

Article

# Evaluation of Automatic Prediction of Small Horizontal Curve Attributes of Mountain Roads in GIS Environments

Sercan Gülci <sup>1,\*</sup> , Hafız Hulusi Acar <sup>2</sup> , Abdullah E. Akay <sup>3</sup>  and Neşe Gülci <sup>1</sup> 

<sup>1</sup> Department of Forest Engineering, Faculty of Forestry, Kahramanmaraş Sutcu Imam University, Kahramanmaraş 46040, Türkiye

<sup>2</sup> Department of Occupational Health and Safety, Faculty of Health Sciences, Istanbul Yeni Yüzyıl University, Zeytinburnu, Istanbul 34010, Türkiye

<sup>3</sup> Department of Forest Engineering, Faculty of Forestry, Bursa Technical University, Bursa 16310, Türkiye

\* Correspondence: [sgulci@ksu.edu.tr](mailto:sgulci@ksu.edu.tr); Tel.: +90-(344)-3001749

**Abstract:** Road curve attributes can be determined by using Geographic Information System (GIS) to be used in road vehicle traffic safety and planning studies. This study involves analyzing the GIS-based estimation accuracy in the length, radius and the number of small horizontal road curves on a two-lane rural road and a forest road. The prediction success of horizontal curve attributes was investigated using digitized raw and generalized/simplified road segments. Two different roads were examined, involving 20 test groups and two control groups, using 22 datasets obtained from digitized and surveyed roads based on satellite imagery, GIS estimates, and field measurements. Confusion matrix tables were also used to evaluate the prediction accuracy of horizontal curve geometry. F-score, Mathews Correlation Coefficient, Bookmaker Informedness and Balanced Accuracy were used to investigate the performance of test groups. The Kruskal–Wallis test was used to analyze the statistical relationships between the data. Compared to the Bezier generalization algorithm, the Douglas–Peucker algorithm showed the most accurate horizontal curve predictions at generalization tolerances of 0.8 m and 1 m. The results show that the generalization tolerance level contributes to the prediction accuracy of the number, curve radius, and length of the horizontal curves, which vary with the tolerance value. Thus, this study underlined the importance of calculating generalizations and tolerances following a manual road digitization.

**Keywords:** spatial data; data quality; field measurement; curve geometry; transportation; line generalization; low-cost



**Citation:** Gülci, S.; Acar, H.H.; Akay, A.E.; Gülci, N. Evaluation of Automatic Prediction of Small Horizontal Curve Attributes of Mountain Roads in GIS Environments. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 560. <https://doi.org/10.3390/ijgi11110560>

Academic Editor: Wolfgang Kainz

Received: 19 September 2022

Accepted: 4 November 2022

Published: 9 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

According to the World Health Organization (WHO), nearly 1.2 million people lose their lives every year in traffic accidents worldwide [1]. EU countries have organized national and international partnerships to reduce traffic accidents on highways and have adopted short- and long-term measures to minimize the number of these accidents (vision zero) [2]. In the next few years, research on the analysis of transportation characteristics will be promoted. Its goal is to reduce traffic accidents in EU countries to minimize them by 2050. Therefore, multi- and interdisciplinary studies on transportation systems and road networks have attracted attention in the scientific community to design more efficient transportation systems around the world.

Several factors (such as vehicle type, road type, environment, and the driver) influence the intensity of traffic accidents and road safety on a highway, directly or indirectly. Some of these factors have been revealed in different empirical and theoretical studies. Research on vehicle accidents or collisions shows that driver characteristics are the most important factors. A driver's psychomotor skills (i.e., reaction speed and eye, hand and foot coordination), mental characteristics (i.e., perception, attention, and memory), education level, driving experience, age, gender, habit, and personality traits (such as risk taking,

aggression, and responsibility) are directly effective variables in vehicle accidents and collisions [3–7]. Some studies benefited from temporal and spatial analysis using different variables to propose different road accident and safety models. In this respect, GIS is widely used for the determination of various variables affecting traffic accidents on a critically dangerous road [8–12]. Kernel density estimation (KDE) and Poisson’s method are also used for traffic accident analysis [13–15], and road geometric parameters are used in GIS environments. For the analysis of road safety and traffic accidents, much spatial data are integrated into different techniques in GIS environments in order to offer effective solutions to reduce and prevent the number of traffic accidents [8–12,16].

Road geometry (i.e., superelevation and lateral and horizontal curves) is a popular research topic as it is considered one of the other determinants of traffic accidents and collisions. The manual calculation of road curvatures may be costly for evaluating accident risks on a road network. Thanks to their cost-effective and time-saving nature, GIS techniques have become an important data production method in various technical and scientific topics (i.e., traffic and wild animal crashes, habitat fragmentation, and setting speed limits) for the prediction of geometric parameters in road networks [17–19]. In particular, the increasing number of GIS-assisted studies in the analysis of accidents occurring on road curves and in the estimation of horizontal curve characteristics offers great advantages in terms of time and cost [20]. For estimating geometric road curve characteristics, several software and add-ons are now available for the automatic, semi-automatic, or manual calculation of different types of road curves, such as “Curve Calculator” (ESRI, USA), “Curve Finder” [20,21], “Curvature Extension” [22], “Road Curvature Analyst” (ROCA) [23], and “CurvS” [24].

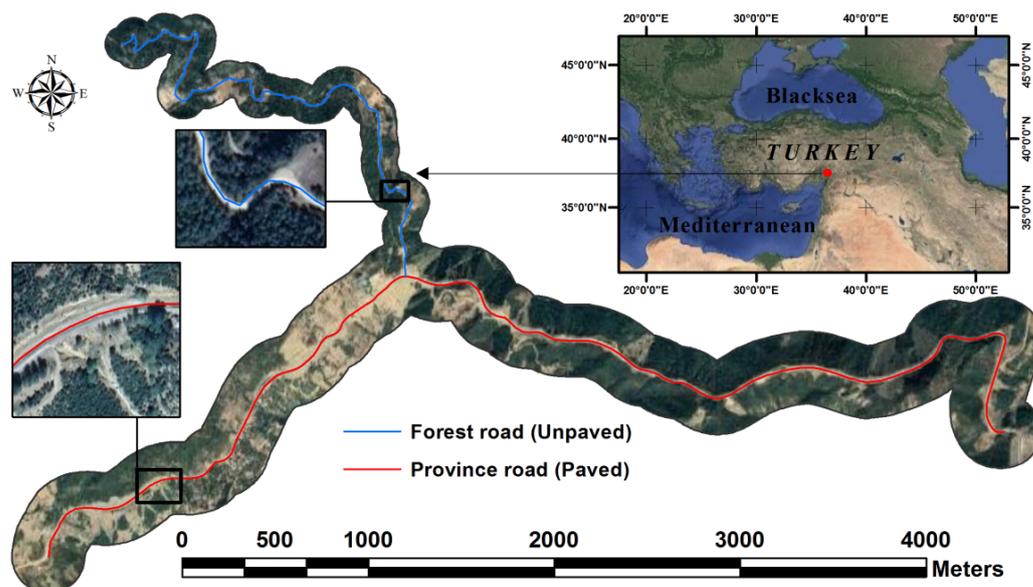
The horizontal curve as one of the geometric road characteristics obtained using low-cost methods often includes irrelevant spatial locations, which may affect the collected spatial data negatively. The impact of generalization in the prediction and calculation accuracy of the elements of the horizontal curve is quite decisive in GIS [25,26]. Thus, digital vector lines such as roads should be generalized/simplified for a higher prediction accuracy rate [17]. The general standards for digitized road networks in the prediction of the characteristics of the horizontal curve using the automatic and semi-automatic methods are usually set without any further details. However, these methods do not offer standard workflows regarding the road vector line obtained using different methods and the impact of generalization tolerance. The data related to linear structures such as a road network are stored in the digital line vector format. Road network vector data are obtained using different methods with different precision and spatial accuracy. Semi-automatic and automatic curve detection software require high-quality data to perform successfully [21]. Therefore, it is essential to pay attention to the performance and accuracy of the small horizontal curves radius analysis software, which will help GIS users to obtain reliable information about the road segments. Accordingly, it is important to deal with the factors that affect the geospatial accuracy of cartographic and digital data obtained from complex GIS analysis [27,28].

