Journal of Al-Azhar University Engineering Sector

Vol. 19, No. 71, April 2024, 475-494

SEASONAL EGYPTIAN ROAD TRAFFIC VOLUME VARIATIONS USING MACHINE LEARNING

Ibrahim M. Ramadan[1], Taha E. Abdelkader[2,*], Sherif A. Kamal[2]

[1] *Department of Civil Engineering, Faculty of Engineering, Shoubra–Benha University, Cairo, Egypt.*

[2] *Department of Building and Construction, Faculty of Engineering, October 6 University, Giza, Egypt.*

***Correspondence:** Taha.Ezzat.eng@o6u.edu.eg

Citation:

I. M. Ramadan, T. E. Abdelkader, and S. A. Kamal, " Seasonal Egyptian Road Traffic Volume Variations Using Machine Learning", Journal of Al-Azhar University Engineering Sector, vol. 19, pp 475 - 494, 2024.

Received: 4 January 2024

Revised: 13 February 2024

Accepted: 20 March 2024

DOI:10.21608/auej.2024.265134.1599

ABSTRACT

Seasonal factors Weather, holidays, and school schedules all contribute to traffic. Seasonal factors have a significant influence on annual average daily traffic (AADT). Traffic volume is measured and predicted using AADT. To make intelligent transportation planning and finance decisions, seasonal factors must be included while examining AADT data since there has been no Egyptian traffic flow research. Machine learning approaches include regression analysis and artificial neural networks (ANNs) to predict seasonal variations. Since 2018, the General Authority for Roads, Bridges, and Land Transport (GARBLT) has not kept traffic records. Seasonal factors were employed to generate a more accurate and understandable AADT figure from 11 stationary monitoring stations monitored from 2013 to 2018. These stations collected socioeconomic data as well as information about the roads, such as lanes and station locations. The artificial neural network (ANN) model seasonal factors were far more precise and reliable when compared to the real values.

KEYWORDS: Traffic Growth Rate, Seasonal Traffic, Monthly Average Daily Traffic, Egyptian Transportation Planning.

التغيرات الموسمية في حجم حركة المرور على الطرق المصرية باستخدام التعلم الآلي

إبراهيم محمد رمضان[1]، طه عزت عبدالقادر[2،*]، شريف أحمد كمال[2]

[1] *قسم الهندسة المدنية، كلية الهندسة بشبرا، جامعة بنها، القاهرة، مصر.*

[2] *قسم التشييد و البناء، كلية الهندسة، جامعة 6 أكتوبر، الجيزة، مصر.*

***البريد الاليكتروني للباحث الرئيسي :** Taha.Ezzat.eng@o6u.edu.eg

الملخص

يؤثر الطقس والإجازات والجداول المدرسية على العوامل الموسمية (SF) وبالتالي تتأثر قيم متوسط حركة المرور اليومية السنوية AADT وحركة المرور المستقبلية، وعليه يجب أن تتضمن بيانات AADT الخصائص الموسمية لفهم تدفق حركة المرور والتخطيط الجيد للنقل. تحسن هذه الدراسة تقدير العوامل الموسمية (SF) في مصر حيث لم يكن لدى الهيئة العامة للطرق والكباري GARBLT بيانات لحصر حركة المرور منذ عام 2018 والذي تسبب في العديد من المشكلات المرورية. تم استخدام البيانات الخاصة بمحطات الرصد المستمرة لعدد 11 محطة رصد ثابتة (التابعة للهيئة العامة للطرق والكباري) في الفترة من عام 2013 الي 2018. بالإضافة إلى البيانات الاجتماعية والاقتصادية القائمة على خصائص الطريق مثل عدد ممرات الطرق

وموقع المحطة، تم حساب قيمة AADT بشكل أكثر دقة وسهولة بعد الحصول على العوامل الموسمية. تنبأت الشبكات العصبية بالعوامل الموسمية بشكل أفضل من طريقة الانحدار. باستخدام قيم SF، يمكن حساب حجم حركة المرور للطرق المتعددة من فترة مراقبة قصيرة، ومن خلال الدراسة يمكن التخطيط الجيد لوسائل النقل المستقبلية للتعامل مع أكبر قضايا النقل في مصر.

الكلمات المفتاحية : معدل نمو حركة المرور ، حركة المرور الموسمية ، متوسط حركة المرور اليومية الشهرية ، تخطيط النقل المصري.

1. INTRODUCTION

Traffic data is gathered to evaluate the utilization and efficiency of the road network. The data gathered is generally utilized in the fields of highway planning, geometrically optimized roadway construction, traffic management, and control, land-use planning, noise computation, analysis of road accidents, and research reasons. Existing traffic monitoring systems gather data on the number of vehicles present [1, 2].

The annual average daily traffic (AADT) is a crucial metric for measuring traffic. It denotes the typical volume of vehicles that pass through a particular segment of road stretch over one year [2]. The variation in the monthly average daily traffic (MADT) over a year is what determines the traffic pattern. By utilizing this approximation, the traffic pattern is summarized as the variation in average daily traffic during each month, referred to (MADT). When the average daily traffic for a month (MADT) is multiplied by a suitable coefficient unique to that month, one can calculate the annual average daily traffic (AADT). The coefficient is the ratio of the MADT to the (AADT). It is commonly referred to as the "seasonal factor" (SF) because it indicates the time-dependent variations in traffic on the route. An approximation of AADT that considers pertinent seasonal elements may be obtained by multiplying the daily traffic count by the ratio of the MADT to the AADT for that specific month. To account for the duration of traffic counts, twelve distinct ratios of MADT to AADT should be computed for each of the twelve months [3].

It is not possible to measure the traffic volume daily for every road in the country. As a result, there are two kinds of traffic counts: permanent and temporary. Every six years, across the country road segments should conduct portable temporary traffic counts, according to the Traffic Monitoring Guide (2001) [4]. Continuous traffic counts are conducted year-round utilizing automatic traffic recorders (ATRs) strategically positioned at specific places. The counts are useful for calculating seasonal adjustment factors that can be used to transform short-term traffic counts into estimates of AADT at any specific location. Automatic traffic recorders (ATRs) aren't needed to measure traffic volume throughout the entire year when temporary traffic counts and "adjustment factors" are used [1].

