

Research Article

Analyzing Traffic Crash Severity in Work Zones under Different Light Conditions

Xinxin Wei,¹ Xiang Shu,² Baoshan Huang,² Edward L. Taylor,³ and Huaxin Chen¹

¹School of Materials Science and Engineering, Chang'an University, Xi'an, Shaanxi 710061, China

²Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN 37996, USA

³Construction Industry Research & Policy Center, Haslam College of Business, University of Tennessee, Knoxville, TN 37996, USA

Correspondence should be addressed to Xiang Shu; xshu@utk.edu and Huaxin Chen; chx92070@163.com

Received 6 March 2017; Revised 24 July 2017; Accepted 12 September 2017; Published 4 December 2017

Academic Editor: Helai Huang

Copyright © 2017 Xinxin Wei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Previous studies have investigated various factors that contribute to the severity of work zone crashes. However, little has been done on the specific effects of light conditions. Using the data from the Enhanced Tennessee Roadway Information Management System (E-TRIMS), crashes that occurred in the Tennessee work zones during 2003–2015 are categorized into three light conditions: daylight, dark-lighted, and dark-not-lighted. One commonly used decision tree method—Classification and Regression Trees (CART)—is adopted to investigate the factors contributing to crash severity in highway work zones under these light conditions. The outcomes from the three decision trees with differing light conditions show significant differences in the ranking and importance of the factors considered in the study, thereby indicating the necessity of examining traffic crashes according to light conditions. By separately considering the crash characteristics under different light conditions, some new findings are obtained from this study. The study shows that an increase in the number of lanes increases the crash severity level in work zones during the day while decreasing the severity at night. Similarly, drugs and alcohol are found to increase the severity level significantly under the dark-not-lighted condition, while they have a limited influence under daylight and dark-lighted conditions.

1. Introduction

Work zones have been an important research topic because they have a substantial effect on a nation's economy and traffic flow. Statistics show the economic cost of a fatal crash was \$1,398,916 in 2010 in the United States [1]. Based on this estimate, the annual cost of work zone fatalities is more than seven billion dollars per year. Moreover, considering the 26,000 nonfatal injury crashes and 60,000 property–damage-only (PDO) crashes that occur in work zones, additional billions of dollars of economic damage occur annually. At this time, the number of work zones is increasing. During peak construction season, approximately 20% of highway system is under construction and motorists may encounter a work zone every 100 miles [2].

To reduce adverse traffic impacts on the public, more and more work zones require night construction. An extensive survey conducted of 175 work zones in 13 states revealed that 58% of the work zones involved mostly daylight construction,

33% involved primarily night work, and the remaining 9% were active day and night [3]. This has raised concerns about whether work zones have influenced traffic safety at night.

Previous studies have found the night crash rate was higher than daylight crash rate in work zones [4–6]. Arditi et al. [7] used Illinois fatality crashes from work zones to investigate safety differences between night and daylight construction in the period 1996–2011, showing that night work zones were more hazardous. In one study, crash rates per million vehicle miles were higher in night work zones by 67 to 156 percent [8]. The differences in crash rate between day and night work zones suggest that both should be examined separately in varying light conditions. Although there is a consensus that the night crash rate has improved, there is debate about improvement in night crash severity.

Knowing crash risk factors is a key to create safe work zones. With more night work zones and the evident differences between day and night work zone crashes, there is an urgent need to investigate work zone crashes under different

light conditions and little research has been undertaken on this topic.

Severity is considered the most important crash outcome and is the core of this paper. This study analyzes crash injury severity in work zones with respect to its inherent casualties and not frequency. A decision tree method is used to model the severity of traffic accidents using available risk factors. Unlike most of the previous studies, in which light condition was treated as a single contributing factor, this study divides light conditions into three categories: daylight, dark-lighted, and dark-not-lighted. Three decision trees are built to reveal the relationship between crash severity and different contributing factors under differing light conditions in work zones.

2. Literature Review

2.1. Effects of Light Condition on Crashes. Light condition is a significant factor affecting traffic crashes. Previous studies have confirmed that adverse light conditions may increase both crash frequency and severity [4–6]. In fact, Gray et al. [4], Abdel-Aty [5], and Huang et al. [6] all reported that injury severity increases during darkness. Pande and Abdel-Aty [9] concluded that there is a significant correlation between lack of illumination and high crash severity. de Oña et al. [10, 11] pointed out that fatal accidents are associated with roadways with no lighting. Wanvik [12] found that good road lighting can decrease road accidents by one-half based upon a study of 763,000 injury accidents and 3.3 million property-damage accidents in Norway. Some studies (e.g., [13]) found that drivers are less likely to be injured in a construction work zone under darkness (with good illumination) than under daylight conditions. Moreover, researchers found that crash prediction models can reveal detailed information about contributing factors [14–16]. For example, Ullman et al. [17] found that some contributing factors are significant in a daytime crash rate model (e.g., low speed limit and the number of entering ramps per lane per mile), while others become significant in a nighttime model (e.g., snow and percentage of trucks). In summary, light condition has an important effect on traffic crashes but little has been done to explore the relationship between work zone crash severity and its contributing factors. More research efforts are needed in this area.

2.2. Decision Tree Method. Crash models are used to investigate the effects of risk factors on crashes in work zones. Regression models (such as logit and probit) have been widely employed [18–21]. In regression models, binary or multiple levels of severity are typically set as dependent variables and the risk factors affecting severity as independent variables. A common assumption for regression models is that there is no dependency among the risk factors. In addition, regression models need to assume a specific functional form to model the relationships between dependent and independent variables. Therefore, use of regression models is limited if the assumptions do not hold well [22].

