

# Derivation of level of service by artificial neural networks at horizontal curves: a case study in Egypt

Ahmed Mohamed Semeida

Received: 7 January 2014 / Accepted: 9 November 2014 / Published online: 27 January 2015  
© The Author(s) 2015. This article is published with open access at SpringerLink.com

## Abstract

**Purpose** Multi-lane highways represent the majority of the total length of highway network at many countries. The geometric design of such facilities and the traffic volume which includes heavy vehicles percentage (HV) are considered the most important factors affecting the level of service (LOS), especially on the horizontal curves.

**Methods** This paper aims to explore the relationship between the road geometric characteristics including horizontal curve properties and traffic volume including average annual daily traffic (AADT) and HV in one hand and LOS in the other hand by two methods. First is generalized linear modeling (GLM) procedure and second is the artificial neural networks (ANNs) procedure. In this research, the traffic and road geometric data are collected on 78 horizontal curves that are distributed on four multi-lane highways in Egypt. Two of them are located in desert area and the others in agricultural area.

**Results** The Analyses showed that the ANNs procedure give the best models for estimating LOS. Also, the most influential variables on LOS are AADT, HV, and radius of horizontal curve (R). Finally, the derived models have statistics within the acceptable regions and, also, they are conceptually reasonable.

**Conclusions** The previous inferences are so important for road authorities in Egypt as they can consider them as a first step for origination of Egyptian highway capacity manual.

**Keywords** Level of service · Highway geometry · Horizontal curves · Generalized linear modeling · Artificial neural networks

A. M. Semeida (✉)  
Civil Engineering Department, Specialization of Transportation and Traffic Engineering, Faculty of Engineering, Port Said University, Port-Said, Port Fouad, Postal No.: 42523, Egypt  
e-mail: semeida77@hotmail.com

## 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 heavy vehicles (HV) in traffic composition on LOS at multi-lane rural highways. Rural multi-lane highways are an important type of uninterrupted flow facilities in which there is no obstructions 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 traffic density (Density), in terms of passenger cars per kilometer per lane as the primary level of service (LOS) measure for multi-lane highways. Abdul-Mawjoud and Sofia [2] indicate that horizontal curves have long been recognized as having a significant effect on vehicle speeds. Therefore, it is necessary and important to be taken into consideration in the present analysis. Hence, this paper aims to evaluate LOS on horizontal curves at multi-lane highways by two modeling techniques. First is generalized linear modeling (GLM) technique and the second is artificial neural networks (ANNs) technique.

Field data on multi-lane highways in Egypt are used in this investigation. The analysis considers 78 horizontal curves that are distributed on four multi-lane highways. Two of them are located in desert area (Cairo-Alexandria and Cairo-Ismailia desert highways), and the other two are located in agricultural area (Cairo- Alexandria and Tanta-Damietta agricultural roads). Then, the paper includes two separate relevant analyses. The first analysis uses GLM procedure to investigate the relationships between Density as dependent variable, and horizontal curve properties, roadway factors, traffic volume, and HV as independent variables. The horizontal curve properties include radius of curve, deflection angle, and superelevation. The road factors for each curve are lane and pavement

width, lateral clearance, and number of lanes in each direction. In addition the traffic volume is expressed by average annual daily traffic (AADT). The second analysis uses ANNs procedure to explore the previous relationships and comparing the results. According to the objectives of this research, road authorities in Egypt can determine LOS for different horizontal curves on multi-lane highways and improve the traffic performance of them in the future.

## 2 Background studies

Several researches have been carried out to analyze the effect of road geometry and traffic composition on LOS for multi-lane highways. Kerner [3] confirmed that the determination of 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 LOS, traffic features and geometric elements on uninterrupted multi-lane highways [4–7]. The Indonesian highway capacity manual [8] mentioned that travel speed as the main measure of LOS of road segments.

Reinfurt et al. [9] tackled the speed changes which lead to LOS variation over the horizontal curves. They found that, as curves become sharper, there is a proportionally greater increase in speed reduction on a curve. The study findings support the debate of drivers cutting the curve short, which can result in run-off road crashes on the inside of the curve as well as head-on and opposite direction-sideswipe crashes with oncoming vehicles.

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 LOS, and subsequent it was concluded that highway performance 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.

Shawky and Hashim [12] studied the effect of horizontal alignment on traffic performance at two-lane rural highways. The follower density was used as a promising measure of traffic performance. The results show that the horizontal alignment characteristics have a significant effect on the follower density, especially curve radius value, by decreasing radius, the follower density increases (i.e., traffic performance decreases).

In Egypt, there were so few concerned with LOS due to lack of road geometric, horizontal curve, traffic flow and speed data. The most important research in this direction is published by Semeida [13]. The analysis in this paper used 45 different tangent sites from four main multi-lane rural

highways. Two modeling techniques were used. First was multiple linear regression and second was ANNs. Results showed that the ANN modeling gives the best models for estimating LOS. The most influential variables on LOS are lane width, HV, and existence of side access. This analysis was limited to tangent sections. Hence, the present paper extends and improves the analysis to include LOS on horizontal curves.

## 3 Study sites and field data

This research concerns with horizontal curves at rural multi-lane highways in Egypt. Therefore the analysis of this paper uses 78 horizontal curves (Sections) from four main multi-lane highways in Egypt. These roads include Cairo-Alexandria Agricultural Highway, Tanta- Damietta Agricultural Highway, Cairo-Alexandria Desert Highway, and Cairo-Ismailia Desert Highway. The collected data are divided into three types as road geometric characteristics including horizontal curves, vehicles speed, and traffic volume data including AADT and HV.