Multiple significant factors contribute to the fluctuating levels of traffic in Egypt's urban and rural throughout the year. The relevant variables considered include roadway characteristics such as station count location, lane number, demographic factors, and socioeconomic factors. If these characteristics are understood and quantified, they can be effectively employed [5]. By utilizing diverse mathematical methodologies, like genetic algorithms, artificial neural networks, cluster analysis, multiple linear regression, or hybrid models that combine fuzzy sets theory with neural networks [6 - 8].

Inaccurate evaluation in road design and management can result in significant and wide-ranging repercussions. According to Spławińska (2015), these defects result in inadequate surface design. Moreover, the implementation of an incorrect number of lanes at entrance and exit locations might cause road segments to be overly wide or too small. Geometrically risky junctions and crossings reduce safety, while inadequate traffic control increases congestion [2]. Moreover, deficient communication system design and insufficient environmental protections against noise and air pollution can detrimentally impact both the environment and road users. Ultimately, inaccurate outcomes in evaluations of traffic safety, economic evaluations, and predictions of road accidents have a significant influence on the process of making decisions. The presence of these inaccuracies highlights the crucial significance of precise estimating and forecasting in the design and operation of road infrastructure.

Typically, a consistent annual traffic growth rate is employed; however, the specific value of this rate varies monthly. This is due to the varying significant factors that can impact the seasonal factors (SFs) of different months in various methods. Therefore, it is necessary to measure the impact of each component on the SFs for various months separately. Hence, this paper utilizes realistic traffic counts to determine the value of seasonal variation. These values will be utilized subsequently to compute MADT and AADT at a later phase.

The study can be categorically separated into three distinct sections. To begin, the seasonal features of Egyptian roadways are assessed based on data collected from 11 different locations throughout Egypt. Furthermore, the growth rates of seasonal components are estimated through a detailed analysis of the available data using linear regression and artificial neural networks (ANN). Additionally, the accuracy of the data obtained from the second technique is also examined.

The American Association of State Highway and Transportation Officials (AASHTO) guideline recommends the utilization of seasonal factors for a more accurate forecast of AADT; however, unfortunately, these parameters are not available on Egyptian roads.

There are various approaches utilized to ascertain the SF value, with cluster analysis being the most commonly utilized method [9, 10]. According to several studies [9, 11, 12], it has been found that regression analysis and neural networks outperform cluster analysis [3, 13]. Nevertheless, there has been a lack of research conducted to investigate the variation of traffic volume on Egyptian roadways during different seasons. This study aims to examine how the flow of traffic on Egyptian highways has changed over time to forecast seasonal factors for Egyptian highways.

2. Regression Analysis

Cluster analysis's poor output prompted the development of a more advanced procedure centered on regression methods. The correlation between physical and functional characteristics and seasonal changes in traffic volume was examined using the regression approach [3]. This is one way to depict the generic regression model:

$$F_{sm} = A_0 + A_1 X_1 + A_2 X_2 \tag{1}$$

Where F_{sm} = seasonal factor in month m; and X_1, X_2, etc. are economic indicators; A_0 = regression constant; and A_1, A_2, etc. are regression coefficients.

The factors that impact the estimations of AADT in Broward County, Florida, were investigated in a series of multiple linear regression experiments by Zhao and Chung (2001). Using GIS technology, land use and accessibility metrics were built. Following the removal of outliers, four models were calibrated. The most significant predictors for calculating annual average daily traffic (AADTs) are functional categorization and lane number, which were recognized as two independent factors [14].

Most studies commonly utilize socioeconomic elements [15, 16]. Employment and per capita income are factors that influence outcomes. Population age, gender distribution, and vehicle ownership are other criteria commonly utilized in practical research. Economic issues have the potential to impact all other factors. Therefore, it is necessary to take economic fluctuations into account while making predictions about traffic growth. The continuous economic expansion in Europe has resulted in an increase in the gross domestic product. The cost of private transportation is also affordable [17]. Research has discovered that this particular combination leads to an increase in traffic expansion [18]. However, other researchers discovered that fuel prices had a minimal effect [19]. Other studies have developed a distinct model and determined that population, PCI (per capita income), and NDP (net domestic product) exert the most influence [20].

There are several advantages to their approach. These include the ease of grouping data, the high level of aggregation, and the simplicity of application and use. Additionally, they discovered that the factoring approach was not as precise as linear regression when applied to daily short-count volumes. Linear regression analysis only requires a small minimum amount of data points, at least in theory. The drawbacks include Not considering how different highway features can change the AADT; normal linear regression methods are usually used to guess future-year AADT; enough high-quality data points are needed to make a strong and statistically significant connection; set a minimum duration for count verification. At least two or more site visits are required to gather the minimum amount of data. Regression algorithms necessitate an important amount of computer resources, resulting in increased complexity [17, 21].

3. Artificial Neural Networks

The goal of computer approaches known as artificial neural networks (ANNs) is to simulate brain activity as closely as possible. ANNs are better than traditional computing methods for pattern classification problems because of their unique characteristics [3]. These characteristics include the system's need to be able to deal with missing input data, process noisy input data, and learn to identify patterns in historical data.

For several road segments, Faghri et al. investigated many methods for determining seasonal factors and, consequently, the AADT. The results showed that the neural network technique performed better than the other two ways after comparing them all. The neural network showed promise for efficiently handling problems with traffic pattern recognition and road categorization with its superior results. The authors opted to test neural networks for this problem rather than cluster analysis, even though regression analysis produced superior results overall [3], as the regression equation is random and there are no natural groupings.

The main model for determining AADT is the multi-layered feed-forward network, which is frequently used. A total of 63 ATR locations on rural and regional roadways in Minnesota were examined in order to ascertain traffic volumes. To examine the data, they used a supervised learning technique that was multi-layered, feed-forward, and back-propagation. The data was gathered throughout the period from May to August 1993. A hidden layer, an output layer, and an input layer were the three neuronal layers that made up the model. Data transmission between the input and output layers was the hidden layer's main function. Since neural network methodology does away with the requirement to classify permanent count stations, this approach is the best choice for calculating AADT [22].

Other researchers utilized a neural network architecture consisting of a three-layer neural network design—an input layer, an output layer, and a hidden layer—in their research. This network's original intent was to classify 1991 traffic flow data collected from thirteen different locations across Hong Kong. Different time intervals throughout the day were used to experiment on counts of different lengths (4, 6, 8, 10, 12, 14, and 16 hours), with numerous matching start times for each. There is a high degree of precision in the findings from the thirteen counting sites when compared to the actual numbers [23].

