A causal discovery approach to study key mixed traffic-related factors and age of highway affecting raveling

Wang, Zili; Krishnakumari, Panchamy; Anupam, Kumar; van Lint, Hans; Erkens, Sandra

DOI
10.1111/mice.13222

Publication date
2024

Document Version
Final published version

Published in
Computer-Aided Civil and Infrastructure Engineering

Citation (APA)
Wang, Z., Krishnakumari, P., Anupam, K., van Lint, H., & Erkens, S. (2024). A causal discovery approach to study key mixed traffic-related factors and age of highway affecting raveling. *Computer-Aided Civil and Infrastructure Engineering*, *39*(19), 2861-2880. https://doi.org/10.1111/mice.13222

Accepted: 18 April 2024

DOI: 10.1111/mice.13222

RESEARCH ARTICLE

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

A causal discovery approach to study key mixed traffic-related factors and age of highway affecting raveling

Zili Wang | **Panchamy Krishnakumari** | **Kumar Anupam** | **Hans van Lint** | **Sandra Erkens**

Faculty of Civil Engineering and Geosciences, Delft University of Technology, South Holland, The Netherlands

Correspondence
Kumar Anupam, Building 23, Stevinweg 1, 2628 CN, Delft, South Holland, The Netherlands.
Email: k.anupam@tudelft.nl

Funding information
Ministry of Infrastructure and Water Management, Grant/Award Number: 31160205; Rijkswaterstaat; Ministerie van Infrastructuur en Waterstaat

Abstract

The relationship between real-world traffic and pavement raveling is unclear and subject to ongoing debates. This research proposes a novel approach that extends beyond traditional correlation analyses to explore causal mechanisms between mixed traffic and raveling. This approach incorporates the causal discovery method, and is applied to five Dutch porous asphalt (PA) highway sites that have substantial data sets. Findings indicate a nonlinear relationship between traffic volume and raveling, with road age emerging as a shared contributor. The results also suggest that the degree to which different vehicle types contribute as a causal factor for raveling varies with carriageway configurations and lane characteristics. This underlines the need for targeted maintenance strategies. Challenges remain due to confounding correlations among traffic variables, necessitating further development of causal discovery models. This study may not conclusively resolve the debate on to what extent traffic contributes to raveling, but we argue we provide sufficient evidence against rejecting this hypothesis.

1 | INTRODUCTION

Maintaining high-quality pavement performance is a fundamental need for facilitating effective road transportation. Because of various kinds of distresses, performance deteriorates over time. Raveling—the dislodgement of aggregates from a pavement surface—stands out as a primary form of damage to porous asphalt (PA) pavements (Zhang et al., 2016). Such pavements are widely used across the Netherlands, Japan, New Zealand, France, Germany, Italy, and Spain. PA pavements are favored for their noise reduction properties (Donavan, 2014; Ghafoori, 2019), enhanced safety in wet conditions (Takahashi, 2013), and other benefits associated with their high porosity (Wang et al., 2021). A comprehensive understanding of raveling is of a vital importance for these road authorities to preserve road quality and optimize maintenance costs.

The susceptibility of pavement to raveling is generally understood to be strongly associated with its mix design, construction quality, and the environmental conditions (Abouelsaad & White, 2021). It is probably for this reason that existing raveling studies have primarily concentrated on material variants (Mo et al., 2010, 2014; Van Loon & Butcher, 2003; You et al., 2018), production and paving specifications (Kuennen, 2013; Van Reisen et al., 2008; You et al., 2018), and aging process (Hagos, 2008; Jing, 2019; Opara et al., 2016).

While the factors have been extensively explored, traffic generating shear forces that can lead to aggregate dislodgement (De Visscher & Vanelstraete, 2017; Kuennen, 2013;

Zhang et al., 2020) has not been deeply delved into apart from the effects of traffic axle loads (Ghadi et al., 2023). Simulating traffic affecting raveling involves challenges (Nicholls et al., 2019). A key difficulty is accurately replicating the traffic and environmental conditions to which a pavement is exposed (Barzegari & Solaimanian, 2019; Zhang et al., 2020). The variability of traffic patterns and real-world conditions, combined with the fact that raveling is a process that unfolds over several years, requires the generation of numerous elements for a realistic simulation. The requirement dramatically increases the demand for computational power and time (Hospodka & Hofko, 2019). Although advancements in collecting observational data offer potential solutions to the issue, the relationship between traffic and raveling remains inconclusive and is subject to ongoing discussion. Many field analyses suggest a positive correlation between traffic flow and raveling (Jain et al., 2005; Miradi, 2009). However, these findings are not undisputed. Henning and Roux (2012) investigated the open-graded PA in New Zealand and did not find significant correlations between axle loads and initiation of raveling. Similarly, Abouelsaad and White (2020) concluded that traffic flow was not necessarily a contributing factor to raveling, as instances of raveling were observed on runway pavements both with and without substantial traffic. Nicholls et al. (2016) presented that traffic played a significant role in early-stage raveling, especially due to intense shear forces at the tire–pavement interface, commonly observed on newly constructed roundabouts. However, its contribution to long-term raveling appeared to diminish, with factors like aging and weather taking precedence. As such, current evidence for the relationship between traffic and raveling is not only inconclusive, but it is also characterized by varied interpretations.

To contribute to the discussion, this study proposes an innovative approach to examine the causal mechanisms between raveling and traffic using (a) large amounts (multiple years) of observational data and (b) a novel combination of causal discovery techniques. Traditionally, the relationship between raveling and traffic are investigated by laboratory tests and correlation-based models, but these approaches have limitations in providing conclusive evidence. Laboratory experiments necessarily oversimplify actual traffic conditions, particularly in terms of simulating a pavement's *long-term* exposure to varying traffic conditions. Observational studies (using field data), even when employing advanced machine learning models as elucidated by Attoh-Okine (2001), primarily find correlations. They are valuable, but also inconclusive in terms of evidence for causal patterns. This limits their ability to differentiate between direct causes, indirect causes, and (confounding) variables that are correlated to the direct or indirect causes, but are not causative themselves.

Therefore, the study proposes causal discovery to delve deeper into causal aspects of the relationship between traffic and raveling. Causal discovery is a data-driven approach to examine whether causal relationships between variables are plausible (Glymour et al., 2019). Limited models have been applied to learn causal effects from performance (Cai et al., 2023) and rutting data (Zhang et al., 2023), but there is a gap in raveling research. Various models have emerged with distinct methodologies. Constraint-based models infer causal structures by testing for (in)dependencies observed in the data set (Guo et al., 2020). Score-based models, on the other hand, evaluate causal structures based on how well they fit the data, assigning scores to reflect this fit (Guo et al., 2020). Both types have their roots in frameworks developed for Gaussian distributions (Shimizu et al., 2006). In contrast, functional causal models (FCMs) are designed with non-Gaussian data in mind, positing that causal relationships manifest as mathematical functions of the variables involved (Glymour et al., 2019). Besides, such a model can also quantify causal effects, which is known as causal inference (Glymour et al., 2019).

Causal discovery and inference models have undergone significant evolution to reduce model requirements and expand their applicability. One of the original constraint-based models, the Peter–Clark (PC) algorithm has been refined to accommodate latent variables (Spirtes et al., 1999), enhance computational speed (Colombo et al., 2012), and adaptively handle time-series data (Huang et al., 2020). The score-based models such as Greedy Equivalence Search (GES) (Chickering, 2002) have been extended to achieve faster computation speed, using the strengths of constraint-based models (Ogarrio et al., 2016). The Linear Non-Gaussian Acyclic Model (LiNGAM) (Shimizu et al., 2006), a fundamental FCM, has been developed into two directions. The first emphasizes rapid computation (Shimizu et al., 2011), and relaxes the constraints on variables, removing the strict necessity for them to be independent and identically distributed (Hyvärinen et al., 2010). The second direction has tailored LiNGAM to handle nonlinear cases (Zhang & Hyvärinen, 2012), without relying on Markov equivalent class (Hoyer et al., 2008; Peters et al., 2014), and incorporating latent variables (Maeda & Shimizu, 2022; Xie et al., 2020).

Based on the model requirements and assumptions, the research outlines criteria for selecting an appropriate model to investigate causal relationships, specifically aimed at assisting researchers in identifying the most suitable causal discovery models for their particular cases. By the criteria, the study identifies the most fitting model for analyzing observational data related to raveling and traffic. The proposed criteria are as follows:

TABLE 1 Characteristics of causal discovery models regarding the selection criteria.

Causal discovery models	I.I.D.	Causal sufficiency	Gaussian	Markov condition	Linearity	NP-hardness
Peter–Clark (PC) (Spirtes et al., 2000)	y	y	y	y		y
CD-NOD (Huang et al., 2020)		y		y		y
Fast Causal Inference (FCI) (Spirtes et al., 1999)	y		y	y		y
Really Fast Causal Inference (RFCI) (Colombo et al., 2012)	y		y	y		
Greedy Fast Causal Inference (GFCI) (Ogarrio et al., 2016)	y		y	y		
Greedy Equivalence Search (GES) (Chickering, 2002)	y	y	y	y		y
Bayesian network structure search (Yuan & Malone, 2013)	y	y	y	y		y
REgression with Subsequent Independence Test (RESIT) (Peters et al., 2014)	y	y	y		y	y
Linear Non-Gaussian Acyclic Model (LiNGAM) (Shimizu et al., 2006)	y	y		y	y	y
Direct LiNGAM (Shimizu et al., 2011)	y	y		y	y	
Vector Autoregression Model LiNGAM (VAR-LiNGAM) (Hyvärinen et al., 2010)		y		y	y	y
Post-NonLinear (PNL) (Zhang & Hyvärinen, 2012)	y	y		y		y
Additive Noise Model (Hoyer et al., 2008)	y	y				y
Repetitive Causal Discovery (RCD) (Maeda & Shimizu, 2022)	y					y
Generalized Independent Noise (GIN) (Xie et al., 2020)	y					y

1. Independent and identically distributed random variables (I.I.D.).
2. Causal sufficiency, that is, no latent variables exist.
3. Gaussian: the variables have Gaussian distributions.
4. Markov condition: variables only depend on the parent variables and not on variables further up the "ancestral" chain.
5. Linearity: it is assumed that the functions of variables are linear. (this criterion only applies to FCMs),
6. non-deterministic polynomial-time hardness: the computation power or time needed to produce a result substantially increases as the number of variables grows. Those models developed with the aim of fast computation are NP-hardness-free.

In Table 1, the "y" cells refer to that the models require the criteria, whereas the blank cells indicate the criteria are not required. Table 1 is by no means an exhaustive list, but a collection of the algorithms that come from a review paper (Glymour et al., 2019). It is highlighted that commonly agreed-upon assumptions in the domain of causal discovery, such as acyclicity (i.e., causal structure can be represented by directed acyclic graph [DAG]) and faithfulness (i.e., all conditional independences in true underlying distribution are represented in DAG) are not in the selection criteria. Because the models share these features and do not differ in these regards. More detailed descriptions of these shared assumptions and DAG can be found in previous studies (Glymour et al., 2019; Spirtes et al., 2000). Apart from these assumptions, certain model variations are excluded. For constraint-based models, these variations include different conditional independence tests, while for score-based models, they involve different score functions and optimization functions. For example, the original GES (Chickering, 2002) is equipped with the score function of Bayesian information criterion and the local optimization of the greedy algorithm, and its variants employ generalized score (Huang et al., 2018), global optimization (Silander & Myllymaki, 2012), and shortest path (Yuan & Malone, 2013). These distinctions, however, are not criteria for selection, as they do not significantly expand model applications.

