TTI: 0-7177

# Enhancing Pavement Safety in Texas: Crash Susceptibility Modeling, GIS Tool, and Hydroplaning Risk Assessment

Technical Report 0-7177-R1

Cooperative Research Program

TEXAS A&M TRANSPORTATION INSTITUTE
COLLEGE STATION, TEXAS

sponsored by the
Federal Highway Administration and the
Texas Department of Transportation
https://tti.tamu.edu/documents/0-7177-R1.pdf

Technical Report Documentation Page

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| FHWA/TX-26/0-7177-R1 | | |

| 4. Title and Subtitle | 5. Report Date |
|---|---|
| ENHANCING PAVEMENT SAFETY IN TEXAS: CRASH SUSCEPTIBILITY MODELING, GIS TOOL, AND HYDROPLANING RISK ASSESSMENT | Published: July 2026 |
| | 6. Performing Organization Code |
| 7. Author(s)<br>Nasir G. Gharaibeh, Lu Gao, Boni Kutela, Yussuf Tarnini, Mohammad Talebzadeh, and Tejaswini Katale | 8. Performing Organization Report No.<br>Report 0-7177-R1 |
| 9. Performing Organization Name and Address<br>Texas A&M Transportation Institute<br>The Texas A&M University System<br>College Station, Texas 77843-3135 | 10. Work Unit No. (TRAIS)<br><br>11. Contract or Grant No.<br>Project 0-7177 |
| 12. Sponsoring Agency Name and Address<br>Texas Department of Transportation<br>Research and Technology Implementation Office<br>125 E. 11th Street<br>Austin, Texas 78701-2483 | 13. Type of Report and Period Covered<br>Technical Report:<br>September 2023–April 2026<br><br>14. Sponsoring Agency Code |

15. Supplementary Notes

Project sponsored by the Texas Department of Transportation and the Federal Highway Administration.
Project Title: Identifying Pavement Improvement Projects for Enhanced Safety
URL: https://tti.tamu.edu/documents/0-7177-R1.pdf

16. Abstract

Traffic crashes and the associated injuries and fatalities remain a major public safety concern in Texas and across the United States. In this project, researchers developed statistical models to assess pavement susceptibility to crashes (i.e., the probability of crash occurrence), considering various pavement conditions, as well as traffic volume, truck percentage, road geometry, and precipitation. The models help determine the effects of these variables on the probability of crash occurrence, which can aid in proactive interventions to improve road safety. Furthermore, researchers analyzed the water film thickness at the time and location of wet-pavement crashes to improve the understanding of the effects of pavement-surface-related factors on the occurrence of crashes caused by hydroplaning. The project developed a geographic information system tool to visualize the safety-related pavement conditions, crash susceptibility predictions, and candidate pavement sections for pavement safety improvement. The project used a large dataset from across Texas consisting of 18,843 on-system roadbed-miles and 24,873 crashes that occurred on them during a 3-year period (Fiscal Year 2021–2023). Finally, researchers developed conclusions and recommendations to support proactive consideration of safety in pavement management decisions.

| 17. Key Words | 18. Distribution Statement |
|---|---|
| Pavement Condition, Crash Susceptibility, Pavement Management, Hydroplaning, Water Film Thickness, Skid Number, International Roughness Index | No restrictions. This document is available to the public through NTIS:<br>National Technical Information Service<br>Alexandria, Virginia<br>https://www.ntis.gov |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 98 | |

**Form DOT F 1700.7 (8-72) Reproduction of completed page authorized**

**ENHANCING PAVEMENT SAFETY IN TEXAS: CRASH SUSCEPTIBILITY MODELING, GIS TOOL, AND HYDROPLANING RISK ASSESSMENT**

by

Nasir G. Gharaibeh, Ph.D., P.E.
Professor and Senior Research Engineer
Texas A&M Transportation Institute

Lu Gao, Ph.D.
Professor
University of Houston

Boni Kutela, Ph.D.
Associate Research Engineer
Texas A&M Transportation Institute

Yussuf Tarnini
Graduate Assistant
Texas A&M Transportation Institute

Mohammad Talebzadeh
Graduate Assistant
Texas A&M Transportation Institute

and

Tejaswini Katale
Graduate Assistant
University of Houston

Report 0-7177-R1B
Project 0-7177
Project Title: Identifying Pavement Improvement Projects for Enhanced Safety

Sponsored by the
Texas Department of Transportation
and the
Federal Highway Administration

Published: July 2026

TEXAS A&M TRANSPORTATION INSTITUTE
College Station, Texas 77843-3135

# DISCLAIMER

# ACKNOWLEDGMENTS

This project was sponsored by TxDOT and FHWA. The authors thank the project manager (Jadé Adediwura) and members of the Project Monitoring Committee (Jenny Li, Lacy Peters, Hui Wu, Taehoon Lim, and James Harris).

# TABLE OF CONTENTS

# LIST OF FIGURES

**Page**

# LIST OF TABLES

**Page**

# CHAPTER 1: INTRODUCTION

Traffic crashes and their associated injuries and fatalities remain a major public safety concern in Texas and across the United States. Crashes are complex events influenced by a wide range of factors related to the driver, vehicle, roadway, traffic stream, and weather. Understanding how these factors contribute to crash occurrence is essential for identifying effective safety improvements.

This project focused on evaluating the impact of pavement-related factors on crash occurrence to support the Texas Department of Transportation (TxDOT) in identifying candidate sites for pavement safety improvements. Pavement surface condition factors are of particular interest to highway agencies because these factors fall directly within their control and can be addressed through maintenance and rehabilitation treatments.

While several studies have attempted to explain the relationship between pavement condition and crash occurrence, their findings have been inconsistent. A large-scale study is needed to assess pavement susceptibility to crashes by considering multiple pavement condition indicators along with other influencing factors, using a statewide, multiyear dataset. Accordingly, the objectives of this project were to develop predictive models for assessing pavement susceptibility to crash occurrence and an online geographic information system (GIS) tool to identify candidate pavement sections for safety improvement.

The datasets used in this project included the Pavement Management Information System (PMIS) for pavement condition, the Crash Records Information System (CRIS) for crash data, TxDOT Roadway Inventory for traffic-related factors, and weather station and National Oceanic and Atmospheric Administration (NOAA) records for precipitation data. Researchers developed statistical models based on pavement type to evaluate the impacts of pavement condition and other factors on crash susceptibility. Moreover, researchers used existing empirical models to estimate the water film thickness (WFT) at the time and location of crashes to evaluate its effect on crashes that occurred due to hydroplaning. Researchers also developed a web-based GIS tool to visualize crash susceptibility predictions and related data at segment, district, and statewide levels, facilitating the identification of candidate pavement sections for safety improvement.

Following this introductory information, subsequent chapters in this report include the following:

- Chapter 2: Literature Review.
- Chapter 3: Data Assembly.
- Chapter 4: Development of Crash Susceptibility Models.
- Chapter 5: Hydroplaning Risk Analysis.
- Chapter 6: Interactive GIS Map Tool for Safety Analysis.
- Chapter 7: Conclusions and Recommendations.

# CHAPTER 2: LITERATURE REVIEW

This chapter summarizes the literature review findings relevant to this project. This chapter first describes existing WFT prediction models used to determine the effect of pavement condition (particularly surface texture) and roadway geometric characteristics (e.g., pavement surface slope and length of flow path) on the WFT and ultimately the susceptibility to hydroplaning. This chapter next describes the effects of other pavement conditions (e.g., roughness, rutting, and skid resistance) and geometric characteristics on crashes. Finally, this chapter concludes with an overview of statistical modeling techniques used in the context of road safety.

## WFT AND HYDROPLANING

### Factors Affecting WFT

Roadway hydroplaning, which could trigger pavement-related crashes, occurs when a layer of water separates the vehicle tire from the pavement surface. As the WFT increases, the risk of hydroplaning increases. The increased uplift forces on the vehicle tires lead to a separation of a portion of the tires from the pavement and a loss of the steering and traction force (Pourhassan et al., 2022). The development and thickness of the water film depend primarily on the rainfall intensity and duration and the pavement geometric and surface characteristics. This section describes these factors and provides a sensitivity analysis of existing WFT prediction models.

The intensity and duration of rainfall play a significant role in water film development. Heavy rainfall increases the accumulated water on the pavement surface, whereas prolonged rainfall extends the rainwater drainage time, leading to increased WFT (Luo et al., 2019). In Texas, rainfall intensity and duration vary greatly by geographic location. For example, for a 2-year, 10-minute storm, the rainfall intensity is estimated (with 90 percent confidence) to be 2.35–4.00 inches/hour near the city center of El Paso, 3.73–6.41 inches/hour near the city center of Atlanta, 4.24–7.26 inches/hour near the city center of Houston, and 4.11–7.03 inches/hour near the city center of Corpus Christi (see Figure 1) (NOAA, 2023). Therefore, it is important for pavement safety improvement programs in Texas to account for such large variability in climatic conditions across the state.

The net WFT that contributes to hydroplaning is the total WFT minus the mean texture depth (MTD). The MTD depends on the macrotexture of the pavement, which is primarily influenced by the pavement's coarse aggregates (Anderson et al., 1998), and the texturing method (Hall et al., 2009). Increasing the pavement surface texture decreases the WFT by increasing the water reservoir inside the pavement texture and improves drainage by providing flow paths for water out of the pavement (Pourhassan et al., 2022). Therefore, highly textured pavements have a lower WFT. Table 1 shows typical MTD ranges for different pavement types. Generally, these ranges are for aged pavements.


**POINT PRECIPITATION FREQUENCY (PF) ESTIMATES**
**WITH 90% CONFIDENCE INTERVALS AND SUPPLEMENTARY INFORMATION**
**NOAA Atlas 14, Volume 11, Version 2**

| PDS-based precipitation frequency estimates with 90% confidence intervals (in inches/hour)[1] | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Duration | Average recurrence interval (years) | | | | | | | | | |
| | 1 | 2 | 5 | 10 | 25 | 50 | 100 | 200 | 500 | 1000 |
| 5-min | **5.77** (4.38-7.63) | **6.79** (5.17-8.86) | **8.41** (6.41-11.1) | **9.78** (7.34-13.0) | **11.7** (8.52-16.1) | **13.2** (9.36-18.6) | **14.7** (10.2-21.3) | **16.3** (10.9-24.2) | **18.3** (11.9-28.2) | **19.9** (12.6-31.4) |
| 10-min | **4.58** (3.47-6.05) | **5.39** (4.11-7.03) | **6.70** (5.10-8.79) | **7.79** (5.85-10.4) | **9.33** (6.81-12.9) | **10.6** (7.49-14.9) | **11.8** (8.11-17.1) | **13.0** (8.70-19.3) | **14.5** (9.41-22.3) | **15.6** (9.89-24.7) |
| 15-min | **3.88** (2.94-5.12) | **4.54** (3.46-5.93) | **5.62** (4.28-7.38) | **6.52** (4.90-8.69) | **7.78** (5.67-10.7) | **8.77** (6.22-12.4) | **9.77** (6.74-14.2) | **10.8** (7.24-16.0) | **12.1** (7.87-18.7) | **13.1** (8.32-20.8) |
| 30-min | **2.76** (2.09-3.65) | **3.22** (2.46-4.21) | **3.97** (3.02-5.22) | **4.60** (3.45-6.13) | **5.47** (3.98-7.51) | **6.14** (4.35-8.67) | **6.83** (4.71-9.90) | **7.55** (5.07-11.2) | **8.52** (5.54-13.1) | **9.28** (5.87-14.7) |
| 60-min | **1.81** (1.37-2.38) | **2.12** (1.62-2.77) | **2.63** (2.00-3.45) | **3.06** (2.30-4.08) | **3.66** (2.66-5.02) | **4.12** (2.92-5.81) | **4.60** (3.17-6.67) | **5.12** (3.44-7.62) | **5.84** (3.80-9.01) | **6.42** (4.06-10.1) |

Source: NOAA's Precipitation Frequency Data Server (*https://hdsc.nws.noaa.gov/pfds/*).

**Figure 1. Estimated Rainfall Intensity for a 2-Year, 10-Minute Storm in the City of Corpus Christi.**

**Table 1. MTD Ranges for Different Pavement Types as Reported in the Literature.**

| Pavement Type | Typical MTD Range (inches) | Source |
|---|---|---|
| Conventional Portland cement concrete | 0.012–0.018 | Hall et al., 2009 |
| Exposed aggregate concrete | 0.035–0.078 | Hall et al., 2009 |
| Dense graded hot-mix asphalt | 0.016–0.024 | Merritt et al., 2015 |
| Open graded friction course | 0.059–0.118 | Merritt et al., 2015 |
| Seal coat | 0.039–0.118 | Buss et al., 2016 |

MTD = mean texture depth.

The flow path length of a pavement surface is defined as the path determined by the pavement surface slope. Therefore, the maximum flow path length is the longest distance a raindrop can flow between the point of contact with the pavement surface and the point of exit from the pavement (Anderson et al., 1998). The presence of a continuous longitudinal gradient increases the flow path length, and consequently, the WFT along the low end of the road (Gallaway et al., 1971). When the pavement flow path slope is steep enough, the water can drain rapidly, avoiding the development of water film on the pavement surface. However, the steepness of the flow path should not be so excessive as to cause a safety hazard. Generally, the TxDOT *Roadway Design Manual* (TxDOT, 2024) recommends a pavement cross slope between 2 and 4 percent (for roads with three or more lanes in one direction in high rainfall intensity locations). This design manual also indicates that longitudinal grades of up to 15 percent may be used under highly constrained conditions, but flatter grades should be used where practicable, with grades less than 5 percent being preferable.

**Chronological Review of WFT Model Development**

Systematic study of roadway hydroplaning began in the mid-20th century, with Ross and Russam's (1968) development of the United Kingdom Road Research Laboratory (UK RRL) model. Their experiments showed that WFT increased with rainfall intensity and drainage length, especially on flat, wide pavements, although the model did not account for surface texture.

Around the same time, Horne's National Aeronautics and Space Administration research introduced a hydroplaning equation, which related critical speed to tire inflation pressure, and classified three types of hydroplaning—viscous, dynamic, and tread rubber reversion (Horne, 1968). By conducting controlled tests on seven bias-ply tires with different tread patterns but equal groove volume in a drum facility, Gengenbach (1967) found that bald tires hydroplaned at water depths as low as 0.3 mm (0.01 inches), whereas treaded tires required about 2 mm (0.08 inches), even on very smooth surfaces (0.13–0.20 mm [0.005–0.008 inches] of macrotexture depth). On highways, however, dynamic hydroplaning of minimally treaded tires (tread depth < 1.6 mm [0.06 inches]) was not observed below a water depth of 1.5 mm (0.06 inches) (Gallaway et al. 1979).

Building upon this foundation, Gallaway et al. (1979) carried out large-scale experiments in Texas to refine WFT prediction by explicitly incorporating pavement macrotexture. Researchers introduced empirical equations linking WFT to rainfall intensity, drainage length, cross slope, and texture depth and also developed a hydroplaning speed model, which showed that greater water depth, worn tread, and low tire pressure significantly reduced hydroplaning speed. A key outcome was the identification of a 2.4-mm (0.095-inch) WFT as the threshold between incipient and full hydroplaning. Huebner et al. (1986) later confirmed this threshold by reconciling Gallaway's Texas data with Agrawal's Pennsylvania study (Agrawal et al., 1977) through a unified regression, showing that 2.4 mm (≈ 0.095–0.10 inches) is the practical breakpoint.

In the late 1990s, the PAVDRN computer tool was developed as part of National Cooperative Highway Research Program (NCHRP) Project 1-29 (Anderson et al., 1998). PAVDRN takes inputs of rainfall intensity, pavement geometry (e.g., cross slope and grade), and surface characteristics (e.g., macrotexture and permeability) to estimate the transverse profile of the WFT and predict hydroplaning speed. It incorporates Gallaway's WFT equation (Gallaway et al., 1979) alongside a dual-regime hydroplaning speed model, using Gallaway's formula for thicker water films and a modified equation for thinner films to ensure continuity. Calibrated with field and laboratory data, PAVDRN proved effective in reproducing observed hydroplaning conditions.

A subsequent evaluation by Jayasooriya and Gunaratne (2014) applied PAVDRN to historical crash data in Florida and confirmed its ability to identify likely hydroplaning crashes under the right rainfall conditions. Researchers in that study compiled a hydroplaning crash database for Florida's interstate highways (2006–2011) and applied multiple WFT and hydroplaning speed models to each crash scenario to assess predictive accuracy.

To further enhance predictive capabilities, the Florida Department of Transportation (FDOT) developed its hydroplaning prediction tool in 2012 (Gunaratne et al., 2012). This tool was later upgraded by Lee and Ayyala (2020) with additional features, including support for continuous roadway geometry inputs (e.g., cross slope, longitudinal grade, and rut depth), batch processing of multiple scenarios, and GIS-based risk mapping. The enhanced FDOT tool integrates four WFT models—Gallaway's, the UK RRL equation, a New Zealand modified equation, and PAVDRN—to provide comprehensive water depth estimates.

Recent studies have improved hydroplaning prediction by addressing the limitations of older models and considering modern pavement surfaces. For example, Pourhassan et al. (2022) developed new empirical WFT equations for surfaces with very high macrotexture (e.g., chip seals). Using laboratory rainfall simulations, they found that conventional models greatly underestimated the effect of macrotexture on reducing WFT. Their revised equations introduced enhanced texture terms and accounted for porous friction course pavements, yielding improved prediction accuracy.

Meanwhile, advances in vehicle dynamics modeling have enabled more nuanced assessments of hydroplaning. Moving beyond a single hydroplaning speed threshold, NCHRP Project 15-55 (National Academies of Sciences, Engineering, and Medicine, 2021) introduced a "performance margin" metric to quantify how close a vehicle is to hydroplaning under specific conditions of water depth, speed, and required traction (during braking or cornering maneuvers). That NCHRP study effort produced a Hydroplaning Potential Assessment Tool that integrates pavement surface hydrology, finite-element tire–pavement interaction analysis, and full-vehicle dynamics simulations to identify road segments with elevated hydroplaning risk.

## Sensitivity Analysis of WFT Prediction Models

This section analyzes the sensitivity of four different WFT prediction models—Gallaway et al. (1979) (Equation 1), PAVDRN (Anderson et al., 1998) (Equation 2), Pourhassan et al. (2022) (Equation 3), and the UK RRL model (Ross & Russam, 1968) (Equation 4):

$$WFT = 0.003726 \cdot (MTD^{0.125} L^{0.519} I^{0.562} S^{-0.364}) - MTD \tag{1}$$

$$WFT = (nLI/36.1S^{0.5})^{0.6} - MTD \tag{2}$$

$$WFT = 0.00639(MTD^{-0.46} L^{0.29} I^{0.37} S^{-0.29}) - MTD \tag{3}$$

$$WFT = 0.046 \cdot (L^{0.5} I^{0.5} S^{-0.2}) - MTD \tag{4}$$

The WFT predicted by these models is a function of the drainage length (*L*), rainfall intensity (*I*), drainage slope (*S*), and *MTD*. The PAVDRN model includes Manning's roughness coefficient (*n*), which considers the hydraulic effect of the pavement surface on the flowing water. For Equations 1 and 3, $L$ is in feet, $I$ is in inches/hour, and *MTD* is in inches. For Equations 2 and 4, *L* is in meters, *I* is in mm/hour, and *MTD* is in mm. *S* is dimensionless for all equations.

Figure 2 through Figure 4 show the WFT sensitivity to changes in rainfall intensity, cross slope, and flow path length for each of the four models for MTD values of 0.02, 0.04, and 0.08 inches, respectively. For each MTD value, researchers considered nine different combinations of rainfall intensities (2, 4, and 6 inches/hour), cross slopes (1, 2, and 3 percent), and flow path lengths (0, 10, and 20 feet). For PAVDRN, the Manning's $n$ value was assumed to be 0.016 for all cases.

Increasing the cross slope and MTD decreased the WFT, while increasing the flow path length and rainfall intensity increased the WFT. Notably, the rate of increase in the WFT decreased as flow path length increased across all models; it was higher for a flow path length of 0 to 10 feet than a flow path length of 10 to 20 feet. Similarly, the rate of decrease in the WFT was lower at higher cross slopes. With all other parameters fixed, the decrease in the WFT was more apparent when the cross slope increased from 1 to 2 percent than when it increased from 2 to 3 percent. All models showed that rainfall intensity has a significant effect on the WFT; any increase in this parameter considerably impacted the WFT.

The effect of MTD on the WFT was more pronounced at lower rainfall intensities and higher cross slopes. This trend was less evident for Pourhassan model, which is most sensitive to the MTD values. For this model, even at higher rainfall intensities and lower cross slopes, changes in MTD had a significant effect on the WFT. The Gallaway, PAVDRN, and UK RRL models are more sensitive to rainfall intensity than the Pourhassan model. When all other variables were held constant, an increase in rainfall intensity resulted in a larger percentage increase in the WFT predicted by these three models than by the Pourhassan model (i.e., the percentage change in the WFT was smaller for the Pourhassan model than for the other models).

