 shows the effects of delay and the heavy truck percentage on NOx emission. The NOx emission can be divided into two categories. The first being 20 s per vehicle; in this region, increasing the percentage of trucks will decreased the NOx emissions. Otherwise, the second region is for delays of more than 20 s: in this region, increasing the percentage of heavy truck increased the NOx emission. Moreover, the highest NOx emission was at the largest delay (55 s) and highest heavy truck percentage. Increasing the heavy truck percentage reduced the traffic movement and increased the waiting time for all vehicles at the road; therefore, emissions increased. In addition, heavy trucks, usually diesel vehicles, comparatively emit higher NOx than gasoline vehicles during times of being stationary. Therefore, its contribution is tangible for NOx emission [21].

4. Conclusions

This study analyses negative implications of traffic congestion on air quality, especially under signalized roundabouts and possibly elsewhere. The developed model could assist municipal planning boards, traffic and environmental engineers to identify planning and management measures and policies for reducing air pollution as a result of traffic congestion in urban zones. The main objective of this research was to investigate the effect of traffic composition variables on CO2 and NOx density on signalized roundabouts, whilst maintaining an accepted degree of accuracy using MANFIS in order to reduce complexity and data collection time. The traffic composition variables included in this study were the percentage of minibuses, percentage of heavy trucks, average delay, and average speed. Proper input selection enhanced model performance. However, environmental systems have limitations to data records. These limitations are essentially related to accuracy, budget, time, and reliability of the data. Therefore, modelling environmental systems saves time, effort and cost whilst maximizing model usability.

The results indicate that vehicle speed and the percentage of heavy trucks are the main input variables to estimate the emission of CO2. Meanwhile, delays and percentage of heavy trucks are the main inputs for NOx modelling. However, there is some suggestion of potential conflicts with some aspects of current planning ideas; specific questions about the advisability of heavy truck entry permission to the city centre, especially during peak periods have arisen. These suggest that environmental and safety consequences of these concepts in specific places should be dealt with. Finally, it is recommended to replicate the investigation of the relation between traffic conditions and air pollutants elsewhere to see if similar findings are obtained to support the confidence in our conclusions.

Acknowledgments

The authors are grateful and thankful for the head of Corum City Council and all the people in the City Council for their help and support during data collection and analysis. Moreover, high appreciation is shown for Philadelphia University and Aqaba University of Technology for funding the project.

References

1. Jin J, Rafferty P. Does congestion negatively affect income growth and employment growth? Empirical evidence from US metropolitan regions. Transport Policy 2017;55:1-8.
2. Suleiman GM, Bezgin NO, Ergun M, Gürsoy M, Karagöz M. Effects of speed management and roadway parameters on traffic flow along arterials. In: Proceedings of the Institution of Civil Engineers-Transport. Thomas Telford Ltd.; 2017.
3. Zhou H, Li Y, Liu H, et al. Temporal distribution, influencing factors and pollution sources of urban ambient air quality in Nanchong, China. Environ. Eng. Res. 2015;20:260-267.
4. Studer L, Ketabdar M, Marchioni G. Analysis of adaptive traffic control systems design of a decision support system for better choices. J. Civil Environ. Eng. 2015;5:195.
5. Solomon S. Segmental assessment of level of traffic congestion on Kality Ring Road to Dukem Bridge [dissertation]. Addis Ababa: Univ. of Addis Ababa; 2015.
6. Sullivan JL, Baker RE, Boyer BA, et al. CO2 emission benefit of diesel (versus gasoline) powered vehicles. Environ. Sci. Technol. 2004;38:3217-3223.
7. Weichenthal S, Ryswyk KV, Kuika R, Sun L, Wallace L, Joseph L. In-vehicle exposures to particulate air pollution in Canadian metropolitan areas: The urban transportation exposure study. Environ. Sci. Technol. 2014;49:597-605.
8. Chen BP, Ma ZQ. Short-term traffic flow prediction based on ANFIS. In: International Conference on Communication Software and Networks; 27-28 February 2009; Macau, China: IEEE.
9. Zengqiang M, Cunzhi P, Yongqiang W. Road safety evaluation from traffic information based on ANFIS. In: Control Conference, 2008. CCC 2008. 27th Chinese. 2008. IEEE.
10. Soh AC, Rahman RZA, Rhung LG, Sarkhan HM. Traffic signal control based on adaptive neural-fuzzy inference system applied to intersection. In: 2011 IEEE Conference on Open Systems (ICOS); 25-28 September 2011; Langkawi, Malaysia: IEEE.
11. Khodayari A, Ghaffari A, Kazemi R, Manavizadeh N. ANFIS based modeling and prediction car following behavior in real traffic flow based on instantaneous reaction delay. In: 2010 13th International IEEE Conference on Intelligent Transportation Systems (ITSC); 19-22 September 2010; Funchal, Portugal: IEEE.
12. Piri J, Kisi O. Modelling solar radiation reached to the Earth using ANFIS, NN-ARX, and empirical models (Case studies: Zahedan and Bojnurd stations). J. Atmos. Sol-Terr. Phy. 2015;123:39-47.
13. Younes MK, Nopiah ZM, Basri NE, Basri H, Abusshammala MF, Maulud KNA. Solid waste forecasting using modified ANFIS modelling. J. Air Waste Manage. Assoc. 2015;65:1229-1238.
14. Karner AA, Eisinger DS, Nienmeier DA. Near-roadway air quality: Synthesizing the findings from real-world data. Environ. Sci. Technol. 2010;44:5334-5344.
15. Willmott CJ, Matsuura K. Advantages of the mean absolute
error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Res.* 2005;30:79.

16. Antanasijević DZ, Pocajt VV, Povrenović DS, Ristić MD, Perić-Grujić AA. PM$_{10}$ emission forecasting using artificial neural networks and genetic algorithm input variable optimization. *Sci. Total Environ.* 2013;443:511-519.

17. Younes MK, Nopiah ZM, Ahmad Basri NE, et al. Landfill area estimation based on integrated waste disposal options and solid waste forecasting using modified ANFIS model. *Waste Manage.* 2016;55:3-11.

18. Pramanik N, Panda RK. Application of neural network and adaptive neuro-fuzzy inference systems for river flow prediction. *Hydrol. Sci. J.* 2009;54:247-260.

19. Khatibinia M, Salajegheh J, Fadaee MJ, Salajegheh E. Prediction of failure probability for soil-structure interaction system using modified ANFIS by hybrid of FCM-FPSO. *Asian J. Civil Eng.* 2012;13:1-27.

20. Lin KP, Pai PF, Lu YM, Chang PT. Revenue forecasting using a least-squares support vector regression model in a fuzzy environment. *Inform. Sci.* 2013;220:196-209.

21. Shancita I, Masjuki HH, Kalam MA, Rizwanul Fattah IM, Rashed MM, Rashedul HK. A review on idling reduction strategies to improve fuel economy and reduce exhaust emissions of transport vehicles. *Energ. Convers. Manage.* 2014;88:794-807.