Classifying seasonal factors (SF) has been the primary emphasis of research so far, with several approaches proposed for more efficient factoring groups. Data and processes are the main differentiators among these methodologies, which cover both more conventional approaches like regression analysis and statistical clustering and more advanced machine learning methods like neural networks and algorithm evolution.

An artificial neural network (ANN) was used in transportation planning. More ANN architectures can be used to predict daily trip flows. A comparison between feed-forward back-propagation (FFBP) and radial basis function (RBF) has been executed. Another comparison was made between FFBP and the generalized regression neural network (GRNN). Notably, RBF neural networks and GRNN were discovered to not generate negative predictions [24, 25], in contrast to FFBP networks. Another paper examined a hybrid prediction model that combined autoregressive integrated moving averages (ARIMA) with multilayer artificial neural networks (MLANN), specifically focusing on feedforward backpropagation (FFBP). unknown aspects are completely random, and there is a lot of complex historical traffic data. To get around this, we did this. Combining the two models can improve the prediction accuracy of each one when applied independently, according to experiments conducted on real data sets [26].

Various studies demonstrate the benefits of the algorithms' ability to accurately simulate complex patterns and forecast issues. The algorithm generation mechanisms are derived from cognitive brain functions. ANN is most effective when applied to tasks such as predicting, image processing, and character recognition. Considers ambiguity, acquires knowledge, and constructs representations of intricate and nonlinear connections. Drawbacks: The reality of this depends on the hardware, specifically central processing units with parallel processing power. The unpredictability of network behavior results in solutions that lack clarity on their causes and mechanisms, leading to reduced trust in the network. There is no established guideline for precisely determining the appropriate network structure of an ANN. ANNs are developed through repeated experimentation and practical knowledge. The transformation of information into numerical values before its use in the ANN has an impact on the network's effectiveness [21, 27, 28].

4. DATA COLLECTION

Establishing a dependable database of traffic volumes for road networks was a crucial component of the investigation. The study utilized data obtained from permanent traffic count sites spanning the years 2013 to 2018. Continuous Automatic Traffic Recorder (ATR) stations record the overall quantity of vehicles passing their designated location. The volume data was condensed into MADT. MADT for each average traveler (AT) was calculated by summing the total amount of traffic for each day in each month and then dividing it by the number of days in that month. This procedure is feasible in cases where there is a lack of data for each weekday within a particular month.

The General Authority for Roads, Bridges, and Land Transport (GARBLT) in Egypt is responsible for managing the intercity road system. They have been gathering statistics on traffic volume at 15 permanent stations. Permanent stations depend on double-loop detectors ("Japan International Cooperation Agency, JICA 2012") [29]. This paper employs a total of 11 permanent monitoring stations. However, four of these stations were not included in the analysis due to issues encountered throughout the monitoring procedure, such as the need for maintenance or the production of inaccurate statistical data. The eleven stations covered in this report are summarized in **Table 1**, and **Fig. 1** shows the study's locations.

Table 1. Stations' Characteristics

Station Number	Location	Road Name	Region
ST. 3	AYT – BNS	Giza - Beni Suef	Greater Cairo
ST. 4	CAI - TNL	Cairo - Suez	Greater Cairo
ST. 10	SHBRA - BNH	Cairo - Qalyubia	Greater Cairo
ST. 15	ALX – DMR	Alexandria - Beheira	Alexandria
ST. 2	DMR – TNT	Beheira – Gharbia	Delta
ST. 8	DFR – ZIT	Defra – Kafr Elziat	Delta
ST. 11	AGA – MTG	Mit Ghamr - Aga	Delta
ST. 5	AHM - ISM	Sharkia - Ismailia	Suez
ST. 7	MST - BLB	Qalyubia-Sharkia	Suez
ST. 13	AF. Toll	Giza - Faiyum	Upper Egypt
ST. 14	MNY - ASU	Minya – Asyut	Upper Egypt

Source: GARBLT

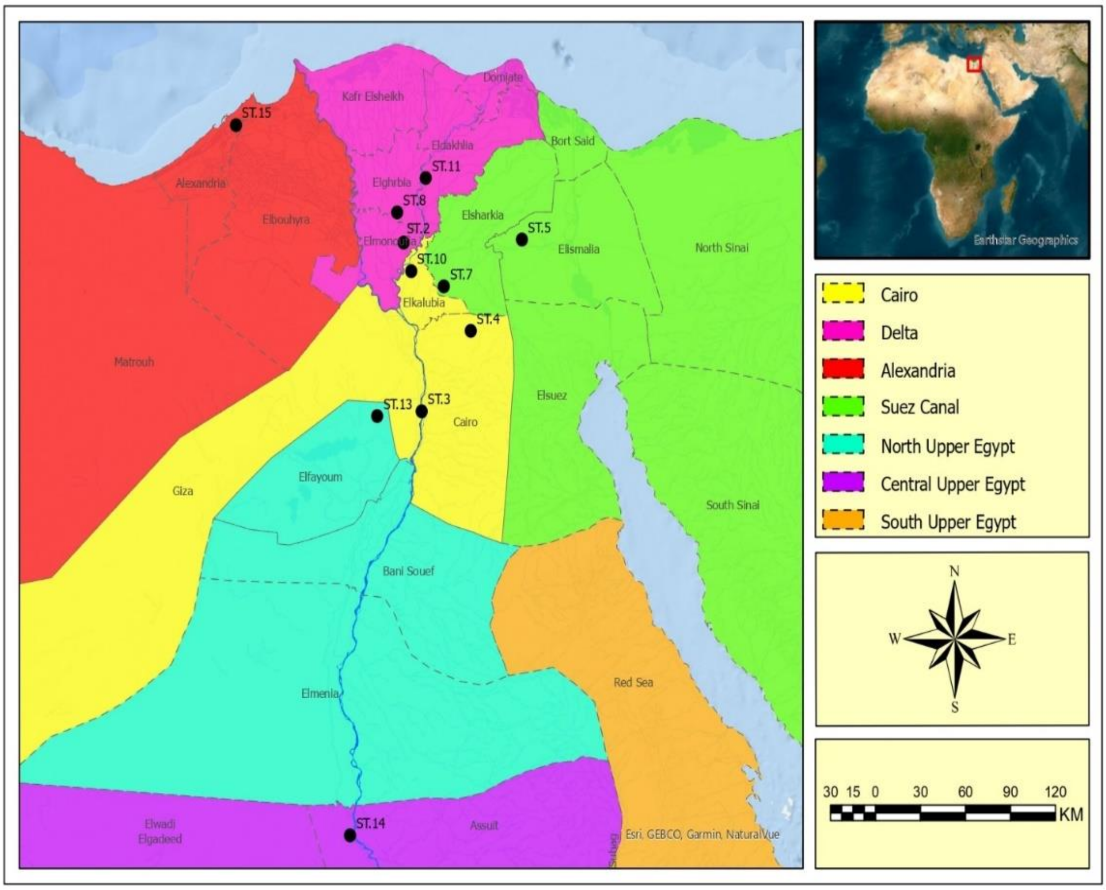


Fig. 1. Study's Locations.

