View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Alexandria Engineering Journal (2013) 52, 455–466

FACULTY OF ENGINEERING
ALEXANDRIA UNIVERSITY

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

ORIGINAL ARTICLE

New models to evaluate the level of service and capacity for rural multi-lane highways in Egypt

Ahmed Mohamed Semeida *

Department of Civil Engineering, Faculty of Engineering, Port Said University, Egypt

Received 5 October 2012; revised 28 February 2013; accepted 6 April 2013
Available online 30 April 2013

KEYWORDS

Multi-lane highways
Heavy vehicles
Level of service
Capacity
Artificial neural networks

Abstract Multi-lane highways represent the majority of the total length of highway network in Egypt. The road geometry and the percentage of heavy vehicles (HVs) are considered the most important factors affecting the level of service (LOS) and capacity for any roadway. Therefore, this paper aims to explore the relationship between the road geometric characteristics and HV, and the LOS and capacity by two ways. First is the statistical modeling and second is the modeling by artificial neural networks (ANNs). In this research, the traffic and road geometric data are collected from mid-tangent points at 45 different sites that are located in desert and agricultural highways. The results showed that the ANN modeling gives the best models for estimating LOS and capacity. Also, it is better for analysis to separate the desert and agricultural sites. In addition, the most influential variables on LOS and capacity in desert sites are HV and lane width (LW), respectively, while in agricultural sites are LW and existence of side access (SA), respectively. These results are so important for road authorities in Egypt as they can determine LOS and capacity for different tangent sections and improve the traffic performance of them in the future.

© 2013 Production and hosting by Elsevier B.V. on behalf of Faculty of Engineering, Alexandria University.

1. Introduction

The transportation system in Egypt is suffered from limited roadway infrastructure and the lack of operation and management experience. Among the most critical issues in highway planning and management is to explore the effectiveness of road geometric characteristics and the percentage of HV in traffic composition on LOS and capacity at multi-lane rural highways. Rural multi-lane highways are an important type of uninterrupted flow facilities in which there is no obstruction to the movement of vehicles along the road. Such facilities represent the majority of the highway system in Egypt. Highway Capacity Manual (HCM) [1] uses density in terms of passenger cars per kilometer per lane as the primary level of service (LOS) measure for multi-lane highways and also [1] uses free flow speed (FFS) as the primary capacity measure for the same type of highways. Therefore, this paper aims to evaluate LOS and capacity on multi-lane highways by two modeling techniques. First is the traditional statistical technique and the second is ANN technique.

* Tel.: +20 1224411977.
E-mail address: semeida77@hotmail.com.

Peer review under responsibility of Faculty of Engineering, Alexandria University.

ELSEVIER Production and hosting by Elsevier

1110-0168 © 2013 Production and hosting by Elsevier B.V. on behalf of Faculty of Engineering, Alexandria University.
http://dx.doi.org/10.1016/j.aej.2013.04.003

Field data on multi-lane highways in Egypt are used in this investigation. The analysis considers 45 tangent sections from two categories of highways. The first consists of two desert roads (Cairo–Alexandria and Cairo–Ismailia desert roads), and the second consists of two agricultural roads (Cairo–Alexandria and Tanta–Damietta agricultural roads). Then, the paper includes two separate relevant analyses. The first analysis uses the regression models to investigate the relationships between LOS and capacity as dependent variables, and roadway factors and HV percentage as independent variables. The road factors are lane, pavement, median width, lateral clearance, number of lanes in each direction, and existence of side access along each section. The second analysis uses the ANN to explore the previous relationships and comparing the results. According to the objectives in this research, road authorities in Egypt can determine LOS and capacity for different multi-lane highway tangent sections and improve the traffic performance of them in the future.

Several researches have been carried out to analyze the effect of road geometry and traffic composition on LOS and Capacity for multi-lane highways. Kerner [2] confirmed that the determination of capacity and LOS for any highway is one of the most important applications of any traffic theory. Some previous theories and empirical researches focused on the interrelationships among the influence of capacity, traffic features, and geometric elements on uninterrupted multi-lane highways [3–6]. Bang [7] in their study for establishing Indonesia HCM mentioned travel speed as the main measure of performance of road segments. Yang and Zhang [8] have established based on their extensive field survey of traffic flow on multi-lane highways in Beijing and subsequent empirical model development that the average roadway capacity per hour per lane on four-lane, six-lane and eight-lane divided carriageways is 2104, 1973 and 1848 passenger car unit, respectively. Velmurugan et al. [9] studied the speed-flow characteristics on varying types of multi-lane highways in India and subsequent capacity of Indian multi-lane highways had been estimated based on traditional and microscopic simulation models. Arasan and Arkatkar [10] studied the effect of variation of traffic composition, road width, magnitude of upgrade and its length on Indian highways capacity, and subsequently, it was concluded that highway capacity significantly changes with change in traffic volume composition, width of roadway, magnitude of upgrade, and its length. Sakai et al. [11] used an empirical approach to produce LOS measure for basic expressway segments in Japan incorporating Customer Satisfaction (CS). It was concluded that LOS and CS were confirmed to have a nonlinear relation. García et al. [12] studied the effect of traffic calming devices on the cross-town roads capacity in Spain based on different type and spacing of devices. The results showed that capacity varied between 810 and 1300 vehicles per hour per lane with traffic calming devices spacing from 25 to 400 m.

2. Data collection and methods of LOS and capacity estimation

This section is divided into two main parts as (1) study sites and field data, and (2) equations of LOS and capacity determination.

2.1. Study sites and field data

This research focuses on the rural multi-lane highways in Egypt. Therefore, the analysis of this paper uses 45 sites (sections) from four main multi-lane highways in Egypt. These roads include Cairo–Alexandria Agricultural Highway (CAA), Tanta–Damietta Agricultural Highway (TDA), Cairo–Alexandria Desert Highway (CAD), and Cairo–Ismailia Desert Highway (CID). These sites are divided into 21 sections in desert roads and 24 sections in agricultural roads. Each section length is 100 m. The chosen sites are located on straight sections with level terrain to avoid the effect of the longitudinal gradient and to be far from the influence of horizontal curves. The collected data are divided into three types as road geometric characteristics, vehicles speed, and traffic volume data.

2.1.1. Road geometric data

These data are collected directly from site investigation that included lane width, lateral clearance, number of lanes in each direction, median width, and pavement width. All the previous variables, their symbols, and statistical analysis are provided in Table 1.

2.1.2. Vehicles speed data

There are two main types of the collected speed. First is FFS data and the second is average travel speed of passenger cars (ATS_{pc}) data. Each of them is measured in field as follows.

2.1.2.1. FFS data. FFS are collected for passenger cars only for 45 sections (the passenger cars include taxis, vans/jeeps, and microbuses). Spot speed data are collected using radar gun (version LASER 500 with ± 1 km/h accuracy) that is placed at midpoint of each section so as to be invisible to drivers [13]. Vehicles traveling in free-flow conditions are considered to have time headways of at least 8 s [14]. The number of speeds collected at each site range from 100 to 160, which led to a total of 6300 spot speeds. Speeds are carried out in

Table 1 Statistical analysis and symbols of independent variables.

Variable	Symbol	Max.	Min.	Avg.	SD
1 – Lane width in meters	LW	3.65	3.00	3.5	0.2
2 – Pavement width in one direction in meters	PW	14.6	7	8.9	2.3
3 – Lateral clearance in meters	LC	3.6	1.5	2.4	0.7
4 – No of lanes in each direction in lanes	NL	4	2	3	–
5 – Median width in meters	MW	16	1	7.2	3.2
6 – Existing of side access (1 if exiting; 0 otherwise)	SA	1	0	–	–
7 – Percentage of heavy vehicles %	HV	28	2.1	8.1	5.9

working days, during daylight hours. During all data collection periods, the weather is clear and the pavement is dry and in a good condition.

