ANALYZING THE DETERMINANTS OF E-BIKE CRASH SEVERITY IN CHINA USING ADVANCED LOGIT MODELING

[1]Sowjanya,[2]Srujana
[12]Students
Department Of Mechanical Engineering

ABSTRACT
The growing popularity of electric bicycles (e-bikes) in China has significantly altered urban mobility patterns but also raised serious road safety concerns. This study employs a Generalized Ordered Logit Model to analyze the factors influencing the severity of e-bike crashes in China, drawing upon a comprehensive crash dataset. The model allows for flexible thresholds between severity levels, accommodating the violation of the proportional odds assumption common in ordinal data analysis. Key variables examined include rider demographics, environmental conditions, road characteristics, crash timing, helmet usage, and vehicle interactions. The results indicate that factors such as age, lighting conditions, intersection type, and involvement of heavy vehicles have a statistically significant impact on crash severity. These findings provide actionable insights for urban planners, traffic safety authorities, and policymakers, highlighting the need for targeted safety interventions, infrastructure enhancements, and enforcement strategies to mitigate crash outcomes and enhance e-bike safety in rapidly urbanizing regions.

I.INTRODUCTION

Electric bicycles (e-bikes) have emerged as a dominant mode of urban transportation in China due to their affordability, convenience, and environmental advantages. However, the rapid proliferation of e-bikes has coincided with a sharp increase in traffic-related injuries and fatalities involving these vehicles. Unlike traditional bicycles, e-bikes can travel at higher speeds and often operate in mixed traffic, leading to a higher risk of severe accidents, especially in densely populated urban areas.

Understanding the factors that influence the severity of e-bike crashes is essential for designing effective safety regulations and infrastructure improvements. Traditional statistical models often assume uniform effects of predictors across outcome categories, which may not hold in real-world crash data. To address this, this study applies a Generalized Ordered Logit Model (GOLM)—an advanced form of ordinal regression that relaxes the parallel lines assumption, allowing for a more nuanced analysis of crash severity levels (e.g., minor, moderate, and severe outcomes).

The primary objective of this research is to identify and quantify the determinants of crash severity among e-bike users in China. By analyzing real-world crash data and applying robust statistical modeling, the study aims to provide empirical evidence to support data-driven policymaking, ultimately contributing to safer urban transportation systems.

II.LITERATURE REVIEW

Weinert et al.6 examined safety perceptions of ebicyclists by surveying e-bicycle riders in Shijiazhuang City in China. Their findings include (1) e-bicyclists feel satisfied and safe when traveling; (2) female riders feel safer riding e-bicycles to cross intersections than bicycles. Feng et al.7 examined the trend of e-bicyclerelated injuries in China. It was found that e-bicyclerelated injuries and deaths were experiencing a significant increase in recent years while overall traffic and bike-related injuries and deaths were decreasing. It was pointed out that e-bicycle safety had become a serious problem that deserves more attention in China. Yao and Wu8 conducted a self-reported questionnaire survey on e-bicyclists in two large cities in China. From the responses of e-bicyclists, it was found that gender and automobile driving experiences were highly correlated with at-fault e-bicycle-related crashes. Males were found to be more likely to be involved in at-fault crashes than females. e-bicyclists with driving licenses were found to be less likely to be involved in crashes. It was also found that risk perceptions and safety attitudes of e-bicyclists significantly impacted their behaviors in e-bicycle crashes. Wu et al.9 examined red-light running behaviors of e-

bicyclists at three signalized intersections in Beijing using video cameras. It was found that age was significantly related to red-light running behaviors: young and middle-aged e-bicyclists were more prone to violate traffic rules. Besides, males were more likely to take risks than females at intersections. Bai et al.10 observed driving behaviors of ebicyclists at signalized intersection and concluded potentially dangerous conflict types between e-bicyclists and drivers. Hu et al.11 explored actors associated with e-bicyclists' injury severity. Age, road user category, traffic rule violation, crash mode, impact type, and vehicle type were found to related to injury severity. However, only hospital data were used and only two injury levels were considered. Bai et al.12 examined the driving behavior of e-bicyclists with the presence of mid-block bicycle lanes. Operating speed of e-bicycles, speed difference between bicycles and e-bicycles, volumes of e-bicycles, and the width of bike lanes were found to be related to e-bicyclist safety. Yang et al.13 presents a hazard-based duration approach to investigate riders' crossing behaviors at signalized intersections. Rider type, gender, waiting position, conformity tendency, and crossing traffic volume were identified to have significant effects on riders' waiting times and violation hazards. Zhang and Wu14 investigated the effect of sunshields on red-light infringement behavior of cyclists and e-bikers in the city of Hangzhou, China. Their results suggested that sunshields installed at intersections can reduce red-light infringement rates of cyclists and e-bikers on both sunny and cloudy days. Zhou et al. examined the relationship between e-bikes' passing rate and operating speed, using field data in Hangzhou. Average speed, 15 percentile speed, and speed variance were found as significant factors.15 Yang et al.16 investigated unsafe driving behaviors of ebicyclists in Suzhou, using a cross-sectional observation study. It was found that speeding, road rule violations and lack of helmet use were frequent unsafe behaviors for e-bicyclists, especially males. Wang et al.17 modeled for e-bicyclists, especially males. Wang et al.17 modeled faulty behaviors among e-bike-related fatal crashes in China. Pre-crash behaviors, bike lane and median type, older e-bike rider, heavy good vehicle drivers, and built
environment were found to be correlated with faulty behaviors of e-bike riders. Xu et al.18 examined spatial interdependence in e-bike choice using spatially autoregressive model. It was found that travelers were more likely to use e-bikes if their neighbors at the trip origin and destination also use e-bikes. Previous studies, mostly based on surveys and observational studies, have provided valuable information
for understanding e-bicyclist safety. However, most focused on specific unsafe driving behaviors. To our
knowledge, there still lacks a thorough crash modeling effort based on historical crash records of e-bicyclist. Thus, in this article, we will examine all possible factors related to injury severity of bicyclists. By modeling injury severity, factors associated with probability of ebicyclists getting different levels of injury outcomes can be explored. In doing so, safety engineers are able to develop safety countermeasures to reduce the occurrence of severe injuries based on such modeling technique.

