Palacký University Olomouc

**Transactions on Transport Sciences**
Peer-Reviewed Open Access Journal
Vol. 2/2024 DOI: 10.5507/tots.2024.001
journal homepage: www.tots.upol.cz

# Estimating Operating Speed on Highways using Multiple Linear Regression and Artificial Neural Network Technique

**KIRAN KUMAR TOTTADI[a*], ARPAN MEHAR[b]**

*a. Research Scholar, Department of Civil Engineering, National Institute of Technology, Warangal, India.*
*b. Associate professor, Department of Civil Engineering, National Institute of Technology, Warangal, India.*

**ABSTRACT:** The speed variation along the successive highway sections is one of the most important factors in assessing geometric design consistency. Therefore, it is necessary to predict operating speed on important highway geometric features involving major safety issues. The present study aims to develop operating speed on the curve and tangent section of four-lane divided highways. For this research, the data is collected on 44 highway sections in India, including 22 curved and 22 tangent sections. The geometric features and free-flow speeds of various vehicles were collected. This study analyzed the speed profiles of vehicles that follow different statistical distribution patterns other than normal distribution. Multiple linear regression (MLR) and Artificial Neural Network (ANN) techniques are adopted to develop operating speed models on curve and tangent sections. The variables like curve radius, curve length and deflection angle are identified as most significant for modelling operating speed on the curves. Similarly, the shoulder width, median width, and access density are found to influence the operating speed on tangent sections. The developed models are successfully validated with field data. The performance measures such as RMSE and MAPE are applied to check degree of accuracy of developed models. The results revealed that the ANN models perform better than MLR models in curved and tangent sections. The developed models are helpful for highway and traffic engineers in establishing posted speed limits on critical sections of highways.

**KEYWORDS:** Design consistency; MLR, ANN, RMSE, MAPE

## 1. INTRODUCTION

India is a vast country with the world's second-largest road network that spans about 6.33 million kilometres (MoRTH, 2022). Road network accounts for over 90 % of passenger traffic and 64.5 % of freight traffic in India. Recent years witnessed tremendous changes in the road infrastructure and vehicular technology, resulting in great vehicle speed changes. The traffic composition dramatically influences the speed variation, as mixed traffic conditions prevail in India. Hence, speed is one of the important factors for road users, highway designers and planners. It is used as a parameter to determine capacity and LOS, accident analysis, traffic calming measures, signal coordination, etc. One major application is setting up speed limits to regulate and propagate the vehicular speeds to reduce the speed difference among the vehicles and reduce severe accidents due to overspeeding. Speed was also used to assess the highway geometric design consistency. As per the MoRTH (2022) report, 4,61,312 accidents occurred due to various reasons. Among them, 1,19,904 deaths occurred on roads due to overspeeding. It indicates that deaths occur in about 26 % of deaths of speeding. By the way, overspeeding constitutes the primary violation of accidents, accident-related deaths, and injuries.

Indian highways have been constructed using design speed as a design criterion. This design speed concept has been applied inflexibly and occasionally without considering adequate geometrical or traffic flow elements. A vehicle can move faster than the design speed for which it was designed when operating on a highway. Due to incorrect speed control estimation and inconsistent design of complicated alignments, such as vertical and horizontal curves, drivers may be in dangerous conditions. As a result, the operational speed idea was applied to determine the maximum speed on roadways. The operating speed is also called 85$^{th}$ percentile speed, defined as "the speed at which vehicles move under free flow conditions". The operating speed depends on the geometric features of highway sections and traffic parameters. The operating speed on the curve and tangent sections are determined based on the fitted cumulative distribution profile (Tottadi & Mehar, 2021). Mostly, the vehicular speed profiles follow the normal distribution under homogeneous conditions or free-flow conditions, whereas in mixed traffic flow conditions, the vehicles might follow other than normal distribution like Beta, log normal, Gamma, Weibull, etc (Tottadi, Varma, & Mehar, 2022). This phenomenon is due to vehicle speed variation in mixed flow conditions. Hence, examining the best-fitted distribution for different speed profiles is necessary.

Tremendous changes in road infrastructure and vehicular technologies resulted in huge variations in vehicular speeds in India. Hence, updated speed studies are a primary need that needs to be carried out to benefit highway designers, planners, engineers, etc. Recently, a few speed studies have been carried out in India by various research institutes, and they were already developing speed prediction models. However, the developed speed equations consider each speed-affecting parameter individually. In India, many studies mainly focus on developing speed models on curves. In the present study, an attempt was made to develop speed prediction models for curved and tangent sections on multi-lane highways. This study is mainly concerned with developing generalized speed prediction models considering distribution type and various parameters affecting speed, such as highway geometry and traffic characteristics.

* *Corresponding author: e-mail: kirant6666@student.nitw.ac.in*

## 2. LITERATURE REVIEW

Glennon, Neuman, & Leisch (1986) studied the safety and operational characteristics of curves of two-lane rural highways. They used four different methodologies: multivariate accident analysis, vehicle behaviour studies, simulation of vehicle/driver operations using the Highway Vehicle Object Simulation Model (HVOSM), and analytical studies. The authors concluded that the radius of the curve, curve length and superelevation significantly affect the operating speed. (Lamm, Choueiri, & Hayward, 1988; Lamm, Psarianos, Drymalitou, & Soilemezglou, 1995; Lamm, Psarianos, & Mailaender, 1999) established an association between $V_{85}$ speed and geometric characteristics of the curve. The authors found curve radius and were significant variables. Morrall & Talarico (1994) developed a relationship between operating speed and geometric parameters on two-lane rural highways. The author concluded that the degree of curvature is the only significant variable that affects the $V_{85}$. Islam & Seneviratne (1994) also studied operating speed on curves. The authors found that the curve radius highly influences the operating speed. Krammes (1995) compared the 85$^{th}$ percentile speed and design speed. It revealed that the 85th percentile speed exceeded the design speed on most curves in each 10 kmph design-speed increment less than or equal to 100 kmph. The operating speed exceeded the design speed of less than 90 kmph. This study shows there was a disparity between operating speed and design speed.