2.1.2.2. ATS_{pc} data. ATS_{pc} is detected during the measurement period in the peak flow of traffic as a worst case for recording. ATS_{pc} is measured by taking a constant distance (100 m) for each section and recording the time during which vehicles traveled at this constant distance. Next, the speed is calculated by dividing the constant distance per the recorded time. The sample of cars at any section is not smaller than 100 vehicles [1].

All speed data are recorded in Table 2.

2.1.3. Traffic volume data

The purpose of collecting traffic volume is to reach the design hourly volume (V) and detect the percentage of HV along each section (HV include semi-trucks, trucks, and trucks trailer that have at least one axle with dual wheels). This can be executed as follows. First, Average Annual Daily Traffic must be estimated at each site as [15]. As there are many sites (45 sections), a manual traffic counting is done for an hour at each site. Surely, the results of any manual counting must be enlarged and corrected to be converted into Average Annual Daily Traffic (AADT) [16]. This needs three factors as follows: hourly factor (HF) (100/% volume in counting hours), Daily Factor (DF), and Seasonally Factor (SF). These factors are constant for each road and are obtained from General Authority of Roads, Bridges and Land Transport (GARBLT) [17]. Therefore, the calculated AADT is given by Eq. (1) [15]. The values of AADT are listed in Table 3.

$$AADT = Counting\ Volume \times HF \times DF \times SF \ (\text{veh/day}) \quad (1)$$

Second, the design hourly volume (V) can be calculated by [1].

$$V = AADT \times K \ (\text{veh/h}) \quad (2)$$

where V is the design hourly volume (typically, the 30th highest annual hourly volume); AADT the Average Annual Daily Traffic in vehicle per day; and K is the factor used to convert annual average daily traffic to a specified annual hourly volume ($K = 0.1$ for rural roads).

During any particular hour, traffic volume will likely be greater in one direction than in the other. Directional distribution (D) is an important factor in capacity and quality of service analysis. To convert hourly volume to hourly directional volumes, the hourly volumes are multiplied by the D-factor as shown in [1].

$$Dir.\ V = V \times D \ (\text{veh/h}) \quad (3)$$

Note that $D = 0.6$ for rural roads as stated by [1]. *Dir. V* values for all sites are recorded in Table 3. Finally, the percentage of HV is extracted from the recorded manual counting as there is a classification of vehicles for passenger cars and HV (Table 3).

2.2. Determination of LOS and capacity for sites under study

2.2.1. LOS determination

To determine LOS for each site, the density is determined first from [1].

$$Density = v_p/\text{ATS}_{pc} \ (\text{pc/km/lane}) \quad (4)$$

where v_p = 15-min passenger-car equivalent flow rate (pc/h/lane) and is computed from [1].

$$v_p = V/(\text{PHF} \times N \times f_{HV} \times f_p) \quad (5)$$

where V is the hourly volume in one direction (veh/h), PHF the peak-hour factor (PHF = 0.88 for rural roads), N the number of lanes in the selected direction, f_{HV} the heavy-vehicle adjustment factor (0.97–0.9) depending on HV percentage, and f_p is the driver population factor (1) as the driver is familiar with the highway.Consequently, LOS is determined for each site from density value which reflects the degree of vehicles congestion on each site. The LOS and density for all sites are determined and listed in Table 3.

2.2.2. Capacity determination

The capacity for the lane of each highway is specified by GARBLT [17] as follows. The average lane capacity for CAA, CAD, CID, and TDA is given as 1800, 2200, 2200, and 1800 vehicle/hour, respectively. Due to different road geometric and traffic conditions along each highway, the previous capacities may not be correct and need to be revised. To rectify these capacities, HCM 2000 [1] gives Eqs. (6) and (7) for capacity determination depending on FFS as follows:

$$Capacity = 1000 + 20 \times \text{FFS}\ (\text{pc/h/lane});\ \text{For FFS} \le 60\,\text{mile/h} \quad (6)$$

$$Capacity = 2200\ (\text{pc/h/lane}); \quad \text{For FFS} > 60\ \text{mile/h} \quad (7)$$

The capacities for all sites are calculated and listed in Table 3.

3. Theory of regression and ANN modeling and calculations

The theory of LOS and capacity modeling in the present research is divided into two main parts: (1) regression models and (2) ANN models.

Table 2 Free-flow speed and average travel speed for passenger cars data.

Site No.	FFS (km/hr)	ATS_{pc}	Site no.	FFS (km/h)	ATS_{pc}
1	73.4	63.92	24	70.12	65
2	66.88	60.05	25	73.24	67.1
3	93	80.73	26	71.95	61.95
4	77.5	63.52	27	94.26	83.17
5	95	84.66	28	90.51	77.93
6	62.5	53.31	29	98.03	84.41
7	57.25	49.36	30	89.6	77.15
8	64.09	58.5	31	98.72	85
9	68.13	63.43	32	93.49	80.5
10	70.81	65.3	33	103.36	89
11	110.76	95	34	100.7	86.71
12	80.4	69.86	35	98.22	84.58
13	71.03	61.54	36	106.27	91.5
14	79.35	68.14	37	96.29	82.91
15	68.75	60.16	38	96.67	83.24
16	85.67	74.41	39	96.81	83.36
17	69.92	60.2	40	105	90.41
18	116.14	100	41	88.62	76.3
19	69.68	60	42	87.33	76.37
20	111.49	96	43	84.11	77.08
21	74.56	63.2	44	90.65	80.18
22	113.82	98	45	68.25	61.53
23	92.91	84.5			

Table 3 Traffic count data, density, LOS, capacity values for all sites.

Site no.	Traffic count in both direction per hour	AADT (veh/day)	*Dir. V* (veh/h)	Overall HV percentage (%)	Density (veh/km/lane)	LOS	Capacity (veh/h/lane)
1	2207	25,022	1501	10	9.34	B	1918
2	2192	24,914	1495	12	11.99	C	1711
3	2135	24,215	1453	5.3	7	A	2163
4	1413	16,017	961	2.5	5.8	A	1969
5	1481	16,790	1007	3.5	4.59	A	2188
6	1712	19,408	1164	4.1	12.67	C	1781
7	1120	12,700	762	3.9	8.94	B	1716
8	2207	25,022	1501	20	17.14	D	1551
9	2222	25,194	1512	22	19.01	D	1477
10	1302	14,758	885	3.2	14.6	C	1510
11	1433	16,249	975	5	5.98	A	2200
12	2106	52,465	3148	24	19.12	D	2005
13	656	16,340	980	12	9.59	B	1888
14	894	22,282	1337	6	7.65	B	1992
15	2106	52,465	3148	28	24.62	E	1784
16	646	15,908	954	4.2	4.96	A	2071
17	645	15,891	953	2.8	9.12	B	1874
18	560	13,782	827	2.7	2.38	A	2200
19	535	13,402	804	2.1	7.69	B	1871
20	781	15,908	954	3.2	2.87	A	2200
21	647	15,920	955	3	8.58	B	1932
22	647	15,932	956	3.3	2.82	A	2200
23	645	15,884	953	4	3.45	A	2161
24	646	15,905	954	16.4	11.17	C	1508
25	401	9786	587	6	7.3	B	1684
26	893	16,206	972	12	9.45	B	1899
27	894	16,227	974	7	7.05	B	2178
28	834	15,132	908	7.2	6.86	B	2131
29	890	16,148	969	6.5	6.73	A	2200
30	895	16,238	974	7.5	7.44	B	2120
31	898	16,302	978	7.1	6.77	A	2200
32	895	16,251	975	7.2	7.13	B	2169
33	877	15,916	955	6.2	6.29	A	2200
34	892	16,193	972	6.8	6.58	A	2200
35	875	15,881	953	6.5	6.61	A	2200
36	893	16,217	973	6.2	6.22	A	2200
37	682	12,370	742	5.8	5.23	A	2200
38	691	12,538	758	6	5.29	A	2200
39	682	12,370	746	6.5	5.21	A	2200
40	691	12,538	752	6.1	4.87	A	2200
41	539	9786	587	6.3	4.5	A	2108
42	1036	25,819	1549	10.8	8.1	B	2092
43	985	24,530	1472	9.75	7.59	B	2051
44	961	23,930	1436	9.5	7.11	B	2133
45	2136	24,215	1453	13	9.53	B	1853

Table 4 Correlations between (density and capacity) and each of independent variables.