### 3.1 Road geometry data

This data presents the key independent variables in the analysis. Some of this data are collected directly from site investigation which includes lane, pavement width, lateral clearance, and number of lanes in each direction. On the other hand, the horizontal curves data are extracted from Abdalla [14] who worked with the survey team of General Authority of Roads, Bridges and Land Transport in Egypt (GARBLT) [15] and obtaining these data. The horizontal curve properties include radius of curve, deflection angle, and superelevation. All the previous variables, their symbols, and statistical analysis are provided in Table 1.

### 3.2 Vehicles speed data

The main type of the collected speed data is average travel speed of passenger cars ( $ATS_{pc}$ ). The passenger cars include taxis, vans, and jeeps. It is measured in field as follows.  $ATS_{pc}$  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 between 50 and 100 m throughout length of each horizontal curve 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 curve is not smaller than 100 vehicles [1]. The speed data is recorded in Table 2.

**Table 1** Statistical analysis and symbols of all collected variables

Variable	Variable Symbol	Max.	Min.	Avg.	S.D.
1- Lane Width in meters	LW	3.65	3	3.61	0.1
2- Pavement Width in one direction in meters	PW	14.6	6	8.36	1.82
3- Lateral Clearance in meters	LC	3.6	1.2	1.84	0.41
4- No of lanes in each direction in lanes	NL	4	2	–	–
5- Radius of curve in meters	R	780	40	402.5	231.7
6- Deflection Angle in Degree	DA	70.6	14.3	27.9	16.7
7- Superelevation %	e	12	2	4.7	3.5

3.3 Traffic volume data

The purpose of collecting traffic volume is to reach AADT, the design hourly volume (V) and detect 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. Average Annual Daily Traffic must be estimated at each site as Semeida [16]. As there are many sites (78 curves), 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) as Heikal and Abdel-latif [17]. 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 (GARBLT) [15]. The values of DF and SF for all roads under study are listed in Table 3. Therefore, the calculated AADT is given by Eq. (1) [16]. The values of AADT are listed in Table 4.

$$AADT = CountingVolume \times HF \times DF \times SF \tag{1}$$

Consequently, the design hourly volume (V) can be calculated by Eq. (2) (J).

$$V = AADT \times K \tag{2}$$

Where

V design hourly volume (typically, the 30th highest annual hourly volume)

**Table 2** Average travel speed data (ATS<sub>pc</sub>) for passenger cars

Curve No.	ATS <sub>pc</sub> (m)	Curve No.	ATS <sub>pc</sub> (m)	Curve No.	ATS <sub>pc</sub> (m)	Curve No.	ATS <sub>pc</sub> (m)
1	44.33	21	78.09	41	58.39	61	56.35
2	36.65	22	79.79	42	67.9	62	39.28
3	47.36	23	47.9	43	73.03	63	80.31
4	46.2	24	62.17	44	74.9	64	76.91
5	26.27	25	53.18	45	72.57	65	88.19
6	46.24	26	54.61	46	78.44	66	75.29
7	26.95	27	38.93	47	25.43	67	86.3
8	29.11	28	61.19	48	65.68	68	72.51
9	39.88	29	85.14	49	70.74	69	62.44
10	69.22	30	83.8	50	48.78	70	79.5
11	78.13	31	61.58	51	64.32	71	71.82
12	81.99	32	73.38	52	66.77	72	74.58
13	79.55	33	68.72	53	70.64	73	61.34
14	75.97	34	73.38	54	50.34	74	70.65
15	73.2	35	71.21	55	50.6	75	78.69
16	77.63	36	74.1	56	66.17	76	64.31
17	79.79	37	73.23	57	66.36	77	58.14
18	25.02	38	73.71	58	73.63	78	24.35
19	25.86	39	68.71	59	65.54		
20	78.59	40	74.45	60	60.17		

**Table 3** Values of DF and SF for roads under research

Factors		Cairo – Alexandria (agricultural highway)	Tanta – Damietta (agricultural highway)	Cairo – Alexandria (desert highway)	Cairo – Ismailia (desert highway)
Daily Factor (DF)	Saturday	1.05	1.03	1.00	1.01
	Sunday	0.96	0.98	1.01	1.03
	Monday	0.98	0.96	0.97	0.99
	Tuesday	0.93	0.94	0.99	0.98
	Wednesday	0.97	0.97	1.03	1.02
	Thursday	1.01	1.05	1.06	1.04
	Friday	1.1	1.13	0.98	0.93
Seasonal Factor (SF)	January	0.82	0.84	0.71	0.74
	February	0.94	0.96	0.65	0.68
	March	0.89	0.88	0.74	0.71
	April	0.91	0.93	0.87	0.84
	May	0.87	0.89	0.94	0.96
	June	1.11	1.08	1.29	1.25
	July	1.21	1.24	1.38	1.41
	August	1.29	1.32	1.61	1.58
	September	1.14	1.18	1.19	1.14
	October	1.08	1.06	1.03	1.01
	November	0.89	0.91	0.84	0.87
	December	0.85	0.82	0.79	0.81

AADT Average Annual Daily Traffic in vehicle per day and

K 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 quality of service analysis. To convert hourly volume to hourly directional volumes (Dir.V), the hourly volumes are multiplied by the D-factor as shown in Eq. (3) [1].

$$Dir.V = V \times D \tag{3}$$

Note that D=0.6 for rural roads as stated by HCM 2010 [1]. Dir. V values for all sites are recorded in Table 4. Finally, HV is extracted from the recorded manual counting as there is a classification of vehicles for passenger cars and HV. These values are indicated in Table 4.