Andueza (2000) developed mathematical models to estimate the operating speeds in mountain roads. Spot speeds were measured at the middle of the curves and tangents with a radar gun. He reported that curve radius was highly influenced by curves, and the previous curve radius had a significant impact on tangents. Schurr, McCoy, Pesti, & Huff (2002) established a relationship between design, operating and posted speeds on curves of two-lane rural highways. They collected passenger cars speeds on tangent and middle of the curve sections. Stepwise multiple linear regression was adopted for modelling. Authors suggested that the radius of the curve greater than or equal to 360 m, the drivers tend not to significantly increase or decrease their speeds when travelling from tangent to curve. Polus, Fitzpatrick, & Fambro (2000) analysed operating speed variability on tangent sections of rural two-lane highways. Results show that the tangent length and radius of the before and after tangent significantly influenced operating speed. Misaghi & Hassan (2005) researched to estimate operating speed and speed differential models for two-lane, rural highways. Curve radius was found to be the most significant parameter. Qureshi, Khakheli, & Memon (2005) developed operating speed models using continuous speed profiles collected with high-end GPS. They found that the deflection angle and the length of the curve are significant variables for predicting operating speed. Memon, Khaskheli, & Qureshi (2008) analyzed the data of speed profile obtained from the vehicle equipped GPS based device called V-Box. Sowmya, Ravishankar, & Anjaneyulu (2012) modeled Operating Speed and Speed Differential on Intermediate Lane Rural Roads. The authors performed stepwise linear regression using SPSS software between the speed and the significant geometric variables, and the curve radius was most influencing factor. Jacob and Anjaneyulu (2013) developed models for predicting operating speed at tangent and mid-curve sections of horizontal alignment of two-lane rural highways and models for estimating the speed reduction from tangent to curve. They revealed that approach tangent length and curve radius influenced the curves.

Morris & Donnell (2014) modelled $V_{85}$ of cars and trucks on multi-lane highways with a combination of curves and steep grades. The high right shoulder width encourages more passenger cars speeds, while trucks don't affect or are not associated. Vertical grades should have more impact on truck speeds. Sil, Nama, Maji, & Maurya (2019b) evaluated geometric design consistency on tangent-curve sections of a four-lane divided highway. They found the radius and preceding tangent length to significantly affect the operating speed at the mid of the curve. The radius of curve is 360 m or less, affecting the 85$^{th}$ percentile speed of passenger cars. Maji, Sil, & Tyagi (2018) modelled $V_{85}$ and $V_{98}$ of cars, LCVs and heavy commercial vehicles on rural four-lane divided highways. They found that the speed on preceding section highly influenced the speeds of mid of the curve, and the curve length and deflection angle parameters significantly influenced car speeds. Maji and Tyagi (2018) estimated the operating speeds of cars and sports utility vehicles on curves of four-lane divided highways. They found that the curve length had a significant impact up to mid of the curve from 50 m from the beginning of the curve. The radius of the curve had the most significant effect on mid of the curve. Sil, Nama, Maji, & Maurya (2020) developed $V_{85}$ models under free-flow conditions. The backward elimination technique was used to develop models. The geometric parameters of curve length, deflection angle and preceding tangent length were found to be significant on operating speed. Sil, Maji, Nama, & Maurya (2019a) studied vehicle speed distribution on four-lane divided highway. They found that the speed distribution follows a similar trend when the radius beyond 400 m and the degree of curvature had significant impact on 85$^{th}$ percentile speed. Summarizing the literature, the majority of studies were developed models using MLR techniques, very few studies were adopted ANN technique for predicting operating speed by considering various geometric and traffic parameters. Hence, the present study used both MLR and ANN technique for finding the better prediction model.

**Objectives**

- To analyze statistical distribution profiles of vehicles under free flow conditions for determining operating speed on curved and tangent sections of four-lane divided highways.
- To develop operating speed models as a function of various geometric parameters such as curve radius, curve length, deflection angle, degree of curvature, shoulder and median width, and access density for accurate estimation of 85$^{th}$ percentile speed.

## 3. SITE SELECTION

The present study collected data from different National highways (NHs) spread across the length and breadth of the country. The study area selection was done in such a way that it contained both tangent and curved sections. Field data collection and analysis were carried out based on the utility and importance of the highways. The parameters such as road characteristics, vehicle characteristics and traffic characteristics were defined. A total of 44 sites were selected, out of which 22 were tangent sections, and 22 were curved sections. Figures 1 (a) and (b) show two study locations of selected sites. Geometric parameters, namely section type, number of lanes, road width, median width, shoulder type and width, etc., were measured and later entered in Excel. The access control type and access point density were also collected along the highway and later verified using Google Maps.

