Clustering Operating Speed Profiles on Horizontal Curves: A Short-term Naturalistic Driving Study

Vinayak Malaghan\textsuperscript{1*}, and Digvijay S. Pawar\textsuperscript{2}

\textsuperscript{1}Postdoctoral Researcher, Transportation Systems Engineering, Department of Civil Engineering, Indian Institute of Technology Hyderabad, Kandi, Medak - 502285, India; E-Mail: vinayakmalagahan@gmail.com

\textsuperscript{2}Assistant Professor, Transportation Systems Engineering, Department of Civil Engineering, Indian Institute of Technology Hyderabad, Kandi, Medak - 502285, India; PH (040) 2301 6167; E-Mail: dspawar@ce.iith.ac.in

\textbf{6\textsuperscript{th} International Symposium on Highways Geometric Design}

Submission Date: January 1\textsuperscript{st}, 2022

Number of words in Abstract = 250

Number of words in Text = 6,093

Number of tables = 3

Number of Figures = 3

Total Word Count = 250 + 6,093 + 3*250 = 7,093

\textsuperscript{*}Corresponding Author
Abstract

The minimum operating speed on the curve is measured to evaluate the geometric design consistency. Several studies measured the operating speed at the midpoint of the curve by approximating equal lengths of deceleration and acceleration in the operating speed profiles for the curves. In contrast, few studies showed different percentage deceleration lengths in operating speed profiles on the curve and measured the operating speed at the deceleration end. A defined pattern(s) in the speed profile was not reported. Therefore, in this study, the operating speed profiles of different drivers on the horizontal curves were studied, and the clustering technique was used to categorize the different patterns in the operating speed profiles on the curves. The optimal number of clusters was determined using four methods: Silhouette, Elbow, Gap statistic, and using NbClust function. The different patterns observed from the clustering results are: (a) complete acceleration/acceleration on the curve, (b) deceleration length slightly greater or lower than acceleration length, (c) longer deceleration/acceleration lengths followed by a shorter acceleration/deceleration lengths. The study results infer that all the operating speed profiles are not symmetric around the MC, and the group of drivers exhibit defined pattern(s) of the operating speed profiles on the curves. This study helps in understanding the different patterns of operating speed profiles exhibited by the drivers and the measurement of the minimum operating speed at the deceleration end to model the operating speed to assess the geometric design consistency.

Keywords: Design consistency; clustering; minimum operating speed; midpoint of the curve.
1. Introduction

The primary components of highway mode of transport that contribute to road crashes are driver, road infrastructure, and vehicle. Specifically, the road infrastructure component contributes to around 2.2 million road crashes in India [1]. A significant number of crashes were reported to be cumulative on the curves and tangent to curve transitions of roads [2]. In India, the fatal and non-fatal crashes on curved roads were reported to be approximately 20,000 and 60,000, respectively, underlining the importance of horizontal alignment design in preventing road crashes [1]. One of the primary reasons for these road crashes is the lack of consistency in the geometric design standards that match the characteristics of the drivers and vehicles. Therefore, the interaction between road users and road infrastructure has been widely studied in recent years. In this context, geometric design consistency is defined as how well the highway geometric design features match the characteristics of the drivers and vehicles that use the road [3, 4]. A consistent roadway designed to meet the characteristics of drivers results in homogenous and harmonious driving operations and thereby safe maneuvers. In contrast, an inconsistent roadway causes sudden variations in the operating speed, causing critical maneuvers and, therefore, increasing the likelihood of crash occurrence [5].

Operating speed is one of the widely used measures in evaluating geometric design consistency [6]. The initial step in the operating speed measure is to develop the operating speed models for tangents, curves, and tangent to curve transitions. In developing the models, the maximum operating speed (the speed at the beginning of deceleration) and minimum operating speed (the speed at the end of deceleration) are measured on the tangent and curve, respectively [4, 7].