### 3.4 Determination of LOS for horizontal curves under study

To determine LOS on each curve, Density is determined first from Eq. (4) [1].

$$Density = v_p / ATS_{pc}(pc/Km/lane) \tag{4}$$

Where

$v_p$  15-min passenger-car equivalent flow rate (pc/h/lane) and is computed from Eq. (5) [1].

$$v_p = Dir.V / (PHF \times N \times f_{HV} \times f_p) \tag{5}$$

Where

- Dir. V hourly volume in one direction (veh/h)
- PHF peak-hour factor (PHF=0.88 for rural roads)
- N number of lanes in the selected direction
- $f_{HV}$  heavy-vehicle adjustment factor (0.97–0.9) depending on HV value, and
- $f_p$  driver population factor (1) as the driver is familiar with the highway.

Consequently, LOS is determined for each curve from Density value which reflects the degree of vehicles congestion on each curve. The LOS and Density for all curves are determined and listed in Table 4.

Thus, the research uses a total number of 10 variables. These variables are divided into dependent and independent variables. Dependent Variable includes one variable. This variable is Density which presents the promising measure for LOS. Independent Variables include 9 Variables as follows,

1. Road Geometric=4 Variables (See Table 1),
2. Horizontal curve=3 Variables (See Table 1), and

**Table 4** Traffic count data, density, and LOS values for all horizontal curves

Curve No.	Traffic count in both direction per hour	AADT	Dir. V	HV (%)	Density	LOS
1	1593	18,060	1084	20.5	18.16	D
2	2298	26,053	1563	32.6	36.09	E
3	641	7267	436	12.5	4.14	A
4	663	7516	451	12.5	4.39	A
5	1200	13,604	816	20.5	15.39	C
6	768	13,341	800	20.5	12.23	C
7	1000	17,371	1042	21	27.5	E
8	1033	17,944	1077	20.5	17.43	D
9	692	12,021	721	13.8	7.87	B
10	3175	55,152	3309	20.5	22.52	E
11	3191	55,430	3326	20.5	19.95	D
12	3163	54,944	3297	20.5	19.02	D
13	3098	53,815	3229	20.5	19.6	D
14	3212	55,795	3348	20.5	20.52	D
15	3181	55,256	3315	20.5	21.3	D
16	3185	55,326	3320	20.5	20.08	D
17	3171	55,083	3305	20.5	19.54	D
18	1640	28,488	1709	22	32.74	E
19	1591	27,637	1658	29.4	33.29	E
20	1348	23,416	1405	12.7	7.67	B
21	3175	55,152	3309	20.5	19.97	D
22	3258	56,594	3396	20.5	19.54	D
23	1665	28,922	1735	20.5	25.6	E
24	1348	23,416	1405	18.4	10.39	B
25	1448	25,153	1509	18.4	12.15	C
26	1376	23,902	1434	18.4	11.83	C
27	1586	27,550	1653	29.3	33.05	E
28	3175	55,152	3309	33.2	43.79	E
29	852	15,902	954	12	3.76	A
30	864	16,126	968	12	3.87	A
31	1037	15,920	955	18.4	11.24	C
32	1056	16,212	973	18.4	9.61	B
33	1037	15,920	955	18.4	10.07	B
34	1101	16,903	1014	18.4	9.61	B
35	1037	15,920	955	18.4	9.72	B
36	1082	16,611	997	16.4	9.29	B
37	1037	15,920	955	18.4	9.45	B
38	1044	16,027	962	16.4	9.34	B
39	1037	15,920	955	18.4	10.08	B
40	1086	16,672	1000	14.8	9.07	B
41	1022	15,690	941	18.4	11.86	C
42	1051	16,135	968	18.4	10.38	B
43	1032	15,843	951	18.4	9.48	B
44	1065	16,350	981	13.8	8.9	B
45	1024	15,720	943	14	9.05	B
46	1063	16,319	979	13.8	8.5	B
47	1076	16,519	991	21	28.57	E
48	311	4774	286	5.7	2.73	A
49	297	4560	274	12	2.76	A

**Table 4** (continued)

Curve No.	Traffic count in both direction per hour	AADT	Dir. V	HV (%)	Density	LOS
50	316	4851	291	12.5	4.02	A
51	291	4467	268	12	3.04	A
52	316	4851	291	7.7	2.76	A
53	310	4759	286	12	2.76	A
54	323	4959	298	12	3.87	A
55	352	5404	324	12.5	3.88	A
56	318	4882	293	5.5	2.7	A
57	312	4790	287	6.5	2.74	A
58	325	4989	299	5	2.41	A
59	330	5066	304	12	2.98	A
60	328	5035	302	12	3.24	A
61	314	4821	289	12	3.47	A
62	802	12,312	739	20.5	13.97	C
63	806	12,374	742	12.7	6.25	A
64	817	12,543	753	12.7	6.62	A
65	806	12,374	742	12.7	5.69	A
66	817	12,543	753	12.7	6.76	A
67	637	9779	587	12.5	4.59	A
68	551	8459	508	12.7	4.74	A
68	637	9779	587	12.7	6.36	A
70	551	8459	508	12.5	4.31	A
71	637	9779	587	12.7	5.53	A
72	551	8459	508	12.7	4.61	A
73	817	12,543	753	13.8	8.41	B
74	823	12,635	758	12.7	7.2	B
75	811	12,450	747	12.7	6.47	A
76	806	12,374	742	12.7	7.81	B
77	817	12,543	753	18.4	9.38	B
78	795	12,205	732	20.5	22.36	E

3. Traffic volume data which includes 2 variables as AADT and HV (See Table 4).

independent variable. ANNs have another advantage in that the ANN modeling approach is fast and flexible. Finally, ANNs model can be easily use by engineers.

