

Examining partial proportional odds model in analyzing severity of high-speed railway accident

Jing Wang, Yinghan Wang, Yichuan Peng and Jian John Lu
*The Key Laboratory of Road and Traffic Engineering, Ministry of Education,
Tongji University, Shanghai, China*

Received 16 October 2020
Revised 22 October 2020
Accepted 22 October 2020

Abstract

Purpose – The operation safety of the high-speed railway has been widely concerned. Due to the joint influence of the environment, equipment, personnel and other factors, accidents are inevitable in the operation process. However, few studies focused on identifying contributing factors affecting the severity of high-speed railway accidents because of the difficulty in obtaining field data. This study aims to investigate the impact factors affecting the severity of the general high-speed railway.

Design/methodology/approach – A total of 14 potential factors were examined from 475 data. The severity level is categorized into four levels by delay time and the number of subsequent trains that are affected by the accident. The partial proportional odds model was constructed to relax the constraint of the parallel line assumption.

Findings – The results show that 10 factors are found to significantly affect accident severity. Moreover, the factors including automation train protection (ATP) system fault, platform screen door and train door fault, traction converter fault and railway clearance intrusion by objects have an effect on reducing the severity level. On the contrary, the accidents caused by objects hanging on the catenary, pantograph fault, passenger misconducting or sudden illness, personnel intrusion of railway clearance, driving on heavy rain or snow and train collision against objects tend to be more severe.

Originality/value – The research results are very useful for mitigating the consequences of high-speed rail accidents.

Keywords High-speed railway, Contributing factors, Accident severity, Partial proportional odds model
Paper type Research paper

Introduction

Safety is the primary and critical consideration in the operation of high-speed railway. However, accidents are inevitable due to the systematic complexity, the unpredictability of

© Jing Wang, Yinghan Wang, Yichuan Peng and Jian John Lu. Published in *Smart and Resilient Transportation*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence maybe seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

This research has been supported by projects of the National Key R&B Program of China (no. 2018YFB1201403).

The authors confirm contribution to the paper as follows: Study conception and design: Yinghan Wang, Yichuan Peng. Data collection: Yichuan Peng. Analysis and interpretation of results: Yinghan Wang, Yichuan Peng, Jian John Lu. Draft manuscript preparation: Yinghan Wang, Yichuan Peng, Jian John Lu, Jing Wang. All the authors reviewed the results and approved the final version of the manuscript.



humans (both operators and passengers) and environment uncertainty. In 2005, Japan's high-speed railway derailment accident caused 107 deaths and 555 injuries. In 2011, the rear-end collision of two high-speed trains in China caused 39 deaths and 191 injuries (Wang, 2014). Similar accidents also occurred in other countries such as South Korea, France and Germany and caused casualties.

Those accidents were attributed to various factors such as train operator's mistake, malfunction of facilities including on-board devices in the train control system and traction power supply equipment and environment conditions such as thunder-strike. If the factors affecting high-speed railways safety are identified, countermeasures can be taken to reduce casualties and property damage. Therefore, plenty of researchers from different institutions and country governments have paid much attention to investigate the contributing factors affecting railway safety to provide theoretical and applicable contributions to safety analysis and accident management. In general, human, equipment and facility and environmental conditions are the most concentrated factors.

Human factors especially the individual characteristics of railway workers have been proved to significantly affect the safety of railway operation (Baysari *et al.*, 2008). The statistics data in Europe pointed out that 75% of fatal railway accidents were caused by human errors (Evans, 2014). In Iran, train drivers' faults account for 47% of human-blamed accidents (Iran Bureau of Rail Mobility, 2015). Guo *et al.* (2016) examined the relationships between train drivers' personalities (openness to experience, neuroticism, extraversion, conscientiousness and agreeableness) and driving safety of the high-speed railway. The results showed that train drivers' personalities are related to the frequency of accidents. Another study conducted by Guo *et al.* (2019) revealed that job insecurity of high-speed railway drivers affected their safety performance. Train drivers' sustained attention also plays a determinative role in railway safety (Hani Tabai *et al.*, 2018). In addition, it should be noted that passengers' inadequate behaviors such as smoking in the train carriage or pulling down the emergency braking valve can also jeopardize the safety operation of high-speed railway. However, these factors were hardly taken into consideration by most researchers.

A high-speed railway is a highly complexed and coupled system involving many diversities of equipment and facilities. In common, infrastructures and on-board equipment involved in the high-speed railway system include track and wheel-rail system, train control system, traction power supply system, signal and communication system, etc. Ensuring the normal running of this equipment and facilities is critical to operation safety. A track plays a fundamental role in the infrastructure of the railway system. A high percentage of failure of tracks such as track geometry degradation requires adequate assessment and preventive maintenance to keep safety and reliable operation (Ižvolta and Šmalo, 2015; Hassankiadeh, 2011). A track circuit is a key component of the signal system of the railway, providing the information of trains' location and movements to ensure operation safety and driving efficiency (Wybo, 2018). The accurate fault prediction of track circuit can improve the consequential accidents (Hu *et al.*, 2019). Wu *et al.* (2018) analyzed the impact of axle fatigue damage on the safety of high-speed railway operation. They assessed the impact of axle surface scratches on fatigue performance by analyzing the depth of scratches. The catenary is an important component of the traction power supply system of the high-speed railway which is responsible for power supply to electric locomotives through pantographs. The structure integrity and dynamic performance of the catenary system directly affect driving safety and speed (Han *et al.*, 2018; Liu *et al.*, 2018).