In order to overcome the limitations of regression models, classification models using data-mining approaches have

been applied to the risk factor analysis problem. Typically, severity level is set as a class variable and risk factors as feature variables [22–26]. The decision tree classifier is one classification model, and the three commonly used decision tree methods are Classification and Regression Trees (CART), the Iterative Dichotomiser 3 (ID3) algorithm [27], and the C4.5 algorithm [28]. The CART method [29] uses a split criterion based on the Gini Index. Quinlan's method uses Information Gain (IG) as a split criterion based on the entropy measure on probabilities [30]. Subsequently Quinlan [31] also presented the algorithm C4.5, which is an advanced version of ID3 with a split criterion, called the Information Gain Ratio (IGR), similar to the one used in the ID3 procedure that penalizes variables with many states. Among these methods, CART is the most widely used [25, 28, 32–36]. It should be noted, however, that compared to conventional statistical models, CART still has its limitations such as simplicity and difficulty in interpreting. Nonetheless, previous studies have confirmed that the CART algorithm can be adopted in crash severity analyses and provide a more precise result compared with other prediction models [34].

3. Method

From the above discussion, many modeling approaches have been used to investigate the effects of risk factors on crashes in work zones. Different modeling approaches have different advantages and disadvantages. In this study, the decision tree method is selected and the CART algorithm is used to generate a decision tree. The split criterion in the CART method is based on gini, which represents the diversity of a factor, and is calculated as follows:

$$\text{gini} = 1 - \sum_i^n p_i^2, \quad (1)$$

where i is the category of the target (injury or PDO), n is the total number of targets, and p is the percentage of injury or PDO. Since CART is a binary tree here, our total number of targets is two.

Gini is used to calculate the diversity of the beginning node, while the Ginidex is created to measure the heterogeneity of the following node. For each node, the Ginidex is calculated as follows:

$$\text{Ginidex}(x) = \sum_j^n p_{xij} \text{gini}(x_{ij}), \quad (2)$$

where x is the contributing factor like lane number and x_{ij} means the severity i of character j . For instance, i can be injury or PDO and j may represent lane number >2 or <2 . Finally p_{xij} represents the percentage of x_{ij} .

In order to determine the next split node, the category with the largest diversity improvement is chosen:

$$\text{Maxnode} = \text{Max} \{ \text{gini}_{\text{parent}} - \text{Ginidex}(x=i)_{\text{childi}} \}, \quad (3)$$

where $\text{gini}_{\text{parent}}$ is the gini of higher layer and Ginidex is the index of second layer. Then this process is repeated several

times until the improvement equals 0 or reaches the maximum level. Since large trees could lead to higher percentage of misclassification [22], decision trees with different layers are tried in this study. When a decision tree with two layers is built, it presents less information. When a four-layer tree is built, it appears that many nodes in the fourth layer contain less than 1% of the total crashes. According to previous research [28], if one node contains less than 1% of samples, the results are not reliable. Therefore, three layers are chosen as the maximum depth for a decision tree in this study.

One of the major advantages of decision tree analysis is the decision rule. Decision rules have logical structure like “If A, then B”. While regression models show the impact of single factors, a decision tree can show the effect produced by a combination of several factors. In this study, decision rules are inferred when injury rate > 50%, or PDO rate > 80% and the population on that node > 1%. This is because the overall injury rate is approximately 20%–30%, and injury rates above 50% are extreme.

In order to feel confident of the results, the data used in this study are divided into two subsets: 70% of the data is used for training the model, while the remaining 30% is used for validation. The accuracy can be calculated as follows:

$$\text{Percentage Correct} = \frac{\sum_{i=1}^n \text{TP}_i}{\sum_{i=1}^n \text{TP}_i + \text{FN}_i} * 100, \quad (4)$$

where TP_i is true positive and FN_i is false negative.

In addition, the CART algorithm can also calculate the importance of each factor based on the improvement in Ginidex. The Importance Index (IM) is defined as follows:

$$\text{IM}(x_j) = \sum_{t=1}^T \frac{n_t}{N} \Delta \text{Ginidex}_{j_t}, \quad (5)$$

where x_j is a variable, $\Delta \text{Ginidex}_{j_t}$ is the reduction in Ginidex, n_t is the number of the observations in the dataset that belong to node t , T is the total number of nodes, and N is the total number of observations. The detailed calculation method and principles can be found in Montella et al. [25], Chang and Chien [35], and J. S. Lee and E. S. Lee [37].

4. Data

The data from the Enhanced Tennessee Roadway Information Management System (E-TRIMS) is used in the study. The crashes that happened in the work zones during 2003–2015 are used. There are five variables describing light conditions in the database including daylight, dark-lighted, dark-not-lighted, dusk, and dawn. The light conditions are based on the fixed streetlights in this study. Since the numbers of crashes under dusk and dawn conditions are relatively small (279 and 215, resp.), these data are excluded. Thus, a total number of 19941 crashes are analyzed. The class variable of the study is accident severity. The injury severity analyzed in the study includes that of drivers, passengers, and pedestrians.

When analyzing injury risk factors, it is desirable to include as many injury severity levels as possible because different factors may have different effects on the injury

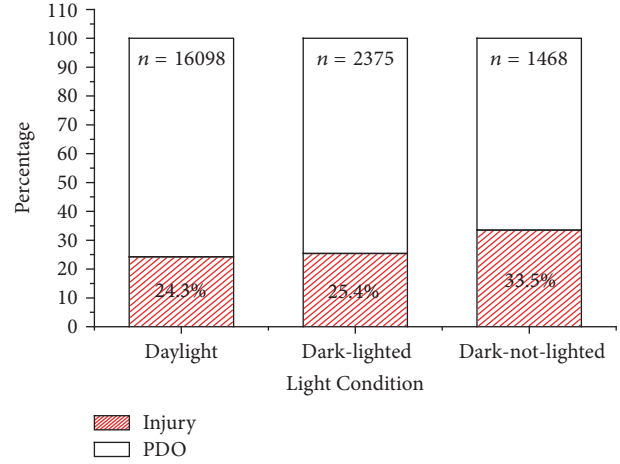


FIGURE 1: Crash severity distribution under different light conditions.

levels. However, among the 19941 crashes used in the study, fatal crashes only account for less than 1%. The number of fatal crashes is not high enough to conduct a reliable analysis. Therefore, several severity levels including fatal, incapacitating, and slight injuries are combined into a single injury level. Similar to previous studies [22], two levels of injury severity are used in the study: injury and property-damage-only (PDO). PDO refers to a crash where no one was injured but only the vehicle was damaged. Figure 1 shows the crash data according to two severity levels and three light conditions.