## 4 Methodology

The methodology of LOS prediction in the present research is divided into two main procedures: GLM procedure and ANNs procedure. It is concluded from the previous researches that the ANNs provide a better model is highlighted by better predictions for lower values, the normality of the residuals and their independence from the predicted variable. Several authors have reported greater performance of ANNs compared to GLM. The advantage of ANNs over GLM is that ANNs can directly take into account any non-linear relationships between the dependent variables and each

### 4.1 GLM procedure

This type of modeling was used successfully in accidents prediction all over the world. Semeida [16] used this technique and concluded important results in this context. In the present research, this model is used but in different form to fit density data. The mathematical form includes all independent variables separately. This is performed by taking the natural logarithm of all these variables to produce a power relationship between density and each of the former items. This allows studying the effect of all independent variables on Density separately. A normal distribution with a log link function is chosen to model these

data. This form takes the following shape in Eq. (6) as Varagouli et al. [18]:

$$Density = \exp(\beta_o) \times X_1^{\beta_1} \times \dots \times X_n^{\beta_n} \tag{6}$$

Where

- $X_n$  Explanatory variables from 1 to 9
- $\beta_o$  Regression constant; and
- $\beta_n$  Regression coefficient.

This form satisfies two main conditions:

- (1). This model must yield logical results (non negative). Also, at  $X_i=0$ ; Density must be zero.
- (2). The existing of logarithmic link function that can linearize this form for the purpose of coefficient estimation.

This model form is executed using the generalized linear model procedure PROC GENMOD in the SAS statistical software [19]. SAS user’s manual [20] applies the maximum log-likelihood technique to estimate the regression coefficients, standard errors, Wald Chi-squared statistics, p-values. In addition, the  $R^2$  (coefficient of determination) and (infinity

norm of error vector)  $\|\delta\|$  values are calculated for each model as Varagouli et al. [18]. Finally, the model with minimum  $\|\delta\|$  and highest  $R^2$  value is selected.

#### 4.2 ANNs procedure

In general, ANNs consist of 3 layers, namely, the input, the hidden and the output layers. In statistical terms, the input layer contains the independent variables and the output layer contains the dependent variables. ANNs typically start out with randomized weights for all their neurons. When a satisfactory level of performance is reached the training is ended and the network uses these weights to make a decision (Singh et al., 2011) [21].

The experience in this field is extracted from Semeida [13, 22, 23]. In his researches, 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) [24], which is also used in this work.

The overall dataset of 78 curve sections is divided into a training dataset and a testing dataset. As in the literature, the

**Table 5** Correlations between Density and independent variables

		D	NL	LW	PW	LC	AADT	R	DA	e	HV
Density	PC	1.000	-0.214	-0.253*	0.190	-0.331**	0.871**	-0.914**	0.887**	0.849**	0.907**
	Sig.		0.060	0.025	0.095	0.003	0.000	0.000	0.000	0.000	0.000
NL	PC		1	-0.285*	0.992**	0.519**	-0.24*	-0.305**	0.179	0.254*	0.195
	Sig.			0.011	0.000	0.000	0.034	0.007	0.116	0.025	0.088
LW	PC			1	-0.165	0.021	0.35**	0.368**	-0.256*	-0.187	-0.291**
	Sig.				0.149	0.855	0.002	0.001	0.024	0.1	0.01
PW	PC				1	0.538**	-0.204	-0.268*	0.151	0.24*	0.166
	Sig.					0.000	0.073	0.018	0.188	0.035	0.146
LC	PC					1	0.195	0.19	-0.343**	-0.31**	0.021
	Sig.						0.088	0.095	0.002	0.006	0.855
AADT	PC						1	0.189	-0.179	-0.204	-0.151
	Sig.							0.094	0.116	0.073	0.188
R	PC							1	-0.811**	-0.823**	0.21
	Sig.								0.000	0.000	0.065
DA	PC								1	0.72**	0.179
	Sig.									0.000	0.116
e	PC									1	0.19
	Sig.										0.095
HV	PC										1
	Sig.										

PC pearson correlation coefficient, Sig. sig. (2-tailed)

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

training data set varies from 70 % to 90 % and the testing data set varies from 10 % to 30 %. Model performances are and  $R^2$  for testing and training data set in one hand and for all data set in the other hand (Voudris, 2006) [25].

So many trials are done to reach the suitable percentage between training and testing data that gives the best performance for cars and trucks speed models. In addition, over fitting can be avoided by randomize the 78 curves 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) (Tarefder et al., 2005) [26].

### 5 Data analysis and results

#### 5.1 Correlation analysis

The correlations among Density on curve sections and the nine independent variables are analyzed. As shown in Table 5, Pearson correlation coefficient and the value of Sig. are calculated by SPSS. It can be seen from this table, there are

relatively significant correlations at 0.01 level among Density and six independent variables. These variables are LC, R, DA, e, AADT, and HV. Then, these variables are introduced into GLM and ANN models. Consequently, there should be no multicollinearity among the selected independent variables in the final models.