Environment conditions also have potential risks to railway safety because the working environment of the railway is outdoor. Natural environment including extreme weather conditions (strong wind, rainstorm, snowstorm, etc.) and geological disasters (soil

settlement, debris flow, landslides, earthquakes, etc.) can directly result in railway accidents such as train derailment or indirectly cause equipment damage, thereby endangering railway safety. One of the reasons for a super railway accident in China in 2011 is that the train control system was damaged by a thunder-strike. In a study did by [Dindar et al. \(2017\)](#), derailments at railway turnout caused by extreme weather conditions were analyzed through Bayesian Network. The natural hazards (temperature change, sea level raise, roadbed settlement, flood, etc.) brought by global warming were also found to have a potential effect on the railway turnout system ([Dindar et al., 2016](#)). In addition to force majeure of the natural environment, the influence of the social environment (public security at stations, management along the railway line, related laws and regulations, etc.) on the operation safety of high-speed railway is becoming more and more prominent ([Guanhua, 2018](#)).

Most of the studies mentioned above were concentrated on a single factor that may result in an accident in the operation of the railway. However, the operation of the railway is a complex system involving many factors working simultaneously. A comprehensive analysis and evaluation of safety factors to examine what extent each factor involved contributes to accident severity is necessary and significant. Besides, previous studies on analyzing accident severity using ordered probability models such as ordered logit are usually constrained by proportional odds assumption (some even neglected it). A generalization method of the ordered logit model, the partial proportional odds model is applied in this research to relax the proportional odds constraint.

This paper aims to identify factors affecting accident severity during the process of high-speed railway operation based on historical data gathered from Railway Bureaus and quantitatively analyze contributions of each factor to accident severity. Accident severity is divided into four categories from low to high (I, II, III and IV) according to the delay time and the number of subsequent trains influenced by the accident. Human, facility and equipment and environment factors are simultaneously considered in the modeling. Considering the high automation and integration of equipment is a technical development in high-speed railway systems ([Crawford and Kift, 2018](#)), human injury or fatality accidents are less reported in these years and can hardly found in the data collected. Therefore, only accidents with no injury or fatality are taken into account in this research.

Methodology

Accident severity in this research is a discrete ordinal categorical variable with four levels. Therefore, an ordered discrete choice model was the primary choice in this study. Ordered probability models were applied in many road traffic crashes severity analysis. Numerical researchers implemented ordered probability models such as ordered probit models ([Abdel-Aty, 2003](#); [Lee and Abdel-Aty, 2005](#); [Ju, 2006](#)) and ordered logit models ([Yasmin and Eluru, 2013](#); [Feng, 2015](#); [Rezapour et al., 2019](#)) to severity research studies. The ordered logit model (also called the proportional odds model) was firstly proposed by [Walker and Duncan \(1967\)](#), which was widely used in accident severity analysis because of the ordinal nature of accident severity. The ordered logit model can be written as:

$$P(Y_i > j) = \frac{\exp(\alpha_j + X_i\beta)}{1 + [\exp(\alpha_j + X_i\beta)]}, \quad j = 1, 2, \dots, M - 1 \quad (1)$$

where $P(Y_i > j)$ is the probability of accident severity for a given accident i . j represents the number of cut points. α_j is the regression intercept of each cut point. β is the vector of the

regression coefficient which does not change across different logits. X is the vector of explanatory variables. M is the number of severity categories.

However, a very strict constraint of using the ordered logit model is proportional odds assumption, also known as parallel line assumption (McCullagh, 1980). This assumption indicates that the effects of explanatory variables on different severity levels of the dependent variable are the same in each cumulative logit. To be more specific, the regression parameter β remains the same across different cut points. Brant test (Brant, 1990) is often used to evaluate if a parallel line assumption is violated.

It is apparent that assuming the effects of explanatory variables remain the same among each severity level cannot be met all the time. That is because the distance between each severity level is not equal sometimes (Wang and Abdel-Aty, 2008). As a consequence of this restrictive assumption, the ordered logit model often fails. Sometimes this constraint is neglected in some research studies. However, it causes a misleading result because the effect of the variable who violates the parallel line assumption is overestimated or underestimated (Williams, 2016). An alternate of the ordered logit model is the generalized ordered logit model which relaxes the restriction of parallel line assumption, and can be written as:

$$P(Y_i > j) = \frac{\exp(\alpha_j + X_i \beta_j)}{1 + [\exp(\alpha_j + X_i \beta_j)]}, \quad j = 1, 2, \dots, M - 1 \quad (2)$$

The only difference between them is that the regression coefficients β_j vary across different equations. Unfortunately, the relaxation of the parallel assumption also brings a new problem that increasing the number of regression parameters to be estimated because it allows β to differ.

An intermediate method between these two models is the partial proportional model (McCullagh and Nelder, 1989; Peterson and Harrell, 1990). In the partial proportional model, the parallel line assumption is only relaxed for some variables. In other words, the regression coefficients β of explanatory variables who violate parallel line assumption change across different logits while others remain the same. The partial proportional model can be written as:

$$P(Y_i > j) = \frac{\exp(\alpha_j + X_{i1} \beta_{j1} + X_{i2} \beta_2)}{1 + [\exp(\alpha_j + X_{i1} \beta_{j1} + X_{i2} \beta_2)]}, \quad j = 1, 2, \dots, M - 1 \quad (3)$$

where X_{i1} is the vector of explanatory variables which violate the constraint of parallel line assumption and accompanied by a vector of regression coefficients β_{j1} which varies across the cut points. X_{i2} is the vector of the rest of the explanatory variables with a vector of regression coefficients β_2 .