As demonstrated in Figure 1, the injury rate increases when light conditions worsen (24.3% < 25.4% < 33.5%). It can be seen that crashes under limited light conditions were more severe than those under dark-lighted condition, which is consistent with findings from other studies [4–6]. In order to investigate factors affecting work zone crashes under different light conditions, 15 variables are identified and presented in Table 1. In order to achieve more concise results, all the covariates are divided into two categories according to previous studies [20, 38, 39].

The variables describe characteristics related to the driver (at fault, drugs and alcohol, etc.), vehicle (body type), road (number of lanes, speed limit, terrain, and operation), and environment (weather condition, crash date, etc.). SPSS 19 is used to build the decision trees in the study.

5. Results and Discussion

5.1. Decision Trees under Different Light Conditions. Figure 2 and Table 2 present the results of the decision tree under the daylight condition. The first node is split by collision type, demonstrating that a head-on collision has a predicted injury probability more than twice that of other noncollision types (51.8% versus 23.5%). This is consistent with the findings from Kockelman and Kweon [19].

The lowest injury probability appears at node 8, with an injury rate of 11.7%. This node represents the most advantageous situation in daytime work zone crashes, indicating

TABLE 1: Description of variables.

Category	Variables	Description: code	Light condition		
			Daylight % (n)	Dark-lighted % (n)	Dark-not-lighted % (n)
Driver	At fault (AF)	No (N)	49.4% (7946)	42.2% (1003)	44.3% (651)
		Yes (Y)	50.6% (8152)	57.8% (1372)	55.7% (817)
	Drugs and alcohol (D&A)	No (N)	99.2% (15974)	93.7% (2225)	92.5% (1358)
		Yes (Y)	0.8% (124)	6.3% (150)	7.5% (110)
		Vehicle maneuver (VM)	Changing lanes or merging (CL)	25.3% (4065)	19.2% (456)
Other (O)	74.7% (12033)		80.8% (1919)	87.9% (1290)	
Vehicle	Body code (BC)	Not-truck-related (NT)	97.1% (15632)	98.8% (2347)	98.9% (1452)
		Truck-related (T)	2.9% (466)	1.2% (28)	1.1% (16)
Road	No. of lanes (NL)	≤ 2	20.0% (3218)	13.1% (312)	23.8% (350)
		> 2	80.0% (12880)	86.9% (2063)	76.2% (1118)
	Speed Limit (SL)	≤ 40	50.9% (8188)	58.0% (1377)	39.2% (576)
		> 40	49.1% (7910)	42.0% (998)	60.8% (892)
	Terrain (TR)	Flat (F)	6.9% (1118)	9.1% (215)	3.7% (54)
		Not flat (NF)	93.1% (14980)	90.9% (2160)	96.3% (1414)
	Operation (OP)	One-way (1)	0.8% (124)	0.7% (16)	0.3% (4)
Two-way (2)		99.2% (15974)	99.3% (2359)	99.7% (1464)	
Environment	Traffic control device (TCD)	At control (AC)	66.65% (10729)	68.67% (1631)	62.36% (918)
		Not at control (NAC)	33.35% (5369)	31.33% (744)	37.64% (554)
	Location (LC)	Along Roadway (AR)	41.4% (6662)	27.0% (641)	46.9% (688)
		At an Intersection (AI)	58.6% (9436)	73.0% (1734)	53.1% (780)
	Date of Crash	Other months (OM)	59.4% (9566)	77.9% (1849)	78.2% (1148)
		Summer vacation (SV)	40.6% (6532)	22.1% (526)	21.8% (320)
	Weather condition (WC)	Clear (C)	90.6% (14592)	85.7% (2035)	80.8% (1186)
		Not clear (NC)	9.4% (1506)	14.3% (340)	19.2% (282)
	Relation to first roadway (RFR)	Off roadway (OR)	4.7% (752)	10.6% (252)	20.0% (294)
		On roadway (OR)	95.3% (15346)	89.4% (2123)	80.0% (1174)
Urban or rural (UR)	Rural (R)	15.8% (2538)	1.7% (40)	28.6% (420)	
	Urban (U)	84.2% (13560)	98.3% (2335)	71.4% (1048)	
Manner of first collision (MFC)	Head-on (HO)	2.2% (358)	7.6% (181)	6.8% (100)	
	Not-head-on (NHO)	97.8% (15740)	92.4% (2194)	93.2% (1368)	

TABLE 2: Decision rules for crash severity of daylight condition.

Node	Rules [if (and...and...)]	Then	P (%)
2	If (CT = HO)	Injury	51.5
6	If (CT = HO) and (LC = AR)	Injury	59.5
8	If (CT = NHO) and (LC = AI) and (TR = FL)	PDO	88.2

that if the collision type is a non-head-on collision at an intersection and if a driver is involved in a collision, there would be an 11.7% chance of an injury and an 88.3% chance of a PDO crash on flat terrain. It should be noted that if the terrain is not flat, the predicted injury rate increases to 22.1%. One possible reason could be that flat terrain provides good visibility, which can slow crash speed and help decrease the severity level.

Of three decision rules inferred from the tree (see Table 3), two are injury rules. Note that both rules contain head-on collision, suggesting that avoiding head-on collisions is critical in lowering the daylight work zone crash

severity. Measures like adopting hard barrier to separate traffic from two directions can be helpful.