Data description

Crash records between e-bicycle and motor vehicle were utilized for crash modeling. The crash records
were drawn from the crash database regularly maintained by the traffic crash treatment center of Guilin
Police Department. Over 31,000 crashes involving about 60,900 individuals from the year 2008 to 2014 were allowed to be accessed. Of these, about 4000 crashes involved e-bicycles. The injury severity of each e-bicyclist involved in the crash is recorded on 4-point ordinal scale: (1) no injury, (2) no treatment injury, (3) treatment injury, and (4) fatal. For the study purpose, we use only e-bicycles crashes that involve a single motor vehicle and an e-bicycle. After checking data validity and consistency, totally 3814 available crash records between e-bicycles and motor vehicles were finally extracted for further analysis. e-bicyclist demographics, crash characteristics, road geometrics, and roadway environments were found to be correlated with e-bicyclist safety, per the literature. More specifically, gender,6

age,9 driving behaviors of ebicyclists, 10 crash pattern,11 impact type,11 vehicle type,11 speed,12 and width of lane12 were found as significant. Thus, these factors need to be considered for severity modeling. Moreover, according to Kim et al.,19 driver demographics, vehicle features, time effects, and land use were found as significant to bicyclists' injury severity. Since bicyclists and e-bicyclists are comparable to two-wheel non-motorized users, those factors also need to be considered. Thus, in general, factors related to e-bicyclists' demographics, driver demographics, vehicle features, crash types, road geometry, time, environment, and land use were extracted from the crash records. The descriptive statistics is showed in Table 1. According to Table 1, e-bicyclists aged from 25 to 54 years comprised 58.1% of the sample. In 21.0% of the crashes, e-bicyclists were found to be older than 55. There was a share difference for gender, that is, about 56.9% of the crash-involved e-bicyclists were males while 43.1% were females. Regarding driver characteristics, in 74.1% of the crashes, motor-vehicle drivers aged from 25 to 54 years are identified. Most crashes were observer for male driver (87.4%). Drivers with less than 3 years of driving experience account for 21.8% of the total crashes, while experienced drivers with driving experiences more than 10 years account for 30.6%. In 4.0% of the crashes, drivers were found to commit drunk-driving offense. Passenger cars were mostly often involved in e-bike crashes (56.5%), followed by motorcycle (17.6%). Note that heavy trucks accounted for only 10.3% of the total crashes, but 20.8% of fatal crashes. About 69.9% of the crashes are side-impact crashes between e-bicycle and motor vehicle head-on and rearend crashes account for 11.6% and 10.5% of the crashes, respectively. In 58.1% of the crashes, motor vehicles were determined to be solely at fault, while in 19.1% of the crashes, e-bicycles were determined to be solely at fault. Although motor vehicle drivers were found to commit speeding offense in 3.6% of the crashes, 23.7% of them were fatal crashes. About 39.3% of the crashes occurred in motorized lanes, while 20.4% of the crashes occurred in non-motorized lanes. In 8.2% of the crashes, drivers or e-bicycle riders were found to escape from crash scene. About 71.4% of the crashes occurred when the ebicyclist was going straight and 15.3% when the vehicle was turning right. 91.7% of the crashes occurred on straight roads and 56.2% happened on asphalt roads. The number of crashes occurred within intersections is close to that in roadway sections. More than half of the crashes occurred on urban arterials (60.8%). 47.1% of the crashes occurred on roadways where motorized and non-motorized vehicles are traveling together. In 35.9% of the crashes, there were traffic line markings between motorized and non-motorized vehicles. Regarding land use characteristics, crashes mostly occurred within residential area (37.9%), followed by commercial area (13.2%). About 39.3% of the crashes occurred within 30%–50% developed area and 50%– 80% developed area accounted for 26.2% of the crashes. About 32.8% of the crashes occurred on weekends and holidays. 12.8% and 13.9% of the crashes happened during morning and evening peak hours, respectively. Most crashes occurred on sunny days (63.4%) and dry roads (77.6%). 72.5% of the crashes occurred during daytime and 49.5% of the crashes occurred when the sight distance is more than 200 m.

