



Inferring heterogeneous treatment effects of crashes on highway traffic: A doubly robust causal machine learning approach

Shuang Li^a, Ziyuan Pu^{a,b,*}, Zhiyong Cui^{c,*}, Seunghyeon Lee^d, Xiucheng Guo^a, Dong Ngoduy^e

^a School of Transportation, Southeast University, China

^b Key Laboratory of Transport Industry of Comprehensive Transportation Theory (Nanjing Modern Multimodal Transportation Laboratory), Ministry of Transport, PR China

^c School of Transportation Science and Engineering, Beihang University, PR China

^d Department of Transportation Engineering, University of Seoul, South Korea

^e Department of Civil Engineering, Monash University, Australia



ARTICLE INFO

Keywords:

Highway crashes
Heterogeneous treatment effect
Causal machine learning
Neyman-rubin causal model
Doubly robust learning

ABSTRACT

Accurate estimating causal effects of crashes on highway traffic is crucial for mitigating the negative impacts of crashes. Previous studies have built up a series of methods via traditional causal inference theory and machine learning methods to estimate the impacts of crashes. Since the structures and variable dimensions of traditional causal inference models are pre-defined, they can not accommodate the characteristics of individual crashes. They only can estimate the average causal effects for the crashes in certain categories, e.g., crash types, crash severity, and occurring locations. For machine learning-based algorithms, they cannot be used for causal reasoning due to their reliance on correlation rather than causation. However, considering the impacts of crashes on traffic status vary across influential factors, such as time periods and locations, heterogeneous causal effects are essential for a better understanding of the effects on traffic status and crash intervention strategy development. To address the aforementioned issues, this study proposes a novel doubly robust causal machine learning framework to infer heterogeneous treatment effects of crashes on highway traffic status. Doubly Robust Learning (DRL), integrating machine learning techniques to perform predictive tasks, is applied into the framework due to its stronger robustness. Considering treatment predictors and colliders may bring bias in estimation results, Conditional Shapley Value Index (CSV) is proposed for selecting confounders from numerous factors. A 3-year crash dataset collected by 3594 real highway crashes in Washington is utilized for demonstrating the designed experiments, including constraining confidence intervals, estimated errors evaluation, and sensitivity analysis of variable selection for various thresholds of CSV. According to the results, the distinctive propagation and dissipation processes of congestion caused by various types of crashes can be achieved. The results also validate the effectiveness of variable selection, and the superiority in estimation accuracy compared to the selected baseline models. Future study includes considering spatial-temporal causal relationships and predicting counterfactual real-time traffic conditions.

* Corresponding authors.

E-mail addresses: ziyuapu@seu.edu.cn (Z. Pu), zhiyongc@buaa.edu.cn (Z. Cui).

1. Introduction

Highway traffic crashes have a pronounced impact on traffic safety and mobility, leading to substantial economic losses and energy consumption. According to the estimates from the National Highway Traffic Safety Administration of the United States, the nationwide cost of car crashes in 2019 was \$340 billion (Blincore et al., 2023). To explore the quantitative effects of crashes on traffic status, numerous researchers conducted scientific analysis (Chung and Recker, 2012, 2013; Adler et al., 2013; Chung, 2017; Benlagha and Charfeddine, 2020), which also builds up a solid foundation for crash countermeasure and intervention strategy development. Previous studies usually employ traditional statistical models as the analytical tools. Due to the limitation of modelling large numbers of parameters and the restrictions caused by distributional assumptions, they only can analyze the average effects of the crashes in certain categories, e.g., crash types, crash severity, and occurring locations. However, the previous studies indicate that the impacts of crashes on traffic status vary across influential factors, such as peak hours, traffic conditions, road alignments, and weather (Snelder et al., 2013; Lin and Li, 2020; Su et al., 2023). Thus, a heterogeneous causal effects estimation method is essential for accommodating the characteristics of individual crashes which is also greatly helpful in having a better understanding of the effects of crashes on traffic status and enhanced crash intervention strategy development.

In previous studies, statistical regression methods are commonly employed to identify the relationship between traffic crashes and the influential factors so that the effects of crashes can be determined (Xie et al., 2012; Yu and Abdel-Aty, 2014; Dabbour et al., 2020; Wen et al., 2021; Ding and Sze, 2022). However, such models can not handle the selection bias of the observational data, which may cause incorrect estimation (Mannering et al., 2020; Pearl, 2014). For instance, during peak hours, rear-end crashes are more likely to occur due to more stop-and-go activities. Traffic volume during this time period is much higher than non-peak hours, resulting in lower traffic speeds. Thus, the rear-end collisions usually may not trigger an additional reduction of traffic speed. In this case, the peak hour is a confounder (i.e., the variables that both influence treatment and outcome) between crashes and traffic speed, which leads to underestimating the Impact of rear-end crashes on traffic speed reduction. Given that a variety of factors can act as confounding variables, including pre-crash traffic conditions, road geometric features, weather, etc., it is imperative to employ causal inference models to mitigate confounding bias and accurately assess the causal effects of traffic crashes (Karwa et al., 2011).

Causal inference models enable us to identify the causal relationships between outcomes and treatments, and understand the modifications of outcomes in response to the changes of causal factors (Morgan and Winship, 2007; Pearl, 2009). Neyman-Rubin Causal Model (RCM, also known as the potential outcome framework) (Rubin, 1974; Splawa-Neyman et al., 1990) is a popular analytical framework of causal inference theory. In the scope of RCM, various models, such as Propensity Score Matching (PSM) (Rosenbaum and Rubin, 1983), Inverse Propensity Weighting (IPW) (Rosenbaum, 1987), and Doubly Robust (DR) (Robins et al., 1995), have been employed in transportation fields, including road safety evaluation, travel behavior, and traffic operations (Zhang et al., 2022a; Li et al., 2019; Yue et al., 2024). These approaches employed in previous studies are under the selection on observables identification strategy which the assumption of no unobservable confounders is one of the important requirements to implement it. In most real cases, collecting all confounders for model development is impossible, however, if the input dataset potentially contains the hidden information of missing variables, it still helps with handling the potential confounding bias bringing by the missing variables (see Section 3 for detailed discussion). Therefore, even some of the variables cannot be collected in this study, e.g., special events, and illumination conditions, by including the lane-based real-time traffic speed, road alignment features, and crash report data, selection-on-observables approach is still applicable for model development.

Usually, the crash indicator and the post-crash traffic speed are defined as “Treatment” and “Outcome” of causal inference models, respectively. The modelling objective is estimating Heterogeneous Treatment Effects (HTE) which refer to the unique impacts of an individual crash on highway traffic speed. In previous studies, the causal inference-based models do not capture the heterogeneous effects of crashes due to the difficulties in modelling excessive variables with complex relationships (Cao et al., 2021; Pasidis, 2019). Levering the support of the computational capability of machine learning (ML) algorithms (Mannering et al., 2020), several recent studies integrate ML with causal inference methods to address the problems of modelling high-dimensional data and relaxing strict structural assumptions, such as Causal Tree (Athey and Imbens, 2016), Causal Forest (Athey and Wager, 2019), Metalearner (Künzel et al., 2017), Double Machine Learning (DML) (Chernozhukov et al., 2016), Generalized Random Forest (GRF) (Athey et al., 2019) and Doubly Robust Learning (DRL) (Foster and Syrgkanis, 2019). These approaches have demonstrated favorable performance for causal inference across various fields, including actionable healthcare, medical intervention, and educational practice (Knaus, 2018; Prosperi et al., 2020; Kristjanpoller et al., 2021). Recently, scholars in the traffic safety field have also attempted to apply ML-based causal models for HTE estimation (Zhang et al., 2021a; Liu et al., 2022; Zhang et al., 2022b). However, their methods heavily rely on the correct specification of statistical models for propensity score or potential outcome regression. If one of them is misspecified, the estimated effects are biased. In addition, since the impact of crashes diminishes with time and distance, it is necessary to estimate the HTE of crashes along with the upstream road segments in the post-crash time period. Besides, there are two more challenges remaining to be solved regarding modelling the causal effects of crashes. Firstly, since containing adverse variables in model input data may increase variations or introduce bias when estimating the causal effects (De Luna et al., 2011; Pearl, 2011), expert judgment to identify likely confounders and exclude adverse variables from the model is essential (Blakely et al., 2021). Secondly, unlike supervised ML algorithms, causal inference methods lack ground truth for validation due to the unobservable nature of counterfactual data. Thus, adequate validation approaches are required to evaluate the accuracy and effectiveness of the estimated results.

Based on the aforementioned considerations, this study proposed a novel causal machine learning framework that facilitates causal inference using high-resolution data to estimate the HTE of crashes on traffic speed. DRL is employed due to its capability of maintaining many favorable statistical properties simultaneously (e.g., small mean squared error, asymptotic normality, construction of confidence intervals) and stronger guarantee of robustness compared to other methods (Foster and Syrgkanis, 2019). To validate the

proposed framework, we conducted designed experiments using real-world datasets and statistical test methods. The major contributions are summarized as follows:

- This study addresses the issue of estimating heterogeneous treatment effects of crashes from a causal perspective grounded on RCM. A causal machine learning framework based on the DRL model is proposed to delve into the intricate ways in which various crashes affect traffic speed under distinct situations.
- A Conditional Shapley Value Index (CSV) based on causal graph theory is proposed as the variable selection module of the framework. The main purpose of CSV is to eliminate adverse null variables and treatment predictors for causal inference, thus enhancing the accuracy of estimated causal effects.
- The data collected from Interstate 5 (I5) in Washington State from 2019 to 2021 are utilized to demonstrate and validate the proposed methods, including 3594 real crash reports, lane-based loop detector traffic conditions, and road alignment features. The conducted experiments include bootstrap for confidence level testing, estimation performance evaluation, and sensitivity analysis of CSV for variable selection. The experimental results indicate the proposed method outperforms all other baseline models in terms of estimation accuracy, effectiveness, and robustness.
- Since it is impossible to obtain counterfactual outcomes from observational datasets, there is no ground truth data to validate the estimated HTE of each crash. Considering the inherent periodicity in traffic flow, a matching algorithm is proposed to search “counterfactual outcomes” from historical non-crash data to help with model validation.

2. Literature reviews

2.1. Traffic crashes impact analysis

Regarding the analysis of the impact of traffic crashes, a considerable amount of research has focused on detecting the impact areas (Pan et al., 2015; Chen et al., 2016; Ou et al., 2020). However, these studies are mainly data-driven and do not provide conclusive evidence on the impact of crashes. Additionally, many scholars have studied the influence of various factors on the severity of crashes (Rolison et al., 2018; Hammond et al., 2019; Zheng et al., 2020; Pu et al., 2020), which play a positive role in understanding the covariates that need to be considered in crashes impact analysis.

When analyzing the impact of crashes on traffic conditions, congestion propagation, duration, and delay after crashes are widely studied in previous researches. For example, Chung et al.(2012; 2017) used survival analysis to analyze the effects of different factors on delays and congestion after crashes. Alder et al. (2013) applied a statistical regression model and estimated the incident duration of non-recurrent congestion. Zheng et al.(2020) identified the determinants of crash-caused congestion by a generalized linear mixed-effects model. With the prevalence of ML, researchers adopted various ML models for predicting the target variables directly. Miller and Gupta(2012) adopted various ML models to predict the cost of delay and incident duration based on calculating economic losses. Xie et al. (2019) built a deep learning model to predict speed after crashes, and a binary classifier was designed to extract the latent impact features for improving performance. Luan et al. (2022) applied a dynamic Bayesian network to predict the congestion of road segments and inferred the propagation. Grigorev et al. (2022) built a bi-level machine learning to predict the duration of incidents with outlier removal, which can deal with the imbalanced data problem.

While there is a wealth of literature on the association between traffic crashes and post-crash outcomes, such as predicting post-crash statements, few studies have focused on the impact of crashes on speed reduction, especially from a causal perspective. Although machine learning models can predict traffic conditions with greater accuracy, their black-box nature does not provide policymakers with a causal explanation of how crashes impact traffic speed.

2.2. Development of causal inference methods

Causal inference is a crucial research area in multiple fields, including statistics, medical science, biology, computer science, and economics (Yao et al., 2020). In RCM, researchers proposed a propensity score to balance the distributions of covariates. The propensity score indicates the probability of a unit being treated under the given covariates. Propensity score matching (PSM) is the most common approach, which involves matching data from control and experimental groups with identical scores and calculating the difference. However, due to the imbalanced problem, PSM may introduce bias into the estimator. To address this, inverse propensity weighting (IPW) is proposed to re-weight each sample (Rosenbaum, 1987). Furthermore, to solve the bias-variance trade-off dilemma, Robins et al.(1995) proposed a Doubly Robust (DR) estimation method, also named Augmented Orthogonal IPW (AIPW).

When employing RCM methods to model numerous variables in practical applications, traditional parametric models face significant constraints concerning the number of parameters and functional forms. Consequently, various ML-based causal inference methods have been proposed to handle high-dimensional datasets and relax structural assumptions like linearity. Athey and Imbens (2016) proposed Causal Trees by applying “two-tree” estimators to estimate control-outcomes and treatment-outcome conditional on features, and the difference between these estimated outcomes calculates the causal effects. Künzel (2017) further proposed X-learner in which various ML can be applied to estimate outcome and imputed treatment effects (i.e., the treatment effects for individuals in the treated/control group calculated by control-outcome/treatment-outcome estimator), and he also called this model along with the former models as Metalearners. Chernozhukov et al. (2016) introduced Double Machine Learning (DML), which utilizes ML to predict both the outcome and treatment variables. The residuals obtained from this prediction are then used to estimate the heterogeneous treatment effects through ML models. Based on DML methods, Athey et al. (2019) proposed GRF models that mainly utilized a random

forest model to balance the covariates, and Foster and Syrgkains proposed DRL(2019) by applying the Doubly Robust form function to estimate the effects with ML models. Leveraging the benefits of causal ML models, researchers have investigated methods to make more precise decisions in fields such as education, medicine, and healthcare (Knaus, 2018; Prosperi et al., 2020; Kristjapoller et al., 2021). Since the application of causal ML models in the traffic safety field has not been fully investigated, it is crucial to investigate this direction to enhance the efficacy of decision-making processes.