Another equivalent form (gamma parameterization) of the partial proportional odds model was proposed by Peterson and Harrell (1990). It can be written as:

$$P(Y_i > j) = \frac{\exp(\alpha_j + X_i \beta + T_i \gamma_j)}{1 + [\exp(\alpha_j + X_i \beta + T_i \gamma_j)]}, \quad j = 1, 2, \dots, M - 1 \quad (4)$$

where T_i is the vector of explanatory variables which violate the assumption of proportional odds and associated with a vector of regression coefficients γ_j , which represents the deviations from proportionality. It can be easily seen from the form that the partial proportional odds model will reduce to proportional odds model (ordered logit model) if $\gamma_j = 0$. It is remarkable

that the interpretation of the model should be careful. The total effect of explanatory variables which violate the proportional odds constraint is the sum of β and γ_j .

Logit models are very sensitive to multicollinearity. Multicollinearity refers to the correlation among explanatory variables. It is quite common during the modeling process of regression analysis. Singularity is the extreme condition of multicollinearity which means one of the explanatory variables is the linear combination of some other explanatory variables in a model. The estimation of the model will generate bias if singularity exists. However, if multicollinearity is not very strong among explanatory variables, whose effect can be neglected and parameter estimation is valid and unbiased. Many criteria can be applied to evaluate multicollinearity (Xiaomu, 2010). Tolerance (TOL) and the variance inflation factor (VIF) are widely used. VIF refers to the ratio of the variance between explanatory variables with multicollinearity and variance without multicollinearity. TOL is the reciprocal of VIF and defined as $TOL = 1 - R_i^2$. Where R_i refers to the coefficient between x_i and the rest of the explanatory variables under the condition that x_i is the dependent variable of other explanatory variables. In general, $VIF > 10$ ($TOL < 0.1$) is an indication of strong multicollinearity.

Pseudo R^2 is a criterion used in logistic regression to describe the proportion of changes in dependent variables explained by explanatory variables just like classical R^2 in linear regression. It is defined as $Pseudo R^2 = 1 - \hat{LL}/\hat{LL}_0$. Where \hat{LL} is the maximum log-likelihood of the fitted model and \hat{LL}_0 is the maximum log-likelihood of the zero model (intercept only).

Data

Accident data was obtained from Railway Bureaus. Accidents in five months were gathered and key information was extracted. Finally, 475 cases of valid data were collected after cleaning the data, excluding invalid and missing values.

One of the most critical information of raw data is the reason for the fault of each accident, which is considered to be the factor resulting in the accident. In total, 14 types of factors affecting high-speed railway accidents were summarized as shown in Table 1. These 14 factors can be summarized into three categories. The statistical distribution of each factor is presented in Figure 1.

Category	Factors	Designation
Facility and equipment factor (F)	ATP system fault	FAC_1
	Turnout indication loss	FAC_2
	Platform screen door or train door fault	FAC_3
	Catenary blackout (trip)	FAC_4
	Objects hanging on catenary	FAC_5
	Pantograph fault	FAC_6
	Traction converter fault	FAC_7
	Carbody's vibration	FAC_8
	Locomotive fault	FAC_9
Human factor (H)	Passenger misconducting or sudden illness	FAC_10
	Personnel intrusion of railway clearance	FAC_11
Environment factor (E)	Driving on heavy rain or snow	FAC_12
	Railway clearance intrusion by objects	FAC_13
	Train collision against objects	FAC_14