Several studies assumed constant operating speed on the horizontal curve and entire acceleration/deceleration to be completed on the departure/approach tangent [8, 9]. Based on the assumption mentioned above, the operating speed was measured at the midpoint of the curve (MC) and tangent to model operating speed for the geometric elements (tangent, curve, and tangent to curve transition) [10, 11]. However, studies invalidated the assumption of operating speed to be constant on the horizontal curve and concluded that the deceleration continued even inside the horizontal curve [4, 7]. In addition, most of the studies assumed the length of deceleration and acceleration to be approximately equal on the horizontal curve [12, 13]. In other words, the pattern of the operating speed profile was assumed to be approximately symmetric about the MC on the horizontal curve [12]. In contrast, several recent studies showed that the deceleration end (or position of minimum operating speed) in every operating speed profile might not coincide with the MC, i.e., every operating speed profile might not be symmetric about the MC [4, 14]. The percentage length of deceleration and acceleration transitions were random and varied across several past and recent studies [12, 14, 15, 16]. Therefore, a defined pattern(s) of the operating speed profile on the horizontal curve was not observed [6, 8]. With this motivation, this study investigates whether the drivers exhibit a defined pattern(s) of operating speed profiles on the horizontal curve. Therefore, in this study, the vehicles were instrumented with a global positioning system (GPS) device to record the continuous operating speeds along the trajectory on the selected road sections. The operating speed profiles on different horizontal curves between PC and PT were extracted, and an unsupervised machine learning approach was used to cluster the operating speed profiles with a similar pattern(s).

The positions of maximum (or beginning of deceleration) and minimum (or end of deceleration) operating speeds on the tangent and curve, respectively, depends on the pattern(s) of the operating speed profiles. Understanding the pattern(s) of the operating speed profiles allow to measure the maximum/minimum operating speed accurately to develop reliable
operating speed models and thereby promote the geometric design consistency evaluation in road safety.

2. Literature Review

Operating speed is the commonly used measure to evaluate the geometric design consistency of rural highways [2, 14, 15]. The speed at which drivers operate their vehicles under free-flowing conditions is referred to as operating speed. It is calculated as the 85th percentile value of the observed operating speeds at a specified location(s) [16]. The device used to measure the operating speed in analyzing the operating speed profiles can be broadly categorized into three types: spot speed devices, driver simulators, and GPS devices equipped in the vehicles. Below subsections present the literature review on operating speed profiles based on the type of device used.

2.1 Operating Speed Profiles Based on Spot Speed Devices

Lamm et al. [17] measured spot speeds at eleven spots on the approach tangent from the point of curvature (PC). They investigated the deceleration and acceleration movements from tangent to curve and curve to tangent, respectively. Their study assumed operating speed to be constant on the horizontal curve and completion of acceleration/deceleration on the departure/approach tangent. The assumption of constant operating speed profile (without any speed variations) on the horizontal curve was supported without any statistical validation, albeit an operating speed difference of 4 to 5 mph at the curve ends. Based on the assumption above and the available length of tangent between the curves, Ottesen and Krammes [8] categorized the patterns of operating speed profiles into three generalized cases. In all three cases, the operating speed on the horizontal curve was assumed constant without any variation in the speed.

The operating speed to be constant on the horizontal curve was not supported in further studies [4, 7, 18]. Jacob and Anjaneyulu [13] conducted a pilot study and concluded that drivers decelerated inside the curve and attained the minimum operating speed around the midpoint of the curve within a 10 m trap length. The authors assumed that the lengths of deceleration and acceleration on the horizontal curve are approximately equal. Figueroa Medina and Tarko [12] generalized the operating speed profile to develop operating speed profile models. The authors assumed the pattern of the operating speed profiles to be symmetric from tangent to curve and curve to tangent about the MC.

As mentioned above, studies inferred that the drivers attained minimum operating speed or ended deceleration at the MC. McFadden and Elefteriadou [18] investigated whether all the drivers attained minimum operating speed at the MC. The authors divided the horizontal curve into four quarters, and the spot speed was measured at five locations to determine the position of the minimum operating speed on the horizontal curve. The results of their study showed that the drivers attained the minimum operating speed at any of the five locations on the horizontal curves.

The spot speed devices allow the measurement of a speed value at a specific point on a given length of the geometric element(s). This limitation can be overcome by using driving simulators or vehicles equipped with a GPS device that allows recording continuous speed data along the vehicle's trajectory.