## 4. DATA COLLECTION

Free flow speed data was collected at all the sections using radar gun where special care was taken to collect data so that the process would not distract drivers. Free-flow speeds of

Figure 1(a). Study location of curve section on NH-7

Figure 1(b). Study location of tangent section on NH-27

various vehicles were collected at mid-point of the curve and tangent (straight) section using a radar guns. The vehicle types, namely, car, 2W, LCV, and HCV, are in these sections, and among them, 85% of cars and the remaining 15% of other vehicles. Surveyors collecting free speeds is shown in Figure 2. Geometric data such as curve radius and curve length were obtained by uploading the Google Earth image into ArcGIS software and rechecking it using AutoCAD software. Three different points on each curve section were identified using AutoCAD software. The curve radius ranges from 180 m to 525 m, and the curve length ranges from 170 m to 870 m. Similarly, the degree of curvature and deflection angle range from $10.91^0$ to $31.83^0$ and $32^0$ to $106^0$, respectively.

Figure 2. Collecting free speeds using Radar Gun

## 5. SPEED DISTRIBUTION ANALYSIS

The speed of a vehicle is a random variable, and it follows some statistical distribution. To analyse speed distribution, the speed data observed for different vehicle types on the highway are arranged to estimate the frequencies at estimated class intervals. The speed frequency observed at a defined class interval was plotted against the speed. The profiles were plotted for each vehicle type, and the combined speed profile was used for mixed traffic. Figures 3 and 4 provides the distribution curves for curve and tangent section, respectively at one location. Usually, the speed profile follows normal distribution under uniform traffic conditions. However, under mixed traffic conditions, different vehicle types operate at different speeds with no lane discipline. Hence, it is necessary to analyse other related distributions to fit the observed speed data. The other distributions include log-normal, beta (with 4 parameters), Weibull (with 3 parameters), Gamma, Generalized Extreme Value (GEV) etc.

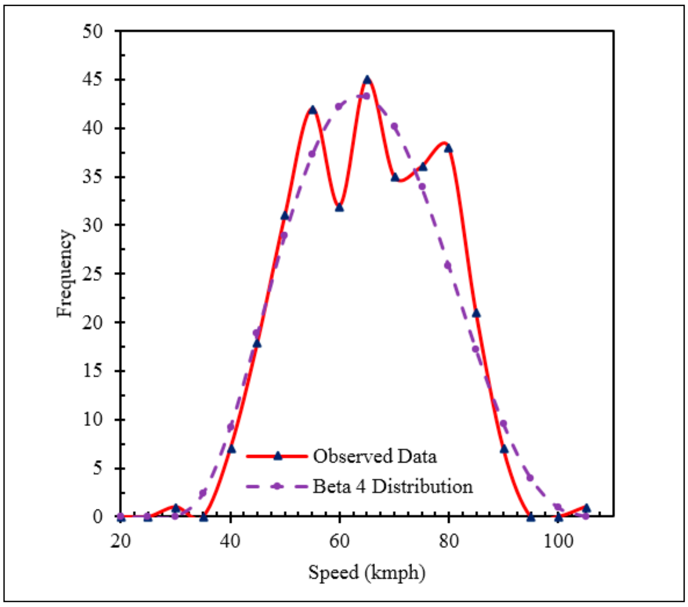


Figure 3. Speed profile of mixed traffic on curved section

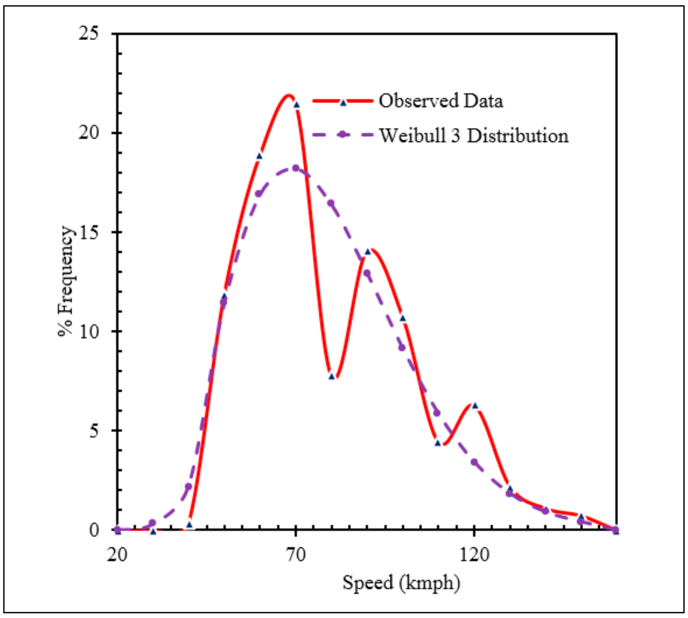


Figure 4. Speed profile of mixed traffic on tangent section

The speed distribution analysis confirms that the vehicles followed continuous distributions other than normal distribution. Statistical tests such as K-S was performed to check whether the data fits to selected distribution. It is based on the maximum vertical distance between the observed and selected cumulative distribution functions (CDFs). The statistical test result showed that the null hypothesis is rejected at a 95% confidence level, as Kolmogorov-Smirnov statistic surpasses the critical value of 0.20517. The best-fit distribution is selected based on maximum *p*-value and K-S test at 95% confidence interval. The results of the best-fit distribution for all sites on curved and tangent sections are depicted in Tables 1 and 2. It is inferred that the speed profiles do not necessarily follow normal distribution where traffic is mixed with different vehicle types. Moreover, other distributions such as log-normal, Weibull, and beta4 distributions also found as best fit distribution in the analysis. It may be inferred that the free speed profiles predominantly followed Beta 4-parameter distribution (α, β, min, max) in most highway sections. The parameters α and β of the Beta distribution are estimated to be in the range, and upper and lower boundary parameters to establish the distribution's form.