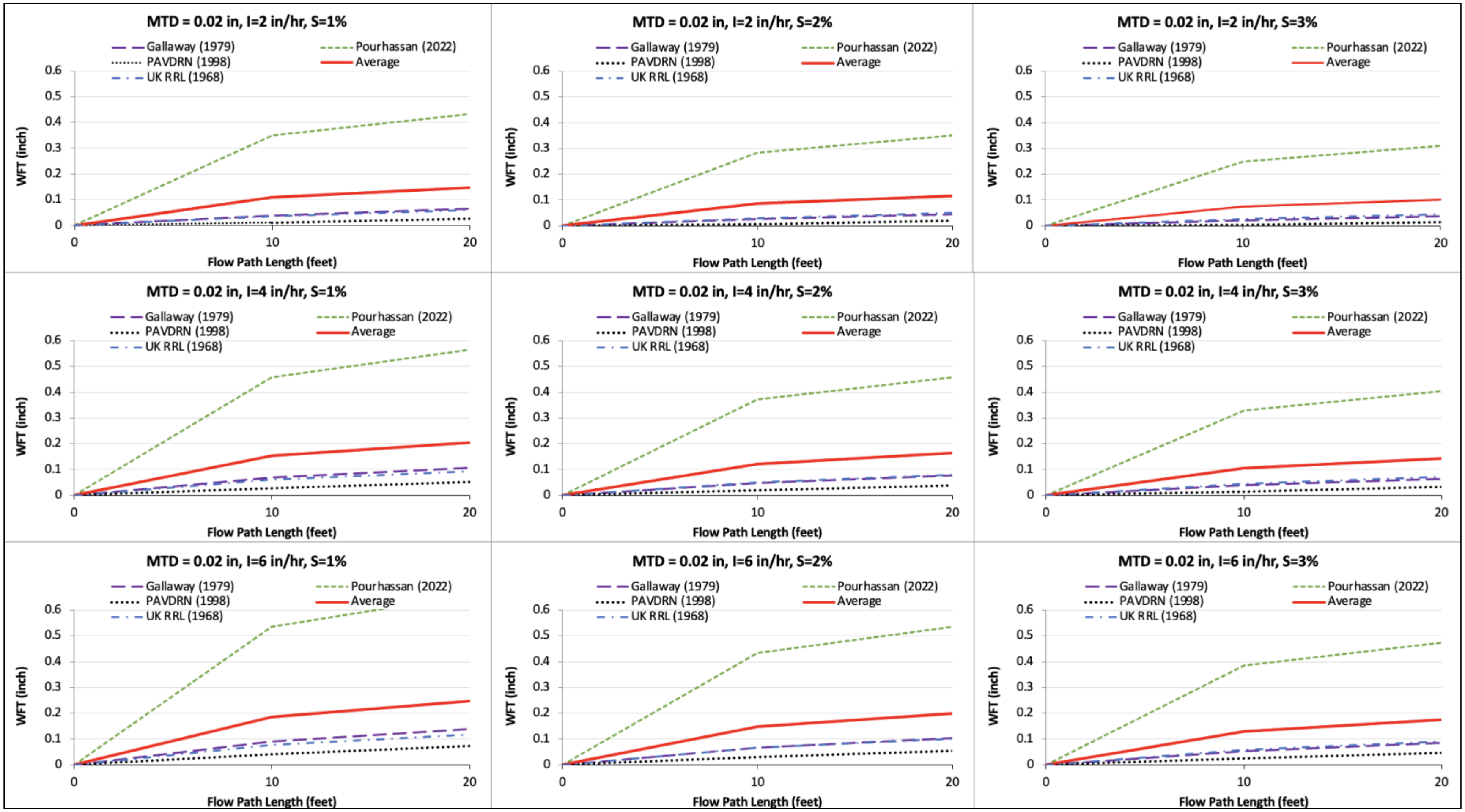



**Figure 2. Sensitivity of WFT to Changes in Rainfall Intensity, Cross Slope, and Flow Path Length—MTD = 0.02 Inches.**

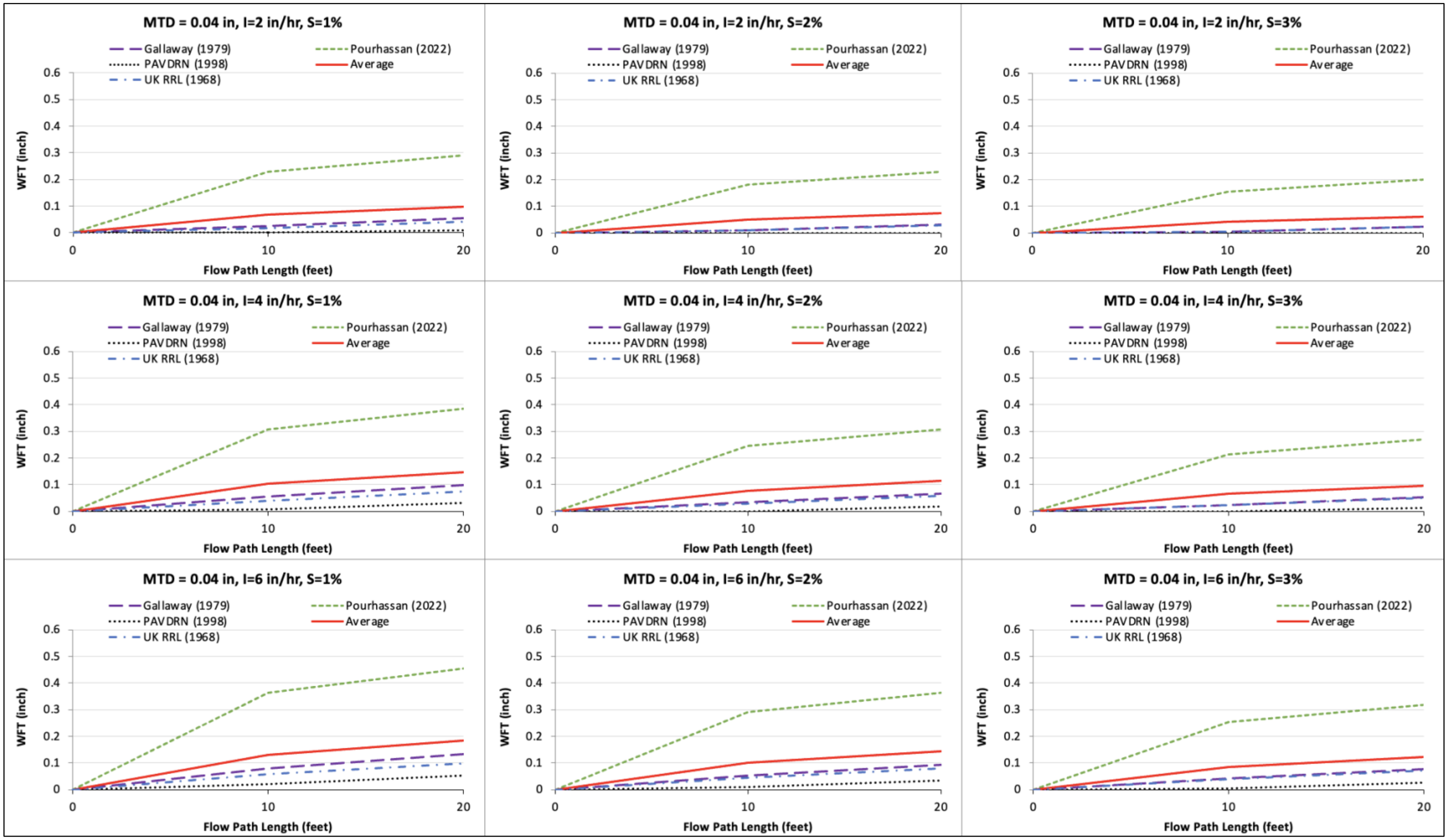



**Figure 3. Sensitivity of WFT to Changes in Rainfall Intensity, Cross Slope, and Flow Path Length—MTD = 0.04 Inches.**

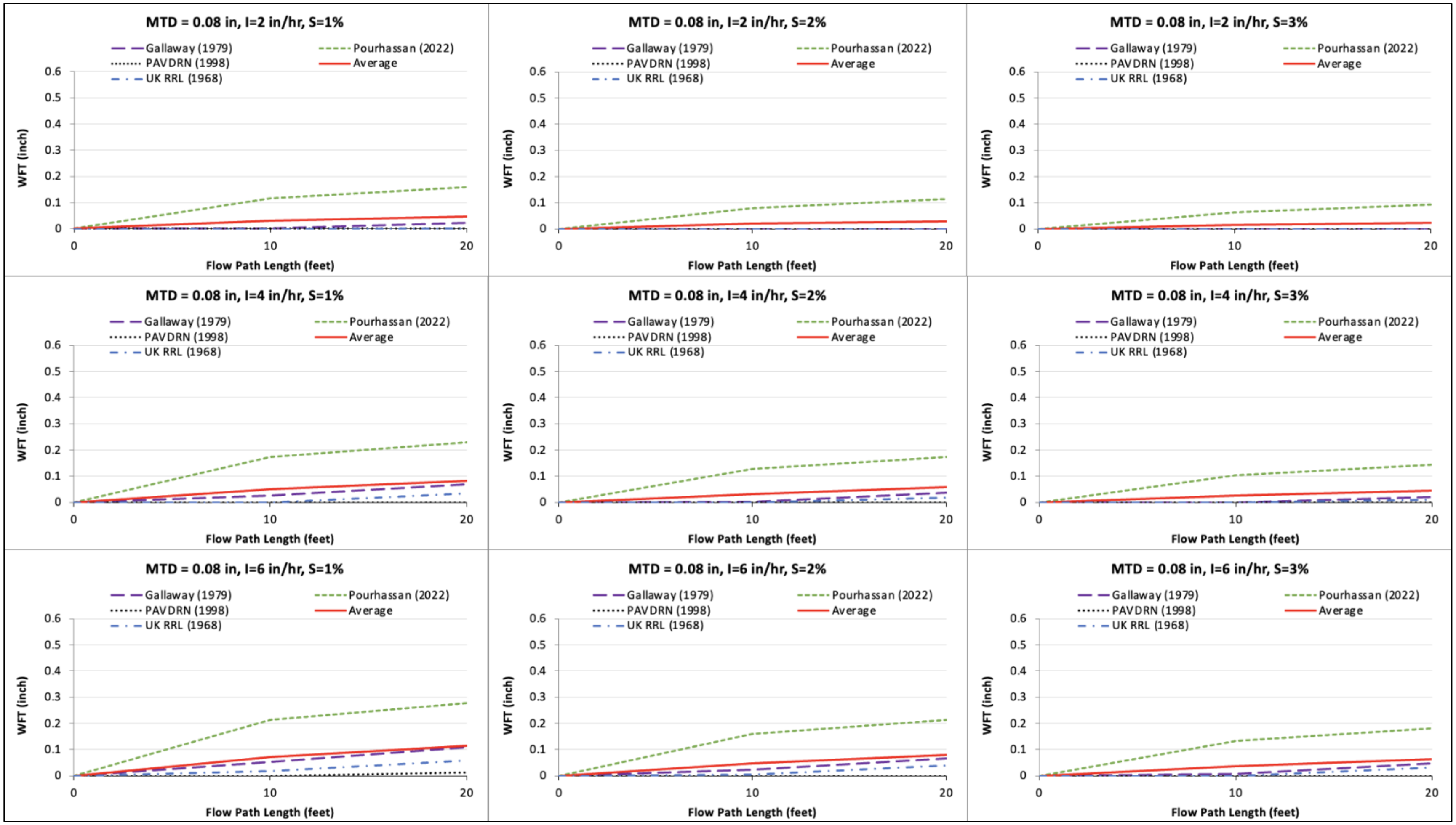


Figure 4. Sensitivity of WFT to Changes in Rainfall Intensity, Cross Slope, and Flow Path Length—MTD = 0.08 Inches.

The WFT predictions were generally lowest in PAVDRN, especially at an MTD value of 0.08 inches, where most of the predictions were zero. Compared to the Gallaway model, the UK RRL model had generally higher WFT predictions at lower MTD values and lower WFT predictions at higher MTD values. Rainfall intensity had the greatest effect on the WFT for the Gallaway, PAVDRN, and UK RRL models, but MTD had the greatest effect on the WFT for the Pourhassan model.

### Crash Narrative Analysis for Identifying Hydroplaning Crashes

Recent research has explored the use of crash narrative text to identify hydroplaning crashes that are often absent from structured databases. The emergence of large language models (LLMs) has advanced this line of work (Arteaga & Park, 2025; Mumtarin et al., 2023; Grigorev et al., 2025; Huang et al., 2024). For example, Mumtarin et al. (2023) tested ChatGPT, Bard, and GPT-4 on 100 police crash reports from Iowa and Kansas and found very high agreement for simple binary queries but much lower agreement for complex classifications such as collision manner. Grigorev et al. (2025) combined LLM-derived features with structured crash data to improve severity classification, while Huang et al. (2024) fine-tuned GPT-3.5-Turbo-1106 on crash descriptions to generate concise reports and achieved strong performance. These studies underscore both the promise of LLMs and the need for careful validation when interpreting nuanced narratives. In the hydroplaning domain, Das et al. (2020) applied natural language processing and machine learning to 7 years of Louisiana crash reports. Hydroplaning was inferred from narrative cues such as “water on roadway” or “loss of control,” with advanced models like XGBoost outperforming traditional classifiers despite the absence of supporting roadway or environmental data.

## EFFECTS OF PAVEMENT CONDITION AND ROAD GEOMETRY ON CRASHES

### Summary of Findings from Prior Studies

Table 2 provides a synopsis of prior studies examining the effects of pavement condition and road geometry on road safety. Pavement conditions examined generally included roughness as determined by the International Roughness Index (IRI), skid resistance as determined by the skid number (SN), rut depth, and overall condition as determined by composite indices. Road geometry factors included horizontal curvature, vertical alignment, intersection presence, and urban/rural location. Most studies relied on statistical analyses of empirical data and considered a wide range of crash-related variables, including factors beyond pavement characteristics.

Although findings varied, most studies reported increased crash rates and frequencies with increased pavement roughness. Most studies examining crash severity indicated that smoother pavements were linked to more severe crashes, including a higher probability of fatal crashes. Rut depth was generally associated with increased crash rates, particularly under specific conditions such as nighttime driving, rainy weather, and severe rutting (> 0.4 inches).

**Table 2. Prior Studies on the Effect of Pavement Condition and Road Geometry on Safety.**

| Safety Metric | Analysis Method and Study Location | Key Findings | Source |
|---|---|---|---|
| Crash frequency | NB (Wyoming) | **Skid resistance:** Decreased pavement friction increased frequency of total and PDO crashes. **Geometry:** Tangent segments increased total crash frequency but decreased PDO crash frequency. Uphill/downhill slopes increased PDO crashes in most roadway classes. | Roy et al. (2023) |
| Crash frequency | Hurdle NB (Oklahoma) | **Skid resistance:** Increased pavement friction decreased crash frequencies. **Geometry:** Shoulders significantly decreased crash frequencies. Steeper longitudinal grades increased crash frequencies. | Li et al. (2022) |
| Crash frequency | Generalized estimating equation with NB (Australia) | **Skid resistance and roughness:** At signalized intersections, roughness and skid resistance effects on crash frequency was dependent on interactions with other influencing variables including day/night light, traffic volume, and wet/dry surface condition. | Hussein et al. (2021) |
| Crash frequency | NB regression (Italy) | **Skid resistance:** Increased skid resistance decreased crash frequency. **Roughness:** Increased roughness increased crash frequency. | Cafiso et al. (2021) |
| Crash frequency (crashes/ 3 years) | NB regression (New Brunswick) | **Rutting:** Rut depth was not a contributing factor to crash occurrence on rural, two-lane, undivided arterial and collector roads. **Roughness:** With all other significant variables held constant, increased IRI decreased crash frequency for single- and multi-vehicle crashes. **Geometry:** Intersections increased crash frequencies. | Ghanbari (2017) |
| Expected crashes in each severity level (fatal, injury, noninjury) | Multivariate random parameter NB model (Indiana) | **Overall pavement condition:** Increased pavement condition (from poor to good) decreased crash propensity. | Chen et al. (2017) |
| Crash frequency | Zero-inflated Poisson regression (Canada) | **Rutting:** Rut depth was not a significant factor in all, daylight, nighttime, peak-/nonpeak-hour crashes. **Roughness:** Increased (rougher) IRI increased number of crashes on rural highways. | Tehrani et al. (2017) |
| Crash rate (crashes per 100 million vehicle-miles traveled per year) | Sigmoidal function regression analysis (Arizona, North Carolina, and Maryland) | **Roughness and rutting:** For IRI $\leq$ 210 inches/mile or rut depth $\leq$ 0.4 inches, crash rate did not change. For IRI > 210 inches/mile or rut depth > 0.4 inches, crash rate increased. | Vinayakamurthy et al. (2017) |

| Safety Metric | Analysis Method and Study Location | Key Findings | Source |
|---|---|---|---|
| Crash severity | Bayesian ordered logistic regression (Florida) | **Overall pavement condition:** Decreased (poor) pavement condition increased multi-vehicle crash severity at all speed levels and single-vehicle crash severity on high-speed roads but decreased single-vehicle crash severity on low-speed roads. **Geometry:** Intersections decreased single-vehicle crash severity on low- and medium-speed roads but increased multi-vehicle crash severity. Crashes tend to be more severe on rural than urban roads. | Lee et al. (2015) |
| Crash frequency | Fisher's least significant difference and Tukey-Kramer test (Texas) | **Ride Score:** Decreased (rougher) ride scores increased mean crash rate by almost three times compared to increased (smoother) ride scores. **Roughness:** Increased (rougher) IRI increased mean crash rate by more than two times compared to decreased (smoother) IRI. **Overall pavement condition:** Decreased (very poor) pavement conditions increased crash rates by more than two times compared to increased (very good) pavement conditions; crash rates on roads with poor to fair pavement conditions were similar. Decreased (poor) pavement conditions increased same-direction, noninjury, favorable light condition, and rural/urban collector with medium speed limits crashes. | Li and Huang (2014) |
| Crash severity | Ordered probit (Texas) | **Roughness:** Decreased (smoother) IRI increased probability of a fatal crash occurrence. **Geometry:** Increased shoulder width on two-lane roads increased likelihood of fatal crashes. Horizontal curves with downward gradients increased likelihood of fatality compared to flat gradients. Rural areas increased crash severity compared to urban areas. | Buddhavarapu et al. (2013) |
| Crash severity | Chi-square tests (Texas) | **Roughness:** Decreased (smoother) pavements increased severe crashes. | Li et al. (2013) |
| Crash rate | Tobit model (Indiana) | **Rutting and roughness:** Increased IRI, rut depth, and surface deflection increased crash rate. **Overall pavement condition:** Decreased pavement condition increased crash frequency. **Geometry:** Horizontal curvature significantly decreased crash rate. | Anastasopoulos et al. (2012) |
| Crash frequency (per 0.1-mile highway segment per year) | NB regression (Tennessee) | **Roughness:** Increasing IRI from 0–100 to 101–200 inches/mile increased crash frequency 1.65 times. **Rutting:** Increasing rut depth by 0.1 inches increased nighttime and wet-weather crash frequencies by 1.51 and 1.684 times, respectively. Rut depth was not significant for daylight condition, good weather condition, or peak-/nonpeak-hour crashes. | Chan et al. (2010) |

| Safety Metric | Analysis Method and Study Location | Key Findings | Source |
|---|---|---|---|
| Crash frequency | Cross-sectional analysis (Spain) | **Skid resistance:** Increased skid resistance decreased wet- and dry-pavement crash frequencies on two-lane rural highways. | Mayora and Piña (2009) |
| Crash Frequency | Comparison of percentage of crashes that occurred on pavement sections within specific ranges of conditions (Alabama) | **Rutting and roughness:** Increased rutting and roughness increased crash frequency. **Geometry:** Grass shoulders disproportionally increased crash frequency compared with other shoulder types. | Graves et al. (2005) |
| Crash rate (crashes per million vehicle-kilometers) | Confidence intervals for rut-related crash rates (Wisconsin) | **Rutting:** Rut depths > 0.30 inches significantly increased crash rates. | Start et al. (1998) |
| Crash rate (crashes per million vehicle-kilometers) | Multiplicative regression equation (Jordan) | **Roughness:** Increased (rougher) IRI increased multi-vehicle crash rate but decreased single-vehicle crash rate. **Geometry:** Increased horizontal curvature increased single-vehicle crash rates. Increased intersection density increased crash rates. Shoulder and lane widths did not significantly affect crashes. Increased vertical curves increased multi-vehicle crash rates. | Al-Masaeid (1997) |

IRI = International Roughness Index; NB = negative binomial; PDO = property damage only.

Increased skid resistance was found to decrease crash frequencies, for both wet and dry pavement conditions. Decreased (very poor) overall pavement condition was found to increase crash frequencies, as well as crash severities for multi-vehicle crashes.