The Central Agency for Public Mobilization and Statistics (CAPMAS) reports the population of different regions, which serves as a representation of socio-economic aspects and has a significant impact on traffic patterns. Furthermore, the Ministry of Planning and Economic Development (MPED) monitors the gross domestic product (GDP) of Egypt from 2013 to 2017, which has a direct influence on the development of infrastructure and transportation networks. The World Bank Group (WBG) assessed per capita income (PCI) over the same period, which reflects people's economic well-being and affects their decision-making regarding movement. Furthermore, the fuel sale data obtained from Trading Economics offers valuable insights into energy usage and its effects on transportation systems. A variety of the road's characteristics were employed, such as the number of lanes and vehicle counting locations, which are represented by the values for the station (X) in the horizontal direction and the elevation (Y) in the vertical direction.

5. METHODOLOGY AND DATA ANALYSIS

Regarding the assessment of seasonal factors SF for Egyptian traffic, there has been no scientific effort documented up until this point in time. Considering this, several different approaches have been utilized to forecast the seasonal factor SF of any given site based on the permanent traffic count PTC data. To conduct this analysis, monthly traffic data from each of the eleven PTCs was utilized. Average seasonal factors, linear regression, and artificial neural networks are the techniques that are utilized in the process of estimating the SF of traffic.

5.1. Estimation of Average Seasonal Factors of Traffic

To get a location's AADT from its short-period traffic count (SPTC), seasonal factors (SF) are required. Monthly SF for Site j in Year k results from dividing $MADT_j^{m,k}$ by $AADT_j^k$. Seasonal variations in traffic are shown in the visualization [30].

$$SF_j^{m,k} = \frac{MADT_j^{m,k}}{AADT_j^{k}} \tag{2}$$

The average seasonal factor approach suggests that the seasonal factors for Month M on any road can be approximated accurately by calculating the average seasonal factors for that month derived from all the PTCs. To calculate the average seasonal factor (ASFm) for Month m, use the following formula [30]:

$$ASF^m = \frac{\sum_{j=1}^{N}\sum_{k=1}^{K_j} SF_j^{m,k}}{\sum_{j=1}^{N} K_j} \tag{3}$$

Here, N is the total count of PTCs, which is eleven in this study. K_j denotes the number of years for which traffic data is accessible at the j-th PTC. **Table 2** displays the seasonal factors for total traffic that have been determined from this technique.

Table 2. Average Seasonal Factors.

ST. No	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
ST. 3	1.0801	0.8801	0.9194	1.0306	0.9045	1.1585	1.1493	1.0311	0.8810	1.0426	1.0695	0.8534
ST. 4	0.9992	0.8984	0.9561	1.0007	0.9918	1.0895	1.1109	0.9658	0.9587	1.0622	0.9823	0.9844
ST. 10	0.9399	0.9971	0.9236	1.0348	1.0091	1.0212	1.1154	1.0672	0.9959	1.1245	0.8326	0.9386
(Greater Cairo)	1.0064	0.9252	0.9330	1.0220	0.9685	1.0897	1.1252	1.0214	0.9452	1.0764	0.9615	0.9255
ST. 15 **(Alex)**	0.8618	0.9445	1.1620	1.1421	1.0700	1.0377	0.9229	1.1290	1.0198	0.8523	0.8714	0.9863
ST. 2	0.7686	0.8877	0.9656	0.8715	0.9000	1.3305	1.0076	0.9884	0.9180	1.1789	1.0246	1.1586
ST. 8	0.8533	0.9403	0.9347	0.9600	1.0443	1.2471	0.8992	0.9716	1.0871	1.1114	0.9722	0.9789
ST. 11	0.9931	0.9806	0.9996	0.9235	1.1732	1.1852	0.8510	1.0385	0.8515	1.0052	1.0437	0.9549
Delta	0.8717	0.9362	0.9666	0.9183	1.0391	1.2543	0.9192	0.9995	0.9522	1.0985	1.0135	1.0308
ST. 5	0.8138	0.9052	1.0033	1.0468	0.9425	1.2006	1.1890	1.2067	0.7756	0.9406	1.1120	0.8638
ST. 7	1.0353	1.0952	0.8306	0.9538	1.0514	0.8905	1.0885	1.0106	1.1485	1.0745	0.9899	0.8313
(Suez)	0.9245	1.0002	0.9169	1.0002	0.9969	1.0455	1.1387	1.1086	0.9620	1.0075	1.0509	0.8475
ST. 13	0.9883	0.9235	1.0246	1.0126	1.0185	1.1096	0.9240	0.9312	1.0056	1.0196	1.0200	1.0225
ST. 14	0.8404	1.0677	1.1027	1.0292	1.0364	0.9596	1.1151	1.1503	0.8557	0.9114	1.1755	0.7560
Upper Egypt	0.9144	0.9956	1.0637	1.0209	1.0274	1.0346	1.0196	1.0407	0.9306	0.9655	1.0978	0.8893
Whole Egypt	0.9249	0.9564	0.9838	1.0005	1.0129	1.1118	1.0339	1.0446	0.9543	1.0294	1.0085	0.9390

Therefore, it is suggested to employ the average seasonal factor approach for calculating the seasonal factors of various months at all locations. The recommendation is based on the fact that the average seasonal factor approach is easier in comparison to regression analysis methods [30]. Moreover, employing the average seasonal factor method enables the utilization of a singular seasonal factor for a specific month across all locations, hence introducing simplicity into the process. It is important to mention that this study solely considered traffic data from rural routes. Therefore, it excluded highways with distinct functional attributes, such as arterial roads or recreational roads. The average seasonal parameters adequately represent the seasonal fluctuations of Egyptian roadways.