III. METHODOLOGY

In this study, injury severity is treated as a response variable with four levels: 0=possible or no injury, 1=no treatment injury, 2=treatment injury, and 3=fatality. Since injury severity is ordinal, ordered logit/proportional odds models are considered instead of multinomial models, which ignore the ordering of categories.20 In the ordered logit model/proportional odds model, there is an observed ordinal variable Y, as a function of another latent variable Y*, that is not measured. The continuous and unmeasured latent variable Y whose values determine what the observed ordinal variable Y equals. The specification of ordered logit model is as follows

$$Y_i^* = a^*X + \varepsilon_i \qquad (1)$$

where $Y^*$is a latent and continuous variable (dependent) measuring the ith observation, that is the injury everity of the e-bicyclist in the ith crash record; X is a vector of explanatory variables describing e- icyclists' demographics, driver demographics, vehicle features, crash types, road geometry, time, nvironment, and land use; $a^*$is a vector of unknown parameters which need to be estimated, that is the coefficients of explanatory variables; e_i is a random error term which is assumed to have a logistic distribution. The observed injury severity variable Y* is determined by the following model

Table 1. Descriptive statics.

Variables	Fatal injury	Treatment injury	No treatment injury	Possible/no injury	Total
e-bicydist characteristics					
Age					
<16	0 (0.0%)	18 (21.4%)	36 (42.9%)	30 (35.7%)	84 (2.3%)
16–24	25 (3.6%)	206 (30.0%)	228 (33.2%)	228 (33.2%)	687 (18.6%)
25–54	78 (3.6%)	706 (33.0%)	756 (35.3%)	600 (28.0%)	2140 (58.1%)
55 +	79 (10.2%)	274 (35.4%)	240 (31.0%)	180 (23.3%)	773 (21.0%)
Gender					
Male	63 (4.0%)	482 (30.3%)	558 (35.1%)	486 (30.6%)	1589 (43.1%)
Female	119 (5.7%)	722 (34.5%)	702 (33.5%)	552 (26.3%)	2095 (56.9%)
Driver characteristics					
Age (years)					
<25	34 (4.9%)	178 (25.5%)	288 (41.3%)	198 (28.4%)	698 (18.9%)
25–54	142 (5.2%)	986 (36.1%)	846 (31.0%)	756 (27.7%)	2730 (74.1%)
55 +	6 (2.3%)	40 (15.6%)	126 (49.2%)	84 (32.8%)	256 (6.9%)
Gender					
Male	179 (5.6%)	1098 (34.1%)	1068 (33.2%)	876 (27.2%)	3221 (87.4%)
Female	3 (0.6%)	106 (22.9%)	192 (41.5%)	162 (35.0%)	463 (12.6%)
Driver years					
<3	66 (8.2%)	274 (34.2%)	258 (32.2%)	204 (25.4%)	802 (21.8%)
3–5	45 (4.5%)	427.8 (43.0%)	264 (26.5%)	258 (25.9%)	994.8 (27.0%)
6–10	57 (7.5%)	96 (12.5%)	342 (44.7%)	270 (35.3%)	765 (20.8%)
>10	14 (1.2%)	406 (36.0%)	402 (35.6%)	306 (27.1%)	1128 (30.6%)
Intoxicated					
Yes	29 (19.7%)	64 (43.5%)	42 (28.6%)	12 (8.2%)	147 (4.0%)
No	153 (4.3%)	1140 (32.2%)	1218 (34.4%)	1026 (29.0%)	3537 (96.0%)
Vehicle characteristics					
Vehicle type					
Car	45 (2.2%)	556 (26.7%)	798 (38.3%)	684 (32.8%)	2083 (56.5%)
Bus	13 (9.6%)	68 (50.4%)	42 (31.1%)	12 (8.9%)	135 (3.7%)
Motorcycle	25 (3.9%)	246 (37.9%)	210 (32.4%)	168 (25.9%)	649 (17.6%)
Minivan	8 (5.4%)	74 (50.0%)	54 (36.5%)	12 (8.1%)	148 (4.0%)
Heavy truck	79 (20.8%)	144 (38.0%)	96 (25.3%)	60 (15.8%)	379 (10.3%)
Other	12 (4.1%)	116 (40.0%)	60 (20.7%)	102 (35.2%)	290 (7.9%)
Crash characteristics					
Type of crash					
Side-swipe	10 (3.9%)	64 (25.2%)	108 (42.5%)	72 (28.3%)	254 (6.9%)
Side-impact	114 (4.4%)	834 (32.4%)	918 (35.7%)	708 (27.5%)	2574 (69.9%)
Rear-end	27 (7.0%)	108 (27.9%)	150 (38.8%)	102 (26.4%)	387 (10.5%)
Head-on	23 (5.4%)	170 (39.8%)	84 (19.7%)	150 (35.1%)	427 (11.6%)
others	8 (19.0%)	28 (66.7%)	6 (14.3%)	0 (0.0%)	42 (1.1%)
Party at fault					
e-Bicyclist	92 (13.1%)	316 (45.0%)	168 (23.9%)	126 (17.9%)	702 (19.1%)
Driver	82 (3.8%)	392 (18.3%)	978 (45.7%)	690 (32.2%)	2142 (58.1%)
Both	8 (1.5%)	196 (36.3%)	114 (21.1%)	222 (41.1%)	540 (14.7%)
Speeding involved					
Yes	31 (23.7%)	76 (58.0%)	12 (9.2%)	12 (9.2%)	131 (3.6%)
No	151 (4.2%)	1128 (31.7%)	1248 (35.1%)	1026 (28.9%)	3553 (96.4%)
Road defects involved					
Yes	3 (6.7%)	42 (93.3%)	0 (0.0%)	0 (0.0%)	45 (1.2%)
No	179 (4.9%)	1162 (31.9%)	1260 (34.6%)	1038 (28.5%)	3639 (98.8%)
Crash location					
Non-motorized lanes	15 (2.0%)	208 (27.7%)	246 (32.8%)	282 (37.5%)	751 (20.4%)
Mixed lanes	71 (5.6%)	354 (27.7%)	432 (33.8%)	420 (32.9%)	1277 (34.7%)
Motorized lanes	96 (6.6%)	566 (39.1%)	540 (37.3%)	246 (17.0%)	1448 (39.3%)
Others	0 (0.0%)	76 (36.5%)	42 (20.2%)	90 (43.3%)	208 (5.6%)
Escape					
Yes	13 (4.3%)	116 (38.3%)	90 (29.7%)	84 (27.7%)	303 (8.2%)
No	169 (5.0%)	1088 (32.2%)	1170 (34.6%)	954 (28.2%)	3381 (91.8%)
e-bicyde behavior					
Going straight	105 (4.0%)	714 (27.1%)	1074 (40.8%)	738 (28.1%)	2631 (71.4%)
Turning Left	33 (7.7%)	220 (51.5%)	78 (18.3%)	96 (22.5%)	427 (11.6%)