Variable	Integral		Inverse		Square		Square root		Inv. square		Inv. square root	
	R^2	P-value	R^2	P-value	R^2	P-value	R^2	P-value	R^2	P-value	R^2	P-value
Mathematical type of variable (density models)												
SA	*0.315*	*0.000*										
LW	0.67	0.000	0.693	0.000	0.656	0.000	0.676	0.000	*0.703*	*0.000*	0.688	0.000
PW	0.036	0.215	0.035	0.221	0.034	0.234	0.037	0.211	0.033	0.24	0.0.035	0.24
LC	0.518	0.000	0.515	0.000	0.503	0.000	*0.522*	*0.000*	0.501	0.000	0.52	0.000
NL	0.00	0.95	0.011	0.853	0.001	0.934	0.09	0.541	0.05	0.663	0.04	0.322
MW	0.373	0.000	0.12	0.02	0.302	0.000	*0.374*	*0.000*	0.03	0.29	0.22	0.001
HV	0.688	0.000	0.218	0.001	*0.756*	*0.000*	0.591	0.000	0.095	0.04	0.326	0.000
Mathematical type of variable (capacity models)												
SA	*0.71*	*0.000*										
LW	*0.241*	*0.000*	0.241	0.001	0.236	0.001	0.24	0.001	0.239	0.001	0.241	0.001
PW	0.007	0.6	0.006	0.63	0.01	0.42	0.008	0.53	0.064	0.65	0.007	0.7
LC	*0.394*	*0.000*	0.357	0.000	0.393	0.000	0.388	0.000	0.33	0.000	0.37	0.000
NL	0.006	0.509	0.01	0.632	0.003	0.634	0.007	0.463	0.013	0.364	0.01	0.392
MW	0.34	0.000	0.168	0.005	0.253	0.000	*0.365*	*0.000*	0.05	0.14	0.267	0.000
HV	0.21	0.002	0.023	0.32	*0.212*	*0.001*	0.17	0.005	0.000	0.9	0.06	0.11

3.1. Regression models

The collected data are used to investigate the relationships between LOS (density) and capacity as dependent variable and road geometric and HV percentage as independent variables. Simple regression is used to check the correlation coefficient (r) between each dependent variable and each of the independent variables. The independent variables that have relatively high r values (>0.5) are introduced into the multiple linear regression models. The form of multiple linear regression models is shown in

$$Y = \beta_o + \sum_{i=1}^{n} \beta_i \chi_i \tag{8}$$

where Y is the density or capacity, X_i the explanatory variables from 1 to n, β_o is the regression constant, and β_i is the regression coefficient.

Then, stepwise regression analysis is used to select the most statistically significant independent variables with dependent variable in one model. Stepwise regression starts with no model terms. At each step, it adds the most statistically significant term (the one with lowest P-value) until the addition of the next variable makes no significant difference. An important assumption behind the method is that some input variables in a multiple regression do not have an important explanatory effect on the response. Stepwise regression keeps only the statistically significant terms in the model. Finally, the R^2 (coefficient of determination) and (Root Mean Square Error) RMSE values are calculated for each model. Several precautions are taken into consideration to ensure integrity of the model as follows [18]:

(1) The signs of the multiple linear regression coefficients should agree with the signs of the simple linear regression of the individual independent variables and agree with intuitive engineering judgment.
(2) There should be no multicollinearity among the final selected independent variables.
(3) The model with the smallest number of independent variables, minimum RMSE, and highest R^2 value is selected.

3.1.1. LOS models

First, each independent variable is modeled with density by Simple regression. The analysis considers many mathematical forms of the independent variables (square, inverse, square root, etc). These results are shown in Table 4. From the results in Table 4, it can be concluded that two variables are excluded from the final model due to poor correlation with density. These variables are NL and PW. On the other hand, the most significant variables are SA, LW, LC, MW, and HV. The forms of each variable which is the most correlated with density are presented in the italicized cells in Table 4. Then, these variables are introduced into the multiple linear regression models. Consequently, stepwise regression analysis is used to select the most statistically significant independent variables with density in one model. The best model and its analysis are presented in Section 4.

3.1.2. Capacity models

To derive these models, the same steps are executed as LOS models. The difference is the exclusion of sections with FFS > 96 km/h as all the observed capacities at these sections are constant (2200 veh/h/ln). The results of simple regression are shown in Table 4. From the results in Table 4, it can be concluded that two variables are excluded from the final model due to poor correlation with capacity. These variables are NL and PW. On the other hand, the most significant variables are SA, LW, LC, MW, and HV. The forms of each variable which is the most correlated with capacity are presented in the italicized cells in Table 4. Consequently, stepwise regression analysis in SSPS Package is used to select the most statistically significant independent variables with capacity in one model. The best model and its analysis are presented in Section 4.

3.2. ANN models

In general, ANNs consist of three layers, namely the input, the hidden, and the output layers. In the input layer, the input variables of the problems are situated. The output layer contains the output variables of what is being modeled. In statistical terms, the input layer contains the independent variables and the output layer contains the dependent variables. The nodes between successive layers are connected by links each carrying a weight that quantitatively describes the strength of those connections, thus denoting the strength of one node to affect the other node [19]. ANNs typically start out with randomized weights for all their neurons. This means that they do not know anything and must be trained to solve the particular problem for which they are intended. When a satisfactory level of performance is reached, the training is ended and the network uses these weights to make a decision [20].

The experience in this field is extracted from researcher [20]. In his research, the multi-layer perceptron (MLP) neural network models give the best performance of all models. In addition, this network is usually preferred in engineering applications because many learning algorithm might be used in MLP. One of the commonly used learning algorithms in ANN applications is back propagation algorithm (BP), which is also used in this research (NeuroSolutions 7) [21].

The overall data set of 45 tangent sections is divided into a training data set and a testing data set. As in the literature, the training data set varies from 70% to 90% and the testing data set varies from 10% to 30%. Model performances are RMSE and R^2 for testing and training data set in one hand and for all data set in the other hand [22].

So many trials are done to reach the suitable percentage between training and testing data that gives the best performance for density and capacity models. In addition, over fitting can be avoided by randomize the 45 sites before training the network to reach the best performance for both training and testing data. The performance of testing data must be good as training data (R^2 must not be smaller than 0.7) [23]. Also, engineers can use ANN for predicting unknown dependent variables as follows. The best network which gives the best correlation and performance between dependent and independent variables for testing data for any problem is saved. Thus, the unknown dependent variables can be evaluated by putting any independent variables in the testing data. Then, the network would be trained and tested again and the unknown dependent variables would be calculated and appeared automatically in the testing results.