#### 5.2 Analysis and results of GLM procedure

There are three models that are statistically significant with Density after the application of Proc GENMOD in SAS software [19]; the best models are as follows in Eqs. (7), (8), and (9). Also, the graphical relations between predicted and observed data are shown in Fig. 1a, b, c, respectively. In addition, Table 6 presents the parameter estimates of the three models.

$$Density = \exp(3.86) \times AADT^{0.136} \times R^{-0.505} \tag{7}$$

(Whereas,  $R^2=0.877$ , and  $\delta=0.258$ )

$$Density = \exp(5.057) \times AADT^{0.086} \times R^{-0.311} \times HV^{0.999} \tag{8}$$

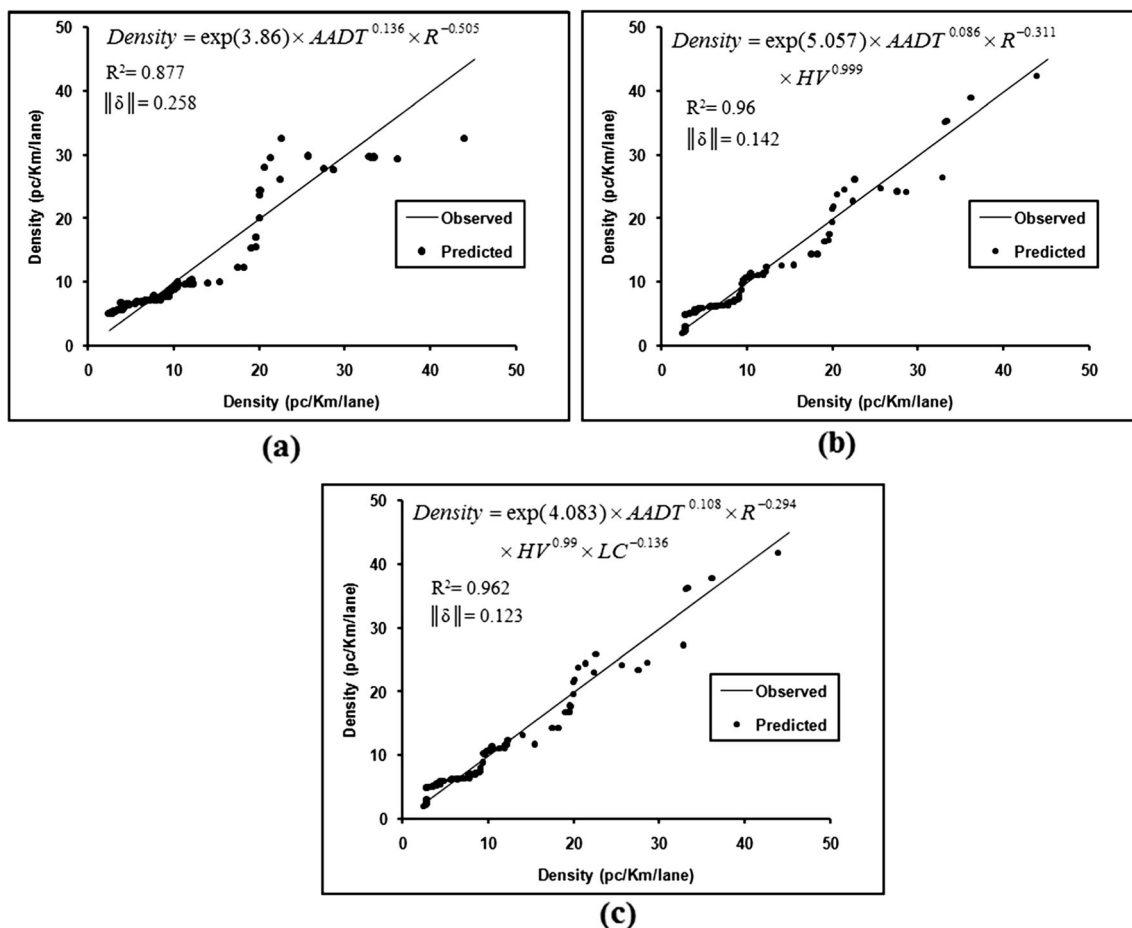


Fig. 1 Observed and predicted Density for the best GLM models



**Table 6** Model estimation results for the best GLM models

GLM Models									
Variable	GLM 1			GLM 2			GLM 3		
	Coeff.	Chi-square	p-value	Coeff.	Chi-square	p-value	Coeff.	Chi-square	p-value
Constant	3.86	531.6	0.0001	5.05	753.4	0.0001	4.8	544.4	0.0001
AADT	0.136	97.1	0.0001	0.086	32.26	0.0001	0.108	39.77	0.0001
R	-0.505	2847	0.0001	-0.311	572.1	0.0001	-0.294	421.3	0.0001
HV	-	-	-	0.999	538.1	0.0001	0.99	514.3	0.0001
LC	-	-	-	-	-	-	-0.136	8.04	0.0046
R <sup>2</sup>	0.877	0.96	0.962	-	-	-	-	-	-
δ	0.258	0.142	0.123	-	-	-	-	-	-

(Whereas, R<sup>2</sup>=0.96, and δ=0.142)

$$Density = \exp(4.083) \times AADT^{0.108} \times R^{-0.294} \times HV^{0.99} \times LC^{-0.136} \tag{9}$$

(Whereas, R<sup>2</sup>=0.962, and δ=0.123)

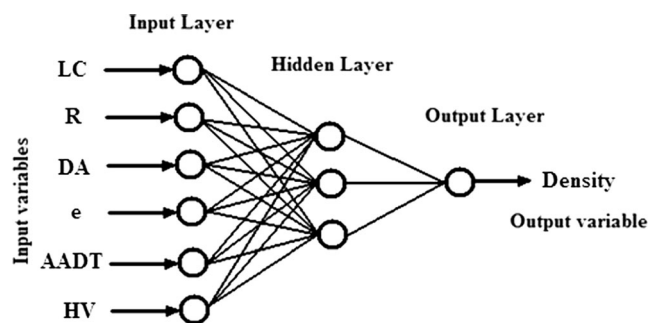
Investigation of the previous results shows that:

- The third model in Eq. (9) is better than the other three as it has higher R<sup>2</sup>, and lower ||δ||.
- In the best model, it is concluded that the negative sign of the coefficient for R means that Density decreases with the increase of R. The horizontal curves that have greater radii encourage the drivers to increase their speed on them than with smaller radii. These curves are safer for drivers. Thus, the greater R improves LOS on horizontal curves.
- In addition for the third model, the positive sign of the coefficient for AADT means that Density increases with the increase of traffic volume. The increase of traffic volume makes the road section more crowded with vehicles and consequently, the number of vehicles per km clearly increases. This implies to the decrease of LOS. Also, this result is rational.
- In the same model, the positive sign of the coefficient for HV means that Density increases with the increase of HV. The horizontal curves with higher HV are more congestive and the drivers of passenger cars are annoyed with HV which force them to decrease their speed. This implies to the decrease of LOS. Also, this result is rational.
- In the best model, it is found that the negative sign of the coefficient for LC means that Density moderately decreases with the increase of LC. In other word, the wider lateral clearance encourages the drivers to increase their speed on horizontal curve as the maneuvering is easier at wider road cross section and then Density decreases.

Consequently, the wider LC slightly improves LOS on horizontal curves. This result is consistent with logic.

### 5.3 Analysis and results of ANNs procedure

As a result of correlation analysis, there are six independent variables are highly correlated with Density. These variables are in input layer. One hidden layer is used, and one desired variable (Density) is in output layer with 78 observations are used. The architecture of the ANN model is shown in Fig. 2. The number of neurons in hidden layer is about half of the total number of neurons at the input and output layers (three neurons), which is set based on generally accepted knowledge in this field. Using of learning rule of (momentum) and the suitable number of epochs (iterations) is 5000. The previous conditions are suitable for quick convergence of the problem as executed by Semeida [13, 22, 23]. So many trials are done to reach the percentage between training and testing data which gives the best model performance in the present case of research. The performances of the best three trials for training and testing data set are presented in Table 7. This table shows that trial two is the best one. In this trial, the curves are divided into training data set that has 63 curves



**Fig. 2** MLP network architecture of Density model

**Table 7** Performances for the best three trials of ANNs model

Trial number	Performance	Number of Training samples	Number of Testing samples	Overall Model
1	$R^2$	68	10	0.928
	$\ \delta\ $	0.931	0.894	
2	$R^2$	63	15	0.993
	$\ \delta\ $	0.134	0.211	
3	$R^2$	58	20	0.879
	$\ \delta\ $	0.996	0.863	
				0.205

(80 % of all observations), and testing data set that has 15 curves (20 % of all sites).

As a result of training and testing processing, the observed versus predicted values are shown in Fig. 3. It is clearly that the ANN models give better and most confident results than GLM models. In order to measure the importance of each explanatory variable, general influence (sensitivity about the mean or standard deviation) is computed based on the trained weights of ANNs. For specified independent variable, if this value (sensitivity about the mean) is higher than other variables. This indicates that the effect of this variable on dependent variable (Density) is higher than other variables. Also, Fig. 3 shows the sensitivity of each explanatory variable in the selected model. It is found that the most influential variable on Density is R, followed by HV and AADT. The relationships between each effective input variable and Density are shown in Fig. 4. It is concluded the following results:

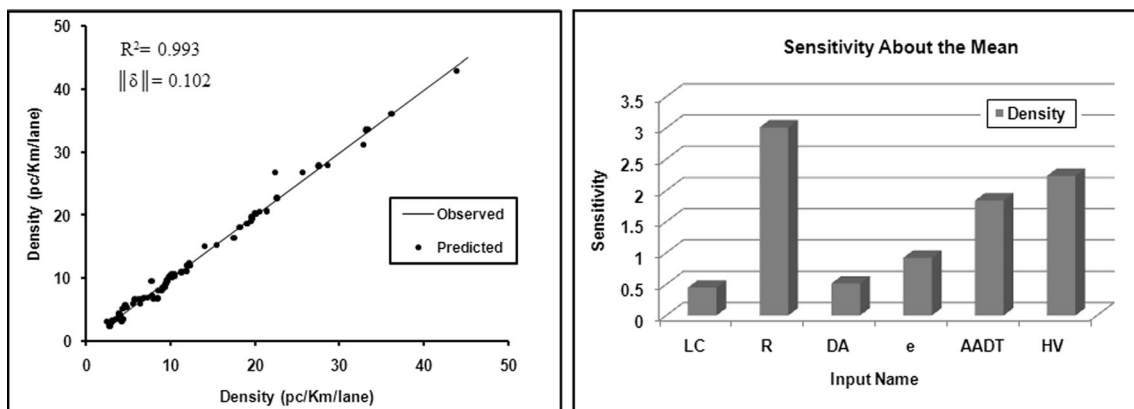
Density decreases with the increase of R. The increase of R from 134 m to 617 m leads to a decrease of Density from 15.7 pc/km/lane to 5.6 pc/km/lane. In other words,

the LOS improves from C to A. The relation between the two variables is nearly linear and clearly inverse.

Density increases with the increase of HV. The increase of HV from 16 % to 22 % leads to an increase of Density from 9.8 pc/km/lane to 18 pc/km/lane. This implies that LOS deteriorates from B to D. Moreover, Density is nearly constant with HV less than 16 %. This implies that HV effect on traffic moving starts at percentage more than 16 %. Less than this value, this effect is feeble.