(continued)

Table 1. Continued

Variables	Fatal injury	Treatment injury	No treatment injury	Possible/no injury	Total
Turning Right	5 (3.9%)	28 (21.7%)	42 (32.6%)	54 (41.9%)	129 (3.5%)
Crossing streets	37 (9.2%)	222 (55.1%)	36 (8.9%)	108 (26.8%)	403 (10.9%)
others	2 (2.1%)	20 (21.3%)	30 (31.9%)	42 (44.7%)	94 (2.6%)
Vehide behavior					
Going straight	122 (5.0%)	822 (33.4%)	840 (34.1%)	678 (27.5%)	2462 (66.8%)
Turning Left	19 (4.4%)	108 (24.9%)	156 (36.0%)	150 (34.6%)	433 (11.8%)
Turning Right	41 (7.3%)	204 (36.2%)	180 (32.0%)	138 (24.5%)	563 (15.3%)
Crossing streets	0 (0.0%)	38 (36.5%)	36 (34.6%)	30 (28.8%)	104 (2.8%)
others	0 (0.0%)	32 (26.2%)	48 (39.3%)	42 (34.4%)	122 (3.3%)
Control characteristics					
Roadway section	87 (4.7%)	618 (33.3%)	696 (37.5%)	456 (24.6%)	1857 (50.4%)
Signal intersection	33 (10.2%)	136 (41.8%)	78 (24.0%)	78 (24.0%)	325 (8.8%)
No signal intersection	40 (4.2%)	264 (27.9%)	324 (34.2%)	318 (33.6%)	946 (25.7%)
Access	27 (4.8%)	186 (33.2%)	162 (28.9%)	186 (33.2%)	561 (15.2%)
Geometry characteristics					
Asphalt road					
Yes	95 (4.6%)	750 (36.2%)	636 (30.7%)	588 (28.4%)	2069 (56.2%)
No	87 (5.4%)	454 (28.1%)	624 (38.6%)	450 (27.9%)	1615 (43.8%)
Road geometry					
Straight	147 (4.4%)	1088 (32.2%)	1158 (34.3%)	984 (29.1%)	3377 (91.7%)
Other	35 (11.4%)	116 (37.8%)	102 (33.2%)	54 (17.6%)	307 (8.3%)
Road class type					
Local roads	13 (3.2%)	138 (34.2%)	102 (25.3%)	150 (37.2%)	403 (10.9%)
Minor arterial	12 (2.0%)	214 (35.4%)	192 (31.8%)	186 (30.8%)	604 (16.4%)
Principle arterial	85 (3.8%)	704 (31.4%)	858 (38.3%)	594 (26.5%)	2241 (60.8%)
Rural highways	72 (16.5%)	148 (33.9%)	108 (24.8%)	108 (24.8%)	436 (11.8%)
No. of traffic lanes					
1	72 (6.8%)	406 (38.3%)	336 (31.7%)	246 (23.2%)	1060 (28.8%)
2	58 (4.6%)	328 (26.1%)	474 (37.7%)	396 (31.5%)	1256 (34.1%)
3	43 (3.9%)	392 (35.4%)	360 (32.5%)	312 (28.2%)	1107 (30.0%)
4 +	9 (3.4%)	78 (29.9%)	90 (34.5%)	84 (32.2%)	261 (7.1%)
Bicycle-motor separation type					
Mixed traffic	71 (4.1%)	428 (24.7%)	636 (36.7%)	600 (34.6%)	1735 (47.1%)
Line separation	15 (2.3%)	142 (22.1%)	282 (43.9%)	204 (31.7%)	643 (17.5%)
Barrier separation	3 (6.4%)	8 (17.0%)	24 (51.1%)	12 (25.5%)	47 (1.3%)
Tree separation	93 (7.4%)	626 (49.7%)	324 (25.7%)	216 (17.2%)	1259 (34.2%)
Motor separation type					
Line separation	82 (6.2%)	424 (32.1%)	438 (33.1%)	378 (28.6%)	1322 (35.9%)
Barrier separation	65 (5.9%)	416 (37.6%)	354 (32.0%)	270 (24.4%)	1105 (30.0%)
Tree separation	8 (2.5%)	68 (21.1%)	138 (42.9%)	108 (33.5%)	322 (8.7%)
other	27 (2.9%)	296 (31.7%)	330 (35.3%)	282 (30.2%)	935 (25.4%)
Land use characteristics					
Land use					
Residential area	51 (3.7%)	428 (30.6%)	504 (36.1%)	414 (29.6%)	1397 (37.9%)
Institutional area	14 (4.5%)	126 (40.9%)	78 (25.3%)	90 (29.2%)	308 (8.4%)
Commercial area	18 (3.7%)	212 (43.4%)	120 (24.6%)	138 (28.3%)	488 (13.2%)
Industrial area	14 (3.6%)	96 (24.9%)	180 (46.6%)	96 (24.9%)	386 (10.5%)
Farm/pasture	54 (12.9%)	96 (22.9%)	150 (35.7%)	120 (28.6%)	420 (11.4%)
Other	31 (4.5%)	246 (35.9%)	228 (33.3%)	180 (26.3%)	685 (18.6%)
Development					
Less than 30%developed	72 (8.7%)	312 (37.7%)	282 (34.1%)	162 (19.6%)	828 (22.5%)
30%–50% developed	67 (4.6%)	468 (32.3%)	516 (35.7%)	396 (27.4%)	1447 (39.3%)
50%–80% developed	38 (3.9%)	308 (32.0%)	312 (32.4%)	306 (31.7%)	964 (26.2%)
More than 80% developed	5 (1.1%)	116 (26.1%)	150 (33.7%)	174 (39.1%)	445 (12.1%)
Urban area					
Yes	90 (3.3%)	908 (32.9%)	1026 (37.1%)	738 (26.7%)	2762 (75.0%)
No	92 (10.0%)	296 (32.1%)	234 (25.4%)	300 (32.5%)	922 (25.0%)
Temporal characteristics					
Weekend					
Yes	45 (3.7%)	306 (25.3%)	498 (41.2%)	360 (29.8%)	1209 (32.8%)
No	137 (5.5%)	898 (36.3%)	762 (30.8%)	678 (27.4%)	2475 (67.2%)