The main contribution of this study was to evaluate whether the automatically predicted horizontal curve geometry attributes in a GIS can be used for road geometry-vehicle-accident modeling, and whether there is an influence of generalization algorithms on the success of horizontal curve prediction. The aim of this study is to analyze the prediction accuracy for small radius horizontal road curves on two digitized roads. An automatic road curve calculation tool, which was used in a GIS environment, was performed and tested for small horizontal road curves. The radius and length of the horizontal curve were also measured and calculated in the field. As the road was a digital line vector, the impact of the generalization factor and the precision of the algorithm for the dataset were also evaluated. Then, using the horizontal curve information obtained from the GIS, the two different roads were grouped and statistically compared. Briefly, the main purpose was also to provide secure, easy, fast, and effective data generation in detecting dangerous horizontal curves using GIS techniques.

## 2. Materials and Methods

### 2.1. Study Area

In this study, two-lane rural and forest road sections with sharp and dangerous curves, low geometric standards, and a low traffic volume were investigated. These two roads are located within the borders of Andirin district in the province of Kahramanmaraş, in the eastern Mediterranean Region of Türkiye. The lengths of low-volume roads in this study, i.e., a rural road and a forest road segment, are approximately  $\cong 6900$  m and  $\cong 3400$  m, respectively, and they are located in a mountainous region with a steep terrain structure (Figure 1).



**Figure 1.** The geographical location of the sample roads and neighboring land use.

### 2.2. Equipment

Leica Disto s910 was used for the measurement of geometric road characteristics, and Garmin Oregon 600 handheld GPS was used to collect spatial data. A drone (DJI Phantom 4) and a 12 MP camera were also used for bird's-eye view images of the study area [29]. ArcGIS software and its add-on (i.e., ROCA) were used for the prediction of curve attributes, storage and processing of vector data, and mapping. An i7 16 Gigabyte (GB) random-access memory (RAM) desktop computer was used for data processing.

A 1/25,000 topographic map and Google Earth satellite images were used as a basis for the digitization of both roads and manual measurement of curve elements. The curves were detected in ArcGIS with Road Curvature Analyst (ROCA) add-on [23] and analyzed using spatial join, intersection, overlay processing, and mapping [30].

### 2.3. Data and Database Preparation

The roads on Google Earth satellite images were digitized in a GIS environment at a drawing scale of 1/4000 by a user with an experience of at least 10 years. Then, the digitized roads were checked via a topographic map with a scale of 1/25,000. WGS84 UTM Zone 37 was used as a spatial projection system. The obtained data were processed and stored as ShapeFile (\*.shp) data.

The horizontal curve data, which were manually obtained from the satellite images (ASat), were attached to feature tables using a field calculator tool. It was necessary to simplify digital line vectors in order to remove drawing vertices on road segments, which were digitized manually in a GIS environment. Therefore, in order to reveal the impact of generalization on the automatic calculation of horizontal curves, digital line vector data with and without generalizations were produced (ANoT) and divided into different data

groups. As a result, 20 different digital line vector data were produced for the rural (A) and forest (B) roads. Douglas-Peucker and Bezier generalization algorithms were used in order to delete drawing vertices while considering a generalization tolerance of 20 cm, 40 cm, 60 cm, 80 cm, 1 m, 2 m, 3 m, and 4 m. The highest horizontal curve radius values to be calculated for the paved rural road (A) and forest road (B) were set to 400 m and 100 m, respectively. Figure 2 shows the general concepts of this study.

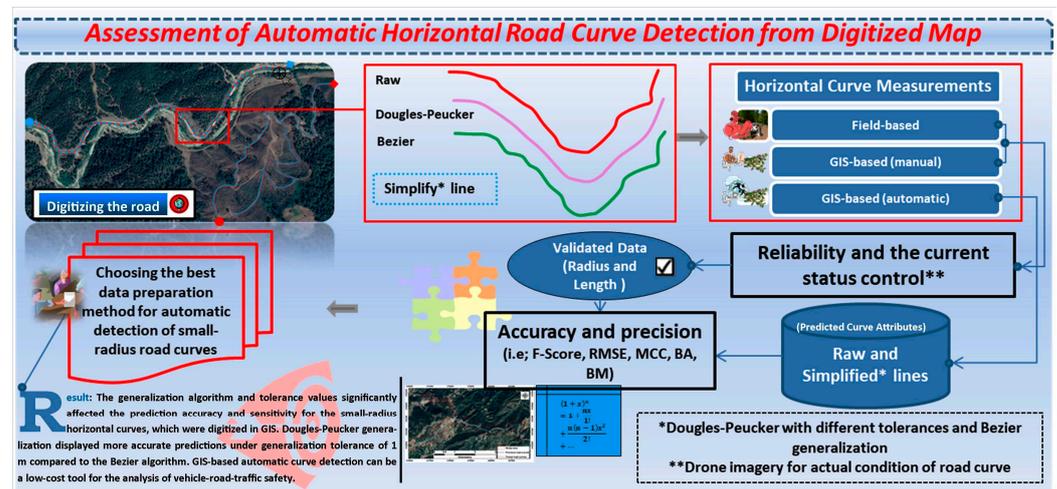


Figure 2. General workflow and brief explanation.

### 2.4. Measuring Horizontal Curves

Satellite image and digital line vectors were overlaid in a GIS environment. Subsequently, the starting point (PC) and ending point (PT) of the curvature were found and marked on the high-resolution image, and the digital point layer was saved. PC and PT were combined to create a chord (C). The chord midpoint (MC) was attached to the farthest point (tangent of curve = ToC) on the road curve. Thus, a polyline was drawn in order to determine a middle ordinate (MO), which is the distance between MC and ToC on the curvature. Then, the relationship between C length (CL), MO, and radius (R) were calculated (Figure 3). Finally, the data were defined in the attribute tables of the related layers.

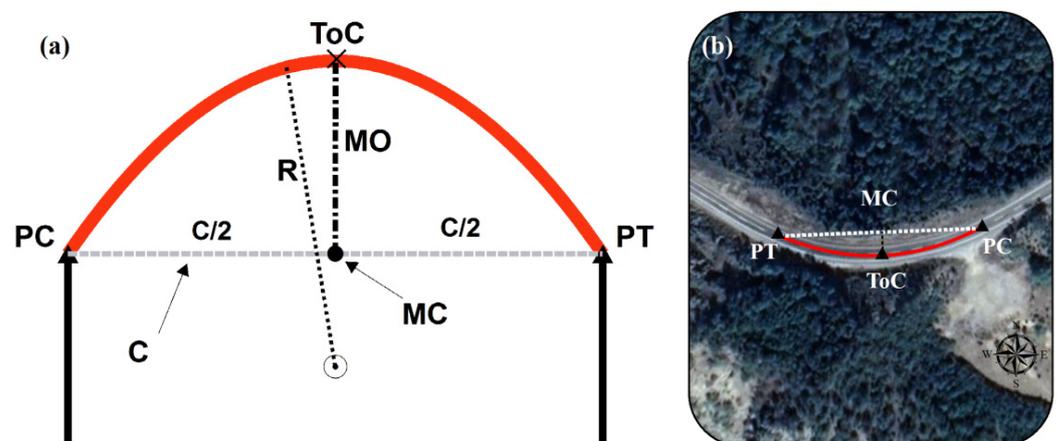


Figure 3. The geometric structure and elements of a horizontal curve (a) and the position of the curve on the satellite image (b).

Straight roads (tangents) and curves were separated from each other using ArcGIS based on PC and PT points. Afterwards, horizontal curve lengths were calculated in the

feature database, and a mathematical radius equation (Equation (1)) was used to calculate the radius values [31].

$$Radius = \frac{CL^2}{8 \times MO} + \frac{MO}{2} \quad (1)$$

### 2.5. Field Measurements

Field measurements were performed for each horizontal curve on the road thanks to a laser range finder. A handheld GPS was used to specify the location on the field map. “Chord length” method was used for calculations, where three opposite measurements were performed in the inner and outer sides of the curve between PC, ToC, and PT points on the horizontal curve (Figure 3). Ambiguous points on the beginning and end of the curve were checked using a drone at an altitude lower than 120 m, as specified by the law (Figure 4). In addition, current conditions of both roads were checked.