2.3. Causal inference in traffic safety analysis

In traffic safety, most studies mainly applied traditional causal inference methods, such as PSM(Li et al., 2013; Song and Noyce, 2019; Zhang et al., 2021b), DR (Li and Graham, 2016; Graham et al., 2019), to evaluate the effect of management measures to the safety level. In the context of studying the causal effect of crashes on traffic speed, Pasidis (2019) conducted an investigation using a modified differences-in-differences model. The aim of their approach was to estimate the average causal effects of traffic crashes on speed. However, their model was limited as it only focused on calculating the median speed of the same day of the week, time, and location for four weeks before and after the accidents. Therefore, this method falls short in satisfying the demand for capturing heterogeneous effects at a fine-grained level. Cao et al. (2021) studied the causal effect of a single crash on traffic speed. They used a dataset with two weeks of traffic speed data, one with the crash and one without it, treating affected road segments as the treatment group and unaffected ones as the control group. They matched control group data to each affected segment using propensity scores and calculated the causal effect as the speed difference between the two groups. However, their study solely concentrated on a single crash, and the applied propensity score matching is not suitable for estimating the HTE of crashes.

Recently, some scholars started paying more attention to causal machine learning methods: Zhang et al., (2021a) applied a GRF model to estimate the road safety level treated by a speed camera. They aggregated the crash data at the site level and combined them with camera site data to estimate the HTE. Liu et al. (2022) estimated the average causal effect of curbside pick-up and drop-offs on speed in different regions based on DML. The model incorporated spatial and temporal features, improving its ability to capture the congestion effect.

However, few studies focused on the HTE of crashes on traffic speed, especially considering the effects changes with time and distance. To address the problem, it is necessary to formalize the estimation issue within a causal inference framework and handle high-resolution datasets. In addition, while many variables were controlled in these studies to mitigate confounding bias, some adverse variables should be excluded from the model to decrease variance and avoid bias.

3. Problem statement

This study explores the causal effect of different types of crashes on highway speed, as congestion levels are primarily linked to crash types (Zheng et al., 2020). The crashes are firstly categorized into rear-end, sideswipe, and crash to objects. Besides, causal inference requires considering individual entities. However, the challenge lies in the continuous nature of traffic conditions. Therefore, the highway is divided into one-mile-long sections and the traffic conditions are aggregated by five-minute intervals to obtain different "individuals" to be estimated.

To identify the causal effects of crashes on speed, we adopt a series strategies to mitigate potential confounding bias. Firstly, we hypothesize that pre-crash traffic conditions may influence the occurrence of crashes, while post-crash traffic conditions are caused by the effect of crashes. Therefore, we incorporate traffic conditions observed 10 min before the crash as confounding variables and the speed after the crash as dependent variable. This avoids the endogeneity problem arising from the outcome (speed) simultaneously affecting the treatment (crashes). Secondly, the traffic conditions include traffic flow, speed, occupancy, congestion index, etc. These variables effectively reflect the specific traffic state at the time, ensuring that the matched control group data are consistent with traffic conditions before the crash. By controlling for these variables, the influence of other potential confounders such as weather, construction works, and special events, can be reduced. In addition, highways exhibit linear spatial features, and traffic speeds on different road sections may respond differently to crashes. To account for this spatial variability, we include the milepost and geometric parameters of each section as control variables to fix individual effects effectively. Finally, considering that traffic flow often exhibits apparent temporal dependence, we further incorporate the day of the week and the time of the day as control variables, which help to reduce the impact of potential factors such as light conditions and holidays. By controlling the above variables, we believe that the influence of other potential confounding variables can be almost entirely avoided.

We introduce our methodology based on RCM. Let $Y_{s,t}$ denotes the traffic speed in s milepost at time interval t , then $Y_{s,t}$ is the potential outcome. Let $T_{s,t}^{\text{type}} \in \{0, 1\}$ be the binary indicator for the treatment variable records whether a crash is observed ($T_{s,t}^{\text{type}} = 1$) or not ($T_{s,t}^{\text{type}} = 0$), where type denotes the type of crash (e.g., rear-end). For further analysis, let $Y_{s,t}^{\text{dur},\text{dis}}$ denotes the outcomes under different scenarios, where dur denotes the duration since the crash time t , and dis denotes s miles upstream from the crash site. Then, based on the RCM framework, the Average Treatment Effect (ATE) can be derived as follows:

$$\begin{aligned} \text{ATE} &= E \left[Y_{s,t}^{\text{dur},\text{dis}}(1) - Y_{s,t}^{\text{dur},\text{dis}}(0) \right] \\ Y_{s,t}^{\text{dur},\text{dis}} &= \begin{cases} Y_{s,t}^{\text{dur},\text{dis}}(1) & T_{s,t}^{\text{type}} = 1 \\ Y_{s,t}^{\text{dur},\text{dis}}(0) & T_{s,t}^{\text{type}} = 0 \end{cases} \end{aligned} \quad (1)$$

However, the impact of highway crashes is heterogeneous at different times and locations. For example, as Fig. 1 shows, all these crashes are rear-end. The crash occurring at morning peak hour caused more serious congestion although at the same place ((1) and (2)), and the crash occurring at 150 milepost had little impact on speed than that at 164 milepost ((1) and (3)). Therefore, we need to introduce covariates $X_{s,t} = (x_{s,t}^1, x_{s,t}^2, \dots, x_{s,t}^k)$ and estimate the Conditional Average Treatment Effect (CATE) to explore HTE of crashes. The CATE is defined as a function of full set of conditioning variables to give the conditional mean of crash effect of any point $x_{s,t}^k$, which can be formulated as follows:

$$\text{CATE} = E[Y_{s,t}^{\text{dur},\text{dis}}(1) - Y_{s,t}^{\text{dur},\text{dis}}(0)|X_{s,t}] \quad (2)$$

where $X_{s,t}$ is composed of k -components, which include features, covariates, or pre-treatment variables that are known not to be affected by crashes (Athey and Imbens, 2016). If $X_{s,t}$ containing all confounders, then we could identify CATE for each subgroup by controlling them. When applying the machine learning, the CATE is conducted for each “individual”, thereby facilitating the realization of HTE estimation.

Different from standard traffic condition prediction models, causal effect models cannot be quickly evaluated against each other using a held-out test set because only one of $Y_{s,t}^{\text{dur},\text{dis}}(1)$ and $Y_{s,t}^{\text{dur},\text{dis}}(0)$ can be observed. Thus, the proper treatment effect itself is never directly observed. This has been termed as the “fundamental problem of causal inference”(Holland, 1985). When estimating the causal effect of treatment, the counterfactual outcomes need to be estimated, which will contribute to bias. Therefore, three key assumptions should be satisfied:

Assumption 1 Conditional Independence Assumption (CIA)

$$(Y_{s,t}^{\text{dur},\text{dis}}(1), Y_{s,t}^{\text{dur},\text{dis}}(0)) \perp T_{s,t}^{\text{type}} | X_{s,t} \quad (3)$$

This assumption states that, given the adjustment set $X_{s,t}$, the treatment variable should be independent of the potential outcomes. This assumption is also known as the “no unobserved confounders” assumption, implying that all variables affecting the crash and speed have been controlled. As illustrated above, we incorporate an adequate number of variables to mitigate bias from any potential unobserved confounders.

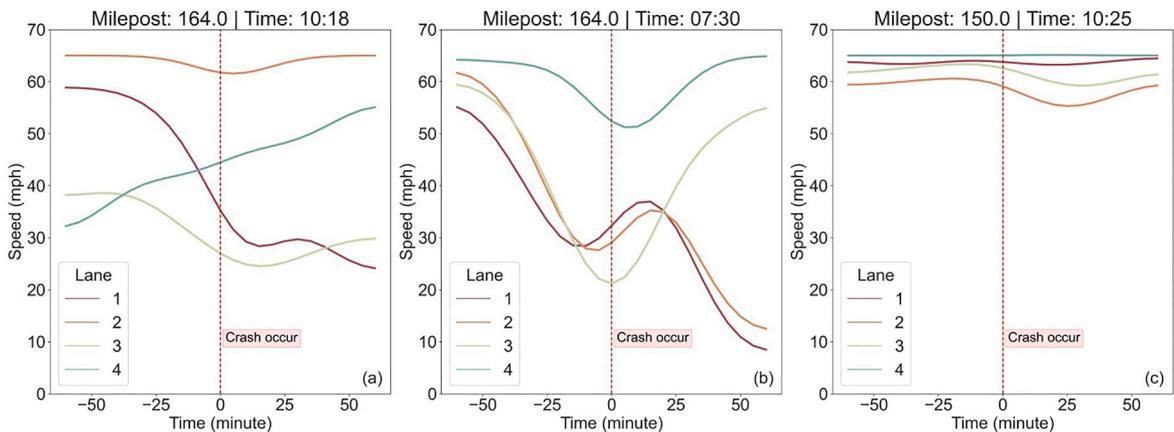
Assumption 2 Stable-unit-treatment-value assumption (SUTVA)

This assumption states that potential outcomes remain unaffected by treatments assigned to other units. However, this can be violated if crashes influence each other or control group traffic states are impacted. To maintain SUTVA, we exclude secondary crashes within 30 min of the initial one. We also remove data one hour before/after the crash and within two miles upstream/downstream to prevent treatment influence on control data.

Assumption 3 Positivity

$$0 < P(T_{s,t}^{\text{type}} = 1 | X_{s,t}) < 1 \quad \text{for all } X_{s,t} \quad (4)$$

This assumption asserts that every combination of covariates in the population has a non-zero probability of receiving the treatment. In other words, there should exist control group data with attributes similar to those of each crash state. However, since the quantity of crash data is significantly less than normal data, random sampling of control data may not satisfy this requirement. To address this, we select normal data with matching period, location, and direction as control, ensuring reasonable satisfaction of the positivity assumption.



(1) High vehicle density section, off-peak hour (2) High vehicle density section, peak hour (3) Low vehicle density section, off-peak hour

Fig. 1. The change of speed caused by rear-end crashes. (1) and (2) occurred at different periods; (1) and (3) occurred at different locations.

4. Methodology

4.1. Traffic crash causal effect analytic framework

This paper presents a causal analytic framework that utilises machine learning techniques to infer the causal effects of crashes on traffic speed, taking into account various influential factors. As illustrated in Fig. 2, firstly, we merge traffic flow, crash, and road alignment data. The crash and non-crash data, along with their corresponding features, are divided into experimental and control groups. The outcome Y , treatment T , and features X can be determined based on the target scenarios. Then, we remove collinearity and eliminate adverse variables based on causal graph theory. Thirdly, we established the causal relationship of the selected variables, treatment, and outcome to determine the statistic estimand for causal effects inference. At last, we chose the most appropriate machine learning model for classification and regression via grid search cross-validation and introduced the data and models into the DRL to estimate the HTE of different crash types at different locations after different time periods. Our framework incorporates causal inference and ML into traffic crash analysis, which could serve as a valuable reference for future research in this area.

4.2. Variables selection based on shapley value

The primary purpose of variable selection is to reduce bias and improve efficiency for causal effect estimation (Witte et al., 2019). Although there are no straightforward methods to ensure the best variable set, we can develop some basic strategies based on causal graph theory to select variables.

4.2.1. Basic variables selection criteria

The causal graph is widely used to reflect the qualitative causal mechanism of factors. The Directed Acyclic Graph (DAG) represents the causal relationship and can be determined according to prior knowledge and data. According to the causal graph theory, we could divide the covariates into four categories $X = \{X_C, X_T, X_O, X_N\}$ (Tang et al., 2020) and define X_C as confounders that have direct causal effects both on speed and crashes. X_O and X_T are used to denote outcome predictors and treatment predictors, respectively. These terms refer to the covariates that exclusively affect speed or crashes causally. X_N denotes as null variables, which represents variables other than the above three types of variables. The corresponding DAG diagram is shown in Fig. 3.

