Pavement condition prediction for communities: A low-cost, ubiquitous, and network-wide approach

Tao Tao[1] and Sean Qian[2]

[1] Ph.D., Carnegie Mellon University, taot@andrew.cmu.edu

[2] Professor, Ph.D., Carnegie Mellon University (Corresponding author), seanqian@cmu.edu

Abstract

Effective prediction of pavement deterioration is critical to forecast infrastructure performance and make infrastructure investment decisions under escalating environmental and traffic change. However, most communities often struggle to undertake such predictive tasks due to limited sensing capacity and lack of granular data. In addition, most studies focus on predicting the absolute pavement conditions rather than monthly/yearly changes, potentially overestimating the effect of current pavement conditions and thus underestimating the perceived effect of underlying driving factors. With the pavement condition rating (PCR) data generated from AI-powered computer vision technologies and multiple openly available datasets, we propose a low-cost and ubiquitous approach to predict system-level pavement conditions, using nine communities across the US as an example. In addition to predicting absolute PCRs as was done in classical models, we develop another set of models to predict the change in PCRs over any time increment (i.e., time lapse between a PCR observation and retrofit decision point). The findings show that the proposed low-cost prediction approach yields results comparable to existing studies, demonstrating its promising application in supporting pavement management. Furthermore, PCR change model indicates that, besides current PCR, weather, road classification, socioeconomics, and built environment attributes are important to predicting PCR change. The interactive impacts also show salient interactive effects between variables and current PCR, offering suggestions on better allocating the limited resources in pavement maintenance projects. Finally, the proposed model could enhance climate resiliency and transportation equity during the pavement management process.

1 Introduction

Effective prediction of pavement conditions plays a critical role in maintaining the performance of road networks and enhancing the resilience of communities. Accurate predictions enable transportation agencies to proactively manage pavement assets, schedule timely maintenance, and allocate resources more efficiently (Piryonesi and El-Diraby 2021). This leads to improved serviceability of roads, extended pavement life, and reduced life-cycle costs. In the context of increasing environmental challenges, accurate pavement condition predictions also support efforts to enhance the resilience of communities. By anticipating and mitigating pavement distresses associated with changing climate and extreme weather events, communities can better prepare for and adapt to these challenges, ensuring the continuity and reliability of their transportation systems.

While some large metropolitan areas may have the capacity and resources to undertake pavement condition prediction tasks and utilize the results to inform their pavement asset management programs, most communities face more significant challenges in doing so mainly because they cannot afford to collect and maintain datasets related to pavement conditions and relevant built/natural environment measures, oftentimes depending on costly sensors and lab experiments. It is a common challenge, for both large and small communities, to measure pavement ratings for more than a small fraction of road segments on a yearly basis. Although studies have been using several datasets maintained by either federal or local transportation agencies to carry out pavement condition predictions, these datasets have several limitations. First, these datasets mainly cover a small fraction of highways and arterial road sections in a community-wide area which provide limited insights for small or underserved communities as most of their roads are local roads. Second, these datasets only encompass a limited number of road sections for each community. Most communities do not have access to a complete set of road inventory data for every road segment, such as pavement materials, pavement ratings, etc. Thus, it is unclear if the resultant predictive model does a comprehensive examination on all the road sections in the whole network. Third, these datasets often contain missing value for variables that are important to predict

pavement conditions, such as pavement characteristics and traffic activities. In addition to the challenges raised by the quality and availability of related datasets, most studies only split the sample once when they train and test models, making the model performance heavily dependent on specific data partitioning. Finally, most studies estimated models to predict absolute pavement condition, which generate disproportionate impact of the initial or previous pavement condition and may lead to a skewed perception of the contributions of underlying driving factors.

With the pavement condition data generated from AI-powered computer vision technologies and multiple openly available datasets, we carry out network-wide pavement condition predictions using a complete coverage of all roads in nine communities across the US. We use a rigorous approach known as repeated nest cross validation to tune parameters and select the best modeling method among multiple machine learning methods. In addition to predicting absolute value of pavement conditions as was done in classical models, we develop another set of models to predict the change of pavement conditions over any time increment (i.e., time lapse between two observations of pavement conditions for road segments). Specifically, we aim to address three research questions:

1) What is the performance of the proposed models in predicting pavement conditions and how do they compare with prevailing models in the literature?
2) How do the outcomes of models that predict the change in pavement conditions differ from those that predict the absolute pavement condition?
3) What are the contributions of variables to pavement condition predictions and how variables impact pavement conditions in estimated models?

Our study contributes to literature in several ways. First, we propose a low-cost and ubiquitous approach to predict system-level pavement conditions in communities with open-source datasets available from a full coverage of geographical locations and a long time span, and the results showed that the model results are comparable to the models for roads/pavements in other studies. Second, with large-scale pavement condition data and other openly available datasets, we model the whole network instead of selected road sections in study areas, which enables better decision-making when it comes to

prioritizing maintenance and retrofit efforts. Third, we innovatively use two types of large-scale and open-source datasets: socioeconomics and built environment variables to account for the impacts of traffic activity and pavement characteristics to address the issue that these two types of variables are often not available. Fourth, in addition to pavement condition models, we estimate pavement condition change across any time increments, offering an unbiased assessment of the variables' contributions to the overall outcome. Fifth, we apply repeated nested cross validation to tune and select best model, which reduces the influence of randomness in sample splitting. Finally, we use two interpretation tools (i.e., SHAP values and accumulated local effect plots) to estimate the contributions of independent variables and their main and interactive effects on estimating pavement conditions, which provide more insights into how different factors affect pavement conditions individually and interactively. The results offer suggestions on efficiently allocating the limited resources in pavement maintenance projects.

The rest of the study is organized as follows. We review the studies on pavement condition predictions and related research gaps in Section 2. We introduce our datasets and methods in Section 3. We present and discuss the results in Section 4. In the last section, we conclude this study and offer implications for pavement condition prediction in communities.

2 Literature review

Pavement condition is an important index to help policymakers plan maintenance projects on the road segments and ensure the roads are reliable and safe for their users. Modeling and predicting of the pavement condition is an important but also challenging task (Sidess, Ravina, and Oged 2020).

The collection of pavement condition data primarily employs two distinct approaches: manual walking survey and automated survey (Attoh-Okine and Adarkwa 2013). The manual walking survey involves trained professionals visually identifying and cataloging pavement distresses using specialized tools. This method, while reliable, can be time-consuming and subject to human error. Conversely, automated surveys involve the use of sophisticated technology to directly record pavement condition data. This method is faster and more accurate than manual surveys, offering more objective results while

requiring fewer human resources. The evolution of computer vision and deep learning technologies has significantly advanced automated surveys. These technologies not only increase the accuracy of distress identification but also reduce the overall cost of surveys by optimizing data collection and analysis processes (Xu and Zhang 2022; Sholevar, Golroo, and Esfahani 2022). With a large dataset of images as training set, deep learning techniques can accurately recognize the specific patterns of pavement distresses and offer an objective rating.

Pavement conditions are affected by multiple factors which studies need to consider in their predictive models. Table 1 summarizes these factors used in selected studies on pavement condition predictions focusing on the areas in the North America. These variables mainly include two distinct types: internal and external factors (Xu and Zhang 2022). Internal factors include initial/current pavement condition, pavement characteristics, and age of the pavement. Pavement characteristics include pavement types and material features (Lou et al. 2001; Kargah-Ostadi, Stoffels, and Tabatabaee 2010). The age of the pavement, also known as its service life, refers to the elapsed time since the pavement was originally constructed (Ziari et al. 2015). External factors include weather/climate, road characteristics, traffic activity, and maintenance history. Weather or climate includes temperature, wind speed, freeze thaw cycles, and precipitation (Barua et al. 2021; Piryonesi and El-Diraby 2020; Hossain, Gopisetti, and Miah 2017). Road characteristics include road classification and geometry (Piryonesi and El-Diraby 2020). Traffic activity, also known as traffic load, primarily include equivalent single axle load (ESAL) and annual average daily traffic (AADT) (Barua et al. 2021; Kargah-Ostadi, Stoffels, and Tabatabaee 2010). Maintenance history refers to elapsed time since last maintenance (Piryonesi and El-Diraby 2020; Yang et al. 2003).

Studies usually utilized two types of datasets: the Long-Term Pavement Performance (LTPP) dataset and locally collected datasets. The LTPP dataset is collected and maintained by the Federal Highway Administration. It has been applied in many studies on pavement condition prediction (Table 1). This dataset includes information of over 2,500 test sections on highways and arterial roads across the US and Canada (Elkins and Ostrom 2021). Besides pavement condition data over time, LTPP also

provides factors related to pavement condition deterioration, such as pavement characteristics, age, weather, traffic activity, and maintenance history. Local transportation agencies also collected their own pavement condition datasets to satisfy the requirements of their own pavement management programs.

These datasets, however, have several limitations. First, the majority of these datasets only contain information for roads of a higher classification, such as freeways and major arterial roads. They are often the focus due to their higher traffic volumes and the more significant consequences should they fail or become excessively deteriorated. However, this leaves gaps for lower-classification roads like local or rural roads. Second, many of these datasets are limited in their geographical scope, typically encompassing only a few road segments from a single location (Table 1). While these samples provide valuable information, their limited coverage presents challenges when attempting to extrapolate findings for network-wide pavement performance predictions. Third, some of the factors that are important for pavement condition predictions are missing in these datasets. For example, AADT is not available for a large portion of the road sections in LTPP (Gong et al. 2018). Finally, the availability of these datasets is often limited in smaller communities, such as townships and rural areas. These communities may have different traffic patterns, climate conditions, and maintenance practices from large metropolitan areas. The lack of comprehensive pavement condition datasets can pose significant challenges for local authorities in effectively managing their pavement infrastructures.