Figure 3 and Table 3 present the results of the decision tree under the dark-lighted condition. In the same manner as discussed previously, collision type is the criterion based on which node 0 was split under the dark-lighted condition (Figure 3). It shows that head-on collisions account for a much higher percentage of injury crashes that occur under the dark-lighted condition.

However, compared to the daylight model, lane number plays a different role in the crash severity in the dark-lighted model. In the daylight model, an increase in lane number

TABLE 3: Decision rules for crash severity of dark-lighted condition.

Rule	Node	Rules [if (and...and...)]	Then	P (%)
(1)	2	If (CT = HO)	Injury	54.5
(2)	5	If (CT = HO) and (NL <= 2)	Injury	100
(3)	8	If (CY = NHO) and (NL > 2) and (WC = NC)	PDO	89.2

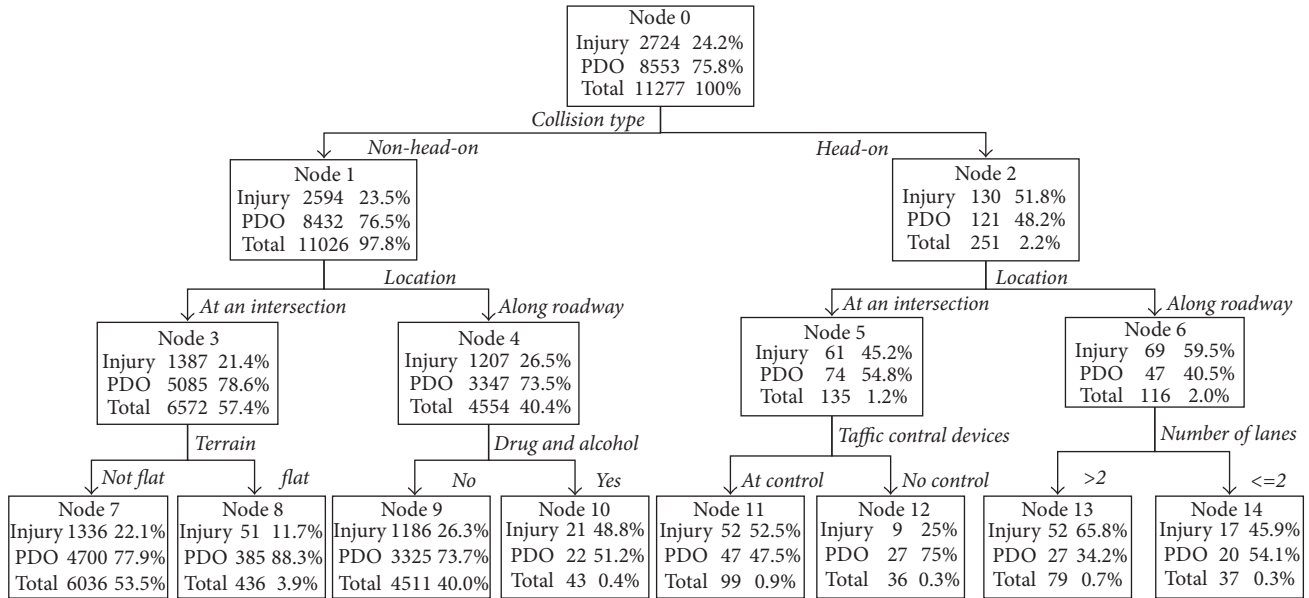


FIGURE 2: Decision tree under daylight condition.

increases the injury percentage of head-on collisions, while an increase in lane number decreases the injury rate in the dark-lighted model. Specifically, crashes occurring on narrow roads (<=2 lanes) predict a 100% injury rate in a head-on collision, whereas the predicted injury rate decreases to 47.4% on multilane roads. At the same time, the rate on narrow roads is 33.3%, while the rate on wider roads is 21.6% for non-head-on collisions. This phenomenon may be attributed to the changes in drivers' maneuvers under different light conditions. Weng and Meng [40] reported that drivers are more likely to be involved in risky driving maneuvers on multilane roads under daylight conditions, whereas at night most risky driving behavior occurred on narrower roads.

Node 8 gave the lowest predicted injury rate of 10.8%, indicating that the minimum severity case happens in non-head-on collisions on multiple lane (>2) roads when the weather condition is not clear under the dark-lighted condition.

Comparing node 9 and node 10, it can be seen that under the same conditions of multilane road and head-on collision, the use of a traffic control device significantly reduces the predicted injury rate by more than one-half (33.8% versus 69.8%). This indicates that a traffic control device is very helpful in lowering head-on crash severity on multilane roads under the dark-lighted condition. Therefore, it is highly recommended that traffic control devices be installed in work zones with multiple lanes under illumination.

Similar to the daylight model, three decision rules are obtained under the dark-lighted condition. However, the injury rate for the dark-lighted condition is slightly higher than that for the daylight condition. The fact that lane number plays an important role in work zone crashes under the dark-lighted condition may be due to decreased visibility (see rule (2) of Table 3).

Figure 4 and Table 4 present the results of the decision tree under the dark-not-lighted condition. Under dark-not-lighted condition, the darkest light condition, some significant changes in crash severity are found. Unlike the daylight and the dark-lighted decision tree models, the first partition node is not based on collision type but on speed limit. A higher speed limit (>40 miles/h) contributes to more severe crashes compared to a lower speed limit (<=40 miles/h) (37.9% versus 23%) for the dark-not-lighted condition. As we know, the dark-not-lighted condition is characterized by a sharp reduction of visibility that affects the driver's ability to perceive obstacles. With the high speed limit and limited visibility, drivers may have insufficient time to stop the vehicle.