Among these variables, including all confounders is the key to correctly identifying the treatment effect. However, even though the outcome and treatment still render the conditional independent requirement, the inclusion of treatment predictors into the conditioning sets may contribute to the increased variance of the estimated treatment effect (Brookhart et al., 2006; Rotnitzky and Smucler, 2020; Schnitzer et al., 2016). Moreover, studies show that this inclusion may cause bias (De Luna et al., 2011; Pearl, 2011). Besides, null variables may function as colliders which should not be controlled in causal inference, because conditioning on colliders may

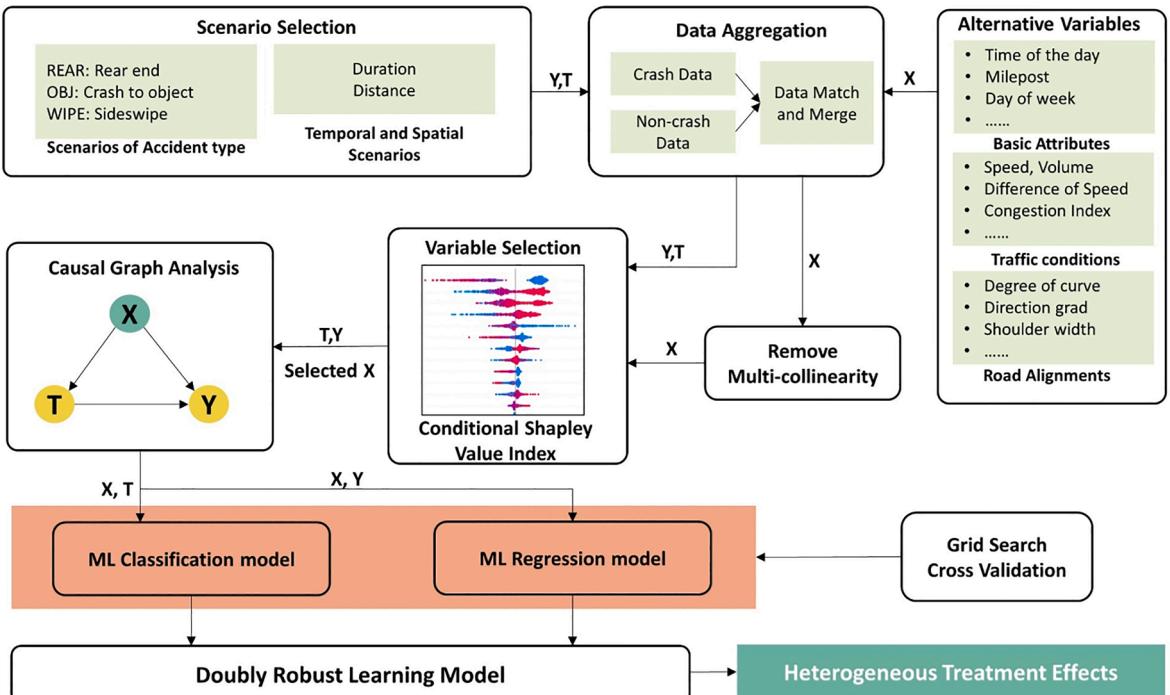


Fig. 2. The doubly robust causal learning framework for highway crash effect inference.

induce association between the covariates and contribute to estimated bias.

4.2.2. Remove multi-collinearity and adverse variables

In causal modeling, it's not uncommon for two or more variables to exhibit a strong correlation, often manifesting as a linear relationship. This deterministic interdependency, known as collinearity, can pose challenges in data analysis (Spirtes and Scheines, 2004). Apart from the demanding task of accurately specifying causal models, bias can be exacerbated with increasing collinearity (Schisterman et al., 2017). The parameters related to traffic conditions often display significant correlations, such as volume and speed. To address this, we first employed the Pearson correlation coefficient to quantify the linear associations between variables, subsequently removing covariates with substantial linear relationships.

Remain covariates may still include adverse variables (i.e. treatment predictor and null variables). According to the causal graph in Fig. 3, if the covariate has no correlation with the outcome conditioned on treatment, it can be removed. To test the conditional correlation, the Shapley value is adopted to test the correlation between remaining covariates and speed conditioned on crash. SHAP is an interpretation method that utilizes the Shapley value based on game theory to combine optimal credit allocation with local explanations (Lundberg and Lee, 2017). Let $Shap(X(t), Y(t))$ denote the absolute Shapley value of variable X for speed Y under crash treatment $T = t$. If $Shap(X(t), Y(t)) = 0$, we consider X is independent of Y when $T = t$.

We propose the Conditional Shapley Value Index (CSVII) to perform conditional independence. Let $\omega = P(T = 1)$ be the probability of receiving a crash, and the CSVII between X and Y given T is defined as follows:

$$CSVII(X, Y|T) = \omega Shap(X(1), Y(1)) + (1 - \omega) Shap(X(0), Y(0)) \quad (5)$$

This index quantifies the importance of variables X for Y when the crash occurs or not. According to the definition of treatment predictors and null variables, if the variable X is not related to Y no matter the value of T , then $Shap(X(1), Y(1)) = Shap(X(0), Y(0)) = 0$. Considering the collected traffic data are not balanced, we add a weight to balance the index. Let $\hat{\omega} = n_t/n = \sum_i^n T_i(1)/n$ be the empirical estimator of ω , where n is the number of all samples and n_t is the number of crash samples. The closer the $CSVII(X, Y|T)$ is to zero, the more likely $X \perp\!\!\!\perp Y|T$, which means the X belongs to X_T or X_N that should be excluded.

Principle 1 If $CSVII(X, Y|T) < \epsilon$, the X should be ignored, where ϵ is a threshold determined according to the experiments

To determine the threshold ϵ , we can apply a series of thresholds to estimate the causal effects and select the best value based on the estimated performance. The validation method to evaluate the performance will be discussed lately (Section 4.5).

4.3. Causal graph analysis of traffic crashes

After the variable selection, we assume there are no adverse variables, and the outcome predictors and confounders can be controlled when estimating the CATE. We determine the controlled variables based on the causal graph to understand the causal relationship between the variables further. To simplify the illustration, the confounders are divided into three categories: 1) basic attributes $X_{s,t}^{Basic}$, such as period, milepost, direction, and city; 2) road alignment $X_{s,t}^{Align}$, referring to the geometric parameters of each road segment, such as gradient and degree of the curve; 3) pre-crash traffic conditions $X_{s,t}^{Condi}$, referring to the traffic conditions (10) minutes before the crash occurrence. The causal graph of these three types of covariates, crash treatment $T_{s,t}^{type}$, and traffic speed $Y_{s,t}^{dur,dis}$ are shown in Fig. 4.

We demonstrate three categories of variables to represent all factors for ease of display, and we claim that all the factors belonging to the corresponding category satisfy the relationship in the diagram. According to the causal graph, the red arrow refers to the causal association of traffic crashes to the traffic speed, which is our estimation target. The paths $T^{type} \leftarrow X \rightarrow Y^{dur,dis}$ are backdoor paths, which should be blocked during estimation. Therefore, the X^{Basic} , X^{Align} and X^{Condi} should be controlled during estimation. Furthermore, in order to estimate the causal effect of crashes on traffic speed at various upstream locations after different periods, we assume as follows:

Assumption 4 For any traffic statement, no additional factor affects the traffic speed dur minutes after the crash and dis miles to the

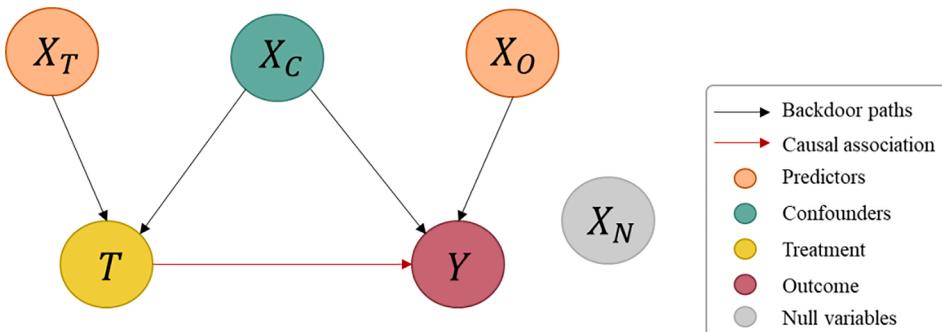


Fig. 3. The causal mechanism of variables demonstrated by the directed acyclic graph.

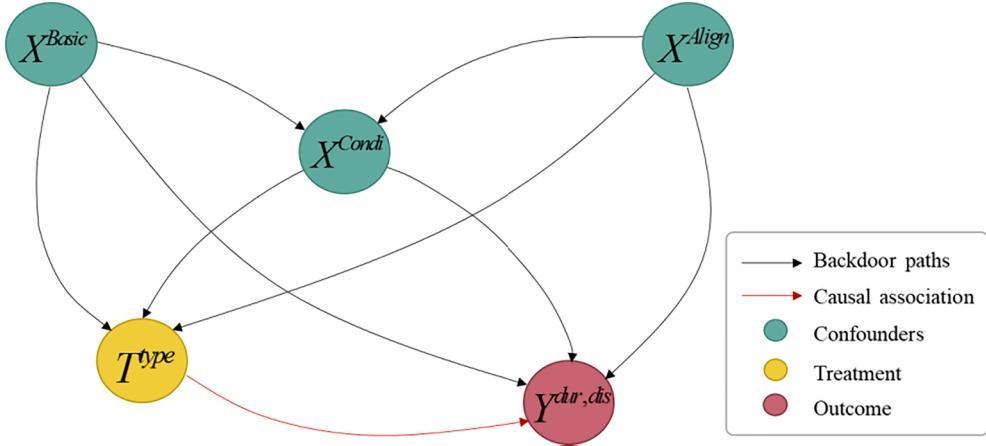


Fig. 4. The causal graph of the covariate, traffic crash, and speed.

crash location

Therefore, based on **Assumption 4**, formulate the statistical estimand of conditional average traffic crash causal effect for different duration and distances in Equation (6).

$$\widehat{CATE}_{dur,dis}^{type} = E_{X_{s,t}} \left[E(Y_{s,t}^{dur,dis} | T_{s,t}^{type} = 1) | X_{s,t} \right] - E(Y_{s,t}^{dur,dis} | T_{s,t}^{type} = 0) | X_{s,t} \right] \quad (6)$$

where $X_{s,t} = (X_{s,t}^{Basic} \cup X_{s,t}^{Align} \cup X_{s,t}^{Condi})$.

4.4. Doubly robust learning causal inference

To improve the robustness of inference, we apply doubly robust learning methods to estimate the causal effects of crashes. As introduced before, the DR estimator is formulated based on the propensity score, defined as the probability of a crash to the traffic state conditional on observed variables as follows:

$$e(x) = P(T = 1 | X = x) \quad (7)$$

The traffic states with the same propensity score are recognized as conditional independence, and their speed from the experiment and control group are the potential outcomes. Thus, the conditional average treatment effect can be calculated as Equation (8), and this process is known as *matching*.

$$\widehat{CATE} = E(Y_{s,t}(1) | e(x_{s,t}) = \gamma) - E(Y_{s,t}(0) | e(x_{s,t}) = \gamma), \quad \gamma \in (0, 1) \quad (8)$$

To solve the problem of imbalanced data, inverse propensity weighting (IPW) is applied to re-weight each sample as:

$$r = \frac{T}{e(x)} - \frac{1-T}{1-e(x)} \quad (9)$$

Therefore, the re-weighted IPW estimator of CATE can be formulated as follows:

$$\widehat{CATE}_{IPW} = E\left(\frac{T_{s,t}Y_{s,t}}{\widehat{e}(x_{s,t})}\right) - E\left(\frac{(1-T_{s,t})Y_{s,t}}{1-\widehat{e}(x_{s,t})}\right) \quad (10)$$

Furthermore, the DR estimator combines the IPW and a regression model to solve the bias-variance problem. As long as one is correctly specified, the estimator can correctly estimate CATE, which is *doubly robust*. According to our problem, the estimator can be formalized as follows:

$$\begin{aligned} \widehat{CATE}_{DR,dur,dis}^{type} &= E[Y_{DR,s,t}^{dur,dis}(1) - Y_{DR,s,t}^{dur,dis}(0) | X_{s,t}] \\ Y_{DR,s,t}^{dur,dis}(1) &= \frac{T_{s,t}^{type}Y_{s,t}^{dur,dis}}{\widehat{e}^{type}(X_{s,t})} - \frac{T_{s,t}^{type} - \widehat{e}^{type}(X_{s,t})}{\widehat{e}^{type}(X_{s,t})} \widehat{m}^{dur,dis}(T_{s,t}^{type} = 1, X_{s,t}) \\ Y_{DR,s,t}^{dur,dis}(0) &= \frac{(1-T_{s,t}^{type})Y_{s,t}^{dur,dis}}{1-\widehat{e}^{type}(X_{s,t})} - \frac{T_{s,t}^{type} - \widehat{e}^{type}(X_{s,t})}{1-\widehat{e}^{type}(X_{s,t})} \widehat{m}^{dur,dis}(T_{s,t}^{type} = 0, X_{s,t}) \end{aligned} \quad (11)$$

where $\hat{e}^{type}(X_{s,t})$ is the propensity score estimated model for indicator variable of $T_{s,t}^{type}$, $\hat{m}^{dur,dis}(T_{s,t}^{type}, X_{s,t})$ is the regression model for the traffic speed $Y_{s,t}^{dur,dis}$ based on the covariates $T_{s,t}^{type}$ and $X_{s,t}$.

The DRL model follows the two-stage process, where ML models estimate the propensity score and outcome in the first stage, and the CATE is estimated in the second stage. In the first stage, based on the causal graph analysis in Section 4.3, we formulate the functions to predict $\hat{m}^{dur,dis}(T_{s,t}^{type}, X_{s,t})$ and $\hat{e}^{type}(X_{s,t})$ as follows:

$$\hat{m}^{dur,dis}\left(T_{s,t}^{type}, X_{s,t}\right) = \hat{Y}_{s,t}^{dur,dis} = \hat{g}_{type}^{dur,dis}\left(T_{s,t}^{type}, X_{s,t}^{Basic}, X_{s,t}^{Condi}, X_{s,t}^{Align}\right) + \varepsilon_{s,t}^{dur,dis} \quad (12)$$

$$\hat{e}^{type}(X_{s,t}) = Pr\left(T_{s,t}^{type}|X_{s,t}\right) = \hat{p}^{type}\left(X_{s,t}^{Basic}, X_{s,t}^{Condi}, X_{s,t}^{Align}\right) + \delta^{type} \quad (13)$$

where $\hat{g}_{type}^{dur,dis}$ predicts the value of $Y_{s,t}^{dur,dis}$ under the crash treat, \hat{p}^{type} predicts the probability of $T_{s,t}^{type} \in \{0, 1\}$, $\varepsilon_{s,t}^{dur,dis}$ and δ^{type} are unobservable noise with $E(\varepsilon_{type}^{dur,dis}|X_i) = 0$ and $E(\delta^{type}|X_i) = 0$.