5.2. Estimation of Seasonal Factors by Linear Regression

Python-based tools are the primary frameworks used for machine learning tasks [31]. The imbalanced-learn module in Python was used to implement each approach. The linear regression graphs were made using the sklearn library in Python. The data is partitioned into two distinct datasets: the training dataset, which accounts for 70% of the data, and the validation dataset, which comprises 15%, as well as the testing dataset, which considers 15%. The model is developed using the training dataset and the appropriate variables. Next, the model is evaluated using the test dataset. Not only does this aid in avoiding overfitting of the model, but it also contributes to ensuring the accuracy of the expected model. The training dataset is used for model development and subsequent predictions.

SF is predicted using multiple linear regression analyses in Python. Twelve monthly seasonal factor (MSF) measurements collected from continuous traffic count locations between 2013 and 2018 constitute the dependent variables. Transportation characteristics (such as location and number of lanes) and socioeconomic factors (such as population POP, per capita income PCI, gross domestic product GDP, and fuel price) are among the independent variables included in regression analysis, which are thought to have a reasonable causal association with MSF.

5.3. Estimation of Seasonal Factors by Artificial Neural Networks

The artificial neural network (ANN) replicates the structure and functioning of a biological neural network in the brain. The learning process can acquire knowledge without the need for specific regulations or instructions. As shown in **Fig. 2**, a neural network consists of neurons, or nodes, and connections, or edges. A numerical number is associated with each neuron, and a weight is provided for each connection. Taking into consideration the weighted values of the previous layer, the calculation takes place by sending data from the input layer to the hidden layer and then, finally, to the output layer. Many non-linear adjustments performed on data acquired from different layers are added together to generate the output [32].

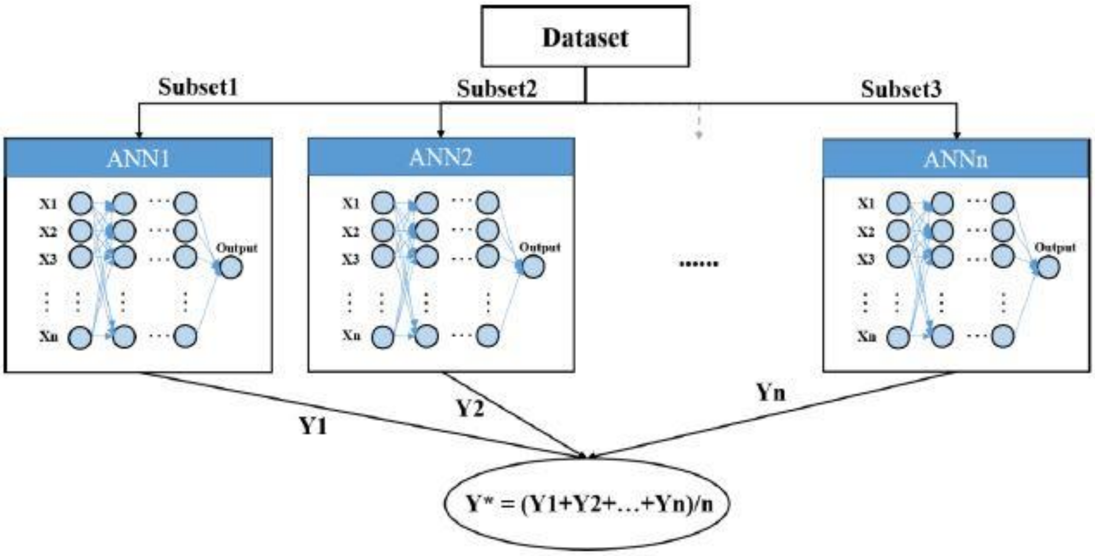


Fig. 2. Structure of Ensembling Artificial Neural Networks [32].

Layered neural networks using backpropagation (BP) learning algorithms that modify weights based on output least mean square error are popular. Use the BP architecture to forecast with neural networks. This study estimated SF using a multilayer feed-forward back-propagation network (BPNN) model utilizing the BP learning method. BPNN models have input, hidden, and output layers. To ensure performance evaluation consistency, the BPNN model was developed using a statistical model database. The network input vector contains all the independent variables. Thus, the input layer has the number of neurons per model as an independent variable. SF is one dependent variable in this research, and the output layer has neurons equal to outputs [33]. To find the value of SF, an artificial neural network (ANN) was used along with the MATLAB program.

Eight separate neural networks, trained with different random subsets of the training dataset, create the artificial neural network ensemble. **Fig. 3** shows the hidden layers, where W is the weight and b is the bias. Databases for validation and testing purposes. Network training received 70% of the data, with the remaining 30% split evenly between model validation (15%) and testing (15%). In each of the three stages of creating the governing model, the following selection criteria are laid out:

• The training group consists of 462 data observations that are utilized to train the artificial ANN. The weights (W) and biases (b) are modified to decrease the overall performance of the ANN.

• Validation group: 99 data observations oversee and evaluate the training process to avoid premature neural network model fitting. It would ensure that the neural network understands the basic functional relationships or patterns, not just memorizes them. A separate set of observations for validation stops overfitting during training. This makes sure that the artificial neural network (ANN) model can predict new observations from different places or years.

• Test group: 99 observations were utilized to assess the predictive capability of the ANN for the output variable, SF. These observations were not included in the model's training or validation datasets.

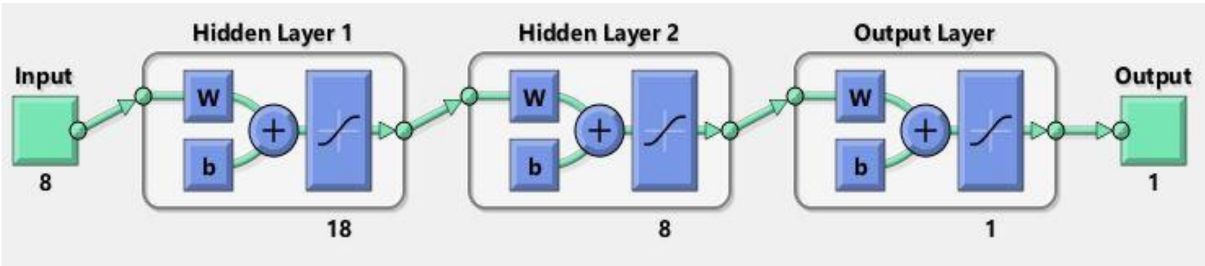


Fig. 3. Detailed Architecture and Overview of the Design of the Governing Model.