Table 1.
Factors resulting in high-speed railway accident

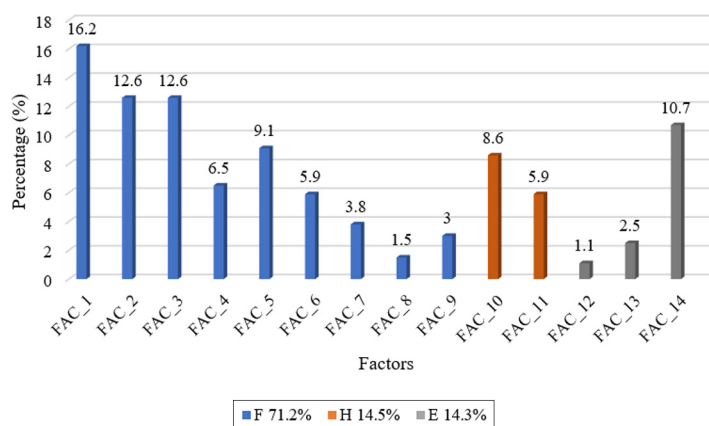


Figure 1.
Distribution of
high-speed railway
accidents by different
factors

Factors related to facility and equipment include nine detailed types. Accidents caused by these factors account for 71.2% of the total. On the one hand, it is because the high-speed railway involves a wide range of equipment and facilities with high technical requirements, complex management and maintenance difficulties. On the other hand, such a high accident rate proves that equipment and facility play a vital role in the operation of high-speed railway. ATP is the abbreviation of an automatic train protection system. It is a safety control system for ensuring train speed is limited to target safety speed. ATP fault (FAC_1) is one of the most frequent risk factors of accidents that must be carefully treated. Turnout indication loss (FAC_2) is a quite common turnout fault referring to no indication of signal when a turnout is a switch to the lateral passing state (Zhang *et al.*, 2018). Platform screen door and train door faults (FAC_3) are mainly linkage failure between them resulted from complex reasons such as signal failure or mechanical failure of equipment. As access for passengers to enter or exit trains, the linkage failure has a direct threat to passengers' safety. Catenary system, pantograph and traction converter are all belong to the traction power supply system of high-speed railway, whose function is to safely and reliably transfer electric energy from the power grid to trains. The main faults of the traction power supply system are concentrated on catenary blackout (FAC_4), objects hanging on catenary (FAC_5), pantograph fault (FAC_6, including an automatic drop of pantograph during operation, objects hanging on the pantograph, etc.) and traction converter fault (FAC_7). Carbody's vibration (FAC_8) refers to the body swaying of a train while the train is in motion (Shanchao *et al.*, 2012). The causation mechanism of this phenomenon is related to the theory of vehicle stability and will not be presented here. The main locomotive faults (FAC_9) include braking failure, halfway parking lead by lack of power, etc.

Accidents caused by human factors account for 14.5% of the total. Two factors including passenger misconducting or sudden illness (FAC_10) and personnel intrusion of railway clearances (FAC_11) are taken into consideration. For FAC_10, some accidents are resulted from improper or even illegal behaviors of passengers, such as smoke alarm triggered by smoking passengers. The condition that passengers with sudden illness need to get off for treatment is also contained in this factor. For FAC_11, residents living along the railway line sometimes invade railway clearance which brings potential dangers to traffic safety. With the improvement of staff management level and automation of high-speed railway, increasingly fewer accidents are caused by operators' errors. In addition, among all the valid accident data, no accident was found to be associated with train operators or other staff of

high-speed railway operation. That is why both human factors are related to non-staff members of high-speed railway.

Accidents led by environment factors account for 14.3% of the total. At present, most of the high-speed railways in China are constructed in plain or hilly areas. The operation safety of the high-speed railway is less affected by topography (Chungang, 2015). Compared with earthquakes, debris flow or other extreme natural disasters, bad weather such as heavy rain or snow (FAC_12) is the most common environment condition to be considered. Railway clearance, which refers in particular to structure clearance here, is a cross-sectional profile perpendicular to the center line of the railway. The intrusion of railway clearance is strictly forbidden to ensure safety operation. In this research, the intrusion of railway clearance by objects (collapsed rock, small animals, etc.) only refers to the intrusion that not collide with trains (FAC_13). If a collision happens, this accident is attributed to train collision against objects (FAC_14).

In China, railway accidents are divided into four severity levels, namely, super major accidents, major accidents, relative major accidents and general accidents according to the standard (Jun *et al.*, 2019). Besides, general accidents are divided into four grades: A, B, C and D. However, no clear standards for the classification of high-speed railway accidents were published in China until now. There are great differences between the high-speed railway and general-speed railway in operation speed, subgrade strength, track radius, traction power supply system and operation safety management. The most intuitive difference is the operation speed. High-speed railway operation speed is generally more than 250 km/h. Faster speed also brings about the improvement of transport capacity. In addition, the high-speed railway is less affected by climate and has a high punctuality rate. High-speed railway equipment is more advanced, automated and safe. Therefore, the original railway accident classification standard cannot fully meet the rapid development of the railway industry, especially the development of the high-speed railway.

The most direct consequences of high-speed railway accidents are the delay of time (DOT) and influence on subsequent trains (IOST). These two effects are easy to be observed and recorded. In addition, it should be clarified that no injury or fatality accidents are observed and recorded. Hence, in this study, a grading method for the severity of general high-speed railway accidents is proposed based on the DOT and the number of trains influenced by the accident (IOST). Severity is categorized into four levels from I to IV with the increase of DOT and IOST. Detailed information of the grading method is given in Table 2. Figure 2 describes the distribution of observed accidents on DOT and IOST. Figure 3 demonstrates the percentage of accidents on each severity level.

Results and discussion

In this research, the partial proportional odds model was fitted in Stata 15.1 with a user-written program, gologit2 (Williams, 2006). After the selection of explanatory variables through mixed stepwise method, 10 variables including FAC_1, FAC_3, FAC_5, FAC_6, FAC_7, FAC_10, FAC_11, FAC_12, FAC_13 and FAC_14 are found to be statistically

Table 2.
Grading method of
high-speed railway
accident

DOT (min)	IOST (train)		
	0 ~ 1	2 ~ 5	>5
0 ~ 10	I	II	III
15 ~ 30	I	II	III
31 ~ 60	I	II	III
>60	I	III	IV