**Figure 4.** The latest drone images for the curve and sharp curve on the rural road (a) and forest road (b).

All curve points were calculated based on the points on a helical spring instead of adjacent tangent lines. The point data of the measured curves were obtained from opposite road sidelines. Two different opposite measurements were performed on the beginning and end of the curve to find the actual curve radius and the middle point of the road. Then, the curve on the road midpoint was used to calculate the curve radius. The value calculated using this method was assumed to be the actual radius and compared with the value predicted using satellite images [31,32].

### 2.6. Main Concept of Automatic Curve Detection Tool

In this study, ROCA, which involves the idea of the naïve Bayes classifier, was examined as a toolbox for ArcGIS [23]. The main steps include digitized road data generalization, computation of explanatory variables, and generalized road network data analysis using training and testing datasets. It is also possible to create different training datasets for the GIS add-on. The model then consists of a classification process, radii calculation, and the use of heuristics [17]. More details about the ROCA add-on and horizontal curve analysis training data can be found at <https://roca.cdvinfo.cz/downloads/> (accessed on 19 September 2022).

### 2.7. Comparison of Data

Field- and GIS-based measurements played a reference role in observing the current conditions on the study area and predicting the curve radius and length accurately. To validate GIS-based manually measured curve attributes, total root mean square error (RMSE) values were considered as an error metric (Equation (2)). Linearity was also considered for the relationship between the radius and length of the road curve for field and image measurement. Thus, the differences between radius and length values in control

and field measurements were defined on a mathematical basis. This step was performed to ensure the reliable calculation of GIS-based automatic curve attribute prediction.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(Measured - Estimated)^2}{n}} \quad (2)$$

Sampling units with a resolution of  $10 \times 10$  m were designed for an evaluation of the prediction performances of GIS-based automatic curve detection. The road and the sample units were intersected to calculate the sensitivity of the predicted curve characteristics. A true/false analysis was performed in the confusion matrix table to calculate the curve prediction sensitivity in sampling units via Equations (3)–(6). Thus, the prediction accuracy was evaluated using *F-Score* analysis as true positive (*TP*), false positive (*FP* = commission error), and false negative (*FN* = omission error). *F-Score* was calculated as given in Equation (3). In this formula, *TP* represents correctly detected curve number, *FP* denotes extra curve segment that is absent in the area, and *FN* refers to curves that are present in the study area but cannot be detected in the study area. As *F-Score* values are sometimes likely to be biased, other metric measurements were also taken into account [33]. Therefore, Mathews Correlation Coefficient (*MCC*) (Equation (4)) [34], Balanced Accuracy (*BA*) (Equation (5)) and Bookmaker Informedness (*BM*) (Equation (6)) were also used as metric criteria.

$$F - Score = 2 \times \frac{\left(\frac{n_{TP}}{n_{TP}+n_{FN}}\right) \times \left(\frac{n_{TP}}{n_{TP}+n_{FP}}\right)}{\left(\frac{n_{TP}}{n_{TP}+n_{FN}}\right) + \left(\frac{n_{TP}}{n_{TP}+n_{FP}}\right)} \quad (3)$$

$$MCC = \frac{n_{TP} \times n_{TN} - n_{FP} \times n_{FN}}{\sqrt{(n_{TP} + n_{FP}) \times (n_{FP} + n_{FN}) \times (n_{TN} + n_{FP}) \times (n_{TN} + n_{FN})}} \quad (4)$$

$$BA = \frac{\frac{n_{TP}}{n_{TP}+n_{FN}} + \frac{n_{TN}}{n_{TN}+n_{FP}}}{2} \quad (5)$$

$$BM = \frac{n_{TP}}{n_{TP} + n_{FN}} + \frac{n_{TN}}{n_{TN} + n_{FP}} - 1 \quad (6)$$

Homogeneity and normality tests were applied to the data groups. When they displayed a homogenous distribution, ANOVA ( $p < 0.05$ ) and Kruskal–Wallis was used for group comparison and non-parametric group data. The following statements were hypothesized to compare the horizontal curve radius and length values.

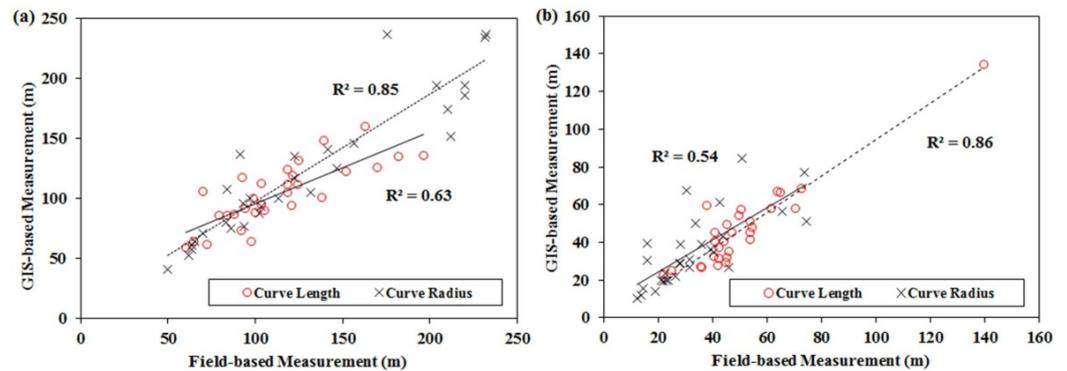
- H0: The curve radii in different groups for the rural road do not differ from each other significantly.
- H1: The curve length values in different groups for the rural road differ from each other significantly.
- H0<sub>a</sub>: The curve radii values in different groups for the forest road do not differ from each other significantly.
- H1<sub>a</sub>: The curve length values in different groups for the forest road differ from each other significantly.

### 3. Results

#### 3.1. The Comparison of Curve Data Obtained from Field- and GIS-Based Measurements

In this study, curve radius and length values obtained from GIS-assisted high resolution satellite images were compared with field measurements (Figure 5). GIS-measured and field measurements for the length and radius of the horizontal curve of both rural (Figure 5a) and forest (Figure 5b) roads show line goodness of fit with  $R^2$ . Total RMSE, which is a parametric value, was calculated for 30 curve values measured on the rural road and 29 curve values measured on the forest road. RMSE for curve radius and length on the rural road were calculated as 22.86 m and 22.99 m, respectively. Alternatively, RMSE for curve radius and length on the forest road were calculated as 16.17 m and 8.63 m, respectively. It was later discovered that some maintenance and repair work was done

in the forest road (such as widening on some curves). Similarly, some asphalt patchwork has been applied to the rural road, which did not lead to any changes in its geometric structure and alignment. Curve calculation and measurement by using GIS-supported satellite images took less time than field studies.



**Figure 5.** The comparison of rural road (a) and forest road (b) curve data obtained from field measurement and GIS-assisted satellite images.

### 3.2. Calculation of Rural Road Curves

As the groups were evaluated based on different generalization tolerance values, different curve values were found in each method, except those with generalization tolerance of 20 cm and 40 cm. Although mean curve radii were different in the data groups except those with generalization tolerance of 20 cm and 40 cm, mean curve length values were different in all groups. Mean curve length values in the groups with generalization tolerance of 20 cm and 40 cm calculated in GIS were partially similar. The number of ASat (satellite-based curve measurements from paved road), A20cm (Douglas-Peucker method with generalization tolerance of 20 cm), and A40cm was 30 curves. The predicted minimum and maximum radius and length values were similar in the data groups with generalization tolerance in centimeters. Based on the data from Bezier algorithm generalization, the number of calculated curves (ABez) was much higher compared to the predicted curves. The statistical curve data calculated and predicted in GIS for the rural road are summarized in Table 1.

**Table 1.** The statistical curve data groups calculated and predicted for the rural road.

Group	Curve N	Min. Radius	Mean Radius	Max. Radius	Min. Length	Mean Length	Max. Length
ASat	30	40.33	122.32	236.63	58.67	103.33	159.90
ANoT	33	31.47	131.34	303.62	31.45	105.34	233.40
A20cm	30	58.47	128.07	253.65	40.07	147.50	288.58
A40cm	30	58.47	128.37	253.65	40.07	152.53	288.58
A60cm	27	60.67	140.27	253.65	40.08	161.73	288.58
A80cm	26	61.96	141.10	253.65	40.08	157.87	288.58
A1m	27	41.97	151.09	303.38	58.40	127.10	245.92
A2m	23	43.42	150.05	326.98	60.31	144.25	283.59
A3m	15	63.33	129.51	314.81	47.01	151.49	267.54
A4m	13	84.42	149.54	298.46	107.24	203.34	365.38
ABez	78	29.79	142.19	357.00	12.33	53.73	119.30

ASat stands for: A = rural road and Sat = control data obtained from the satellite images.