3.2.1. LOS models

First, each independent variable is training and testing with density. Therefore, the effective variables with density are as follows ($r > 0.5$): SA, LW, PW, LC, MW, and HV (6 variables) that have r values of 0.56, 0.86, 0.9, 0.73, 0.71, and 0.87, respectively. Then, the input variables (6 variables) are in input layer. One hidden layer and 1 desired variable (density) are in output layer with 45 observations used. The architecture of the ANN model is shown in Fig. 1. Observations are divided into training data set and testing data set. The results of many trials show that training data set of 38 observations (85% of all observations), and testing data set of 7 observations (15% of all observations) give the best performance for density model. The number of neurons in hidden layer is nearly half the total number of neurons at the input and output layers (3 neurons), which is set based on generally accepted knowledge in this field. The learning rule of Momentum is used. In addition, the suitable number of epochs (iterations) is 10,000. The previous conditions are suitable for quick convergence of the problem [20]. The best model and its analysis are presented in Section 4.

3.2.2. Capacity models

The same steps are executed as density models. The difference is the exclusion of sections with FFS > 96 km/h as all the observed capacities at these sections are constant (2200 veh/h/ln). The effective variables with capacity are as follows ($r > 0.5$): SA, LW, PW, LC, MW, and HV (6 variables) that have r values of 0.84, 0.61, 0.63, 0.69, 0.66, and 0.71, respectively. The results of many trials show that training data set of 24 observations (78% of all observations) and testing data set of 7 observations (22% of all observations) give the best performance for capacity model. In addition, the architecture of the ANN model is the same as density model. The best model and its analysis are presented in Section 4.

4. Results and discussions

4.1. Results of regression models

4.1.1. LOS model

The modeling is divided into three models. The first uses the 45 sections in the model. The second uses only the desert sections (21 sections). The last one uses only the agricultural sections (24 sections).

4.1.1.1. All roads model. There are four models that are statistically significant with density after stepwise regression using SSPS Package. All of the variables are significant at the 5% significance level for these four models (P-value is < 0.05). Finally, many models are excluded due to poor significance with density. Thus, the best model is as follows in Eq. (9) and shown in Fig. 2:

$$\text{Best model} \quad Density_{All} = 14.34 + 2.107(SA) - 5.55\sqrt{LC} + 186.4(HV^2) \quad (\text{Whereas, } R^2 = 0.905,\ \text{RMSE} = 1.44) \tag{9}$$

4.1.1.2. Desert roads model (21 sections). After stepwise regression, there are two models that are statistically significant with density. The best of them is as follows in Eq. (10) and shown in Fig. 2:

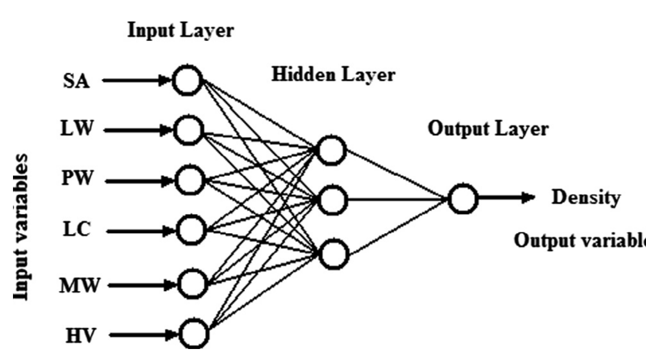

Figure 1 MLP network architecture of density model.

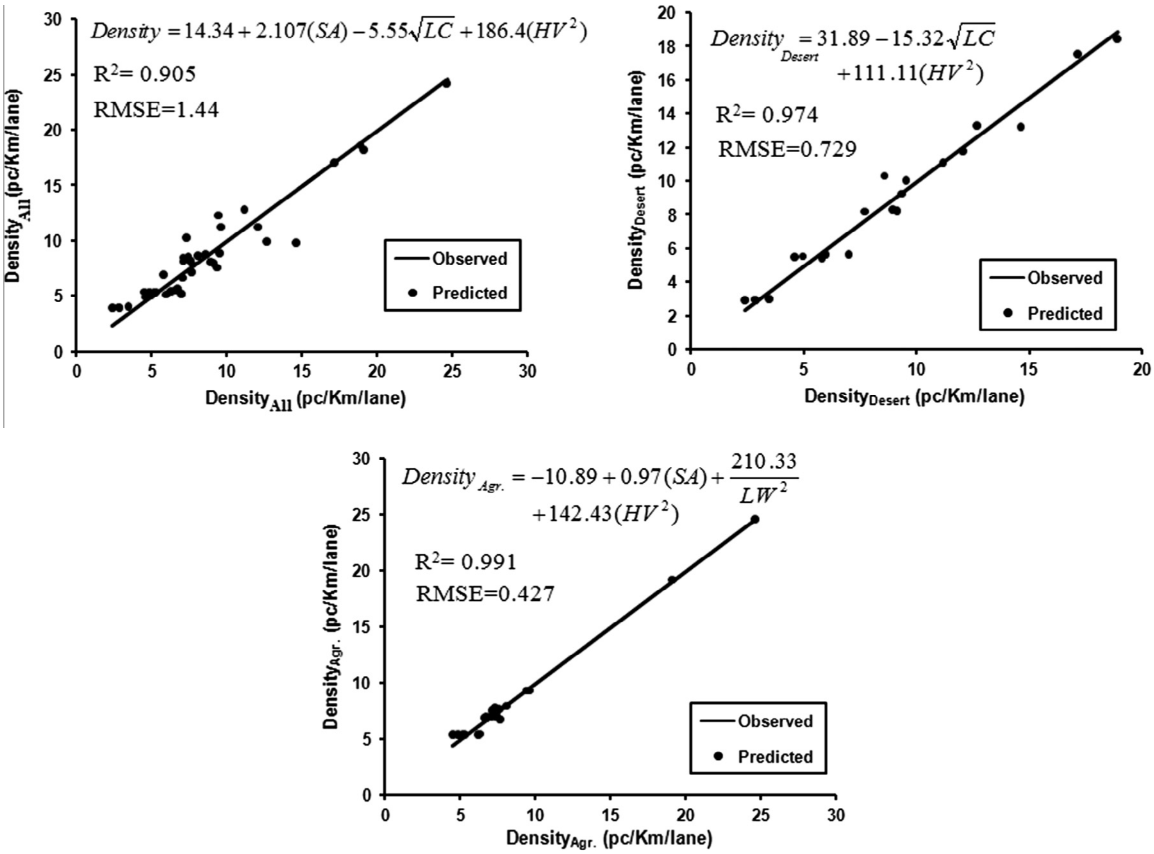

Figure 2 Observed and predicted density for the best three regression models.