The research develops the causal discovery model to discern plausible causal relationships between the factors, offering a methodological advancement over conventional statistical approaches. The traditional methods widely adopted by road agencies fall short in distinguishing between mere correlation and genuine causation, a limitation that the proposed model aims to rectify. The approach determines whether traffic mix and overall volume directly cause raveling in a given field. By controlling for influencing factors unrelated to traffic, such as those tied to materials, structure, and weather, the research pinpoints the traffic cause of raveling, potentially enhancing raveling prediction and maintenance strategies.

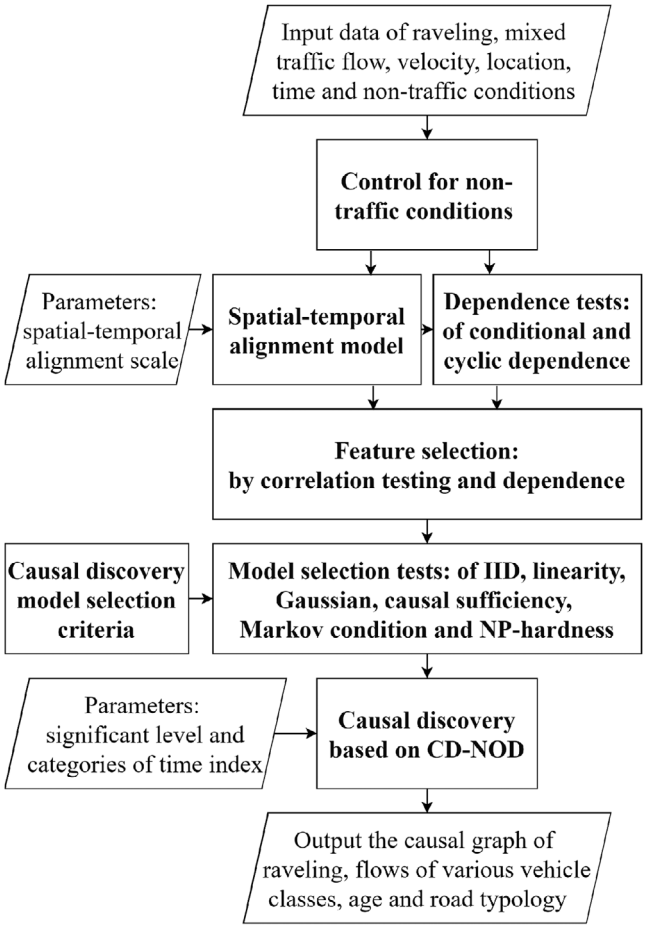

FIGURE 1 Methodological framework.

1.1 | Paper organization

This paper is organized to showcase the novel approach in analyzing the causal mechanisms between raveling and traffic: It begins with a comprehensive review of prominent causal discovery models, establishing a selection criteria tailored to address the specific needs of model application. This sets the stage for the methodological part of the study, where models and tests are meticulously detailed. Next, the paper delves into the data section, explaining the processes of data extraction, cleaning, and selection of the study areas. Following this, the results and discussions section interprets the findings and contextualizes them within the existing body of knowledge. Lastly, conclusions are presented.

2 | METHODOLOGY

This study presents a comprehensive framework to examine the causal relationship between raveling and mixed traffic, as depicted in Figure 1. The requisite data encompass variables such as raveling, mixed traffic flow, velocity, location, time, and nontraffic conditions, including but not limited to material composition, and structural and climatic factors.

The methodology begins with the control for nontraffic conditions, through the selection of pavements that exhibit uniform nontraffic attributes. The step is essential in causal inference as it mitigates the influence of potential confounders on the primary relationship being examined. The control for nontraffic conditions allows for the isolation of the traffic-related effects on raveling. This approach is particularly pertinent given that the real-world factors contributing to raveling are diverse and simultaneously present in road infrastructure.

With the varied spatial–temporal scales of input data sets in mind, the framework includes the spatial–temporal alignment model. This model is to tackle the challenge that arises from the heterogeneity of the spatial–temporal scales of raveling and traffic measurements, especially the considerable difference in their temporal resolutions (Wang et al., 2022). Traffic data are dynamic in the short term, which is usually collected per minute or hour. Raveling, indicating long-term road deterioration, is measured over seasons and years. The use of adaptive smoothing and alignment techniques are necessary to address this discrepancy (Wang et al., 2022), ensuring that the temporal and spatial dimensions of the data are compatible for analysis.

The alignment can introduce noise due to the loss of information from long-term aggregation and potential confounding variables over extended observation periods, which may impact the outcomes of causal discovery. To estimate these effects, (conditional) independence tests are conducted on high-resolution temporal variables. Additionally, cyclic dependence tests are applied to address the possibility of cyclic dependencies influencing causal discovery. The estimated dependence relationships are further considered to guide feature selection. Moreover, to identify traffic-related variables that may causally influence raveling, feature selection also incorporates domain knowledge and correlation testing.

Following feature selection, the methodological framework proceeds with the selection and application of the most appropriate causal discovery model. Model selection is based on the criteria in Table 1, guiding the choice of the model fitting the determined features. This research employs the Constraint-based causal Discovery model with heterogeneous/NOnstationary Data (CD-NOD) proposed by Huang et al. (2020), for its robustness in handling nonlinearity, non-Gaussianity, and nonstationary data. The justification for preferring CD-NOD over alternative models is detailed in Section 4.3, with the model parameters set according to significance levels and categories of time indices. Finally, this methodological

framework yields the causal graph that reveals the causal relationships between raveling, traffic flow for different vehicle classes, pavement age, and road type, offering a nuanced understanding of the interplay between these factors.

2.1 | Control for nontraffic conditions

In raveling studies, attention is mainly drawn to two primary phenomena. The recent article by Abouelsaad and White (2021) underscores these occurrences. The first phenomenon involves the loss of fine aggregate, binder, and/or filler, a process commonly attributed to aging and referred to as fretting. The second phenomenon is identified by the loss of coarse aggregate, which predominantly occurs due to the action of vehicle tires. Fretting is seen as the less severe version of raveling, whereas the loss of coarse aggregate is considered more severe (Abouelsaad & White, 2021; Nicholls et al., 2015). It is obvious that in fields these phenomena coexist, involving both traffic and nontraffic factors. The nontraffic factors influencing raveling include oxidative aging and environmental conditions, such as ultraviolet radiation (Abouelsaad & White, 2020, 2021; Hagos, 2008), rainfall (Abouelsaad & White, 2020, 2021; Kringos & Scarpas, 2005; Nicholls et al., 2015; Thube et al., 2006), and temperature variations (Abouelsaad & White, 2021; De Visscher & Vanelstraete, 2017; Huurman et al., 2010; Mo et al., 2010). Material and structure properties also play a vital role, encompassing the adhesion properties of the bitumen-aggregate system, mixture gradation, and binder film thickness (Abouelsaad & White, 2021; Nicholls et al., 2015; Voskuilen et al., 2004).

To focus on the effects of mixed highway traffic, pavements exhibiting similar nontraffic conditions are selected. Considering the practical availability of the aforementioned nontraffic variables, the requirements for the selected pavements are as follows:

(1) The chosen pavements should exhibit uniformity in both structure and material composition.
(2) The configurations of the selected pavements should remain consistent throughout the observational period.
(3) The selected pavements should undergo identical construction and maintenance processes to have comparable levels of quality. When construction, maintenance details, and quality assessments are challenging to access, these criteria can be interpreted as selecting pavements constructed and maintained within the same project time frame, under the assumption that these pavements would have experienced similar procedures and met the same quality standards.
(4) The chosen pavements should share similar exposure to weather conditions. These include factors such as temperature fluctuations, sunlight intensity, moisture, precipitation, and frost.

This approach minimizes the effects of nontraffic factors, thereby yielding a more credible causal relationship between traffic factors and raveling.

2.2 | Spatial–temporal alignment model

The spatial–temporal alignment model, depicted in Figure 1, is utilized to address the issues arising from the heterogeneous spatial–temporal scales of raveling and traffic measurements. The authors have previously proposed a method to align data of raveling and traffic to identical space and time units (Wang et al., 2022). These units correspond to the larger resolutions of spatial and time measurements. To provide a concise overview and spare readers from referencing the original work, this paper presents the main equations, encompassing adaptive smoothing method (ASM) introduced by Treiber and Helbing (2003), temporal alignment, and spatial alignment. For an in-depth explanation, readers are referred to the prior study (Wang et al., 2022).

The main equations, including nonlinear weight function, traffic count, and mean velocity calculations, are central to ASM. Equation (1) formulates the continuous traffic flow by integrating congested and free-flow conditions through an adaptive weighting function. Traffic counts for various vehicle types and their mean velocities are then combined as per Equations (2) and (3). These counts are further detailed in matrix form in Equation (4), while the velocity data are represented in Equation (5).

$$\dot{q}(\eta, t^q) = \omega(q^{\text{cong}}, q^{\text{free}}) q^{\text{cong}}(q(s,t), v(s,t), \eta, t^q) + (1 - \omega(q^{\text{cong}}, q^{\text{free}})) q^{\text{free}}(q(s,t), v(s,t), \eta, t^q) \tag{1}$$

$$q(s,t) = \sum_{k=1}^{n} q^k(s,t) \tag{2}$$

$$v(s,t) = \frac{\sum_{k=1}^{n} q^k(s,t) v^k(s,t)}{\sum_{k=1}^{n} q^k(s,t)} \tag{3}$$

$$\left\{q^k(s,t)\right\} = \mathbf{Q}(s,t) \tag{4}$$

$$\left\{v^k(s,t)\right\} = \mathbf{V}(s,t) \tag{5}$$

The notation used throughout the above equations is consistent. The symbol $\dot{q}$ denotes continuous traffic flow with respect to space η and time t^q. ω is the adaptive weight

factor, q^{cong} and q^{free} are the spatial–temporal lowpass filters applied specifically to congested and free traffic, respectively, q denotes the total number of the vehicles at the location s during the time t, and v indicates the mean velocity of q. The weight $\omega \in [0,1]$ depends upon the filters q^{cong} and q^{free}, according to the traffic theory that a wave propagates upstream at a constant speed when traffic is congested, and downstream at free speed in free traffic. Both q^{cong} and q^{free} are derived from q and its mean speed v, and the alignment scale of space η and time t^q. η is uniformly distributed, ranging from $s+\Delta s, s+2\Delta s, s+3\Delta s$, and so forth. t^q is also uniformly distributed, ranging from $t+\Delta t, t+2\Delta t, t+3\Delta t$, and so forth. Δs and Δt are the spatial and temporal resolutions, respectively. q^{k} represents the vehicle counts of a kind k at the location s during the time t. Usually, traffic data collection systems define s and t. v^{k} is the mean speed of a kind k of vehicles q^{k} counted at the location s during the time t, and n denotes the number of vehicle types. The matrices $\mathbf{Q}$ and $\mathbf{V}$ represent the vehicle counts and velocities, respectively, for various vehicle types recorded at location s and time t. These matrices can be exported from traffic data collection systems.