Regarding roadway geometry, the presence of horizontal curves decreased overall crashes but increased single-vehicle crashes. Vertical curves with a higher average slope were associated with increased crash frequencies on any road type and increased crash severities on horizontal curves. The presence of shoulders was found to decrease crash frequencies, but an increased shoulder width led to increased crash severities. Crash frequencies and injury severities were also higher at intersections and in rural areas.

### Statistical Models for Evaluating the Effect of Contributing Factors on Crashes

Pavement-related crash frequencies are commonly analyzed using count models, with Poisson and negative binomial (NB) models being the most widely applied (Cafiso et al., 2021; Chan et al., 2010; Hussein et al., 2021; Kassu & Anderson, 2019). Due to the inherent overdispersion in crash data, the NB model is often preferred. Kassu and Anderson (2019) compared Poisson and NB models for wet- and dry-pavement crashes on highways and found the NB model to provide superior performance.

The use of either Poisson or NB depends on the variability of the number of crashes. The Poisson model can estimate the number of pavement-related crashes, $y_i$, for a given segment, $i$, given a vector of explanatory variables, $X_i$. Assuming that the mean, $\mu_i$, equals the variance, the number of crashes is assumed to follow a Poisson distribution with a likelihood calculated as follows:

$$y_i \sim \text{Pois}(\mu_i) = f(y_i|X_i) = \frac{e^{-\mu i}{\mu_i}^{yi}}{y_i!} \tag{5}$$

In most cases, the average number of crashes is not equal to the variance; instead the crashes are considered overdispersed. This scenario violates the regular Poisson distribution assumption, and the variability is included as a mixture distribution with conjugate gamma mixing (Booth et al., 2003). After the introduction of the overdispersion parameter, $\alpha$, the NB model for crash frequency estimation is formulated as follows:

$$\Pr(y_i|\alpha, \mu_i) = \frac{\Gamma(y_i+\alpha)}{\Gamma(\alpha)y_i!}\left(\frac{\alpha}{\mu_i+\alpha}\right)^{\alpha}\left(\frac{\mu_i}{\mu_i+\alpha}\right)^{y_i} \tag{6}$$

where the expected mean and variance can be computed as $\mu_i$ and $\mu_i + \mu_i^2 / \alpha$, respectively.

Because the Poisson model exists as a log-linear relationship, this tendency prevails between $y_i$, which is negatively binomially distributed with $\mu_i$, and $X_i$ based on the following relationship:

$$ln(E(y_i|x_i) = X_{ij}\beta \tag{7}$$

where $X_{ij}$ indicates the covariates row vector, $y_i$ is the number of crashes, and $\beta$ is the matrix of regression coefficients.

The effects of these pavement condition indicators can be quantified by estimating their parameters/coefficients, revealing how a one-unit increase in the condition indicator corresponds to an increase in the number of crashes. This modeling approach offers the advantage of incorporating all relevant variables and identifying their effects through estimated parameters. However, it has the disadvantage of providing monotonic estimates that may not capture dichotomous effects.

Binary models can be used to evaluate the impacts of contributing factors on the probability of crash occurrence or severity. Logit and probit models are most commonly used in this case. These models are forms of regression models, where the dependent variable has a binary outcome. While the two types are similar, logit models are preferred due to the easy interpretation of the model coefficients using odds ratios (Kutela et al., 2022; Norton & Dowd, 2018). The logit regression model is formulated as follows:

$$logit(P_i) = \beta X + \varepsilon \tag{8}$$

where $P_i$ is the probability of crash occurrence, $\beta$ is a fixed effects parameters matrix, $X$ is an explanatory variables matrix, and $\varepsilon$ is an error terms matrix.

In the context of pavement safety, most prior studies have utilized binary outcome models to predict the effect of pavement-related variables on injury severity, not crash occurrences. For example, Ma et al. (2009) associated pavement conditions and injury severity of crashes using logit models. Their study used road cross section, crash location, road alignment, road type, and lighting condition as the key predictors for crash severity. Similarly, Nowakowska (2010) used logistic models to predict crash severity classification based on road design elements, such as vertical and horizontal curves. Li et al. (2019) evaluated the injury severity of single-vehicle crashes under rain conditions with mixed logit and latent class models.

In this study, the research question to be answered is, “What factors affect the likelihood that a crash will occur?” To answer this question, researchers used logistic regression to estimate crash susceptibility as a probability bounded between 0 and 1 (probability of a crash occurrence). This approach was advantageous because it was intuitive and useful for identifying relative risk factors. The details of this modeling approach are discussed in Chapter 4 of this report.

# CHAPTER 3:
# DATA ASSEMBLY

This chapter summarizes the research dataset assembled for the pavement safety analysis and describes the process used to assemble it. The dataset (hereafter referred to as the integrated dataset) contains information on the pavement condition, geometry, precipitation, and traffic for 18,843 on-system roadbed-miles across 3 years (Fiscal Year [FY] 2021–2023) and 24,873 crashes that occurred on them during the 3-year period. The information in the integrated dataset was primarily obtained from existing TxDOT databases, except for the precipitation data, which was obtained from publicly available weather station data.

## DATA ASSEMBLY PROCESS

The integrated dataset consisted of multiple data elements representing variables that described the roadway, traffic, and weather (precipitation) characteristics at the time of the crash, as well as the crash counts and characteristics (see Table 3).

**Table 3. Primary Elements of the Integrated Dataset.**

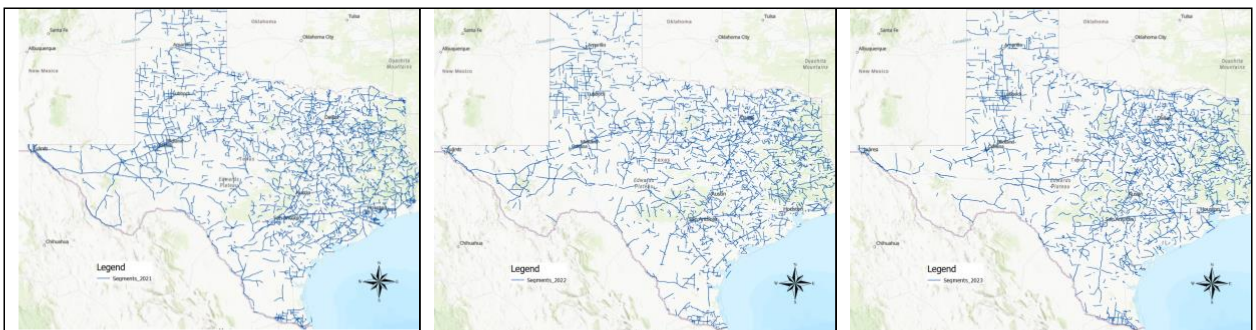
| Pavement Condition | Pavement Geometry | Traffic | Weather | Crashes Counts |
|---|---|---|---|---|
| • Pavement type<br>• SN<br>• Distress score<br>• Condition score<br>• Average IRI<br>• Cracking percentage<br>• Mean MPD<br>• Maximum MPD<br>• Minimum MPD<br>• MPD standard deviation<br>• ACP Average rut depth<br>• ACP percentage of area cracked<br>• ACP raveling<br>• ACP flushing<br>• Potholes<br>• ACP failures<br>• JCP Average faulting<br>• JCP percentage of slabs with cracks<br>• JCP failures<br>• CRCP percentage of area cracked | • Number of lanes<br>• Lane width<br>• Roadbed width<br>• Segment length<br>• Beginning elevation<br>• Ending elevation<br>• Longitudinal grade<br>• Beginning bearing<br>• Ending bearing<br>• Deflection angle<br>• Curve radius<br>• Estimated as-designed cross slope<br>• Water flow path slope | • Speed limit<br>• ADT<br>• Truck ADT percentage | • Precipitation intensity at crash time | • Total crashes<br>• Fatal crashes<br>• Serious injury crashes<br>• Nonserious injury crashes<br>• Possible serious injury crashes<br>• Property damage or other/unknown damage crashes<br>• Daytime crashes<br>• Nighttime crashes<br>• Dry-pavement crashes<br>• Wet-pavement crashes |

ACP = asphalt concrete pavement; ADT = average daily traffic; CRCP = continuously reinforced concrete pavement; JCP = jointed concrete pavement; MPD = mean profile depth; IRI = International Roughness Index; SN = skid number.

The assembly of these data elements into an integrated dataset consisted of four primary steps—identifying pavement sections with complete pavement condition data and subsequently and separately joining the roadway inventory/geometry, crash, and precipitation data. The pavement condition data elements were obtained from TxDOT's PMIS and Pavement Analyst system.

### Step 1: Identify Pavement Sections with Complete Pavement Condition Data

Step 1 included identifying on-system 0.1-mile pavement sections with complete pavement condition data for FY 2021–2023. In this context, complete pavement condition data consisted of distress indicators (e.g., cracking, rutting, raveling, flushing, potholes, faulting, failures, distress score, and condition score), SN, IRI, and mean profile depth (MPD). These data were collected for the entire length of the outer lane of the 0.1-mile sections, except for the SN, which represents a test point within the 0.1-mile section. Figure 5 shows the pavement sections identified in this step (i.e., sections with complete pavement condition data).



FY 2021 | FY 2022 | FY 2023

**Figure 5. Spatial Distribution of Pavement Sections in the Integrated Dataset (FY 2021, 2022, and 2023).**

### Step 2: Join Roadway Inventory and Geometry Data

In Step 2, roadway inventory and geometric characteristic data deemed relevant to safety analysis were obtained from TxDOT's open data portal or estimated based on other existing data. The annual average daily traffic (AADT), number of lanes, roadbed width, lane width, and speed limit were obtained from TxDOT's open data portal. The inventory and condition datasets were joined by matching highway roadbed IDs and overlapping the beginning and ending distances from origin (DFOs) between the road sections from the inventory data and the 0.1-mile sections from the condition data.

Pavement geometric characteristics, including longitudinal slope and deflection angle, were calculated using the beginning and ending elevations and bearings of the 0.1-mile sections. Because the actual (in-service) pavement cross slope was not available in existing databases, the as-designed cross slope was estimated based on TxDOT's *Roadway Design Manual* (TxDOT, 2024) and used as a surrogate for the actual cross slope.

A 2 percent cross slope was assigned as the as-designed cross slope for tangent pavements with two lanes or less inclined in the same direction. For tangent pavements with three or more lanes inclined in the same direction, the as-designed cross slope depended on the number of lanes. For example, for four lanes inclined in the same direction in a tangent pavement, the cross slope was 1.5 percent for the first two lanes and 2 percent for the third and fourth lanes, with a weighted average of 1.75 percent. For horizontal curves, the as-designed superelevation was estimated based on the minimum radius and design speed. The radius of each horizontal curve was estimated using Equation 9 as follows:

$$R = \frac{C}{2\sin\left(\frac{\Delta}{2}\right)} \tag{9}$$

where *R* is the radius of the horizontal curve (feet), *C* is the cord length of the curve (feet), and $\Delta$ is the deflection angle (degrees). The chord length was assumed to equal the pavement section length. The defection angle was estimated from the beginning and ending bearings of the pavement section. Once the radius of the curve was computed, the superelevation was obtained directly from tables in TxDOT's *Roadway Design Manual* (TxDOT, 2024). If the superelevation was below 2 percent or was identified as a normal or reverse crown in the design manual, a 2 percent cross slope was used in the integrated database.

**Step 3: Join Crash Data**

Data on crashes that occurred on the 0.1-mile sections in the integrated database in FY 2021–2023 were obtained from TxDOT's CRIS.

Crashes were spatially joined if their longitude and latitude coordinates were located on the roadbed of the pavement sections. However, most crashes on divided roads are located at the centerline of roads (rather than on a specific roadbed). In these cases, the crash attributes—street name, DFO, vehicle direction, and road part (if frontage roads exist)—were used to determine the specific roadbed where the crash occurred (see Figure 6). The street name attribute indicated the highway system and number, and the DFO specified the exact location of the crash along the highway. To identify the roadbed where the crash occurred, researchers matched the vehicle direction and road part (main or frontage road) with the corresponding roadbed.

If it was not possible to locate the specific roadbed where the crash occurred, then the pavement sections at the crash location were removed from the integrated dataset to avoid attributing crashes to the wrong sections and creating potential inaccuracies in the data analysis. Such exclusions occurred when:

- The vehicle's direction before the crash was unknown or inconsistent with the direction of the road (e.g., the vehicle's direction was east, but the roadbed runs north and south).

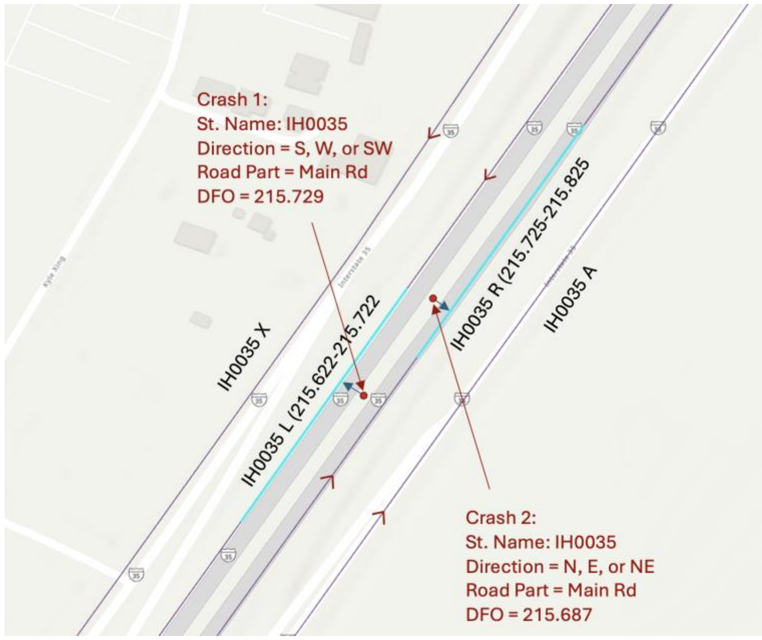



**Figure 6. Example of How Crashes with Centerline Coordinates Were Joined to Roadbeds.**

- The crash occurred on a supplemental roadbed. On supplemental roadbeds, it was not possible to determine the exact roadbed where the crash occurred because the road part attribute in the crash dataset did not specify whether the crashes occurred on supplemental roadbeds.
- The crash occurred at an intersection, an intersection-related location, or a driveway access or on an on-ramp, a connector, or detour route. Only crashes on non-intersection main and frontage roads were considered.

Overall, these exclusions represented approximately 5.4 percent of the initial integrated dataset.

**Step 4: Join Precipitation Data**

The precipitation intensity data were sourced from TexMesonet, a network of weather stations distributed across Texas. These stations provided real-time data on various climatic conditions, including precipitation, temperature, humidity, and wind speed. This study focused on estimating precipitation intensity at the crash time using data from these weather stations.

Given Texas' vast area and diverse geography, matching crash locations with the nearest weather stations was crucial for accurate analysis. Researchers used the ArcGIS Pro software to spatially join the crash data with the weather station data (based on the longitude and latitude coordinates of the crash and weather station locations). This step was critical in ensuring that each crash was accurately matched with the closest possible weather station. However, the precipitation data from some weather stations were missing or incomplete. To address this issue, the five nearest weather stations to each crash site were identified and ranked based on their proximity to the crash location. If precipitation data from the closest weather station were missing or incomplete, data from the next nearest station were used. Most crashes had precipitation data from the closest or second closest weather station (i.e., first or second proximity ranking) (see Figure 7).

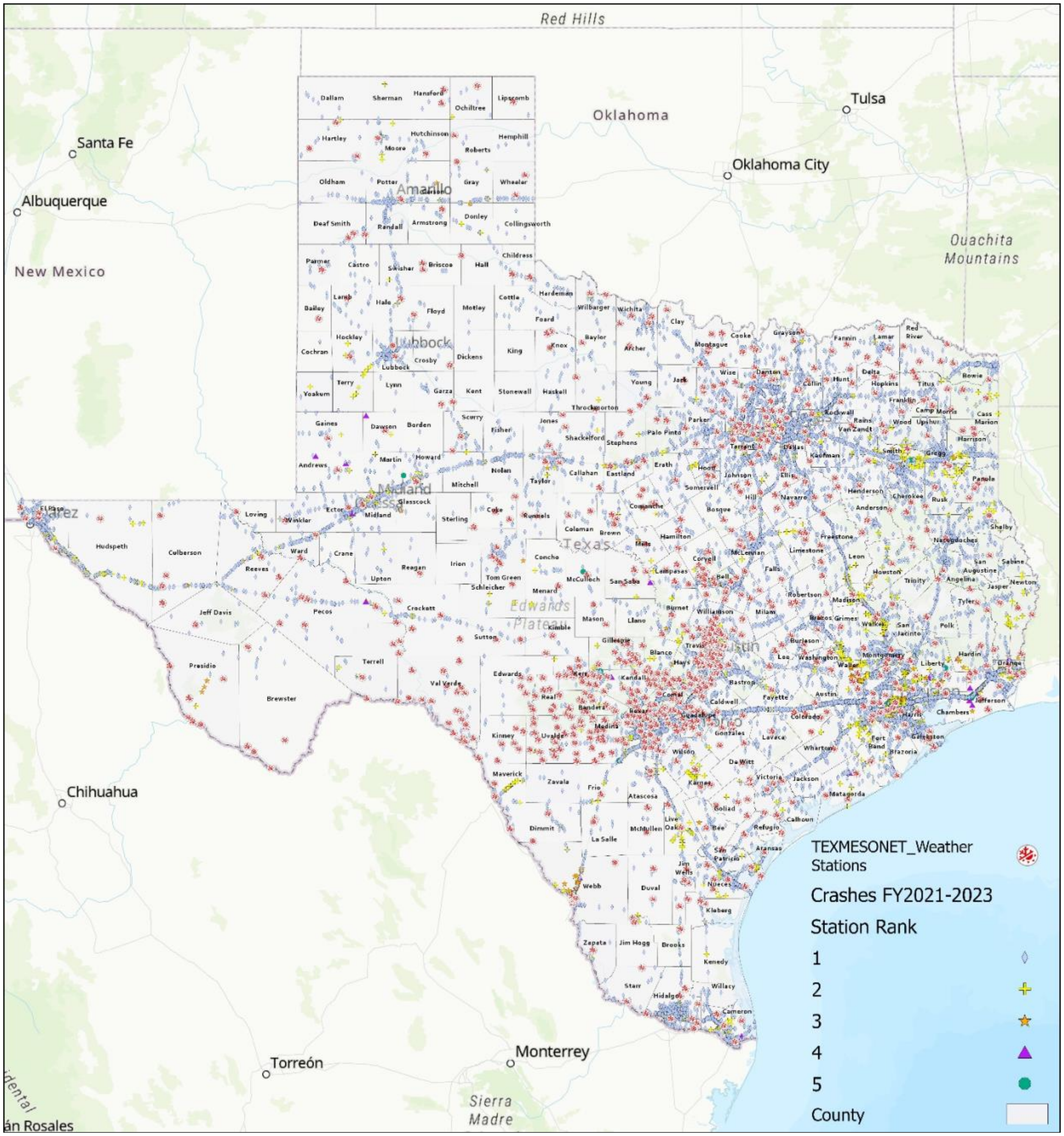



**Figure 7. Color-Coded Crashes Based on Proximity Rank of Weather Station for Precipitation Data (FY 2021–2023).**

Figure 8 shows the distribution of crashes relative to the distance from the nearest weather station. The x-axis represents the distance in miles between the crash site and the weather station from which the precipitation data were obtained. The y-axis represents the corresponding number of crashes. This chart shows that 83 percent of crashes occurred within a relatively short distance (0–10 miles) from the corresponding weather station. A small percentage of crashes (less than 1 percent) obtained their precipitation data from distant weather stations (> 25 miles).