The methods applied to predict the road pavement conditions mainly have two types: deterministic and probabilistic models (Sidess, Ravina, and Oged 2020). Deterministic models include mechanistic, empirical, mechanistic-empirical models. Mechanistic models are based on the principle of mechanics and make predictions of the pavement responses. Empirical models predict the pavement conditions based on experimental and observed data. Mechanistic-empirical models are the combinations of both. Probabilistic models mainly use probabilistic distributions to model the pavement conditions. Markov chain models and survivor curves are examples of probabilistic models (George, Rajagopal, and Lim 1989). Deterministic models, especially empirical ones, are most widely applied in literature. As empirical models usually do not rely on the principle of mechanics and probabilistic distributions, their

capability to fit the data, especially the relationships between pavement conditions and other factors, are limited. Machine learning approaches have been applied to address this limitation. As shown in Table 1, several approaches including artificial neural networks (ANN), gradient boosting machines (GBM), and random forests (RF) were used by these studies for either regression or classification tasks.

To tune or select the best model, researchers commonly use train-test split, train-test-validation split, and cross-validation (Table 1). Train-test split involves dividing the sample into training and testing sets to assess model performance. An improved approach is train-test-validation split, which adds a validation set for final evaluation. Cross-validation splits the sample into multiple subsets, repeatedly training and testing the model on different fold combinations. However, these methods rely on a single split, making model performance susceptible to specific data partitioning.

All studies enumerated in Table 1 use pavement condition as their dependent variable. This is intuitive as the forecasted pavement condition can be directly applied to inform decision-making processes in pavement management. However, one potential drawback of this research design lies in the disproportionate influence of the initial or current pavement condition in models predicting future pavement states. This factor often plays a dominant role, thereby overshadowing the contributions of other variables. Consequently, this might lead to a skewed perception of the relative importance of these other factors, potentially underestimating their influence on pavement deterioration over time. For example, Piryonesi and El-Diraby (2020) estimated a GBM model to predict pavement condition index (PCI) and found that initial PCI significantly dominated the model, with all other variables contributing less than 20% of the influence of initial PCI. Evers, Bashar, and Torres-Machi (2023) applied a RF model to predict international roughness index (IRI) and had a similar finding.

Table 1. Model information for selected studies on future pavement condition prediction with advanced machine learning approaches in North America

Author	Location	Sample size and coverage	Model [a]	Model tuning/selection approach	Outcome [b]	Initial/ current pavement condition	Pavement features	Age	Weather /climate	Road characteristics	Traffic activity	Maintenance history	Model fitness
Lou et al. (2001)	Florida, US	Unknown, selected segments	ANN	Train-test-validation split	CI		✓	✓				✓	1-year R^2: 0.9-0.91
Yang et al. (2003)	Florida, US	Unknown, selected segments	ANN	Train-test split	PCR	✓		✓				✓	1-year R^2: 0.79-0.88 2-year R^2: 0.55-0.76 3-year R^2: 0.52-0.59 4-year R^2: 0.4-0.48 5-year R^2: 0.2-0.38
Kargah-Ostadi, Stoffels, and Tabatabaee (2010)	North America (LTPP)	1072, selected segments	ANN	Train-test split	IRI	✓	✓	✓	✓		✓	✓	1-year R^2: 0.96
Ziari et al. (2015)	North America (LTPP)	205, selected segments	ANN	Train-test-validation split	IRI		✓	✓	✓		✓		1-year R^2: 0.96 2-year R^2: 0.97
Hossain, Gopisetti, and Miah (2017)	North America (LTPP)	Unknown, selected segments	ANN	Train-test split	IRI			✓	✓		✓		
Piryonesi and El-Diraby (2020)	North America (LTPP)	943, selected segments	GBM	Cross-validation	PCI	✓	✓	✓	✓	✓	✓	✓	3-year Classification accuracy: 82%
Barua et al. (2021)	North America (LTPP)	1091-2090, selected segments	GBM	Train-test-validation split	PCI	✓	✓	✓	✓		✓	✓	Runway R^2: 0.9 Taxiway R^2: 0.86
Evers, Bashar, and Torres-Machi (2023)	Colorado, US	Unknown, selected segments	RF	Train-test split	IRI	✓	✓		✓		✓	✓	
This study	Nine communities, US	1938-11963, network-wide segments	GBM/RF	Repeated nested cross-validation	PCR	✓	✓ (Proxy)		✓	✓	✓ (Proxy)		1-year R^2: 0.90 2-year R^2: 0.84 3-year R^2: 0.85 4-year R^2: 0.78

Note: [a] ANN: Artificial neural network; GBM: Gradient boosting machine; RF: Random Forest;
[b] CI: Cracking index; PCR: Pavement condition rating; IRI: International roughness index; AC: Alligator cracking; PCI: Pavement condition index

3 Data and method

3.1 Data

The pavement condition rating (PCR) was obtained from RoadBotics Inc. RoadBotics uses AI-powered computer vision technologies to evaluate the performance of road infrastructure (RoadBotics 2023). They develop mobile device applications to visualize and analyze roads and generate ratings. This approach is low-cost and easy to carry out.

RoadBotics categorizes their PCR into five levels, ranging from one to five. A lower PCR indicates the corresponding road pavement condition is better. Figure 1 presents two example road segments. The left one is a road segment with its PCR equal to one, which is in good condition. The right one is five, which has cracks and small potholes. For each road segment, RoadBotics collects ratings of multiple locations along it. Then, they calculate the average value of these ratings as the rating for the corresponding road segment. Besides PCR, RoadBotics documents the date when the PCR is collected.

(a) PCR = 1 (b) PCR = 5

Figure 1. Two example road segments with different PCRs by RoadBotics

The PCR data provided by RoadBotics contains road segments from multiple communities and multiple years. The data include Dormont, Robinson, Shaler, and Upper St Clair in Pennsylvania, Alma in Arkansas, Cumberland in Maryland, Sedona in Arizona, South Bend in Indiana, and Longboat Key in Florida (Figure 2). In the US, these are in general small communities (e.g., small cities and townships), with population ranging from 5,850 to 103,353 (Table 2).

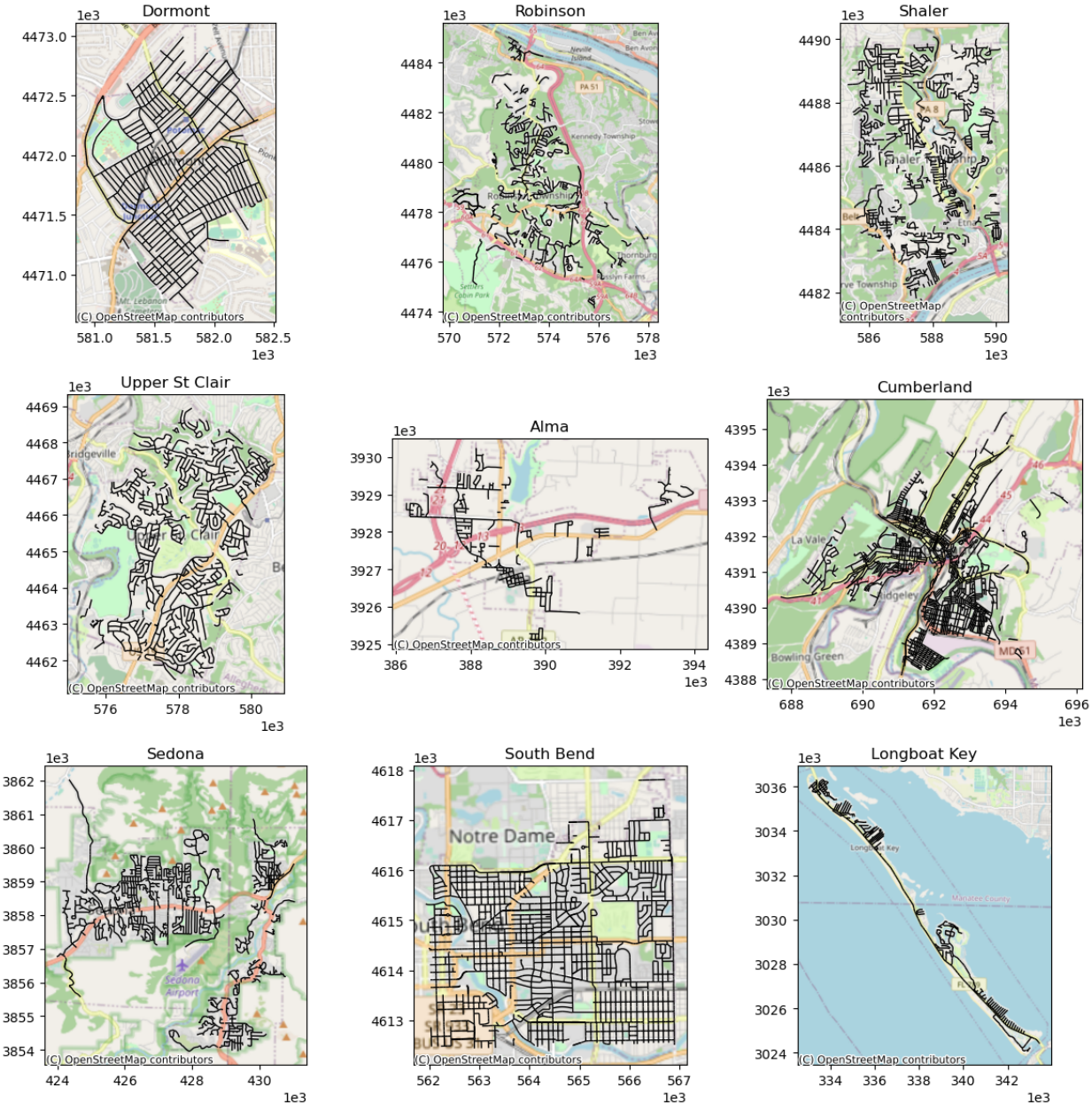

Figure 2. Study areas

Table 2. Number and mileage of road segments with valid PCR in the communities from 2018 to 2022