| Section | *p*-value estimated for Distribution | | | | | | Best fit distribution |
|---|---|---|---|---|---|---|---|
| | Normal | Log normal | Beta4 | Gamma | Weibull3 | GEV | |
| 1 | **0.9862** | 0.8465 | 0.9256 | 0.8126 | 0.8359 | 0.8012 | **Normal** |
| 2 | 0.3215 | 0.5695 | **0.9611** | 0.6589 | 0.8654 | 0.7986 | **Beta-4** |
| 3 | 0.4153 | 0.3156 | **0.9724** | 0.8456 | 0.2156 | 0.4562 | **Beta-4** |
| 4 | 0.8691 | 0.2562 | **0.8942** | 0.1168 | 0.8446 | 0.8211 | **Beta-4** |
| 5 | 0.2689 | 0.1156 | **0.9326** | 0.1026 | 0.8217 | 0.3423 | **Beta-4** |
| 6 | 0.4516 | 0.2215 | **0.9513** | 0.3226 | 0.8914 | 0.3281 | **Beta-4** |
| 7 | 0.3814 | 0.7981 | **0.8996** | 0.8655 | 0.8769 | 0.7986 | **Beta-4** |
| 8 | 0.4199 | 0.2899 | **0.9265** | 0.1159 | 0.3315 | 0.1886 | **Beta-4** |
| 9 | **0.9128** | 0.3911 | 0.8922 | 0.8451 | 0.8766 | 0.8466 | **Normal** |
| 10 | 0.3155 | 0.2566 | 0.8326 | 0.7955 | **0.8432** | 0.8112 | **Weibull** |
| 11 | 0.9166 | 0.8966 | **0.9365** | 0.3662 | 0.8866 | 0.8544 | **Beta-4** |
| 12 | **0.9158** | 0.2249 | 0.8811 | 0.7915 | 0.4558 | 0.2116 | **Normal** |
| 13 | 0.4119 | 0.8622 | **0.9048** | 0.8412 | 0.8862 | 0.8796 | **Beta-4** |
| 14 | 0.8599 | 0.7899 | **0.8763** | 0.1023 | 0.8336 | 0.7999 | **Beta-4** |
| 15 | 0.4219 | 0.2388 | **0.9478** | 0.1165 | 0.9148 | 0.1089 | **Beta-4** |
| 16 | 0.9689 | 0.3771 | **0.9896** | 0.8459 | 0.9666 | 0.4115 | **Beta-4** |
| 17 | 0.8777 | 0.8122 | **0.8992** | 0.3015 | 0.8605 | 0.3308 | **Beta-4** |
| 18 | 0.2089 | 0.3707 | **0.9768** | 0.8955 | 0.9548 | 0.9401 | **Beta-4** |
| 19 | 0.8716 | **0.8924** | 0.8842 | 0.2566 | 0.8713 | 0.2351 | **Log-Normal** |
| 20 | 0.8991 | 0.1886 | **0.9112** | 0.3112 | 0.8896 | 0.2533 | **Beta-4** |
| 21 | 0.4566 | 0.3692 | **0.9468** | 0.4122 | 0.9369 | 0.2333 | **Beta-4** |
| 22 | 0.9122 | 0.4555 | **0.9244** | 0.8699 | 0.9015 | 0.2212 | **Beta-4** |

**Table 1. Summary of speed distribution of mixed traffic on curved section**

| Section | *p*-value estimated for Distribution | | | | | | Best fit distribution |
|---|---|---|---|---|---|---|---|
| | Normal | Log normal | Beta4 | Gamma | Weibull | GEV | |
| 1 | **0.9996** | 0.1229 | 0.4733 | 0.2177 | 0.8695 | 0.1222 | **Normal** |
| 2 | 0.9662 | 0.3566 | **0.9869** | 0.8865 | 0.9548 | 0.1123 | **Beta-4** |
| 3 | 0.9659 | 0.1189 | **0.9964** | 0.3559 | 0.9785 | 0.1148 | **Beta-4** |
| 4 | **0.9972** | 0.2251 | 0.3955 | 0.1114 | 0.3223 | 0.1869 | **Normal** |
| 5 | 0.9556 | 0.9215 | **0.9762** | 0.8912 | 0.8814 | 0.8622 | **Beta-4** |
| 6 | 0.9443 | 0.2319 | **0.9625** | 0.9115 | 0.9539 | 0.2236 | **Beta-4** |
| 7 | 0.4322 | 0.9122 | **0.9645** | 0.9555 | 0.3998 | 0.9144 | **Beta-4** |
| 8 | 0.3966 | 0.9667 | **0.9992** | 0.9511 | 0.3448 | 0.9556 | **Beta-4** |
| 9 | 0.9322 | 0.8974 | **0.9421** | 0.8666 | 0.9399 | 0.9228 | **Beta-4** |
| 10 | 0.4139 | 0.8169 | **0.9687** | 0.9558 | 0.9559 | 0.9177 | **Beta-4** |
| 11 | 0.9266 | 0.2259 | **0.9301** | 0.3422 | 0.8669 | 0.8999 | **Beta-4** |
| 12 | 0.2533 | 0.3015 | 0.3908 | 0.2287 | **0.9011** | 0.1183 | **Weibull** |
| 13 | 0.4406 | 0.9663 | **0.9987** | 0.9551 | 0.9589 | 0.8996 | **Beta-4** |
| 14 | 0.9666 | 0.9503 | **0.9921** | 0.9644 | 0.9708 | 0.8866 | **Beta-4** |
| 15 | 0.9886 | 0.8744 | **0.9998** | 0.9168 | 0.9703 | 0.9168 | **Beta-4** |
| 16 | 0.9648 | 0.1126 | **0.9721** | 0.9382 | 0.9666 | 0.2301 | **Beta-4** |
| 17 | 0.2739 | 0.9382 | **0.9548** | 0.9169 | 0.9326 | 0.3609 | **Beta-4** |
| 18 | 0.3719 | 0.8631 | **0.9248** | 0.8829 | 0.9067 | 0.8999 | **Beta-4** |
| 19 | 0.4269 | **0.9587** | 0.4139 | 0.3927 | 0.9919 | 0.2968 | **Log-Normal** |
| 20 | **0.9788** | 0.1639 | 0.9839 | 0.3723 | 0.9667 | 0.1639 | **Normal** |
| 21 | 0.3762 | 0.3633 | **0.9279** | 0.8999 | 0.2089 | 0.1934 | **Beta-4** |
| 22 | 0.4551 | 0.2639 | **0.9968** | 0.1979 | 0.3999 | 0.2738 | **Beta-4** |