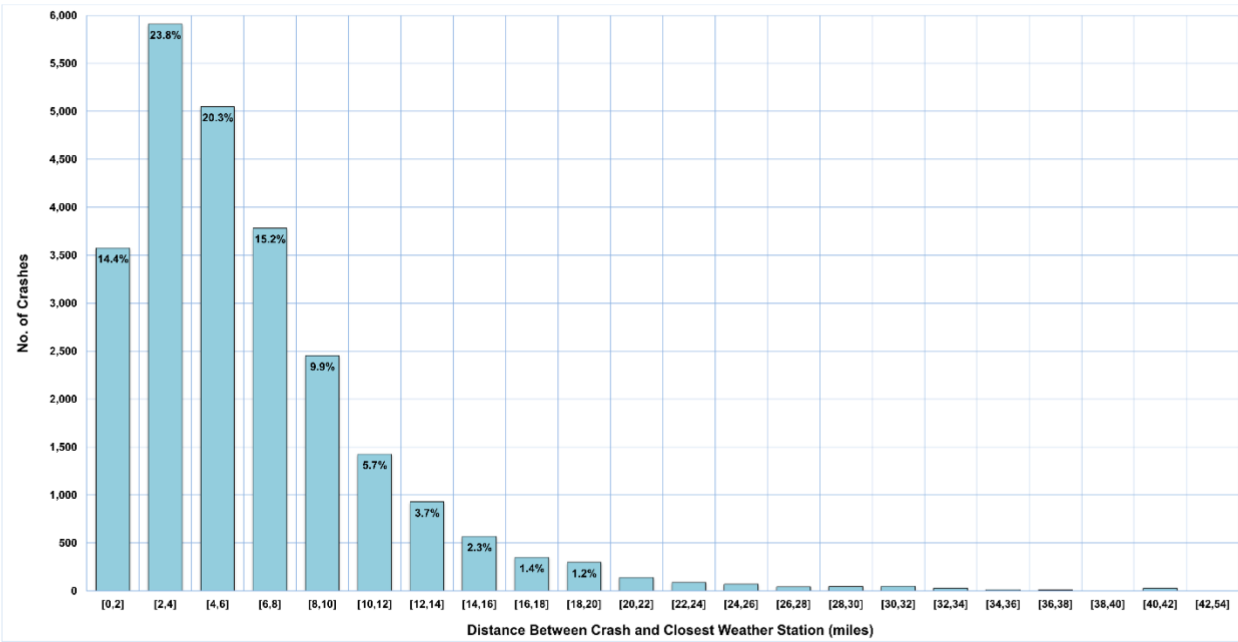



**Figure 8. Distance between Crash and Closest Weather Station for Precipitation Data (FY 2021–2023).**

The weather stations recorded precipitation depths across various time intervals (e.g., 5, 10, 15 minutes). Therefore, it was necessary to define an optimal time interval that represented the actual precipitation intensity at the time of the crash. After several trials, researchers found that a 1-hour interval (30 minutes before and 30 minutes after the crash time) yielded the maximum agreement between the precipitation intensity (from weather station data) and the weather condition (e.g., clear, rain, snow, etc.) at the crash site, as reported in CRIS. Table 4 illustrates the process of computing precipitation intensity at crash time for the following scenario:

- Weather station = Terrell Municipal Airport weather station (Station ID = KTRL).
- Crash date and time = 2020-01-09, 17:00:00.
- Precipitation data 1-hour interval = 16:30:00 to 17:30:00.
- Precipitation intensity within interval = 0.231 inches/hour.

**Table 4. Sample Precipitation Data from Terrell Municipal Airport Weather Station for a Hypothetical Crash Time.**

| Date and Time (UTC) | Precipitation (inches) | Crash Time and Precipitation Interval |
|---|---|---|
| 2020-01-09T15:50:00Z | 0.001 | |
| 2020-01-09T15:53:00Z | 0.001 | |
| 2020-01-09T15:55:00Z | | |
| 2020-01-09T16:00:00Z | | |
| 2020-01-09T16:05:00Z | | |
| 2020-01-09T16:10:00Z | | |
| 2020-01-09T16:15:00Z | | |
| 2020-01-09T16:20:00Z | | |
| 2020-01-09T16:25:00Z | | |
| 2020-01-09T16:30:00Z | [ | ←Start of 1-hour precipitation interval |
| 2020-01-09T16:35:00Z | | |
| 2020-01-09T16:40:00Z | | |
| 2020-01-09T16:45:00Z | | |
| 2020-01-09T16:50:00Z | | |
| 2020-01-09T16:53:00Z | 0.001 | |
| 2020-01-09T16:55:00Z | | |
| 2020-01-09T17:00:00Z | | ←Crash time |
| 2020-01-09T17:05:00Z | 0.02 | |
| 2020-01-09T17:10:00Z | 0.03 | |
| 2020-01-09T17:12:00Z | 0.03 | |
| 2020-01-09T17:15:00Z | 0.03 | |
| 2020-01-09T17:20:00Z | 0.03 | |
| 2020-01-09T17:21:00Z | 0.03 | |
| 2020-01-09T17:25:00Z | 0.03 | |
| 2020-01-09T17:30:00Z | 0.03] | ←End of 1-hour precipitation interval |
| 2020-01-09T17:35:00Z | 0.03 | |
| 2020-01-09T17:40:00Z | 0.03 | |
| 2020-01-09T17:45:00Z | 0.03 | |
| 2020-01-09T17:50:00Z | 0.03 | |
| 2020-01-09T17:53:00Z | 0.03 | |

UTC = Coordinated Universal Time.

## SUMMARY OF THE INTEGRATED DATASET

Table 5 shows the roadbed-miles, lane-miles, and number of crashes for each fiscal year in the integrated dataset. In total, the integrated dataset consisted of 18,843 roadbed-miles for FY 2021–2023, of which 17,765 roadbed-miles were unique (nonoverlapping through the 3-year period). The total number of crashes that occurred on these road segments during the 3-year period was 24,873 crashes, with the number of crashes per year varying between 6,912 in FY 2021 and 8,264 in FY 2023.

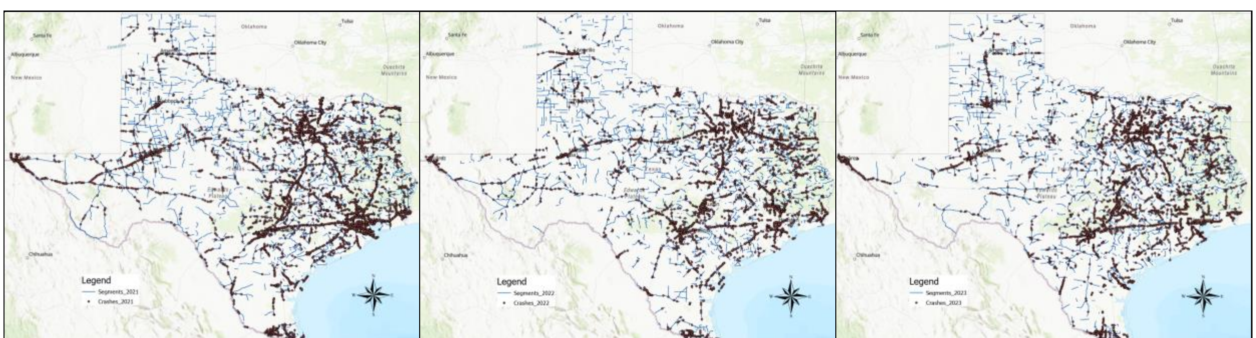
Considering that the current on-system network consists of approximately 103,052 roadbed-miles (per TxDOT's roadway inventory data and Open Data Portal), the integrated dataset for the 3 years combined represented 17.2 percent of TxDOT's on-system roadway network (overall sampling rate = 17,765 / 103,052 = 17.2 percent). The sampling rate for the individual years was 6.57, 5.83, and 5.89 percent for FY 2021, 2022, and 2023, respectively.

**Table 5. Number of Roadbed-Miles, Lane-Miles, and Crashes in the Integrated Dataset (FY 2021, 2022, and 2023).**

| FY | Roadbed-Miles | Lane-Miles | Crashes |
|---|---|---|---|
| 2021 | 6,771 | 14,975 | 9,697 |
| 2022 | 6,005 | 13,004 | 6,912 |
| 2023 | 6,067 | 13,270 | 8,264 |
| Total | 18,843 | 41,249 | 24,873 |

FY = fiscal year.

Figure 9 displays a series of maps showing road segments and crashes contained in the integrated dataset for each fiscal year in the analysis. These maps delineate crash sections from non-crash sections. Crash sections are those sections where one or more crashes occurred during a particular year, whereas non-crash sections are those sections where no crashes occurred during the year. As illustrated in these maps, the road sections and crashes were distributed throughout Texas, allowing for the analysis of pavement safety in varying geographic contexts and climatic conditions. Naturally, concentrations of crashes were evident in major urban areas compared to nonurban areas due to higher traffic volumes.



| FY 2021 | FY 2022 | FY 2023 |
|---|---|---|

**Figure 9. Spatial Distribution of Pavement Sections and Crashes in the Integrated Dataset (FY 2021, 2022, and 2023).**

Figure 10 shows the number of pavement sections in the integrated dataset with varying crash frequencies (number of crashes in FY 2021–2023). Approximately 91 percent of the sections in the integrated dataset (173,099 sections of 189,697 total sections) had no crashes during the 3-year period, and approximately 6.51 percent (12,345 sections of the 189,697 sections) had only a single crash. The number of sections with 2 and 3 crashes dropped sharply to 2,551 and

836 sections, respectively (1.78 percent of all sections). The number of sections that had more than 3 crashes was very low (less than 0.5 percent of all sections).

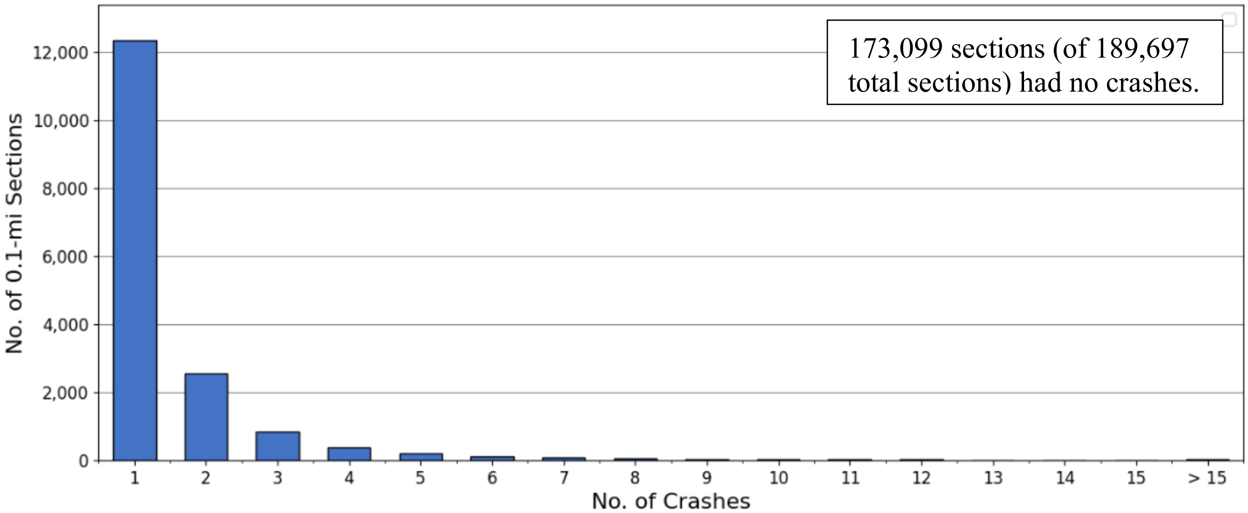



**Figure 10. Crash Frequency in the Integrated Dataset (FY 2021–2023).**

# CHAPTER 4:
# DEVELOPMENT OF CRASH SUSCEPTIBILITY MODELS

## DATA AND METHODS USED FOR CRASH SUSCEPTIBILITY MODELING

This chapter summarizes the dataset and discusses the approach for modeling pavement susceptibility to crashes. Because surface characteristics and deterioration mechanisms vary across pavement types, separate crash susceptibility models were developed for asphalt concrete pavement (ACP) with and without a surface treatment (ST) or seal coat (hereafter referred to as ACP with ST and ACP without ST, respectively), continuously reinforced concrete pavement (CRCP), and jointed concrete pavement (JCP).

### Data Summary

To develop pavement crash susceptibility models, researchers considered several pavement condition indicators including SN, IRI, MPD, cracking percentage, and distress score. Each of these indicators was evaluated for all pavement types. In addition to these crosscutting pavement condition variables, researchers evaluated other pavement type-specific condition indicators. ACP-specific conditions included rut depth, raveling, flushing, failures, and potholes. Similarly, CRCP-specific conditions included punchout and spalled cracks, and JCP-specific conditions included faulting, failures, and failed joint cracks. The IRI and rutting were evaluated using the average of the left and right wheel path measurements.

In addition to pavement conditions, researchers considered traffic, roadway geometry, and precipitation factors in the modeling process. Traffic factors included the AADT, along with the percentage of trucks in the AADT. Because AADT is an exposure variable, researchers used a log transformation of this variable in the analysis. For roadway geometry, variables included longitudinal grade, as-designed cross slope, and deflection angle. Lastly, the precipitation normal was considered to account for the influence of weather factors on crash occurrence. Table 6 shows the ACP variable summary statistics for 168,890 sections used in the analysis, and Table 7 shows CRCP/JCP variable summary statistics for 15,925 sections used in the analysis. Each pavement section was approximately 0.1-mile long.

### General Form of Crash Susceptibility Model

Researchers in this study introduced crash susceptibility as a new parameter useful for incorporating safety in pavement management decisions. It was defined as the probability of crash occurrence at any given pavement segment in any given year. The research team utilized logistic regression to predict the probability of crash occurrence with two possible outcomes—a pavement segment had at least 1 crash in any year during the 3-year study period (FY 2021–2023) or it did not. Thus, the dependent variable had two possible outcomes $\{0, 1\}$, where 0 indicated that no crash and 1 indicated at least 1 crash for a given segment.

**Table 6. Summary Statistics for ACP Variables.**

| Variable | | ACP without ST | | | | ACP with ST | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation | Mean | Minimum | Maximum | Standard Deviation |
| Crashes | Crash frequency (number of crashes per year) | 0.16 | 0 | 23 | 0.59 | 0.03 | 0 | 14 | 0.21 |
| Pavement condition | SN | 35.7 | 1.1 | 99 | 13.8 | 41.2 | 1.1 | 99 | 15.5 |
| | Distress score (0–100) | 91.3 | 2 | 100 | 14 | 91.1 | 1 | 100 | 13.8 |
| | Average IRI (inches/mile) | 95.5 | 15 | 774 | 47.4 | 124 | 23 | 491 | 48.5 |
| | Cracking percentage | 0.61 | 0 | 65 | 2.99 | 0.56 | 0 | 60.3 | 2.55 |
| | MPD (inches) | 0.04 | 0.006 | 0.21 | 0.02 | 0.05 | 0.006 | 0.32 | 0.03 |
| | Average rut depth (inches) | 0.12 | 0.001 | 0.79 | 0.07 | 0.12 | 0.001 | 0.76 | 0.07 |
| | Raveling (percentage of rated lane surface area) | 4.08 | 0 | 100 | 12.5 | 8.24 | 0 | 100 | 17.5 |
| | Flushing (percentage of rated wheelpath length) | 3.49 | 0 | 100 | 17.5 | 2.79 | 0 | 100 | 15.4 |
| | Failure quantity (number of failures in rated section) | 0.007 | 0 | 13 | 0.15 | 0.01 | 0 | 13 | 0.22 |
| | Pothole quantity (number of potholes in rated section) | 0.008 | 0 | 13 | 0.14 | 0.005 | 0 | 7 | 0.1 |
| Traffic | AADT (vpd) | 7,426 | 51 | 136,648 | 10,639 | 1494 | 51 | 39,450 | 2,565 |
| | Percentage trucks in AADT (%) | 19.9 | 0 | 94.6 | 15.2 | 19 | 0 | 96.1 | 13.4 |
| Roadway geometry | Absolute longitudinal grade (%) | 0.9 | 0 | 12.6 | 1.02 | 1.06 | 0 | 11.2 | 1.19 |
| | As-designed cross slope (%) | 2.32 | 1.62 | 6 | 0.93 | 2.53 | 1.62 | 6 | 1.18 |
| | Deflection angle (degrees) | 2.46 | 0 | 90 | 6.12 | 3.51 | 0 | 89.9 | 7.93 |
| Weather | Precipitation normal (inches/year) | 35.2 | 8.28 | 70.5 | 13.9 | 32.1 | 8.28 | 70.5 | 11.5 |

AADT = annual average daily traffic; ACP = asphalt concrete pavement; IRI = International Roughness Index; MPD = mean profile depth; SN = skid number; ST = surface treatment.

**Table 7. Summary Statistics for CRCP/JCP Variables.**

| Variable | | CRCP | | | | JCP | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Minimum | Maximum | Standard Deviation | Mean | Minimum | Maximum | Standard Deviation |
| Crashes | Crash frequency (number of crashes per year) | 0.47 | 0 | 26 | 1.21 | 0.33 | 0 | 9 | 0.75 |
| General pavement condition | SN | 30 | 4.7 | 88.5 | 9.08 | 29.7 | 6.6 | 93.6 | 8.49 |
| | Distress score (0–100) | 92.9 | 1 | 100 | 19.1 | 74.6 | 1 | 100 | 35.1 |
| | Average IRI (inches/mile) | 111 | 19 | 602 | 43.8 | 154 | 37 | 615 | 65 |
| | Cracking percentage | 0.15 | 0 | 36 | 1.29 | 2.77 | 0 | 81.4 | 7.34 |
| | MPD (inches) | 0.02 | 0.006 | 0.2 | 0.007 | 0.02 | 0.008 | 0.09 | 0.007 |
| CRCP-specific condition | Punchout quantity (number of punchouts in rated section) | 0.02 | 0 | 6 | 0.21 | - | - | - | - |
| | Spalled crack quantity (number of spalled cracks in rated section) | 0.11 | 0 | 30 | 0.8 | - | - | - | - |
| JCP-specific condition | Faulting (inches) | - | - | - | - | 0.05 | 0.005 | 0.65 | 0.04 |
| | Failure quantity (number of pavement failures in rated section) | - | - | - | - | 1.45 | 0 | 56 | 4.08 |
| | Failed joint/crack quantity (number of slabs with failed joints/cracks in rated section) | - | - | - | - | 0.52 | 0 | 20 | 1.34 |
| Traffic | AADT (vpd) | 21,877 | 116 | 151,498 | 21,095 | 14,389 | 106 | 81,682 | 11,770 |
| | Percentage trucks in AADT (%) | 19.2 | 0.3 | 90.6 | 16.9 | 11.6 | 0.2 | 74.8 | 10.3 |
| Roadway geometry | Absolute longitudinal grade (%) | 0.86 | 0 | 7.67 | 0.93 | 0.89 | 0 | 6.48 | 0.99 |
| | As-designed cross slope (%) | 2.13 | 1.62 | 6 | 0.72 | 2.12 | 1.62 | 6 | 0.69 |
| | Deflection angle (degrees) | 1.89 | 0 | 77.2 | 4.1 | 2.85 | 0 | 86.9 | 6.19 |
| Weather | Precipitation normal (inches/year) | 39.3 | 8.28 | 70.5 | 15.3 | 46.7 | 9.24 | 70.5 | 10.6 |

AADT = annual average daily traffic; CRCP = continuously reinforced concrete pavement; IRI = International Roughness Index; JCP = jointed concrete pavement; MPD = mean profile depth; SN = skid number.

The logistic regression model follows a Bernoulli distribution and can be used to estimate the probability of crash occurrence ($Y$=1) as a function of the independent variables ($X$) as follows:

$$P(Y = 1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_p X_p)}} \tag{10}$$

where P($Y$=1|$X$) is the probability that a crash occurs in a given year on a given pavement section (i.e., outcome = 1); $X_1$, $X_2$, … $X_n$ are the predictor variables; $\beta_1$, $\beta_2$, … $\beta_n$ are the coefficients of the predictor variables determined using maximum likelihood estimation; and $\beta_0$ is the intercept.

Figure 11 illustrates a fitted logistic regression model with a single independent variable. In this example, the probability of crash occurrence had a positive correlation with the variable (i.e., the model predicted a higher crash probability for observations with higher independent variable values). This study considered multiple independent variables; therefore, researchers used a multidimensional hyperplane rather than a single curve to represents the correlation between the independent variables and the probability of crash occurrence.

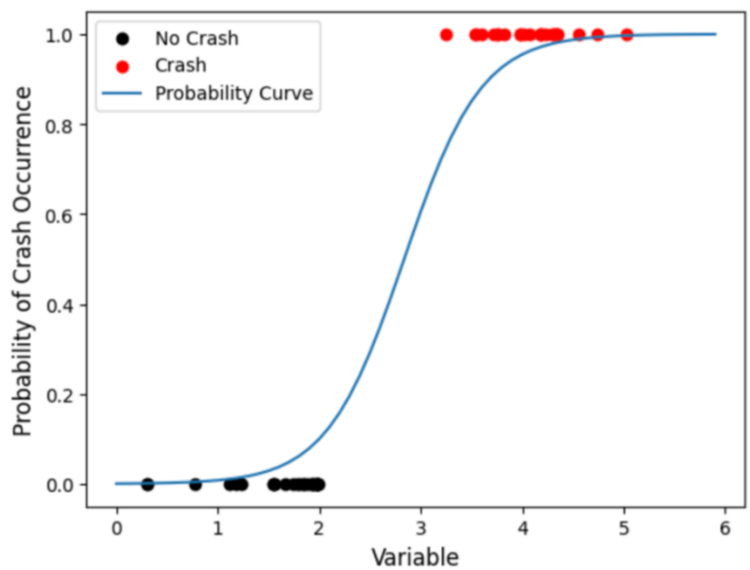



**Figure 11. Example Fitted Logistic Regression Model with a Single Independent Variable.**

The predictor variables in the final models were selected to optimally balance the model's complexity (i.e., number of predictor variables) and accuracy by minimizing the Akaike information criterion (Akaike, 1974). For each of the final models, researchers also compared the log likelihood (LL) of the final model ($LL_{Full}$), which included all predictors and the intercept, to the LL of the baseline model, which included only the intercept ($LL_{Null}$). Through these comparisons, researchers were able to assess how much better the full logistic regression model fit the data than the intercept-only model.