### 3.3. The Calculation of Forest Road Curves

Similar to the rural road, the curve numbers, radii and lengths in different data groups were also calculated for the forest road. Since this study focuses on curves with a maximum

curve radius of 100 m, the statistical data related to the control groups are given in Table 2. It can be seen that different numbers of curves and curve length values were calculated because of generalization values on the road line, except those without generalizations and with a generalization tolerance of 20 cm, 60 cm and 1 m. The statistical relationship analysis between the curve radii calculated using satellite images and digitized line vectors with generalization tolerance of 20 cm, 40 cm, 60 cm and 1 m yielded similar predictions with partial mean differences. The highest number of curves was predicted on the road with Bezier generalization. On the other hand, the number of curves predicted on the roads with generalization tolerance of 2 m or higher was lower.

**Table 2.** The statistical data related to the predicted and calculated curves on the forest road.

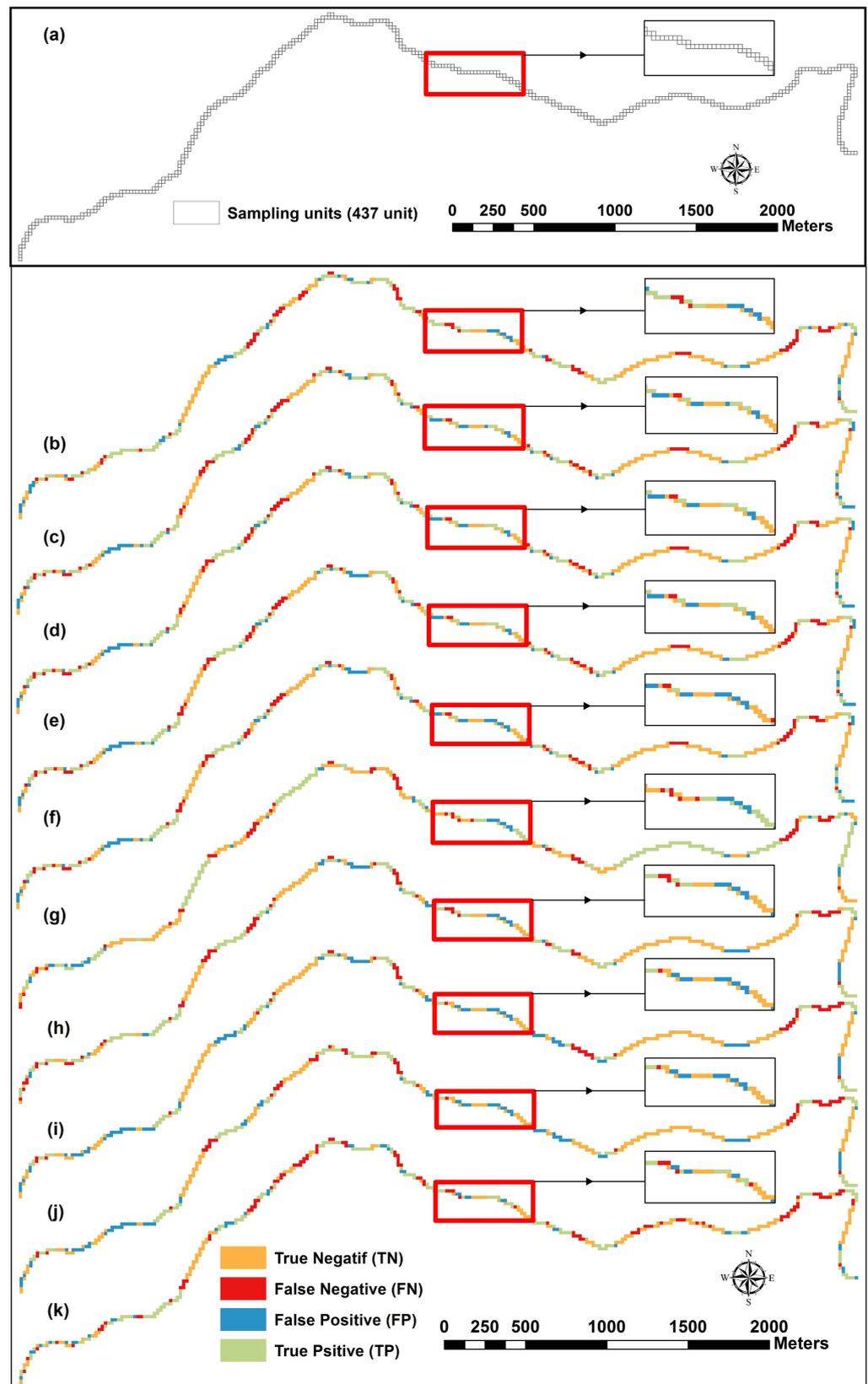
Group	Curve N	Min. Radius	Mean Radius	Max. Radius	Min. Length	Mean Length	Max. Length
BSat *	29	10.50	35.46	84.64	22.95	46.87	134.65
BNoT	29	12.54	33.26	70.95	17.86	48.01	120.90
B20cm	27	12.54	35.71	70.95	17.86	51.11	120.90
B40cm	26	12.54	36.46	70.95	17.86	52.73	120.90
B60cm	27	12.49	39.21	75.23	17.86	54.65	120.90
B80cm	25	12.54	39.84	82.86	17.86	63.60	186.85
B1M	29	12.54	33.26	70.95	17.86	48.01	120.90
B2M	15	14.58	36.01	73.76	28.83	68.44	117.47
B3M	10	14.58	39.73	70.43	55.04	98.45	196.53
B4M	7	21.71	47.81	79.85	64.78	109.00	193.27
BBez	68	5.40	36.32	72.26	6.21	25.61	89.50

\* BSat stands for: B = forest road and Sat = control data obtained from the satellite images.

### 3.4. The Prediction Accuracy for the Rural Road

The roads selected for the test groups were divided into 437 equal sampling units for the rural road (Figure 6). An error matrix was created in the GIS environment for the prediction accuracy rates in all groups. The rural roads were divided into sampling units in order to see any curve predictions in each unit. The spatial distributions of true-false analysis for the rural road were shown in Figure 6. We observed that the accuracy rates of curve prediction in the groups (Figure 6b–f,h–k) with and without generalization tolerance increased significantly in small curve standards, except for those with generalization tolerance of 1 m (Figure 6g).

F-Score, (Equation (3)), Mathews Correlation Coefficient (MCC) (Equation (4)), Balanced Accuracy (BA) (Equation (5)) and Bookmaker Informedness (BM) (Equation (6)) were taken into account in the analysis. Based on these metrics, the curves measured manually on satellite images and the prediction accuracy rates of 10 test groups obtained from ROCA analysis of curves with and without a generalization tolerance value were calculated. The results demonstrated that the curve data using GIS environment displayed different prediction accuracy rates depending on different generalization coefficients and algorithms. It was also observed that the highest prediction accuracy rates in the digitized road curves were obtained with generalization tolerance of 20 cm and 1 m using Douglas-Peucker method (Table 3). The most successful test group data, on the other hand, belonged to Douglas-Peucker method, with generalization tolerance of 1 m (A1m).



**Figure 6.** Sampling units ( $10 \times 10$  m) for the control and test groups in the study area (a). Spatial data accuracy without generalization tolerance (b), Douglas-Peucker generalization tolerance values: 20 cm (c), 40 cm (d), 60 cm (e), 80 cm (f), 1 m (g), 2 m (h), 3 m (i), 4 m (j) and the map displaying true/false distributions for the forest road with Bezier generalization algorithm (k).