**Table 2. Summary of speed distribution of mixed traffic on tangent section**

The cumulative distribution curves of best fitted are drawn to estimate the 85th percentile speed at each section. Figure 5 shows the estimation of 85th percentile value using distribution curve. After estimation of 85th percentile speeds at all sections, the statistics, such as maximum, average, and minimum values of 85th percentile speed, were calculated. The standard deviation of speed data was also obtained to observe the variability. The descriptive statistics of 85th percentile speed data is presented in Table 3. The difference between maximum speeds on curve and tangent section is 20 kmph. Also, it can be observed that the difference of average speed variation is 5 kmph only.

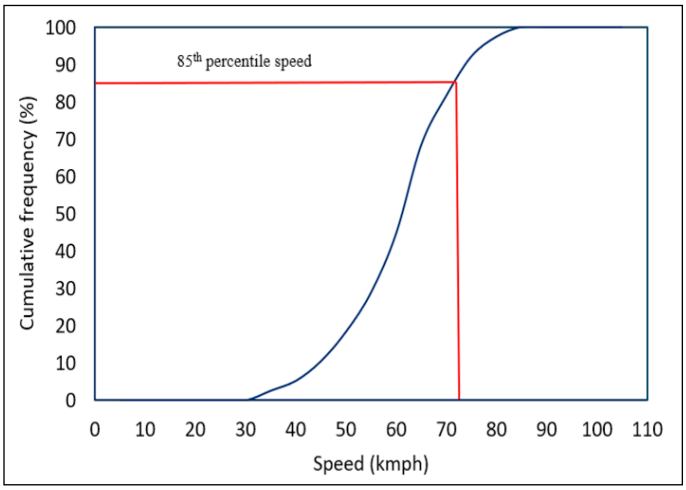


**Figure 5. Cumulative distribution curve of Beta4 for mixed traffic on curved section**

| Statistic | 85th percentile speed on curved section | 85th percentile speed on tangent section |
|---|---|---|
| Maximum (kmph) | 100 | 120 |
| Average (kmph) | 82 | 87 |
| Minimum (kmph) | 66 | 73 |
| Standard deviation (kmph) | 11.2 | 11.35 |
| Sample size | 2200 | 2200 |

**Table 3. Statistics of 85th percentile speed**

## 6. MODEL DEVELOPMENT

The present study develops an operating speed model for vehicles observed on curve and tangent sections of a four-lane divided highway. Multiple linear regression (MLR) and Artificial Neural Network techniques are used for this research. The model forms are discussed in the following section.

### 6.1. MLR model formulation

MLR is the extension of simple linear regression to the case of multiple independent variables. The basic form of the MLR equation is presented in Eq. 1.

(1) $Y_i = a_0 + a_1x_1 + a_2x_2 + \text{-----}\ a_ix_i + \varepsilon_i$

where $Y_i$ is the dependent variable of ith observed values, $a_0$ is the intercept, $a_1$, $a_2$, and $a_i$ are the regression coefficients, $x_1$, $x_2$, and $x_i$ are the independent variables, $\varepsilon_i$ is the regression model random error. In equation (1), Y represents the operating speed ($V_{85}$) as the dependent or response variables and ($x_1$, $x_2$, …… $x_i$) are the geometric variables such as curve radius (R), curve length (CL), deflection angle (D), degree of curvature (DC) shoulder and median (SW & MW) and Access density (AD) considered as the independent or explanatory variables.

### 6.2. ANN model formulation

Artificial neural network is one of the most popular machine learning techniques for non-linear analysis. The ANN structure consists of neurons, weights, and activation functions at various layers. The architecture of the basic ANN processing unit with all elements is shown in Figure 6. The present study adopted a feed-forward with backpropagation neural network (BPNN) modeling technique for estimating operating speed on curves. BPNN is a multi-layer perceptron consisting of an input, hidden, and output layer interconnected with neurons. The ANN architecture uses the Trial and error method to select the required number of hidden layers and neurons. The optimal network structure has a high correlation coefficient and low mean square error value between observed and predicted values.