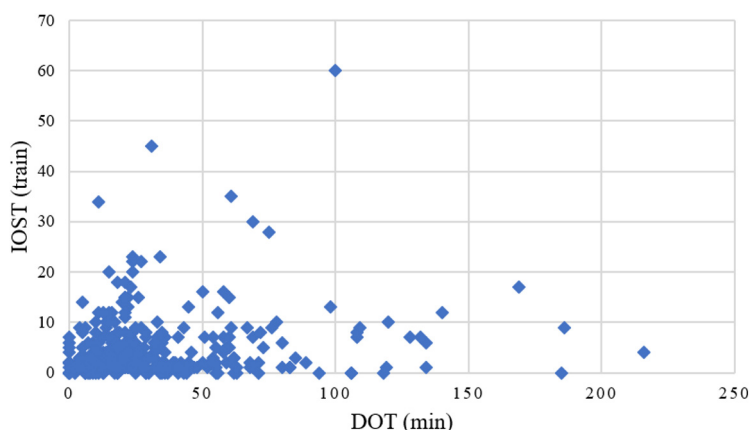


Figure 2. Distribution of observed accidents on DOT and IOST

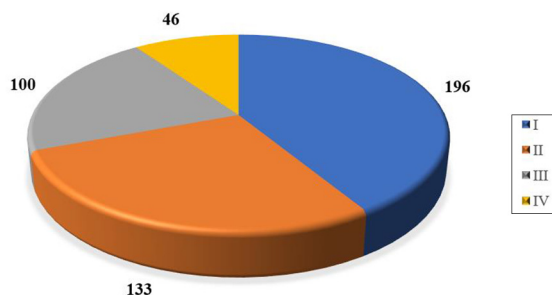


Figure 3. Frequencies of each severity level of accidents

Explanatory variables	TOL	VIF
FAC_1 (Yes = 1; No = 0)	0.707	1.414
FAC_3 (Yes = 1; No = 0)	0.745	1.342
FAC_5 (Yes = 1; No = 0)	0.795	1.259
FAC_6 (Yes = 1; No = 0)	0.850	1.176
FAC_7 (Yes = 1; No = 0)	0.895	1.117
FAC_10 (Yes = 1; No = 0)	0.801	1.248
FAC_11 (Yes = 1; No = 0)	0.850	1.176
FAC_12 (Yes = 1; No = 0)	0.967	1.034
FAC_13 (Yes = 1; No = 0)	0.927	1.079
FAC_14 (Yes = 1; No = 0)	0.770	1.299

Table 3. Multicollinearity check of explanatory variables

significant (p -value < 0.05). All the explanatory variables are binary. The value = 1 indicates the existence of this factor in an accident (Yes). The value = 0 indicates the inexistence of this factor in an accident (No).

The results of the multicollinearity check of explanatory variables are given in Table 3. It can be seen clearly that all the values of VIF are around 1 and far less than 10, which proves that weak multicollinearity exists and the impact of multicollinearity can be neglected. Thus, the parameter estimation results are valid.

A Brant test was conducted to assess if explanatory variables violate the parallel line assumption. The results show that the only FAC_14 violates the proportional odds assumption. The estimation results of the partial proportional model are illustrated in Table 4. All explanatory variables have one beta. Variable FAC_14 who violates the parallel line assumption has two gamma coefficients. Three alpha represents the regression intercepts of each logit function. Table 5 gives the marginal effects of each explanatory variable for accident severity.

The most effective factor identified to affect accident severity is passenger misconducting or sudden illness (FAC_10, Coef. = -2.8364). The negative value indicates that the occurrence of this type of accident brings a relatively low level of accident severity. It can also be proved from

Partial proportional odds model estimates				No. of observations = 475		
Log likelihood = -513.31358				LR χ^2 (11) = 185.40		
				Prob > χ^2 = 0.0000		
				Pseudo R^2 = 0.1530		
Variables	Coef.	Std. err.	z	P > z	[95% Conf. interval]	
<i>Beta</i>						
FAC_1	-0.7257	0.2730	-2.66	0.008	-1.2608	-0.1906
FAC_3	-2.3854	0.3773	-6.32	0.000	-3.1249	-1.6458
FAC_5	0.7453	0.3328	2.24	0.025	0.0931	1.3975
FAC_6	0.8353	0.3875	2.16	0.031	0.0759	1.5947
FAC_7	-1.1391	0.5131	-2.22	0.026	-2.1448	-0.1334
FAC_10	-2.8364	0.5097	-5.57	0.000	-3.8354	-1.8375
FAC_11	0.9757	0.3921	2.49	0.013	0.2072	1.7442
FAC_12	2.3862	0.9102	2.62	0.009	0.6022	4.1701
FAC_13	-1.2388	0.5924	-2.09	0.037	-2.3998	-0.0777
FAC_14	0.9823	0.4461	2.20	0.028	0.1079	1.8567
<i>Gamma_2</i>						
FAC_14	0.1552	0.3854	-2.64	0.008	-0.6001	0.9106
<i>Gamma_3</i>						
FAC_14	-2.8618	1.0840	0.40	0.687	-4.9863	-0.7373
<i>Alpha</i>						
_cons_1	0.8560	0.1829	4.68	0.000	0.4975	1.2145
_cons_2	-0.6993	0.1802	-3.88	0.000	-1.0525	-0.3461
_cons_3	-2.0325	0.2209	-9.20	0.000	-2.4656	-1.5995

Table 4. Estimation results of the partial proportional odds model

Explanatory variables	Severity levels			
	I	II	III	IV
FAC_1	0.1303	-0.0099	-0.0709	-0.0495
FAC_3	0.4133	-0.0047	-0.2190	-0.1896
FAC_5	-0.1291	0.0015	0.0684	0.0592
FAC_6	-0.1447	0.0017	0.0767	0.0664
FAC_7	0.1974	-0.0023	-0.1046	-0.0905
FAC_10	0.4914	-0.0056	-0.2604	-0.2254
FAC_11	-0.1690	0.0019	0.0896	0.0775
FAC_12	-0.4134	0.0047	0.2191	0.1896
FAC_13	0.2146	-0.0025	-0.1137	-0.0985
FAC_14	-0.1702	-0.0246	0.3442	-0.1494

Table 5. Marginal effects of explanatory variables for accident severity

the marginal effect of FAC_12 (0.4914 for Level I and -0.0056 , -0.2604 , -0.2254 for Levels II, III and IV, respectively), which means that the greatest effect of FAC_10 is to reduce accident severity to a low level. In general, the most common measure to deal with passengers with a sudden illness or misconducting behavior in the course of traveling is to contact with corresponding members (police or medical care) in the next station and transfer the passenger to them when the train arrives. Emergency parking is hardly required which saves a lot of time and causes less IOST. In contrast of FAC_10, the other human factor, personnel intrusion of railway clearance (FAC_11) indicates an increase of accident severity level (Coef. = 0.9757). A train driver is required to slow down the running speed to a limited level and report the dispatching center as long as he finds someone intruding railway clearance. Only when the dispatching center confirms that the unrelated personnel have been taken away from the scene will the train speed return to normal. The whole process may cost a relatively long time which consequently causes serious delays.