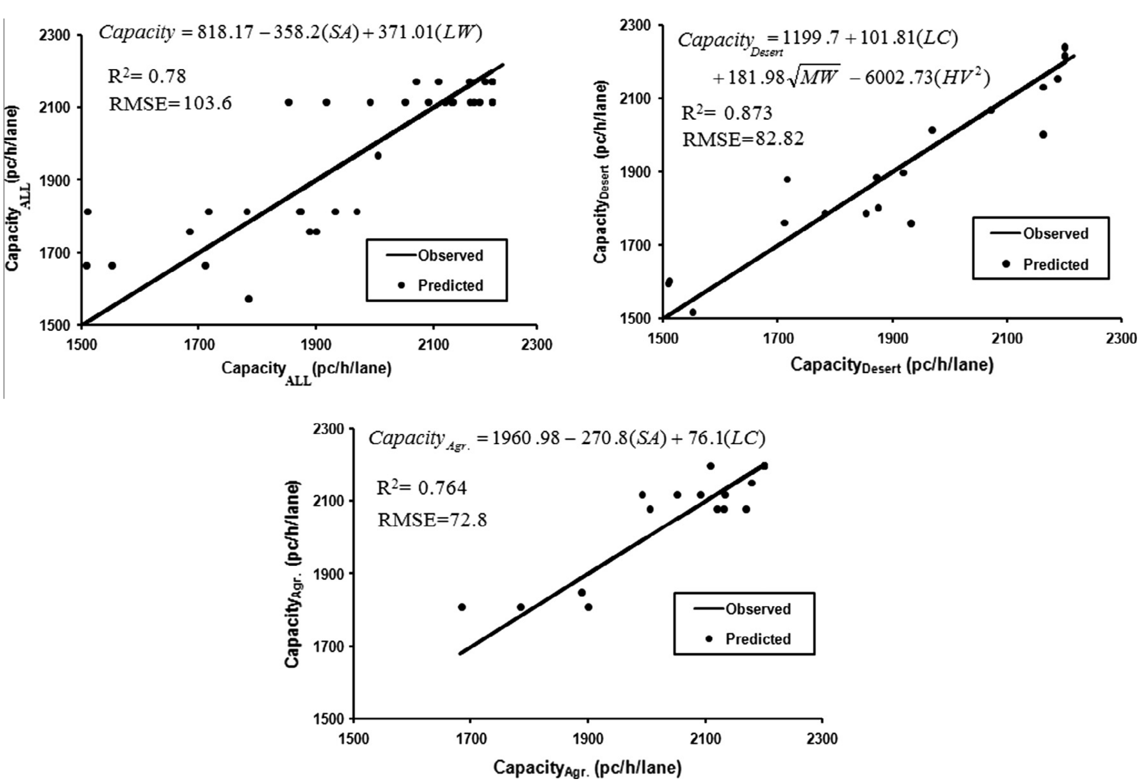

Figure 3 Observed and predicted capacity for the best three regression models.

Best model $$Density_{Desert} = 31.89 - 15.32\sqrt{LC} + 111.11(HV^2) \quad \text{(Whereas, } R^2 = 0.974, \text{ RMSE} = 0.729) \tag{10}$$

4.1.1.3. Agricultural roads model (24 sections). After stepwise regression, there are three models that are statistically significant with density. The best of them is as follows in Eq. (11) and shown in Fig. 2:

Table 5 Performances for the best ANN model (density and capacity).

Model type	Performance	Training	Testing	Overall model
Density	R^2	0.93	0.913	0.926
(All sections)	RMSE	1.26	1.00	1.22
Density	R^2	0.998	0.998	0.992
(Desert roads)	RMSE	0.199	0.913	0.437
Density	R^2	0.959	0.987	0.997
(Agricultural sections)	RMSE	0.963	0.537	0.26
Capacity	R^2	0.894	0.874	0.88
(All sections)	RMSE	72.5	83	78.2
Capacity	R^2	0.97	0.984	0.961
(Desert sections)	RMSE	42	35.8	45.34
Capacity	R^2	0.979	0.999	0.93
(Agricultural sections)	RMSE	22	59.2	41.96

Best model $$Density_{Agr.} = -10.89 + 0.97(SA) + \frac{210.33}{LW^2} + 142.43(HV^2) \quad \text{(Whereas, } R^2 = 0.991, \text{ RMSE} = 0.427) \tag{11}$$

Investigating the best models for the three cases, it is found that:

- The desert and agricultural models are so better than the collective model as they give higher statistical performance. Then, it is better to separate analysis for each road type.
- For agriculture model, the positive sign of the coefficient for SA means that the density increases with the existence of side access. The vehicles are to be more adjacent in these sections due to merging of side traffic flow with main traffic flow. Consequently, LOS becomes poorer than closed side sections. In addition, the positive sign of the coefficient for the inverse of square LW means that the wider LW decreases the density of this section and consequently improves LOS. The previous results are consistent with logic.
- For desert model, the negative sign of the coefficient for the square root of LC means that the density decreases with the increase in LC. Consequently, the wider LC improves LOS. Also, this result is rational.
- For both road models, the positive sign of the coefficient for square HV means that the density increases with the increase in HV. The sections with higher HV are more

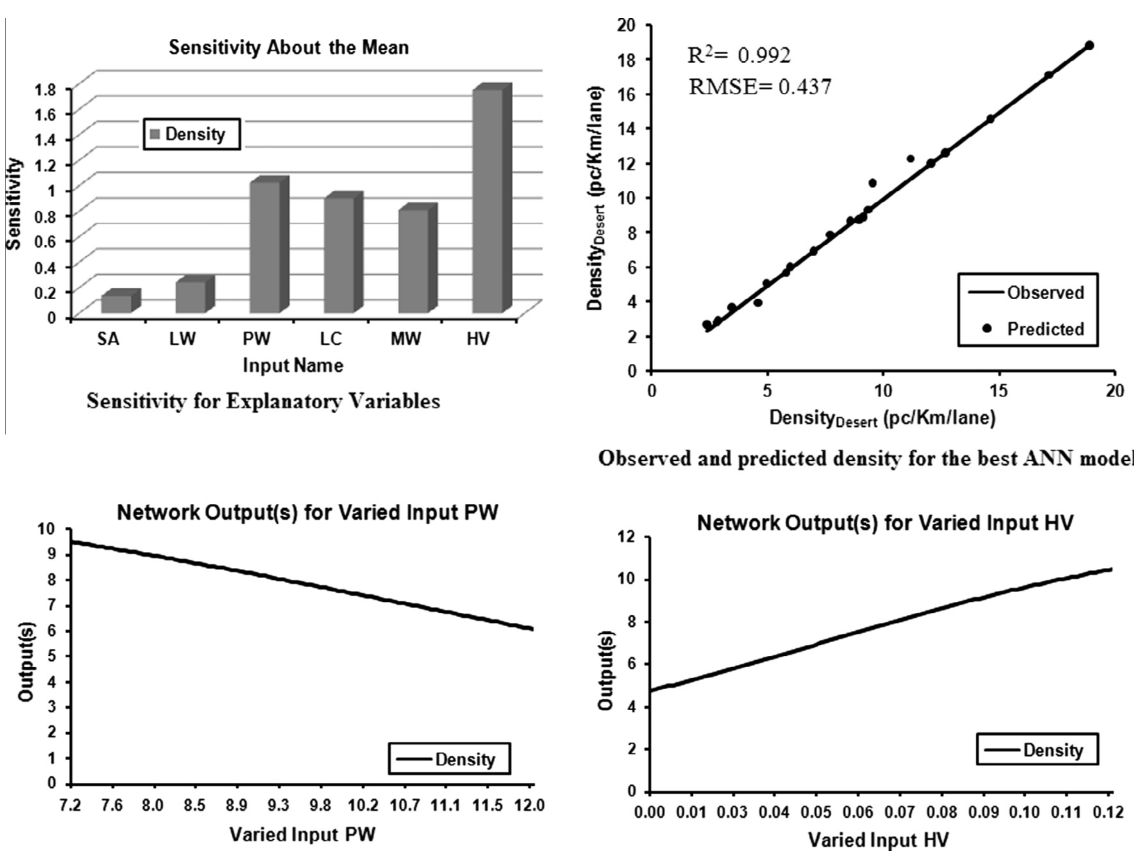

Figure 4 MLP results for the best density model (desert sections).

congestive and the drivers of passenger cars are annoyed with HV which force them to decrease their speed. This implies to the decrease in LOS.