(continued)

Table 1. Continued

Variables	Fatal injury	Treatment injury	No treatment injury	Possible/no injury	Total
Time					
0:00–6:59	15 (5.3%)	106 (37.5%)	90 (31.8%)	72 (25.4%)	283 (7.7%)
7:00–8:59	22 (4.7%)	132 (28.0%)	180 (38.1%)	138 (29.2%)	472 (12.8%)
9:00–11:59	31 (6.0%)	186 (36.0%)	162 (31.3%)	138 (26.7%)	517 (14.0%)
12:00–13:59	33 (6.1%)	148 (27.4%)	192 (35.5%)	168 (31.1%)	541 (14.7%)
14:00–16:69	31 (4.7%)	246 (37.6%)	216 (33.0%)	162 (24.7%)	655 (17.8%)
17:00–18:59	15 (2.9%)	162 (31.6%)	180 (35.1%)	156 (30.4%)	513 (13.9%)
19:00–23:59	35 (5.0%)	224 (31.9%)	240 (34.1%)	204 (29.0%)	703 (19.1%)
Environmental characteristics					
Weather					
Clear	117 (5.0%)	742 (31.8%)	792 (33.9%)	684 (29.3%)	2335 (63.4%)
Cloudy	32 (5.6%)	202 (35.4%)	198 (34.7%)	138 (24.2%)	570 (15.5%)
Other (fog, rain, etc.)	33 (4.2%)	260 (33.4%)	270 (34.7%)	216 (27.7%)	779 (21.1%)
Light					
Daylight	122 (4.6%)	828 (31.0%)	942 (35.3%)	780 (29.2%)	2672 (72.5%)
Dark-streetlight	31 (4.0%)	244 (31.8%)	258 (33.6%)	234 (30.5%)	767 (20.8%)
Dark-no streetlight	29 (11.8%)	132 (53.9%)	60 (24.5%)	24 (9.8%)	245 (6.7%)
Road surface					
Dry	139 (4.9%)	968 (33.9%)	942 (32.9%)	810 (28.3%)	2859 (77.6%)
Other	43 (5.2%)	236 (28.6%)	318 (38.5%)	228 (27.6%)	825 (22.4%)
Visibility					
<50 m	33 (9.4%)	150 (42.7%)	120 (34.2%)	48 (13.7%)	351 (9.5%)
50–100 m	37 (4.7%)	306 (38.9%)	264 (33.5%)	180 (22.9%)	787 (21.4%)
100–200 m	28 (3.9%)	216 (29.8%)	258 (35.6%)	222 (30.7%)	724 (19.7%)
>200 m	84 (4.6%)	532 (29.2%)	618 (33.9%)	588 (32.3%)	1822 (49.5%)
Total	182 (4.9%)	1204 (32.7%)	1260 (34.2%)	1038 (28.2%)	3684

$$Y_i = \begin{cases} 0\,if\ -\infty < Y_i^* < \omega_1 & possible\ and\ no\ injury \\ 1\,if\ \omega_1 < Y_i^* < \omega_2 & no\ treatment\ injury \\ 2\,if\ \omega_2 < Y_i^* < \omega_3 & treatment\ injury \\ 3\,if\ Y_i^* > \omega_3 & fatality \end{cases} \quad (2)$$

where vk is an unknown threshold to be estimated along with parameter a, aexplains how explanatory variables affect the underlying injury severity level

$$\omega_k < Y_n^* < \omega_{k+1} \Rightarrow \omega_k < a^*X + \in \leq \omega_{k+1}$$
$$\Leftrightarrow \omega_k - a^*X < \in \leq \omega_{k+1} - a^*X \quad (3)$$