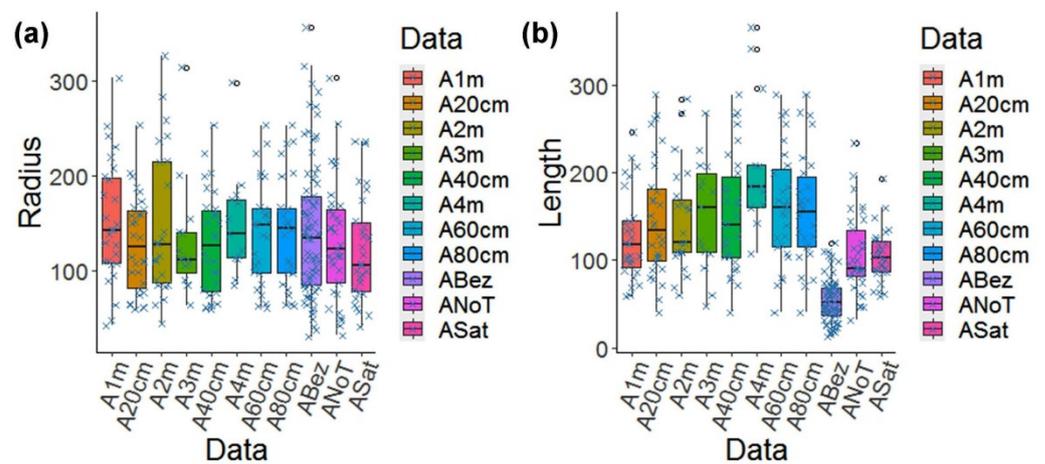
**Table 3.** The comparison of curve prediction accuracy rates for the rural road in terms of generalization tolerance values.

Group (A)	TN	FN	FP	TP	N	Recall	Precision	F-Score	MCC	BA	BM
ANoT-ASat	168	81	51	137	437	0.63	0.73	0.67	0.40	0.70	0.40
A20cm-ASat	167	82	62	126	437	0.61	0.67	0.64	0.34	0.67	0.34
A40cm-ASat	169	80	59	129	437	0.62	0.69	0.65	0.36	0.68	0.36
A60cm-ASat	167	82	64	124	437	0.60	0.66	0.63	0.33	0.66	0.32
A80cm-ASat	169	80	69	119	437	0.60	0.63	0.61	0.31	0.65	0.31
A1m-ASat	181	68	40	148	437	0.69 *	0.79 *	0.73 *	0.51 *	0.75 *	0.50 *
A2m-ASat	178	71	49	139	437	0.66	0.74	0.70	0.45	0.72	0.45
A3m-ASat	193	56	104	84	437	0.60	0.45	0.51	0.24	0.62	0.25
A4m-ASat	188	61	88	100	437	0.62	0.53	0.57	0.29	0.65	0.30
ABez-ASat	139	110	46	142	437	0.56	0.76	0.65	0.31	0.66	0.31

\* is the highest score.

3.5. The Statistical Relationships between Control and Test Groups for the Rural Road Curves

The curve radius obtained from the data groups on the rural road displayed a homogeneous distribution ( $p > 0.05$ ,  $p = 0.752$ ,  $F = 0.620$ ), which verifies  $H_0$ . In other words, no statistically significant differences were observed among the data groups. However, as far as the mean values are concerned, a statistically significant difference was observed between the data groups with and without Douglas-Peucker and Bezier generalization (Figure 7).



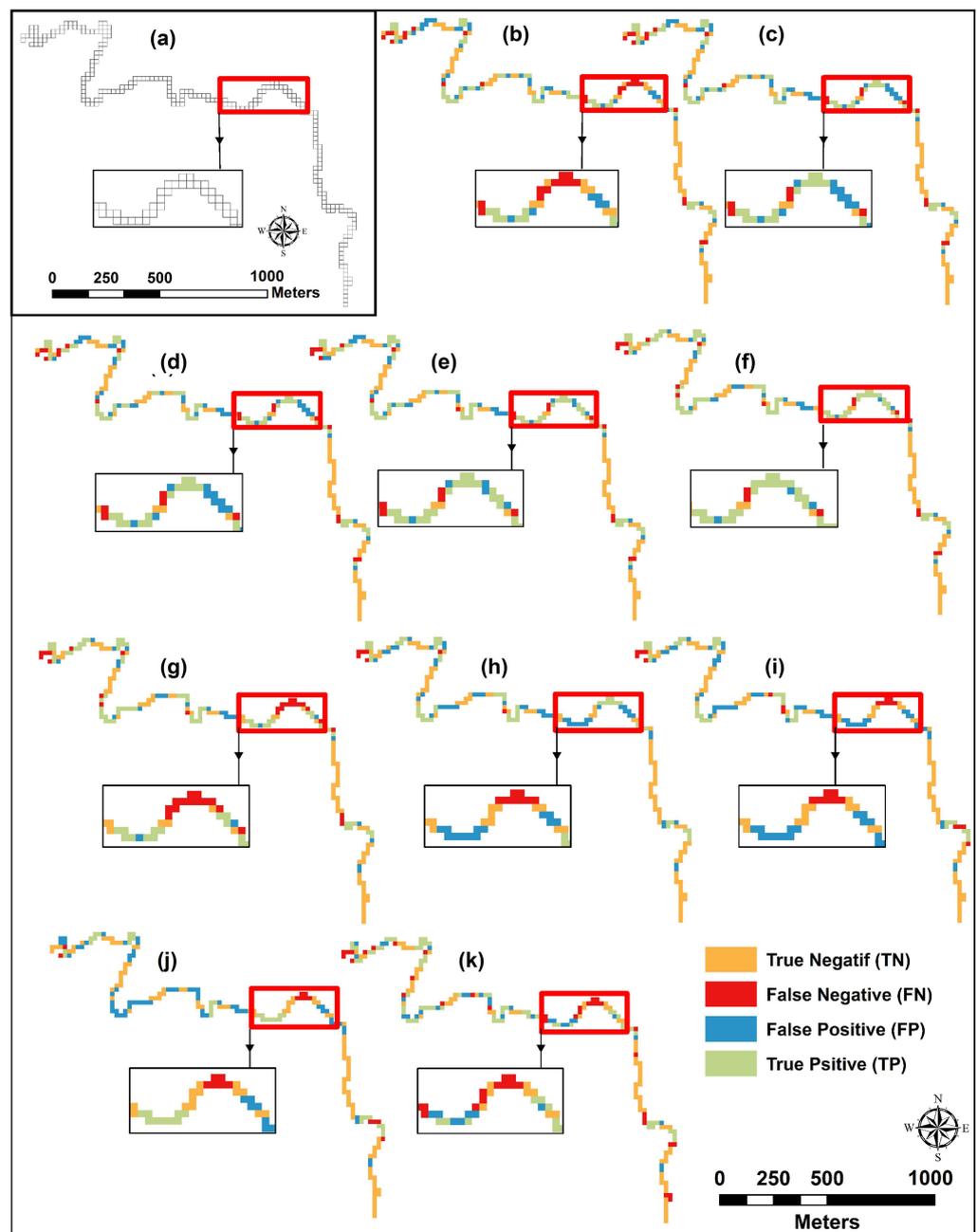
**Figure 7.** The mean curve radius (a) and length (b) values for the rural road and a box plot view of outliers.

Similar to the forest road, the curve length values for the rural road did not display a homogenous distribution. Therefore, Kruskal-Wallis test was applied instead of ANOVA analysis. The test results ( $\chi^2 = 158.561$ ,  $p < 0.005$ ) indicated a statistically significant difference among the data groups, which contradicts  $H_1$ . Thus, it can be stated that a statistically significant difference was found among the data groups. The standard deviations among control groups, i.e., ASat, A20cm, A1m and ANoT, were partially similar. Therefore, when the standard deviations in all groups were calculated, a statistically significant difference was observed among all groups (Figure 6).

The difference between the mean radius length calculated using ANoT and ASat for the rural road ( $\pm 2.01$  m) was lower compared to other control groups, whereas the difference between the mean curve radius calculated using A20cm and ASat ( $\pm 5.75$  m) was lower compared to other groups.