## Model Evaluation

Researchers assessed the performance of the developed crash susceptibility models using out-of-sample data. In this case, the models were trained on a training dataset and evaluated on a separate test dataset (unseen data), with a train/test split of 0.75/0.25. Researchers evaluated the classification performance of the models based on recall, precision, accuracy, and the F-score using Equations 11–14 as follows:

$$\text{Recall} = \frac{TP}{TP+FN} \tag{11}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{12}$$

$$\text{Accuracy} = \frac{TP+TN}{\text{Total population}} \tag{13}$$

$$F_\beta = \frac{(\beta+1)^2 \times \text{precision} \times \text{recall}}{\beta^2 \times \text{precision} + \text{recall}} \tag{14}$$

where *TP* is the true positive (both actual and predicted values indicate a crash), *FP* is the false positive (actual value indicates no crash but predicted value indicates a crash), *TN* is the true negative (both actual and predicted values indicate no crash), *FN* is the false negative (actual value indicates a crash but predicted value indicates no crash), and $\beta$ is a weighting factor for precision and recall.

Researchers used a modified F2-score equation that combined precision and recall but more heavily weighted recall over precision to minimize the likelihood of false negatives. Correctly classifying the crash instances (*TP*) was more important than correctly classifying the no crash instances (*TN*).

## ACP CRASH SUSCEPTIBILITY MODELS

This section presents the statistical models developed to assess ACP susceptibility to crash occurrence.

### ACP Model Parameters

Table 8 presents the results for two separate logistic regression models—ACP with ST and ACP without ST. The predictors of crash occurrence in each of the models included pavement-condition-, traffic-, roadway-geometry-, and weather-related predictors. The remainder of this subsection describes the effects of each of these predictors on crash occurrence on ACP sections with and without ST.

**Table 8. Logistic Regression Model Results for ACP.**

| Variable | ACP without ST | | | ACP with ST | | |
|---|---|---|---|---|---|---|
| | Estimate | Standard Error | Z-score | Estimate | Standard Error | Z-score |
| Constant | –10.71 | 0.153 | –69.81 | –9.61 | 0.270 | –35.65 |
| SN | –0.0060 | 0.0013 | –4.78 | –0.0038 | 0.0019 | –2.02 |
| IRI (inches/mile) | 0.0032 | 0.0003 | 10.89 | 0.0011 | 0.0006 | 1.90 |
| MPD (inches) | –1.52 | 0.734 | –2.07 | - | - | - |
| Pothole present (no = 0/yes = 1) | - | - | - | 0.602 | 0.344 | 1.75 |
| Log of AADT | 2.22 | 0.0315 | 70.47 | 1.83 | 0.0567 | 32.21 |
| Percentage of trucks in AADT (%) | –0.0099 | 0.0010 | –10.05 | –0.0064 | 0.0026 | –2.52 |
| Downhill (downhill = 1/uphill = 0) | 0.0780 | 0.0256 | 3.05 | - | - | - |
| Deflection angle (degrees) | 0.0140 | 0.0021 | 6.61 | 0.0269 | 0.0024 | 11.06 |
| Precipitation normal (inches) | 0.0077 | 0.0010 | 7.37 | 0.0148 | 0.0027 | 5.44 |
| **Observations (number of sections)** | 72,519 | | | 54,147 | | |
| **$LL_{Null}$** | –24,957 | | | –7,239 | | |
| **$LL_{Full}$** | –20,775 | | | –6,318 | | |

AADT = annual average daily traffic; ACP = asphalt concrete pavement; IRI = International Roughness Index; $LL_{Full}$ = log likelihood of full model; $LL_{Null}$ = log likelihood of intercept-only model; MPD = mean profile depth; SN = skid number; ST = surface treatment.

*Pavement-Condition-Related Predictors*

SN and IRI were statistically significant in both the ACP with ST and ACP without ST models. MPD was a significant factor only for the ACP without ST model, whereas the potholes variable was significant only for the ACP with ST model. The positive correlation of IRI and potholes with crash susceptibility indicated that rougher roads increase the likelihood of crash occurrence (greater roughness may reduce the driver's ability to control their vehicle). SN had a negative correlation with crash susceptibility, suggesting that crashes are more likely to occur on pavements with low skid resistance. Lower skid resistance increases skidding events, which increase the probability of crashes, especially in wet conditions. MPD was a significant factor only in the ACP without ST model. Its negative coefficient indicates that pavements with lower macrotexture are associated with increased crash susceptibility. Lower macrotexture delays water drainage during rain, which can increase the risk of hydroplaning.

*Traffic-Related Predictors*

AADT and percentage of trucks were statistically significant in both the ACP with ST and ACP without ST models. AADT showed a strong positive correlation with crash susceptibility; higher exposure to traffic naturally leads to an increased likelihood of crashes. In contrast, the percentage of trucks was negatively correlated with crash susceptibility, suggesting that a higher proportion of trucks in the traffic stream reduces crash susceptibility. The presence of trucks may encourage driver caution and discourage reckless driving behavior.

*Road-Geometry-Related Predictors*

For roadway geometry, deflection angle showed a significant positive correlation in both the ACP with ST and ACP without ST models, with a higher statistical significance in the ACP with ST model. Road segments with high deflection angles have lower sight distances, which can reduce a driver's ability to anticipate and react to potential hazards, thereby increasing the likelihood of crashes. Additionally, sharper deflection angles require drivers to reduce speed to counteract the greater centrifugal forces acting on the vehicle.

*Weather-Related Predictors*

The precipitation normal was significant in both the ACP with ST and ACP without ST models, with higher precipitation associated with higher crash susceptibility.

## ACP Model Evaluation

Table 9 shows the performance metrics for the ACP with ST and ACP without ST models using the test dataset (unseen data) at a class prediction threshold that maximized the F2-score. When comparing the performance of the ACP models, the ACP without ST model achieved a higher recall and F2-score. This outcome is likely because the dataset for the ACP with ST model was more unbalanced, with only 2.8 percent of the sections having crashes compared to 10.8 percent for the ACP without ST dataset. The higher percentage of crash instances in the ACP without ST dataset allowed the model to better learn crash patterns, resulting in better model performance. The classification accuracy was higher for the more imbalanced dataset because accuracy measures overall prediction correctness and is strongly influenced by the correct classification of the majority class (non-crash instances).

**Table 9. ACP Model Performance.**

| Evaluation Metric | ACP without ST | ACP with ST |
|---|---|---|
| F2-score | 0.509 | 0.309 |
| Recall | 0.722 | 0.550 |
| Precision | 0.233 | 0.113 |
| Accuracy | 0.708 | 0.864 |

ACP = asphalt concrete pavement; ST = surface treatment.

## ACP Model Sensitivity Analysis

Researchers performed a sensitivity analysis to evaluate how changes in each variable affected the predicted crash susceptibility, while keeping all other variables fixed at their mean values. The analysis highlighted the relative importance of each variable and its contribution to overall crash susceptibility. Figure 12 and Figure 13 present the sensitivity analysis results for ACP without ST and ACP with ST, respectively.

Each chart shows how changing the magnitude of one variable, while keeping all other variables at their mean values (listed previously in Table 6), affected the probability of crash occurrence. The 2.5th and 97.5th percentiles are displayed on the sensitivity charts to represent the predictor bounds, highlighting the effect of each variable on crash susceptibility between these extreme values. The probability of crash occurrence in any given year, with all independent variables set at their mean values, was 3.1 percent for ACP with ST and 12.9 percent for ACP without ST.

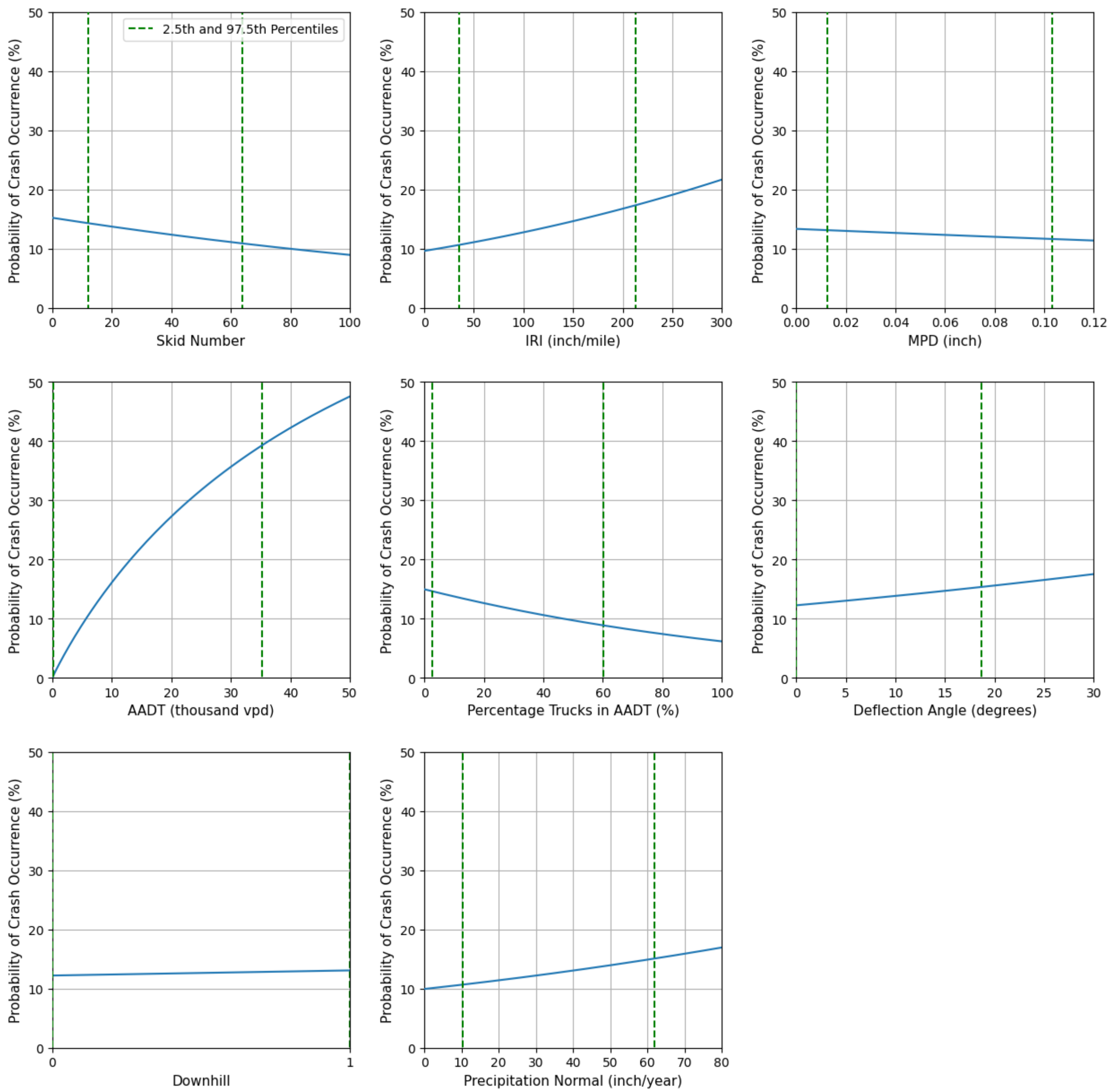



**Figure 12. Sensitivity Analysis for ACP without ST Model.**

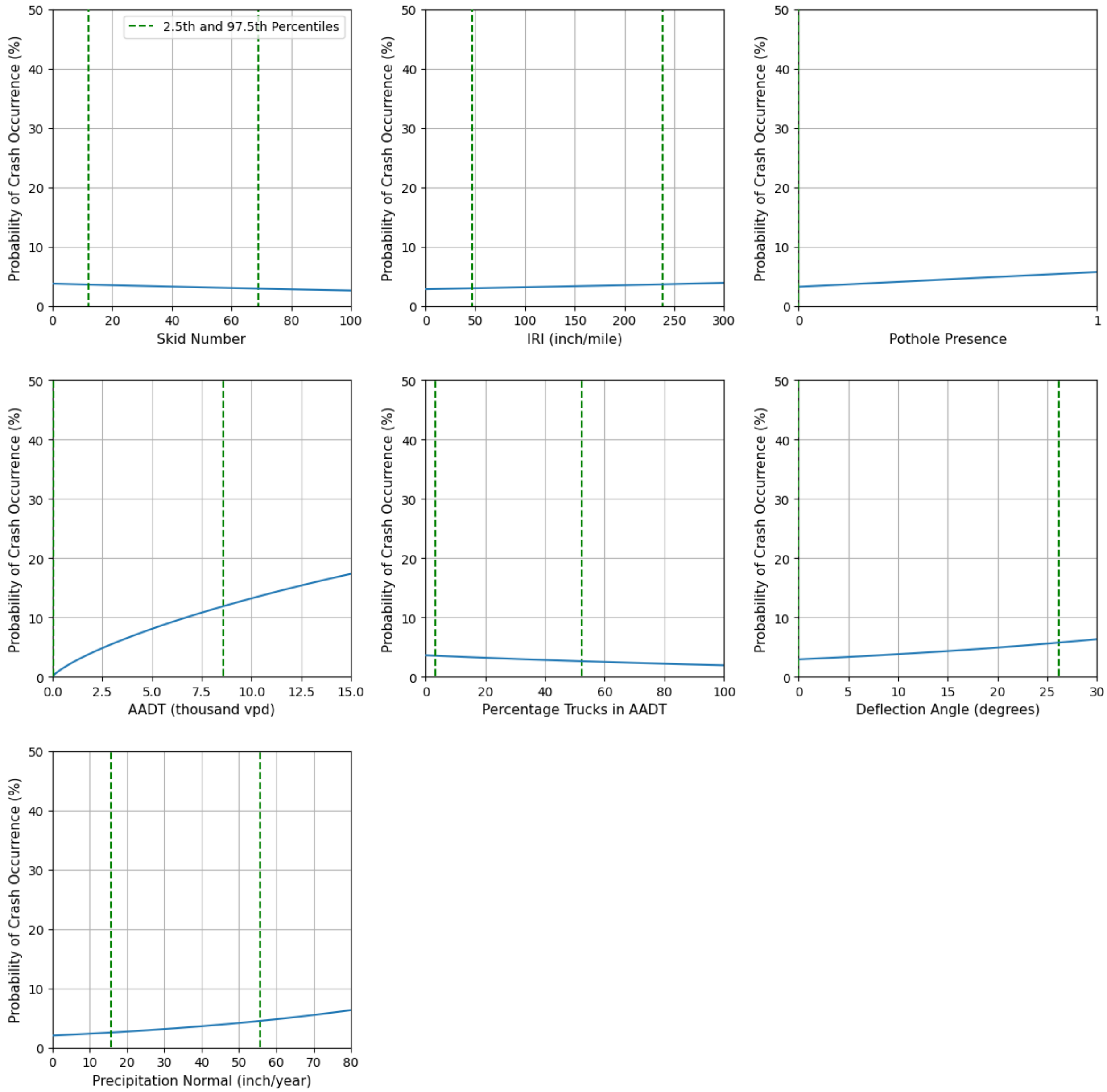



**Figure 13. Sensitivity Analysis of ACP with ST Model.**

For both models, AADT had the greatest impact on crash occurrence likelihood. Deflection angle was also a major factor in both models; increasing it from 0º to 30º increased the probability of crash occurrence by nearly 6 percent. IRI was a critical factor for ACP without ST; at very high values (300 inches/mile), the probability of crash occurrence exceeded 20 percent. The effect of IRI was less pronounced on ACP with ST; crash probability increased only slightly at high IRI values. Precipitation had a moderate effect on crash susceptibility in both models. SN and truck percentage also had moderate effects on crash probability for ACP without ST, but their impact was less clear for ACP with ST. Moreover, MPD and longitudinal slope, which were only significant in the ACP without ST model, did not have a big impact on crash susceptibility.

For longitudinal slope, the probability of crash occurrence showed only a slight variation between uphill and downhill conditions. Potholes greatly increased crash susceptibility for ACP with ST; the presence of potholes increased the crash likelihood from 3.1 to 6.1 percent.

## CRCP/JCP CRASH SUSCEPTIBILITY MODELS

This section discusses the statistical models developed to assess the susceptibility of CRCP and JCP to crash occurrence.

### CRCP/JCP Model Parameters

Table 10 and Table 11 present the logistic regression model parameters for CRCP and JCP, respectively. As in the ACP models, the predictor variables included in the final models were selected by minimizing the Akaike information criterion. The inclusion of these variables significantly improved model fit, as evidenced by the increase in LL.

**Table 10. Logistic Regression Model Results for CRCP.**

| Variable | Estimate | Standard Error | Z-score |
|---|---|---|---|
| (Intercept) | –11.52 | 0.380 | -30.28 |
| IRI (inches/mile) | 0.0045 | 0.0006 | 7.21 |
| MPD (inches) | –8.96 | 3.98 | -2.25 |
| Log AADT (vpd) | 2.32 | 0.0767 | 30.24 |
| Percentage trucks in AADT (%) | –0.0076 | 0.0021 | -3.64 |
| Absolute longitudinal slope (%) | –0.0505 | 0.0280 | -1.80 |
| Precipitation normal (inches/year) | 0.0105 | 0.0020 | 5.30 |
| **Observations (number of sections)** | 10,062 | | |
| **$LL_{Null}$** | –5,650 | | |
| **$LL_{Full}$** | –4,815 | | |

AADT = annual average daily traffic; IRI = International Roughness Index; $LL_{Full}$ = log likelihood of full model; $LL_{Null}$ = log likelihood of intercept-only model; MPD = mean profile depth.

**Table 11. Logistic Regression Model Results for JCP.**

| Variable | Estimate | Standard Error | Z-score |
|---|---|---|---|
| (Intercept) | –9.73 | 0.741 | -13.12 |
| SN | –0.0184 | 0.0062 | -2.95 |
| Faulting (inches) | 3.22 | 1.33 | 2.42 |
| Log AADT (vpd) | 2.09 | 0.159 | 13.14 |
| Percentage trucks in AADT (%) | –0.0107 | 0.0052 | -2.07 |
| Precipitation normal (inches/year) | 0.0099 | 0.0049 | 2.03 |
| **Observations (number of sections)** | 1,881 | | |
| **$LL_{Null}$** | –1,030 | | |
| **$LL_{Full}$** | –927 | | |

AADT = annual average daily traffic; $LL_{Full}$ = log likelihood of full model; $LL_{Null}$ = log likelihood of intercept-only model; SN = skid number.

Similar to the ACP models, the predictors of crash occurrence in the CRCP and JCP models included pavement-condition-, traffic-, roadway-geometry-, and weather-related predictors.

*Pavement-Condition-Related Predictors*

For the CRCP model, IRI and MPD were statistically significant variables. As with the ACP models, IRI showed a positive correlation with crash susceptibility, while MPD showed a negative correlation. For the JCP model, SN and faulting were significant variables. SN was negatively associated with crashes, consistent with the ACP model findings. Faulting had a positive correlation with crash occurrence in the JCP model; faulting may introduce abrupt surface roughness, potentially decreasing driver control and increasing crash likelihood.

*Traffic-Related Predictors*

Similar to the ACP models, both AADT and truck percentage were statistically significant in both the CRCP and JCP models. AADT showed a positive correlation with crash occurrence and was the most significant variable based on the Z-score. This outcome was intuitive because higher exposure would naturally lead to more crashes. In contrast, truck percentage was negatively correlated with crash occurrence, possibly because the presence of trucks encourages greater driver caution and reduces reckless behavior.

*Roadway-Geometry-Related Predictors*

The effect of roadway geometry on crash susceptibility was less pronounced in JCP and CRCP models compared to the ACP models. The only significant variable was the absolute longitudinal slope in the CRCP model, which showed a negative correlation with crash occurrence. This pattern may reflect the fact that CRCP and JCP are predominantly used on major highways, where roadway alignment is generally subjected to higher design standards.

*Weather-Related Predictors*

The precipitation normal was significant in both the JCP and CRCP models, with higher precipitation associated with higher crash susceptibility.

### CRCP/JCP Model Evaluation

Table 12 shows the performance of the CRCP and JCP models based on the F2-score, recall, precision, and accuracy. Similar to the approach used when evaluating the ACP models, this evaluation was performed using the test dataset (unseen data) at a prediction threshold that maximized the F2-score. The CRCP model had slightly better performance than the JCP model, as evidenced by a higher F2-score (0.665 vs. 0.576 for the JCP model). Precision and accuracy were also higher for the CRCP model compared to the JCP model, with precision values of 0.322 vs. 0.240 and accuracy values of 0.514 vs. 0.458. However, the JCP model marginally outperformed the CRCP model regarding recall, with recall values of 0.883 vs. 0.888.