Since e is assumed to obey logistic distribution in ordered logit model, the probability of the response variable at level k can be computed as

$$P(Y_i = k) = P(\omega_k < Y_i^* \leq \omega_{k+1}) = 1 - P(Y_i^* > \omega_{k+1}) - 1 + P(Y_i^* > \omega_k) \quad (4)$$

where

$$P(Y_i^* > \omega_k) = \frac{\exp(aX'_i - k_j)}{1 + \exp(aX'_i - k_j)} \quad k = 1, 2, 3 \quad (5)$$

For the ordered logit model, the relationship between any pairs of outcome categories is assumed to be equal. In other words, the thresholds have to be fixed for all explanatory variables. If this assumption (i.e. the parallel line assumption) is sometimes violated by some explanatory variables that they may have different effects on thresholds with other variables, the ordered logit model could result in biased estimates.21 To address this, a generalized ordered logit/partial proportional odds model can be used instead that relax the parallel line assumption. Moreover, it is more parsimonious than traditional multinomial logit/probit models that are often suffer from independence of irrelevant alternative (IIA) issues.20 The probability of Y* larger than a specific threshold can be specified as

$$P(Y_i^* > \omega_k) = \frac{\exp(X'_{1i}b_1 + X'_{2i}b_{2j} - k_j)}{1 + \exp(X'_{1i}b_1 + X'_{2i}b_{2j} - k_j)} \quad k = 1, 2, 3 \quad (6)$$

where b1 is a vector of parameters of variables that follow the parallel line assumption and is associated (Xli), and b2j is a vector of parameters of variables that vary across different severity levels (X2i). If no variables were found violating the parallel line assumption, the model will degenerate to the ordinary ordered logit model (i.e.equations (1)–(5)) and all variables will have constant coefficients across all injury levels. The probability of injury severity has a closed-form expression20 and the log-likelihood function can be expressed by

$$\ln(L) = \sum_{n=1}^{N} \sum_{k=1}^{3} K_{nk} \ln P(Y_i^* > \omega_k) \quad (7)$$

where ln(L) is the log-likelihood function, Knk is equal to 1 if crash n results in severity level k and 0 otherwise, and N is the number of crashes. Variable coefficients were obtained by maximize the likelihood function, given the observations. The marginal effect of a variable indicates how the factor affects the response variable on the underlying scale. For continuous variables, marginal effects can be obtained by calculating derivatives. For dummy variables, a difference rather than the derivative is calculated. The goodness of fit for the partial proportional odds model is measured by the likelihood ratio index. The likelihood ratio index can be calculated by

$$R^2 = 1 - \left[\frac{\ln L_a}{\ln L_0}\right] \quad (8)$$

where ln La is the log-likelihood with all predictors included in the model (at convergence); ln L0 is the loglikelihood value of the models containing only intercept (at constants); The likelihood ratio is improved when the value of index increases from 0 to 1.22 Akaike information criterion (AIC) can be utilized to evaluate model's performance. The specification of AIC can be written as

$$AIC = -2 \ln L_a + 2P \quad (9)$$

where P is the number of parameters estimated.

IV. RESULTS

Table 2 shows the estimated e-bicyclist injury severity model, including the parameter estimates, standard errors, and level of significance. According to modeling results, a number of factors have been identified to be significantly associated with injury severity sustained by e-bicyclists. A series of Wald tests were applied to determine the validity of parallel line assumption.20 Variables with p value less than 0.05 are considered to have heterogeneous effects across different severity levels. Table 2 shows that all variables have p values larger than 0.05, indicating that no variable included in the final model violates the parallel line assumption. In this case, the generalized ordered logit model provides the same estimates with an ordinary ordered logit model. Despite of that, the generalized ordered logit model was still considered as superior than the ordered logit model, because it was able to detect the existence of heterogeneity effects across observations (no such existencewas found in this study). Moreover, considering injury severity as an ordinal instead of nominal variable is intuitively a more reasonable model choice, since there is a clear ordering from non-injury to fatality Thus, the generalized ordered logit model is a better choice than traditional multinomial models, in terms of modeling crash severity.

e-bicyclist characteristics

e-bicyclists aged 55 years and over are more prone to be injured than younger e-bicyclists. This result is consistent with other studies that older adults are more likely to suffer fatal injuries in bicycle crashes.23–25 However, many unobserved variables could be correlated with age.26 For example, reaction time, physical fragility, and bone density could also contribute to higher injury probability. The result shows that males are associated with lower injury severity. The reason, however, could not be revealed based on data. A possible reason could be the difference in physical fragility between male and female.

Driver characteristics

When drivers are intoxicated, the probabilities of ebicyclists' injuries largely increase. Similar results can be found in Noland and Quddus'27 research, they found that drinking drivers had significant associations with bicycle injury severity. Drivers with less than 2 years of experience are more likely to be severely injured. Lack of experience could be a reason: inexperienced drivers could be more likely to mistakenly conduct improper maneuvers in emergency conditions.

Vehicle characteristics

The vehicle type variables were also found to be significant correlated to e-bicyclists' injury severity. Unsurprisingly, heavy trucks could be expected to be related to higher injury severity. The similar result is found by McCarthy and Gilbert28 that the heavy trucks were more likely to be related to fatal bicycle crashes, since heavy trucks have greater momentum than passenger car. Motorcycles are more likely to be involved in severe e-bicycle crashes. The reason could be due to frequent speeding offense of motorcycles in China, as claimed by Guilin Police.