Driving on heavy rain or snow (FAC_12) is another very effective factor (Coef. = 2.3862). Unlike other natural disasters such as earthquake, flood or landslide which may lead to a great number of injuries or fatalities, heavy rain or snow mainly reduce the running speed of high-speed railway trains. High-speed railway trains are forced to limit running speed if subsystems of the natural disaster monitoring system detect the excess of rainfall or snow depth. Besides, heavy rain or snow usually lasts in a relatively long duration for several hours. Therefore, it mainly causes severe delays and consequently impacts the following trains. That is why in this research, the severity level is considered to be high under the circumstances of bad weather like heavy rain or snow.

The intrusion of a foreign object (FAC_13) and train collision with a foreign object (FAC_14) both refer to the structural clearance of high-speed railway is invaded by other structures, collapsed rocks or small animals, but the key difference is that if they collide with trains. That is the reason for the totally opposite effect (Coef. = -1.2388 and 0.9823 , respectively) of these two factors on severity level. For FAC_13, when the foreign object intruding railway clearance is detected by the monitoring system, if not serious, trains usually go through the blocking section at limited speed. For FAC_14, if a train collides with an object (mostly birds), the train must take braking measures and stop immediately until the mechanical engineer check if the collision is severe, which costs considerable time. The subsequent train also has to limit its running speed when passing through the collision area to ensure safety. It should be emphasized that FAC_14 violates the parallel line assumption, so γ_2 of FAC_14 (0.1552) should be added to beta (0.9823) to get the coefficient of FAC_14 in the second logit function ($0.9823 + 0.1552 = 1.1375$). The positive value of beta and γ_2 illustrates that the severity of accidents associated with FAC_14 is more severe, which can be proved from the marginal effects of I and II for FAC_14 (-0.1702 and -0.0246). The same process on γ_3 is repeated to get the coefficient in the third logit equation ($0.9823 - 2.8618 = -1.8795$). The two results seem to be inconsistent. Because the p -value of γ_3 is $0.687 > 0.05$, which means this coefficient is insignificant. Thus, the effect of γ_3 is not reliable.

As a key element of the train control system, the ATP system is used for real-time control of running speed to ensure operational safety. However, ATP fault (FAC_1) is one of the most frequent accidents in high-speed railway operation. Fortunately, the negative effect of FAC_1 (Coef. = -0.7257) indicates the severity level tends to be relatively low because most ATP faults can be resolved by restarting the ATP system in a short time.

The coefficient value of platform screen door and train door fault (FAC_3, Coef. = -2.3854) is very similar to that of FAC_12, however, which has an opposite direction of effect. As is mentioned before, linkage failure at the high-speed railway station is the main fault of the platform screen door and train door. Although these accidents occurred when trains parked in

stations, statistical results show that nearly 80% of this kind of accident affect no more than one train and can be fixed in 10 min. The marginal effect of FAC_3 on severity I (0.4133) explains its effect on reducing the severity level of accidents.

Objects hanging on catenary (FAC_5), pantograph fault (FAC_6) and traction converter fault (FAC_7) are all faults of the traction power supply system of high-speed railway. However, accidents led by the first two factors tend to be more severe while the traction converter fault has an effect on reducing accident severity. Reducing running speed and drop the pantograph are the correct measures to dodge the foreign object when a train driver observes a foreign object hanging on the catenary system if the train can go through it by dropping the pantograph. However, if an object hangs on its pantograph or the pantograph automatically drops during traveling, the train must stop and a mechanical engineer on board will check the pantograph or clean up the foreign object. That is why the effect of FAC_6 (Coef. = 0.8353) is a little bigger than FAC_5 (Coef. = 0.7453). Trains of China railway high-speed are all power-distributed which means traction power systems are distributed in multiple carriages. The failure of a traction power system has little effect on the whole train. Thus, FAC_7 has a positive effect on reducing the severity level (Coef. = -1.1391).

Conclusions

This study investigates the impact factors affecting the accident severity of high-speed railway. A total of 14 factors including ATP system fault, turnout indication loss, platform screen door or train door fault, catenary blackout (trip), foreign object hanging on the catenary, pantograph fault, traction converter fault, Carbody's vibration, locomotive fault, passenger misconducting or sudden illness, personnel intrusion of railway clearance, driving on heavy rain or snow, railway clearance intrusion by a foreign object, train collision against foreign object are taken into consideration.

A grading method for accident severity by DOT and the IOST is proposed. To be more specific, accident severity is classified into I, II, III, IV with severity level more and more severe. Ordered probability models are firstly selected as a methodology for the ordinal nature of accident severity. For the purpose of relaxing the constraint of the parallel line assumption, a gamma parameterization of a generalized ordered logit model named the partial proportional odds model is established.