RESULTS AND DISCUSSION

The seasonal factors were forecasted using two methodologies: regression analysis and artificial neural networks. After calibrating the models, their efficiency was evaluated with the help of the test dataset. Using four performance indicators, we compared the model: the coefficient of determination (R2), the coefficient of correlation (R), the mean absolute error (MAE), and the mean square error (MSE). These metrics were used to assess the outcomes of the predictions, which are shown in Table 3. The following equations present the formulation of the measuring formula:

$$R^2 = 1- \frac{\sum_1^n (Y_i - \hat{y}_i)^2}{\sum_1^n (Y_i - \mathrm{m})^2} \tag{4}$$

$$\mathrm{MAE} = \frac{\sum_1^n |Y_i - \hat{y}_i|}{n} \tag{5}$$

$$\mathrm{MSE} = \frac{\sum_1^n (Y_i - \hat{y}_i)^2}{n} \tag{6}$$

yi, which corresponds to the actual value, represents the observed seasonal factor on segment i. The predicted value is denoted by $\hat{y}$, the sample size is represented by n, and the average seasonal factors in the dataset are denoted by m.

There is a comparison in Table 3 between the outcomes of using R2, R, MAE, and MSE to show the outcomes of using artificial neural networks and regression.

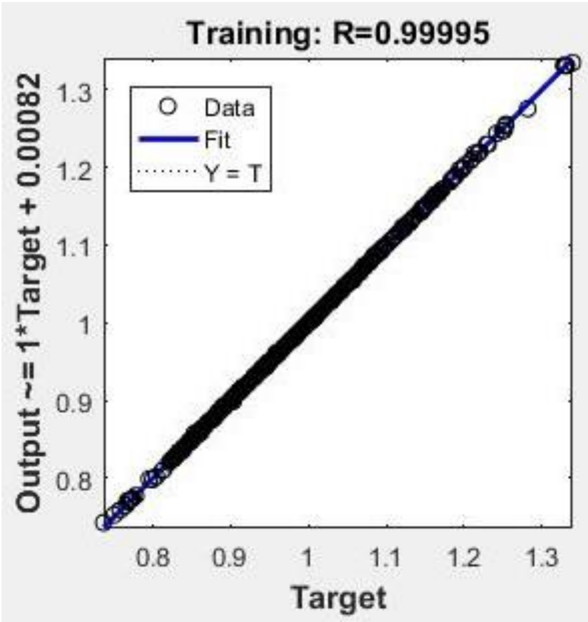
Table 3. Accuracy Measurements

Accuracy Measurements	R2	R	MAE	MSE
Regression	0.971630	0.990086	0.010911	0.000401
Artificial Neural Networks	0.999712	0.99983	0.000784	0.000002

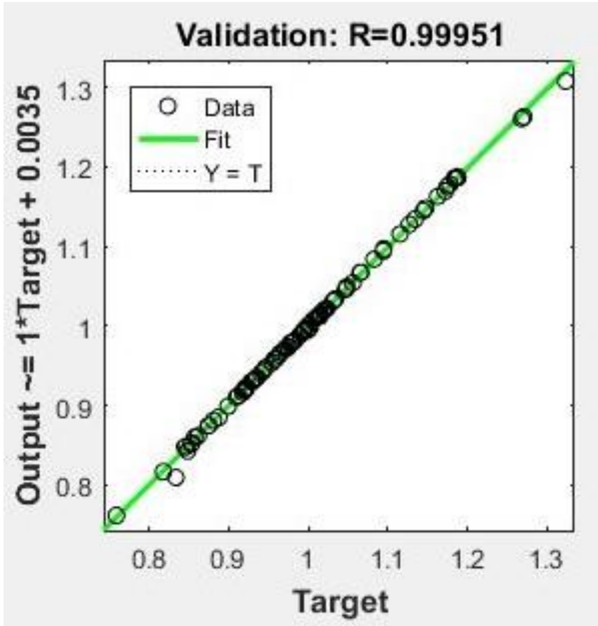
The statistical metric known as R2 allows one to quantify how well the regression model describes the data. The range of values is from 0 to 1, where higher values indicate a more optimal fit. The R2 for the artificial neural network model is 0.999712, signifying that the model accounts for 99% of the variability in the data. R measures the strength of a linear relationship between two variables in statistics. The range of values is from -1 to 1, where larger absolute values indicate a more pronounced association. The artificial neural network model has an R-value of 0.99983, indicating a robust positive correlation between the variables.

The mean absolute error (MAE) is a statistical term that quantifies the average absolute deviation when comparing predicted and measured outcomes. This metric measures the degree of precision that the model exhibits. The MAE for the artificial neural network model is 0.000784. Mean square error (MSE) is a statistical measure that quantifies the average of the squared differences between the expected and actual values. Additionally, it serves to measure the accuracy of the model. The MSE value for the artificial neural network model is 0.000002.

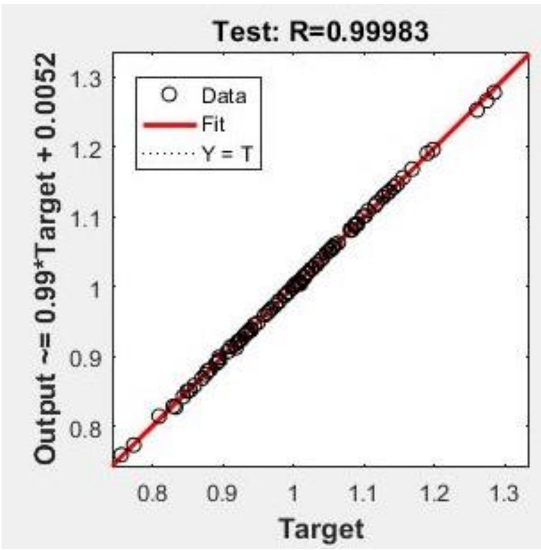
The artificial neural network model demonstrates more accuracy in comparison to the regression model, as seen by its elevated R^2 value as well as its lowest MAE and MSE values. The findings suggest that the artificial neural network model is better suited for the data in comparison to the regression model. **Fig. 4** exhibits correlation plots of the models that were accomplished using artificial neural networks.