Considering T_i^{treat} is a binary variable and $Y_i^{dur,dis}$ is a continuous variable, we apply classification ML models and regression ML models to predict them, respectively.

In the second stage, it should be noted that the true CATE cannot be obtained since the counterfactual outcomes are predicted based on observational data. Thus, it is not necessary to pay much attention to the accuracy of the final model since the input variables are estimated. This study applies Ordinary Least Square (OLS) to estimate the $\widehat{CATE}_{DR,dur,dis}^{type}$ according to Equation (11) by regressing $\hat{Y}_{DR,s,t}^{dur,dis}(1) - \hat{Y}_{DR,s,t}^{dur,dis}(0)$ on $X_{s,t}$. After fitting the causal machine learning model, the heterogeneous treatment effect can also be estimated when inputting the data of each “individual”.

4.5. Validation based on matching algorithm

Because the counterfactual outcomes cannot be observed, there is no ground truth data to evaluate the estimated errors. Therefore, we propose a matching algorithm to search the reference data that functioned as “counterfactual outcomes” for verification. Due to the inherent periodicity in traffic flow, it is possible to identify numerous non-crash traffic conditions that exhibit spatiotemporal consistency with crash states. These conditions can serve as a reference group for comparative analysis. Pasidis (2019) applied matched data as counterfactuals in the control group, demonstrating the validity of using matched data for estimating the causal effects of crashes. Differing from his study, matching data is not applied for the control group during the modeling task but for validation. The average non-crash conditions for all locations and time periods are calculated and input into the model for training, therefore, the matched data for validation does not overlap with training data to a certain extent.

During the matching process, several non-crash conditions are selected for each crash. To be exact, for a given crash data c_i from experimental data group S_{expri} , we search the corresponding k non-crash traffic states $N_i = \{n_1, n_2, \dots, n_k\}$ from the control group S_{ctrl} based on the *time*, *week*, *milepost*, and *direction* of c_i . Each data c_i or n_i is a vector data containing all variables (such as pre-treatment traffic conditions) and outcomes. **Algorithm 1** shows the procedure of “counterfactual outcomes” matching.

Algorithm 1 Data Matching: match “counterfactual outcomes” for crash conditions

Input: S_{expri} : dataset of crash data; S_{ctrl} : dataset of non-crash data; *TimePeriod*: a given time range for search; K : number of expected matched data
Output: \mathbb{N} : the matched non-crash dataset, $\mathbb{N} = \{N_1, N_2, \dots, N_i\}$

```

 $\mathbb{N} = \{\}$ 
For  $c_i$  in  $S_{expri}$  do
     $t = c_i.time$ ,  $w = c_i.week$ ,  $mp = c_i.milepost$ ,  $dir = c_i.direction$ 
     $k = 0$ ,  $N_i = []$ 
    For  $n_j$  in  $S_{ctrl}$  do
        If  $n_j.time$  in Range(Timeperiod) then
            If  $t == n_j.time$  and  $w == n_j.week$  and  $mp == n_j.milepost$  and  $dir == n_j.direction$  then
                 $N_i.add(n_j)$ ,  $k += 1$ 
        End if
    End for
    If  $k == K$  then // Avoid spending too much time on matching data
        End for
    End for
End for

```

For each crash c_i , we define the original speed as osp_i and estimated individual treatment effect is ice_i . After we matched the non-crash dataset N_i , the “ground truth” speeds $SP_i = [sp_1, sp_2, \dots, sp_k]$ can be obtained, and the matched causal effects are:

$$mce_i = \sum_k (csp_i - sp_k) / k \quad (14)$$

where csp_i is the post-crash speed of c_i calculated by $osp_i + ice_i$. Based on these matched causal effects, we can calculate the performance of estimation, which will be discussed in Section 5.5.2.

It should be noted that different from the before-after analysis that compares historical observational data to calculate the effects, our method is to build a model that can be applied to predicting HTE. Therefore, this method exhibits superior efficacy in simultaneously analyzing diverse crashes and accurately estimating HTE.

5. Experiments and results

5.1. Data source and preparation

The main road of Interstate 5 (I5) in Washington is selected as our research object. The data used in this study range from Nov. 2019 to Jun. 2021, and the milepost range from 139 to 178, consisting of three data sets:

Traffic data set: The traffic data record the traffic conditions by three parameters detected by the loops: Volume (vol), Occupancy (occ), and Speed (spd), and also provide information including time, location (milepost), road type, direction and the number of lanes.

The raw data is recorded by detectors installed unevenly on the road. We pre-process this data to obtain traffic conditions at each milepost using two rules: 1) Calculate the mean values of the speed and occupancy and the sum of the volume recorded by the detectors within one mile. 2) Interpolate the nearest two data points for unrecorded mileposts. Furthermore, all data is aggregated in 5-minute intervals.

Incident data set: The incident data are collected by the Highway Safety Information System (HSIS), focusing on the 4815 crashes within the scope of research. This data contains basic information about the crashes (e.g., time, location, weather), collision type (e.g., sideswipe, rear-end), and subtypes (e.g., vehicle type, driver status). We match the crash data with the traffic data based on location and time.

Road Alignment data set: The alignment data comes from roadway inventory files collected by HSIS and contains three subfiles (Roadlog, Curve, and Grade) “section” files. Each homogeneous roadway section is defined by a beginning and ending milepost. The basic Roadlog file contains information like lane, shoulder width, and type, while the supplemental files provide horizontal and vertical alignment information. To match traffic data, we rearranged the data as the “point data” so that the alignment characteristics are described at a given milepost.

We chose one and two miles upstream and downstream of the given milepost as the related positions and used their six parameters 10 min before a given time. The variables set is outlined in [Table 1](#).

5.2. Traffic crash scenarios definition and characteristics

After conducting a preliminary analysis of the crash types in the dataset, this study classifies 3594 traffic crashes into three main categories: Rear-end (REAR), Crash to object (OBJ), and Sideswipe (WIPE). The sample sizes for each type are 1959, 713, and 922, respectively, after removing crash types with less than ten occurrences.

To explore the causal effects of crashes on different locations after different durations, we estimated the causal effect 5 min to 30 min after the crash happened from the place of the crash to the 5 miles upstream. Therefore, we can define that

Table 1

The description of candidate variables.

Category	Variables	Values	Description
Basic attributes	<i>time</i>	0: night, 1: off-peak, 1: peak hour	The period during the day,
	<i>milepost</i>	[139,178]	The milepost of the road
	<i>direction</i>	0: Northward, 1: Southward	The direction of the road
	<i>week</i>	{0, 1, ..., 6}	The day of the week
	<i>aadt</i>	R*	Average annual daily traffic
	<i>city</i>	R*	City number
	<i>pop_grp</i>	R*	City population
Traffic conditions*	<i>vol</i>	R*	The volume of vehicles
	<i>occ</i>	R*	The occupancy time
	<i>spd</i>	R*	The average speed during the time interval within 1 milepost
	<i>std_spd</i>	R*	The standard deviation of speed
	<i>spd_diff</i>	R*	The max speed difference between lanes
	<i>ci</i>	[0, 1]	Congestion Index
Road Alignment Features	<i>lanewid</i>	R*	The width of the lane
	<i>medwid</i>	R*	The width of median
	<i>shlwid</i>	R*	The width of the shoulder
	<i>no_lane</i>	{2, 3, 4, 5}	The number of lanes
	<i>curv_max</i>	R*	The max degree of the curve
	<i>deg_curv</i>	R*	The degree of the curve
	<i>pct_grad</i>	R*	The percentage of gradient
	<i>dir_grad</i>	0: up, 1: down	The direction of the gradient

*Note: All variables of traffic conditions includes the variables upstream and downstream to the occurrence place and are denoted in models with suffix down_1, down_2, up_1, and up_2.

$dur \in [5, 10, 15, 20, 25, 30]$, $dis \in [0, -1, -2, -3, -5]$, $type \in [REAR, OBJ, WIPE]$. Here, the values in dis denote the relative distance to the place where the crash occurred.

As shown in Fig. 5, The spatial distribution of crashes depends on whether the road passes through an urban area. There is a particular section in Seattle, from 162 to 167 milepost, where crashes are more likely to occur. REAR and WIPE occur more often during the day, especially during peak hours, while OBJ crashes happen more frequently at night due to poor visibility and driver fatigue.

According to the data in Fig. 6, each type of crash has a unique effect on traffic conditions compared to non-crash data. REAR crashes result in the highest means (red dots) of occupancy and volume, leading to the lowest means of traffic speed. WIPE has a similar effect but less impact on traffic conditions. Traffic conditions associated with OBJ crashes exhibit greater similarity to non-crash conditions. Particularly noteworthy are the boxplots illustrating the lane-based speed difference and congestion index for OBJ and non-crash scenarios, both of which are close to zero. This pattern is significantly different from the other two crash types. Therefore, these two parameters, namely spd_diff and ci , are potential useful for models to distinguish between different types of crash.

5.3. Variable selection results

We conducted a test for multicollinearity using the Pearson correlation coefficient. Fig. 7 displays the results, with each square representing a coefficient between two variables. The colors correspond to the coefficient values, as indicated by the color bar legend. We observed a high correlation between occupancy and speed, as well as between the standard deviation of speed and the congestion index. Consequently, we retained only one of these correlated variables in the model. We retained ci since it has a lower coefficient with vol than the other three variables and shows higher differentiation for different types of crashes (Fig. 6). We removed occ , spd , and std_spd from the model. Additionally, we deleted $medwid$ and pct_grad since they have high correlations to $milepost$ and dir_gard , respectively.

Considering the large computational burden of our experiments, we use Light Gradient Boosted Machines (LGBM) to calculate the Shapley values due to its implementation of two techniques: Gradient-Based One-Side Sampling and Exclusive Feature Bundling. These techniques enable faster execution and higher accuracy (Ke et al., 2017). Fig. 8 summarizes the effects of all features. The horizontal axis represents the SHAP values, and the left vertical axis shows the 20 most critical variables, sorted by their mean absolute Shapley value, and the right vertical axis shows the value of the feature. The redder the color, the higher the value of the feature, and vice versa. For instance, for the control group of the REAR crash, the ci_up1 has the highest mean absolute SHAP value, indicating its significant contribution to predicting speed, and a higher ci_up1 leads to a lower speed. Fig. 8 reveals that the variables of traffic conditions contribute significantly to the outcome, and the difference in SHAP values between the control and experiment groups is apparent. To eliminate bias and variance, we calculated the CSVI for each variable (Fig. 9) and removed variables such as no_lane , $lanewid$, $city$, dir_grad , pop_grp , $aadt$, and $curv_max$, based on Principle 1, with a threshold of $\epsilon = 0.15$.

5.4. Classification and regression machine learning estimation results

We conduct specific screening for the ML models in the first stage. We take several alternative ML models for classification and regression prediction, and the alternative models are listed as follows: 1) extreme gradient boosting (XGBoost) (Chen and Guestrin, 2016), 2) Random Forest (RF) (Breiman, 2001), 3) Light gradient boosted machines (LGBM) (Ke et al., 2017), 4) Linear Regression (LR) and 5) Support Vector Machine (SVM).

To select the best models, we must choose a series of hyper-parameters to train and test the data, and the searched hyperparameters are outlined in Table 2. Considering large quantities of parameters in ML models, especially XGBoost and LGBM, grid searching for all parameters will consume much time. Thus random grid search was applied, which could randomly choose hyper-parameters more efficiently within a small fraction of the computation time (Bergstra and Bengio, 2012). After grid searching, the best parameters were input into the model for cross-validation. In this study, k -fold cross-validation was employed, enabling all observations to be used both for training and testing by splitting the sample data into k equal-sized subsamples. We applied 10-fold cross-validation for model

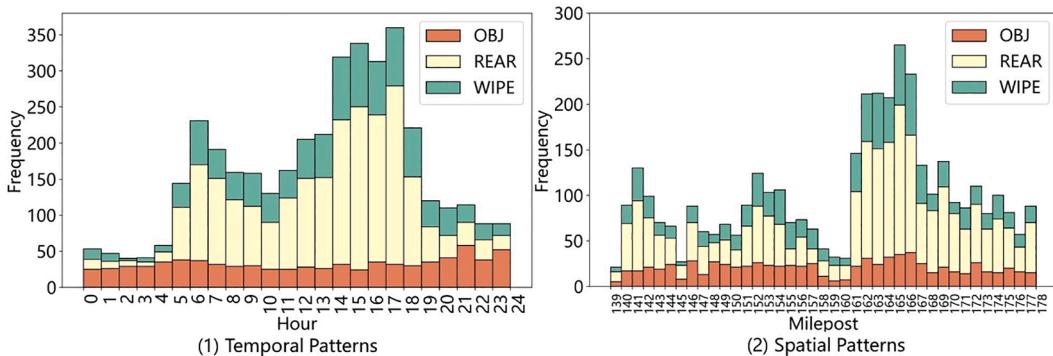


Fig. 5. The spatial and temporal frequency distribution of different crash types.