A set of equations are used to align temporal scales of data. The temporal alignment model aggregates short-time traffic flow into the target time resolution τ, which corresponds to raveling progression. The primary Equation (6) computes the cumulative traffic flow $\ddot{q}$, consisting of all the observed flows within the raveling time and an estimate of flows without observation during this time. Equation (7) calculates the average flow $\bar{q}$, representing the mean traffic volume over a section during the observed period.

$$\ddot{q}(\eta,\tau) = \sum_{t^q=t^q_{\min}}^{\min(t^q_{\max},\tau)} \dot{q}(\eta,t^q) + \bar{q}(\eta)\left(\tau - \sum_{t^q_{\min}}^{\min(t^q_{\max},\tau)} t^q\right) \quad (6)$$

$$\bar{q}(\eta) = \frac{\sum_{t^q=t^q_{\min}}^{\mathrm{T}} \dot{q}(\eta,t^q)}{\sum_{t^q=t^q_{\min}}^{\mathrm{T}} t^q},\ \mathrm{T} \in (t^q_{\min}, t^q_{\max}] \quad (7)$$

where $\dot{q}$ exported from ASM, denotes continuous traffic flow regarding space η and time t^q, $t^q_{\min}$ and $t^q_{\max}$ are the first and last measurements of t^q, and $\bar{q}$ is an average flow traversing on space η of a given period T within the entire observation from $t^q_{\min}$ to $t^q_{\max}$.

The subsequent Equation (8) determines the cumulative flow for each vehicle type $\ddot{q}^{\text{k}}$, assuming that the ratio of a specific vehicle type to the total flow in a short but sufficient observation period stays consistent with that of a long term. This assumption is validated by the research data and by empirical knowledge of vehicle distribution throughout road typology. The ratio $\ddot{\text{W}}^{\text{k}}$ in Equation (9) specifies the proportion of each vehicle type, calculated by the p-th percentile function of the observed ratios $\mathbf{W}^{\text{k}}$.

$$\ddot{q}^k(\eta,\tau) = \ddot{q}(\eta,\tau)\ddot{W}^k(\eta,\tau) \quad (8)$$

$$\ddot{W}^k(\eta,\tau) = f(\mathbf{W}^k, p),\ (\forall s \in \eta, \forall t \in \tau, w^k(s,t) \in \mathbf{W}^k) \quad (9)$$

where $\ddot{q}$ resulting from the temporal alignment model represents the cumulative flow of all vehicles at the space η and τ, and $\ddot{W}^{\text{k}}$ denotes a set of the vehicle ratios of vehicle type k corresponding to η and τ. f is the p-th percentile function, and $\mathbf{W}^{\text{k}}$ is a set of vehicle ratios w^{k} at locations s within η and time t within τ. The computation of w^{k} is to divide q^{k} by q. It is noted that not all the data stations supporting this research are upgraded to include the function of exporting q^{k}. In such cases, $\ddot{W}^{\text{k}}$ is estimated by multiplying annual vehicle ratios with a lane vehicle distribution index derived from the data (Van Beinum et al., 2018).

The spatial alignment model is employed to convert raveling of a high spatial dimension to the resolution η in agreement with the traffic variables. The representative value of raveling r̈ is computed using the formula given in Equation (10) by a sample of raveling measured within the defined space η.

$$\ddot{r}(\eta,\tau) = f(\mathbf{R}, p)\ (\forall s^r \in \eta, r(s^r,\tau) \in \mathbf{R}) \quad (10)$$

where f is the p-th percentile function of a sample according to descriptive statistics. In the computation of r̈, the sample $\mathbf{R}$ is a set of raveling measurements r at time τ and locations s^r, where s^r are within space η.

2.3 | Dependence tests

The alignment potentially influences the outcomes of causal discovery due to the loss of information from long-term aggregation and the emergence of confounding variables over extended observation periods, introducing data noise. To estimate the potential impacts and guide the selection and redefinition of features in next step, dependence tests are undertaken in two main steps. Initially, influences of temporal aggregation on the variables are estimated by (conditional) independence tests. Second, the extent of dependencies, particularly cyclic ones, among traffic variables are evaluated across varied time frames through cyclic dependence tests.

2.3.1 | (Conditional) independence tests

(Conditional) independence tests are conducted of the variables with high dimensions, pertaining to traffic flows.

(Conditional) independence is set as the null hypothesis H_0 as shown in Equation (11), and the alternative hypothesis H_A in Equation (12) indicates (conditional) dependence.

$$\mathrm{H}_0 : x \perp\!\!\!\perp y \mid z, \quad x \neq y \neq z \tag{11}$$

$$\mathrm{H}_\mathrm{A} : x \not\perp\!\!\!\perp y \mid z, \quad x \neq y \neq z \tag{12}$$

where x, y denote variables from the flow matrix $\mathbf{Q}$ or the cumulative traffic flow $\ddot{q}$, and z denotes a variable from $\mathbf{Q}$ or $\ddot{q}$ in a conditional independence test or an empty set in an independence test.

To assess the strength of the evidence against the null or alternative hypothesis, a significance level is set (typically 5% in this research), which is the threshold for determining statistical significance. This level, chosen based on conventional standards for hypothesis testing, is then compared with the p-value of the tests. The p-value is derived using the kernel conditional independence (KCI) test, as introduced by Zhang et al. (2011). The KCI test is selected for its proficiency in managing high-dimensional data and its nonreliance on any assumptions regarding the underlying data distributions. It functions by calculating the kernels of data. The test statistic is calculated using the trace of the matrix product by normalizing and multiplying the kernels, and the p-value of the test is determined by a gamma approximation according to the mean and variance of the independent sample kernel values (Zhang et al., 2011).

2.3.2 | Cyclic dependence tests

Cyclic dependence tests are conducted to detect dependencies that form cycles, which could potentially distort the outcomes of causal discovery. Identifying such recurring dependencies is crucial for accurate causal analysis. For these tests, the PC algorithm is utilized, as introduced by Spirtes et al. (2000), which is grounded on hypothesis testing exemplified by Equations (11) and (12). Under the premise that the data satisfy the Markov condition and faithfulness, these tests enable the construction of an undirected independence graph **G*** and its DAG **DAG*** that represents the interdependencies of traffic flow as:

(1) if $x \perp\!\!\!\perp y \mid z$, x and y are separated by z in **G***,
(2) if $y \notin PA_x, \forall e_{xy}, e_{yz} \in \mathbf{G}^*, d_{xy}, d_{zy} \in \mathbf{DAG}^*$,

where x, y, z are nonidentical variables from the flow matrix $\mathbf{Q}$ or the cumulative traffic flow $\ddot{q}$, PA_x denotes the parent variable(s) of x in the undirected independence graph $\mathbf{G}*$, e_{xy} and e_{yz} indicate the edges $x - y$ and $y - z$, respectively, and d_{xy} and d_{zy} are the directed edges $x \rightarrow y$ and $z \rightarrow y$, respectively. The edge such as e_{xy} illustrates the existence of a causal relationship between variables, and the directed edge d_{xy} shows x is a direct cause of y. It is highlighted that **G*** and **DAG*** have no unique solution if a Markov equivalence class exists.

2.4 | Feature selection

The process of feature selection serves as a critical step in the methodological framework in Figure 1. The primary objective of feature selection in this study is to identify variables that demonstrate a potential causal relationship with the outcome variable. Such identification aids in uncovering the fundamental causes of the phenomena under study. The secondary objective is to enhance the overall performance of the causal discovery process by focusing on the features that exhibit robustness against both temporal aggregation and cyclic dependencies. To this end, this section describes the step-by-step process involved in selecting features for the causal discovery model. The proper consideration of domain knowledge is essential in identifying and classifying the pertinent features linked to traffic and pavement raveling in the initial phase of the feature selection process. Correlation testing is then performed to further identify the variables showing significant correlation with raveling. The final stage comprises a thorough examination of dependent tests to detect any potential distortion due to the cumulative nature of this study.

2.4.1 | Domain knowledge

The initial phase of the feature selection process begins with an appropriate consideration of domain knowledge, which helps discerning and categorizing the relevant features related to raveling and traffic factors. This step is guided by established theoretical frameworks of statistical causal inference within the literature (Guo et al., 2020; Peng & Xu, 2023), which delineate the key variables of a causal model into treatments, outcomes, and confounders. A treatment is assumed in a studied question to cause the outcome, and a confounder refers to the confounding variable that causally influences both the treatment and the outcome (Guo et al., 2020; Peng & Xu, 2023). Within this study framework, traffic flow is defined as the treatment, while pavement raveling is characterized as the outcome. A significant step is to identify potential confounders that can influence both the treatment and the outcome.

The identification of confounders is according to domain knowledge on raveling, as detailed in the recent review articles (Abouelsaad & White, 2021; Nicholls et al.,

2015). A combination of traffic loads and aging is expected to impact most raveling cases (Abouelsaad & White, 2021), leading to the identification of vehicle classes and pavement age as confounders. Therefore, a preliminary set of features is determined: raveling, traffic flow, age, and vehicle classes.

2.4.2 | Correlation testing

Correlation testing is performed to ensure that the features used in the model have strong correlation with raveling. Spearman's correlation coefficient is utilized to assess the strength of the relationships among the preliminary features, because the coefficient is well-suited to both linear and nonlinear relationships, given its rank-based nature.

2.4.3 | Analysis of dependence tests

A thorough analysis of the dependence tests is to verify if the study's cumulative nature might have introduced any distortions. Given the relationships of dependence and conditional dependence between traffic variables across different levels of temporal aggregation, the focus is on those features exhibiting consistent dependence relationships, regardless of time aggregation. Two scenarios are of particular interest. In one scenario, short-term traffic variables may show dependence relationships that become insignificant as aggregation extends toward long-term pavement raveling progress. This phenomenon could potentially be associated with information loss due to temporal aggregation. In another scenario, relationships between certain short-term traffic variables, initially independent, may display dependence when aggregated over the long term. This may be indicative of confounding variables surfacing due to extended observation periods.

In terms of cyclic dependence, the ideal case is to select features with none cyclic dependencies. However, in instances where such dependence is inevitably present, as is in this study, the strategy involves redefining the features to reduce cyclic dependencies. For example, the analysis uncovers cyclic dependencies among various vehicle classes within traffic flow, requiring a decrease in the number of considered vehicle classes to limit these cyclic dependencies.

2.5 | Model selection tests

This section presents the tests in relation to the proposed model selection criteria in Table 1. The criteria and the corresponding tests are designed for identifying the most appropriate causal discovery models for a particular scenario. Via the method, the study pinpoints the most suitable model for analyzing observational data concerning ravel and traffic. The tests leverage correlation analyses and normality assessments. The tests can confirm whether a case study satisfies the conditions of I.I.D., Gaussian distribution, and linearity. Regarding the other three criteria (i.e., causal sufficiency, Markov condition, and NP-hardness), evaluations and certain assumptions are formulated based on existing knowledge.

2.5.1 | Independent and identically distributed random variables

Two statistical tests are proposed to help differentiate between I.I.D. and time-series characteristics: the correlation between two consecutive measurements of a feature, and the correlation between measurements of a feature and its time sequence of measurements. Spearman's correlation coefficient is used for these tests because it is not confined to linear relationships. If the correlations are weak, the case study is assumed to fulfill the I.I.D. requirement. Identifying I.I.D. characteristics is usually also feasible based on the data generation process.