Community	State	Population (year)	Number of road segments (Mileage)				
			2018	2019	2020	2021	2022
Dormont	Pennsylvania	8117 (2021)	470 (26.5)	436 (25.5)	467 (26.3)		482 (26.9)
Robinson	Pennsylvania	13673 (2018)		677 (71.8)	695 (72.8)		
Shaler	Pennsylvania	27963 (2018)	1031 (93.2)	1054 (95.1)	1055 (94.5)		
Upper St Clair	Pennsylvania	19737		784	780		

		(2019)		(92.4)	(92.4)		
Alma	Arkansas	5850 (2021)		264 (26.8)			306 (29.2)
Cumberland	Maryland	18736 (2021)	2041 (114.4)	2249 (125.4)		2102 (115.6)	230 (12.3)
Sedona	Arizona	9763 (2021)		988 (86.4)	931 (83.6)	966 (84.8)	
South Bend	Indiana	103353 (2021)			2164 (125.4)	2022 (119.8)	
Longboat Key	Florida	7592 (2021)		375 (37.4)		349 (36.2)	

RoadBotics collected PCRs for most of the road segments for each community. However, from 2018 to 2022, RoadBotics collected data in different years for different communities since the pavement management plans of these communities are different. Dormont has the most temporal coverage, including 2018, 2019, 2020, and 2022. Robinson, Upper St Clair, Alma, South Bend, and Longboat Key have the least coverage, with two years of data available. In addition, the coverage of road segments is slightly different across years. Totally, all nine communities have 22918 road segments across five years, which encompass 1714.7 miles.

We requested weather data from the Local Climatological Data (LCD) tool hosted by the National Centers for Environmental Information. The stations for data collection are located across the US, mostly in airports. We assigned the weather data from the closest station to each community. The LCD data include hourly records of multiple measurements, such as temperature, wind speed, and precipitation.

We extracted the road classification data from Open Street Map (OSM) road networks. We aggregated the OSM labels of road classification into three levels: primary road, secondary road, and local and other roads. Primary roads include *motorway*, *primary*, *trunk* and related links. Secondary roads include *secondary* and *secondary_link*. Local and other roads include all other types of roads. We then matched the road classification of OSM network to the road segments by RoadBotics.

Another important type of variable that influences pavement condition is traffic activity (or traffic load). Variables such as Annual Average Daily Traffic (AADT) and equivalent single axle load (ESAL) are frequently utilized in research to account for the impact of traffic activity. However, these types of

variables are usually only observed for primary roads in major urban areas. Estimates of AADT for all roads may be available but oftentimes are inaccurate. Most communities do not have the capacity to collect and maintain this information. In addition, the road segments in the communities considered in this study are mostly local roads, which are not covered by the traffic data collection. To account for the impact of traffic activity more effectively, we incorporated socioeconomic and built environment variables derived from two open data sources. The strength of traffic activity is closely related to socioeconomics and built environment variables. In this study, socioeconomic variables include total population, car to work percentage, labor percentage, and median household income, which were downloaded from the National Historical GIS database hosted by University of Minnesota. Socioeconomics are available each year for all census block groups (CBGs) in the US. Built environment variables include employment density, network density, job accessibility, and land use mix, which were downloaded from the Smart Location database hosted by the Environmental Protection Agency. Built environment features are available in 2019 for all CBGs in the US. In this study, we assumed that socioeconomics changed by year. Furthermore, we assumed that built environment variables remained consistent across the years, which is justifiable considering the relatively short study period of five years, spanning from 2018 to 2022.

In addition to traffic activity, socioeconomic factors are also correlated with pavement characteristics. This association arises primarily because an important portion of investment into pavement infrastructure is sourced from local budgets, which are inherently linked to the socioeconomic status of the respective areas. Therefore, incorporating socioeconomic factors into pavement prediction models can help account for variations in infrastructure investment and its subsequent impact on pavement characteristics.

Table 3 lists the variables and their description, data source, and time. Some of the variables, including weather and socioeconomics, were aggregated to certain periods according to the research design, which is covered in the next section.

Table 3. Variable description, source, and time

Variable	Description	Source	Time
PCR	Pavement condition rating of the road segment in the corresponding date	RoadBotics [a]	2018-2022
PCR change	Change of pavement rating condition between current and previous dates when PCR was collected	This study	
Road classification			
Road classification (Dummies)	Road classification of the road segment, including primary road, secondary road, and local/other road	Open Street Map	-
Weather (aggregate for the days between the two dates of PCRs)			
Maximum temperature	Average maximum daily temperature in Fahrenheit	LCD [b]	2018-2022
Minimum temperature	Average minimum daily temperature in Fahrenheit		
Average temperature	Average daily temperature in Fahrenheit		
Frozen days	Number of days with daily minimum temperature lower than 32 Fahrenheit		
Hot days	Number of days with daily maximum temperature higher than 86 Fahrenheit		
Freeze thaw cycles	Freeze thaw cycles, calculated based on the half of the number that hourly temperature crossed 32 Fahrenheit		
Total precipitation	Total precipitation in inches		
Wind speed	Average daily wind speed in miles per hour		
Windy days	Number of days with daily maximum wind speed larger than 15 miles per hour		
Socioeconomic characteristics (weighted by area percentage of CBGs within one-mile buffer of the road segment)			
Total population	Average total population of the years between the dates of the PCRs	NHGIS [c]	2018-2022
Car to work percentage	Average percentage of population commuting with car of the years between the dates of the PCRs		
Labor percentage	Average percentage of population in labor force of the years between the dates of the PCRs		
Change of total population	The difference of total population between the years of the dates of the PCRs		
Change of car to work percentage	The difference of car to work percentage between the years of the dates of the PCRs		
Change to labor percentage	The difference of labor percentage between the years of the dates of the PCRs		
Median household income	Average median household income of the between the years of the dates of the PCRs		
Change of income	The difference of median household income between the years of the dates of the PCRs		
Built environment characteristics (weighted by area percentage of CBGs within one-mile buffer of the road segment)			
Employment density	Gross employment density (jobs/acre) on unprotected land	EPA [d]	2019
Land use mix	Land use entropy calculated based on the 5-tier employment (retail, industrial, service, entertainment, office) numbers on unprotected land		
Network density	Total road network density (miles per square mile)		
Job accessibility	Jobs within 45 minutes auto travel time, time-decay (network travel time) weighted		
Other			
Community (Dummies)	Community where the road is located	RoadBotics	-
Number of days	Number of days between the two dates of the PCRs	This study	

Note: [a] RoadBotics: https://www.roadbotics.com/

[b] Local Climatological Data: https://www.ncdc.noaa.gov/cdo-web/datatools/lcd

[c] National Historical GIS: https://www.nhgis.org/

[d] US Environmental Protection Agency Smart Location Database: https://www.epa.gov/smartgrowth/smart-location-mapping#SLD

3.2 Method

3.2.1 Research design

In this study, we predicted absolute PCR and PCR change in short term with variables including current PCR, road classification, weather, socioeconomics, built environment attributes, and several other variables. We used Equation (1) below to carry out the prediction for absolute PCR.

$$PCR = f(PCR_0, R, W, S, B, O) \quad (1)$$

In the equation, PCR indicates the future PCR to be estimated and PCR_0 is the current PCR. For each road segment, it has PCRs from multiple years as shown in Table 2. Based on the present research design, one PCR was chosen as the current PCR and another PCR from a future year was chosen for prediction. For instance, a road segment in Shaler has three PCRs collected in 2018, 2019, and 2020. We constructed three observations and incorporated them into the final analysis: 1) using the 2018 PCR as the current PCR and the 2019 PCR as the future PCR; 2) using the 2019 PCR as the current PCR and the 2020 PCR as the future PCR; and 3) using the 2018 PCR as the current PCR and the 2020 PCR as the future PCR. Given this approach, for road segments with PCRs available over N years, each segment can construct C_N^2 (i.e., the number of unique combinations of 2 items chosen from a set of N items, also known as N choose 2) observations. The key strength of this research design lies in its versatility. It can seamlessly accommodate any assortment of road segments and any selection of dates for PCR measurements.

In this study, we focused solely on observations where the current PCR is lower than the future PCR, indicating pavement deterioration. We excluded observations where the current PCR is higher than the future PCR, representing pavement improvement, to minimize the effects of unaccounted-for pavement maintenance and potential inaccuracies in PCR measurements on our estimations.