2.2 Operating Speed Profiles Based on Driving Simulators Studies

Several studies based on driving simulators analysed the operating speed profiles on the curve and from tangent to the curve and curve to the tangent [16, 19, 20]. Montella et al. [19] obtained 357 operating speed profiles and analysed the deceleration and acceleration behaviour from tangent to curve transition and curve to tangent transition, respectively. The average percent proportion of deceleration length on the geometric elements is: (a) 59 percent on the
approach tangent, 30 percent on the spiral transition, and 11 percent on the curve. In the case of acceleration, the average percent proportion on the geometric elements include: (a) 2 percent afore the circular curve, (b) 47 percent on the circular curve, (c) 18 percent on the spiral transition after the curve; and (d) 33 percent on the departure tangent. Similarly, Bella [16] analysed 856 operating speed profiles of a single driver from tangent to curve and curve to tangent and developed operating speed profile models. The authors observed that the prevailing behaviour of drivers on the geometric elements constituted as follows:

- The drivers keep their preferred operating speed on the approach tangent;
- The drivers (83 %) commence deceleration as they approach the curve and complete the deceleration at some point (minimum operating speed position) inside the curve (74 %);
- The attainment of the minimum operating speed position depends on the features of the curve; and
- The drivers start the acceleration inside the curve and continue the acceleration (95 %) until their preferred operating speed is attained on the tangent.

The studies based on driving simulators have limitations such as awareness of low-risk, lack of dynamic visualization, and lack of motion perception. Therefore, the vehicles are instrumented with GPS devices to overcome these limitations and obtain driver data in the real-world environment.

2.3 Operating Speed Profile Studies Based on Instrumented Vehicle Studies

Pérez Zuriaga et al. [7] used a pocket-sized GPS device equipped in the passenger cars to obtain continuous operating speed profiles along the trajectory of the vehicles. Using the speed profiles, the authors determined the operating speed at the beginning and end of deceleration and developed operating speed profile models. Their study results showed that 45 percent of the curve length was traversed in deceleration, and in 58 percent of the curve sites, the deceleration ended before the MC. A study by Montella et al. [21] analysed 2,720 operating speed profiles from tangent to curve (and curve to tangent) and developed operating speed profile models. The authors observed that 52 percent of the transitions in deceleration started on the approach tangent and ended inside the curve. Whereas 90 percent of the transitions in acceleration started inside the curve. Among 90 percent of the transitions, 27 and 69 percent started before and after the MC, respectively. The most frequent transition on the curve was complete deceleration. Studies showed random percentage variation in the beginning or end of deceleration/acceleration transitions on the geometric elements. A recent study by Malaghan et al. [4] showed that the distribution of minimum operating speed positions followed normal distribution on the greater number of curves.

3. Research Gaps and Objective

To summarize, the earlier studies in the literature assumed constant operating speed on the horizontal curve. While fewer past and recent studies concluded that the deceleration might end around the midpoint of the curve, i.e., the acceleration and deceleration lengths of the operating speed profiles on the horizontal curve were assumed to be approximately equal. With advancements in technology, the studies based on driving simulators and instrumented vehicles showed random variations in the percentage acceleration and deceleration lengths on the horizontal curves. A defined pattern(s) of the operating speed profiles on the horizontal curve was not observed. Therefore, the primary objective of this study is to determine the pattern(s) of the operating speed profiles on the horizontal curve(s) using an unsupervised machine learning approach.

4. Research Methodology
4.1 Test Routes and Geometric Data

Five two-lane rural highways, a national highway (NH), three state highways (SH), and a major district road (MDR) segments in the states of Karnataka and Telangana, India were selected to perform the field experiment. The chosen road segments were located in the plain terrain. The existing horizontal alignment design features of the selected road sections were determined using a highway geometric design tool, ‘AutoCAD Civil 3D’ and ‘Google Earth Pro’ software. The summary of the descriptive statistics of the horizontal alignment design features of the chosen road segments is same as presented in a study by Malaghan et al. [22]. The secondary data of the selected road segments, such as lane width, number of lanes, type of shoulder, and width of paved/unpaved shoulder were measured during the field visits. The lane width of the selected road segments ranged from 3.0 m to 3.5 m. The paved and earthen shoulder-width ranged from 1.0 m to 1.5 m and 1.0 m to 2.0 m, respectively. The following criteria to ensure the influence of road geometry on the operating speed were confirmed during the field visits:

- The curve sites were away from the major/minor intersection;
- The curve sites were away from the merging/diverging approaches;
- The curve sites were away from the influence of narrow bridges, culverts, and the road over/under bridge;
- The selected road segments represented low traffic volume; and
- The selected road segments represented good pavement condition with proper road markings.