**Table 12. CRCP and JCP Model Performance.**

| Evaluation Metric | CRCP | JCP |
|---|---|---|
| F2-score | 0.655 | 0.576 |
| Recall | 0.883 | 0.888 |
| Precision | 0.322 | 0.240 |
| Accuracy | 0.514 | 0.458 |

### CRCP/JCP Model Sensitivity

Researchers performed a sensitivity analysis on the CRCP and JCP models to determine how changes in each predictor variable affected the predicted crash susceptibility, while keeping all other variables fixed at their mean values. Figure 14 and Figure 15 show the results of this analysis for the CRCP and JCP models, respectively. The 2.5$^{th}$ and 97.5$^{th}$ percentiles are displayed on the sensitivity charts to represent the predictor bounds, highlighting the effect of each variable on crash susceptibility between these extreme values. The probability of crash occurrence in any given year, with all predictor variables set at their mean values, was 26.3 percent for CRCP and 28.5 percent for JCP.

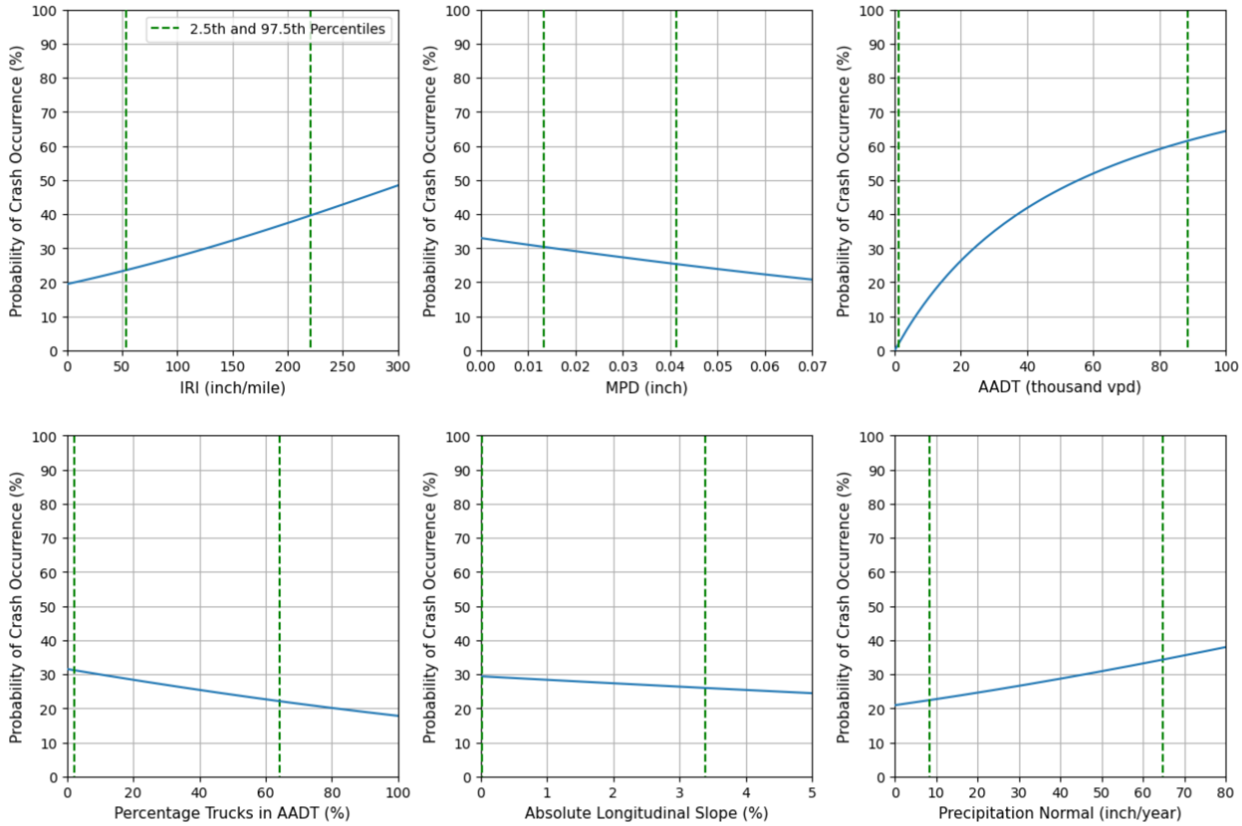



**Figure 14. Sensitivity Analysis of CRCP Model.**

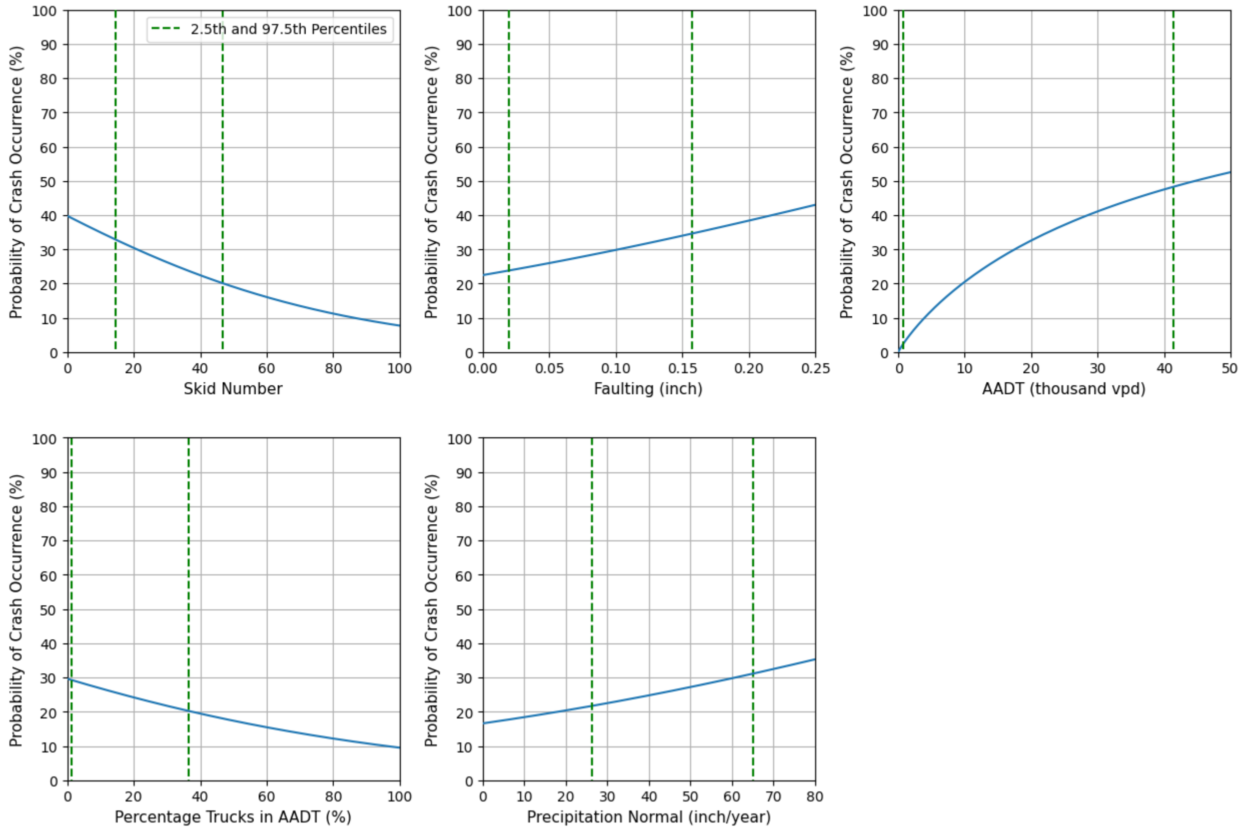



**Figure 15. Sensitivity Analysis of JCP Model.**

In both models, AADT had the greatest effect on crash susceptibility, increasing from nearly 0 at the 2.5$^{th}$ percentile (very low traffic volume) to more than 50 percent at the 97.5$^{th}$ percentile (very high traffic volume), while holding all other predictors at their mean values. The percentage of trucks had a weaker effect. Increasing the truck percentage from 0 to 50 percent reduced crash susceptibility by approximately 9 percent for CRCP and 12 percent for JCP.

Pavement conditions had moderate to strong impacts on crash susceptibility in both CRCP and JCP models. IRI and MPD had moderate to strong impacts on crash susceptibility in the CRCP model. For example, the probability of crash occurrence rose to approximately 40 percent at the 97.5$^{th}$ percentile IRI (≈ 220 inches/mile), demonstrating a strong impact. MPD had a moderate impact in the CRCP model, with a noticeable decrease in the crash susceptibility at higher values. In the JCP model, faulting and SN significantly impacted crash likelihood. Increasing SN from 15 to 45 decreased the JCP crash susceptibility from 45 to 20 percent, and increasing faulting from 0 to 0.25 inches increased JCP crash susceptibility from 23 to 43 percent.

The precipitation normal had a moderate influence on the probability of crash occurrence in both CRCP and JCP models. For both pavement types, as the precipitation normal increased from 0 to 80 inches/year, the probability of crash occurrence rose by nearly 17 percent. The impact of longitudinal slope appeared low; increasing longitudinal slope from 0 to 5 percent only slightly decreased the probability of crash occurrence in the CRCP model.

# CHAPTER 5:
# HYDROPLANING RISK ANALYSIS

Hydroplaning crash risk is influenced by WFT, driver behavior, and vehicle condition, but only WFT can be managed through pavement design and maintenance. Researchers combined WFT modeling with artificial-intelligence-driven crash narrative analysis to identify and evaluate pavement-related factors contributing to hydroplaning crashes.

## DATA AND METHODS USED FOR HYDROPLANING RISK ANALYSIS

### Data Processing

As noted in Chapter 3, 24,873 crashes were recorded in FY 2021–2023. Of these crashes, 4,376 (17.6 percent) occurred under wet-pavement conditions. To ensure that precipitation data are reasonably accurate, only wet-pavement crashes with available precipitation data within a distance of 10 miles were considered. Only 1,835 crashes of the 4,376 wet-pavement crashes met this criterion.

To identify crashes that involved hydroplaning, researchers applied an LLM to the crash narratives text, inferring potential hydroplaning involvement from context. By interpreting the intent and semantics of each description, this LLM approach avoided the false positives and missed cases that occurred with simple keyword searches. This method thus improved the identification of hydroplaning crashes, particularly when crash reports did not explicitly use the term “hydroplaning” or contained ambiguous language.

Police crash reports seldom include a structured field for hydroplaning, even though the narrative may describe wet conditions or loss-of-control events. Researchers used advanced natural language processing to extract such details reliably. Prior studies found that hydroplaning is often underreported in coded crash data but can be identified through narrative analysis (Das et al., 2020; National Academies of Sciences, Engineering, and Medicine, 2021). Building upon this foundation, researchers employed a prompt-based GPT-4 classifier to label each crash narrative as either hydroplaning or not hydroplaning, based solely on the written description.

Researchers compiled a dataset of 24,873 police-reported crash narratives (FY 2021–2023). Each record included a unique crash ID, crash date, and investigator narrative. Using a standardized GPT-4 prompt, each narrative was then classified as “hydroplaning” or “not hydroplaning.” The prompt was designed to focus on factual road and vehicle conditions, while disregarding unverified driver claims.

The model demonstrated strong contextual reasoning capabilities. In one case, a driver’s narrative mentioned “hydroplaned” even though no rain had occurred. Another narrative described a car losing control on standing water without using the words “hydroplane” or “hydroplaning.” The LLM correctly interpreted the context in both instances. In the first case, the method avoided a false hydroplaning flag, while in the second, it identified an implied hydroplaning event that a keyword-based approach would likely have missed.

After classifying all 24,873 crash narratives using the LLM, researchers focused on a filtered subset of 1,835 wet-pavement crashes selected for integration with WFT modeling. These crashes occurred under wet pavement conditions and had valid rainfall data from a TexMesonet weather station located within 10 miles of the crash site. Within this subset, the LLM classified 627 crashes (34.2 percent) as involving hydroplaning, while the remaining 1,208 crashes (65.8 percent) were classified as "other wet-pavement crashes". This classification was based on contextual cues in the narratives, such as references to standing water, loss of traction, or vehicle behavior, rather than on the presence of specific keywords. Overall, the narrative-based classification revealed that roughly one-third of the wet-pavement crashes in the subset used for WFT modeling involved hydroplaning.

**WFT Estimation**

Researchers used the Gallaway, PAVDRN, Pourhassan, and UK RRL prediction models shown in Equations 1–4 in Chapter 2 to estimate the WFT for wet-pavement crashes. In addition to pavement slope and precipitation intensity, the pavement flow path length, surface texture, and Manning's *n* needed to be estimated to estimate the WFT.

Pavement surface drainage parameters were computed using established geometric relationships (Wang & Li, 2017). The pavement flow path slope, $S_f$, represents a composite gradient that accounts for both the longitudinal slope, $S_L$, and the cross slope, $S_c$, of the roadway surface, as follows:

$$S_f = \sqrt{S_L^2 + S_c^2} \tag{15}$$

The pavement flow path length, $L_f$, is the effective length over which water travels across the pavement surface. It is determined as a function of the pavement drainage width, $W_p$, and the ratio of flow path slope to cross slope, which can be determined as follows:

$$L_f = W_p \left(\frac{S_f}{S_c}\right) \tag{16}$$

where $S_f$ is the flow path slope (feet/feet), $S_L$ is the longitudinal grade (feet/feet), $S_c$ is the cross slope (m/m), $L_f$ is the flow path length (feet), and $W_p$ is the pavement drainage width (feet). The pavement drainage width, $W_p$, was determined based on the roadway's geometric configuration. For horizontal curve sections and tangent segments on divided roads, the full pavement width was assumed to contribute to surface drainage, calculated as the number of lanes multiplied by the standard lane width. In contrast, for tangent segments on undivided roads, only half of the pavement width was considered effective. Water typically drains symmetrically from the crown toward both shoulders, so $W_p$ was equal to half the product of the number of lanes and lane width. This approach accounted for the directionality of surface runoff and yielded a realistic effective drainage width for WFT modeling.

MPD is commonly used for texture evaluation, but the empirical models applied in this study for predicting WFT required the MTD as input. Therefore, researchers converted the MPD values from TxDOT's PMIS to MTD using the following relationship proposed by Lee and Ayyala (2020) and adopted by FDOT:

$$MTD = 0.007 + 0.805 \cdot MPD \tag{17}$$

where both MPD and MTD are in inches.

To account for surface roughness in the PAVDRN model, Manning's *n* was assigned based on pavement type and macrotexture observed in the study dataset. For asphalt pavements with MPD values ≥ 0.79 mm (0.0309 inches), a coefficient of 0.016 was assigned, reflecting a higher macrotexture and a higher hydraulic resistance. Asphalt sections with MPD values below this threshold were assigned a lower coefficient of 0.013, indicating smoother surfaces. For concrete pavements, sections with MPD values ≥ 0.56 mm (0.0219 inches) were assigned a coefficient of 0.016, while sections with lower MPD values were assigned a coefficient of 0.014. These assigned *n* values align with guidance in Table 4-8 of the TxDOT *Hydraulic Design Manual* (TxDOT, 2019).

## RESULTS AND DISCUSSION

This section presents the findings from an integrated analysis of WFT, rainfall intensity, and pavement condition for 627 narrative-confirmed hydroplaning crashes. It assesses the contribution of these factors to hydroplaning risk, quantifies their influence on crash occurrence, and identifies the critical thresholds at which the likelihood of hydroplaning markedly increases. In addition, MPD and SN were analyzed across three crash categories that included 627 hydroplaning crashes, 1,208 other wet-pavement crashes, and 20,497 dry-pavement crashes. Including dry-pavement crashes as a baseline provides broader context for interpreting the distinctive pavement surface characteristics of hydroplaning crashes relative to dry-road crashes.

### WFT Estimation Results

To evaluate the role of WFT in hydroplaning, researchers estimated WFT for each of the 627 narrative-confirmed hydroplaning crashes using the four models, as well as a model-averaged value (the mean of the four predictions). Table 13 summarizes the average, maximum, and 90$^{th}$ percentile WFT values (inches) from each model. To probe model behavior under hydroplaning-prone conditions, Figure 16 shows the cumulative distribution of WFT for the 627 hydroplaning crashes across the four models (and their average).

**Table 13. WFT Estimates by Prediction Model for Hydroplaning Crashes.**

| WFT Prediction Model | Hydroplaning (n=627) | | |
|---|---|---|---|
| | Average (inches) | Maximum (inches) | 90th Percentile (inches) |
| UK RRL | 0.0039 | 0.0640 | 0.0143 |
| Gallaway | 0.0032 | 0.0640 | 0.0111 |
| PAVDRN (analytical form) | 0.0004 | 0.0244 | 0 |
| Pourhassan | 0.0867 | 0.3624 | 0.1812 |
| Model-averaged WFT | 0.0235 | 0.1287 | 0.0516 |

UK RRL = United Kingdom Road Research Laboratory; WFT = water film thickness.

The four WFT prediction models exhibited distinct performance patterns. The PAVDRN model predicted a WFT value of zero for more than 90 percent of the hydroplaning crashes. The UK RRL and Gallaway models produced flatter cumulative curves than the PAVDRN model, predicting a WFT value < 0.01 for more than 85 percent of the hydroplaning crashes. This finding suggested that these two models (UK RRL and Gallaway) were more responsive to varying conditions, with estimates that were seemingly more realistic.

The Pourhassan model yielded a much flatter curve in the low-WFT range. It predicted fewer hydroplaning crashes at very low water depths and shifted the cumulative increase toward higher WFT levels. This model produced substantially higher WFT estimates (up to 0.36 inches; see Table 13), but its gradual rise implied reduced sensitivity to incipient hydroplaning. Notably, a sharp uptick in the Pourhassan model's cumulative curve occurred between WFT values of 0.10 and 0.25 inches, indicating that the model attributed a large share of hydroplaning crashes to relatively deep water films.

The model-averaged composite fell between these extremes, smoothing out the individual model biases. A key finding from Figure 16 was that 90 percent of hydroplaning crashes occurred at model-averaged WFT values < 0.05 inches. Traditionally, a WFT of 0.095–0.10 inches has long been recognized as the approximate breakpoint between incipient and full dynamic hydroplaning, where hydrodynamic lift fully supports the tire's weight (Anderson et al., 1998; Gallaway et al., 1979; Huebner et al., 1986). However, the observed distribution suggested that a high proportion of hydroplaning crashes occurred when the WFT was < 0.095 inches. Instead, a threshold near 0.05 inches may represent a more practical operational indicator of hydroplaning risk, effectively lowering the traditional dynamic hydroplaning threshold. The model-averaged WFT across four models showed that 90 percent of the 627 narrative-confirmed hydroplaning crashes occurred below this value.

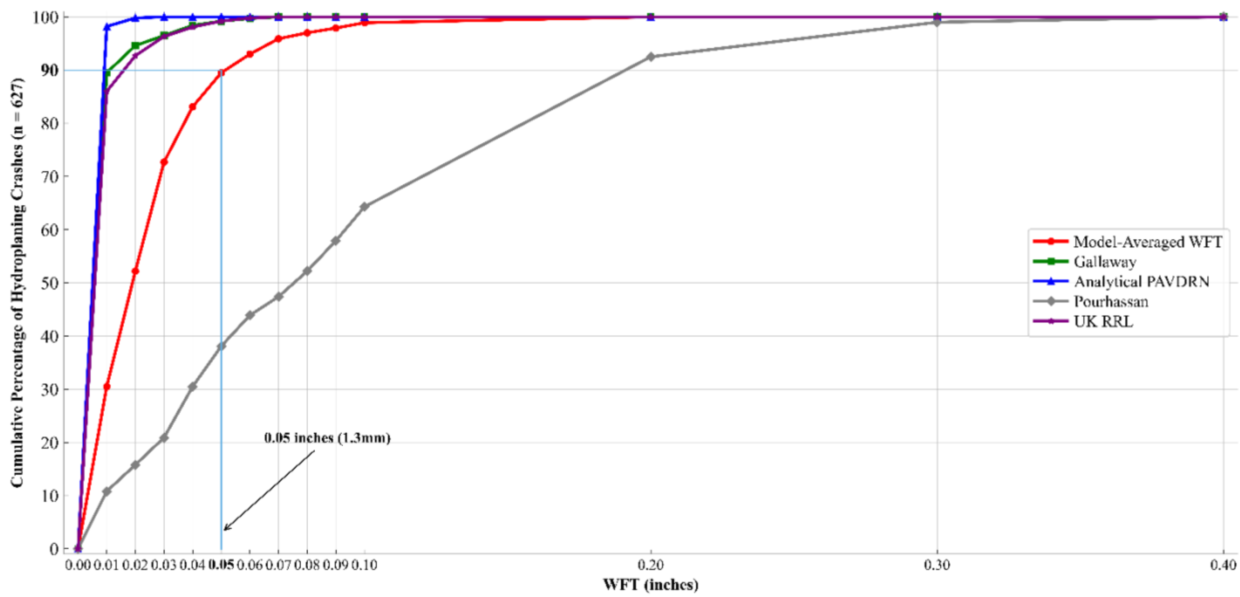



**Figure 16. Cumulative Distribution of Estimated WFT Values for Hydroplaning Crashes from Four Models and Their Average.**

Figure 17 reinforces this interpretation, indicating that 30.5 percent of hydroplaning crashes occurred at model-averaged WFT values ≤ 0.01 inches, which points to other contributing factors such as excessive speed, inadequate tire tread depth, or driver overcorrection when water depth is minimal. These results underscored that hazardous hydroplaning conditions frequently develop well before the classical WFT 0.095–0.10-inch threshold is reached. The concentration of crashes for WFT values < 0.05 inches suggested that this lower value may represent a more practical and safety-relevant trigger point for roadway monitoring, vehicle warning systems, and targeted countermeasures in real-world operational settings.