4.1.2. Capacity model

As density model, the modeling is divided into three models excluding the sections with FFS > 96 km/h.

4.1.2.1. All roads model (31 sections). There are six models that are statistically significant with capacity after stepwise regression using SSPS Package. All of the variables are significant at the 5% significance level for these four models (*P*-value is <0.05). Thus, the best model is as follows in Eq. (12) and shown in Fig. 3:

$$\text{Best model} \quad Capacity_{ALL} = 818.17 - 358.2(SA) + 371.01(LW) \quad (\text{Whereas, } R^2 = 0.78,\ \text{RMSE} = 103.6) \tag{12}$$

4.1.2.2. Desert roads model (17 sections). After stepwise regression, there are three models that are statistically significant with capacity. The best of them is as follows in Eq. (13) and shown in Fig. 3:

$$Capacity_{Desert} = 1199.7 + 101.81(LC) + 181.98\sqrt{MW} - 6002.73(HV^2) \quad (\text{Whereas, } R^2 = 0.873,\ \text{RMSE} = 82.82) \tag{13}$$

4.1.2.3. Agricultural roads model (14 sections). After stepwise regression, there are three models that are statistically significant with capacity. The best of them is as follows in Eq. (14) and shown in Fig. 3:

$$Capacity_{Agr.} = 1960.98 - 270.8(SA) + 76.1(LC) \quad (\text{Whereas, } R^2 = 0.764,\ \text{RMSE} = 72.8) \tag{14}$$

Investigating the best models for the three cases, it is found that:

- As density model, the desert and agricultural models give more accurate results than the collective road model.
- For desert model, the positive sign of the coefficient of the square root of MW means that the wider MW improves the capacity of road. Generally in desert roads, the wider cross-section of road leads to widening of all the elements of highway cross-section that include MW. In addition, the negative sign of the coefficient for square HV means that the capacity decreases with the increase in HV. The sections with higher HV are more congestive and the drivers of passenger cars are annoyed with HV which force them to decrease their speed. The previous results are rational.
- For agriculture model, the negative sign of the coefficient for SA means that the capacity decreases with the existence of side access. The moving of vehicles is more difficult due

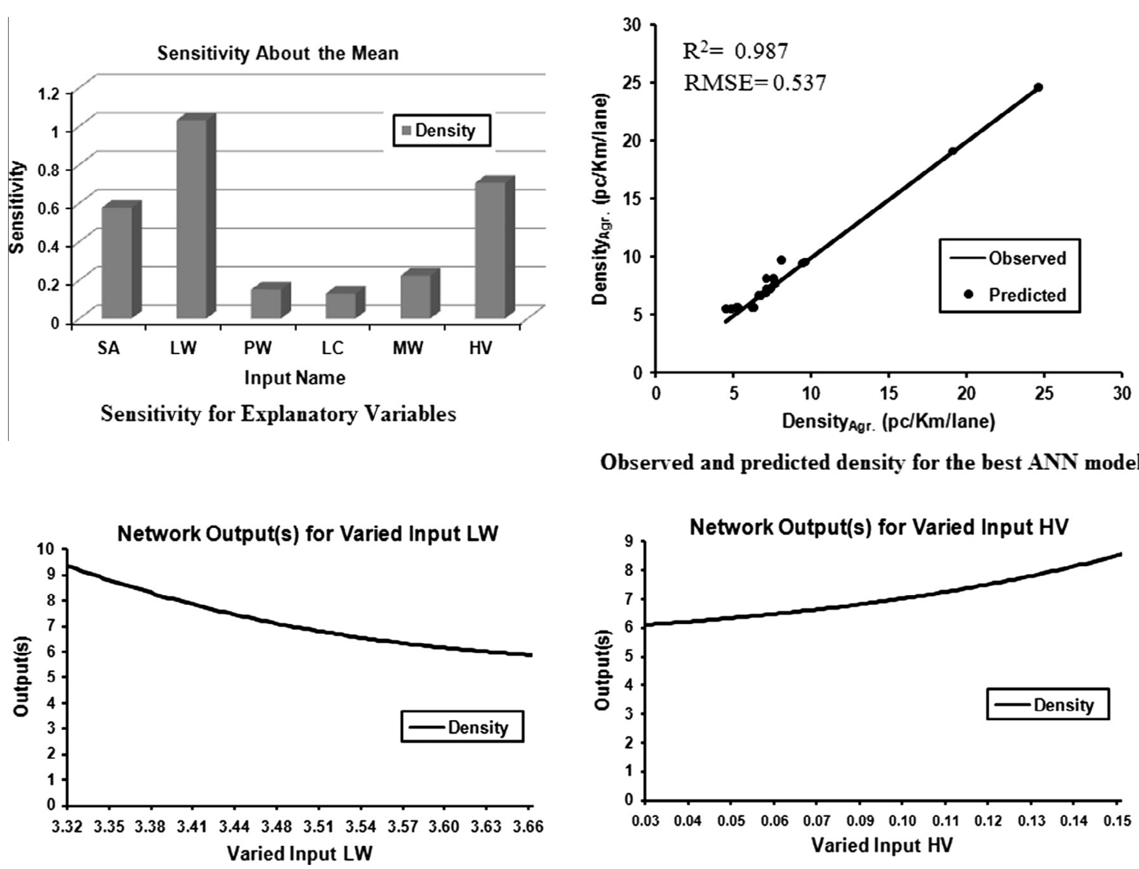

Figure 5 MLP results for the best density model (agricultural sections).

to merging of side traffic flow with main traffic flow. Consequently, capacity becomes worse than closed side sections. This result is consistent with logic.

- For both road models, the positive sign of the coefficient for LC means that the capacity increases with the increase in LC. Consequently, the wider LC improves the road capacity.