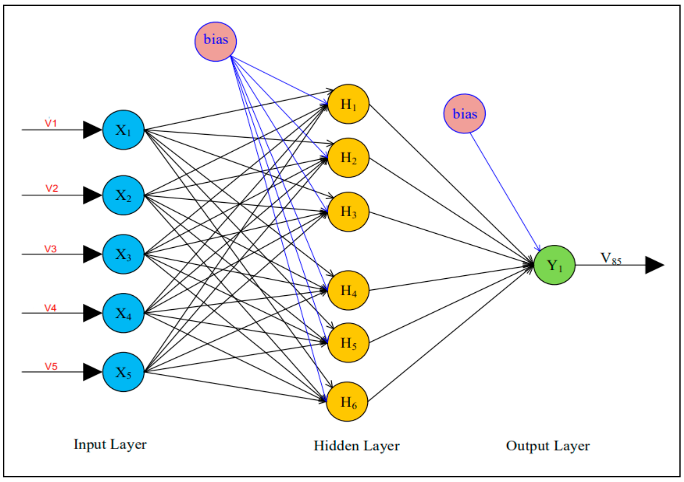


**Figure 6. Architecture of basic ANN structure**

### 6.3 Modeling operating speed on curves using MLR

A correlation test was performed between the geometric parameters and 85th percentile operating speed. The correlation results between the variables give an idea about the strength of the relationship and form the basis for developing the model. The correlation results are shown in Table 4. The negative value infers the decreasing trend, and the positive value infers the increasing trend. It has been clear that the $V_{85}$ has a good correlation with curve radius, curve length and deflection angle. These three variables are selected as independent variables for model development.

| | $V_{85}$ | R | CL | D | DC | SW | MW | AD |
|---|---|---|---|---|---|---|---|---|
| V85 | 1 | | | | | | | |
| R | 0.857* | 1 | | | | | | |
| CL | 0.527 | 0.758 | 1 | | | | | |
| D | -0.794 | -0.906 | -0.698 | 1 | | | | |
| DC | -0.205 | -0.064 | 0.420 | -0.019 | 1 | | | |
| SW | 0.108 | 0.139 | 0.101 | -0.300 | -0.057 | 1 | | |
| MW | -0.300* | -0.031 | -0.162 | -0.057 | -0.172 | -0.058 | 1 | |
| AD | -0.355* | -0.127 | 0.353 | 0.101 | 0.638 | -0.092 | -0.424 | 1 |

* Significant variables; Level of Significance = 0.05

**Table 4. Results of Pearson correlation matrix**

The multiple Linear regression (MLR) method was used to develop the models by selecting all three independent variables. The best model is proposed to estimate operating speed on curves. The results of the developed model are presented in Table 5. It is found that the model shows a good fit with identified variables, providing lower values of standard error

| Parameters | Coefficient | Standard error | t-value | p-value | $R^2$ | Adjusted $R^2$ | Mallow's $C_p$ |
|---|---|---|---|---|---|---|---|
| Intercept | 48.28 | 0.226 | 8.86 | 0.000 | 0.821 | 0.786 | 3.41 |
| R | 0.082 | 0.002 | 5.58 | 0.001 | | | |
| MW | -0.012 | 0.024 | -4.02 | 0.000 | | | |
| AD | -0.12 | 0.056 | -4.86 | 0.003 | | | |

**Table 5. Results of the developed model**

and higher t-values. The *p*-value is found to be significant at a 95% confidence interval. The $R^2$ value and adjusted $R^2$ value are high, and Mallow's $C_p$ value is less. Mallows' coefficient helps to strike a significant balance with the number of predictors in the model. It compares the accuracy and bias of the full model to models with a subset of the predictors. A small Mallow's coefficient value designates that the model is comparatively accurate in assessing the factual regression coefficients and forecasting future responses. It is found that curve length and curve radius are positively correlated, which indicates that the curve sharpness decreases, resulting in more operating speed. The deflection is negatively correlated, which increases curve sharpness, resulting in less operating speed.

## 6.4 Modeling operating speed on curves using ANN

In this study, an ANN-based operating speed model was built using the R STUDIO "neuralnet" package. As a result of correlation analysis, three explanatory parameters (R, MW, and AD) are highly correlated with $V_{85}$ of all vehicle types used in input layer. For network development, the curve data is divided into training data set (75%) and testing data set (25%). To develop an optimal model, the number of neurons in the hidden layer changes according to the performance of the maximum correlation coefficient and minimum error. Several trials were done to reach the optimum model performance; among them, some models were presented in Table 6.

| Architecture | Maximum correlation coefficient | Error (%) |
|---|---|---|
| 3-1-1 | 0.9268 | 0.1358 |
| 3-2-1 | 0.8652 | 0.1963 |
| 3-3-1 | 0.9123 | 0.1632 |
| **3-4-1** | **0.9486** | **0.1257** |
| 3-5-1 | 0.9369 | 0.1423 |

**Table 6. ANN architecture with correlation coefficient and error value**

Table 6 shows that the error value decreases by increasing the number of neurons in the hidden layer in all model structures, but the correlation coefficient randomly changes. The best ANN structure consists of 3 input neurons, 4 hidden neurons, and one output neuron. The correlation coefficient and error values are less than the remaining structures, which are 0.9486 and 0.1257, respectively.