The modeling results show that 10 factors significantly affect the severity level of high-speed railway accidents (p -value < 0.05). Among these factors, nine factors pass the Brant test while only one factor, train collision against foreign object violates the parallel line assumption. Hence, the proportional odds model is established with three alpha coefficients which represent regression intercepts of each logit function, one beta coefficient for each explanatory variable and two gamma coefficients for FAC_14 who are relaxed from proportional odds constraint. The marginal effects of each explanatory variable on severity level are also given in detail. Passenger misconducting or sudden illness and platform screen door or train door fault are the two factors which have the biggest effect on decreasing severity level of accidents while driving on heavy rain or snow has the biggest effect on aggravating accidents.

In summary, the estimated results demonstrate that factors consist of an ATP system fault, platform screen door or train door fault, traction converter fault and railway clearance intrusion by foreign objects have an effect on reducing the severity level whose regression coefficients are negative. On the contrary, the regression coefficients of a foreign object hanging on the catenary, pantograph fault, passenger misconducting or sudden illness, personnel intrusion of railway clearance, driving on heavy rain or snow and railway clearance intrusion by the foreign object are positive, which proves that accidents caused by these factors tend to be more severe.

References

- Abdel-Aty, M. (2003), "Analysis of driver injury severity levels at multiple locations using ordered probit models", *Journal of Safety Research*, Vol. 34 No. 5, pp. 597-603, doi: [10.1016/j.jsr.2003.05.009](https://doi.org/10.1016/j.jsr.2003.05.009).
- Baysari, M.T., McIntosh, A.S. and Wilson, J.R. (2008), "Understanding the human factors contribution to railway accidents and incidents in Australia", *Accident Analysis and Prevention*, Vol. 40 No. 5, pp. 1750-1757, doi: [10.1016/j.aap.2008.06.013](https://doi.org/10.1016/j.aap.2008.06.013).
- Brant, R. (1990), "Assessing proportionality in the proportional odds model for ordinal logistic regression", *Biometrics*, Vol. 46 No. 4, pp. 1171-1178, doi: [10.2307/2532457](https://doi.org/10.2307/2532457).
- Chungang, H. (2015), *Operation and Maintenance Requirement Research of Lanzhou-Xinjiang High-Speed Railway for CRH*, China Academy of Railway Sciences.
- Crawford, E.G.C. and Kift, R.L. (2018), "Keeping track of railway safety and the mechanisms for risk", *Safety Science*, Vol. 110, pp. 195-205, doi: [10.1016/j.ssci.2018.07.004](https://doi.org/10.1016/j.ssci.2018.07.004).
- Dindar, S., Kaewunruen, S., An, M. and Osman, M.H. (2016), "Natural hazard risks on railway turnout systems", *Procedia Engineering*, Vol. 161, pp. 1254-1259, doi: [10.1016/j.proeng.2016.08.561](https://doi.org/10.1016/j.proeng.2016.08.561).
- Dindar, S., Kaewunruen, S., An, M. and Sussman, J.M. (2017), "Bayesian network-based probability analysis of train derailments caused by various extreme weather patterns on railway turnouts", *Safety Science*, Vol. 110, pp. 20-30, doi: [10.1016/j.ssci.2017.12.028](https://doi.org/10.1016/j.ssci.2017.12.028).
- Evans, A.W. (2014), *Fatal Train Accidents on Europe's Railways: 1980-2013*, Imperial College London.
- Feng, S., Li, Z., Ci, Y. and Zhang, G. (2015), "Risk factors affecting fatal bus accident severity: their impact on different types of bus drivers", *Accident Analysis and Prevention*, Vol. 86, pp. 29-39.
- Guanhua, C. (2018), "Analysis of external environment problems of high-speed railway with measures and suggestions", *China Safety Science Journal*, Vol. 28, pp. 145-148.
- Guo, M., Liu, S., Chu, F., Ye, L. and Zhang, Q. (2019), "Supervisory and coworker support for safety: buffers between job insecurity and safety performance of high-speed railway drivers in China", *Safety Science*, Vol. 117, pp. 290-298, doi: [10.1016/j.ssci.2019.04.017](https://doi.org/10.1016/j.ssci.2019.04.017).
- Guo, M., Wei, W., Liao, G. and Chu, F. (2016), "The impact of personality on driving safety among Chinese high-speed railway drivers", *Accident Analysis and Prevention*, Vol. 92, pp. 9-14, doi: [10.1016/j.aap.2016.03.014](https://doi.org/10.1016/j.aap.2016.03.014).
- Han, Y., Liu, Z., Lyu, Y., Liu, K., Li, C. and Zhang, W. (2018), "Deep learning-based visual ensemble method for high-speed railway catenary clevis fracture detection", *Neurocomputing*, Vol. 396, doi: [10.1016/j.neucom.2018.10.107](https://doi.org/10.1016/j.neucom.2018.10.107).
- Hani Tabai, B., Bagheri, M., Sadeghi-Firoozabadi, V. and Sze, N.N. (2018), "Evaluating the impact of train drivers' cognitive and demographic characteristics on railway accidents", *Safety Science*, Vol. 110, pp. 162-167, doi: [10.1016/j.ssci.2018.03.027](https://doi.org/10.1016/j.ssci.2018.03.027).
- Hassankiadeh, S.J. (2011), "Failure analysis of railway switches and crossings for the purpose of preventive maintenance", (Master's thesis), KTH, Stockholm.
- Hu, L.Q., He, C.F., Cai, Z.Q., Wen, L. and Ren, T. (2019), "Track circuit fault prediction method based on grey theory and expert system", *Journal of Visual Communication and Image Representation*, Vol. 58, pp. 37-45, doi: [10.1016/j.jvcir.2018.10.024](https://doi.org/10.1016/j.jvcir.2018.10.024).
- Iran Bureau of Rail Mobility (2015), "Statistical data of Iran railway accident (2009-2014). Iran ministry of roads and urban development", Rail accidents caused by the train driver.
- Ižvolta, L. and Šmalo, M. (2015), "Assessment of the track geometry quality from the aspect of safe and reliable operation of the railway track", *Procedia Engineering*, Vol. 111, pp. 344-350, doi: [10.1016/j.proeng.2015.07.099](https://doi.org/10.1016/j.proeng.2015.07.099).
- Ju, T.O. (2006), "Development of severity models for vehicle accident injuries for signalized intersections in rural areas", *KSCE Journal of Civil Engineering*, Vol. 10 No. 3, pp. 219-225.
- Jun, Q., Xuelei, M., Huiping, S. and Lin, T. (2019), "A study on classification and gradation of railway crash", *Railway Transport and Economy*, Vol. 41 No. 4, pp. 118-122.