The factor of drugs and alcohol shows a significant impact on crash severity under the dark-not-lighted condition. If a driver is under the influence, the predicted injury percentage climbs to 72.7% even with the speed limit lower than 40 miles/h. This rate reduces dramatically to 20.1% if the driver is not under the influence. This verifies the findings from

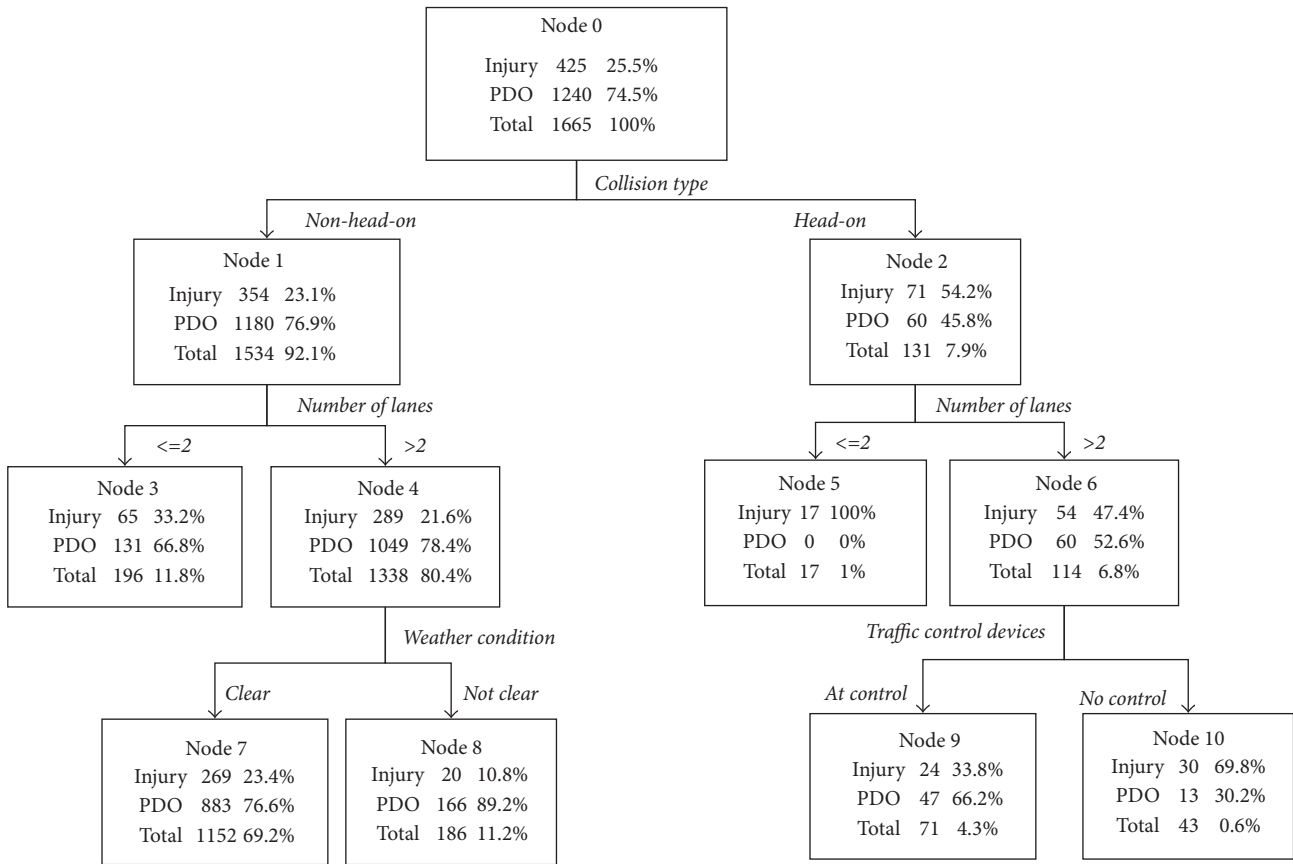


FIGURE 3: Decision tree under dark-lighted condition.

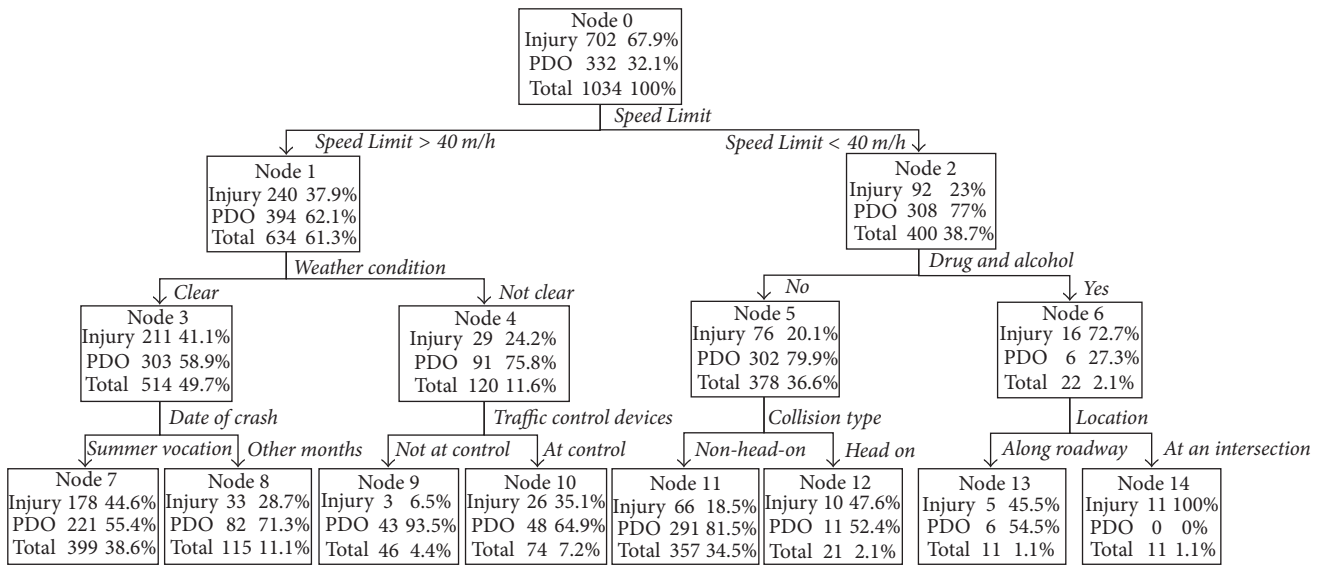


FIGURE 4: Classification tree under dark-not-lighted condition.