W indicates weather variables. All of these variables were aggregated over the time period between the two PCR dates (i.e., current and future PCRs) for each constructed observation, as demonstrated in the previous examples. The dates of the PCRs were originally documented by

RoadBotics. Weather variables include three main types: temperature, precipitation, and wind speed. Temperature includes average maximum temperature, average minimum temperature, and average temperature, which represent the average-level temperature condition. These variables can depict the diverse climate conditions found in different communities. However, although the time period between two PCR dates can vary significantly (e.g., ranging from approximately one year to four years), these average-level temperature variables may be similar across different time periods. To address this issue, we supplemented the dataset with variables reflecting the length of periods, including number of frozen days, number of hot days, and freeze thaw cycles. Similarly, we included both average wind speed and number of windy days for wind speed. Additionally, we incorporated total precipitation into the analysis.

S represents socioeconomic attributes, including total population, car to work percentage, labor percentage, and median household income. For each variable, we included their average value which was calculated through years between the two PCR dates and difference between the two years of the current and future PCRs.

R indicates road classification and B indicates built environment variables. These two types of variables are assumed as consistent through years. Additionally, we incorporated two further variables into our analysis. One is community, which serves to control for the effects associated with varying contexts of communities. The other variable is the number of days between the two PCR dates, providing a measure of time elapsed between assessments.

f represents the model used to estimate the relationships between future PCR and multiple types of factors. Model selection is an important task as it significantly affects the prediction performance. We illustrate the approach of model selection in the next section.

Besides the set of models to predict absolute values of future PCR, we also estimate another set of models to predict PCR change. These two sets of models include the same types of independent variables. As discussed in the literature review section, we expect the set of PCR change models could provide new insights into how different factors contribute to pavement condition deterioration. Table 4 presents the descriptive statistics of the variables considered in the modeling process.

Table 4. Descriptive statistics of the observations

Variable	Mean	Standard Deviation	Minimum	Maximum
Future PCR	2.94	0.95	1	5
PCR change	0.55	0.45	0	3.67
Current PCR				
Current PCR	2.39	0.9	1	5
Road classification				
Road classification (Dummies)	Primary: 3% Secondary: 6% Local and other: 92%			
Weather				
Maximum temperature	65.38	4.51	57.29	83.92
Minimum temperature	47.01	3.65	41.24	68.59
Average temperature	55.91	3.81	49.31	75.8
Frozen days	154	94	0	392
Hot days	86	64	17	369
Freeze thaw cycles	275.43	190.4	0	842
Total precipitation	80.02	50.9	9.12	207.52
Wind speed	6.5	1.45	4.39	8.4
Windy days	246	107	102	519
Socioeconomic characteristics				
Total population	1330	342	420	3017
Car to work percentage	0.83	0.09	0.46	1
Labor percentage	0.59	0.1	0.14	0.83
Median household income	63513	22992	30099	169371
Change of total population	-20	263	-1660	956
Change of car to work percentage	-0.01	0.03	-0.34	0.15
Change of labor percentage	0	0.03	-0.13	0.12
Change of income	4031	6477	-58544	45165
Built environment characteristics				
Employment density	2.34	2.18	0.03	13.49
Land use mix	0.67	0.12	0.17	0.91
Network density	17.54	7.55	1.14	30.92
Job accessibility	44052	40831	1434	134332
Other				
Community (Dummies)	Dormont: 16% Robinson: 3% Shaler: 15% Upper St Clair: 4% Alma: 2% Cumberland: 37% Sedona: 11% South Bend: 12% Longboat Key: 1%			
Number of days	697	329	325	1500

3.2.2 Method

Within the existing literature, several models demonstrate exceptional performance in predicting pavement conditions (Table 1), including random forest (RF), gradient boosting machine (GBM), and multiple-layer perception (MLP, i.e., artificial neural network). Therefore, in this study, we employed these methods, in addition to using linear regression model (LM) as a benchmark, to conduct our estimations.

Both RF and GBM are models based on decision trees. Decision trees work by splitting the sample into subsamples according to various decision rules based on the values of independent variables, and then using the average value of the dependent variable in subsamples for predictions. One limitation of decision trees is their relatively weak predictive capability. To address this issue, there are two popular approaches: bagging and boosting. Bagging combines the results of multiple decision trees in parallel, which is applied by RF (Breiman 2001). Boosting is an iterative process that combines decision trees in a sequential order. Boosting is used in GBM (Friedman 2001, 2002). Both RF and GBM are capable of modeling complex relationships between dependent and independent variables and exhibit strong predictive performance.

MLP consists of multiple layers of interconnected nodes (i.e., artificial neurons or perceptions). These layers include an input layer, one or more hidden layers, and an output layer. The neurons of the hidden layers apply nonlinear activation functions to the results from the previous layer. These nonlinear activation functions, including logistic function, hyperbolic tangent function, and rectified linear unit (ReLU), allow MLP to model complex, nonlinear relationships in the data.

We employed repeated nested cross validation to compare the prediction performance of the selected models. Nested cross validation, as its name suggests, consists of two nested loops of cross validation (Figure 3). The inner loop aims to select the best combination of parameters for the models. The outer loop evaluates the performance of the tuned models. The nested cross validated is repeated multiple times and the performance of the models is recorded, including R-squared, mean absolute error (MAE), and root mean squared error (RMSE). Introducing repetition to this process could help mitigate

the impact of having a specific random split of the data that can potentially bias the model performance. After the best modeling method is selected, another round of repeated cross validation is used to search for the best combination of parameters for the selected modeling method.

In this study, we ran 10-time repeated cross validation on the sample to select the best modeling method among LM, RF, GBM, and MLP (Figure 3). The outer loop is 5-fold cross validation. The inner loop is 3-fold cross validation and chooses the best combination of parameters based on RMSE. After the best modeling method was selected, we used 10-time repeated 5-fold cross validation to search for the best combination of parameters of this method based on RMSE.

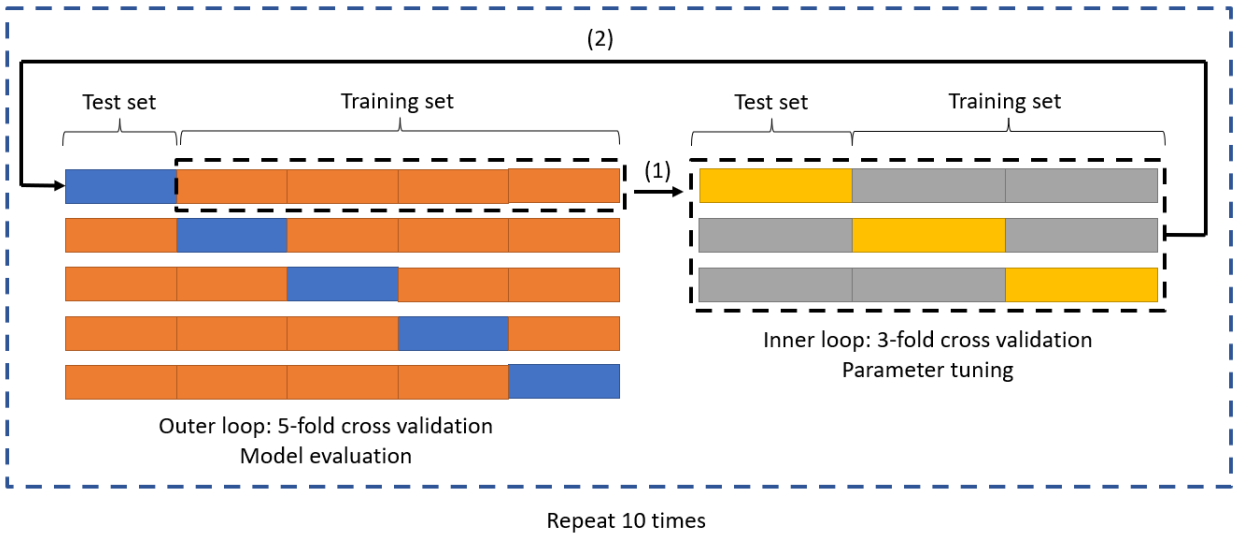

Figure 3. Repeated nested cross validation

Finally, we measured the contributions of independent variables in PCR and PCR change models. We applied SHAP (Shapley Additive Explanations) value (Chen, Lundberg, and Lee 2022), which is based on the concept of Shapely value from cooperative game theory (Shapley 1953). For each observation, the inclusion of one independent variable could have a marginal impact on the prediction of future pavement condition and this marginal impact is calculated as SHAP value. The average absolute SHAP value of one independent variable in all observations of the sample is the corresponding contribution. SHAP value has been widely used in explaining the contributions of independent variables in black-box machine learning methods (Xiao et al. 2021; Wagner et al. 2022). Furthermore, we used the

accumulated local effect (ALE) plots to show the main effect of important independent variables and interactive effect between current PCR and other important variables in PCR change model. ALE could present the marginal impact by one or two independent variables on the dependent variable after controlling for the effects of other independent variables (Apley and Zhu 2020; Molnar 2020). Compared with the other similar interpretation tools (e.g., partial dependence plots in Barua et al. (2021)), it can avoid the influence of multicollinearity among independent variables (Molnar 2020).

4 Results

4.1 Model performance

Each set of models comprises five distinct models, categorized as all-year, one-year, two-year, three-year, and four-year models. The all-year model encompasses all observations, regardless of the duration between the current and future PCR dates. In contrast, the other four models consist of specific observation subsets. In the literature, scholars usually estimated models to predict the pavement conditions in specific years. We followed and constructed the other four models. The one-year model contains observations with a time span of less than one year (365 days) between the PCR dates. The two-year model includes observations with a time period between one and two years (365-730 days). Similarly, the three-year and four-year models incorporate observations with time spans specific to their respective categories. With these four models, we can compare our model performance with those in the literature.