Density increases with the increase of AADT. The increase of AADT from 8000 veh. /h to 38,000 veh. /h leads to an increase of Density from 7.7 pc/km/lane to 12.8 pc/km/lane. This indicates that LOS retreats from B to C. The relation between the two variables in this zone is linear. Density is nearly constant with AADT less than 8000 veh./ h. This proves that AADT is not effective on traffic flow less than this value.

Although the effect of the other variables (LC, DA, and e) on Density is limited, but it is discussed for more knowledge. Density decreases with the increase of LC. The increase of LC from 1.4 m to 2.3 m leads to a decrease of Density from 10.5 pc/km/lane to 9 pc/km/lane. This implies that no change in LOS B. In addition, the same



**Fig. 3** Measured and predicted Density and sensitivity analysis for the best ANN Model

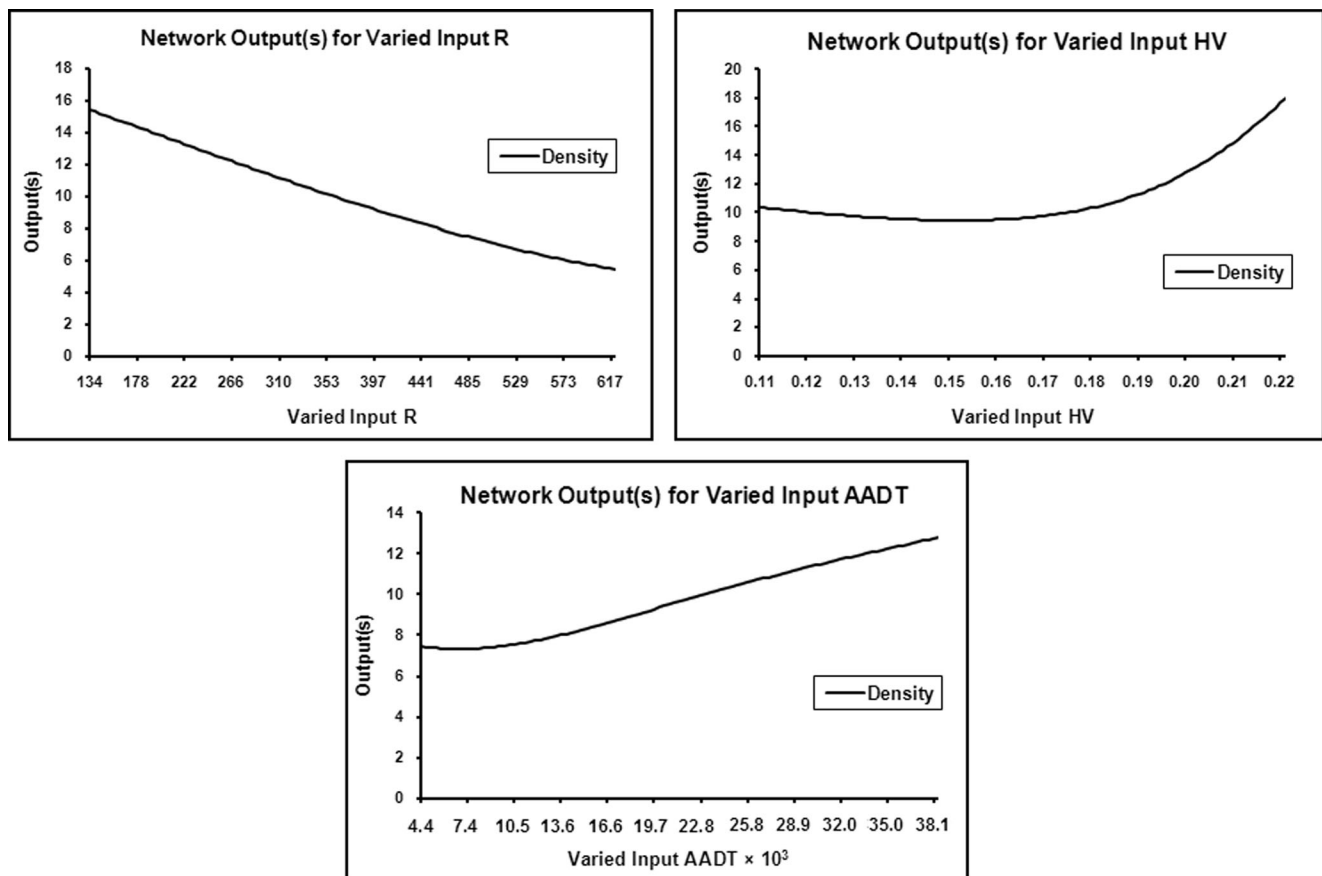


Fig. 4 Relations between the effective explanatory variables and Density for the best ANNs model

conclusion is resulted for DA and e as the same LOS B is existed with variable change. Therefore, the effect of these variables on Density is weak and can be neglected.

In addition, Density increases with the increase of both HV and AADT. These results are more accurate than the regression models and rational.

### 6 Conclusions

The current paper explores the impact of highway geometry, horizontal alignment, and traffic characteristics on density at multilane highways in Egypt. The most important findings of this paper are:

1. The ANNs procedure gives so better, more confident and consistent with logic results than GLM procedure in terms of predicting Density.
2. The best ANNs model gives  $R^2$  and  $\delta$  equal to 0.993 and 0.102, respectively compared with the best GLM model for the same horizontal curves gives  $R^2$  and  $\|\delta\|$  equal to 0.962, and 0.123, respectively.