4.2 In-Vehicle Data

The in-vehicle data were collected using passenger cars instrumented with a 10 Hz GPS data logger similar to study by Malaghan et al. [4]. The in-vehicle data such as speed, position (latitude and longitude), longitudinal acceleration, lateral acceleration were recorded at every 0.1 s and stored in the memory card inserted in the GPS box. The camera recorded the high definition (HD) video (1080p video at 60 frames per second) of roadside and weather conditions of the selected road segments during the data collection period. The data were collected during the daytime under dry weather conditions. The number of the trips on the selected road segments ranged between 48 and 91, as shown in Table 1. The data collection extended for six months, and the total time spent was 126 man-hrs.

Table 1. Summary of the number of trips on the selected road segments

<table>
<thead>
<tr>
<th>Road segments</th>
<th>No. of curves</th>
<th>No. of tangents</th>
<th>No. of independent tangents</th>
<th>No of trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDR</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>54</td>
</tr>
<tr>
<td>SH 17</td>
<td>13</td>
<td>20</td>
<td>20</td>
<td>70</td>
</tr>
<tr>
<td>SH 53</td>
<td>17</td>
<td>24</td>
<td>17</td>
<td>87</td>
</tr>
<tr>
<td>SH 135</td>
<td>11</td>
<td>15</td>
<td>15</td>
<td>91</td>
</tr>
<tr>
<td>NH 161</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>48</td>
</tr>
</tbody>
</table>

4.3 Participants Profiles

To investigate the influence of road geometry on the operating speed of the drivers, a total of 49 participants voluntarily participated in the experiment. The age of the participants ranged from 21 to 59 years, with a mean and standard deviation (SD) of 33.02 years and 9.58 years. The participants’ driving experience ranged from 1 to 38 years (counted from the driving license), with a mean of 12.60 years and a standard deviation of 9.77 years.

4.4 Data Extraction and Reduction
The speed data recorded at 10 Hz along the trajectory of the selected road sections were used to draw the operating speed profiles for the horizontal curves. Each of the selected road segments was subdivided into tangents and curves by referencing the position coordinates (latitude and longitude) at the beginning and end of the tangents and curves, respectively (see Figure 1 (b)). Figure 1 (b) is a schematic representation of the geometric elements in the direction of travel as the vehicle traverses from section 1-1 to 4-4. Sections 1-1 to 2-2 represent approach tangent, sections 2-2 to 3-3 represent horizontal curve, and sections 3-3 to 4-4 represent departure tangent (see Figure 1 (b)). The speed data corresponding to the distance at 0.1 s on the horizontal curve was extracted by referencing the position coordinates at section 2-2 (beginning of the curve) and 3-3 (end of the curve) (refer to Figure 1 (a and b)). The data obtained were used to plot the operating speed profiles on the horizontal curves. An example operating speed profile for a given curve features is shown in Figure 1 (a). In total, 3,798-speed profiles on the horizontal curves were extracted for the analysis. The data extraction and plotting of the profiles were performed using the code developed in R Studio.

![Figure 1](image)

**Figure 1.** (a) Operating speed profile for the horizontal curve, and (b) representation of the geometric elements.

Several studies recommended a time headway between the host vehicle and the following vehicle ranging from 4 to 6 s to ensure the influence of road geometry on the operating speed [6, 22]. Hence, in this study, the speed profiles of the vehicles in which time headway ≥ 5 s were considered for the analysis. The operating speed profiles on the curves influenced by factors other than road geometry were removed as explained in a study by Malaghan et al. [5]. Hence, 2,310 speed profiles affected alone due to the road geometry were considered in further analysis.

### 4.5 Clustering

Clustering is an unsupervised machine learning approach used to partition the observations into different groups so that observations in a group are quite similar to each other. In contrast, the observations in different groups are quite different from each other. In this study, $k$-means clustering approach is used to partition the data into $k$ different sets. The observations are partitioned in to $k$ – clusters such that the total within-cluster, added over all $k$ – clusters is minimized. The most common choice to define the within-cluster variation involves squared Euclidean distance. In order to identify the pattern(s) of the operating speed profiles on the horizontal curves with different curve features, this study used $k$ – means clustering approach.