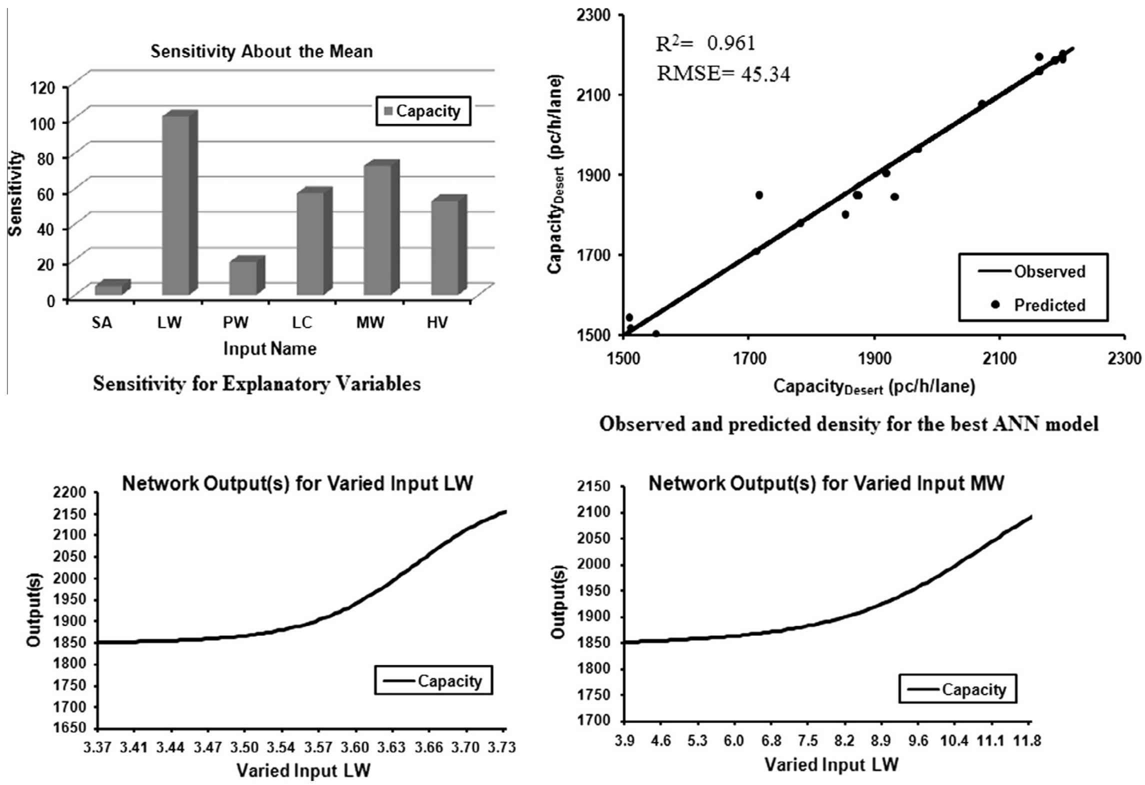

Figure 6 MLP results for the best capacity model (desert sections).

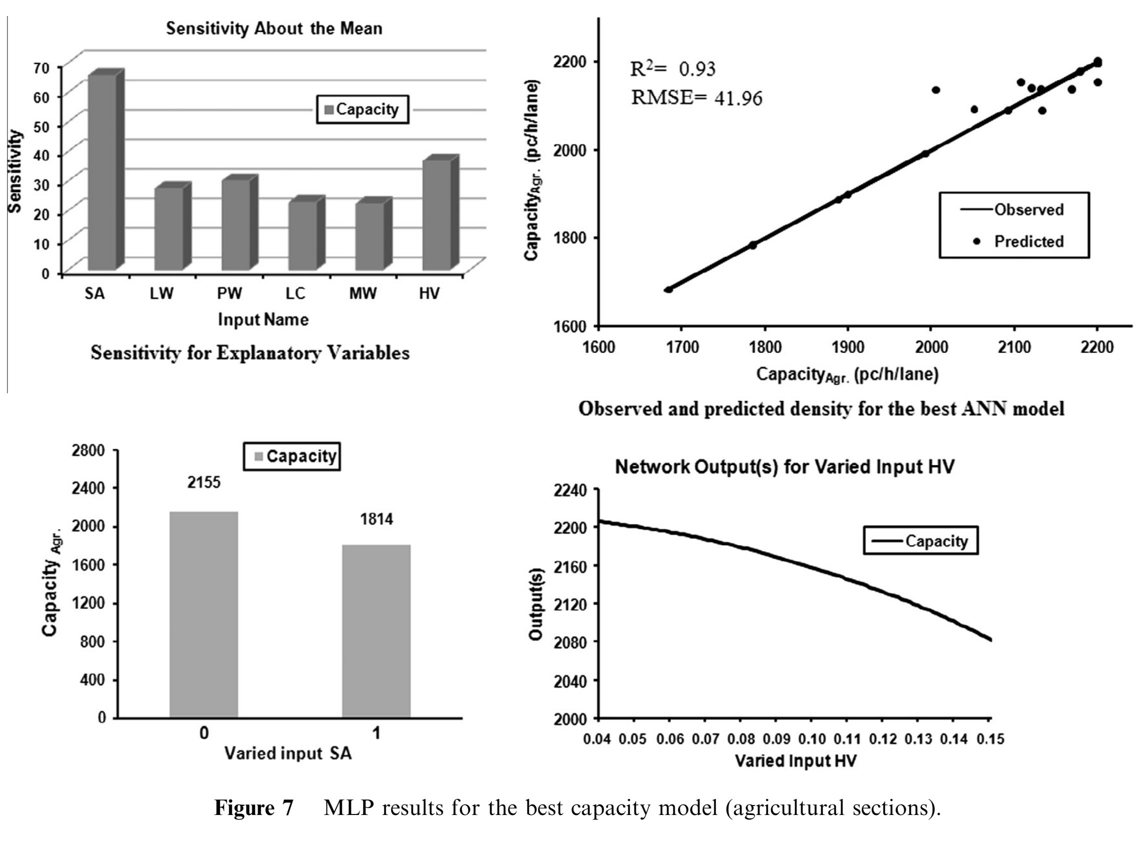

Figure 7 MLP results for the best capacity model (agricultural sections).

4.2. Results of ANN models

4.2.1. LOS model

As regression modeling, the ANN modeling is divided into three models. As a result of training and testing processing of all sections (45 sections), the performances of the best model for training (38 samples) and testing (7 samples) data set are presented in Table 5. Also, as a result of training and testing processing of desert sections (21 sections), the performances of the best model for training (17 samples) and testing (4 samples) data set are presented in Table 5. Finally, as a result of training and testing processing of agricultural sections (24 sections), the performances of the best model for training (19 samples) and testing (5 samples) data set are presented in Table 5. From the results in this table, it is found that the desert and agricultural models are better than the collective model as they give higher statistical performance. Then, it is better to separate the analysis for each road type.

4.2.1.1. Desert roads model (21 sections). The observed versus predicted values are shown in Fig. 4. Also, the same figure shows the sensitivity of each explanatory variable in the selected model. It is found that the most influential variable on density is HV, followed by PW. The relationships between each effective input variable and density are shown in Fig. 4.

4.2.1.2. Agricultural roads model (24 sections). The observed versus predicted values are shown in Fig. 5. Also, the same figure shows the sensitivity of each explanatory variable in the selected model. It is found that the most influential variable on density is LW, followed by HV. The relationships between each effective input variable and density are shown in Fig. 5.

Investigation of the two models shows that:

- For desert model, it is concluded that density decreases with the increase in PW. Consequently, the wider PW leads to LOS improvement.
- For agriculture model, it is found that density decreases with the increase in LW. Consequently, the wider LW leads to LOS improvement.
- For both models, the HV percentage leads to increase in Density. Then, the higher HV leads to LOS weakening. Finally, all the previous results are better than regression models in terms of model performance. Also, they are rational and consistent with logic.

4.2.2. Capacity model

As LOS models, the ANN modeling is divided into three models. The performances of the three models are presented in Table 5. From the results in this table, it is found that the desert and agricultural models are better than the collective model as they give higher performances. As before, the sections are divided in the analysis.

4.2.2.1. Desert roads model (17 sections). The observed versus predicted values are shown in Fig. 6. Also, the same figure shows the sensitivity of each explanatory variable in the selected model. It is found that the most influential variable on capacity is LW, followed by MW. The relationships between each effective input variable and capacity are shown in Fig. 6.

4.2.2.2. Agricultural roads model (14 sections). The observed versus predicted values are shown in Fig. 7. Also, the same figure shows the sensitivity of each explanatory variable in the selected model. It is found that the most influential variable on capacity is SA, followed by HV. The relationships between each effective input variable and capacity are shown in Fig. 7.

Investigation of the two models shows that:

- For desert model, it is concluded that capacity increases with the increase in both LW and MW. Consequently, the wider LW and MW lead to capacity improvement.
- For agriculture model, it is found that capacity decreases with the increase in HV. So, the higher HV leads to capacity reduction. In addition, the existence of SA leads to a considerable decrease in capacity. Although the average capacity at sites without SA is 2155 veh/h/ln, the average capacity at sites with SA is 1814 veh/h/ln.
- Finally, all the previous results are better than regression models in terms of models performance. Also, they are rational and consistent with logic.