## 6.5 Modeling operating speed on tangent section using MLR

Pearson correlation test was conducted between the 85th percentile speed and geometric variables to find the correlation of these geometric variables on the 85th percentile speed and their internal correlation. The correlation results between the variables give an idea about the strength of the relationship and form the basis for developing the model. The correlation results are shown in Table 7. The negative value implies the decreasing trend, and the positive value implies the increasing trend. Clearly, the $V_{85}$ correlates well with shoulder width, median width, and access density. These three variables are selected as independent variables for model development.

| | V85 | SW | MW | AD | AT | ST |
|---|---|---|---|---|---|---|
| V85 | 1 | | | | | |
| SW | 0.468* | 1 | | | | |
| MW | 0.471* | 0.096 | 1 | | | |
| AD | -0.661* | -0.103 | -0.223 | 1 | | |
| AT | 0.208 | 0.470 | 0.552 | -0.678 | 1 | |
| ST | 0.234 | 0.195 | 0.181 | -0.238 | 0.243 | 1 |

*significant variables Level of Significance = 0.05

**Table 7. Results of Pearson correlation matrix**

The Multiple Linear regression (MLR) method was used to develop the models using three independent variables. Finally, the best-fit model is proposed to predict operating speed on the tangent section. Table 8 presents the summary of developed models. The *p*-value is significant at a 95% confidence interval. The $R^2$ value and adjusted $R^2$ value are high, and Mallow's $C_p$ value is less. It can be observed that shoulder width is positively associated with 85th percentile speed, indicating that the tangent section offers higher operating speed. The median width and access density are negatively correlated. If the median width increases, the operating speed also increases due to avoiding opposite vehicle movement and glare, and vice versa. The access density increases, and the operating speed decreases due to a number of minor road junctions.

## 6.6 Modeling operating speed on tangent section using ANN

This study used the R STUDIO "neuralnet" package to create an ANN-based operating speed model. As a result of correlation analysis, three independent variables (SW, MW, and AD) are highly correlated with $V_{85}$ of all vehicle types used in input layer. For network development, the curve data is divided into training data set (75%) and testing data set (25%). To develop an optimal model, the number of neurons in the hidden layer changes according to the performance of the maximum correlation coefficient and minimum error. Several trials were done to reach the optimum model performance; among them, some models were presented in Table 9.

| Parameters | Coefficient | Standard error | t-value | p-value | $R^2$ | Adjusted $R^2$ | Mallow's $C_p$ |
|---|---|---|---|---|---|---|---|
| Intercept | 94.48 | 1.221 | 9.26 | 0.040 | 0.825 | 0.796 | 2.86 |
| SW | 4.42 | 0.002 | 5.33 | 0.002 | | | |
| MW | -1.30 | 0.023 | -6.23 | 0.031 | | | |
| AD | -1.44 | 0.066 | -4.28 | 0.001 | | | |

**Table 8. Results of the developed model**

| Architecture | Maximum correlation coefficient | Error (%) |
|---|---|---|
| **3-1-1** | **0.8154** | **0.1023** |
| 3-2-1 | 0.2338 | 0.6941 |
| 3-3-1 | 0.7004 | 0.2652 |
| 3-4-1 | 0.4657 | 0.1973 |
| 3-5-1 | 0.0548 | 0.1133 |

**Table 9. ANN architecture with correlation coefficient and error value**

Table 9 shows that the error value decreases by increasing the number of neurons in the hidden layer in all model structures, but the correlation coefficient randomly changes. The best ANN structure consists of 3 input neurons, 1 hidden neuron and one output neuron. The correlation coefficient and error values are less than the remaining structures, which are 0.8154 and 0.1023, respectively.

## 7. VALIDATION

The validation of the model is performed using field data obtained from different sections of highway sections. The validity of the model is based on various performance measures like root mean square error (RMSE) and mean absolute percentage error (MAPE). RMSE is adopted to calculate the differences between model and field data values (sample and population values). MAPE measures the prediction accuracy of a forecasting method in statistics. The RMSE and MAPE values are presented in Table 10. The validation showed a good match between predicted and observed values of operating speeds. The performance measures are within permissible limits (RMSE is 0.2 to 0.5 and MAPE if less than 10%). Moreover, the values of RMSE and MAPE are less for ANN models than MLR in both curved and tangent section models. Hence, the ANN model is the best model for predicting operating speed on curved and tangent sections of Four-lane divided highways.

| Section | Model | RMSE | MAPE |
|---|---|---|---|
| Curved | MLR | 0.386 | 2.680 |
| | ANN | 0.127 | 1.311 |
| Tangent | MLR | 0.412 | 1.953 |
| | ANN | 0.132 | 1.412 |

**Table 10. Performance measures of developed models**

## 8. CONCLUSIONS

The following conclusions are drawn from the findings of the present study:

- Speed data collected at different sections were analyzed to estimate vehicle speed parameters. Statistical speed distribution pattern analysis confirmed that the speed profiles developed are predominantly followed Beta 4 distribution type in most of the curved and tangent sections.
- Three geometric parameters, such as curve radius, median width, and access density, are found to affect the operating speed on curves significantly. The increase in curve radius and median width has increased the operating speed. In contrast, an increase in access density decreased the operating speed.
- The primary influence parameters for developing the operating speed model for tangent section are shoulder width, median width, and access density. The shoulder width correlates positively with operating speed, whereas median width and access density correlate negatively.
- Performance measures like RMSE and MAPE used to validate models are found within the allowable limits shows accuracy of models. Moreover, the ANN models showed more accuracy than MLR models for predicting operating speed.

The developed models are helpful for highway designers to fix the speed limits on curve and tangent sections of four-lane divided highways. This study does not consider the effect of the number of lanes on operating speed because of the limitation of field data. The operating speed model was proposed based on the geometric parameters only. The effect of gradient, road roughness, and degree of super elevation on operating speed should be considered for future studies.