### 5. Analysis of Operating Speed Profiles on Horizontal Curves

The operating speed profiles on the five horizontal curves with different geometric features were analysed to identify the pattern(s) in the operating speed profiles. This study identified and considered five primary operating speed profile variables: operating speed at PC ($V_{PC}$),
operating speed at PT ($V_{pt}$), minimum operating speed on the curve ($V_{cm\text{in}}$), deceleration length on the curve ($D_{cl}$), and acceleration length on the curve ($A_{cl}$) to cluster the operating speed profiles with a similar pattern(s). The $k$-means clustering approach was used to understand if the operating speed profiles on the horizontal curves followed a defined pattern or set of patterns. The cluster analysis was performed separately on the variables in the operating speed profiles for each of the five horizontal curves having different geometric design features. To begin with, a curve site 1 (CS1) having a curve radius of 60 m and a curve length of 50 m, the operating speed profiles on this horizontal curve were clustered into three groups (see Figure 2 and Table 2). The optimal number of clusters were determined using four methods: Elbow method, Silhouette method, Gap statistic method, and using NbClust function [23]. The Silhouette and Gap statistic methods compute the optimal number of clusters as four and one, respectively (see Figure 2 (a and b)). The location on the knee in the elbow method suggests the optimal number of clusters as three (see Figure 2 (c)). The fourth method using NbClust () function, computes the optimal number of clusters for each of the 30 indices. For example, one of the indices is a Dindex, and the plots are shown in Figure 2 (d). The significant knee in the Dindex can be obtained from the significant peak in the second differences Dindex plot (from Figure 2 (d)) that corresponds to a significant increase in the value of this index. The Dindex proposed an optimal number of clusters as four. Out of 30 indices, 4 indices proposed 2 as the optimal number of clusters, 7 proposed 3 as the optimal number of clusters, 3 proposed 5 as the optimal number of clusters, 1 proposed seven as the optimal number of clusters, 1 proposed 10 as the optimal number of clusters, 3 proposed 11 as the optimal number of clusters, 1 proposed 14 as the optimal number of clusters, and 3 proposed 15 as the optimal number of clusters. The optimal number of clusters is three as per the majority rule from all the four methods, as shown in the frequency histogram (see Figure 2 (e)). Therefore, the optimal number of clusters is three for the CS1. For CS1, the percentage of operating speed profiles in cluster 1, cluster 2, and cluster 3 are 20 percent, 41 percent, and 39 percent, respectively (see Table 2). Several studies assumed the occurrence of minimum operating speed or the end of the deceleration at the midpoint of the curve [13, 24]. In other words, the deceleration and acceleration lengths were assumed to be approximately equal, and the operating speed profile to be symmetric around the MC [12]. Hence, the percentage of operating speed profiles ending the deceleration before and after the MC for each cluster in different curve sites are shown in Table 2. All the operating speed profiles ended deceleration after the MC in cluster 1 for the CS1. The average percentage deceleration length on the curve is 85 percent, and the average percentage acceleration length is 15 percent for the operating speed profiles in cluster 1 for the CS1. The pattern of the operating speed profiles in curve site 1 for cluster 1 (CS1- C1) is shown in Figure 3 (CS1 – C1). In the notation CS1 – C1, the first word CS1 indicates the curve site number, which is one in this case, and the second word followed by hyphen indicates the cluster number, which is also one in this case. These notations are followed and presented further to identify the cluster number and curve site in the top left corner for each figure in Figure 3. The operating speed profiles in CS 1-C1 show a continuous decrease in operating speed from PC with a steep slope followed by a constant operating speed or marginal increase in the operating speed till the PT. Overall, this variation in the operating speed is represented by the average operating speed profile shown in the dashed line (see Figure 3 (CS1-C1)). In the case of CS1-C2, 96 percent of the operating speed profiles ended deceleration after the MC and the remaining 4 percent before the MC (see Figure 3 (CS1-C2) and Table 2). The average percentage of deceleration and acceleration lengths are 74 and 26 percent, respectively. The operating speed values decrease continuously with a steeper slope followed by a constant or mild slope with a marginal increase in the operating speed values. This variation in the operating speed values for the operating speed profiles in cluster 2 is shown by an average operating profile (dashed line as shown in Figure 3 (CS1 - C2)). While in the case of CS1 –
C3, the operating speed profiles show a decrement in the speed values near PC for a smaller distance followed by an increase in the operating speed values continuously till the PT (see Figure 3 (CS1 – C3)). The average percentage deceleration length of the operating speed profiles in this cluster is 25 percent, and the average percentage acceleration length is 75 percent (see Table 2). All the operating speed profiles show the end of deceleration before the MC, and the drivers begin the acceleration before the MC, which is continued till PT. The results of the summary statistics of the acceleration and deceleration lengths of the operating speed profiles for the respective clusters in the five curve sites is presented in Table 2.

![Figure 2](image_url)

**Figure 2.** Four methods showing optimum number of clusters: a). Silhouette method, b). Elbow method, c). Gap statistic method, and d). Using NbClust function.