2.5.2 | Linearity

R^2 of linear regression is a common measure of linearity. To identify if a study meets the criterion of linearity, the linear correlation of every dependent variable and the independent variable is tested. When all the R^2 are (roughly) equal to 1, the study case has the characteristics of linearity.

2.5.3 | Gaussian

A normality test is to identify if a sample follows Gaussian distribution. Skewness and kurtosis are the widely used indicators. According to the Fisher–Pearson definition, skewness S is formulated as the ratio of the third cumulant to the 1.5th power of the second cumulant. The kurtosis K (based on Fisher's formula) is the fourth central moment divided by the square of the variance and subtracted by three. A standard Gaussian distribution has both skewness and kurtosis of zero, and the closer the values are to zero, the more analogous the variable is to Gaussian distribution. Moreover, the condition of Gaussian is met when all the features in a study case are (nearly) normally distributed.

2.5.4 | Causal sufficiency

In terms of causal sufficiency, no latent variable is the key to satisfy this condition. According to the recent review papers of raveling (Abouelsaad & White, 2021; Nicholls et al., 2015), a few traffic factors are identified to be related to raveling. This research supposed the knowledge is sufficient, and investigates these factors as well as the confounders.

2.5.5 | Markov condition

The Markov condition assumes that the state of an entity at a given moment depends only on its previous stage, disregarding states from further in the past. The assumption simplifies the analysis of time-series data by mitigating the complexity involved in modeling temporal dependencies. The assumption aligns with modeling pavement raveling, as it mirrors the actual deterioration process: The extent of future raveling is directly influenced by the current condition, which itself evolved from previous states. The assumption especially fits roads with maintenance histories. Because after maintenance, the future condition depends on the new state rather than the historical condition sequence.

Building on the concept, causal discovery models operate on the premise that variables acting as children and ancestors are independent, given the variables identified as parents. The models equipped with Markov condition construct a unique DAG when the assumption holds (Spirtes et al., 2000). Thus, the suitability of a case for the Markov condition can also be validated from the model outcomes. The model selection for this study is made based on the potential to choose the models that have the restriction of Markov condition.

2.5.6 | NP-hardness

The NP-hardness problem indicates that finding a solution becomes quickly infeasible as the number of variables increases. Koivisto and Sood (2004) and Silander and Myllymaki (2012) have presented the feasible cases with 26 and 33 variables. In practice, the influence of an NP-hardness problem is related to computation power, which is not a significant restriction with regard to the workstations supporting this research.

2.6 | Causal discovery based on CD-NOD

Based on the model selection test previously discussed, the CD-NOD proposed by Huang et al. (2020) is chosen as the suitable model to uncover the causal relationships between traffic, raveling, and potential confounding variables. It has the capacity to analyze nonstationary data where the underlying process changes across domains and/or over time. Emerging from the foundational structure of the PC algorithm (elaborated in Section 2.3), this model is designed to address changes in causal structures over domain and/or time index. It functions based on the model in Equation (13).

$$v_i = f_i\left(PA_{v_i},\ g^i(\tau^*),\ \theta_i(\tau^*),\ \epsilon_i\right) \tag{13}$$

where for every variable v_i, it is formulated as a function f_i of its parent variable(s) PA_{v_i}, confounder(s) g^i, parameters θ_i relating to time index τ^*, and a disturbance term ϵ_i. In this research, v_i denotes the flow of each vehicle type and the traveling resulting from the spatial–temporal alignment model. τ^* is the time index, obtained by discretizing time τ into given levels. For instance, given a 5-year interval, a road section in service for less than 5 years receives a label of 1. If the section has been in service for more than 5 but less than 10 years, it is labeled as 2, and so forth.

This research implements three steps of the CD-NOD developed by Huang et al. (2020). At first, it identifies the (conditional) independence among variables and between variables and confounders. As expressed in Equation (11) and Equation (12), it conducts hypothesis tests for conditional independence. In these tests, the variables x, y, z refer to v_i; or one of the variables refer to η^* or τ^* and the other two denote v_i. Although the original CD-NOD established by Huang et al. (2020) does not specify a method of such a test, this study uses KCI (elaborated in Section 2.3) because it is advanced with regard to handling dimensionality and free from data distribution assumptions. Second, it identifies the undirected independence graph **G**. Section 2.3 describes how to illustrate an undirected independence graph based on (in)dependence relationships. CD-NOD adjusts the corresponding edges in the graph when the changes in distribution shifts over domain and time index is detected according to Equation (13). The last step is to estimate the causal directions of **G**. Grounded in the PC algorithm, it determines causal directions based on the detected graph and the conditional independence, as elaborated in Section 2.3. Additionally, CD-NOD constructs the changing causal module(s) in the existence of domain and/or time indexes based on Hilbert Schmidt Independence Criterion (HSIC) introduced by Gretton et al. (2007).

2.7 | Sensitivity analysis

The purpose of the sensitivity analysis in this study is to ascertain how specific model inputs and parameters

TABLE 2 Summary of data used in the analysis.

Data	Source	Data type	Sampling frequency	Unit	Features
Raveling	(Van Aalst et al., 2023)	Time series	Annually	%	The ratio of the total lost aggregate area to the measurement area per lane
Road	(Rijkswaterstaat, 2016)	Single data	Once	-	Construction and maintenance date, structure, material, geographical location, number of lanes, and lane type
Traffic	(NDW, 2023)	Time series	Minutely	Veh/h, km/h	Flow, speed, and vehicle types
Vehicle	Van Beinum et al. (2018)	Time series	Minutely	m	Vehicle positions and lengths

influence the results. The method used to assess the effect of each road section is a comparative analysis of the causal graphs obtained when a specific road section is included versus when it is excluded. The similarity between the graphs is estimated based on the cosine similarity of their corresponding matrices. The sensitivity of the model to the duration of the data collection period is also assessed by a comparative study. It involves comparing the cosine similarity of the causal graphs based on the data with varying observation duration, ranging from 3 to 9 years. The sensitivity analysis related to the CD-NOD parameter addresses the time interval, which determines the time index τ^*. The sensitivity analysis employs intervals of 1, 3, and 5 years to encompass the spectrum from short-term changes through medium-term developments to long-term trends.

3 | DATA

The research incorporates a wide range of raveling, road, traffic, and vehicle data, shown in Table 2. The data used in this study correspond to the roadways and time periods where raveling data are available. In the supporting database, the raveling data majorly come from the PA networks given that PA is vulnerability to raveling. As a result, the study areas are all PA pavements. The data corresponding to the roadways and raveling measurement periods are extracted from the other databases to ensure consistency and validity. The extraction process generates 4302 raveling records, 492 road data entries, 1,895,287,910 traffic data entries, and 30,975 vehicle data entries.

The pavement selection is made to align with the defined requirements in Section 2.1. Pavements are grouped into subsets based on the same structural and material types at the same construction time. To maintain consistency in configurations, the pavements with geographical location changes are excluded. To meet the condition of consistent environmental factors, the Dutch climate database (Institute, 2023) is consulted to ensure that the selected pavements are within one climatic zone. The data include five areas of Dutch highways shown in Figure 2,

FIGURE 2 Five study areas on Dutch highways. *Source*: Background map source: Rijkswaterstaat (2016).

representing a total distance of 18.6 km and encompassing information gathered from 2012 to 2020. It is noted that different areas are considered independently in the model.

It is highlighted that while the methodology is applied to the PA sites in the study, it is not limited to any particular material type. The application requirement regarding materials is that the chosen pavements exhibit uniformity in both structure and material composition, as the first requirement in Section 2.1.

3.1 | Data cleaning

A data cleaning process is implemented to guarantee the data consistency, validity, and relevance to the

proposed causal discovery model. This process is pivotal in fostering a robust analysis of the intricate relationship between mixed highway traffic and raveling. The procedure unfolds in four distinct steps, which are delineated in the succeeding subsections.

3.1.1 | Data formatting

All data are processed to maintain consistent formatting, especially the date format across various data sets. This procedural uniformity enhances the ease of data manipulation and integrity throughout the research.

3.1.2 | Handling missing data, duplicate data, and outliers

Data irregularities can impede the accuracy of analyses. The road data have some missing values, particularly in road construction and maintenance records and road surface material data, with 49 missing entries in the total 492 data points. Additionally, the traffic data contain minor missing values—144 days missing in the 9-year total data. The data are repaired by referring to more original data, and as a result, 49 entries are replenished with pertinent information. The missing values of the traffic data are filled by referencing the data of the previous week if the duration of gaps are less than a week. Longer missing periods are addressed using a defined process encapsulated in Equation (6). Another problem is a small number of duplicate values, predominantly in traffic data. These duplications representing less than 0.1% of the total data volume are removed. Outliers are detected in both road and traffic data, especially where they do not conform to reasonable maintenance times and speed values. These are treated appropriately, with removal for road data and correction of traffic data.

3.1.3 | Data integration

Due to the multisource nature of the data, it is necessary to amalgamate data from various sources into a unified data set. An approach is employed where road segment information is converted into BPS coding (WegenWiki, 2023), a concatenation of road, direction, distance, lane, and time fields. Through the method, each data set can be accurately matched according to the corresponding fields, enabling effective data synthesis across multiple databases.

3.1.4 | Data transformation

To better fit the proposed causal discovery model, specific transformations are performed. Measurement dates of raveling and traffic data are converted to a corresponding time variable, considering the specific time period for each variable. For instance, the time period of raveling takes into account the operation start of the road until the measurement, incorporating road construction and maintenance records, and the time corresponding to traffic flow is the interval between two successive measurements by an inductive-loop traffic detector. Another transformation converts the traffic flow per vehicle composition ratio into the traffic flow per lane per vehicle composition ratio, thus giving a lane-specific vehicle distribution that aligns with the perceived traffic pattern on Dutch highways.

4 | RESULTS AND DISCUSSION

The section presents the results from the main steps examining the causal relationship between mixed highway traffic and raveling. This investigation is based on 9-year field measurements of the selected five PA highway sites.

4.1 | Spatial–temporal alignment of raveling and traffic data

The spatial–temporal alignment model is used to reconcile the considerable discrepancies in the spatial–temporal units of raveling and traffic data. Raveling measurements, being one-square-meter zonal data with cumulative time series, are transformed to a spatial scale of 100-m sections per lane, maintaining its temporal scale. Traffic data, originally point-based and noncumulative time series, are adjusted to the same spatial scale while matching the temporal scale of raveling data. The choice of 100-m segments is because the corresponding pavement management systems commonly use hectometer (100-m) units, and traffic variability is relatively low at smaller spatial scales. A detailed discussion of alignment scales is available in the previous work (Wang et al., 2022).