In Table 5, we present the results of PCR models based on repeated nested cross validation. For the all-year model, GBM has a better performance than the other three models in terms of the three measures. Note that RMSE and MAE are presented in negative form to facilitate comparison. Its R-squared, RMSE, and MAE are approximately 0.85, -0.37, and -0.28, respectively.

For the one-year model, RF has the best performance in fitting the data, with its R-squared value as around 0.90, RMSE as -0.26, and MAE as -0.19. The performance of the one-year model, as measured by R-squared, is comparable to the findings of several prior studies. For instance the artificial neural

network (ANN) models used by Lou et al. (2001) to predict the cracking index in one year have an R-squared value ranging from 0.9 to 0.91. The ANN models to predict the PCR[1] in Yang et al. (2003) exhibit R-squared values ranging from 0.79 to 0.88.

As to the two-year model, GBM is better than the other models. It has an R-squared of 0.84, an RMSE of -0.36, and an MAE of -0.27. The results are mixed when comparing our RF two-year model with similar models in other literatures. For example, the R-squared value of our model is much higher than that of the ANN model to predict PCR by Yang et al. (2003), ranging from 0.55 to 0.76. However, the ANN model by Ziari et al. (2015) to predict the international roughness index in two years has a larger R-squared (0.97) than our model.

GBM also performs better than other models in the three-year and four-year models. The GBM three-year model has an R-squared of 0.85, an RMSE of -0.39, and an MAE of -0.30. The GBM four-year model has an R squared of 0.78, an RMSE of -0.47, and an MAE of -0.36. In terms of R-squared, the two models have a better fitness than the ANN models estimated by Yang et al. (2003). The R-squared values are 0.4 to 0.48 and 0.2 to 0.38 for their three-year and four-year models, respectively.

Table 5. Model performance based on repeated nested cross validation (PCR)

Model	R-squared		RMSE		MAE	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
All-year predication (N=11963)						
LM	0.8100	0.0002	-0.4149	0.0002	-0.3295	0.0001
RF	0.8391	0.0005	-0.3817	0.0006	-0.2876	0.0005
GBM	**0.8491**	**0.0011**	**-0.3696**	**0.0012**	**-0.2790**	**0.0011**
MLP	0.8308	0.0011	-0.3914	0.0013	-0.3035	0.0011
One-year predication (N=1024)						
LM	0.8636	0.0143	-0.3053	0.0130	-0.2453	0.0022
RF	**0.8980**	**0.0031**	**-0.2644**	**0.0043**	**-0.1875**	**0.0029**
GBM	0.8960	0.0019	-0.2669	0.0024	-0.1923	0.0020
MLP	0.8737	0.0062	-0.2943	0.0055	-0.2353	0.0022
Two-year prediction (N=5763)						
LM	0.8074	0.0003	-0.3901	0.0002	-0.3146	0.0002
RF	0.8362	0.0013	-0.3595	0.0014	-0.2729	0.0013
GBM	**0.8394**	**0.0008**	**-0.3559**	**0.0010**	**-0.2713**	**0.0007**
MLP	0.8247	0.0015	-0.3721	0.0015	-0.2928	0.0013
Three-year prediction (N=3238)						
LM	0.8241	0.0008	-0.4130	0.0007	-0.3251	0.0007
RF	0.8428	0.0016	-0.3904	0.0021	-0.2955	0.0015

[1] Yang et al. (2003) defined PCR as the minimum value of crack rating, ride rating, and rut rating, which is different from the PCR in this study.

GBM	**0.8461**	**0.0015**	**-0.3864**	**0.0017**	**-0.2928**	**0.0014**
MLP	0.8414	0.0007	-0.3923	0.0009	-0.3037	0.0007
Four-year prediction (N=1938)						
LM	0.7407	0.0016	-0.5089	0.0017	-0.4025	0.0014
RF	0.7748	0.0030	-0.4743	0.0027	-0.3643	0.0020
GBM	**0.7811**	**0.0035**	**-0.4676**	**0.0035**	**-0.3602**	**0.0034**
MLP	0.7770	0.0019	-0.4717	0.0020	-0.3636	0.0020

Table 6 presents the results of PCR change models. GBM has a better performance than other models for the all-year model. However, RF performs better than other models for the one-year, two-year, three-year, and four-year models. The GBM all-year model has an R-squared of 0.34, an RMSE of -0.37, and an MAE of -0.28. The R-squared of the other four models ranges from 0.25 to 0.31. The corresponding RMSE ranges from -0.46 to -0.26. The corresponding MAE ranges from -0.35 to -0.19. In general, the fitness of PCR change models is lower than the PCR models, which is consistent with our expectation. In PCR models, most of the variance in future PCRs could be explained by the current PCRs, resulting in a higher goodness of fit.

Table 6. Model performance based on repeated nested cross validation (PCR change)

Model	R-squared		RMSE		MAE	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
All-year predication (N=11963)						
LM	0.1665	0.0008	-0.4149	0.0002	-0.3295	0.0001
RF	0.3135	0.0021	-0.3765	0.0006	-0.2881	0.0004
GBM	**0.3402**	**0.0041**	**-0.3692**	**0.0011**	**-0.2806**	**0.0013**
MLP	0.2674	0.0045	-0.3889	0.0012	-0.3010	0.0018
One-year predication (N=1024)						
LM	0.0589	0.1049	-0.3053	0.0130	-0.2453	0.0022
RF	**0.3134**	**0.0189**	**-0.2623**	**0.0034**	**-0.1924**	**0.0029**
GBM	0.2918	0.0132	-0.2659	0.0026	-0.2002	0.0024
MLP	0.1580	0.0133	-0.2903	0.0021	-0.2318	0.0015
Two-year prediction (N=5763)						
LM	0.0881	0.0010	-0.3901	0.0002	-0.3146	0.0002
RF	**0.2537**	**0.0030**	**-0.3529**	**0.0006**	**-0.2717**	**0.0005**
GBM	0.2489	0.0041	-0.3540	0.0010	-0.2731	0.0007
MLP	0.1767	0.0047	-0.3706	0.0010	-0.2915	0.0018
Three-year prediction (N=3238)						
LM	0.1214	0.0038	-0.4130	0.0007	-0.3251	0.0007
RF	**0.2492**	**0.0037**	**-0.3818**	**0.0010**	**-0.2913**	**0.0009**
GBM	0.2453	0.0086	-0.3827	0.0021	-0.2920	0.0020
MLP	0.2068	0.0076	-0.3922	0.0016	-0.3042	0.0018
Four-year prediction (N=1938)						
LM	0.0851	0.0070	-0.5089	0.0017	-0.4025	0.0014
RF	**0.2616**	**0.0060**	**-0.4573**	**0.0018**	**-0.3522**	**0.0011**
GBM	0.2519	0.0123	-0.4600	0.0035	-0.3546	0.0027

MLP	0.2229	0.0103	-0.4691	0.0030	-0.3618	0.0026

Visualization of the errors in the road network is another way to demonstrate the model performance. Due to limitation of space, we used Dormont in Pennsylvania as an example. Figure 4 and Figure 5 present the distribution of prediction error across selected periods (one year and four years) in Dormont for all-year PCR model and PCR change model, respectively. Consistent with the results in Table 5 and Table 6, as the projected period lengthens, the prediction error correspondingly rises. For instance, within the PCR model, the inaccuracies observed for projections set one year in the future are significantly less pronounced than those for projections set four years ahead. In addition, some road segments in the lower part of the area have larger errors in both PCR and PCR change models, potentially due to certain variables missing in the models.

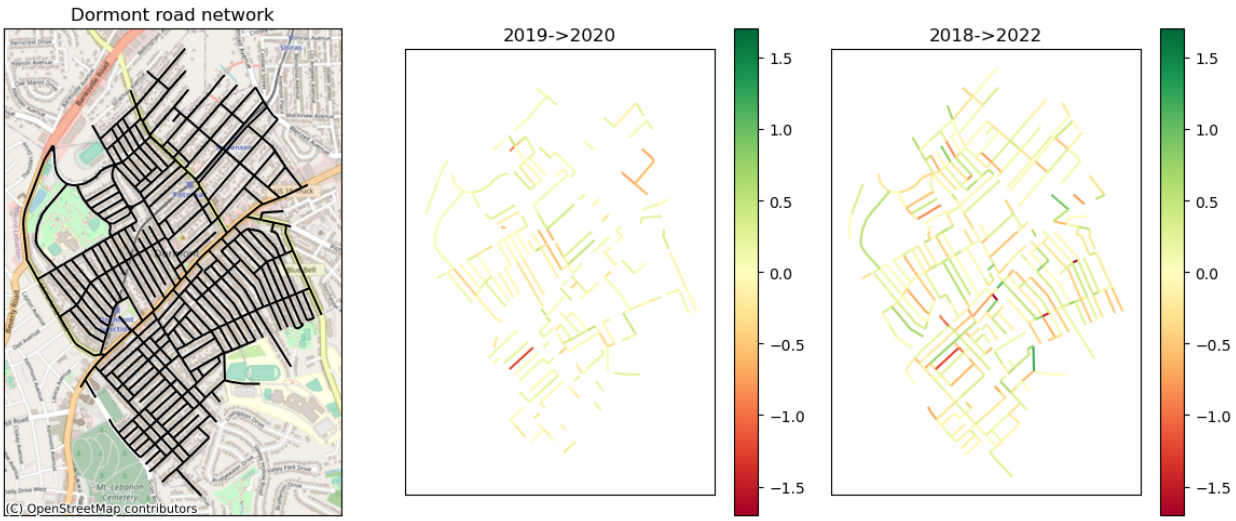

Figure 4. The distribution of prediction error for the road segments across selected periods in Dormont, PA (all-year PCR model)