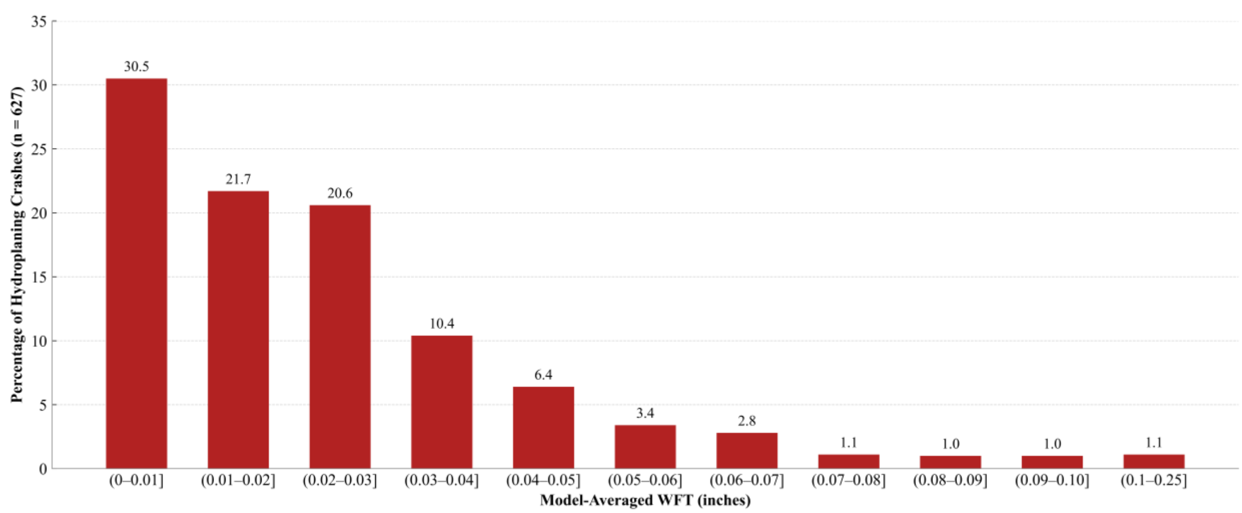



**Figure 17. Distribution of Hydroplaning Crashes by Model-Averaged WFT.**

### Influence of Rainfall Intensity on Hydroplaning Risk

Rainfall intensity is a critical environmental factor influencing hydroplaning risk. Higher precipitation rates result in deeper water accumulation on the pavement surface, increasing the likelihood of tire-pavement separation. In this study, rainfall intensity was categorized as light (≤ 0.01 inches/hour), moderate (0.01–0.25 inches/hour), and heavy (> 0.25 inches/hour). These categories aligned with classifications used in previous wet-weather traffic safety research (Gunaratne, 2012; Hranac et al., 2006).

As shown in Figure 18, most hydroplaning crashes (73 percent) occurred during moderate to heavy rainfall (≥ 0.01 inches/hour), with 18.3 percent of crashes taking place in intense downpours exceeding 0.25 inches/hour. Light rain conditions (≤ 0.01 inches/hour) still accounted for 27 percent of hydroplaning crashes—a pattern consistent with the WFT analysis, where nearly 30 percent of crashes occurred at model-averaged WFT values ≤ 0.01 inches. This overlap indicates that hazardous hydroplaning conditions are not exclusive to severe rainfall and can arise across a broad range of precipitation intensities. Notably, most hydroplaning risk models, including those applied in this study, were developed under high rainfall intensities. While such conditions accelerate water buildup, they may overlook risks from lower-intensity conditions, where poor drainage, worn pavement, or unfavorable vehicle conditions can still trigger hydroplaning. The notable share of crashes in light rain underscored the need for models that explicitly account for such conditions.

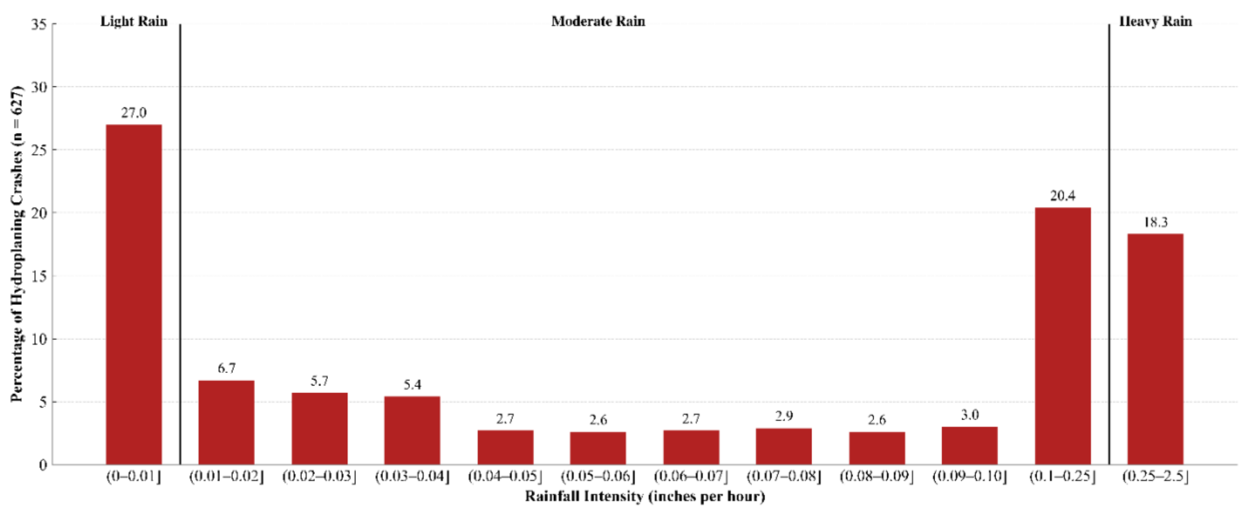



**Figure 18. Distribution of Hydroplaning Crashes by Rainfall Intensity.**

### Pavement Texture and Friction

Researchers examined how MPD and SN varied across hydroplaning crashes (n = 627), other wet-pavement crashes (n = 1,208), and dry-pavement crashes (n = 20,497), normalizing distributions within each crash type to facilitate comparison. Across the statewide network of 189,700 segments, average values for MPD and SN were 0.046 inches and 37.7, respectively.

These baseline values provided context for interpreting crash distributions relative to typical pavement conditions.

Figure 19 shows that hydroplaning crashes are disproportionately concentrated on pavements with low macrotexture. Specifically, 40.8 percent of hydroplaning crashes occurred on road segments with MPD values ≤ 0.020 inches, compared to 32.4 percent of other wet-pavement crashes and 27.9 percent of dry-pavement crashes in that same texture range. The highest proportion of hydroplaning crashes (27.4 percent) had MPD values of 0.015–0.020 inches, indicating that many crashes occurred on minimally textured surfaces. This finding aligns with tire-pavement interaction principles—low macrotexture impedes drainage, increasing WFT and hydroplaning risk (Federal Highway Administration, 2023). Henry (2000) reported that reduced macrotexture raised hydroplaning propensity, and Roe et al. (1998) found wet-weather crash rates nearly doubled below MPD values of ≈ 0.015–0.020 inches, which underscored the importance of adequate macrotexture.

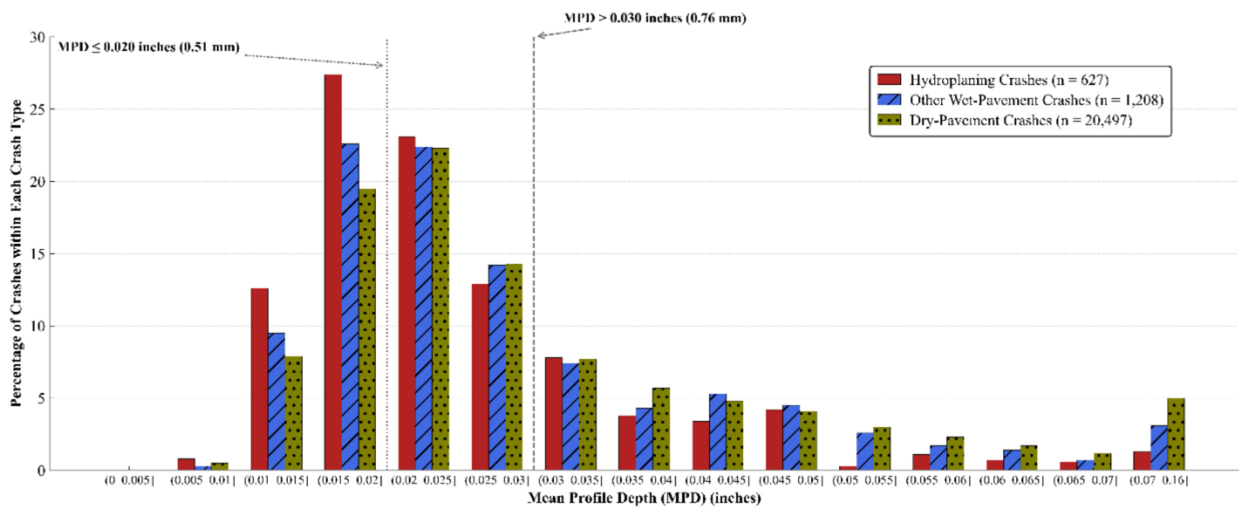



**Figure 19. Distribution of Hydroplaning, Other Wet-Pavement, and Dry-Pavement Crashes by MPD.**

At the opposite end of the texture spectrum, very few hydroplaning crashes occurred on high-texture pavements. Importantly, all three crash categories showed a noticeable decline in crash frequency when MPD values exceeded 0.030 inches, suggesting that a 0.030-inch MPD may represent a critical macrotexture threshold for safety (when texture is above this level, both wet- and dry-pavement crashes became less common). The consistency of this trend across hydroplaning, other wet-pavement, and dry-pavement crashes indicated that increasing macrotexture beyond 0.030 inches could yield broad safety benefits. In other words, ample macrotexture can enhance safety in all weather conditions, not just during rain. A high-texture surface provides drainage channels for water, maintains tire contact and friction even at highway speeds, and thereby prevents both hydroplaning and ordinary skidding.

Figure 20 shows the effects of SN on hydroplaning, other wet-pavement, and dry-pavement crashes. Most wet-weather crashes, including both hydroplaning and other wet-pavement crashes, occurred on low-friction segments. Specifically, 60.5 percent of hydroplaning crashes happened on segments with SN values ≤ 30, compared to 58.8 percent of other wet-pavement crashes and 50.4 percent of dry-pavement crashes. This overrepresentation of wet-pavement crashes on low-friction surfaces suggested that low skid resistance was a general contributing factor in wet-weather crash risk, not only in hydroplaning-specific crashes. Conversely, high-friction pavements were associated with far fewer crashes. For SN values > 40, crash frequencies dropped significantly across all categories. These results confirmed that maintaining SN values of 40 or higher can substantially reduce both wet- and dry-pavement crashes.

Overall, these findings highlighted that both pavement texture and pavement friction are key determinants of crash risk across all weather conditions. Hydroplaning crashes, in particular, were concentrated in areas with insufficient macrotexture (MPD ≤ 0.020 inches) and low skid resistance (SN ≤ 30). Conversely, areas with sufficient macrotexture (MPD > 0.030 inches) and high-friction surfaces (SN > 40) experienced significantly fewer crashes, with wet or dry pavements.

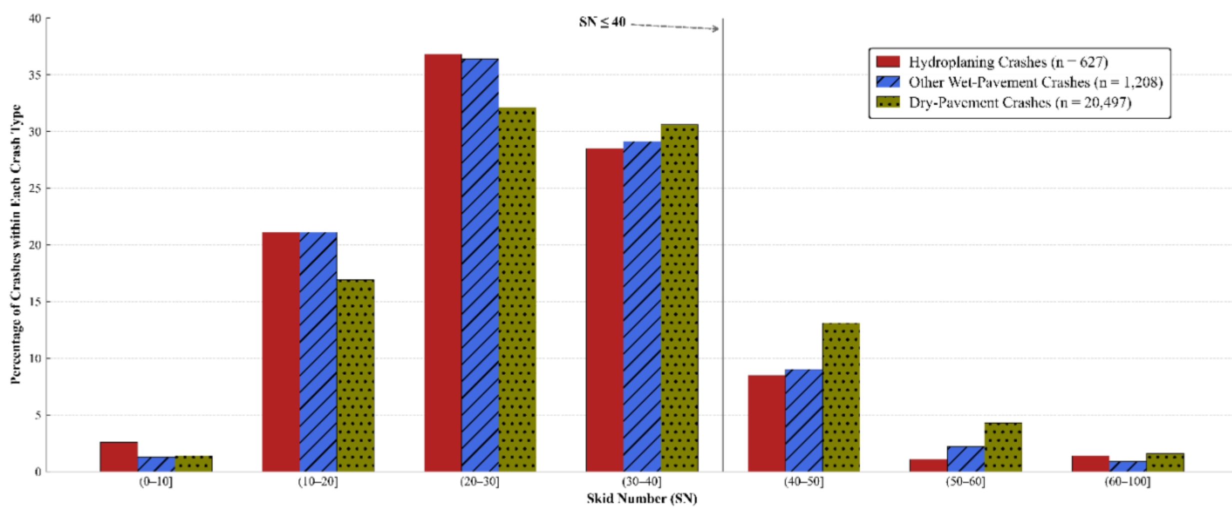



**Figure 20. Distribution of Hydroplaning, Other Wet-Pavement, and Dry-Pavement Crashes by SN.**

Collectively, these findings highlighted the importance of maintaining pavement macrotexture and friction to reduce both dry- and wet-pavement crash risk. Frequent pavement surface inspection and timely application of surface treatments can substantially reduce the occurrence of hydroplaning crashes. Overall, combining physics-based WFT models with artificial-intelligence-driven crash analysis provided guidance for identifying high-risk pavement segments and pavement improvement decisions. Expanding weather station coverage, leveraging additional pavement condition data, and evaluating different pavement types under varied rainfall and geometric conditions could further refine pavement safety improvement strategies.

# CHAPTER 6:
# INTERACTIVE GIS MAP TOOL FOR SAFETY ANALYSIS

As part of this study, the University of Houston research team developed a web-based map application to help users visualize and analyze pavement safety and crash data in a user-friendly way. This tool allows engineers and researchers to explore how road conditions impact safety at both the pavement section and network levels across various TxDOT districts. It provides interactive controls and filters that allow users to navigate to different districts and visualize the safety-related pavement conditions, crash susceptibility predictions, and crash locations. This visualization aims to assist the user in identifying candidate pavement sections for pavement safety improvement. Users can also use filters such as pavement type, condition indicator, traffic, and crash frequency to modify pavement sections displays. The appendix provides additional information about the design and functionalities of this tool. The tool is accessible at https://www.uh.edu/~lgao5/geomap/ (see Figure 21).

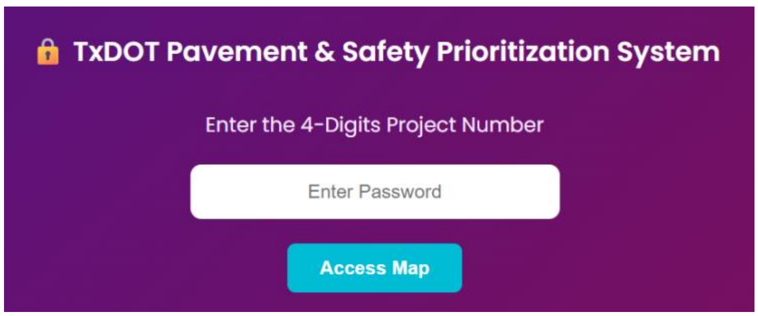



**Figure 21. Password-Protected Entry to the GIS Map Tool.**

The tool includes a District drop-down menu to select any of TxDOT's 25 districts. Selecting a district displays all pavement sections in it. TxDOT's Atlanta District was used as an example in this report to demonstrate the tool's functionality. Figure 22 displays the pavement segments in TxDOT's Atlanta District as an example.

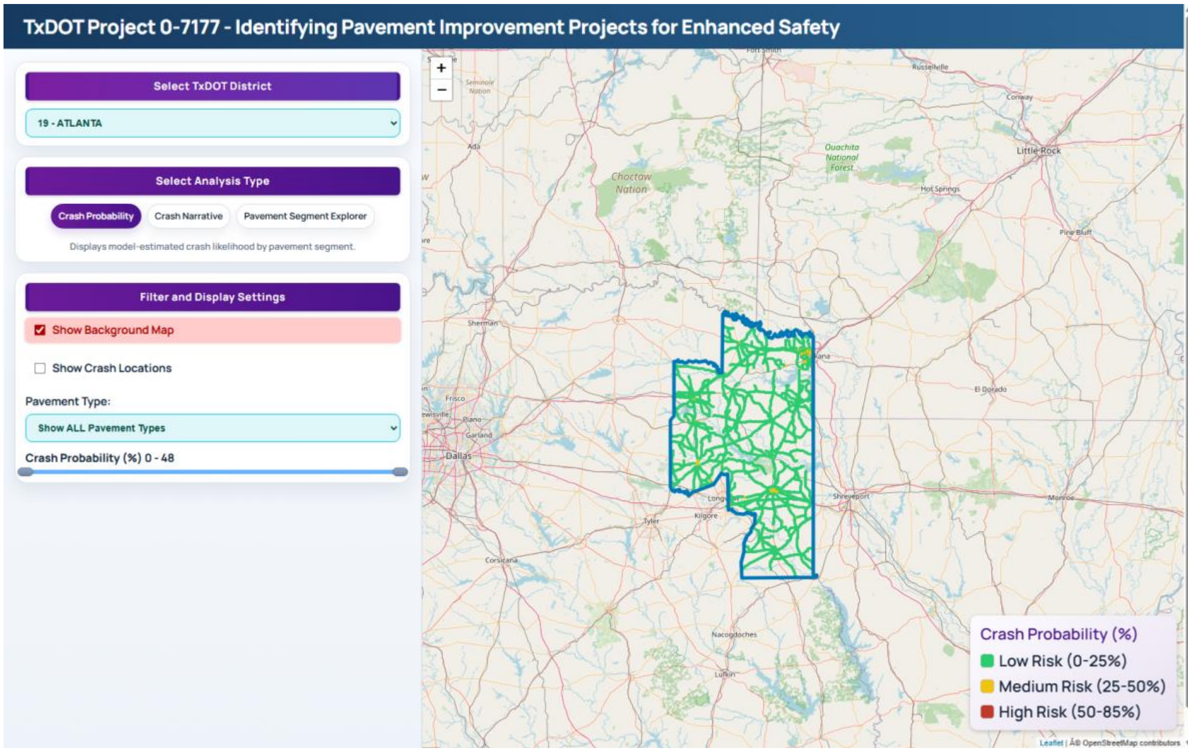



**Figure 22. Map Interface Showing District Selection, Analysis Selection, Filter Controls, and District-Level Pavement Sections.**

## CRASH PROBABILITY

The tool's Crash Probability function allows users to visualize predicted crash risk across pavement sections based on the crash susceptibility models developed in this project. Users can refine the display using crash-probability-specific controls, including pavement type and crash probability range. These controls restrict the displayed network to the selected pavement families and selected crash-probability ranges. Figure 23 shows the segment-level crash probability predictions for TxDOT's Atlanta District as an example. Pavement sections are labeled as low (0–25 percent), medium (25–50 percent), and high (50–85 percent) risk and color-coded in green, yellow, and red, respectively. The crash probability represents the chance that at least one crash will occur on that segment in a given year. Users can select the Street View link from a crash or pavement-segment popup to view the pavement section in Street View (Figure 24).