Figure 3 demonstrates the transformation of the raw raveling and traffic measurements into the aligned data. Raw raveling data and their proceeded data have similar patterns, as Figure 3a and Figure 3b indicate. The traffic data transformation from Figure 3c to Figure 3d shows that, at the raveling progression time scale, traffic characteristics become simpler. These characteristics include accumulation increases with service time, and traffic variations mainly correspond to distinct road configurations, such as additional lanes for exiting or merging into the main traffic stream. Other fundamental characteristics, such as periodicity, are generally observed at shorter scales, making it challenging to estimate the relationship between them and road deterioration. The method necessitates that raw traffic data incorporate information both upstream and

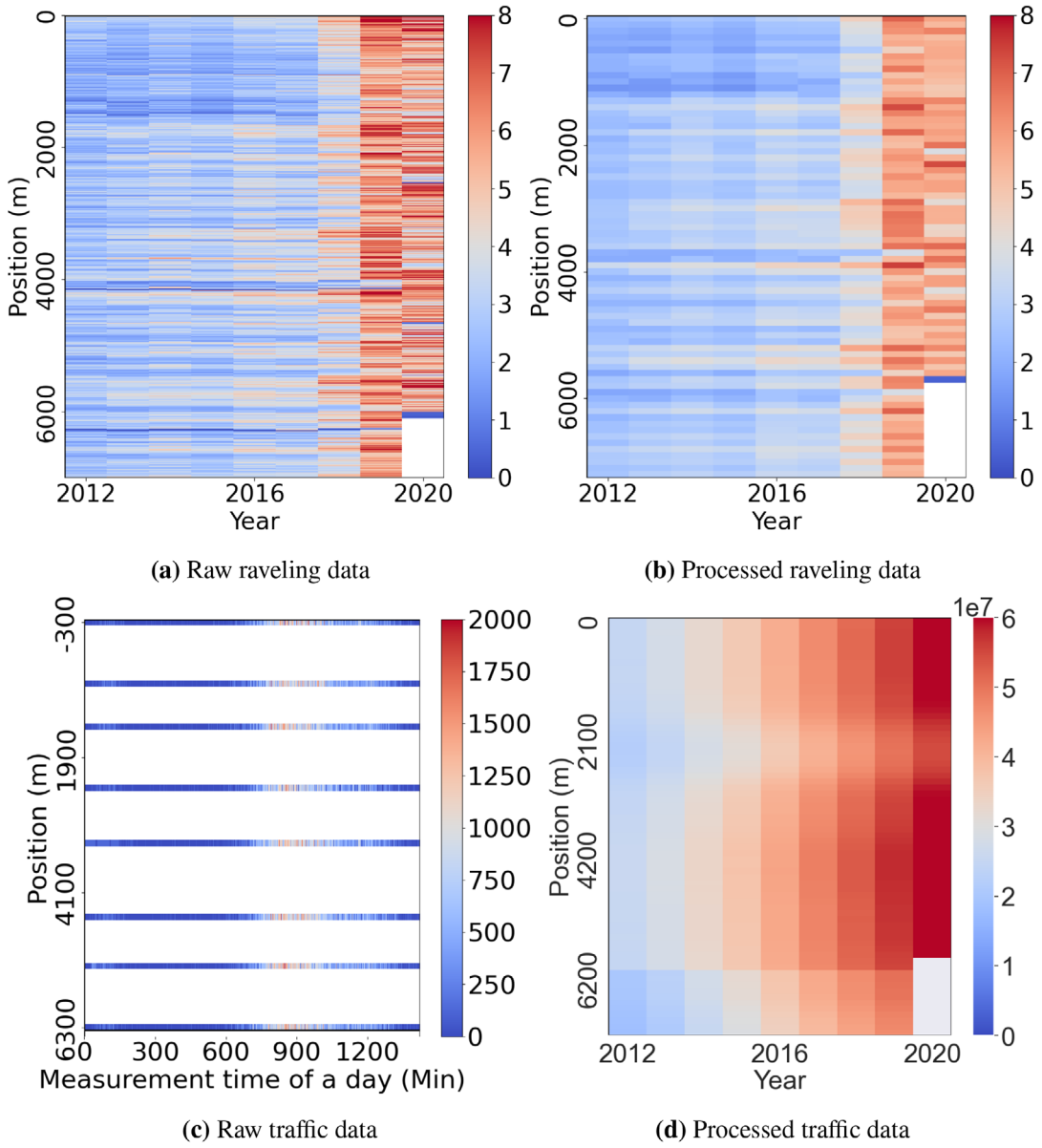

FIGURE 3 Raw data (a, c) and processed (spatial–temporally aligned and smoothed) data (b, d) pertaining to raveling and traffic. An example of the overtaking lane in Area I.

downstream of a study area (e.g., the raw data are from six sensors inside Area I and two sensors at 300-m upstream and downstream, respectively, in Figure 3c coordinates −300 m and 6300 m).

4.2 | Correlation and dependence relationships between raveling and traffic variables

The variables correlated with raveling that exhibit minimal cyclic dependencies, and thus are less influenced by temporal aggregation are selected. Table 3 shows nonlinear correlations between raveling and various traffic quantities calculated at the previous step. Strong correlations (coefficients greater than.7) emerge when data are lane-categorized and primarily on the main lanes. Most correlations for the overtaking lanes are moderate, between.20 and.65 suggesting influence of driving behavior, the main difference between an overtaking and main lane. However, data for long-term behavior across lanes are not available for this study. To address this, the categorization of lanes (i.e., overtaking or main through lanes) is considered, which determines driving behavior. Variables correlating with raveling, such as road age, total traffic volume, individual vehicle class volumes, and lane categorization, are included in the subsequent analysis.

TABLE 3 Spearman correlation coefficients and *p*-values for the relationships between raveling and traffic of all study areas.

Study area	Lane	Total traffic		Total cars		Total trucks		Total short trucks		Total long trucks		Road age	
		ρ	*p*-value	ρ	*p*-value	ρ	*p*-value	ρ	*p*-value	ρ	*p*-value	ρ	*p*-value
I	Overtaking	.77	<.001	.77	<.001	.78	<.001	.78	<.001	.78	<.001	.78	<.001
I	Main	.80	<.001	.80	<.001	.80	<.001	.80	<.001	.80	<.001	.81	<.001
II	Overtaking	.21	.002	.21	.002	.23	<.001	.20	.002	.28	<.001	.50	<.001
II	Main	.69	<.001	.68	<.001	.71	<.001	.70	<.001	.71	<.001	.70	<.001
III	Overtaking	.60	<.001	.60	<.001	.67	<.001	.67	<.001	.59	<.001	.81	<.001
III	Left main	.84	<.001	.84	<.001	.86	<.001	.87	<.001	.84	<.001	.82	<.001
III	Right main	.54	<.001	.54	<.001	.54	<.001	.54	<.001	.54	<.001	.53	<.001
IV	Overtaking	.63	<.001	.63	<.001	.62	<.001	.62	<.001	.62	<.001	.65	<.001
IV	Left main	.36	<.001	.37	<.001	.34	<.001	.34	<.001	.35	<.001	.38	<.001
IV	Right main	.08	.282	.08	.271	.11	.150	.09	.227	.12	.104	.08	.276
V	Overtaking	.27	.012	.27	.012	.28	.010	.28	.010	.27	.013	.21	.055
V	Left main	.61	<.001	.61	<.001	.59	<.001	.59	<.001	.59	<.001	.63	<.001
V	Right main	.85	<.001	.85	<.001	.84	<.001	.84	<.001	.84	<.001	.88	<.001
I	All	.42	<.001	.50	<.001	.02	.499	.31	<.001	.01	.641	.75	<.001
II	All	.14	.007	.19	<.001	−.05	.280	.05	.308	−.04	.453	.56	<.001
III	All	.34	<.001	.36	<.001	.04	.314	.27	<.001	−.02	.640	.77	<.001
IV	All	−.15	<.001	−.11	.001	−.32	<.001	−.22	<.001	−.38	<.001	.56	<.001
V	All	.53	<.001	.53	<.001	.27	<.001	.43	<.001	.12	.031	.59	<.001

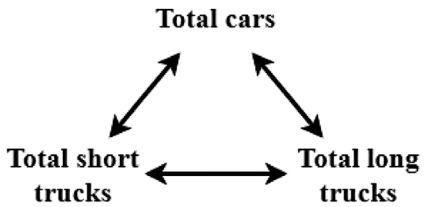

FIGURE 4 Cyclic dependencies of vehicle types in the study areas using the Peter–Clark (PC) algorithm.

Followed by correlation tests, conditional independence tests are performed for different vehicle types and cumulative time periods. The results consistently indicate an inherent dependence between the quantity of one type of vehicle within the traffic flow and that of other vehicle types. This holds true irrespective of investigating hourly, daily, weekly, monthly traffic patterns, or even fluctuations between peak and off-peak traffic periods. Despite the aggregation of traffic flow necessitated by the nature of raveling development, such accumulation does not alter the independent and conditionally independent relationships among the traffic volumes. These observed dependence relationships are cyclic. They interlink various categories of vehicles within the traffic flow, as shown in Figure 4.

This is supported by traffic flow theory that the presence of multiple factors affects the relationship between different vehicle classes. Factors that attribute to the whole traffic volume as well as each vehicle classes include traffic demand, time of day, and economic conditions. Besides, the influencing factors of local traffic patterns act as confounders in the dynamic interplay among vehicle classes, like weather, accidents, road works, and congestion. According to traffic flow theory, these codeterminants and influencers are the underlying causes of the observed dependencies. However, it is challenging to collect sufficient information to capture these confounders interwoven within multiyear accumulations to achieve the ideal feature selection free of cyclic dependencies. Nevertheless, to obtain a robust causal discovery outcome, it is imperative to curtail the complexity of cyclic dependencies, which necessitates considering a reduced selection of vehicle classes. Thus, the selected features are road age, total volumes of cars and trucks, and lane categorization.

4.3 | Selection of the CD-NOD model: Tackling nonlinearity, non-Gaussianity, and time-series data characteristics

According to the proposed model selection criteria, the characteristics of the selected features are estimated. Raveling, traffic, and road age exhibit time-series properties. Because traffic manifests a periodic pattern, and both the amount of dislodged aggregates on the road and its age accrue over time. This accumulation suggests a temporal dependence, whereby subsequent observations are

influenced by prior ones—a perspective substantiated by the tests. Throughout the various study areas, a marked correlation is observed among these variables in tandem with the sequence of their recordings, particularly in the linkage between successive observations. Given the data attributes, the preferred causal discovery models should eschew I.I.D. assumption. It is highlighted that the tests utilized in this study violate I.I.D. assumption by detecting time-series characteristics. While the data indeed exhibit these characteristics, leading to the choice of a model that does not rely on this assumption, the absence of such characteristics does not always imply that the I.I.D. criteria are met. Other scenarios might necessitate rigorous tests like independence and comparative distribution tests to affirm I.I.D. assumption.

Furthermore, the data exhibit non-Gaussian characteristics. Raveling and traffic data across all study areas display a positive skewness and tend toward platykurtic distributions. On the other hand, road age showcases kurtosis, while lane categorization follows a uniform distribution. The relationships between raveling and other features lack linearity, as evidenced by the R^2 values registering below.54 across all study areas. Only in Area I, a linear relationship—signified by an R^2 exceeding.8—is discerned between car flow and age. Therefore, the selected model ought to be able to accommodate non-Gaussian variables and nonlinear relationships.

The causal discovery model of CD-NOD is selected. It specifically deals with time-series data, non-Gaussian data, and nonlinear relationships by employing two approaches. One is integrating a time index into the identification of changing causal modules and the estimation of causal structure, and the other is developing an enhanced constraint-based and nonparametric method to perform causal discovery.