TABLE 4: Decision rules for crash severity of dark-not-lighted condition.

Rules	Node	Rules [if (and. . and. .)]	then	P (%)
(1)	6	If (SL <= 40) and (D & A = Y)	Injury	72.7
(2)	9	If (SL > 40) and (WC = NC) and (TCD = N)	PDO	93.5
(3)	11	If (SL <= 40) and (D & A = N) and (CT = NHO)	PDO	81.5
(4)	14	If (SL <= 40) and (D & A = Y) and (LC = AI)	Injury	100

TABLE 5: Rank of the importance of variables under different light conditions.

Rank	Daylight	Dark-lighted	Dark-not-lighted
(1)	Manner of first collision	Manner of first collision	Drugs and alcohol
(2)	Location	No. of lanes	Speed limit
(3)	Drugs and alcohol	Traffic control devices	Weather cond.
(4)	Terrain	Location	Traffic control devices
(5)	Urban or rural	Weather cond.	Date of crash
(6)	Traffic control devices	Drugs and alcohol	Manner of first collision
(7)	Speed limit	Relation to first roadway	Location
(8)	Date of crash	Urban or rural	No. of lanes
(9)	No. of lanes	Speed limit	Operation
(10)	Body code	Terrain	Body code
(11)	Relation to first roadway		Relation to first roadway
(12)	At fault		Urban or rural
(13)	Vehicle maneuver		

previous studies that drug and alcohol intake significantly increases the likelihood of severe injuries [41, 42] and further reveals that the impact of drugs and alcohol is significantly higher under the dark-not-lighted condition. One possible reason may be that drivers need to be more alert under this condition and impaired drivers under the influence are likely not to stop, change lanes to avoid obstacles, or otherwise avoid severe crashes.

Also, under the influence of drugs and alcohol, the injury rate of traffic crashes doubles at intersections compared to long open roadways at a speed limit less than 40 miles/h, indicating that intersections introduce additional risk under the dark-not-lighted condition. Node 9 shows the lowest injury rate of 6.5% under the dark-not-lighted condition, at a speed limit above 40 miles/h without traffic control devices. This is significantly different from the findings from Chang and Chien [35], which concluded that traffic control devices can enhance traffic safety. The reason for the seemingly contradictory findings may be due to the fact that traffic control devices usually appear at crash-prone sites where road conditions are more complex. It is improper to compare two places with different geometry features and traffic patterns. The effect of traffic control devices can only be validated by a before-after test in a future study.

Table 4 shows the decision rules inferred from the dark-not-lighted condition. Two of the rules are injury rules, indicating that the involvement of drugs and alcohol and a speed limit less than 40 miles/h are common in severe crashes. It is highly recommended that fines for driving under the influence be increased significantly in work zones including dark-not-lighted zones.

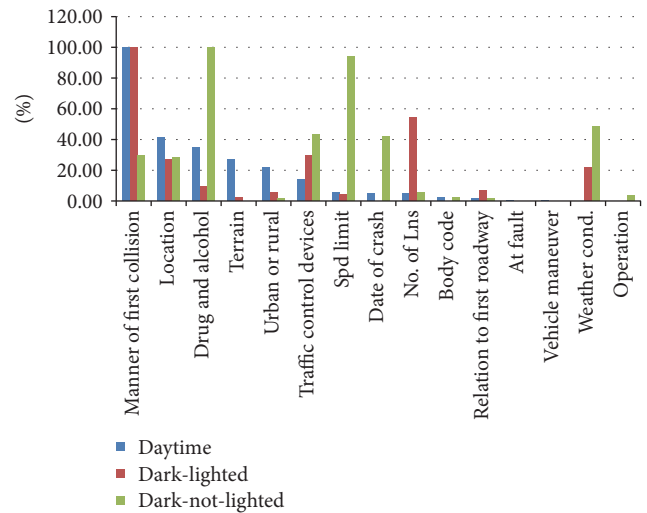


FIGURE 5: The importance of variables under different light conditions.

5.2. Comparison of the Importance of Variables. Table 5 and Figure 5 rank and compare the importance of various risk factors under different light conditions. For the daylight decision tree model, the top five factors contributing to crash severity are collision type, location, drugs and alcohol, terrain, and urban or rural location. For the dark-lighted model, the top five factors are collision type, number of lanes, traffic control devices, location, and weather condition. From the dark-not-lighted model, the top five factors are drugs and alcohol, speed limit, weather condition, traffic control devices, and

TABLE 6: Prediction Performance.

Sample	Observed	Predicted		
		Injury	PDO	Percentage correct
Daylight	Injury	46	1134	3.9%
	PDO	32	3609	99.1%
	Overall percentage	1.6%	98.4%	75.8%
Dark-lighted	Injury	19	159	10.7%
	PDO	9	523	98.3%
	Overall percentage	3.9%	96.1%	76.3%
Dark-not-lighted	Injury	5	155	3.1%
	PDO	0	274	100.0%
	Overall percentage	1.2%	98.8%	64.3%

date of crash. The ranking of the top five contributing factors is different under the three light conditions, indicating that the effects of the major contributing factors vary significantly under different light conditions within work zones. Table 5 also shows that vehicle body type, location (rural or urban), and road operation (one-way or two-way) do not show an important effect on crash severity in work zones. In order to achieve safer work zones, safety guidelines should be established according to different light conditions.