### 3.6. The Prediction Accuracy for the Forest Road

The spatial distribution of true/false analysis for the forest road is shown in Figure 8. Two hundred and thirteen (213) sample sites were selected for the forest road (Figure 8a). An error matrix was produced in a GIS environment for the prediction accuracy rates in all data groups. The data groups with generalization tolerances of 20 cm and 40 cm in the sampling units modeled in GIS displayed similar sensitivity for road alignment and curve predictions (Figure 8c,d). However, it was observed that the road alignment and curve prediction accuracy rates decreased for generalization tolerance higher than 1 m and Bezier algorithm generalization (Figure 8h–k). The sampling units may display differences in the map analysis.



**Figure 8.** Sampling units ( $10 \times 10$  m) for the control and test groups in the study area (a). Spatial data accuracy without generalization tolerance (b), Douglas-Peucker generalization tolerance values: 20 cm (c), 40 cm (d), 60 cm (e), 80 cm (f), 1 m (g), 2 m (h), 3 m (i), 4 m (j) and the map displaying true/false distributions for the forest road with Bezier algorithm generalization (k).

Similar to the rural road, since the statistical analysis for the forest road is limited to a maximum curve radius of 100 m, the highest prediction accuracy rates were obtained in the groups with generalization tolerance of 80 cm and 1 m using Douglas-Peucker algorithm. Additionally, the curve prediction accuracy rates in the data groups without generalization and with generalization tolerance of 20 cm and 40 cm were similar to each other. According to true/false analysis of the sampling units, the data groups with generalization tolerance of 0 m and 2 m displayed a higher prediction accuracy rate compared to other generalization tolerance values (Table 4).

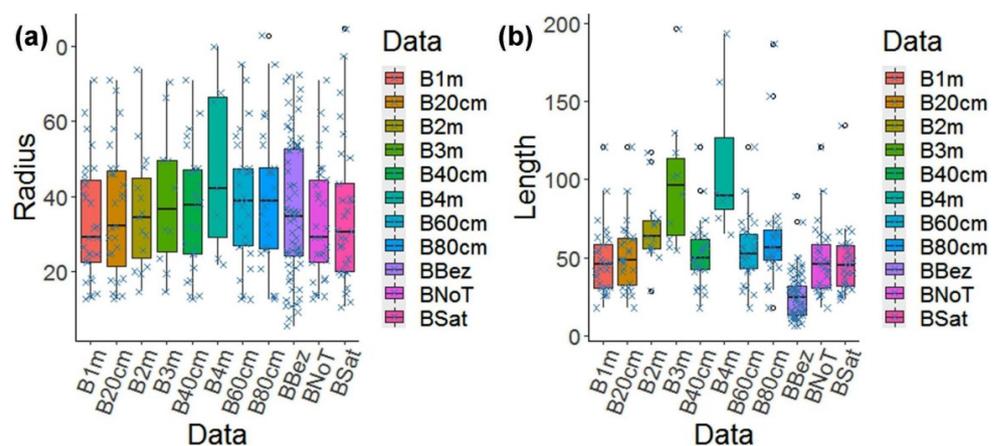
**Table 4.** The comparison of curve prediction accuracy rates for the forest road in terms of generalization tolerance values.

Group (B)	TN	FN	FP	TP	N	Recall	Precision	F-Score	MCC	BA	BM
BNoT-BSat	85	19	41	68	213	0.78	0.62	0.69	0.45	0.73	0.46
B20cm-BSat	84	20	42	67	213	0.77	0.61	0.68	0.43	0.72	0.44
B40cm-BSat	85	19	42	67	213	0.78	0.61	0.69	0.44	0.72	0.45
B60cm-BSat	85	19	35	74	213	0.80	0.68	0.73	0.50	0.75	0.50
B80cm-BSat	86	18	26	83	213	0.82	0.76 *	0.79 *	0.59	0.79	0.59
B1m-BSat	87	17	26	83	213	0.83	0.76 *	0.79 *	0.60 *	0.80 *	0.60 *
B2m-BSat	97	7	53	56	213	0.89 *	0.51	0.65	0.49	0.77	0.54
B3m-BSat	92	12	62	47	213	0.80	0.43	0.56	0.35	0.70	0.39
B4m-BSat	98	6	69	40	213	0.87	0.37	0.52	0.38	0.73	0.46
BBez-BSat	74	30	35	74	213	0.71	0.68	0.69	0.39	0.70	0.39

\* is the highest score.

### 3.7. The Statistical Relationships between Control and Test Groups for the Forest Road Curves

The curve radii obtained from the data groups for the forest road displayed a homogeneous distribution ( $p > 0.05$ ,  $p = 0.733$ ,  $F = 0.691$ ), which verifies  $H0_a$ . The statistical difference between the average values in the control group data increases in proportion to the Douglas-Peucker algorithm. In line with this, the most similar mean curve values were observed in the data groups with generalization tolerance of 1 m (Figure 8). A statistically significant difference was observed among the mean values because of similar standard deviations in all data groups. The curve length values among the data groups did not display a homogenous distribution, and Kruskal-Wallis was applied instead of ANOVA. The test results ( $\chi^2 = 127.152$ ,  $p < 0.005$ ) contradicted  $H1_a$ . Thus, no statistical relationships were found among the curve length values in the data groups. As far as mean values are concerned, the data groups without generalization and with generalization tolerance of 0.2 m and 1 m using Douglas-Peucker algorithm were more similar compared to the curve length values in the other data groups (Figure 9).



**Figure 9.** The mean curve radius (a) and length (b) values for the forest road and a box plot view of outliers.

On the other hand, the difference between the mean curve length values calculated using BNoT and B1m for the forest road ( $\pm 1.14$  m), which had a lower horizontal curve radius, was lower compared to other control groups. However, compared with the other groups, the difference between the average curve radii calculated using B20cm and BSat ( $\pm 0.25$  m) is lower.

#### 4. Discussion

The assessment of this study was carried out with the use of existing road digitizing data according to the estimation of success in the road curve parameters. We suggested accuracy control method with different generalization/simplified effect on the automatic estimation of road attributes (curve radius, length, and number).

It usually takes a long time and a high cost to use the manual/traditional methods to perform on-site measurements of the road network. Photogrammetry, aerial laser scanners, and terrestrial laser scanners performing in the study area, which also increased the number of high-cost and time-consuming steps in the field measurement, will contribute greatly to the field measurements [35]. Alternatively, vision learning techniques can detect the road curve characteristics from aerial or satellite images [36]. However, curve geometry detection by using the current advanced computer vision techniques for large-scale road networks are still limited. In short, all techniques used in curve detection have pros and cons [16]. Analyzing vector data in digitized road networks in a GIS needs lower-capacity processors and costs rather than advanced image processing techniques and expensive surveying equipment.

There is no doubt that the prediction accuracy of automatic and semi-automatic detection for geometric road elements is affected by spatial data measurement technology. Ensuring the accuracy, resolution, scale, and precision of the data used in the GIS is crucial for obtaining geometric characteristics of roads. Because the measurements of the geometric characteristics of roads are obtained from light detection and ranging (LiDAR), such a high-spatial accuracy and precise tool, they are still controversial today [37–39]. GIS-assisted curve detection, which is still an alternative to LiDAR technology, is a cost-effective and time-saving technique. Innovative and traditional methods (Digital imagery-, map-, and Global positioning system-based road mapping) [40–42] also run the risk of producing relatively low-quality and less sensitive data, making it necessary to conduct detailed studies on the feasibility.

The use of technology in scientific research makes it easier to obtain data from the field. For example, several types of geometric road features, such as curve radius, length, midpoint, and angle, can be automatically generated on the road sampling unit within a few seconds [43]. When compared with the duration and cost of field measurements, the automatic calculation of horizontal curve data in GIS has great advantages for experts studying road vehicle safety.

The number of spatial points in the calculation of prediction accuracy and sensitivity for horizontal curve radius in GIS was directly proportional to the prediction accuracy of Curve Calculator, whereas Curve Finder and Curvature Extension were not affected by the sequence of spatial points (vertex) [28,44]. It was observed in the sensitivity analysis of ROCA that generalization algorithms and tolerances have a significant impact on the results (Tables 3 and 4) [17]. Therefore, for the comparison of the performance of curve geometry estimation, it can be explained that it is more effective to perform on-site measurement along the route instead of at a single point. The Douglas–Peucker algorithm, which was used in the analysis of small radius curves in this study, increased the curve prediction accuracy rate for the digital vector line, with generalization tolerance of 20 cm and 1 m. Different performance metrics (such as F-Score, MCC, BA, and BM) demonstrated that the prediction accuracy of the Bezier algorithm for mean curve radius and length values was relatively low. This may be attributed to the frequency of points applied in the GIS or the impact of the drawing scale. In this respect, a bigger drawing scale ( $>1/4000$ ) for the road network lines in mountainous areas is likely to increase the prediction performance. For

the statistical relationship between the test and the control groups, there is a statistically significant relationship between the calculated horizontal curve radius, which verifies the  $H_0$  and  $H_{0a}$ . However, no statistically significant relationship was observed between the curve lengths, which contradicted the  $H_0$  and  $H_{0a}$  of the two roads.