For the curve site 2 \((R = 80 \text{ m}, L_c = 60 \text{ m})\), the operating speed profiles in the cluster 1 are less in number and a definite pattern could not be observed from this cluster (see Figure 3 (CS2 – C1) and Table 2). In cluster 2, the operating speed profiles showed decrement in the speed over major section of the curve from PC and then increase in the speed till the curve end (see Figure 3 (CS2 – C2)). In cluster 3 and 4, the pattern of the operating speed profiles is similar, i.e., driver decelerate initially after entering the curve and followed by acceleration in the major part of the curve (see Figure 3 (CS2 – C3) and Figure 3 (CS2 – C4)). However, the percentage of deceleration length in the operating speed profiles in cluster 3 is lesser than that in the cluster 4 (see Table 2).

For the curve site 3 \((R = 80 \text{ m}, L_c = 87 \text{ m})\), the operating speed profiles in cluster 1 show that drivers decelerate in major portion of the curve length from PC and followed by acceleration in minor portion of the curve length till PT (see Figure 3 (CS3 – C1) and Table 2). The average percentage acceleration and deceleration lengths are 75 percent and 25 percent.

For the curve site 3 \((R = 80 \text{ m}, L_c = 87 \text{ m})\), the operating speed profiles in cluster 1
Table 2. Summary statistics of the operating speed profiles in different clusters for the curves with different geometric features

<table>
<thead>
<tr>
<th>CS</th>
<th>GF</th>
<th>Clusters</th>
<th>Deceleration</th>
<th>Acceleration</th>
<th>Percentage length</th>
<th>Percentage of Observations</th>
<th>Percentage of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Mean</td>
<td>SD</td>
<td>Min</td>
</tr>
<tr>
<td>CS1</td>
<td>R = 60, L_c = 50</td>
<td>Cluster 1</td>
<td>28.825</td>
<td>49.715</td>
<td>40.404</td>
<td>5.757</td>
<td>1.265</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 3</td>
<td>1.347</td>
<td>24.538</td>
<td>12.926</td>
<td>6.907</td>
<td>24.997</td>
</tr>
<tr>
<td>CS2</td>
<td>R = 80, L_c = 60</td>
<td>Cluster 1</td>
<td>34.060</td>
<td>57.127</td>
<td>47.753</td>
<td>9.870</td>
<td>5.182</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 2</td>
<td>32.059</td>
<td>48.594</td>
<td>37.996</td>
<td>4.899</td>
<td>9.279</td>
</tr>
<tr>
<td>CS3</td>
<td>R = 80, L_c = 87</td>
<td>Cluster 1</td>
<td>49.193</td>
<td>80.000</td>
<td>63.034</td>
<td>8.282</td>
<td>1.530</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 2</td>
<td>19.090</td>
<td>56.044</td>
<td>39.646</td>
<td>9.012</td>
<td>27.671</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 3</td>
<td>12.300</td>
<td>50.552</td>
<td>32.079</td>
<td>13.438</td>
<td>32.776</td>
</tr>
<tr>
<td>CS4</td>
<td>R = 90, L_c = 64</td>
<td>Cluster 1</td>
<td>41.105</td>
<td>63.651</td>
<td>56.636</td>
<td>8.536</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 2</td>
<td>16.063</td>
<td>38.558</td>
<td>29.236</td>
<td>6.394</td>
<td>24.904</td>
</tr>
<tr>
<td>CS5</td>
<td>R = 95, L_c = 47</td>
<td>Cluster 1</td>
<td>23.048</td>
<td>46.463</td>
<td>32.128</td>
<td>6.200</td>
<td>0.597</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 3</td>
<td>1.371</td>
<td>25.004</td>
<td>12.986</td>
<td>6.183</td>
<td>22.727</td>
</tr>
</tbody>
</table>

Note: CS – Curve Site; CS1, CS2...CS13 – Curve Site1, Curve Site2,..., Curve Site13; GF – Geometric Features; D – Deceleration; A – Acceleration; MC – midpoint of curve.
Figure 3. Clusters showing the patterns of the speed profiles for different curves show that drivers decelerate in major portion of the curve length from PC and followed by acceleration in minor portion of the curve length till PT (see Figure 3 (CS3 – C1) and Table 2). The average percentage acceleration and deceleration lengths are 75 percent and 25 percent, respectively (see Table 2). While in the case of cluster 2, the average percentage deceleration length and acceleration length are 46 percent and 54 percent, respectively. The acceleration lengths are slightly higher than the deceleration lengths, and that the positions of minimum operating speeds in the operating profiles are located near the MC (see Figure 3 (CS3 – C2)).
In cluster 3, the number of operating speed profiles are less, and a defined pattern of the operating speed profiles was not observed.