The characteristics of the observed data highlight the potential for enhancing raveling prediction and pavement management through models capable of addressing nonlinear and non-Gaussian data. Given the time-series nature of raveling and traffic characteristics, incorporating a systematic inspection regimen within pavement management practices is recommended. The strategy can make the comprehensive monitoring of traffic influence and pavement wear over extended periods, facilitating the development of proactive and informed maintenance strategies.

4.4 | Causal relationships between raveling and traffic variables

The causal graphs resulting from the CD-NOD model, such as Figure 5 and Figure 6, represent the causal relationships between mixed traffic flows and raveling, as identified

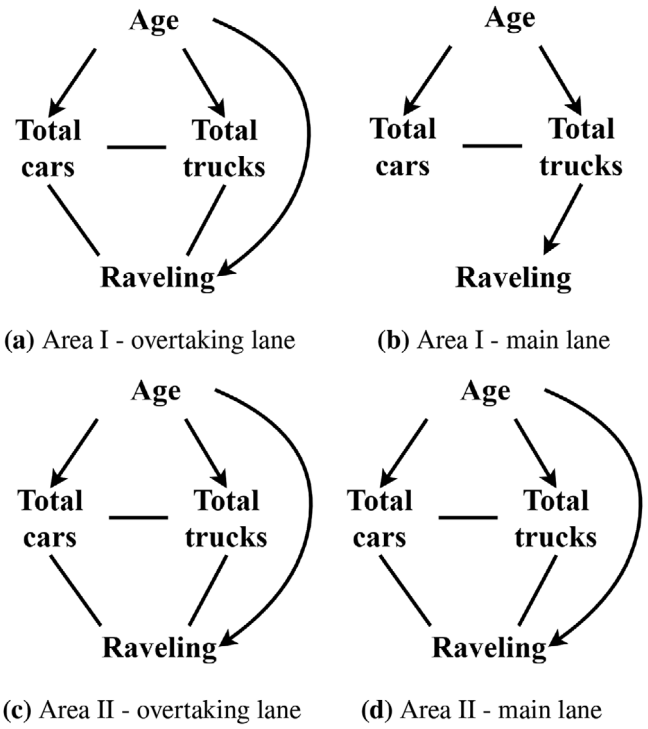

FIGURE 5 Graphical representation of causal relationships based on the two-lane study areas (95% significance). Arrows point from causes to effects. Undirected links indicate bidirectional relationships, suggesting the conditions to estimate causal directions are insufficient.

through the observation data of the study areas. Figure 5 and Figure 6 display undirected edges linking the total counts of cars, trucks, and raveling. In accordance with traffic flow theory, the model accurately represents the bidirectional relationship between car and truck volumes. There is a complex interplay between traffic flows of different user classes and overall traffic conditions as illustrated in the aforementioned dependence test. However, the model limitations in handling cyclic dependencies become evident with plausible causality directions between vehicle volumes and raveling.

Analyzing the causal graphs across diverse road configurations offers valuable insights into the correlations between traffic variables and raveling. The consistency between these graphs validates the causal discovery model. Graphs of lanes with the same configuration and functionality exhibit similarities. These similar characteristics discovered from the two-lane highway scenarios are highlighted as follows:

(1) In the two-lane highways, raveling and truck volumes have a direct causal relationship, and in the case of overtaking lanes, the same relationship exists between raveling and car volumes.

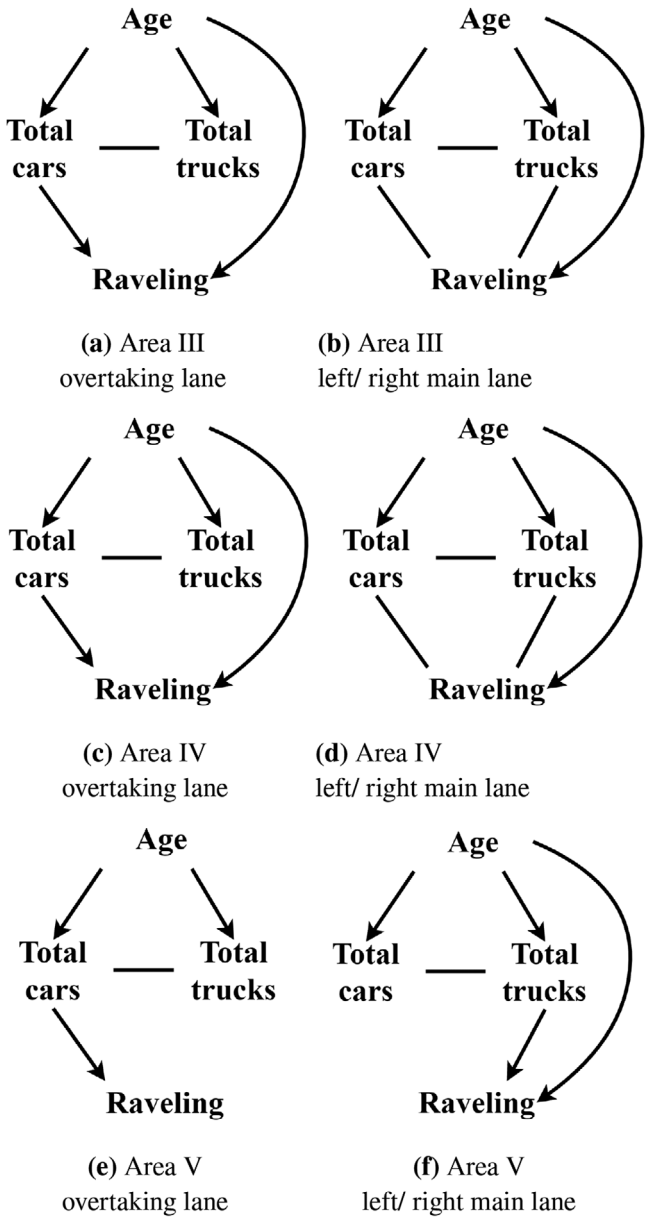

FIGURE 6 Graphical representation of causal relationships based on the three-lane study areas (95% significance). Arrows point from causes to effects. Undirected links indicate bidirectional relationships, suggesting the conditions to estimate causal directions are insufficient.

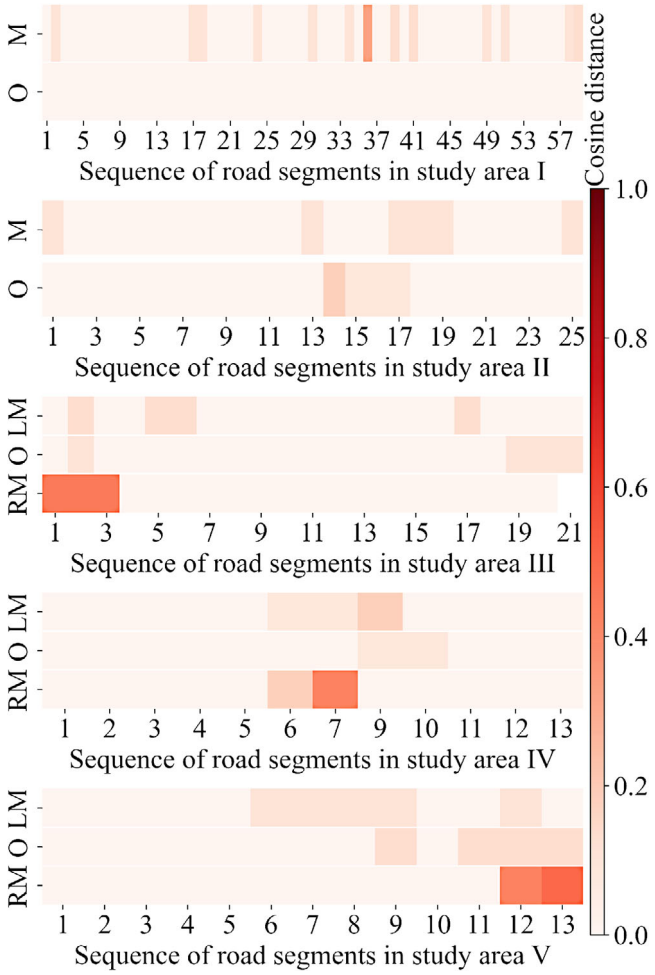

FIGURE 7 The model sensitivity to individual road sections across all the study areas is analyzed by comparing the causal graphs generated with and without the specific sections. The similarity of the graphs is estimated by the cosine distance of their corresponding matrices. (O, M, LM, and RM indicate overtaking, main, left main, and right main lanes, respectively).

(2) In the two-lane highways, road age can impact raveling in two ways: as a direct cause, through the natural aging process of a road; and as an indirect cause, due to longer use and increased accumulated traffic volumes causing a road to ravel.
(3) Car and truck flows affect each other.

Despite observed similarities in the two-lane highway scenarios, distinctions exist specifically in the main through lane of Area I. The causal relationship between raveling and both car volumes and road age lack statistical significance, while in the other two-lane scenarios, these relationships are statistically significant. Two potential factors can account for the observed differences. The one is that Area I has a considerably higher volume of trucks, approximately twice as much as in Area II. This larger volume of trucks could cause more wear on the road, resulting in more severe raveling. The other is that the findings for Area I main lane may be affected by biases in multiple specific sections as shown in Figure 7.

The three-lane highway scenarios present similarities across the overtaking and main lanes, as shown in Figure 6 as noted below:

(1) In the three-lane highways, car volumes serve as a direct cause of raveling, while truck volumes contribute as a direct cause of raveling only for the main lanes.

(2) In the three-lane highways, both direct and indirect ways that road age causes raveling are observed.
(3) Car and truck flows affect each other.

While similarities exist in the three-lane highway scenarios, the results of the main lanes have a difference, particularly in Area V. No direct causal relationship is found between the raveling and road age for its overtaking lane, and between raveling and car volumes for its main lanes. However, both direct and indirect causal relationships in the other areas are observed. These different relationships observed between traffic variables and raveling could be due to the higher truck volume in Area V. This area has twice the truck traffic on its left main lane than the other study areas, and double on its right main lane compared to Area III. Although the right main lane of Area V has the similar total truck volumes as that of Areas V and IV, its daily truck volume is 1.2 times bigger with shorter total service time. Moreover, its overtaking lane has experienced a maintenance during the observation, which could significantly slow the degradation, obscuring the correlation between road age and raveling.

The main divergence between two-lane and three-lane scenarios stems from the differing composition of vehicle classes as the direct cause of raveling in the overtaking lanes. In three-lane highway overtaking lanes, car volumes directly affect raveling, with truck volumes acting indirectly, whereas both vehicle classes have direct causal relationships with raveling in two-lane scenarios. Trucks constitute a smaller portion of overtaking lane traffic, especially in three-lane highways with the "keep right" traffic rule, as in this study. The finding could guide the allocation of maintenance resources more effectively, prioritizing areas identified as high risk due to traffic patterns.

The causal relationships identified among various road configurations, lane types, and key variables highlight the necessity of formulating tailored maintenance strategies. The significance of lane categorization points toward adopting lane-specific approaches to enhance pavement management. Utilizing the direct causes of raveling allows for customized intervention strategies. For example, the impact of road age and the volumes of specific vehicle classes indicates the need to adjust the scheduling of inspections and preventive maintenance based on real deterioration patterns and the predominant vehicle types.