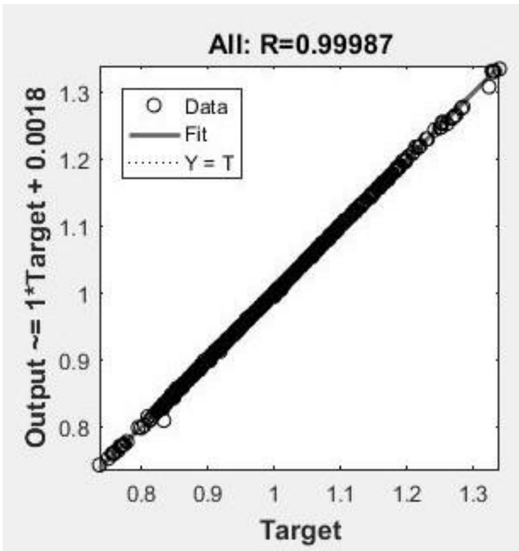
a) Correlation Plots For Training Data



b) Correlation Plots For Validation Data



a) Correlation Plots For Testing Data



b) Correlation Plots For All Data

Fig. 4. Correlation Plots of the Developmental Phases of the Model.

CONCLUSIONS AND RECOMMENDATIONS

The topic of seasonal variables is complicated. Although numerous research studies have employed diverse methodologies to ascertain seasonal elements, understanding the root causes of seasonal changes in traffic and constructing models for forecasting seasonal factors has posed substantial difficulty. This dissertation research aims to improve the accuracy of current methods used to estimate seasonal factors by implementing a more data-driven and impartial methodology for Egyptian road analysis.

This paper will be pioneering work in the field of Egyptian transportation planning in this specific direction.

The limits of the regression approach were identified after an examination of the method. Consequently, a fresh strategy based on the artificial neural network was proposed. An explanation of the method for determining the seasonal components was provided. Using constantly recorded ATR data from eleven ATR sites in Egypt, we compared the two techniques. In terms of performance, the neural network approach outperformed the regression method. The neural network's better results showed that it could solve problems with traffic pattern recognition.

Thus, it can be inferred that artificial neural networks (ANN) can be utilized to forecast traffic seasonal factors on Egyptian roadways. The neural network approach offers an advantage by eliminating the need for classifying ATR sites and assigning sample sites to ATR groups. The writers of this research concluded that individuals with limited experience can utilize neural networks as an alternative to complex statistical methods in their analysis.

After acquiring the SF values, it's clear that the suggested method, which is based on socio-economic and road features, makes estimating AADT easier and more accurate than traditional methods. ATRs are not necessary for data acquisition. The neural network application is suitable for enhancing predictions, while statistical models are suitable for mathematical formulation. The research article contributes to the design of effective transportation for the future to address the most pressing transportation issues that Egypt is presently facing, including traffic congestion.

ACKNOWLEDGMENTS

The authors wish to acknowledge the members of the General Authority for Roads, Bridges, and Land Transport (GARBLT) in Egypt for providing their data for this research.

CONFLICT OF INTEREST

The authors have no financial interest to declare in relation to the content of this article.

NOTATIONS

AADT: Annual Average Daily Traffic

AASHTO: American Association of State Highway and Transportation Officials

AF. Toll: Faiyum Toll

AGA – MTG: Aga Mit Ghamr

AHM - ISM: Abu Hamad Ismailia

ALX – DMR: Alexandria Damanhur

ANN: Artificial Neural Network

ATR: Automatic Traffic Recorder

AYT – BNS: Ayyat Beni Suef

CAI – TNL: Cairo Tunnel

DFR – ZIT: Defra Kafr Elziat

DMR – TNT: Damanhur Tanta

GARBLT: General Authority for Roads, Bridges, and Land Transport

GDP: Gross Domestic Product

MADT: Monthly Average Daily Traffic

MF: Monthly Factor

MNY – ASU: Minya Asyut

MST – BLB: Mostorod Belbeis

PCI: Per Capita Income

POP: Population

PTCs: Permanent Traffic Counts

SF: Seasonal Factor

SHBRA – BNH: Shubra Banha

SPTCs: Short Period Traffic Counts

References

[1] Bassan, Shy. "A statistical practical methodology of statewide traffic pattern grouping and precision analysis." Canadian Journal of Civil Engineering 36.3 (2009): 427-438. Retrieved from: https://cdnsciencepub.com/doi/abs/10.1139/L08-137

[2] Spławińska, M. "Models for determining annual average daily traffic on the national roads." Archives of Civil Engineering 61.2 (2015). Retrieved from: https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-fb9b8be6-6ad2-40a2-af24-ae6583a37f08

[3] Faghri, A. R. D. E. S. H. I. R., and J. I. U. Y. I. Hua. "Roadway seasonal classification using neural networks." Journal of Computing in Civil Engineering 9.3 (1995): 209-215. Retrieved from: https://ascelibrary.org/doi/abs/10.1061/(ASCE)0887-3801(1995)9:3(209)

[4] Guide, Traffic Monitoring. "Traffic Monitoring Guide." Federal Highway Administration (2001). Retrieved from: https://rosap.ntl.bts.gov/view/dot/41607

[5] Yang, Shanshan. "Improving Seasonal Factor Estimates for Adjustment of Annual Average Daily Traffic." (2012). Retrieved from: https://digitalcommons.fiu.edu/etd/709/

[6] Li, Min-Tang, Fang Zhao, and Yifei Wu. "Application of regression analysis for identifying factors that affect seasonal traffic fluctuations in Southeast Florida." Transportation Research Record 1870.1 (2004): 153-161. Retrieved from: https://doi.org/10.3141/1870-20

[7] Gastaldi, Massimiliano, Gregorio Gecchele, and Riccardo Rossi. "Estimation of Annual Average Daily Traffic from one-week traffic counts. A combined ANN-Fuzzy approach." Transportation Research Part C: Emerging Technologies 47 (2014): 86-99. Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S0968090X14001818

[8] Spławińska, M. "Factors determining seasonal variations in traffic volumes." Archives of Civil Engineering 63.4 (2017): 35-50. Retrieved from: http://archive.sciendo.com/ACE/ace.2017.63.issue-4/ace-2017-0039/ace-2017-0039.pdf

[9] Sharma, Satish C. "Improved classification of Canadian primary highways according to type of road use." Canadian Journal of Civil Engineering 10.3 (1983): 497-509. Retrieved from: https://cdnsciencepub.com/doi/abs/10.1139/l83-077

[10] Ritchie, Stephen G. "Statistical approach to statewide traffic counting." Transportation Research Record 1090 (1986): 14-21. Retrieved from: https://www.safetylit.org/citations/index.php?fuseaction=citations.viewdetails&citationIds[]=citjournalarticle_688624_38

[11] Yang, Shanshan, et al. "Estimation for seasonal factors of similarity-based traffic for urban roads in Florida." Transportation research record 2121.1 (2009): 74-80. Retrieved from: https://doi.org/10.3141/2121-08