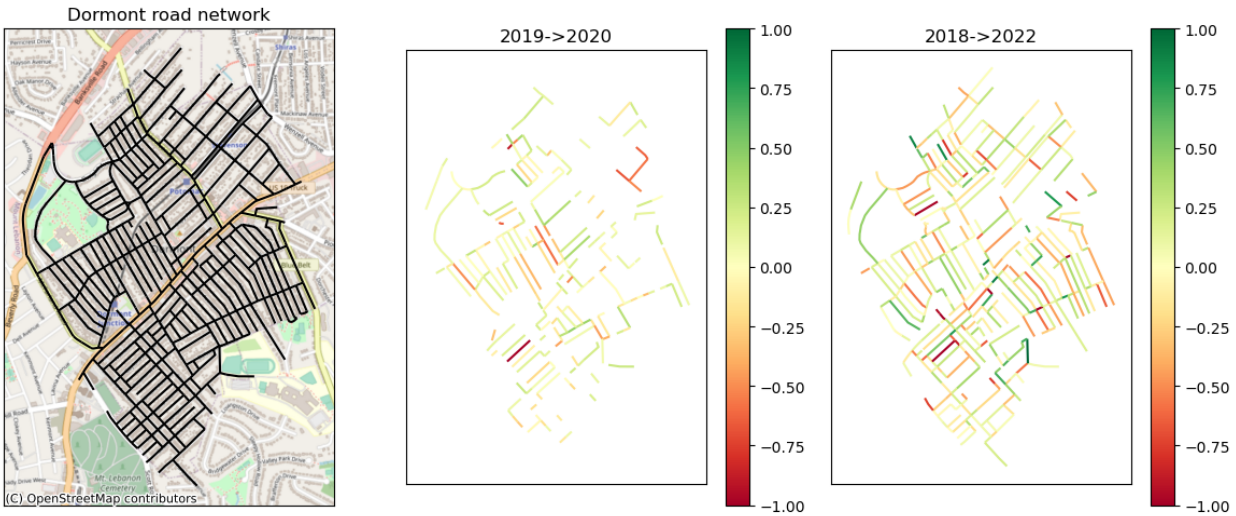

Figure 5. The distribution of prediction error for the road segments across selected periods in Dormont, PA (all-year PCR change model)

4.2 Variable contributions

We estimated the contributions of independent variables to predicting future PCR and PCR change in all-year models, respectively. Figure 6 presents the top ten independent variables based on their contributions (i.e., average absolute SHAP value) to estimating future PCR. Current PCR is the dominant factor in predicting future PCR and its contribution is much larger than other variables. This is consistent with previous studies (Piryonesi and El-Diraby 2020; Evers, Bashar, and Torres-Machi 2023; Gong et al. 2018). Its average absolute SHAP value is close to 0.7, indicating that inclusion of current PCR in the all-year model could have an average marginal impact of 0.7 on the future PCR. The two variables following current PCR are freeze thaw cycles and frozen days. Their average absolute SHAP value is 0.05 and 0.04, respectively. The other important variables include number of days, change of income, land use mix, wind speed, median household income, minimum temperature, and network density. Their contributions are similar, with the average absolute SHAP value ranging from 0.015 to 0.022. As illustrated in Figure 6, the prominent role of current PCR could bias people's perceptions of the contributions made by other factors. Statistically, current PCR is highly correlated with future PCR, and thus it accounts for most of the variance in future PCR predictions.

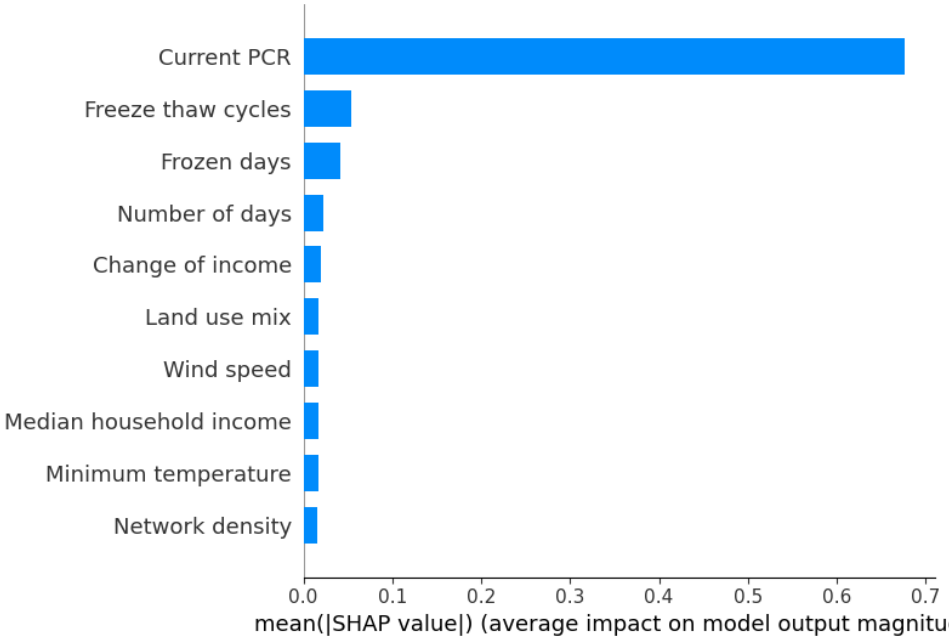

Figure 6. Average absolute SHAP value of independent variables considered in GBM all-year model to predict PCR

We present the top ten independent variables employed in the GBM all-year model, ranked based on their contributions to predicting PCR changes. Current PCR is still the most important variable based on its average absolute SHAP value (approximately 0.09). Nonetheless, in comparison to the GBM all-year PCR model, current PCR does not show the same dominant role in influencing predictions. The following four variables are frozen days (0.06), freeze thaw cycles (0.03), and number of days (0.03). The other important variables include minimum temperature, land use mix, local or other road, change of income, wind speed, and network density. They share a similar average absolute SHAP value of 0.02. Although most of these ten variables are the same as those in the all-year PCR model, their rankings are different.

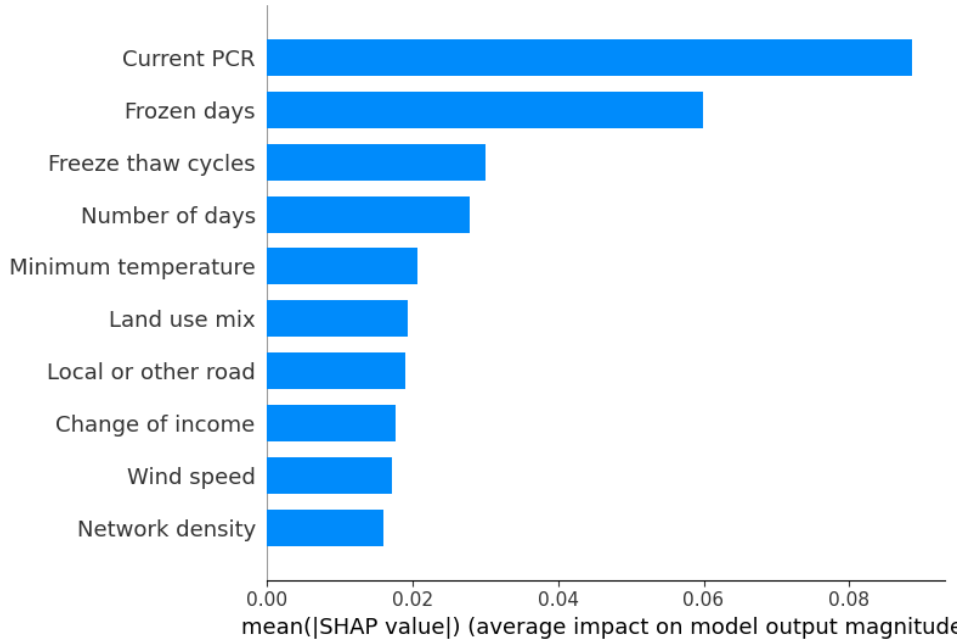

Figure 7. Average absolute SHAP value of independent variables considered in GBM all-year model to predict PCR change

4.3 Main and interactive effects

As shown above, the results of PCR model potentially bias people's impression of other variables' contribution to predicting pavement conditions and PCR change model helps fixing this issue. Therefore, when discussing the main and interactive effects of important variables on pavement conditions, we focus on PCR change model. Figure 8 presents the main effect of top seven important variables. Figure 9 presents the interactive effect between current PCR and other six important variables. Note that interactive effect does not include the main effect and only indicates the marginal effect resulting from the interaction between two independent variables. To facilitate comparison, we used the same scale for the y axes of the main effect plots and for the z axes of the interactive effect plots, respectively. In main effect plots, we used rugs on the x axes to present the distribution of the corresponding independent variables.

As shown in Figure 8, current PCR is negatively correlated with PCR change. In addition, frozen days, number of days, and freeze thaw cycles have positive relationships with PCR change. Minimum temperature and land use mix are negatively correlated with PCR change. Local or other road is

associated with higher PCR change than other road classifications. These relationships are consistent with our expectations.