The performance of software in the automatic or semi-automatic calculation of horizontal curve data in a GIS environment has been evaluated in some studies. The prediction accuracy rates of models such as Curve calculator, Curve Finder, Curvature extension, and ROCA are reported as 78%, 69%, 80%, and 95%, respectively [23,28]. According to the results of the smaller radius curve detection, the detection performance of ROCA has a downward trend. The trained data, which was provided by the owner of the extension, may be revised for small curve radii. We assume that more reliable results can then be obtained. We observed that the increasing number of curves detected using the digital line vector with Bezier algorithm generalization increased the prediction sensitivity in small curves. Douglas–Peucker with preferred tolerance value is crucial in the estimation accuracy of geometric parameters. Due to high tolerance value usage during the generalization process, very close road segments with small radii can result in self-intersection problems [26]. Further, the estimated curve length errors are larger than the estimated radius. Thus, this study may show different performance in the detection and prediction of the digitized road networks on a straight or flat alignment or horizontal road curve [45].

## 5. Conclusions and Future Work

This study shows that the automatic curve detection based on the GIS concept and the point density of the digitized lines representing the road lanes directly affect the prediction accuracy and sensitivity of the detection of small-radius horizontal curves. It was found that the prediction accuracy for the number of horizontal curves, along with curve radius and length values, varied depending on different tolerance values. Therefore, different generalization tolerance values can be used to improve the prediction accuracy and sensitivity for curve radius and length. For small radius horizontal curves, the Douglas–Peucker algorithm outperforms Bezier.

The results of this study revealed that the generalization algorithm and tolerance values significantly affected the prediction accuracy and sensitivity for the small radius horizontal curves, which were digitized in GIS. It can be seen that in the horizontal curve detection performance of the two roads, MC and BM are more sensitive to errors than other metrics (Tables 3 and 4). In this respect, future studies must benefit from the metric criteria for the true/false analysis in their respective sample sites to find an optimal generalization method and tolerance. Thus, it will be more likely for them to predict curve geometry more accurately in accordance with a generalization tolerance value suitable for the specific objectives of their study. It is a practical and cost-effective approach that increases the potential of using road geometric data as a decision variable in GIS-based road traffic accident models. GIS-assisted tools, which automatically calculate the geometric parameters of small radius curves, are a promising technique for transportation-related studies. The automatic calculation of horizontal curve data using GIS tools has great advantages. However, the GIS-based method has a limitation of not effectively detecting the types of road curves. Data that can be used in the analysis of safe curve crossing speed estimates and transportation times of emergency response vehicles in the GIS environment can be produced on low radius roads where field work is difficult and dangerous. It can be added to vehicle navigation systems as road information datasets.

In a follow-up study, curves that are potentially dangerous for the road network will be identified. Then, the curve geometric information, which is considered as a variable, will be analyzed by vehicle type, road surface, speed, and rollover and skidding. As a result, by detecting horizontal curves that do not comply with the standards, it will be aimed at revealing a cost-effective method for taking necessary precautions for emergency and heavy vehicles.

**Author Contributions:** Conceptualization, Methodology, Software, Sercan Gülci; Formal Analysis, Sercan Gülci and Neşe Gülci; Investigation, Sercan Gülci and Hafız Hulusi Acar; Resources, Sercan Gülci and Neşe Gülci; Data Curation, Neşe Gülci and Hafız Hulusi Acar; Writing—Original Draft Preparation, Sercan Gülci; Abdullah E. Akay, Writing—Review & Editing, Sercan Gülci and Abdullah E. Akay; Visualization, Neşe Gülci and Sercan Gülci; Supervision, Hafız Hulusi Acar and Abdullah E. Akay; Project Administration, Sercan Gülci, Neşe Gülci and Hafız Hulusi Acar; Funding Acquisition, Sercan Gülci. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding for the publication process.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This study is a part of a project funded by The Scientific and Technological Research Council of Türkiye (TUBITAK) with the project number 120O954. The authors would like to thank CDV–Transport Research Centre for providing the non-commercial use of ROCA extension. The authors also wish to thank anonymous reviewers and editors for their insightful suggestions and comments, which led to an improved manuscript.

**Conflicts of Interest:** No potential conflict of interest was reported by the authors.

## References

- Toroyan, T.; Peden, M.M.; Iaych, K. WHO Launches Second Global Status Report on Road Safety. *Inj. Prev.* **2013**, *19*, 150. [[CrossRef](#)] [[PubMed](#)]
- European Commission. *EU Road Safety Policy Framework 2021–2030-Next Steps towards “Vision Zero.” SWD(2019) 283 Final*; European Commission: Brussels, Belgium, 2019; 32p.
- Rolison, J.J.; Regev, S.; Moutari, S.; Feeney, A. What Are the Factors That Contribute to Road Accidents? An Assessment of Law Enforcement Views, Ordinary Drivers’ Opinions, and Road Accident Records. *Accid. Anal. Prev.* **2018**, *115*, 11–24. [[CrossRef](#)] [[PubMed](#)]
- Vanlaar, W.; Yannis, G. Perception of Road Accident Causes. *Accid. Anal. Prev.* **2006**, *38*, 155–161. [[CrossRef](#)] [[PubMed](#)]
- Shankar, V.; Mannering, F.; Barfield, W. Statistical Analysis of Accident Severity on Rural Freeways. *Accid. Anal. Prev.* **1996**, *28*, 391–401. [[CrossRef](#)]
- Wang, B.; Hallmark, S.; Savolainen, P.; Dong, J. Crashes and Near-Crashes on Horizontal Curves along Rural Two-Lane Highways: Analysis of Naturalistic Driving Data. *J. Saf. Res.* **2017**, *63*, 163–169. [[CrossRef](#)]
- Kronprasert, N.; Boontan, K.; Kanha, P. Crash Prediction Models for Horizontal Curve Segments on Two-Lane Rural Roads in Thailand. *Sustainability* **2021**, *13*, 9011. [[CrossRef](#)]
- Yalcin, G.; Duzgun, H.S. Spatial Analysis of Two-Wheeled Vehicles Traffic Crashes: Osmaniye in Turkey. *KSCE J. Civ. Eng.* **2015**, *19*, 2225–2232. [[CrossRef](#)]
- Satria, R.; Castro, M. GIS Tools for Analyzing Accidents and Road Design: A Review. *Transp. Res. Procedia* **2016**, *18*, 242–247. [[CrossRef](#)]
- Dereli, M.A.; Erdogan, S. A New Model for Determining the Traffic Accident Black Spots Using GIS-Aided Spatial Statistical Methods. *Transp. Res. Part A Policy Pract.* **2017**, *103*, 106–117. [[CrossRef](#)]
- Rodrigues, D.S.; Ribeiro, P.J.G.; da Silva Nogueira, I.C. Safety Classification Using GIS in Decision-Making Process to Define Priority Road Interventions. *J. Transp. Geogr.* **2015**, *43*, 101–110. [[CrossRef](#)]
- Wang, C.; Li, S.; Shan, J. Non-Stationary Modeling of Microlevel Road-Curve Crash Frequency with Geographically Weighted Regression. *ISPRS Int. J. Geo Inf.* **2021**, *10*, 286. [[CrossRef](#)]
- Bíl, M.; Andrášik, R.; Sedoník, J. A Detailed Spatiotemporal Analysis of Traffic Crash Hotspots. *Appl. Geogr.* **2019**, *107*, 82–90. [[CrossRef](#)]
- Erdogan, S.; Yilmaz, I.; Baybura, T.; Gullu, M. Geographical Information Systems Aided Traffic Accident Analysis System Case Study: City of Afyonkarahisar. *Accid. Anal. Prev.* **2008**, *40*, 174–181. [[CrossRef](#)] [[PubMed](#)]
- Mohaymany, A.S.; Shahri, M.; Mirbagheri, B. GIS-Based Method for Detecting High-Crash-Risk Road Segments Using Network Kernel Density Estimation. *Geo Spat. Inf. Sci.* **2013**, *16*, 113–119. [[CrossRef](#)]
- Budzynski, M.; Jamroz, K.; Pyrchla, J.; Kustra, W.; Ingot, A.; Pyrchla, K. Automated Parameter Determination for Horizontal Curves for the Purposes of Road Safety Models with the Use of the Global Positioning System. *Geosciences* **2019**, *9*, 397. [[CrossRef](#)]
- Andrášik, R.; Bíl, M. Efficient Road Geometry Identification from Digital Vector Data. *J. Geogr. Syst.* **2016**, *18*, 249–264. [[CrossRef](#)]
- Xu, H.; Wei, D. Improved Identification and Calculation of Horizontal Curves with Geographic Information System Road Layers. *Transp. Res. Rec.* **2016**, *2595*, 50–58. [[CrossRef](#)]
- Ma, Q.; Yang, H.; Wang, Z.; Xie, K.; Yang, D. Modeling Crash Risk of Horizontal Curves Using Large-Scale Auto-Extracted Roadway Geometry Data. *Accid. Anal. Prev.* **2020**, *144*, 105669. [[CrossRef](#)]