For the curve site 4 ($R = 80$ m, $L_c = 87$ m), the number of operating speed profiles are less in cluster 1 and 2. A defined pattern of the operating speed profiles was not observed in cluster 2; however, most of the operating speed profiles in cluster 1 showed a complete deceleration with a steeper slope (see Figure 5 (CS4 – C2)). In cluster 3, the average acceleration length of the operating speed profiles is marginally greater than the deceleration length. The cluster of the operating speed profiles observed to be approximately to be symmetric around the MC (see Figure 5 (CS4 – C3)).

For the curve site 5 ($R = 95$ m, $L_c = 47$ m), the operating speed profiles in cluster 1 show that the major portion of the curve is traversed in deceleration with the mild slope in both acceleration and deceleration (see Figure 5 (CS5 – C1)). While clusters 2 and 3 showed similar pattern, i.e., a smaller portion of the curve is traversed in deceleration from PC and major portion of the curve traversed in acceleration (see Figure 5 (CS5 – C2) and Figure 5 (CS5 - C3)). However, the average deceleration length in cluster 2 is higher than that in the cluster 1 (see Table 4).

One-way Analysis of Variance (ANOVA) test was conducted to check if any statistically significant difference existed in the means of the deceleration and acceleration lengths of the operating speed profiles in each cluster for five different horizontal curves.

Null hypothesis: the means of the deceleration and acceleration lengths in a cluster are the same

Alternative hypothesis: the means of the deceleration and acceleration lengths are different in a cluster

The results of the ANOVA test are presented for the deceleration and acceleration lengths in each cluster for five curves with different geometric design features in Table 3. From Table 3, it can be seen that the means of the deceleration and acceleration lengths for the operating speed profiles in fourteen clusters are statistically different at a 0.05 significance level. Thus, the test rejected the null hypothesis stating that the means of the deceleration and acceleration lengths are the same in the operating speed profiles for 87.5 percent of the clusters. However, it can be seen from the results of Table 3 (in bold) that the means of the deceleration and acceleration lengths for the operating speed profiles in three clusters are not statistically different at a 0.05 significance level. Thus, the test failed to reject the null hypothesis that the

### Table 3. Summary of the ANOVA results for deceleration and acceleration lengths of operating speed profiles in each cluster for curves with different design features.

<table>
<thead>
<tr>
<th>Sr.No</th>
<th>Geometric Features</th>
<th>Clusters</th>
<th>$F$ – value</th>
<th>$p$ – value</th>
<th>$F$ – critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$R = 60, L_c = 50$</td>
<td>Cluster 1</td>
<td>156.672</td>
<td>&lt; 0.001</td>
<td>4.260</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 2</td>
<td>141.710</td>
<td>&lt; 0.001</td>
<td>4.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 3</td>
<td>172.919</td>
<td>&lt; 0.001</td>
<td>4.034</td>
</tr>
<tr>
<td>2.</td>
<td>$R = 80, L_c = 62$</td>
<td>Cluster 1</td>
<td>52.031</td>
<td>&lt; 0.001</td>
<td>4.965</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 2</td>
<td>144.401</td>
<td>&lt; 0.001</td>
<td>4.085</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 3</td>
<td>60.885</td>
<td>&lt; 0.001</td>
<td>4.113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 4</td>
<td>105.778</td>
<td>&lt; 0.001</td>
<td>4.062</td>
</tr>
<tr>
<td>3.</td>
<td>$R = 80, L_c = 87$</td>
<td>Cluster 1</td>
<td>313.969</td>
<td>&lt; 0.001</td>
<td>4.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 2</td>
<td>7.161</td>
<td>0.010</td>
<td>4.034</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 3</td>
<td>13.802</td>
<td>0.001</td>
<td>4.351</td>
</tr>
<tr>
<td>4.</td>
<td>$R = 90, L_c = 64$</td>
<td>Cluster 1</td>
<td>150.620</td>
<td>&lt; 0.001</td>
<td>4.494</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 2</td>
<td>2.483</td>
<td>0.135</td>
<td>4.494</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 3</td>
<td>0.343</td>
<td>0.559</td>
<td>3.967</td>
</tr>
<tr>
<td>5.</td>
<td>$R = 95, L_c = 47$</td>
<td>Cluster 1</td>
<td>66.674</td>
<td>&lt; 0.001</td>
<td>4.113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 2</td>
<td>63.134</td>
<td>&lt; 0.001</td>
<td>4.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cluster 3</td>
<td>120.739</td>
<td>&lt; 0.001</td>
<td>4.098</td>
</tr>
</tbody>
</table>
means of the deceleration and acceleration lengths were the same in the operating speed profiles for 12.5 percent of the clusters. Therefore, assuming equal acceleration and deceleration lengths to measure the minimum operating speed at the midpoint of the curve is not appropriate in the majority of the operating speed profiles.