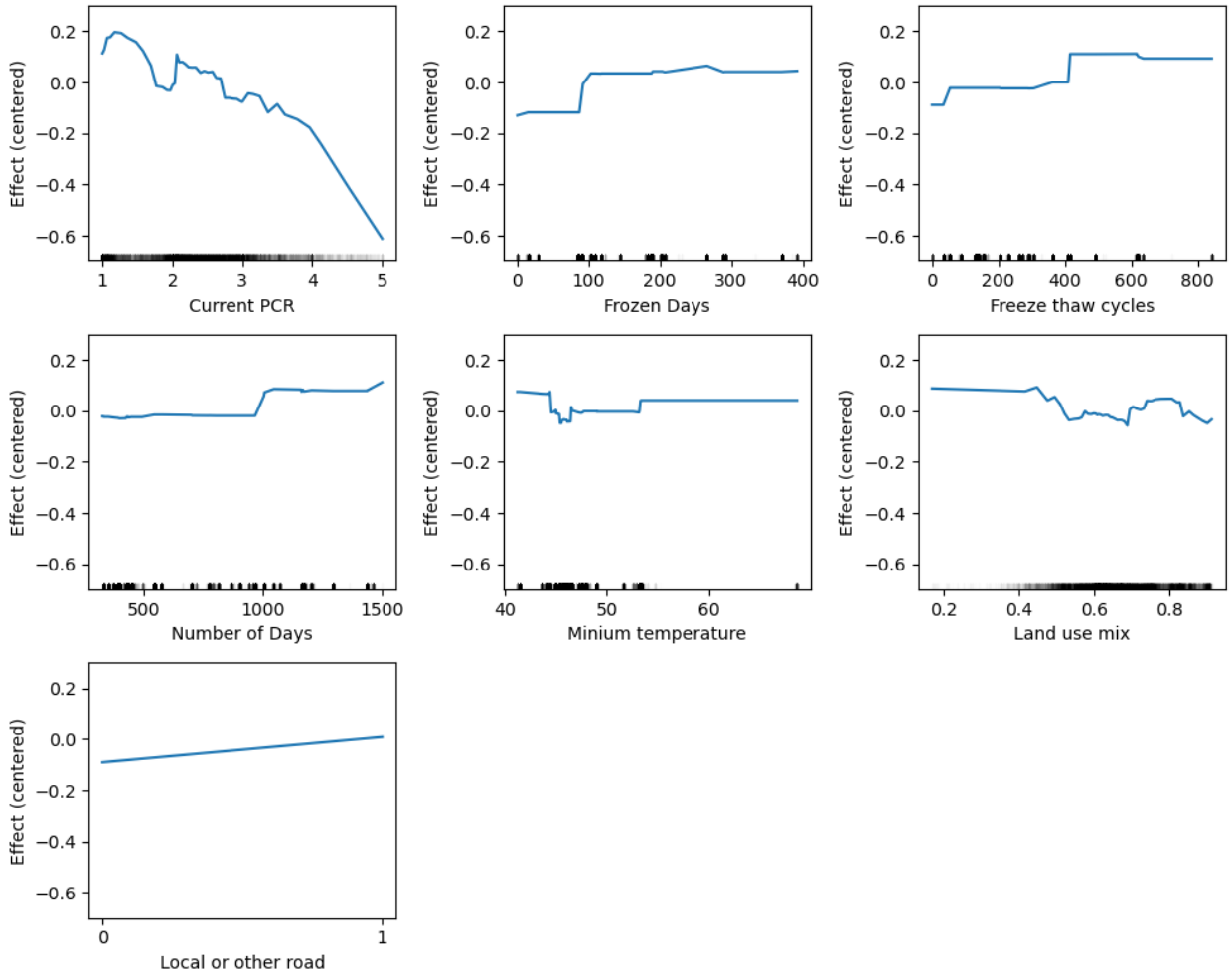

Figure 8. The main effects of important variables in all-year PCR change model

In Figure 9, current PCR and frozen days show strong interactive effects. When current PCR and frozen days are low, there is a positive impact on PCR change. When current PCR is higher than 4 and frozen days is larger than about 100 days, there is a negative impact on PCR change. As to the interactive impact between current PCR and free thaw cycles, there is a negative impact on PCR change where both current PCR and free thaw cycles are high. The interactive effect between current PCR and number of days shows a similar pattern. For other variables, their interactive effects are mostly zero or do not suggest a clear pattern.

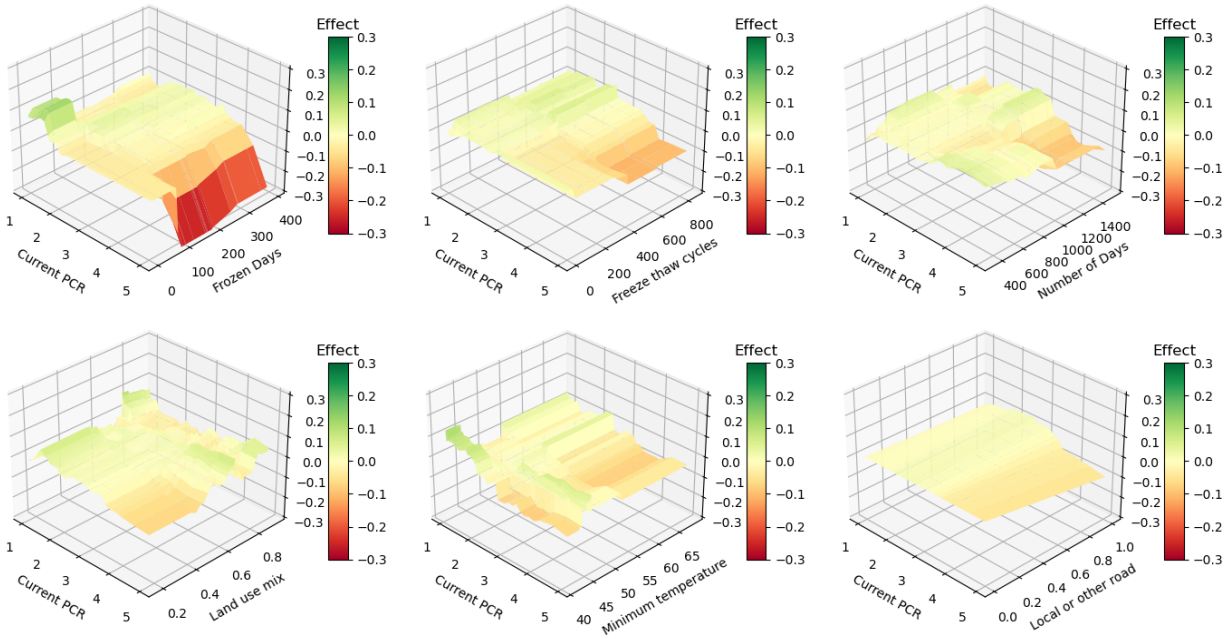

Figure 9. Interactive effects between current PCR and other important variables in all-year PCR change model

5 Conclusion

In this study, we proposed a low-cost, network-wide, and ubiquitous approach to predict the pavement condition in nine communities in the US. Specifically, we applied the pavement condition rating (PCR) data collected by AI-powered computer vision technologies from RoadBotics and several types of variables from open-sourced datasets, including road classification, weather, socioeconomics, and built environment variables. We used several machine learning models, including random forest (RF), gradient boosting machine (GBM), and multiple-layer perception (MLP), to estimate two sets of models to predict absolute PCR and PCR change, respectively. The results of this study contribute significantly to literature.

This study utilized several open-sourced datasets which cover most of the areas in the US and multiple years and the PCR models estimated in this study show a good performance in predicting future PCRs. Based on the results of repeated nested cross validation, GBM was selected for the all-year PCR model. RF was selected for the one-year PCR model and GBM was selected for the two-year, three-year, and four-year models. Compared with the similar models in the literature, the one-year model has a comparable performance, the two-year model has mixed results, and the three-year and four-year models

498 have better performance in terms of the values of R-squared. In terms of the contributions of single
499 independent variables, our all-year model shows that current PCR has a dominant role in estimating future
500 PCRs. This result is also consistent with the results in other studies.

501 We include socioeconomic variables from National Historical GIS and built environment
502 attributes from EPA Smart Location Database to account for the impact of traffic activity and pavement
503 characteristics, and some of these variables show good contributions to model performance. In both all-
504 year PCR and PCR change models, land use mix, network density, and change of income are among the
505 top ten independent variables according to their contributions. In addition, median household income is
506 among the top ten variables in all-year PCR. These results suggest that these variables are effective in
507 estimating either PCR and PCR change in addition to variables usually considered in the models such as
508 current PCR, weather, and road classification.

509 Besides the PCR models, we estimated another set of models to predict the PCR change. In PCR
510 models, current PCR explains most of the variance in future PCR and the contributions of other variables
511 become trivial. Modeling PCR change avoids this issue and helps demonstrating the contributions of
512 other variables. The all-year PCR change model indicates that, although current PCR makes the most
513 contribution, other variables including frozen days, freeze thaw cycles, and number of days have
514 substantial contributions to predicting PCR change.