20. Li, Z.; Chitturi, M.; Bill, A.; Noyce, D. Automated Identification and Extraction of Horizontal Curve Information from Geographic Information System Roadway Maps. *Transp. Res. Rec.* **2012**, *2291*, 80–92. [[CrossRef](#)]
21. Li, Z.; Chitturi, M.V.; Bill, A.R.; Zheng, D.; Noyce, D.A. Automated Extraction of Horizontal Curve Information for Low-Volume Roads. *Transp. Res. Rec. J. Transp. Res. Board* **2015**, *2472*, 172–184. [[CrossRef](#)]
22. Linear Referencing System (LRS) Handbook, Curvature Extension for ArcGIS. 2020. Available online: [https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/statistics/docs/lrs-handbook-20200000cd0628604ebc9e3de2a438395b08.pdf?sfvrsn=ca6083b2\\_2](https://fdotwww.blob.core.windows.net/sitefinity/docs/default-source/statistics/docs/lrs-handbook-20200000cd0628604ebc9e3de2a438395b08.pdf?sfvrsn=ca6083b2_2) (accessed on 19 September 2022).
23. Bil, M.; Andrášik, R.; Sedoník, J.; Cícha, V. ROCA—An ArcGIS Toolbox for Road Alignment Identification and Horizontal Curve Radii Computation. *PLoS ONE* **2018**, *13*, e0208407. [[CrossRef](#)] [[PubMed](#)]
24. Bartin, B.; Demiroglu, S.; Ozbay, K.; Jami, M. Automatic Identification of Roadway Horizontal Alignment Information Using Geographic Information System Data: *CuroS* Tool. *Transp. Res. Rec. J. Transp. Res. Board* **2022**, *2676*, 532–543. [[CrossRef](#)]
25. Weibel, R.; Dutton, G. Generalising Spatial Data and Dealing with Multiple Representations. *Geogr. Inf. Syst.* **1999**, *1*, 125–155.
26. Liu, B.; Liu, X.; Li, D.; Shi, Y.; Fernandez, G.; Wang, Y. A Vector Line Simplification Algorithm Based on the Douglas–Peucker Algorithm, Monotonic Chains and Dichotomy. *ISPRS Int. J. Geo Inf.* **2020**, *9*, 251. [[CrossRef](#)]
27. Veregin, H. Data Quality Parameters. *Geogr. Inf. Syst.* **1999**, *1*, 177–189.
28. Rasdorf, W.; Findley, D.J.; Zegeer, C.V.; Sundstrom, C.A.; Hummer, J.E. Evaluation of GIS Applications for Horizontal Curve Data Collection. *J. Comput. Civ. Eng.* **2012**, *26*, 191–203. [[CrossRef](#)]
29. DJI Phantom 4 Specs. Available online: <https://www.dji.com/phantom-4/info> (accessed on 10 January 2019).
30. Esri, R. *ArcGIS Desktop: Release 10*; Environmental Systems Research Institute: Redlands, CA, USA, 2011.
31. Price, M. Under Construction: Building and Calculating Turn Radii. *ArcUser Mag.* **2010**, *13*, 50–56.
32. Carlson, P.J.; Burris, M.; Black, K.; Rose, E.R. Comparison of Radius-Estimating Techniques for Horizontal Curves. *Transp. Res. Rec.* **2005**, *1918*, 76–83. [[CrossRef](#)]
33. Chicco, D.; Jurman, G. The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genom.* **2020**, *21*, 6. [[CrossRef](#)]
34. Matthews, B.W. Comparison of the Predicted and Observed Secondary Structure of T4 Phage Lysozyme. *BBA Protein Struct.* **1975**, *405*, 442–451. [[CrossRef](#)]
35. Rodríguez-Cuenca, B.; García-Cortés, S.; Ordóñez, C.; Alonso, M.C. Morphological Operations to Extract Urban Curbs in 3d Mls Point Clouds. *ISPRS Int. J. Geo Inf.* **2016**, *5*, 93. [[CrossRef](#)]
36. Kuo, C.L.; Tsai, M.H. Road Characteristics Detection Based on Joint Convolutional Neural Networks with Adaptive Squares. *ISPRS Int. J. Geo Inf.* **2021**, *10*, 377. [[CrossRef](#)]
37. Gargoum, S.; El-Basyouny, K.; Sabbagh, J. Automated Extraction of Horizontal Curve Attributes Using LiDAR Data. *Transp. Res. Rec.* **2018**, *2672*, 98–106. [[CrossRef](#)]
38. Gargoum, S.A.; El Basyouny, K. A Literature Synthesis of LiDAR Applications in Transportation: Feature Extraction and Geometric Assessments of Highways. *GIScience Remote Sens.* **2019**, *56*, 864–893. [[CrossRef](#)]
39. Gézero, L.; Antunes, C. Automated Road Curb Break Lines Extraction from Mobile LiDAR Point Clouds. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 476. [[CrossRef](#)]
40. Maboudi, M.; Amini, J.; Malihi, S.; Hahn, M. Integrating Fuzzy Object Based Image Analysis and Ant Colony Optimization for Road Extraction from Remotely Sensed Images. *ISPRS J. Photogramm. Remote Sens.* **2018**, *138*, 151–163. [[CrossRef](#)]
41. Yang, B.; Dong, Z.; Liu, Y.; Liang, F.; Wang, Y. Computing Multiple Aggregation Levels and Contextual Features for Road Facilities Recognition Using Mobile Laser Scanning Data. *ISPRS J. Photogramm. Remote Sens.* **2017**, *126*, 180–194. [[CrossRef](#)]
42. Zhang, Y.; Zhang, Z.; Huang, J.; She, T.; Deng, M.; Fan, H.; Xu, P.; Deng, X. A Hybrid Method to Incrementally Extract Road Networks Using Spatio-Temporal Trajectory Data. *ISPRS Int. J. Geo Inf.* **2020**, *9*, 186. [[CrossRef](#)]
43. Bil, M.; Andrášik, R.; Sedoník, J. Which Curves Are Dangerous? A Network-Wide Analysis of Traffic Crash and Infrastructure Data. *Transp. Res. Part A Policy Pract.* **2019**, *120*, 252–260. [[CrossRef](#)]
44. Findley, D.J.; Zegeer, C.V.; Sundstrom, C.A.; Hummer, J.E.; Rasdorf, W.; Fowler, T.J. Finding and Measuring Horizontal Curves in a Large Highway Network: A GIS Approach. *Public Work. Manag. Policy* **2012**, *17*, 189–211. [[CrossRef](#)]
45. Bogenreif, C.; Souleyrette, R.R.; Hans, Z. Identifying and Measuring Horizontal Curves and Related Effects on Highway Safety. *J. Transp. Saf. Secur.* **2012**, *4*, 179–192. [[CrossRef](#)]