- Lee, C. and Abdel-Aty, M. (2005), "Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida", *Accident Analysis and Prevention*, Vol. 37 No. 4, pp. 775-786, doi: [10.1016/j.aap.2005.03.019](https://doi.org/10.1016/j.aap.2005.03.019).
- Liu, Z., Stichel, S. and Rønnquist, A. (2018), "Application of tuned-mass system on railway catenary to improve dynamic performance", *Engineering Structures*, Vol. 165, pp. 349-358, doi: [10.1016/j.engstruct.2018.03.060](https://doi.org/10.1016/j.engstruct.2018.03.060).
- McCullagh, P. (1980), "Regression models for ordinal data[J]", *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 42 No. 2, pp. 109-142.
- McCullagh, P. and Nelder, J.A. (1989), *Generalized Linear Models*, 2nd ed., London, UK, Chapman and Hall.
- Peterson, B. and Harrell, F.E. Jr (1990), "Partial proportional odds models for ordinal response variables", *Applied Statistics*, Vol. 39 No. 2, pp. 205-217, doi: [10.2307/2347760](https://doi.org/10.2307/2347760).
- Rezapour, M., Moomen, M. and Ksaibati, K. (2019), "Ordered logistic models of influencing factors on crash injury severity of single and multiple-vehicle downgrade crashes: a case study in Wyoming", *Journal of Safety Research*, Vol. 68, pp. 107-118, doi: [10.1016/j.jsr.2018.12.006](https://doi.org/10.1016/j.jsr.2018.12.006).
- Shanchao, S., Weidong, W., Jinzhao, L. and Hailang, L. (2012), "Study of carbody's severe vibration based on stability analysis of vehicle system", *China Railway Science*, Vol. 33 No. 2, pp. 82-88.
- Walker, S.H. and Duncan, D.B. (1967), "Estimation of the probability of an event as a function of several independent variables", *Biometrika*, Vol. 54 Nos 1/2, pp. 167-179, doi: [10.2307/2333860](https://doi.org/10.2307/2333860)
- Wang, J. (2014), *Correlation Analysis and Its Practice of Safety Factors for High-Speed Rail Operation Process - A Case Study in Taiyuan Railway*, Dalian Jiaotong University.
- Wang, X. and Abdel-Aty, M. (2008), "Modeling left-turn crash occurrence at signalized intersections by conflicting patterns", *Accident Analysis and Prevention*, Vol. 40 No. 1, pp. 76-88.
- Williams, R. (2006), "Generalized ordered logit/partial proportional odds models for ordinal dependent variables", *The Stata Journal*, Vol. 6 No. 1, pp. 58-82, doi: [10.1177/1536867X0600600104](https://doi.org/10.1177/1536867X0600600104).
- Williams, R. (2016), "Understanding and interpreting generalized ordered logit models", *The Journal of Mathematical Sociology*, Vol. 40 No. 1, pp. 7-20.
- Wu, S.C., Xu, Z.W., Kang, G.Z. and He, W.F. (2018), "Probabilistic fatigue assessment for high-speed railway axles due to foreign object damage", *International Journal of Fatigue*, Vol. 117, pp. 90-100, doi: [10.1016/j.ijfatigue.2018.08.011](https://doi.org/10.1016/j.ijfatigue.2018.08.011).
- Wybo, J.L. (2018), "Track circuit reliability assessment for preventing railway accidents", *Safety Science*, Vol. 110, pp. 268-275, doi: [10.1016/j.ssci.2018.03.022](https://doi.org/10.1016/j.ssci.2018.03.022).
- Xiaomu, Y. (2010), *Study on Diagnostic Methods of Multicollinearity in Logistic Regression*, Dalian Medical University.
- Yasmin, S. and Eluru, N. (2013), "Evaluating alternate discrete outcome frameworks for modeling crash injury severity", *Accident Analysis and Prevention*, Vol. 59, pp. 506-521, doi: [10.1016/j.aap.2013.06.040](https://doi.org/10.1016/j.aap.2013.06.040).
- Zhang, F., Huang, S., Guo, Q., Liu, H. and Dong, D. (2018), "Research on turnout fault diagnosis and reliability assessment technology based on fault tree analysis", *Urban Mass Transit*, Vol. 21 No. 10, pp. 52-56.

Corresponding author

Jian John Lu can be contacted at: jianjohnlu@tongji.edu.cn

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com