## REFERENCES


Andueza, P. J. (2000). Mathematical models of vehicular speed on mountain roads. *Transportation Research Record*, *1701*, 104–110. https://doi.org/10.3141/1701-13

Glennon, J. C., Neuman, T. R., & Leisch, J. E. (1986). *Safety and Operational Considerations for Design of Rural Highway Curves*. https://trid.trb.org/view/273415

Islam, M., & Seneviratne, P. (1994). Evaluation of design consistency of two-lane rural highways. *ITE J*, *64*(2), 28–31. http://worldcat.org/oclc/614107147

Jacob, A., & Anjaneyulu, M. V. L. R. (2013). Operating speed of different classes of vehicles at horizontal curves on two-lane rural highways. *Journal of Transportation Engineering*, *139*(3), 287–294. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000503

Krammes, R. A. (1995). Geometric Inconsistencies and Accident Experience on Two-Lane Rural Highways. *Transportation Research Record*, *(1356)*, 1–10. http://onlinepubs.trb.org/Onlinepubs/trr/1992/1356/1356-001.pdf

Lamm, R., Choueiri, E. M., & Hayward, J. C. (1988). Tangent as an independent design element. *Transportation Research Record*, *1195*, 123–131. http://onlinepubs.trb.org/Onlinepubs/trr/1988/1195/1195-011.pdf

Lamm, R., Psarianos, B., Drymalitou, D., & Soilemezglou, G. (1995). Guidelines For the Design of Highway Facilities. *Ministry for Environment, Regional Planning and Public Works, Athens, Greece*, *3*.

Lamm, R., Psarianos, B., & Mailaender, T. (1999). *Highway Design and Traffic Safety Engineering Handbook*. http://worldcat.org/isbn/0070382956

Maji, A., Sil, G., & Tyagi, A. (2018). 85th and 98th percentile speed prediction models of car, light, and heavy commercial vehicles for four-lane divided rural highways. *Journal of Transportation Engineering Part A: Systems*, *144*(5), 04018009-1–8. https://doi.org/10.1061/JTEPBS.0000136

Maji, A., & Tyagi, A. (2018). Speed prediction models for car and sports utility vehicle at locations along four-lane median divided horizontal curves. *Journal of Modern Transportation*, *26*(4), 278–284. https://doi.org/10.1007/s40534-018-0162-1

Memon, R. A., Khaskheli, G. B., & Qureshi, A. S. (2008). Operating speed models for two-lane rural roads in Pakistan. *Canadian Journal of Civil Engineering*, *35*(5), 443–453. https://doi.org/10.1139/L07-126

Misaghi, P., & Hassan, Y. (2005). Modeling operating speed and speed differential on two-lane rural roads. *Journal of Transportation Engineering*, *131*(6), 408–418. https://doi.org/10.1061/(ASCE)0733-947X(2005)131:6(408)

Morrall, F. J., & Talarico, J. R. (1994). Side Friction Demanded and Margins of Safety on Horizontal Curves. *Transportation Research Record*, *1435*, 145–152. http://onlinepubs.trb.org/Onlinepubs/trr/1994/1435/1435-019.pdf

Morris, C. M., Asce, M., & Donnell, E. T. (2014). Passenger Car and Truck Operating Speed Models on Multilane Highways with Combinations of Horizontal Curves and Steep Grades. *Journal of Transportation Engineering*, *140*(11), 1–10. https://doi.org/https://doi.org/10.1061/(ASCE)TE.1943-5436.0000715

MoRTH. (2022). Road accidents in india 2022. In *2022*.
Polus, A., Fitzpatrick, K., & Fambro, D. B. (2000). Predicting operating speeds on tangent sections of two-lane rural highways. *Transportation Research Record*, *1737*, 50–57. https://doi.org/10.3141/1737-07
Qureshi, A. S., Khakheli, G.B., & Memon, R, A. (2005). Operating speed prediction model for existing two-lene two-way old alignments. *Mehran University Research Journal of Engineering & Technology*, *24*(4). http://worldcat.org/issn/02547821
Schurr, K. S., McCoy, P. T., Pesti, G., & Huff, R. (2002). Relationship of design, operating, and posted speeds on horizontal curves of rural two-lane highways in Nebraska. *Transportation Research Record*, *1796*, 60–71. https://doi.org/10.3141/1796-07
Sil, G., Maji, A., Nama, S., & Maurya, A. K. (2019a). *Operating Speed Prediction Model as a Tool For Consistency Based Geometric Design Of Four-Lane Divided Highways*. *34*(4), 425–436. https://doi.org/https://doi.org/10.3846/transport.2019.10715
Sil, G., Nama, S., Maji, A., & Maurya, A. K. (2019b). Effect of horizontal curve geometry on vehicle speed distribution: a four-lane divided highway study. *Transportation Letters*, *7867*. https://doi.org/10.1080/19427867.2019.1695562
Sil, G., Nama, S., Maji, A., & Maurya, A. K. (2020). Modeling 85th Percentile Speed Using Spatially Evaluated Free-Flow Vehicles for Consistency-Based Geometric Design. *Journal of Transportation Engineering, Part A: Systems*, *146*(2), 04019060-1–12. https://doi.org/10.1061/JTEPBS.0000286.
Sowmya, N. J., Ravishankar, A. U., & Anjaneyulu, M. V. L. (2012). Modelling operating speed and speed differential on intermediate lane rural roads. *International Journal of Earth Sciences and Engineering*, *5*(1), 1408–1414.
Tottadi, K.K., Mehar, A. (2021). Operating speed: review and recommendations for future research. *Innovative Infrastructure Solutions*, 7(67), https://doi.org/10.1007/s41062-021-00669-9.
Tottadi, K.K., Varma, C. & Mehar, A. (2022). Influence of Horizontal Curve Geometry on Operating Speeds on Four-Lane Divided Highways Under Heterogeneous Traffic Conditions. *Journal of Institution of Engineers India Series A*, 103, 1135–1145. https://doi.org/10.1007/s40030-022-00677-7.