Crash characteristics

Head-on crashes have a positive relationship with the probability of severely injured in crashes. When considering fault, Table 1 show that e-bicyclists are found jointly at fault most of the time. This result is consistent with Kim et al.'s19 study that cyclists are more likely to commit faults than motorists. The model results from Table 2 indicate that e-bicyclists are more likely to be severely injured when they are solely at fault. Crashes with e-bicyclists at fault often occur at motorized lanes, probably resulting in more severe injuries. When driving on motorized lanes, e-bicyclists have higher risk of being injured. e-bicycle riders were found to frequently drive in motorized lanes, increasing their exposure of being crashed by motor vehicles.

Table 2. Parameter estimation of generalized ordered logit model.

	Estimates	Standard deviation	Significance	Wald tests (p value)
Threshold parameters				
Fatal injury	–0.439	0.199	0.027	
Treatment injury	1.572	0.211	0.000	
No treatment injury	2.895	0.513	0.008	
e-bicyclist characteristics				
e-bicyclist age 55 + years	1.523	0.545	0.005	0.2123
e-bicyclist is female	0.249	0.110	0.023	0.3476
Driver characteristics				
Driver is intoxicated	1.429	0.315	0.000	0.4511
Driver years <3	0.525	0.231	0.033	0.8781
Vehicle characteristics				
Motorcycle	0.870	0.221	0.004	0.2310
Heavy truck	2.703	0.268	0.000	0.5623
Crash characteristics				
Head-on crash	0.543	0.207	0.047	0.5521
e-Bicyclist at fault	1.782	0.188	0.000	0.3295
Speeding	0.191	0.083	0.032	0.1233
Crash in motorized lanes	1.186	0.287	0.000	0.1587
Behavior characteristics				
e-bicyclist turning left	0.988	0.404	0.014	0.4368
e-bicyclist crossing the road	1.692	0.497	0.001	0.5629
Driver turning right	1.225	0.209	0.000	0.7731
Geometry characteristics				
Signal intersection	0.304	0.1 23	0.014	0.2361
Curved road	0.727	0.224	0.001	0.1137
Rural highways	1.254	0.330	0.000	0.0945
No. of traffic lanes = 1	0.279	0.068	0.031	0.2466
Trees separation between motor vehicle and e-bicycle	1.470	0.560	0.009	0.4821
Two-way divided by trees separation	–0.700	0.231	0.002	0.2264
Two-way divided by barriers	–0.855	0.322	0.008	0.0742
Temporal characteristics				
Weekday	0.434	0.148	0.003	0.5366
p.m. peak	–0.643	0.297	0.030	0.8615
Land characteristics				
Development	–0.462	0.122	0.000	0.4672
Residential area	–0.716	0.273	0.009	0.8463
Industrial area	0.543	0.207	0.047	0.5923
Log-likelihood at convergence		–3887.1		
Log-likelihood at constant		–5592.6		
Pseudo R^2		0.315		
AIC		7830.2		

AIC: Akaike information criterion.

When speeding offense is found in e-bicycle crashes, the injury severity of e-bicyclists significantly increased.This finding is reasonable since, intuitively, larger speed could result in higher impact force, causing more severe injuries to e-bicyclists.

Behavior characteristics

e-bicyclists are more likely to be severely injured when they are making left turns or crossing streets. When making such moves, e-bicyclists have to pass motorized lanes, increasing their exposure. Meanwhile, drivers could underestimate the speed of e-bicycles, resulting in severe crashes. When e-bicycle riders make right turns, their injury risk also increases. This could be partly attributed to the right-turn-on-red rules. Drivers could ignore the presence of e-bicycle riders when they make right turns on red. Another possible reason could be that the right-turn radius of intersections is designed to be relatively larger in China, causing relatively higher turning speed. In this case, e-bicyclists could underestimate the operating speed of vehicles. Moreover, the difference of radius between inner wheels could be another factor. Especially for large trucks, drivers could consider that they successfully overtake e-bicycles, after the front wheels of vehicles passing e-bicycles. However, the rear wheel could still crush e-bicyclists due to the difference of radius between inner wheels.

Geometry characteristics

Signalized intersections are correlated with increased risk of high injury severity of e-bicyclists. Red-light running is a frequent illegal behavior conducted by e-bicycle riders, which could significantly increase their risk. Ebicyclists are more likely to be injured when they are driving on curves. The increased maneuvering difficulty and reduced vision on curves could be a factor. Rural highways are correlated with higher injury severity of ebicycle riders. Rural highways are set to have relatively higher speed limits than urban roads. Moreover, lighting conditions in rural highways could be worse than urban highways. Single-lanes are correlated with higher probability of severe injuries. The possible reason could be due to the limited room increasing potential conflicts between e-bicycles and motor vehicles. With the presence of trees separation between motorized and non-motorized vehicles, severe crashes are more likely to occur. With trees separation, drivers could be less cautious to e-bicyclists, increasing traffic risk when e-bicyclists are crossing streets. To tackle this, deploying warning signs at the mid-block openings could be a viable option. Two-way divided roadways are associated with higher probability of low injury severity of e-bicyclists. It could be that e-bicyclists are less likely to driver in opposite directions on divided roads which leads to fewer head-on crashes. Barriers are correlated with less severe injuries, compared to trees separation. With barriers, both e-bicycle riders and drivers have better sights and visions, when they make turns or crossing streets. Temporal characteristics e-bicyclists are more likely to be severely injured on weekdays. This result is contradictory with the previous research in other countries that weekend crashes were found to result in a higher probability of increased injury severity.22,24 e-bicycles are extensively used for daily commuting in China. Thus, e-bicycle riders could have higher exposure during weekdays, which could be a potential reason. e-bicycle riders were less likely to be severely injured during evening peak hours (17:00–19:00). Thanks to high traffic volume, motor vehicles move slowly during this period, probably resulting in lower pre-crash speed. Land use characteristics In this article, a land development intensity variable was defined as an index that ranges from 0 to 1, with 1 indicating high intensity of development and 0 indicating low intensity of development. The model results show that land development intensity is negatively associated with the injury severity of e-bicyclists. Within highly-developed area, traffic volume is relatively large, possibly resulting in low vehicle speed. Moreover, in such area, walking and public transit facilities are also well developed, lowering the share of e-bicycle usage. Commercial areas are associated with the decreased injury severity of e-bicycle riders. This is presumably due to well-developed walking and public transportation facilities within such area, such as quality streetlights roadway markings and dedicated e-bike facilities, lowering e-bicycle usage. Residential area is associated with the decreased injury severity of e-bicyclists. It could be the reason that
the speed limit within such area is set to be relatively low due to the high population and building density.