515 The main and interactive effects showed that the variables not only affect PCR change
516 independently but also manifest salient effects when interacting with current PCR. The interaction
517 between current PCR and frozen days increases PCR change when both variables are low and decreases
518 PCR change when both two variables are higher than certain thresholds. There is a positive impact on
519 PCR change when both current PCR and freeze thaw cycles exceed certain thresholds. The interaction
520 between PCR change and number of days shows a similar effect. The findings suggest that frozen days
521 will be more harmful to pavement conditions when the road segments are in relatively good condition.
522 However, when the road segments are in poor condition, frozen days, freeze thaw cycles, and number of
523 days seem to have diminished impact.

This study provides three important implications for pavement management in communities. First, leveraging various datasets from open sources could address the data shortage faced by most communities. Open-source data provides a wealth of information to supplement and enhance the pavement monitoring process. Also, by leveraging open-source data, which has large spatial and temporal coverage, communities can tap into a wider range of information and improve the accuracy of their pavement assessments. This, in turn, enables better decision-making when it comes to prioritizing maintenance and repair efforts, ultimately leading to more efficient allocation of resources.

Second, this study demonstrates that the results of modeling PCR and PCR change both have their unique advantages and could complement each other. Modeling pavement condition directly is easy to interpret, and the results could be applied directly to prioritize road segments for maintenance. Modeling the change of pavement conditions, on the other hand, is more sensitive to changes in the factors and provides additional insights on how factors influence pavement conditions. Their results could help understand which factors contribute more to pavement condition deterioration and thus allocate resources to the road sections influenced most by these factors. Decision makers could benefit from both models for pavement management.

Third, pavement management policies need to consider not just the individual effects of factors, but also their interactive impacts, which helps better allocating the limited resources. For example, when pavement is in good condition, more aggressive maintenance measures could be employed to reduce their exposure to frozen environments. This could involve increasing the frequency of inspections during cold weather periods or utilizing more resistant materials during initial construction. On the other hand, when pavement conditions are poor, our results show that these variables have a diminished impact. This might suggest a more conservative use of resources. Considering the diminished effect of frozen days and freeze-thaw cycles under poor conditions, these periods might not warrant an immediate intensive repair, but rather a scheduled complete rehabilitation.

Finally, the proposed model could enhance climate resiliency and transportation equity during pavement management process. By incorporating weather variables, the model can account for extreme

climate scenarios when predicting future pavement conditions. This proactive approach helps address potential climate change-related challenges in pavement conditions. Moreover, considering the pavement condition across the entire community ensures the inclusion of all populations, resulting in more equitable outcomes in pavement management decision making.

This study presents two limitations. Firstly, factors such as age and maintenance history are not incorporated into the analysis due to the unavailability of related datasets for the selected communities. Since including these factors could further enhance model performance, we encourage future studies to consider them when the necessary datasets become available. Secondly, we combined observations from all communities into one model rather than modeling them individually. This decision was made because the temporal coverage of PCR data varied between communities, making separate models incomparable and unable to provide additional insights. For future research, we recommend developing individual models when the temporal data coverage is consistent across all communities.

6 Data Availability Statement

All data utilized in this study have been responsibly attributed to their respective sources in the article. At the same time, we shared the processed datasets and codes through the GitHub repository: https://github.com/vtao1989/AssetPerformance. The pavement condition data were generously provided by RoadBotics. It's important to note that these raw data cannot be shared with third parties.

7 Acknowledgement

This project was funded by Pennsylvania Infrastructure Technology Alliance (PITA). The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein.

References

Apley, Daniel W., and Jingyu Zhu. 2020. "Visualizing the effects of predictor variables in black box supervised learning models." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82 (4):1059-1086. doi: 10.1111/rssb.12377.

Attoh-Okine, Nii, and Offei Adarkwa. 2013. Pavement Condition Surveys – Overview of Current Practices. Newark, Delaware: Delaware Center for Transportation.

Barua, Limon, Bo Zou, Mohamadhossein Noruzoliaee, and Sybil Derrible. 2021. "A gradient boosting approach to understanding airport runway and taxiway pavement deterioration." *International Journal of Pavement Engineering* 22 (13):1673-1687. doi: 10.1080/10298436.2020.1714616.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45 (1):5-32. doi: 10.1023/a:1010933404324.

Chen, Hugh, Scott M. Lundberg, and Su-In Lee. 2022. "Explaining a series of models by propagating Shapley values." *Nature Communications* 13 (1). doi: 10.1038/s41467-022-31384-3.

Elkins, Gary E., and Barbara Ostrom. 2021. Long-Term Pavement Performance Information Management System User Guide. Federal Highway Administration.

Evers, Elijah, Mohammad Z. Bashar, and Cristina Torres-Machi. 2023. "Improving the Prediction of Full-Depth Reclamation Pavements' Performance using Random Forests." Transportation Research Board Annual Conference, Washington, DC.

Friedman, Jerome H. 2001. "Greedy function approximation: A gradient boosting machine." *The Annals of Statistics* 29 (5):1189-1232. doi: 10.1214/aos/1013203451.

Friedman, Jerome H. 2002. "Stochastic gradient boosting." *Computational Statistics & Data Analysis* 38 (4):367-378. doi: 10.1016/s0167-9473(01)00065-2.

George, K P, A S Rajagopal, and L K Lim. 1989. "Models for predicting pavement deterioration." *Transportation Research Record: Journal of the Transportation Research Board*.

Gong, Hongren, Yiren Sun, Xiang Shu, and Baoshan Huang. 2018. "Use of random forests regression for predicting IRI of asphalt pavements." *Construction and Building Materials* 189:890-897. doi: 10.1016/j.conbuildmat.2018.09.017.

Hossain, M. I., L. S. P. Gopisetti, and M. S. Miah. 2017. "Prediction of International Roughness Index of Flexible Pavements from Climate and Traffic Data Using Artificial Neural Network Modeling." Airfield and Highway Pavements 2017, 2017-08-24.

Kargah-Ostadi, Nima, Shelley M. Stoffels, and Nader Tabatabaee. 2010. "Network-Level Pavement Roughness Prediction Model for Rehabilitation Recommendations." *Transportation Research Record: Journal of the Transportation Research Board* 2155 (1):124-133. doi: 10.3141/2155-14.

Lou, Z., M. Gunaratne, J. J. Lu, and B. Dietrich. 2001. "Application of Neural Network Model to Forecast Short-Term Pavement Crack Condition: Florida Case Study." *Journal of Infrastructure Systems* 7 (4):166-171. doi: 10.1061/(asce)1076-0342(2001)7:4(166).

Molnar, Christoph. 2020. "Interpretable Machine Learning - A Guide for Making Black Box Models Explainable." In: lulu.com. https://christophm.github.io/interpretable-ml-book/.

Piryonesi, S. Madeh, and Tamer El-Diraby. 2021. "Climate change impact on infrastructure: A machine learning solution for predicting pavement condition index." *Construction and Building Materials* 306. doi: 10.1016/j.conbuildmat.2021.124905.

Piryonesi, S. Madeh, and Tamer E. El-Diraby. 2020. "Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index." *Journal of Infrastructure Systems* 26 (1). doi: 10.1061/(asce)is.1943-555x.0000512.

RoadBotics. 2023. "About RoadBotics by Michelin." https://www.roadbotics.com/company-about/.

Shapley, Lloyd S. 1953. "A value for n-person games." In *Contributions to the Theory of Games, Volume II, Annals of Mathematics Studies*, 307-317. Princeton University Press.

Sholevar, Nima, Amir Golroo, and Sahand Roghani Esfahani. 2022. "Machine learning techniques for pavement condition evaluation." *Automation in Construction* 136. doi: 10.1016/j.autcon.2022.104190.

Sidess, Arieh, Amnon Ravina, and Eyal Oged. 2020. "A model for predicting the deterioration of the pavement condition index." *International Journal of Pavement Engineering* 22 (13):1625-1636. doi: 10.1080/10298436.2020.1714044.

623 Wagner, Felix, Nikola Milojevic-Dupont, Lukas Franken, Aicha Zekar, Ben Thies, Nicolas Koch, and Felix
624 Creutzig. 2022. "Using explainable machine learning to understand how urban form shapes
625 sustainable mobility." *Transportation Research Part D: Transport and Environment* 111. doi:
626 10.1016/j.trd.2022.103442.
627 Xiao, Longzhu, Siuming Lo, Jixiang Liu, Jiangping Zhou, and Qingqing Li. 2021. "Nonlinear and synergistic
628 effects of TOD on urban vibrancy: Applying local explanations for gradient boosting decision
629 tree." *Sustainable Cities and Society* 72. doi: 10.1016/j.scs.2021.103063.
630 Xu, Yang, and Zhanmin Zhang. 2022. "Review of Applications of Artificial Intelligence Algorithms in
631 Pavement Management." *Journal of Transportation Engineering, Part B: Pavements* 148 (3). doi:
632 10.1061/jpeodx.0000373.
633 Yang, Jidong, Jian John Lu, Manjriker Gunaratne, and Qiaojun Xiang. 2003. "Forecasting Overall
634 Pavement Condition with Neural Networks: Application on Florida Highway Network."
635 *Transportation Research Record: Journal of the Transportation Research Board* 1853 (1):3-12.
636 doi: 10.3141/1853-01.
637 Ziari, Hasan, Jafar Sobhani, Jalal Ayoubinejad, and Timo Hartmann. 2015. "Prediction of IRI in short and
638 long terms for flexible pavements: ANN and GMDH methods." *International Journal of*
639 *Pavement Engineering* 17 (9):776-788. doi: 10.1080/10298436.2015.1019498.
640