5.3. Validation. Table 6 presents the results from the validation of the decision trees. For all three models, the percentage predicted correctly was above 60%. This rate is higher than those reported in previous studies using the same method [35]. All three models are capable of predicting PDO crashes with an accuracy rate above 98%. On the other hand, these models show a poor ability to predict injury crashes with an accuracy ranging from 3.1% to 10.7%. This rate for predicting injuries is close to that reported in other studies using CART decision method [36]. The reason may lie in the dataset itself. There are more than 70% of the crashes that are PDO while only less than 30% belong to injury. In order to achieve the minimum error of the whole dataset, models or prediction methods tend to classify a result to PDO which is the major crash type.

6. Conclusions

In this study, three decision trees are generated using the CART method to investigate the factors contributing to crashes in work zones. The light-based individual decision tree models use detailed information of work zone crashes to identify risk factors. Identification of these factors then suggests mitigation measures that may help establish safer work zones. In the study, the following are found.

- (i) The daylight model indicates that in head-on crashes occurring along roadways, as opposed to intersections, drivers are at a higher risk, up to 59.7%, of being involved in injury crash.
- (ii) The dark-lighted model demonstrates that the injury rate of head-on crashes occurring on a narrow road (≤ 2 lanes) could reach 100%.

- (iii) Under the dark-not-lighted condition, a combination of a speed limit less than 40 miles/h and drivers being under the influence of drugs and alcohol could lead to an injury rate of up to 72.7%.

By examining the effects of specific light conditions on crash severity, this study reveals some new findings never reported before. The study shows that if drivers are under the influence of drug/alcohol, they have a larger chance of being involved in severe crashes when passing a work zone without street light than a work zone with street light. This study reveals that collision type is the most important risk factor under daylight and dark-lighted conditions but not under dark-not-lighted condition. On the other hand, the study suggests that traffic control devices do not reduce crash severity under the dark-not-lighted condition, yet, they do under the dark-lighted condition. This implies that traffic control devices should be designed and used differently according to light conditions. Additionally, the number of roadway lanes shows opposite effects on crash severity under the daylight and the dark-not-lighted conditions. Specifically, under the daylight condition, an increase in the number of lanes may increase crash severity, whereas it may help reduce crash severity under the dark-not-lighted condition.

The CART decision tree method was found to be useful in revealing crash severity characteristics and the factors contributing to traffic crashes in work zones. In future work, these results may be helpful in developing work zone safety guidelines to mitigate crash severity. In addition, it will be of practical significance to use the decision tree method to investigate drivers' behavior under different light conditions.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the Construction Industry Research and Policy Center (CIRPC) for funding this research project. The first author would also like to thank the China Scholarship Council (CSC) for its financial support.