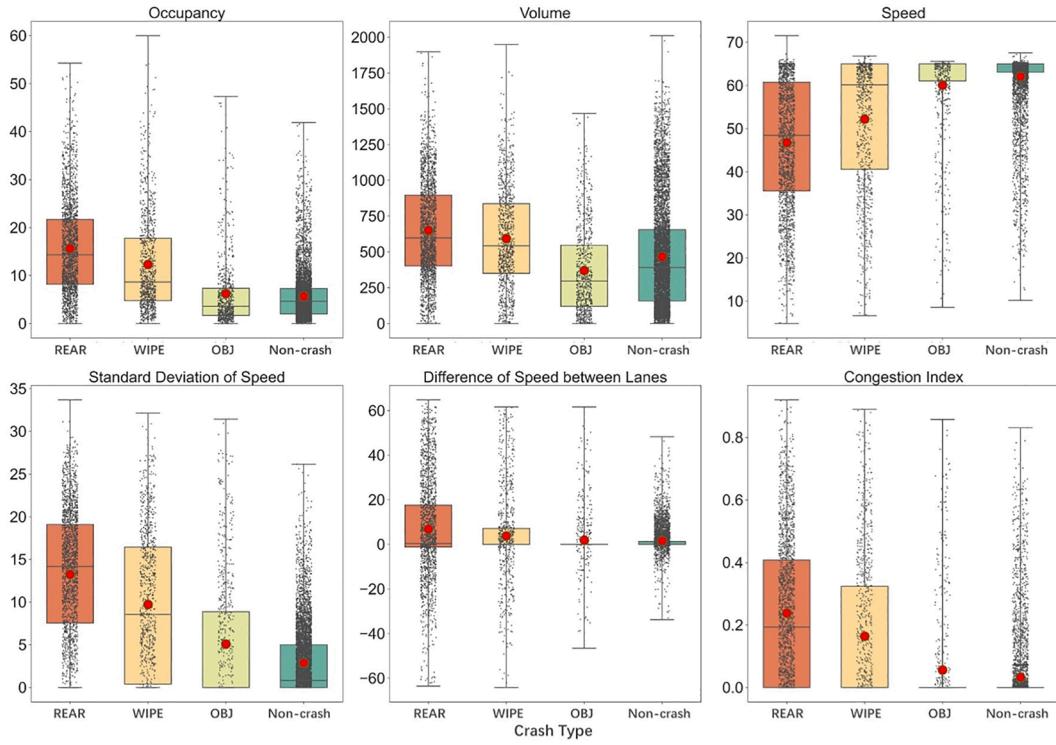


Fig. 6. The statistic distributions of basic traffic parameters based on the types (red dots are mean values).

training, dividing the data into ten equal subsets. Among the ten subsamples, one functions as the validation data, and the remaining nine subsamples were used as the training data. To select the best models, several indicators for classification and regression models were applied to evaluate the predicting accuracy. For classification, Precision, Accuracy, Recall, and F1-score are used to quantify the error. These indicators are determined as follows:

$$precision = \frac{TP}{TP + FP} \quad (15)$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

$$recall = \frac{TP}{TP + FN} \quad (17)$$

$$F1 = \frac{2TP}{2TP + FN + FP} \quad (18)$$

where TP , TN , FP , FN denote the true positives, true negatives, false positives, and false negatives, respectively.

As for regression, we used Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2). The formulas are shown as follows:

$$MAE = \frac{1}{n} \sum_i^n |\hat{y}_i - y_i| \quad (19)$$

$$MSE = \frac{1}{n} \sum_i^n (\hat{y}_i - y_i)^2 \quad (20)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_i^n (\hat{y}_i - y_i)^2} \quad (21)$$

$$MAPE = \frac{1}{n} \sum_i^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (22)$$

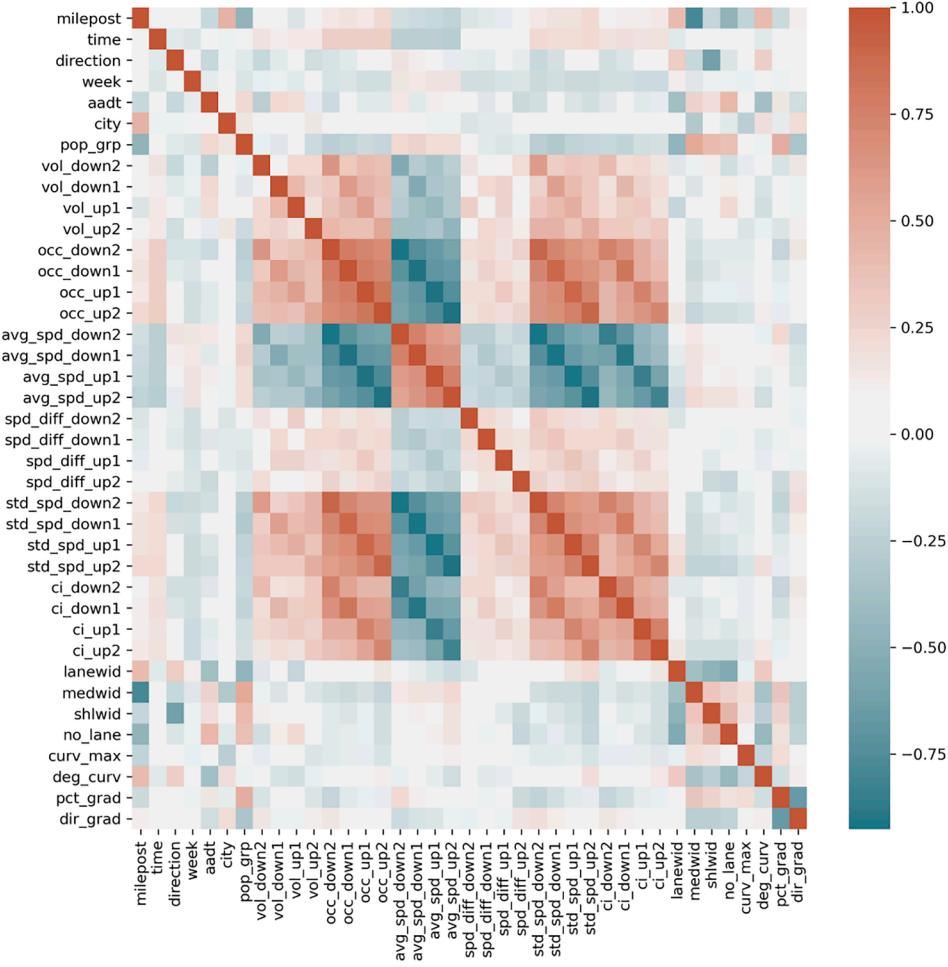


Fig. 7. The Pearson correlation coefficient of all candidate variables.

$$R^2 = 1 - \frac{\sum_i^n (\hat{y}_i - y_i)^2}{\sum_i^n (y_i - \bar{y})^2} \quad (23)$$

where y_i denotes the actual value of the target variable, \hat{y}_i denotes the corresponding predicted value, and \bar{y} denotes the average value of the target variable.

Table 3 and **Table 4** present the performance results. The Random Forest model is best suited for classification and regression, even though sometimes LGBM performs better in specific error parameters. Therefore, we chose the Random Forest model as the base ML model to predict propensity scores and outcomes.

5.5. Conditional average treatment effect of traffic crashes

In this section, we applied EconML(Battocchi et al., 2019), a causal inference Python package that utilizes ML techniques, to estimate causal effects. We presented the results of CATE for each type of crash considering different spatial and temporal conditions, and we also discussed the validation of the results.

5.5.1. Analysis and discussion of conditional average treatment effect results

After we obtained $\widehat{e}^{type}(x_{s,t})$ and $\widehat{m}^{dur,dis}(t_{s,t}^{type}, x_{s,t})$, the final $\widehat{CATE}_{dur,dis}^{type}$ were estimated by OLS, and the results are outlined in **Table 5**, and we highlight the highest value for each distance, except those more than -1 mph. To conduct statistical hypothesis tests for validation, we apply the bootstrap method to calculate the p-value and confidence interval. The results show that the closer to the crash's time and location, the higher the confidence level, and when the effects are nearly zero, the p-values are high, meaning that there is no treatment effect of crashes.

In terms of the estimated CATE, it can be found that the degree of crash effect is REAR > WIPE > OBJ, which is consistent with

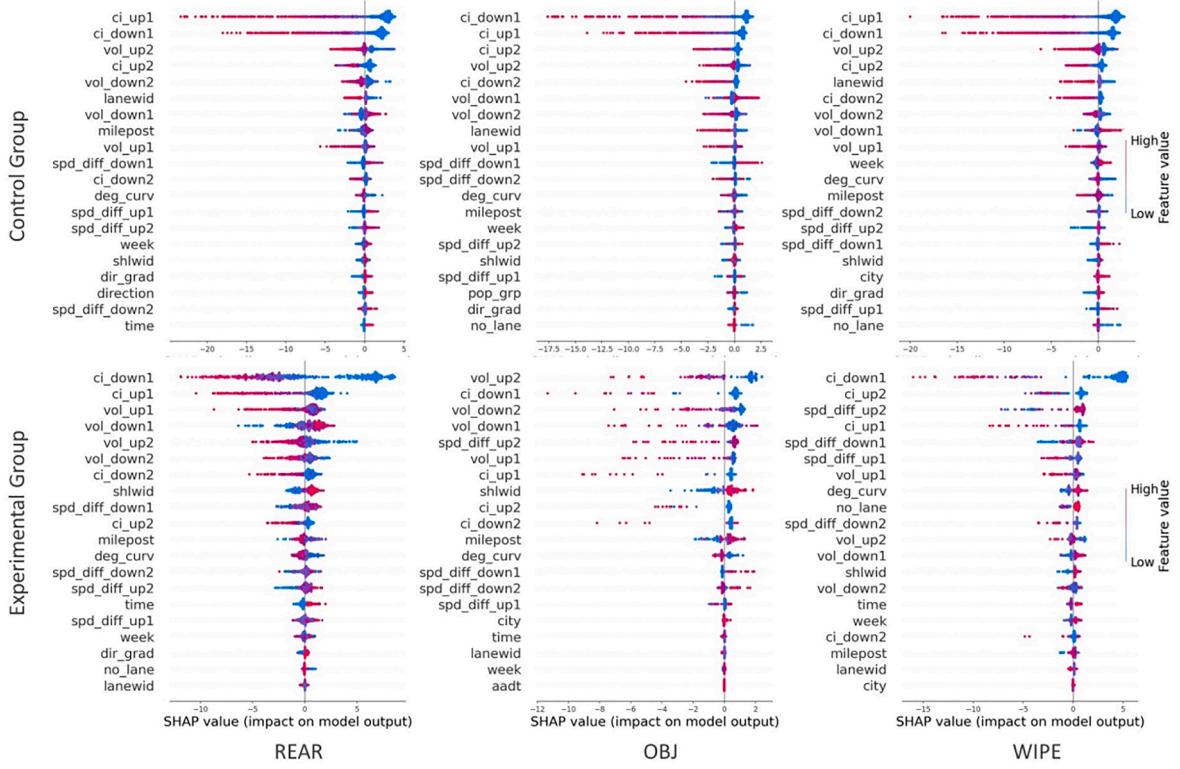


Fig. 8. Shapley values of the variables for the speed in experimental and control group.

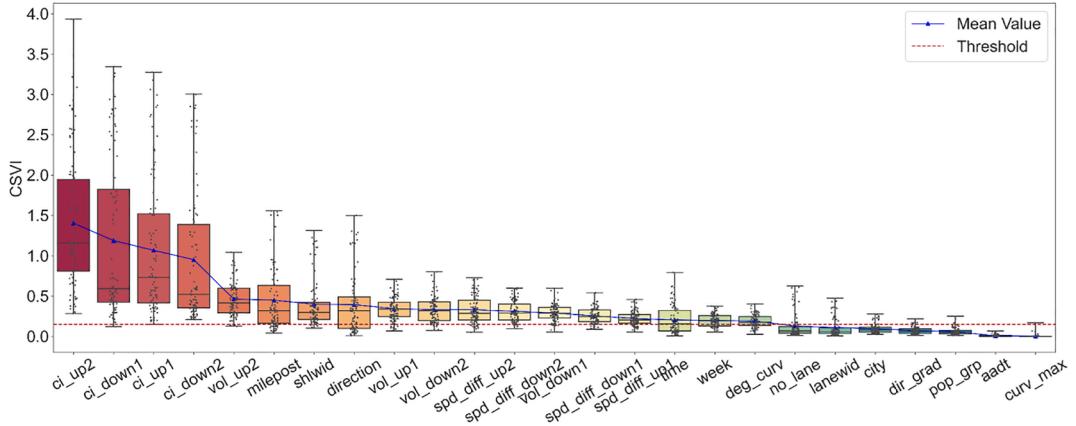


Fig. 9. Distribution of CSVI results for each variable.

previous findings(Garib et al., 1997; Chung, 2017). Regarding the spatial and temporal impact of crashes on highway traffic speed, our analysis reveals that crashes can cause an average speed reduction of around 10 to 15 miles per hour. Additionally, congestion propagates about 2–3 miles upstream and lasts over half an hour. We observed an apparent propagation and dissipation process of congestion, with different crash types exhibiting varying characteristics.

Specifically, REAR crashes cause the most severe and lasting congestion compared to the other two crash types. The lowest speeds are observed at 0, -1, and -2 miles, occurring 10, 15, and 20 min after the occurrence time, respectively, demonstrating the propagation of a backward wave.

OBJ crashes contribute to the shortest congestion propagation distance since their CATE is nearly zero at -1 mile. OBJ crashes also exhibit the fastest recovery speed, with their CATE increasing from -10.91 to -3.61 mph within 30 min at 0 miles. This may be because OBJ crashes mostly occur at night, when fewer vehicles are on the road, and therefore, OBJ crashes do not affect many vehicles.

Table 2

The searched hyperparameters of each model.