5. Conclusions

The most important conclusions of the current paper are as follows: first, the ANN models give so better and most confidence results than regression models in terms of predicting density and capacity. Also, the desert and agricultural models are better than the collective model as they give higher statistical performance. Then, it is better to separate the analysis for each road type. For density model, the best ANN model for desert sections gives R^2 and RMSE equal to 0.992 and 0.437 for overall data set compared with the best regression model for the same sections gives R^2 and RMSE equal to 0.974 and 0.729 for all data set. For agricultural sections, the best ANN model gives R^2 and RMSE equal to 0.997 and 0.26 for overall data set compared with the best regression model for the same sections gives R^2 and RMSE equal to 0.991 and 0.427 for all data set. Also, for capacity model, the best ANN model for desert sections gives R^2 and RMSE equal to 0.961 and 45.34 for overall data set compared with the best regression model gives R^2 and RMSE equal to 0.873, and 82.82 for all data set. For agricultural sections, the best ANN model gives R^2 and RMSE equal to 0.93 and 41.96 for overall data set compared with the best regression model for the same sections gives R^2 and RMSE equal to 0.764 and 72.8 for all data set.

The second conclusion shows that the most influential variables on density of desert sections are HV, followed by PW. The increase in HV by 5% leads to an increase in density by nearly 2.3 veh/km/ln. Also, the increase in PW by 2 m leads to a decrease in density by nearly 1.5 veh/km/ln. For agricultural sections, the most influential variables on density are LW, followed by HV. The increase in LW by 20 cm leads to a decrease in density by nearly 2.5 veh/km/ln. In addition, the increase in HV by 7% leads to an increase in density by 1 veh/km/ln.

The third conclusion clarifies that the most influential variables on capacity of desert sections are LW, followed by MW. The increase in LW from 3.6 m to 3.7 m leads to an increase in capacity from 1940 to 2115 veh/h/ln. In addition, the increase in MW from 8 m to 10 m leads to an increase in capacity from 1900 to 1990 veh/h/ln. For agricultural sections, the most influential variables on density are SA, followed by HV. The existence of SA leads to a considerable decrease in capacity. Although the average capacity at sites without SA is 2155 veh/h/ln, the average capacity at sites with SA is 1814 veh/h/ln. In addition, the increase in HV by 10% leads to a decrease in capacity by 60 veh/h/ln.

The previous results are so important and useful for road authorities in Egypt as they can determine LOS and capacity for different tangent sections and improve the traffic performance of them in the future. Finally, future research should be conducted to extend all aspects of this research using comprehensive field data from various rural roads to increase number of sites to more than 100 sites in order to reach more accurate modeling and analysis of LOS and capacity. In addition, the use of curved and sloping sections in order to explore the impact of them on LOS and capacity for rural multi-lane highways in Egypt.

Acknowledgements

The author acknowledges the support of the General Authority of Roads, Bridges, and Land Transport, (GARBLT) for their assistance with the acquisition of traffic count data. Also, the author acknowledges Eng. Nasser Abdalla, Department of Civil Engineering, Faculty of Engineering, Al-Azhar University for his assistance with the acquisition of spot speed data at sites under research.

References

[1] Transportation Research Board (TRB), Highway Capacity Manual (HCM), fourth ed., Transportation Research Board, National Research Council, Washington, DC, 2000.
[2] B.S. Kerner, Three phase traffic theory and highway capacity, Physica A 333 (2004) 379–450.
[3] C.J. Hoban, Evaluating Traffic Capacity and Improvements to Road Geometry, World Bank, Washington, DC, 1987.
[4] M. Iwasaki, Empirical analysis of congested traffic flow characteristics and free speed affected by geometric factors on an intercity expressway, Transport Res. Rec. 1320 (1991) 242–250.
[5] A.T. Ibrahim, F.L. Hall, Effect of adverse weather conditions on speed-flow-occupancy relationships, Transport Res. Rec. 1457 (1994) 184–191.
[6] V. Shankar, F. Mannering, Modeling the endogeneity of lane-mean speeds and lane-speed deviations: a structural equations approach, Transport Res. A – Policy and Pract. 32 (1998) 311–322.
[7] K.L. Bang, Indonesian Highway Capacity Manual, Department of Public Works, Directorate General Highways, Jakarta, Indonesia, 1997.
[8] X. Yang, N. Zhang, The marginal decrease of lane capacity with the number of lanes on highway, EASTS 5 (2005) 739–749.
[9] S. Velmurugan, E. Madhu, K. Ravinder, K. Sitaramanjaneyulu, S. Gangopadhyay, Critical evaluation of roadway capacity of multi-lane high speed corridors under heterogeneous traffic conditions through traditional and microscopic simulation models, J. Indian Road Congr. 556 (2010) 235–264.
[10] V.T. Arasan, S.S. Arkatkar, Derivation of capacity standards for intercity roads carrying heterogeneous traffic using computer simulation, Procedia 16 (2011) 218–229.
[11] T. Sakai, K. Yamada-Kawai, H. Matsumoto, T. Uchida, New measure of the level of service for basic expressway segments incorporating customer satisfaction, Procedia 16 (2011) 57–68.
[12] A. García, A.J. Torres, M.A. Romero, A.T. Moreno, Traffic microsimulation study to evaluate the effect of type and spacing of traffic calming devices on capacity, Procedia 16 (2011) 270–281.
[13] N. M. Abdalla, Relationship between Traffic Flow Characteristics and Pavement Conditions on Roads in Egypt. M.Sc. Thesis, Dept. of Civil Eng., Faculty of Eng., Al-Azhar University, 2010.
[14] I.H. Hashim, T.A. Abdel-Wahed, Evaluation of performance measures for rural two-lane roads in Egypt, AEJ 50 (2011) 245–255.
[15] A.M. Semeida, Analysis and Evaluation of Road Safety in Egypt Using Conventional and Non-Conventional Modeling Techniques, PhD Thesis, Dept. of civil Eng., Faculty of Eng., Port Said University, 2011.
[16] A.S. Heikal, H.M. Abdel-latif, Principals of Traffic, Engineering, 2009, pp. 4–11, 4–12, ISBN-977-19-7061-5.
[17] General Authority of Roads, Bridges and Land Transport, "GARBLT", System of Traffic counting data, El-Nasr St., Cairo, Egypt, 2009.
[18] A.L. Simpson, J.B. Rauhut, P.R. Jordahl, E. Owusu-Antwi, M.I. Darter, R. Abroad, O.J. Pendleton, Y. Lee, Sensitivity Analyses for Selected Pavement Distresses, SHRP-P-393, 1994.
[19] D. Singh, M. Zaman, L. White, Modeling of 85th Percentile Speed for Rural Highways for Enhanced Traffic Safety. Report submitted to Oklahoma Department of Transportation, FHWA 2211, 2011.
[20] A.M.M.A. Semeida, Impact of Highway Geometry and Posted Speed on Operating Speed at Multilane Highways in Egypt, J. Adv. Res. (2012), <http://dx.doi.org/10.1016/j.jare.2012.08.014>.
[21] User's Manual "NeuroSolutions 7", NeuroDimension, Inc., Gainesville, 2010.
[22] A.V. Voudris, Analysis and Forecast of Capsize Bulk Carriers Market using Artificial Neural Networks, M-D Thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, June 2006.
[23] R.A. Tarefder, L. White, M. Zaman, Neural network model for asphalt concrete permeability, J. Mater. Civil Eng. 17 (2005) 19–27.