[12] Zhao, Fang, Min-Tang Li, and Lee-Fang Chow. Alternatives for estimating seasonal factors on rural and urban roads in Florida. No. Final Report. 2004. Retrieved from: https://trid.trb.org/view/740949

[13] Lingras, Pawan. "Classifying highways: hierarchical grouping versus Kohonen neural networks." Journal of Transportation Engineering 121.4 (1995): 364-368. Retrieved from: https://ascelibrary.org/doi/abs/10.1061/(ASCE)0733-947X(1995)121:4(364)

[14] Zhao, Fang, and Soon Chung. "Contributing factors of annual average daily traffic in a Florida county: exploration with geographic information system and regression models." Transportation research record 1769.1 (2001): 113-122. Retrieved from: https://doi.org/10.3141/1769-14

[15] Selby, Brent, and Kara M. Kockelman. "Spatial prediction of traffic levels in unmeasured locations: applications of universal kriging and geographically weighted regression." Journal of Transport Geography 29 (2013): 24-32. Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S0966692312003079

[16] Fu, Miao, J. Andrew Kelly, and J. Peter Clinch. "Estimating annual average daily traffic and transport emissions for a national road network: A bottom-up methodology for both nationally-aggregated and spatially-disaggregated results." Journal of Transport Geography 58 (2017): 186-195. Retrieved from: https://www.sciencedirect.com/science/article/pii/S0966692316307244

[17] Apronti, Dick, et al. "Estimating traffic volume on Wyoming low volume roads using linear and logistic regression methods." Journal of traffic and transportation engineering (English edition) 3.6 (2016): 493-506. Retrieved from: https://www.sciencedirect.com/science/article/pii/S209575641530670X

[18] Sfyridis, Alexandros, and Paolo Agnolucci. "Annual average daily traffic estimation in England and Wales: An application of clustering and regression modelling." Journal of Transport Geography 83 (2020): 102658. Retrieved from: https://www.sciencedirect.com/science/article/pii/S096669231930568X

[19] Nandakumar, Ramjit, and Mithun Mohan. "Analysis of traffic growth on a rural highway: A case study from India." Journal of European Transport, Vol., 74, (2019). Retrieved from: https://idr.nitk.ac.in/jspui/handle/123456789/10345

[20] Reddy, I. Srinivasula, and Vikas Mendi. "Future Traffic Prediction from Short Period Traffic Data." IOP Conference Series: Materials Science and Engineering. Vol. 1006. No. 1. IOP Publishing, 2020. Retrieved from: https://iopscience.iop.org/article/10.1088/1757-899X/1006/1/012029/meta

[21] Baffoe-Twum, Edmund, Eric Asa, and Bright Awuku. "Estimation of annual average daily traffic (AADT) data for low-volume roads: a systematic literature review and meta-analysis." Emerald Open Research 1.5 (2023). Retrieved from: https://www.emerald.com/insight/content/doi/10.1108/EOR-05-2023-0010/full/html

[22] Sharma, Satish C., et al. "Neural networks as alternative to traditional factor approach of annual average daily traffic estimation from traffic counts." Transportation research record 1660.1 (1999): 24-31.Retrieved from: https://doi.org/10.3141/1660-04

[23] Lam, William HK, and Jianmin Xu. "Estimation of AADT from short period counts in Hong Kong—A comparison between neural network method and regression analysis." Journal of Advanced Transportation 34.2 (2000): 249-268. Retrieved from: https://onlinelibrary.wiley.com/doi/abs/10.1002/atr.5670340205

[24] Celikoglu, Hilmi Berk, and Hikmet Kerem Cigizoglu. "Modelling public transport trips by radial basis function neural networks." Mathematical and computer modelling 45.3-4 (2007): 480-489. Retrieved from: https://www.sciencedirect.com/science/article/pii/S0895717706002639

[25] Celikoglu, Hilmi Berk, and Hikmet Kerem Cigizoglu. "Public transportation trip flow modeling with generalized regression neural networks." Advances in Engineering Software 38.2 (2007): 71-79. Retrieved from: https://www.sciencedirect.com/science/article/abs/pii/S0965997806001153

[26] Zeng, Dehuai, et al. "Short term traffic flow prediction using hybrid ARIMA and ANN models." 2008 workshop on power electronics and intelligent transportation system. IEEE, 2008. Retrieved from: https://ieeexplore.ieee.org/abstract/document/4634929/

[27] Mahanta, Jahnavi. "Introduction to neural networks, advantages and applications." Towards Data Science 13 (2017).

[28] Mijwel, Maad M. "Artificial neural networks advantages and disadvantages." Mesopotamian Journal of Big Data 2021 (2021): 29-31. Retrieved from: https://journals.mesopotamian.press/index.php/bigdata/article/view/225

[29] Japan International Cooperation Agency (JICA): Oriental Consultants Co., Ltd. ALMEC Corp.: Katahira & Engineers International Y. "The comprehensive study on the master plan for nationwide transport system in the Arab republic of Egypt Final Report (2012)." Retrieved from: https://open jicareport.jica.go.jp/

[30] Chakraborty, Pranamesh, and Partha Chakroborty. Estimation of Annual Average Daily Traffic (AADT) for Indian Highways. No. 15-3896. 2015. Retrieved from: https://pranamesh.github.io/files/2015-TRB-SPTC.pdf

[31] Kaggle. "State of Machine Learning and Data Science 2020." Retrieved from: https://storage.googleapis.com/kagglemedia/surveys/Kaggle%20State%20of%20Machine%20Learning%20and%20Data%20Science%202020.pdf

[32] Sun, Qianqian. Nationwide Annual Average Daily Traffic (AADT) Estimation on Non-Federal Aid System (NFAS) Roads by Machine Learning with Data Mining of Built-In Environment. Diss. University of Maryland, College Park, 2020. "Retrieved from: https://www.proquest.com/openview/9acd2f0b1aed9474d31fa6cbb62a2964/1?pqorigsite=gscholar&cbl=18750&diss=y

[33] Duddu, Venkata Ramana, and Srinivas S. Pulugurtha. "Principle of demographic gravitation to estimate annual average daily traffic: Comparison of statistical and neural network models." Journal of Transportation Engineering 139.6 (2013): 585-595. Retrieved from: https://ascelibrary.org/doi/abs/10.1061/(ASCE)TE.1943-5436.0000537