The presented results above are from the CD-NOD model given the 9-year observed data and the setting of the 1-year interval. As the data collection period increases, Figure 8 illustrates an upward trend with intermittent fluctuations in the cosine similarity. With data duration less than 3 years of the study areas, the model does not

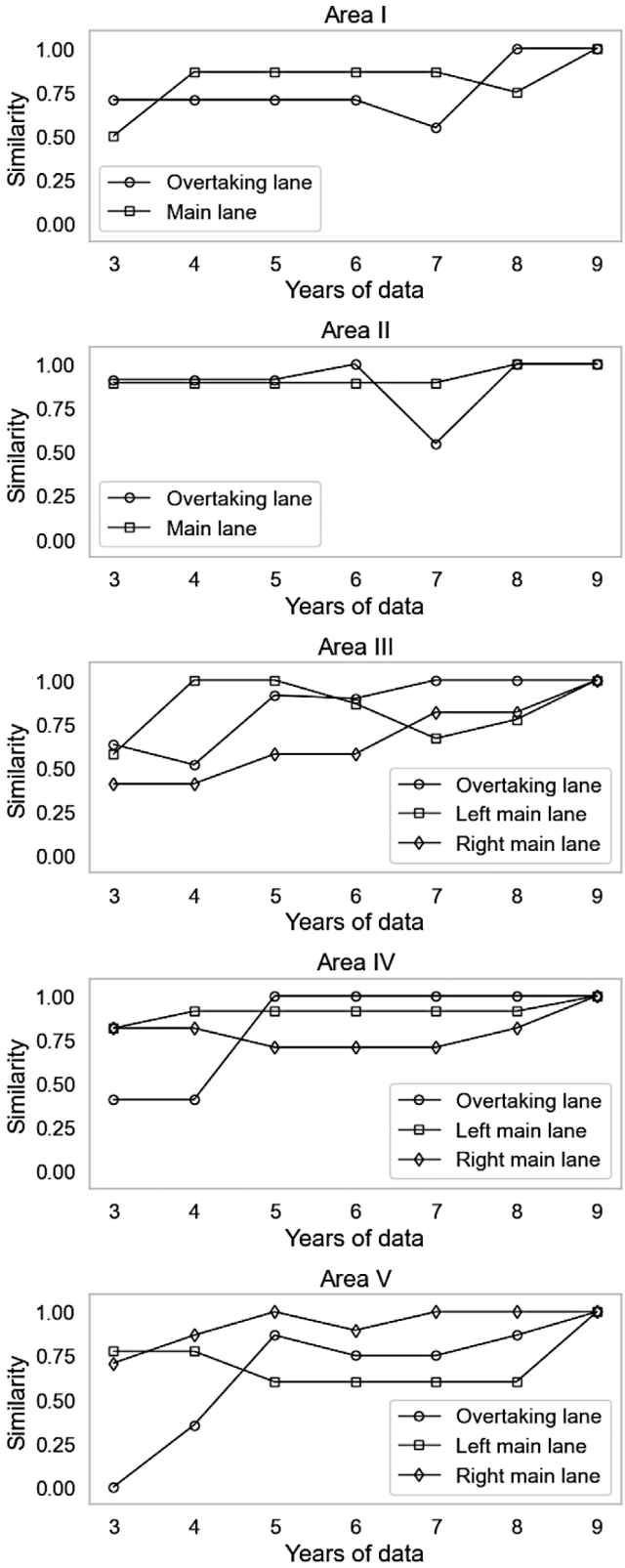

FIGURE 8 The model sensitivity to data collection duration is evaluated by the comparative study of each study area. It entails comparing the cosine similarity of causal graphs derived by the data spanning various periods, specifically from 3 to 9 years.

yield to reliable results (i.e., with 95% significance). It indicates that the model exhibits heightened sensitivity to data duration, with longer periods potentially yielding more reliable outcomes. In addition, compared to the intervals of 3 years and 5 years, this model finds more causal mechanisms directly related to road age, including the direct relationship between age and traffic flow and the relationship between age and raveling. In very few cases, it results in less direct relationships between vehicle volumes and raveling. This may be due to the fact that roads require exposure to many years of traffic volumes before any observable distress becomes apparent. Selecting an appropriate time interval is thus important, which is recommended to be based on observation duration and pavement usage.

It is highlighted that causal model outcomes can be noticeably influenced by the specific characteristics of one or more road segments. Figure 7 visually represents the influence of each road segment on the model outcomes. When certain road segments are included in the model, the causal graphs are changed by more than 40%, meaning nearly half of the causal links are inconsistent with causal graphs created when these segments are excluded. These important segments are mainly located downstream of the on-ramps or upstream of the off-ramps. The distinguishing feature of these segments suggests a unique mechanism of raveling compared to other sections. Specifically, the causal graphs incorporating these segments exhibit more direct causal relationships between raveling and traffic. This implies that the direct correlation between traffic flow and raveling appears to be more pronounced within these segments. Given the potential for increased turbulence and congestion in segments downstream of an on-ramp or upstream of an off-ramp, these traffic patterns might significantly contribute to raveling. However, an important caveat is that not all road segments positioned downstream of the on-ramps or upstream of the off-ramps significantly impact the results. These findings call for further research into the influence of road segment characteristics and traffic patterns on raveling.

5 | CONCLUSIONS

The relationship between real traffic and raveling is unclear and characterized by different interpretations of the previous field studies using correlation approaches. To contribute to the discussion, the study develops a new model to delineate plausible causal relationships between mixed traffic and raveling using field data. A thorough review of popular causal discovery models was conducted, and the CD-NOD model was selected. The research methodology is developed based on this model, and designed specifically to tackle the challenges related to confounding variables, the large disparity between the spatial–temporal scales of raveling and traffic measurements (years vs. minutes), and the relationship between traffic composition and traffic volume. The methodology is applied to five Dutch highway sections after meticulous selection of identical nontraffic conditions, and identifying traffic-related variables with potential causal influence.

The work reveals several important findings that enhance the understanding of raveling and can guide maintenance strategies. The connection between cumulative traffic volume and raveling is nonlinear, exhibiting a cause-and-effect relationship. In the relationship, road age acts as a shared contributor, both intensifying the effect of cumulative traffic volume on raveling through prolonged use and escalating the raveling process itself through natural aging. This nuanced understanding can inform predictive raveling models. This study also highlights that not only trucks but cars particularly on overtaking lanes, can cause raveling, implying the need for targeted maintenance strategies. Furthermore, the type of carriageway configuration and specific lane categories play a crucial role in establishing significant correlations and meaningful causation between mixed traffic and raveling.

The proposed methodology marks a substantial advancement, allowing for a clear distinction between mere correlations and actual causal links affecting raveling due to mixed traffic. The findings from the study not only provide a deeper understanding of the multifaceted relationship between traffic variables and the distress but also offer practical insights for infrastructure management. The findings facilitate the development of more targeted maintenance strategies, which could result in economic savings and enhanced roadway safety. The model application to the Dutch network underscores its practical relevance and potential for broader implementation. Thus, the research contributions extend from theoretical modeling to actionable solutions, with the potential to influence both future research directions and policy making in civil infrastructure.

Nonetheless, challenges persist in understanding the isolated impact of specific types of vehicles on pavement distress due to the complex interplay between traffic flows of different user classes, leading to unavoidable cyclic dependencies. This obscures the underlying mechanisms. Therefore, the acyclic nature of causal discovery models, including the CD-NOD model, limits their ability to uncover these cyclic dependencies. This limitation clearly opens up an avenue for future development of causal discovery models.

ACKNOWLEDGMENTS

This research was supported by the Ministry of Infrastructure and Water Management in the Netherlands (project No. 31160205).

REFERENCES

Abouelsaad, A., & White, G. (2020). Fretting and ravelling of asphalt surfaces for airport pavements: A load or environmental distress. In *Proceedings of the 19th annual international conference on highways and airport pavement engineering, asphalt technology and infrastructure* (pp. 11–12). Liverpool John Moores University, Liverpool, UK. https://research.usc.edu.au/esploro/outputs/conferencePresentation/Fretting-and-ravelling-of-asphalt-surfaces/99575308702621

Abouelsaad, A., & White, G. (2021). Review of asphalt mixture ravelling mechanisms, causes and testing. *International Journal of Pavement Research and Technology*, *15*, 1448–1462.

Attoh-Okine, N. O. (2001). Grouping pavement condition variables for performance modeling using self-organizing maps. *Computer-Aided Civil and Infrastructure Engineering*, *16*(2), 112–125.

Barzegari, S., & Solaimanian, M. (2019). A method for determining the appropriate frequency for testing asphalt mixtures in the laboratory. In A. F. Nikolaides & E. Manthos (Eds.), *Bituminous mixtures and pavements VII* (pp. 262–270). CRC Press.

Cai, W., Song, A., Du, Y., Liu, C., Wu, D., & Li, F. (2023). Fine-grained pavement performance prediction based on causal-temporal graph convolution networks. *IEEE Transactions on Intelligent Transportation Systems*, 1–14. https://ieeexplore.ieee.org/abstract/document/10311071

Chickering, D. M. (2002). Optimal structure identification with greedy search. *Journal of Machine Learning Research*, *3*, 507–554.

Colombo, D., Maathuis, M. H., Kalisch, M., & Richardson, T. S. (2012). Learning high-dimensional directed acyclic graphs with latent and selection variables. *The Annals of Statistics*, *40*(1), 294–321.

De Visscher, J., & Vanelstraete, A. (2017). Ravelling by traffic: Performance testing and field validation. *International Journal of Pavement Research and Technology*, *10*(1), 54–61.

Donavan, P. R. (2014). Effect of porous pavement on wayside traffic noise levels. *Transportation Research Record*, *2403*(1), 28–36.

Ghadi, M. Q., Ahmad, H., & Jannoud, I. (2023). Influence of traffic characteristics on pavement performance of parking lots. *Infrastructures*, *8*(4), 65.

Ghafoori, E. (2019). *Noise reducing asphalt pavements: A literature review on requirements, evaluating methods and recent developments*. Technical report, Swedish National Road and Transport Research Institute (VTI).

Glymour, C., Zhang, K., & Spirtes, P. (2019). Review of causal discovery methods based on graphical models. *Frontiers in Genetics*, *10*, 524.

Gretton, A., Fukumizu, K., Teo, C., Song, L., Schölkopf, B., & Smola, A. (2007). A kernel statistical test of independence. In J. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Proceedings of the 20th International Conference on Neural Information Processing Systems* (Vol. 20, pp. 585–592). Curran Associates Inc. https://dl.acm.org/doi/abs/10.5555/2981562.2981636

Guo, R., Cheng, L., Li, J., Hahn, P. R., & Liu, H. (2020). A survey of learning causality with data: Problems and methods. *ACM Computing Surveys (CSUR)*, *53*(4), 1–37.

Hagos, E. T. (2008). *The effect of aging on binder properties of porous asphalt concrete*. PhD thesis, Delft University of Technology.

Henning, T., & Roux, D. (2012). A probabilistic approach for modelling deterioration of asphalt surfaces. *Journal of the South African Institution of Civil Engineering*, *54*(2), 36–44.