3. For ANN model: The most influential variable on Density is R, followed by HV and AADT.
4. Also, the increase of R from 134 m to 617 m leads to a decrease of Density from 15.7 pc/km/lane to 5.6 pc/km/lane. In other words, the LOS improves form C to A.
5. In addition, the increase of HV from 16 % to 22 % leads to an increase of Density from 9.8 pc/km/lane to 18 pc/km/lane. This implies that LOS deteriorates from B to D.
6. Finally, the increase of AADT from 8000 veh. /h to 38,000 veh. /h leads to an increase of Density from 7.7 pc/km/lane to 12.8 pc/km/lane. This indicates that LOS retreats from B to C.

These results are so important for road authorities in Egypt as they can determine LOS for different horizontal curves and improve the traffic performance of them in the future. Also, they can consider these results as a first step for origination of Egyptian highway capacity manual. Finally, future research should be conducted to extend all aspects of this research using comprehensive field data including more multi-lane rural roads and vertical alignment. In addition, capacity for multi-lane highways is the main target in the Egyptian future research.

**Acknowledgments** 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 horizontal curve properties at sites under research.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

1. Transportation Research Board (TRB) (2010) Highway capacity manual (HCM), 1st edn. Transportation Research Board, National Research Council, Washington
2. Abdul-Mawjoud A, Sofia G (2008) Development of models for predicting speed on horizontal curves for Two-lane rural highways. *Arab J Sci Eng* 33(2):365–377
3. Kemer B (2004) Three phase traffic theory and highway level of service. *J Phys A* 333:379–450
4. Hoban C (1987) Evaluating traffic capacity and improvements to road geometry. World Bank, Washington
5. Iwasaki M (1991) Empirical analysis of congested traffic flow characteristics and free speed affected by geometric factors on an intercity expressway. *Transp RES REC* 1320:242–250
6. Ibrahim A, Hall F (1994) Effect of adverse weather conditions on speed-flow- occupancy relationships. *Transp RES REC* 1457:184–191
7. Shankar V, Mannering F (1998) Modeling the endogeneity of lane-mean speeds and lane- speed deviations: a structural equations approach. *J Transp RES A* 32:311–322
8. Manual IHC (1997) Department of public works. Directorate General Highways, Jakarta
9. Reinfurt W, Zegeer C, Shelton B, Neuman T (1991) Analysis of vehicle operations on horizontal curves. *Transp RES REC* 1318:43–50
10. Arasan V, Arkatkar S (2011) Derivation of capacity standards for intercity roads carrying heterogeneous traffic using computer simulation. *J Procedia Sci* 16:218–229
11. Sakai T, Yamada-Kawai K, Matsumoto H, Uchida T (2011) New measure of the level of service for basic expressway segments incorporating customer satisfaction. *J Procedia Sci* 16:57–68
12. Shawky M, Hashim I (2010) Impact of horizontal alignment on traffic performance at rural two-lane highways, In: Proceedings of 4th International Symposium on Highway Geometric Design, Valencia, Spain, June 2–5
13. Semeida AM (2013) New models to evaluate the level of service and capacity for rural multi-lane highways in Egypt. *A Eng J* 52(3):455–466. doi:10.1016/j.aej.2013.04.003
14. Abdalla N (2010) Relationship between traffic flow characteristics and pavement conditions on roads in Egypt. M.Sc. Dissertation, Department of civil Engineering, Faculty of Engineering, Al-Azhar University
15. General Authority of Roads, Bridges and Land Transport (GARBLT) (2009) System of traffic counting data, El-Nasr St., Cairo, Egypt. <http://www.garblt.gov.eg/>. Accessed 12 Feb 2009
16. Semeida A. Analysis and Evaluation of Road Safety in Egypt Using Conventional and Non-Conventional Modeling Techniques (2011) Ph.D. Dissertation, Department of civil Engineering, Faculty of Engineering, Port Said University
17. Heikal A, Abdel-latif H (2009) Principals of traffic engineering. Alhakim press, Inc., (ISBN-977-19-7061-5), Cairo
18. Varagouli A, Simos T, Xeidakis G (2005) Fitting a multiple regression line to travel demand forecasting: the case of the prefecture of xanthi, Northern Greece. *J Math Comp Model* 42:817–836. doi:10.1016/j.mcm.2005.09.010
19. SAS Institute Inc (2003) Version 9 of the SAS system for windows. SAS, Cary
20. SAS User's Manual (2003) (Chapter 29): the GENMOD procedure. SAS, Cary, pp 1365–1462
21. Singh D, Zaman M, White L (2011) Modeling of 85th percentile speed for rural highways for enhanced traffic safety. No. FHWA 2211, Oklahoma Department of Transportation
22. Semeida AM (2013) Impact of highway geometry and posted speed on operating speed at multilane highways in Egypt. *J Adv Res* 4(6): 515–523. doi:10.1016/j.jare.2012.08.014
23. Semeida AM (2014) Application of artificial neural networks for operating speed prediction at horizontal curves: a case study in Egypt. *J Mod Transp* 22(1):20–29. doi:10.1007/s40534-014-0033-3
24. User's Manual "NeuroSolutions 7" (2010) NeuroDimension, Inc. Gainesville
25. Voudris AV (2006) Analysis and forecast of capsized bulk carriers market using artificial neural networks. M.Sc. Dissertation, Massachusetts Institute of Technology, USA
26. Tarefder RA, White L, Zaman M (2005) Neural network model for asphalt concrete permeability. *M Civ Eng J* 17:19–27