Model	Hyperparameters	Search Range
XGBoost	n_estimators	[10, 40, 60, 80, 100]
	max_depth	[3, 4, 5, 6, 7, 8, 9, 10, None]
	learning_rate	[0.01, 0.02, 0.03, 0.04, 0.05]
	subsample	[0.5, 0.6, 0.7, 0.8, 0.9, 1.0]
	min_child_weight	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
	max_bin	[10, 40, 60, 80, 100]
LGBM	n_estimators	[10, 40, 60, 80, 100]
	max_depth	[3, 4, 5, 6, 7, 8, 9, 10, -1]
	learning_rate	[0.01, 0.02, 0.03, 0.04, 0.05]
	subsample	[0.5, 0.6, 0.7, 0.8, 0.9, 1.0]
	min_child_samples	[10, 20, 30, 40, 50]
	min_child_weight	[1e-3, 1e-2, 1e-1, 1, 1e1, 1e2]
RF.	max_bin	[255, 510, 765, 1020]
	n_estimators	[10, 15, 20, 25, 30]
	max_depth	[3, 5, 7, 9, 11, None]
	min_samples_split	[2, 3, 4, 5]
LR.	min_samples_leaf	[1, 2, 3, 4, 5]
	max_features	['auto', 'sqrt', 'log2', None]
	fit_intercept	[True, False]
	normalize	[True, False]
SVM	n_jobs	[-1, 1, 2, 4, 8]
	positive	[True, False]
	C	[0.001, 0.01, 0.1, 1, 10, 100]
	kernel	['linear', 'poly', 'rbf', 'sigmoid']
	degree	[1, 2, 3, 4, 5, 6]
	gamma	['scale', 'auto', 0.01, 0.1, 1, 10, 100]

Table 3

Performance results of classification for different types of crash type.

Model	Criterion	XGBoost	LGBM	RF.	SVM	Best Model
REAR	Accuracy	0.9688	0.9692	0.9704	0.5363	RF.
	Recall	0.9577	0.9552	0.9580	0.5414	
	Precision	0.9786	0.9820	0.9817	0.5212	
	F1	0.9679	0.9683	0.9696	0.4483	
OBJ	Accuracy	0.9930	0.9932	0.9938	0.6317	RF.
	Recall	0.9760	0.9771	0.9765	0.3348	
	Precision	0.9970	0.9967	0.9996	0.2671	
	F1	0.9862	0.9867	0.9878	0.2313	
WIPE	Accuracy	0.9740	0.9755	0.9770	0.5752	RF.
	Recall	0.9341	0.9372	0.9349	0.3327	
	Precision	0.9822	0.9841	0.9912	0.2726	
	F1	0.9572	0.9597	0.9619	0.2215	

Table 4

Performance results of regression for different types of crash type.

Model	Criterion	XGBoost	LGBM	RF.	LR.	Best Model
REAR	MAE	3.3800	3.2483	3.1836	4.8665	RF.
	MSE	46.3607	42.2842	42.3343	65.8335	
	RMSE	6.7393	6.4333	6.4389	8.0456	
	R ²	0.6403	0.6725	0.6734	0.4757	
	MAPE	0.1021	0.1004	0.0998	0.1425	
	MAE	1.6401	1.6005	1.5504	2.5869	
OBJ	MSE	17.3722	15.9016	15.8896	26.2485	RF.
	RMSE	4.0902	3.9126	3.9103	5.0345	
	R ²	0.6981	0.7241	0.7244	0.5309	
	MAPE	0.0409	0.0407	0.0402	0.0616	
	MAE	2.2389	2.1737	2.1321	3.3788	
	MSE	25.7684	23.4542	23.8749	37.7180	
WIPE	RMSE	5.0087	4.7752	4.8176	6.0562	RF.
	R ²	0.6922	0.7202	0.7248	0.5311	
	MAPE	0.0585	0.0576	0.0570	0.0847	

Table 5

The estimated causal effects of crashes on speed (mph) for different duration and distance.

Crash Type	disdur	0	-1	-2	-3	-5
REAR	5	-14.88***±2.78	-9.31***±2.49	-3.56***±1.95	-0.73 ± 1.51	0.36 ± 0.75
	10	-15.88***±2.49	-11.47***±3.06	-3.66***±2.43	-1.21**±1.66	0.48 ± 0.74
	15	-15.63***±2.70	-12.84***±3.42	-6.51***±2.83	-2.22 ± 2.09	0.42 ± 2.07
	20	-14.98***±2.74	-12.38***±3.15	-8.12***±3.14	-2.38***±2.12	0.35 ± 0.72
	25	-13.67***±2.63	-10.55***±2.86	-7.91***±3.24	-1.75 ***±1.21	0.54 ± 0.67
	30	-13.19***±2.87	-10.16***±2.95	-7.62***±2.97	-2.32**±1.79	0.55 ± 0.72
	5	-9.09***±5.65	-0.97 ± 3.55	-0.68 ± 2.54	-0.37 ± 1.99	-0.30 ± 2.44
OBJ	10	-10.91***±6.83	-1.81 ± 4.21	-0.79 ± 2.86	-0.39 ± 1.71	-0.31 ± 2.31
	15	-9.91*±6.69	-3.42 ± 4.86	-0.65 ± 2.22	-0.38 ± 1.83	-0.21 ± 1.73
	20	-9.62 ± 7.50	-2.95 ± 5.67	-0.57 ± 2.21	-0.55 ± 1.52	-0.14 ± 1.51
	25	-4.52*±3.22	-2.02 ± 5.99	-0.68 ± 1.81	-0.35 ± 1.77	-0.14 ± 1.59
	30	-3.61 ± 4.03	-1.27 ± 4.14	-0.71 ± 4.18	-0.45 ± 1.98	0.00 ± 1.97
	5	-10.40***±3.63	-2.06**±1.90	-1.31 ± 1.53	-0.75 ± 1.67	0.23 ± 1.15
	10	-11.29***±4.51	-3.25**±2.92	-1.41*±1.32	-0.55 ± 1.47	0.46 ± 1.30
WIPE	15	-12.08***±4.68	-4.22***±3.89	-1.72***±2.08	-0.43 ± 1.56	0.49 ± 1.04
	20	-12.96***±4.75	-4.67**±4.11	-2.42*±1.92	-0.79 ± 1.69	0.23 ± 1.11
	25	-11.46***±4.48	-3.30*±4.20	-2.33*±2.15	-0.72 ± 1.82	0.36 ± 1.24
	30	-10.8***±4.90	-2.80*±3.40	-2.70 ± 2.06	-0.70 ± 1.81	0.20 ± 1.37

Note: ***P < 0.001; **P < 0.01; *P < 0.05; ±95 %Confidence Interval.

WIPE crashes result in slightly less severe congestion than REAR crashes but have the longest hysteresis of congestion. The maximum deceleration occurs 20 min after the occurrence of WIPE crashes, while it takes only 10 min for other crash types. This indicates that sideswipe crashes have obviously delayed impact and need to be addressed more timely and efficiently, such as by deploying clean-up crews and tow trucks.

The impact of time on the causal crash effect is explored in Fig. 10. The peak hours (6:30–9:00 and 16:40–19:30) and night periods (23:00–4:00) are found to have different effects on crashes compared to off-peak hours (the rest time). According to Fig. 10, it can be found that the trends between REAR and OBJ in different periods are opposite, while the effect of WIPE remains stable. The underlying factors contributing to this phenomenon are that REAR and WIPE tend to occur during peak hours with higher volume. Additionally, because average speeds at night surpass those during peak hours, the impacts of REAR and WIPE are more pronounced during the nighttime hours. Conversely, OBJ is usually triggered by poor lighting conditions at night and tends to occur when traffic volume is relatively low. As a result, OBJ crashes occurring during peak hours exert a more substantial influence on average speeds. These findings provide insights into the heterogeneous effects of crashes on traffic speed, which can be valuable for policymakers in developing effective traffic management strategies.

5.5.2. Validation of estimated causal effect

This study validates the proposed method from two aspects to demonstrate its rationality. Firstly, Fig. 11 illustrates the change process of traffic speeds affected by crashes through several individual examples, where the occurrence times of crashes are set as 0 on the x-axis. The orange lines represent the actual speeds before and after crashes, and the green lines represent the average speed without crash treatment, determined from the traffic condition one week after the crashes. The green area denotes the individual treatment effect since the effect is inferred for a single crash. The figure shows that the counterfactual speeds demonstrate reasonable trends and quantities compared to average speeds, proving that the proposed method provides an acceptable estimation of causal effects.

Secondly, we evaluate the overall estimate performance by calculating error metrics based on the matched “counterfactual outcomes” generated by matching algorithm. Furthermore, we used GRF and DML as comparisons, as shown in Table 6. The main difference between GRF and DML lies in their use of decision trees. GRF is an extension of the traditional random forest algorithm that uses a double-robust estimation procedure to adjust for confounding variables. As for DML, it also estimates the causal effects by two stages process but predicts the outcome from just the controls rather than both treatment and controls in the first stage. Because DML does not use double-robust estimation, more bias may be introduced into the final regression. To calculate the error metrics, we substituted the estimated value \hat{y}_i and actual value y_i in Equation (19)–(22) with ice_i and mce_i (Equation (14)), respectively. We found that the DRL model outperforms the other methods in terms of accuracy. This further validates the rationality and effectiveness of our proposed method.

5.5.3. Validation of variables selection

We validated the effectiveness of variable selection by sensitivity analysis from two perspectives. Firstly, we compared the estimated results and error performance before and after the selection process by DRL. Fig. 12 illustrates that without variable selection, the CATE is nearly zero, which is inconsistent with real-world phenomena and indicates the failure of causal inference. Table 7 shows the error performance and validates the improvement due to variable selection. Additionally, when estimating the errors for different times and locations, Fig. 13 indicates that the effectiveness of variable selection reduces as time and distance increase. This means that the proposed method has no advantages in terms of spatial and temporal causal inference.

Secondly, the sensitivity analysis for CSVI is conducted. We validated the estimated performance based on different threshold

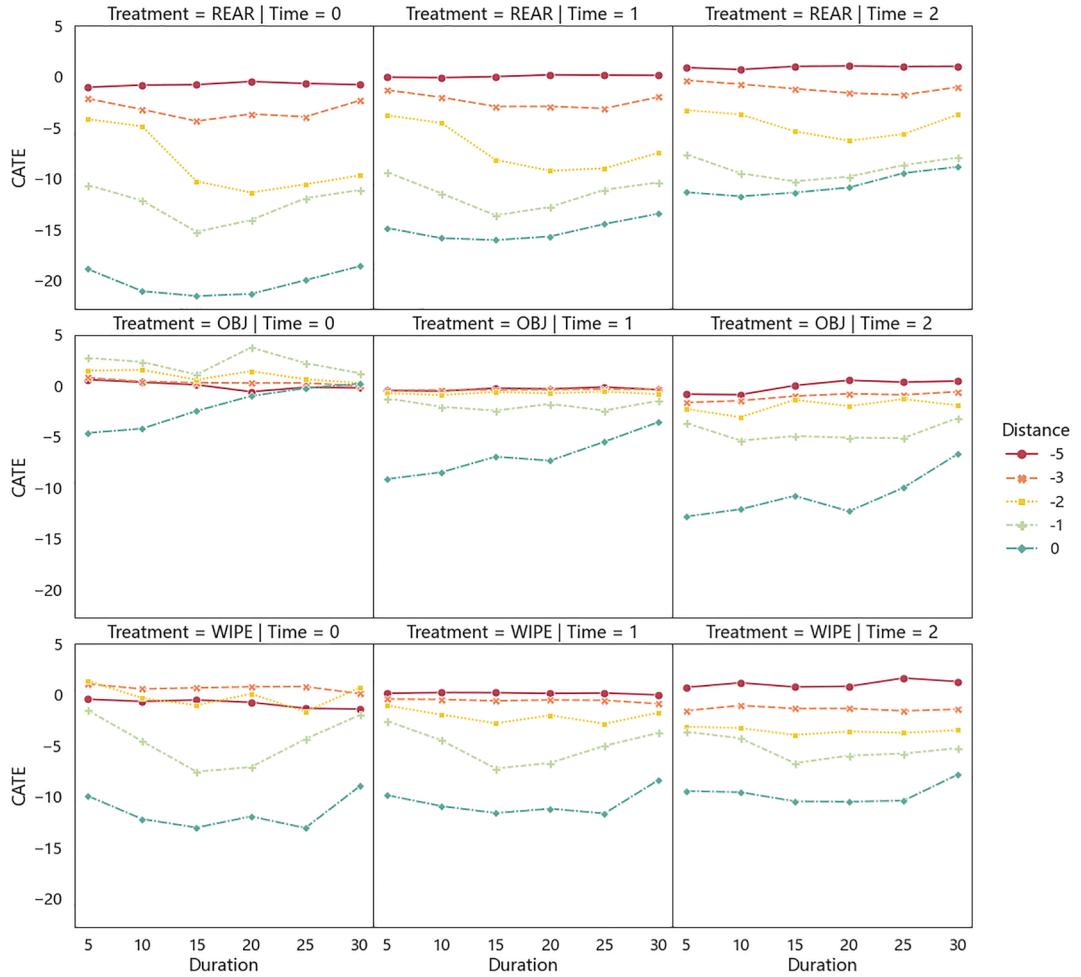


Fig. 10. Comparison of CATE in different time periods (0: night, 1: off-peak, 2: peak hour).