Hospodka, M., & Hofko, B. (2019). Fatigue resistance of asphalt mastic by dynamic shear rheometer testing. In A. F. Nikolaides & E. Manthos (Eds.), *Bituminous mixtures and pavements VII* (pp. 27–31). CRC Press.

Hoyer, P., Janzing, D., Mooij, J. M., Peters, J., & Schölkopf, B. (2008). Nonlinear causal discovery with additive noise models. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Proceedings of the 21th International Conference on Neural Information Processing Systems* (Vol. 21, pp. 689–696). Curran Associates Inc. https://dl.acm.org/doi/abs/10.5555/2981780.2981866

Huang, B., Zhang, K., Lin, Y., Schölkopf, B., & Glymour, C. (2018). Generalized score functions for causal discovery. In Y. Guo & F. Farooq (Eds.), *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 1551–1560). Association for Computing Machinery.

Huang, B., Zhang, K., Zhang, J., Ramsey, J. D., Sanchez-Romero, R., Glymour, C., & Schölkopf, B. (2020). Causal discovery from heterogeneous/nonstationary data. *Journal of Machine Learning Research*, *21*(1), 1–53.

Huurman, M., Mo, L., & Woldekidan, M. F. (2010). Porous asphalt ravelling in cold weather conditions. *International Journal of Pavement Research and Technology*, *3*(3), 110.

Hyvärinen, A., Zhang, K., Shimizu, S., & Hoyer, P. O. (2010). Estimation of a structural vector autoregression model using non-gaussianity. *Journal of Machine Learning Research*, *11*(5), 1709–1731.

KNMI. (2023). *KNMI data platform*. https://dataplatform.knmi.nl/

Jain, S., Aggarwal, S., & Parida, M. (2005). HDM-4 pavement deterioration models for Indian national highway network. *Journal of Transportation Engineering*, *131*(8), 623–631.

Jing, R. (2019). *Ageing of bituminous materials: Experimental and numerical characterization*. PhD thesis, Delft University of Technology.

Koivisto, M., & Sood, K. (2004). Exact Bayesian structure discovery in Bayesian networks. *Journal of Machine Learning Research*, *5*, 549–573.

Kringos, N., & Scarpas, A. (2005). Raveling of asphaltic mixes due to water damage: Computational identification of controlling parameters. *Transportation Research Record*, *1929*(1), 79–87.

Kuennen, T. (2013). *Loss of cover-decoding the secrets of asphalt stripping, raveling*. Equipment World.

Maeda, T. N., & Shimizu, S. (2022). Repetitive causal discovery of linear non-Gaussian acyclic models in the presence of latent confounders. *International Journal of Data Science and Analytics*, *13*(2), 77–89.

Miradi, M. (2009). *Knowledge discovery and pavement performance: Intelligent data mining*. PhD thesis, Faculty of Civil Engineering and Geosciences, Delft University of Technology.

Mo, L., Huurman, M., Woldekidan, M. F., Wu, S., & Molenaar, A. A. (2010). Investigation into material optimization and development

for improved ravelling resistant porous asphalt concrete. *Materials & Design*, *31*(7), 3194–3206.

Mo, L., Huurman, M., Wu, S., & Molenaar, A. (2014). Mortar fatigue model for meso-mechanistic mixture design of ravelling resistant porous asphalt concrete. *Materials and Structures*, *47*, 947–961.

NDW. (2023). *Dutch national road traffic data portal*. NDW Data Portal. https://www.ndw.nu/

Nicholls, J. C., Jacobs, M. M. J., Schoen, E., van Vliet, D., Mookhoek, S., Meinen, N., van Bochove, G. G., De Visscher, J., Vanelstraete, A., Hammoum, F., Blumenfeld, T., Böhm, S., & Schulze, C. (2019). Development of a ravelling test for asphalt. In A. F. Nikolaides & E. Manthos (Eds.), *Bituminous mixtures and pavements VII* (pp. 144–152). CRC Press.

Nicholls, J. C., de Visscher, J., Hammoum, F., & Jacobs, M. (2015). Review of parameters influencing the propensity of asphalt to ravel. In *CEDR call 2014: Asset management and maintenance—DRaT development of the ravelling test*.

Nicholls, J. C., de Visscher, J., Hammoum, F., & Jacobs, M. (2016). Compendium of sites and the extent of ravelling. In *CEDR call 2014: Asset management and maintenance—DRaT development of the ravelling test* (pp. 1–26). https://www.cedr.eu/peb-research-call-2014-asset-management-and-maintenance

Ogarrio, J. M., Spirtes, P., & Ramsey, J. (2016). A hybrid causal search algorithm for latent variable models. In A. Antonucci, G. Corani, & C. P. Campos (Eds.), *Conference on probabilistic graphical models* (pp. 368–379).

Opara, K. R., Skakuj, M., & Stöckner, M. (2016). Factors affecting raveling of motorway pavements—A field experiment with new additives to the deicing brine. *Construction and Building Materials*, *113*, 174–187.

Peng, C., & Xu, C. (2023). A coordinated ramp metering framework based on heterogeneous causal inference. *Computer-Aided Civil and Infrastructure Engineering*, *38*(10), 1365–1380.

Peters, J., Mooij, J. M., Janzing, D., & Schölkopf, B. (2014). Causal discovery with continuous additive noise models. *Journal of Machine Learning Research*, *15*, 2009–2053.

Rijkswaterstaat. (2016). *Strategic planning road surfaces 2016–2022*. (Tech. Rep.). Ministry of Infrastructure and Water Management in the Netherlands.

Shimizu, S., Hoyer, P. O., Hyvärinen, A., Kerminen, A., & Jordan, M. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, *7*(10), 2003–2030.

Shimizu, S., Inazumi, T., Sogawa, Y., Hyvärinen, A., Kawahara, Y., Washio, T., Hoyer, P. O., & Bollen, K. (2011). Directlingam: A direct method for learning a linear non-Gaussian structural equation model. *Journal of Machine Learning Research*, *12*, 1225–1248.

Silander, T., & Myllymaki, P. (2006). A simple approach for finding the globally optimal Bayesian network structure. In R. Dechter & T. Richardson (Eds.), In *Proceedings of the 22nd conference on uncertainty in artificial intelligence* (pp. 445–452). AUAI Press.

Spirtes, P., Glymour, C. N., Scheines, R., & Heckerman, D. (2000). *Causation, prediction, and search*. MIT Press.

Spirtes, P., Meek, C., & Richardson, T. (1999). An algorithm for causal inference in the presence of latent variables and selection bias. *Computation, Causation, and Discovery*, *21*, 211–252.

Takahashi, S. (2013). Comprehensive study on the porous asphalt effects on expressways in Japan: based on field data analysis in the last decade. *Road Materials and Pavement Design*, *14*(2), 239–255.

Thube, D. T. (2012). Artificial neural network (ANN) based pavement deterioration models for low volume roads in India. *International Journal of Pavement Research and Technology*, *5*(2), 115–120.

Treiber, M., & Helbing, D. (2003). An adaptive smoothing method for traffic state identification from incomplete information. In H. Emmerich, B. Nestler, & M. Schreckenberg (Eds.), *Interface and transport dynamics* (pp. 343–360). Springer.

Van Aalst, W., Piscaer, P., Vreugdenhil, B., Bouman, F., van Antwerpen, M., van Antwerpen, G., & Baan, J. (2023). Raveling algorithm for PA and SMA on 3D data. In M. Crispino & E. Toraldo (Eds.), *Roads and airports pavement surface characteristics* (pp. 382–392). CRC Press.

Van Beinum, A., Farah, H., Wegman, F., & Hoogendoorn, S. (2018). Driving behaviour at motorway ramps and weaving segments based on empirical trajectory data. *Transportation Research Part C: Emerging Technologies*, *92*, 426–441.

Van Loon, H., & Butcher, M. (2003). *Analysis of a thirteen year old rap site in South Australia*. In 21st ARRB Transport Research Conference.

Van Reisen, F., Erkens, S., Ven, M., Voskuilen, J., & Hofman, R. (2008). *Synthesis report of improving service life double-layer ZOAB* (Verbetering levensduur tweelaags ZOAB synthese rapport in Dutch). Technical report, Foundation Eurasphalt. https://trid.trb.org/View/743917

Voskuilen, J., Tolman, F., & Rutten, E. (2004). Do modified porous asphalt mixtures have a longer service life? In *Proceedings of the 3rd Eurasphalt and Eurobitume Congress* (Vol. *2*), Vienna, May 2004.

Wang, M., Zhang, Y., Zhang, D., Zheng, Y., Zhou, S., & Tan, S. K. (2021). A Bayesian decision model for optimum investment and design of low-impact development in urban stormwater infrastructure and management. *Frontiers in Environmental Science*, *9*, 713831.

Wang, Z., Krishnakumari, P., Anupam, K., van Lint, J., & Erkens, S. (2022). Spatial-temporal analysis of road raveling and its correlation with traffic flow characteristics. In *2022 IEEE 25th international conference on intelligent transportation systems (ITSC)* (pp. 3609–3616). IEEE.

WegenWiki. (2023). *Beschrijvende plaatsaanduidingsystematiek*. https://www.wegenwiki.nl/Beschrijvende_Plaatsaanduiding Systematiek

Xie, F., Cai, R., Huang, B., Glymour, C., Hao, Z., & Zhang, K. (2020). Generalized independent noise condition for estimating latent variable causal graphs. *Advances in Neural Information Processing Systems*, *33*, 14891–14902.

You, Q., Zheng, N., & Ma, J. (2018). Study of ravelling failure on dense graded asphalt pavement. *Proceedings of the Institution of Civil Engineers-Transport*, *171*(3), 146–155.

Yuan, C., & Malone, B. (2013). Learning optimal Bayesian networks: A shortest path perspective. *Journal of Artificial Intelligence Research*, *48*, 23–65.

Zhang, H., Anupam, K., Skarpas, A., Kasbergen, C., & Erkens, S. (2020). Simple homogenization-based approach to predict raveling in porous asphalt. *Transportation Research Record*, *2674*(12), 263–277.

Zhang, J., Cao, J., Huang, W., Shi, X., & Zhou, X. (2023). Rutting prediction and analysis of influence factors based on multivariate transfer entropy and graph neural networks. *Neural Networks*, *157*, 26–38.

Zhang, K., & Hyvarinen, A. (2016). Nonlinear functional causal models for distinguishing cause from effect. In W. Wiedermann & A. von Eye (Eds.), *Statistics and Causality: Methods for Applied Empirical Research* (pp. 185–201). John Wiley & Sons, Inc.

Zhang, K., Peters, J., Janzing, D., & Schölkopf, B. (2011). Kernel-based conditional independence test and application in causal discovery. In F. Cozman & A. Pfeffer (Eds.), *27th conference on uncertainty in artificial intelligence (UAI 2011)* (pp. 804–813). AUAI Press.

Zhang, Y., van de Ven, M., Molenaar, A., & Wu, S. (2016). Preventive maintenance of porous asphalt concrete using surface treatment technology. *Materials & Design*, *99*, 262–272.

How to cite this article: Wang, Z., Krishnakumari, P., Anupam, K., van Lint, H., & Erkens, S. (2024). A causal discovery approach to study key mixed traffic-related factors and age of highway affecting raveling. *Computer-Aided Civil and Infrastructure Engineering*, 1–20. https://doi.org/10.1111/mice.13222