7. Discussion

This study used a clustering approach to investigate the pattern(s) of the operating speed profiles on the horizontal curves. Ottesen and Krammes [8] and Fitzpatrick and Collins generalized the operating speed profiles on a tangent to curve and curve to tangent into three different cases. The authors assumed operating speed to be constant on the horizontal curves without any speed variation. On the other hand, several other studies approximated the acceleration and deceleration length to be equal and that the minimum operating speed occurs at the midpoint of the curve to model geometric design consistency measures in the evaluation of design consistency [12, 13, 24]. The results from the clusters analysis of this study showed the variation in the operating speed on the horizontal curves with different patterns. The operating speed profiles in most of the clusters showed a major portion of the curve traversed in either acceleration or deceleration. Besides, the results of the ANOVA test showed that in 85 percent of the clusters, the mean lengths of acceleration and deceleration were found to be statistically different at 5 percent significance level.

Pérez Zuriaga et al. [7] found that the average percentage deceleration length on the curve was 45 percent and that the deceleration ended before MC in 58 percent of the curve sites. Montella et al. [21] determined that in 27 percent of the transitions, the deceleration started before the MC, and in 63 percent of the transitions, the acceleration started after the MC. The most frequent pattern observed was the continuous deceleration in 22 percent of the transitions on the entire curve. The studies mentioned above did not define any pattern in the operating speed profiles on the horizontal curve. Also, the studies presented the overall average results of the deceleration and acceleration maneuvers. In the present study, the patterns were clustered, and the descriptive statistics of the acceleration and deceleration maneuvers for each of the clusters are presented in Table 4.

8. Summary, Conclusions and Future Directions

This study analysed the pattern(s) in the operating speed profiles on the horizontal curves of the rural roadway segments located in the plain terrain. The operating speed profiles were recorded on the selected road segments using a GPS device instrumented in the passenger cars. The operating speed profiles between PC and PT were extracted for all the data samples for each curve site on each of the selected road sections. In total, the operating speed profiles on the five horizontal curves were studied. The operating speed profiles on each curve were clustered separately using the k-means clustering approach. This study analysed the pattern(s) in the sixteen clusters developed from the operating speed profiles on the five horizontal curves.

The key findings of the research study are presented below:

- Most of the clusters show that the mean lengths of acceleration and deceleration are statistically different. This infers that measurement of operating speed at the end of deceleration results in accurate determination of minimum operating speed on the horizontal curve in evaluating design consistency. In others words, the deceleration end will not coincide with the MC in all the operating speed profiles; the measurement of the operating speed at MC might not result in the accurate measurement of minimum operating speed in design consistency evaluation.

- The operating speed profiles showed defined patterns on the horizontal curves. The primary patterns observed included a) longer deceleration length followed by a marginal acceleration length for a shorter distance, b) shorter deceleration length
followed by longer acceleration length, c) deceleration length slightly greater than acceleration length, and d) deceleration length slightly shorter than the acceleration length. In addition to the four patterns, fewer operating speed profiles showed continuous acceleration and continuous deceleration between PC and PT.

- The study did not observe any constant operating speed profiles without variation in the operating speed on the horizontal curves.

This study has a few limitations which can be addressed in future research work. In this study, the variation in the operating speed only on the horizontal curves is studied. Since the maximum operating speed on the tangent is measured in developing the variation in the operating speed only on the horizontal curves is studied. Since the tangent can be identified using the clustering approach. The study identified the patterns in the operating speed profiles on the curves sites located in plain terrain. It will be interesting to identify the patterns of the operating speed profiles in rolling and mountainous terrain along with the downgrade and upgrade on the selected road sections.

References