Another possible reason could be drives' compensatory behaviors with the anticipation of encountering nonmotorists (e.g. pedestrians and cyclists) within such area. Industrial area is linked to high injury severity. Based on our observations, these areas lack of traffic facilities for non-motorists, which could be a possible reason. Since commercial large trucks are predominant in this area, they could also be associated with increased injury severity of e-bicyclists.

V. CONCLUSION

This study provides a comprehensive assessment of the factors influencing the severity of e-bike crashes in China using a Generalized Ordered Logit Model. The results reveal that crash severity is significantly affected by a variety of factors, including rider age, helmet use, lighting conditions, intersection presence, and the type of colliding vehicle. In particular, crashes involving older riders or heavy vehicles, as well as those occurring at night or at complex intersections, tend to result in more severe injuries.

The flexibility of the GOLM allows for a more accurate representation of how each variable influences the likelihood of progressing from one level of severity to another, offering a valuable improvement over traditional ordered logit models. The insights generated from this model can inform targeted interventions, such as helmet mandates, improved street lighting, e-bike lane segregation, and enhanced education for vulnerable rider groups.

In conclusion, this research underscores the importance of context-specific, data-driven strategies to enhance e-bike safety in China. Future work could integrate spatial analysis and real-time traffic data to further refine crash prediction models and support adaptive safety policy design.

REFERENCES

1. Xinhua News. Electric bike ownership in China was over 200 million [DB/OL], http://news.ddc.net.cn/newsview_65356.html (in Chinese)
2. Cherry C, Weinert J and Yang X. Comparative environmental impacts of electric bikes in China. Transp Res Part D 2009; 14: 281–290.
3. China Road Traffic Accidents Statistics Report (CRTASR). Traffic administration bureau of China state security ministry. Beijing, China: CRTASR, 2014.
4. Transport Research and Injury Prevention Programme. The road traffic injury prevention training manual (publicationWA 275). New Delhi, India: Transport Research and Injury Prevention Programme, 2006.
5. Muhammad AH and Justin JF. Modeling bicyclists' injury severity levels in the province of Nova Scotia, Canada using a generalized ordered probit structure. In:Proceedings of the 93th annual meeting of transportation research board, Washington, DC, 2014, http://docs.trb.org/prp/14-5303.pdf
6. Weinert JX, Ma C, Yang X, et al. Electric two-wheelers in China: effect on travel behavior, mode shift, and user safety perceptions in a medium-sized city. J Transp Res Board 2007; 2038: 62–68.
7. Feng Z, Raghuwanshi RP and Xu Z. Electric-bicyclerelated injury: a rising traffic injury burden in China.Injury Preven 2010; 16: 417–419.
8. Yao L and Wu C. Traffic safety of e-bike riders in China: safety attitudes, risk perception, and aberrant riding behaviors. In: Proceedings of the 91th annual meeting compendium transportation research board of the national academies, 2012, http://docs.trb.org/prp/12-4599.pdf
9. Wu C, Yao L and Zhang K. The red-light running behavior of electric bike riders and cyclists at urban intersections in China: an observational study. Accident Anal Prev 2012; 9: 186–192.
10. Bai L, Liu P, Chen Y, et al. Comparative analysis of the safety effects of electric bikes at signalized intersections. Transp Res Part D 2013; 20: 48–54.
11. Hu F, Lv D, Zhu J, et al. Related risk factors for injury severity of E-bike and bicycle crashes in Heifei. Traffic Inj Prev 2014; 15: 319–323.
12. Bai L, Chan CY, Liu P, et al. Identifying factors affecting the safety of mid-block bicycle lanes considering mixed 2-wheelel traffic flow. Traffic Inj Prev 2017; 18: 761–766.
13. Yang X, Huan M, Aty M, et al. A hazard-based duration model for analyzing crossing behavior of cyclists and electric bike riders at signalized intersections. Accident Anal Prev 2015; 74: 33–41.
14. Zhang YQ and Wu CX. The effects of sunshields on red light running behavior of cyclists and electric-bike riders. Accid Anal Prev 2013; 52: 210–218.
15. Dan Z, Ma XL, Jin S, et al. Modeling influencing factors of vehicle passing rate in mixed bicycle traffic flow. J Zhejiang Univ 2015; 49: 1672–1678.