References

- [1] L. J. Blincoe, T. R. Miller, E. Zaloshnja, and B. A. Lawrence, "The economic and societal impact of motor vehicle crashes, 2010," Report No. DOT HS 812 013, National Highway Traffic Safety Administration, Wash, USA, 2015.
- [2] K. Wunderlich and D. Hardesty, *A Snapshot of Summer 2001 Work Zone Activity*, Federal Highway Administration, U.S. Department of Transportation, Wash, USA, 2003.
- [3] J. E. Bryden and D. J. Mace, "A procedure for assessing and planning nighttime highway construction and maintenance," NCHRP Report 475, Transportation Research Board, National Research Council, Wash, USA, 2002.
- [4] R. C. Gray, M. A. Quddus, and A. Evans, "Injury severity analysis of accidents involving young male drivers in Great Britain," *Journal of Safety Research*, vol. 39, no. 5, pp. 483–495, 2008.
- [5] M. Abdel-Aty, "Analysis of driver injury severity levels at multiple locations using ordered probit models," *Journal of Safety Research*, vol. 34, no. 5, pp. 597–603, 2003.
- [6] H. Huang, H. C. Chin, and M. M. Haque, "Severity of driver injury and vehicle damage in traffic crashes at intersections: a Bayesian hierarchical analysis," *Accident Analysis & Prevention*, vol. 40, no. 1, pp. 45–54, 2008.
- [7] D. Arditi, M. A. Ayrancioglu, and J. Shi, "Effectiveness of safety vests in nighttime highway construction," *Journal of Transportation Engineering*, vol. 130, no. 6, pp. 725–732, 2004.
- [8] E. C. Sullivan, "Accident Rates during Nighttime Construction," Report No. UCB-ITS-RR-89-11, Institute of Transportation Studies, University of California-Berkeley, Berkeley, Calif, USA, May 1989.
- [9] A. Pande and M. Abdel-Aty, "Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool," *Safety Science*, vol. 47, no. 1, pp. 145–154, 2009.
- [10] J. de Oña, R. O. Mujalli, and F. J. Calvo, "Analysis of traffic accident injury on Spanish rural highways using Bayesian networks," *Accident Analysis & Prevention*, vol. 43, pp. 402–411, 2011.
- [11] J. de Oña, G. López, and J. Abellán, "Extracting decision rules from police accident reports through decision trees," *Accident Analysis & Prevention*, vol. 50, pp. 1151–1160, 2013.
- [12] P. O. Wanvik, "Effects of road lighting: an analysis based on Dutch accident statistics 1987-2006," *Accident Analysis & Prevention*, vol. 41, no. 1, pp. 123–128, 2009.
- [13] J. Weng and Q. Meng, "Analysis of driver casualty risk for different work zone types," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1811–1817, 2011.
- [14] J. D. Bullough, E. T. Donnell, and M. S. Rea, "To illuminate or not to illuminate: roadway lighting as it affects traffic safety at intersections," *Accident Analysis & Prevention*, vol. 53, no. 1, pp. 65–77, 2013.
- [15] R. R. Dinu and A. Veeraragavan, "Random parameter models for accident prediction on two-lane undivided highways in India," *Journal of Safety Research*, vol. 42, no. 1, pp. 39–42, 2011.
- [16] E. T. Donnell, R. J. Porter, and V. N. Shankar, "A framework for estimating the safety effects of roadway lighting at intersections," *Safety Science*, vol. 48, no. 10, pp. 1436–1444, 2010.
- [17] G. L. Ullman, M. D. Finley, J. E. Bryden, R. Srinivasan, and F. M. Council, "Traffic Safety Evaluation of Nighttime and Daytime Work Zones," NCHRP Report 627, Transportation research board, Wash, USA, 2008.
- [18] A. S. Al-Ghamdi, "Using logistic regression to estimate the influence of accident factors on accident severity," *Accident Analysis & Prevention*, vol. 34, no. 6, pp. 729–741, 2002.
- [19] K. M. Kockelman and Y.-J. Kweon, "Driver injury severity: an application of ordered probit models," *Accident Analysis & Prevention*, vol. 34, no. 3, pp. 313–321, 2002.
- [20] N. N. Sze and S. C. Wong, "Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes," *Accident Analysis & Prevention*, vol. 39, no. 6, pp. 1267–1278, 2007.
- [21] C. Dong, D. B. Clarke, X. Yan, A. Khattak, and B. Huang, "Multivariate random-parameters zero-inflated negative binomial regression model: an application to estimate crash frequencies at intersections," *Accident Analysis & Prevention*, vol. 70, pp. 320–329, 2014.
- [22] L.-Y. Chang and H.-W. Wang, "Analysis of traffic injury severity: an application of non-parametric classification tree techniques," *Accident Analysis & Prevention*, vol. 38, no. 5, pp. 1019–1027, 2006.
- [23] T. Beshah and S. Hill, "Mining road traffic accident data to improve safety: role of road-related factors on accident severity in ethiopia," in *Proceedings of AAAI Artificial Intelligence for Development (AI-D '10)*, pp. 22–24, 2010.
- [24] A. T. Kashani and A. S. Mohaymany, "Analysis of the traffic injury severity on two-lane, two-way rural roads based on classification tree models," *Safety Science*, vol. 49, no. 10, pp. 1314–1320, 2011.
- [25] A. Montella, M. Aria, A. D'Ambrosio, and F. Mauriello, "Data-mining techniques for exploratory analysis of pedestrian crashes," *Transportation Research Record*, vol. 2237, no. 1, pp. 107–116, 2011.
- [26] A. Montella, M. Aria, A. D'Ambrosio, and F. Mauriello, "Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery," *Accident Analysis & Prevention*, vol. 49, pp. 58–72, 2012.
- [27] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.
- [28] J. Abellán, G. López, and J. de Oña, "Analysis of traffic accident severity using decision rules via decision trees," *Expert Systems with Applications*, vol. 40, no. 15, pp. 6047–6054, 2013.
- [29] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*, Chapman and Hall, Belmont, Calif, USA, 1984.
- [30] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 4, pp. 623–656, 1948.
- [31] J. R. Quinlan, *C4.5: Programs for machine learning*, Morgan Kaufmann Publishers, San Mateo, Calif, USA, 1993.
- [32] P. M. Kuhnert, K.-A. Do, and R. McClure, "Combining non-parametric models with logistic regression: an application to motor vehicle injury data," *Computational Statistics & Data Analysis*, vol. 34, no. 3, pp. 371–386, 2000.
- [33] A. Pakgozar, R. S. Tabrizi, M. Khalili, and A. Esmaeili, "The role of human factor in incidence and severity of road crashes based on the CART and LR regression: a data mining approach," *Procedia Computer Science*, vol. 3, pp. 764–769, 2010.
- [34] O. H. Kwon, W. Rhee, and Y. Yoon, "Application of classification algorithms for analysis of road safety risk factor dependencies," *Accident Analysis & Prevention*, vol. 75, pp. 1–15, 2015.
- [35] L. Chang and J. Chien, "Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model," *Safety Science*, vol. 51, no. 1, pp. 17–22, 2013.

- [36] A. T. Kashani, A. Shariat-Mohaymany, and A. Ranjbari, "A data mining approach to identify key factors of traffic injury severity," *Promet - Traffic - Transportation*, vol. 23, no. 1, pp. 11–17, 2011.
- [37] J. S. Lee and E. S. Lee, "Exploring the usefulness of a decision tree in predicting people's locations," *Procedia - Social and Behavioral Sciences*, vol. 140, pp. 447–451, 2014.
- [38] R. Tay, "A random parameters probit model of urban and rural intersection crashes," *Accident Analysis & Prevention*, vol. 84, pp. 38–40, 2015.
- [39] D. M. Cerwick, K. Gkritza, M. S. Shaheed, and Z. Hans, "A comparison of the mixed logit and latent class methods for crash severity analysis," *Analytic Methods in Accident Research*, vol. 3–4, pp. 11–27, 2014.
- [40] J. Weng and Q. Meng, "Effects of environment, vehicle and driver characteristics on risky driving behavior at work zones," *Safety Science*, vol. 50, no. 4, pp. 1034–1042, 2012.
- [41] S. P. Baker, E. R. Braver, L.-H. Chen, G. Li, and A. F. Williams, "Drinking histories of fatally injured drivers," *Injury Prevention*, vol. 8, no. 3, pp. 221–226, 2002.
- [42] B. E. Smink, B. Ruiters, K. J. Lusthof, J. J. De Gier, D. R. A. Uges, and A. C. G. Egberts, "Drug use and the severity of a traffic accident," *Accident Analysis & Prevention*, vol. 37, no. 3, pp. 427–433, 2005.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