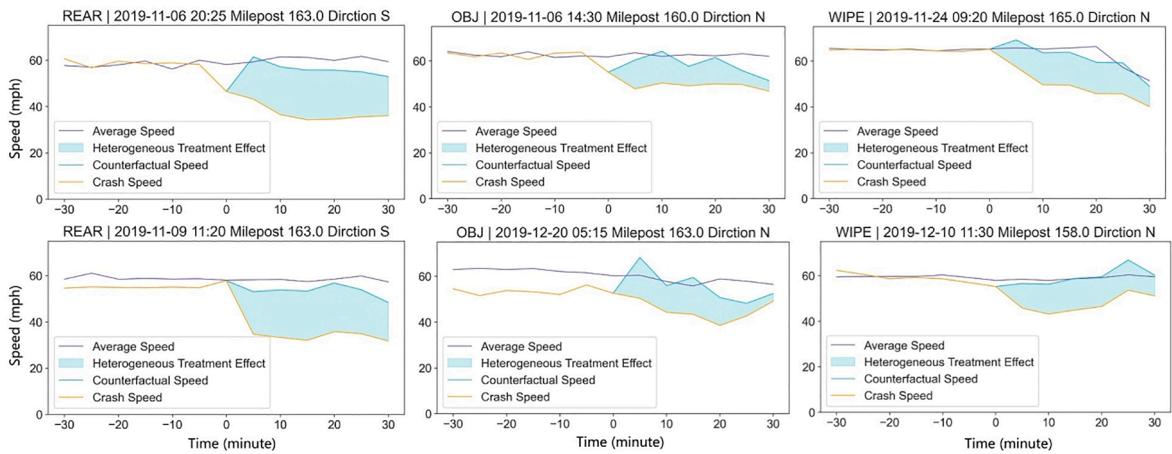


Fig. 11. Examples of individual treatment effects compared by matched average speed. The green areas between green and orange lines refer to estimated effects.

Table 6

Performance comparison based on matched data.

Methods	MAE	MSE	RMSE	MAPE
GRF	12.60	345.18	17.76	0.22
DML	12.81	322.54	16.88	0.24
DRL	10.73	244.59	14.76	0.20

values, ranging from 0 to 0.40, with a 0.05 step. Based on the results in [Table 8](#), it is evident that the optimal performance is achieved when the threshold is set to 0.15. In addition, the relatively small variation in prediction errors to some extent reflects the robustness of our method.

6. Conclusions

This paper presents a novel causal machine learning framework for estimating heterogeneous treatment effect of traffic crashes on speed. The framework utilizes Doubly Robust Learning models, which combine machine learning and doubly robust inference methods. Additionally, the framework incorporates causal inference theory and notations to guide variable selection and causal structure analysis. In order to showcase the effectiveness and precision of the proposed framework, real-world data encompassing incident records, traffic flow information, and road geometry details were utilized from the Interstate 5 freeway located in Washington.

The results of the estimated CATE by proposed method reveal heterogeneous effects of crashes on speed. Overall, REAR exhibits the most significant impact, followed by WIPE, and finally OBJ. The estimated effects also reveal the propagation and dissipation mechanism of traffic congestion. Specifically, REAR has a greater range of traffic impacts upstream, while OBJ has the shortest range of impacts. The delayed effects of WIPE are particularly pronounced. Furthermore, REAR exhibits more adverse effects at night, whereas OBJ demonstrates more adverse effects during peak hours, which can be attributed to the traffic flow conditions. The results have been validated by comprehensive tests, supporting their effectiveness and accuracy. This proposed model and the findings can significantly aid authorities in developing more effective emergency countermeasures for highway safety.

This study offers a new perspective for analyzing the causality between highway crashes and traffic speed. However, there are still avenues for future research, including analyzing the causality of other factors that are hard to observe in normal traffic flow, such as drinking, gender, and age. Additionally, the proposed model cannot deal with the spatial and temporal causal relationship since traffic data are time series and have spatial dependencies. Therefore, modifying the causal inference methods to improve the ability for spatiotemporal data and predicting counterfactual real-time traffic conditions would be worthwhile.

CRediT authorship contribution statement

Shuang Li: Writing – original draft, Visualization, Software, Methodology, Data curation, Conceptualization. **Ziyuan Pu:** Writing – original draft, Supervision, Resources, Methodology, Investigation, Conceptualization. **Zhiyong Cui:** Writing – review & editing,

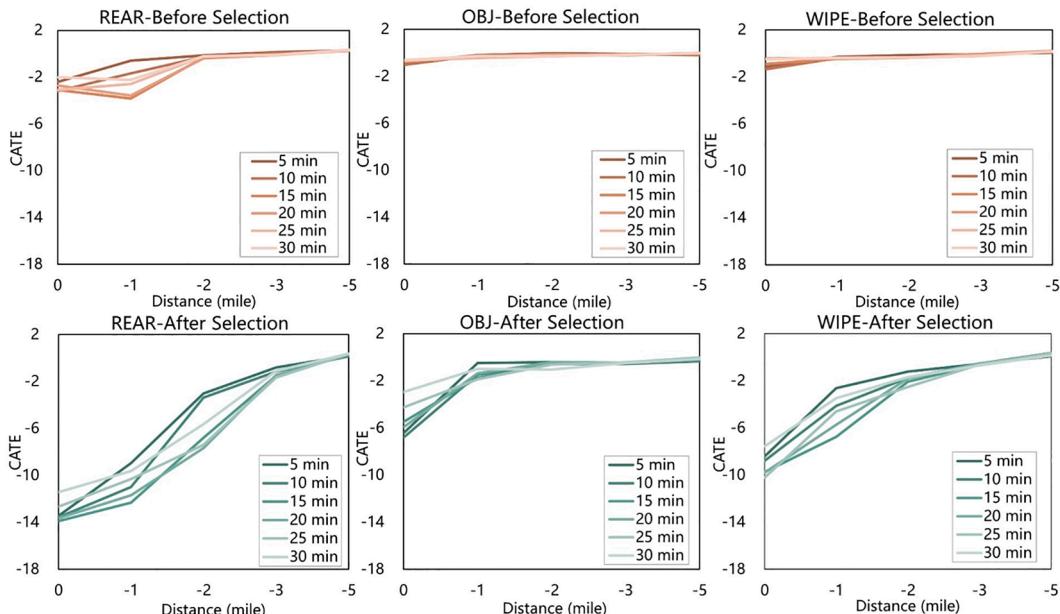
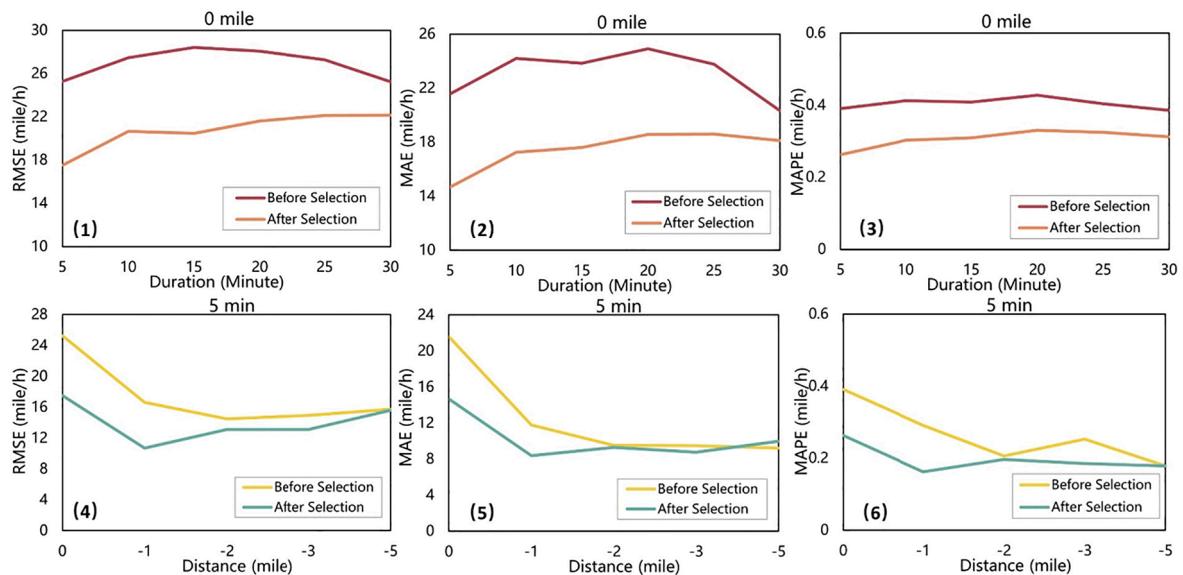


Fig. 12. Estimated CATE before and after variables selection. Each line represents the CATE 5–30 min after the crash, respectively.

Table 7

Average performance of DRL before and after variables selection.

Type	Method	MAE	MSE	RMSE	MAPE
REAR	Before selection	10.74	259.12	15.01	0.21
	After selection	9.17	158.94	12.37	0.17
OBJ	Before selection	12.6	375.93	17.7	0.23
	After selection	11.01	367.72	16.3	0.21
WIPE	Before selection	11.40	293.48	15.99	0.23
	After selection	10.03	203.2	13.66	0.18

**Fig. 13.** Estimated error before and after variables selection. (1–3) are the errors of crash sites changing with time; (4–6) are the errors 5 min after crashes changing with distance.**Table 8**

The average error performance for different threshold of CSVI.

Threshold	MAE	MSE	RMSE	MAPE
0.00	7.82	129.26	10.49	0.1380
0.05	7.81	129.27	10.49	0.1379
0.10	7.79	128.57	10.46	0.1376
0.15	7.77	127.98	10.44	0.1372
0.20	7.82	129.38	10.50	0.1381
0.25	7.80	128.96	10.49	0.1378
0.30	7.85	130.38	10.58	0.1383
0.35	8.33	135.32	11.01	0.1462
0.40	8.56	133.52	10.92	0.1496

Supervision, Methodology, Investigation. **Seunghyeon Lee:** Writing – review & editing, Methodology, Conceptualization. **Xiucheng Guo:** Writing – review & editing, Methodology, Conceptualization. **Dong Ngoduy:** Writing – review & editing, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (project number: 52202378), the

Postgraduate Research&Practice Innovation Program of Jiangsu Province (project number: KYCX22_0271), the Ministry of Transport of PRC Key Laboratory of Transport Industry of Comprehensive Transportation Theory (Nanjing Modern Multimodal Transportation Laboratory) (project number: MTF2023002).

References

- Adler, M.W., van Ommeren, J., Rietveld, P., 2013. Road congestion and incident duration. *Econ. Transp.* 2, 109–118. <https://doi.org/10.1016/j.ecotra.2013.12.003>.
- Athey, S., Wager, S., 2019. Estimating Treatment Effects with Causal Forests: An Application. <https://doi.org/10.48550/arXiv.1902.07409>.
- Athey, S., Imbens, G.W., 2016. Recursive partitioning for heterogeneous causal effects. *PNAS* 113, 7353–7360. <https://doi.org/10.1073/pnas.1510489113>.
- Athey, S., Tibshirani, J., Wager, S., 2019. Generalized random forests. *Ann. Stat.* 47, 1148–1178. <https://doi.org/10.1214/18-aos1709>.
- Battocchi, K., Dillon, E., Hei, M., Lewis, G., Oka, P., Oprescu, M., Syrgkanis, V., 2019. EconML: A Python Package for ML-Based Heterogeneous Treatment Effects Estimation.
- Benlagha, N., Charfeddine, L., 2020. Risk factors of road accident severity and the development of a new system for prevention: new insights from China. *Accid. Anal. Prev.* 136, 105411 <https://doi.org/10.1016/j.aap.2019.105411>.
- Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* 13, 281–305. <https://doi.org/10.5555/2503308.2188395>.
- Blakely, T., Lynch, J., Simons, K., Bentley, R., Rose, S., 2021. Reflection on modern methods: when worlds collide—prediction, machine learning and causal inference. *Int. J. Epidemiol.* 49 <https://doi.org/10.1093/ije/dyz132>.
- Blincore, L., Miller, T., Wang, J.-S., Swedler, D., Coughlin, T., Lawrence, B., Guo, F., Klauer, S., Dingus, T., 2023. The Economic and Societal Impact of Motor Vehicle Crashes 2019 (Revised) (No. DOT HS 813 403). National Highway Traffic Safety Administration.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/a:1010933404324>.
- Brookhart, M.A., Schneeweiss, S., Rothman, K.J., Glynn, R.J., Avorn, J., Stürmer, T., 2006. Variable selection for propensity score models. *Am. J. Epidemiol.* 163, 1149–1156. <https://doi.org/10.1093/aje/kwj149>.
- Cao, D., Wu, J., Dong, X., Sun, H., Qu, X., Yang, Z., 2021. Quantification of the impact of traffic incidents on speed reduction: a causal inference based approach. *Accid. Anal. Prev.* 157, 106163 <https://doi.org/10.1016/j.aap.2021.106163>.
- Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. *KDD* 785–794. <https://doi.org/10.1145/2939672.2939785>.
- Chen, Z., Liu, X.C., Zhang, G., 2016. Non-recurrent congestion analysis using data-driven spatiotemporal approach for information construction. *Transport. Res. Part C: Emerg. Technol.* 71, 19–31. <https://doi.org/10.1016/j.trc.2016.07.002>.
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W.K., 2016. Double machine learning for treatment and causal parameters. *arXiv: Machine Learning*. <https://doi.org/10.1920/wp.cem.2016.4916>.
- Chung, Y., 2017. Identification of critical factors for non-recurrent congestion induced by urban freeway crashes and its mitigating strategies. *Sustainability* 9, 2331. <https://doi.org/10.3390/su9122331>.
- Chung, Y., Recker, W.W., 2012. A methodological approach for estimating temporal and spatial extent of delays caused by freeway accidents. *IEEE Trans. Intell. Transport. Syst.* 13, 1454–1461. <https://doi.org/10.1109/TITS.2012.2190282>.
- Chung, Y., Recker, W.W., 2013. Spatiotemporal analysis of traffic congestion caused by rubbernecking at freeway accidents. *IEEE Trans. Intell. Transport. Syst.* 14, 1416–1422. <https://doi.org/10.1109/TITS.2013.2261987>.
- Dabbour, E., Dabbour, O., Martinez, A.A., 2020. Temporal stability of the factors related to the severity of drivers' injuries in rear-end collisions. *Accid. Anal. Prev.* 142, 105562 <https://doi.org/10.1016/j.aap.2020.105562>.
- De Luna, X., Waernbaum, I., Richardson, T.S., 2011. Covariate selection for the nonparametric estimation of an average treatment effect. *Biometrika* 98, 861–875. <https://doi.org/10.1093/biomet/asr041>.
- Ding, H., Sze, N.N., 2022. Effects of road network characteristics on bicycle safety: a multivariate poisson-lognormal model. *Multimodal Transport.* 1, 100020 <https://doi.org/10.1016/j.multra.2022.100020>.
- Foster, D.J., Syrgkanis, V., 2019. Orthogonal Statistical Learning. *arXiv: Statistics Theory*.
- Garib, A., Radwan, A.E., Al-Deek, H., 1997. Estimating magnitude and duration of incident delays. *J. Transport. Eng.-ASCE* 123, 459–466. [https://doi.org/10.1061/\(asce\)0733-947x\(199712\)123:6\(459\)](https://doi.org/10.1061/(asce)0733-947x(199712)123:6(459)).
- Graham, D.J., Naik, C., Naik, C., McCoy, E.J., Li, H., 2019. Do speed cameras reduce road traffic collisions. *PLoS One* 14. <https://doi.org/10.1371/journal.pone.0221267>.
- Grigorev, A., Mihaita, A.-S., Lee, S., Chen, F., 2022. Incident duration prediction using a bi-level machine learning framework with outlier removal and intra-extra joint optimisation. *Transportation Research Part c: Emerging Technologies* 141, 103721. <https://doi.org/10.1016/j.trc.2022.103721>.
- Hammond, R.L., Soccolich, S.A., Hanowski, R.J., Hanowski, R.J., 2019. The impact of driver distraction in tractor-trailers and motorcoach buses. *Accid. Anal. Prev.* 126, 10–16. <https://doi.org/10.1016/j.aap.2018.03.015>.
- Holland, P.W., 1985. Statistics and causal inference. *ETS Res. Report Series* 1985. <https://doi.org/10.1080/01621459.1986.10478354>.
- Karwa, V., Slavkovic, A.B., Donnell, E.T., 2011. Causal inference in transportation safety studies: comparison of potential outcomes and causal diagrams. *Ann. Appl. Stat.* <https://doi.org/10.1214/10-aosa440>.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Chen, W., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: a highly efficient gradient boosting decision tree. *NIPS* 30, 3149–3157.
- Knaus, M.C., 2018. A Double Machine Learning Approach to Estimate the Effects of Musical Practice on Student's Skills. *arXiv: Econometrics*. <https://doi.org/10.1111/rssc.12623>.
- Kristjanpoller, W., Michell, K., Michell, V.K., Minutolo, M.C., Minutolo, M.C., 2021. A causal framework to determine the effectiveness of dynamic quarantine policy to mitigate COVID-19. *Appl. Soft Comput.* 104, 107241. <https://doi.org/10.1016/j.asoc.2021.107241>.
- S. Künzel, J. Sekhon, P. Bickel, Bin Yu, 2017. Meta-learners for estimating heterogeneous treatment effects using machine learning.
- Li, H., Graham, D.J., 2016. Quantifying the causal effects of 20 mph zones on road casualties in London via doubly robust estimation. *Accid. Anal. Prev.* 93, 65–74. <https://doi.org/10.1016/j.aap.2016.04.007>.
- Li, H., Graham, D.J., Majumdar, A., 2013. The impacts of speed cameras on road accidents: an application of propensity score matching methods. *Accid. Anal. Prev.* 59, 148–155. <https://doi.org/10.1016/j.aap.2013.08.003>.
- Li, H., Graham, D.J., Ding, H., Ren, G., 2019. Comparison of empirical Bayes and propensity score methods for road safety evaluation: a simulation study. *Accid. Anal. Prev.* 129, 148–155. <https://doi.org/10.1016/j.aap.2019.05.015>.
- Lin, Y., Li, R., 2020. Real-time traffic accidents post-impact prediction: based on crowdsourcing data. *Accid. Anal. Prev.* 145, 105696 <https://doi.org/10.1016/j.aap.2020.105696>.
- Liu, X., Qian, S., Ma, W., 2022. Estimating and Mitigating the Congestion Effect of Curbside Pick-ups and Drop-offs: A Causal Inference Approach. <https://doi.org/10.48550/arXiv.2206.02164>.
- Luan, S., Ke, R., Huang, Z., Ma, X., 2022. Traffic congestion propagation inference using dynamic Bayesian graph convolution network. *Transport. Res. Part c: Emerg. Technol.* 135, 103526 <https://doi.org/10.1016/j.trc.2021.103526>.
- Manning, F.L., Bhat, C.R., Shankar, V., Abdel-Aty, M., 2020. Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis. *Anal. Methods Accident Res.* 25, 100113 <https://doi.org/10.1016/j.amar.2020.100113>.
- Miller, M., Gupta, C., 2012. Mining traffic incidents to forecast impact. In: Proceedings of the ACM SIGKDD International Workshop on Urban Computing, UrbComp '12. Association for Computing Machinery, New York, NY, USA, pp. 33–40. <https://doi.org/10.1145/2346496.2346502>.
- Morgan, S.L., Winship, C., 2007. Counterfactuals and Causal Inference: Methods and Principles for Social Research.

- Ou, J., Jishun Ou, Jishun Ou, Xia, J., Wang, Y., Wang, C., Chen Wang, Lu, Z., Zhenbo Lu, 2020. A data-driven approach to determining freeway incident impact areas with fuzzy and graph theory-based clustering. *Computer-aided Civil and Infrastructure Engineering* 35, 178–199. <https://doi.org/10.1111/mice.12484>.
- Pan, B., Demiryurek, U., Gupta, C., Shahabi, C., 2015. Forecasting spatiotemporal impact of traffic incidents for next-generation navigation systems. *Knowl. Inf. Syst.* 45, 75–104. <https://doi.org/10.1007/s10115-014-0783-6>.
- Pasidis, I., 2019. Congestion by accident? a two-way relationship for highways in England. *J. Transp. Geogr.* 76, 301–314. <https://doi.org/10.1016/j.jtrangeo.2017.10.006>.
- Pearl, J., 2009. Causal inference in statistics: an overview. *Statistics Surveys* 3, 96–146. <https://doi.org/10.1214/09-ss057>.
- Pearl, J., 2011. Invited commentary: understanding bias amplification. *Am J Epidemiol* 174, 1223–1227. <https://doi.org/10.1093/aje/kwr352>.
- Pearl, J., 2014. Comment: understanding Simpson's paradox. *Am. Stat.* 68, 8–13. <https://doi.org/10.1080/00031305.2014.876829>.
- Prosperi, M., Guo, Y., Sperrin, M., Koopman, J.S., Min, J., He, X., Rich, S.N., Wang, M., Buchan, I., Bian, J., Bian, J.-G., 2020. Causal inference and counterfactual prediction in machine learning for actionable healthcare. *Nat. Mach. Intell.* 2, 369–375. <https://doi.org/10.1038/s42256-020-0197-y>.
- Pu, Z., Li, Z., Ke, R., Hua, X., Wang, Y., 2020. Evaluating the nonlinear correlation between vertical curve features and crash frequency on highways using random forests. *J. Transport. Eng., Part A: Syst.* 146, 04020115. <https://doi.org/10.1061/JTEPBS.0000410>.
- Robins, J.M., Rotnitzky, A., Zhao, L.P., 1995. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J. Am. Stat. Assoc.* 90, 106–121. <https://doi.org/10.1080/01621459.1995.10476493>.
- Rolison, J.J., Regev, S., Moutari, S., Feeney, A., 2018. What are the factors that contribute to road accidents? an assessment of law enforcement views, ordinary drivers' opinions, and road accident records. *Accid. Anal. Prev.* 115, 11–24. <https://doi.org/10.1016/j.aap.2018.02.025>.
- Rosenbaum, P.R., 1987. Model-based direct adjustment. *J. Am. Stat. Assoc.* 82, 387–394. <https://doi.org/10.1080/01621459.1987.10478441>.
- Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55. <https://doi.org/10.1093/biomet/70.1.41>.
- Rotnitzky, A., Smucler, E., 2020. Efficient adjustment sets for population average causal treatment effect estimation in graphical models. *J. Mach. Learn. Res.* 21, 1–86.
- Rubin, D.B., 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* 66, 688–701. <https://doi.org/10.1037/h0037350>.
- Schisterman, E.F., Perkins, N.J., Perkins, N.J., Kannan, K., Mumford, S.L., Ahrens, K.A., Mitchell, E.M., Mitchell, E.M., 2017. Collinearity and causal diagrams: a lesson on the importance of model specification. *Epidemiology* 28, 47–53. <https://doi.org/10.1097/ede.0000000000000554>.
- Schnitzer, M.E., Lok, J.J., Gruber, S., 2016. Variable selection for confounder control, flexible modeling and collaborative targeted minimum loss-based estimation in causal inference. *Int. J. Biostatistics* 12, 97–115. <https://doi.org/10.1515/ijb-2015-0017>.
- Snelder, M., Bakri, T., van Arem, B., 2013. Delays caused by incidents: data-driven approach. *Transp. Res.* 2333, 1–8. <https://doi.org/10.3141/2333-01>.
- Song, Y., Noyce, D.A., 2019. Effects of transit signal priority on traffic safety: interrupted time series analysis of Portland, Oregon, implementations. *Accid. Anal. Prev.* 123, 291–302. <https://doi.org/10.1016/j.aap.2018.12.001>.
- Spirites, P., Scheines, R., 2004. Causal inference of ambiguous manipulations. *Philos. Sci.* 71, 833–845. <https://doi.org/10.1086/425058>.
- Splawa-Neyman, J., Dabrowska, D.M., Speed, T.P., 1990. On the application of probability theory to agricultural experiments. essay on principles. section 9. *Stat. Sci.* 5, 465–472. <https://doi.org/10.1214/ss/1177012031>.
- Su, X., Zhi, D., Song, D., Tian, L., Yang, Y., 2023. Exploring weather-related factors affecting the delay caused by traffic incidents: mitigating the negative effect of traffic incidents. *Sci. Total Environ.* 877, 162938. <https://doi.org/10.1016/j.scitotenv.2023.162938>.
- Tang, D., Kong, D., Pan, W., Wang, L., 2020. Ultra-high Dimensional Variable Selection for Doubly Robust Causal Inference.
- Wen, X., Xie, Y., Jiang, L., Pu, Z., Ge, T., 2021. Applications of machine learning methods in traffic crash severity modelling: current status and future directions. *Transp. Rev.* 41, 855–879. <https://doi.org/10.1080/01441647.2021.1954108>.
- Witte, J., Didelez, V., Didelez, V., 2019. Covariate selection strategies for causal inference: classification and comparison. *Biom. J.* 61, 1270–1289. <https://doi.org/10.1002/bimj.201700294>.
- Xie, Q., Guo, T., Chen, Y., Xiao, Y., Wang, X., Zhao, B.Y., 2019. "How do urban incidents affect traffic speed?" A Deep Graph Convolutional Network for Incident-driven Traffic Speed Prediction. <https://doi.org/10.48550/arXiv.1912.01242>.
- Xie, Y., Zhao, K., Huynh, N., 2012. Analysis of driver injury severity in rural single-vehicle crashes. *Accid. Anal. Prev.* 47, 36–44. <https://doi.org/10.1016/j.aap.2011.12.012>.
- Yao, L., Zhixuan Chu, Chu, Z., Li, S., Yaliang Li, Li, Y., Gao, J., Zhang, A., 2020. A Survey on Causal Inference. *arXiv: Methodology*. <https://doi.org/10.1145/3444944>.
- Yu, R., Abdel-Aty, M., 2014. Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Saf. Sci.* 63, 50–56. <https://doi.org/10.1016/j.ssci.2013.10.012>.
- Yue, Q., Feng, Z., Shao, C., Huang, Z., Ruan, X., 2024. Factors impacting bus selection: differences between the middle and later stages of COVID-19. *Multimodal Transport.* 3, 100106. <https://doi.org/10.1016/j.multra.2023.100106>.
- Zhang, Z., Akinci, B., Qian, S., 2022a. Inferring the causal effect of work zones on crashes: methodology and a case study. *Anal. Methods Accident Res.* 33, 100203. <https://doi.org/10.1016/j.amar.2021.100203>.
- Zhang, Z., Akinci, B., Qian, S., 2022b. Inferring heterogeneous treatment effects of work zones on crashes. *Accid. Anal. Prev.* 177, 106811. <https://doi.org/10.1016/j.aap.2022.106811>.
- Zhang, Y., Li, H., Ren, G., 2021a. Estimating heterogeneous treatment effects in road safety analysis using generalized random forests. *Accid. Anal. Prev.* 165, 106507. <https://doi.org/10.1016/j.aap.2021.106507>.
- Zhang, Y., Li, H., Sze, N.N., Ren, G., Ren, G., 2021b. Propensity score methods for road safety evaluation: practical suggestions from a simulation study. *Accid. Anal. Prev.* 158, 106200. <https://doi.org/10.1016/j.aap.2021.106200>.
- Zheng, Z., Wang, Z., Zhu, L., Jiang, H., 2020. Determinants of the congestion caused by a traffic accident in urban road networks. *Accid. Anal. Prev.* 136, 105327. <https://doi.org/10.1016/j.aap.2019.105327>.