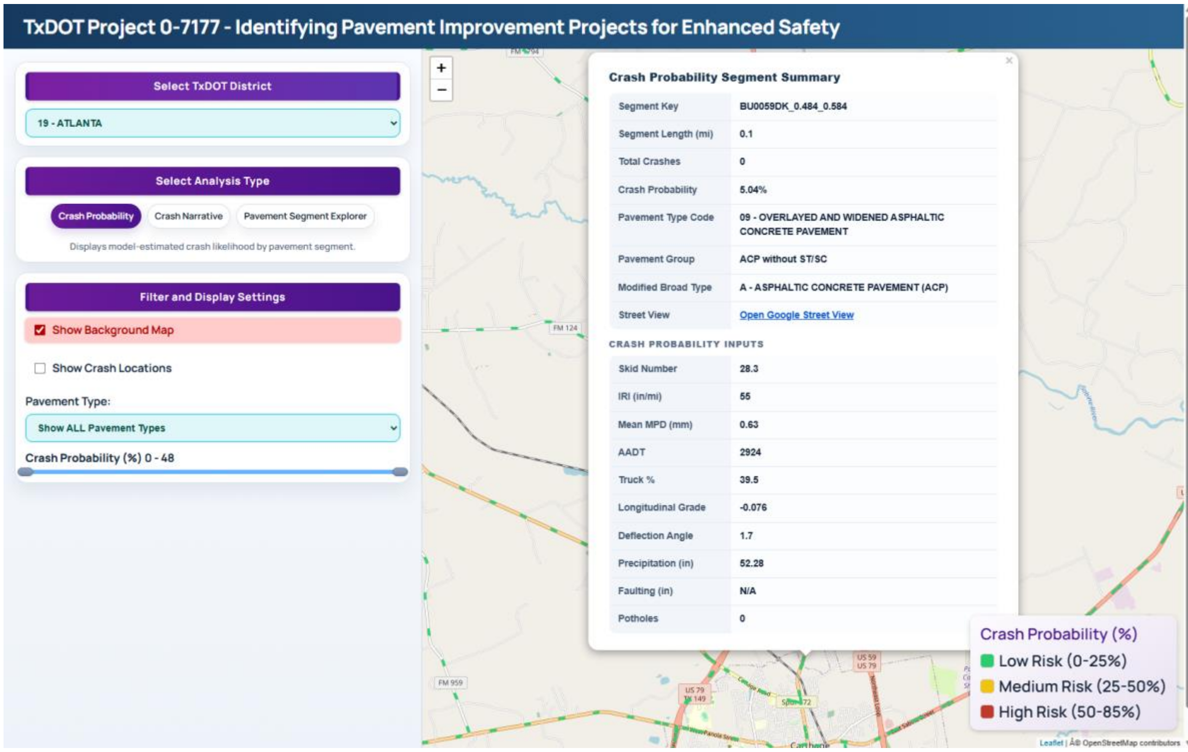



**Figure 23. Predicted Crash Probabilities for a Pavement Section in TxDOT's Atlanta District.**

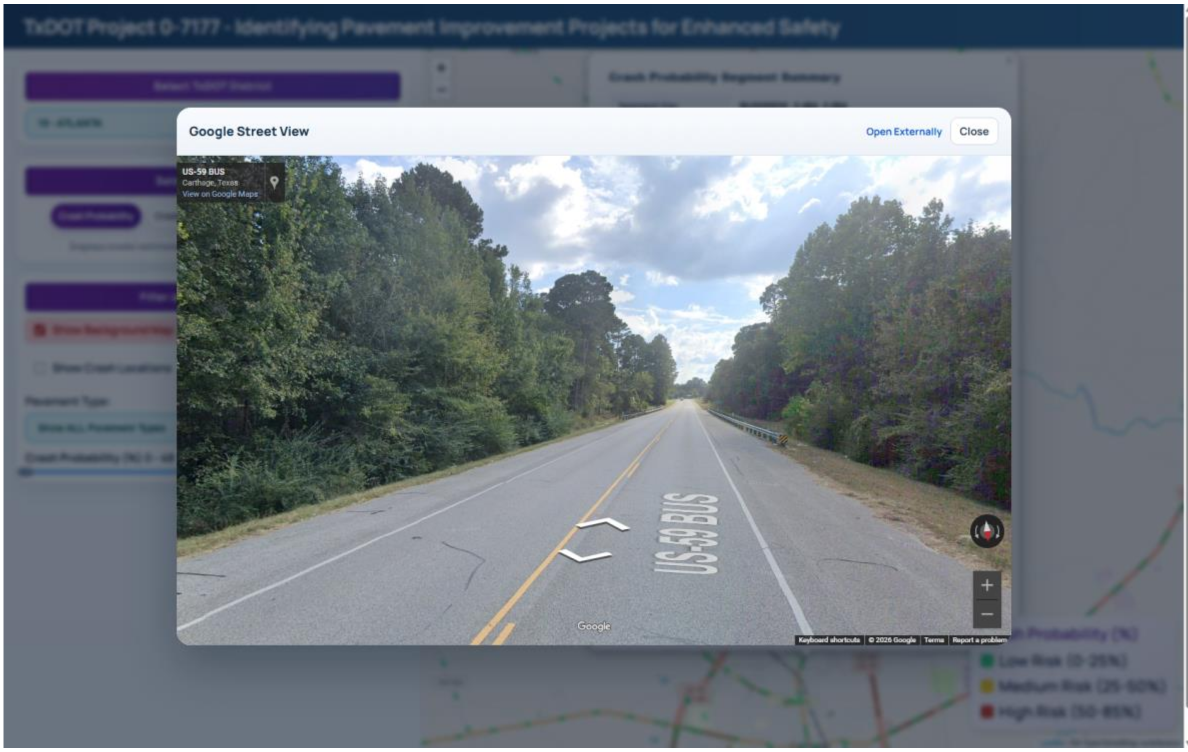



**Figure 24. In-Application Google Street View Window Displaying the Selected Pavement Section.**

The tool allows users to display crash locations for a selected district via the Show Crash Locations check box. Figure 25 displays the crash locations for TxDOT's Atlanta District, where the red circles represent the crash locations.

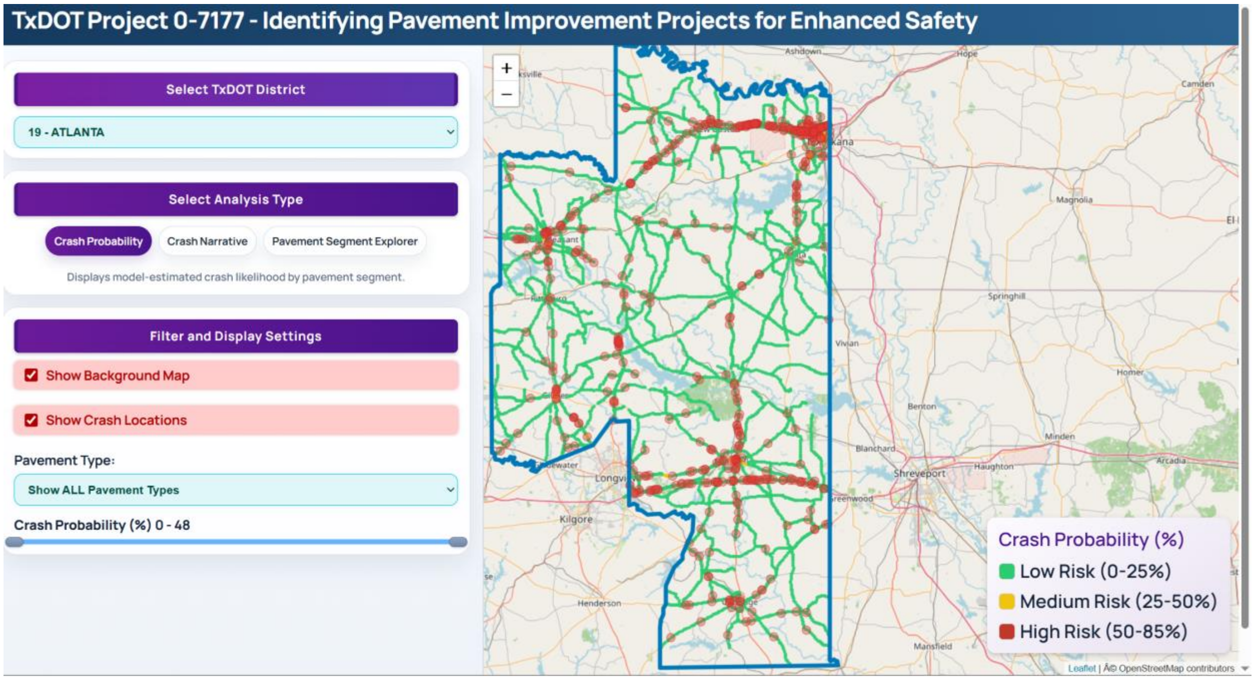



**Figure 25. Crash Locations in TxDOT's Atlanta District.**

## CRASH NARRATIVE

The tool includes a Crash Narrative Analysis function that can display pavement sections where hydroplaning-, rough-pavement-, curve-, or pothole-related crashes were documented in crash narratives. For example, the hydroplaning slider, ranging from 0 to 10, allows users to filter pavement sections based on the number of hydroplaning-related crashes occurring on that section in a given period. Figure 26 displays pavement sections where at least 1 hydroplaning-related crash was recorded in FY 2021–2023. Figure 27 displays pavement sections where at least 1 loss-of-control-at-curves crash was recorded in FY 2021–2023, based on the crash narrative analysis.

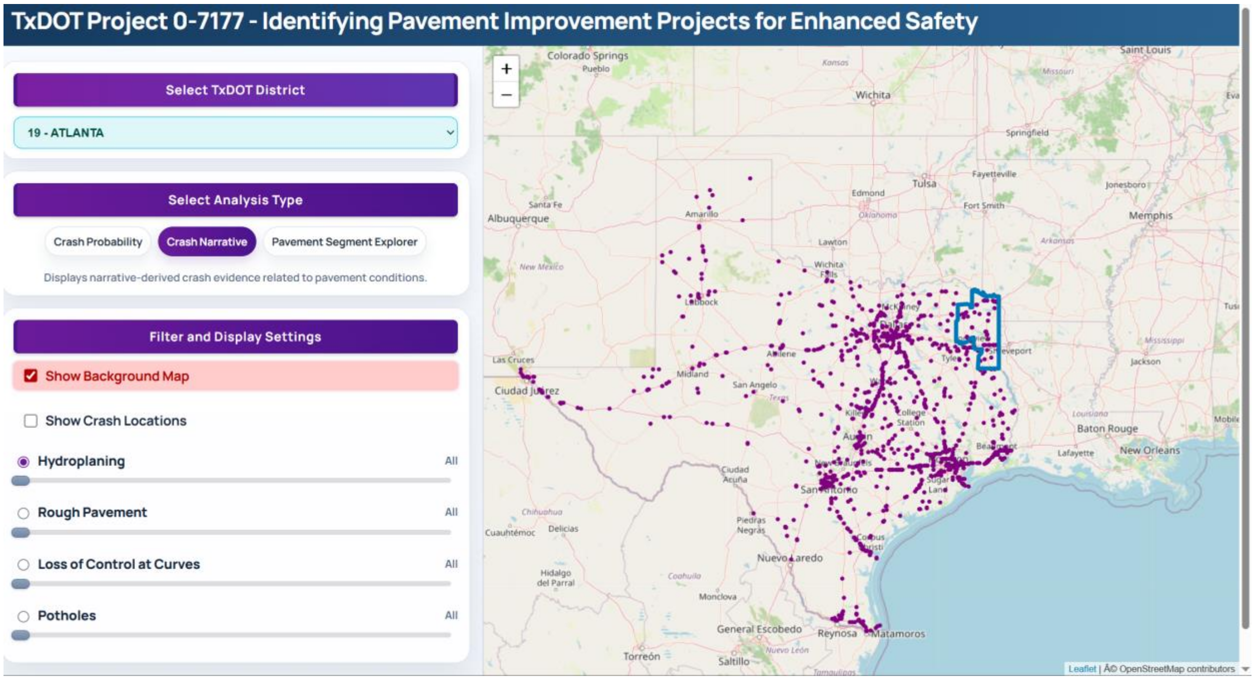



**Figure 26. Pavement Sections with at Least 1 Hydroplaning-Related Crash in FY 2021–2023.**

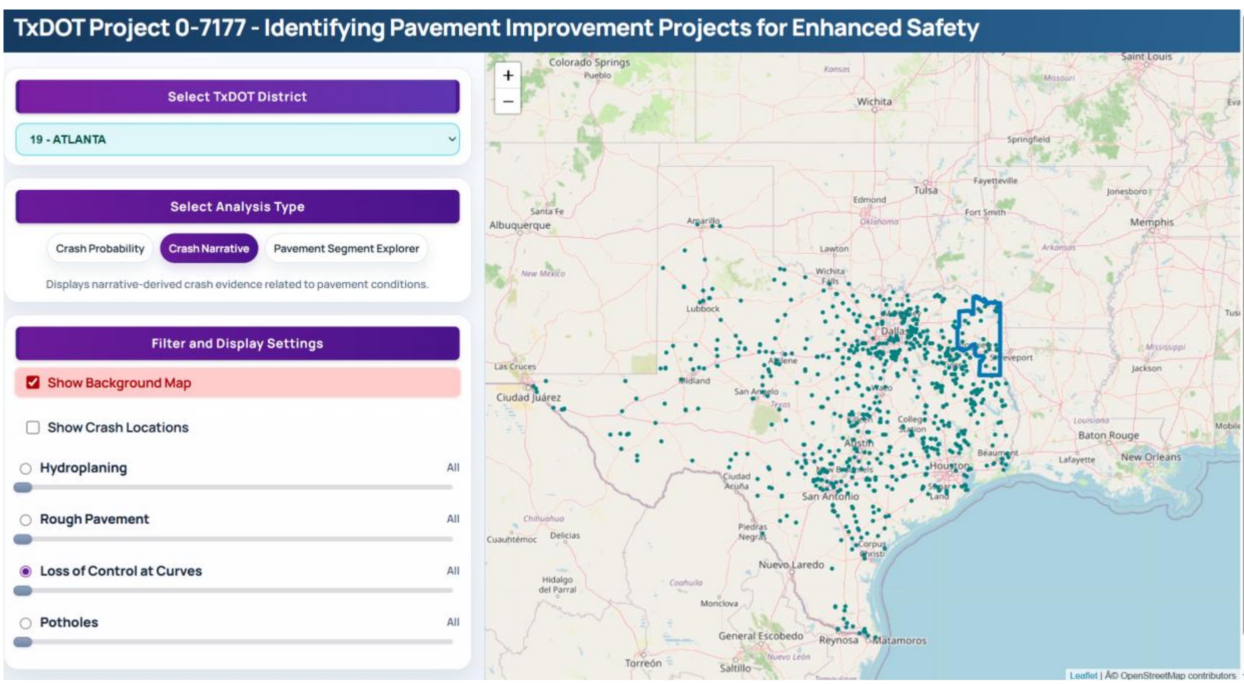



**Figure 27. Pavement Sections with at Least 1 Loss-of-Control-at-Curve Crash in FY 2021–2023, Based on the Crash Narrative Analysis.**

The tool provides a section-level pop-up window that displays segment-level narrative summary information, crash probability context, possible crash contributing factors, and a recommended safety-improvement direction. The pop-up presents the predicted crash probability for the pavement sections, along with the average crash probability of nearby segments along the same route. The average is computed using 20 upstream and 20 downstream segments (approximately 2 miles upstream and 2 miles downstream). It also summarizes possible contributing factors and provides a recommended safety improvement (Figure 28).

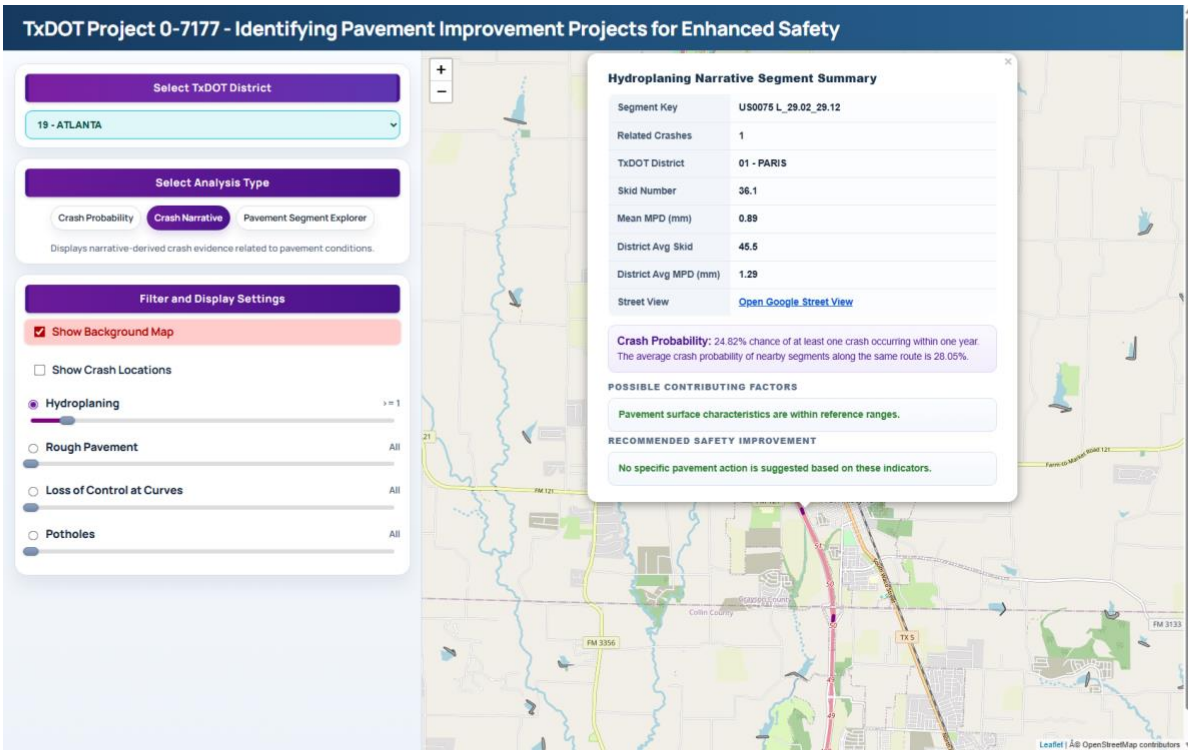



**Figure 28. Narrative Popup for a Pavement Segment with a Hydroplaning-Related Crash.**

## PAVEMENT SEGMENT EXPLORER

The Pavement Segment Explorer mode displays pavement indicators and crash-history variables across the district network and supports filtering based on combinations of those variables. For example, the user can filter pavement sections based on MAP-21 Condition Category. Pavement sections are color-coded as green (good), orange (fair), and red (poor) according to the MAP-21 overall condition category. Figure 30 displays the MAP-21 condition categories for all pavement sections in TxDOT's Atlanta District as an example.

Similarly, the tool includes a MAP-21 IRI Condition Category drop-down menu to filter pavement sections based on IRI. Pavement sections are categorized as good, fair, or poor according to the MAP-21 IRI classification criteria (Figure 30).

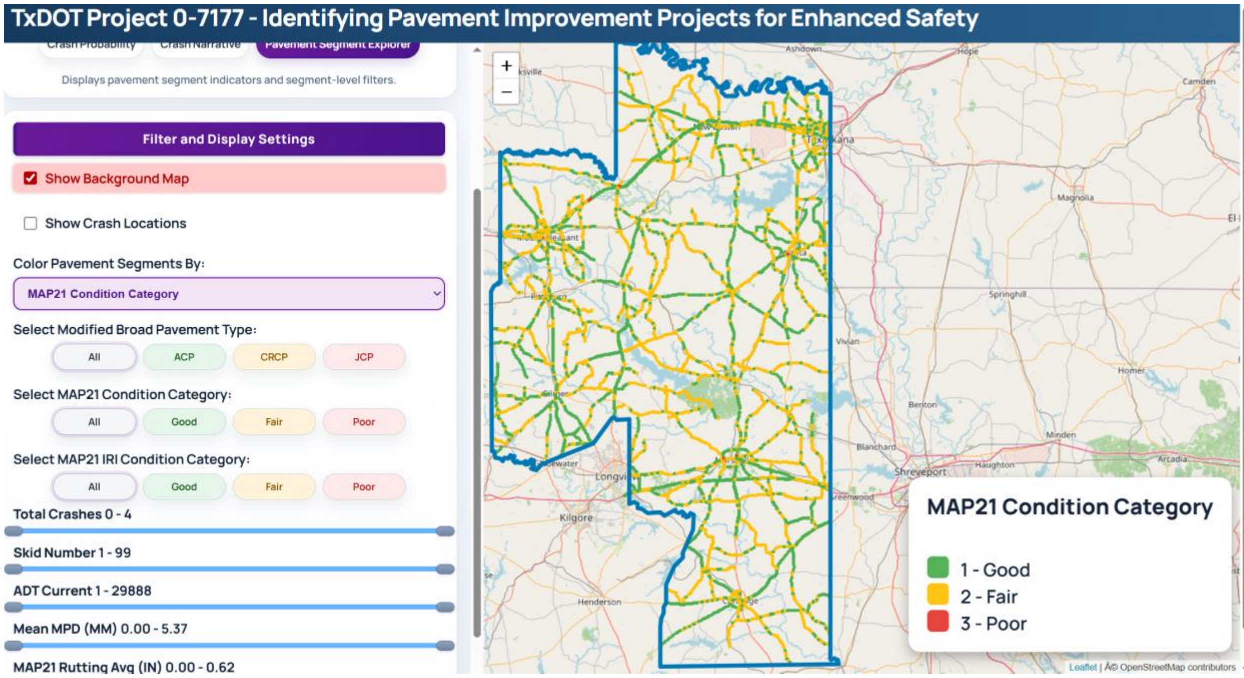



**Figure 29. Pavement Sections in TxDOT's Atlanta District by MAP-21 Condition Category.**

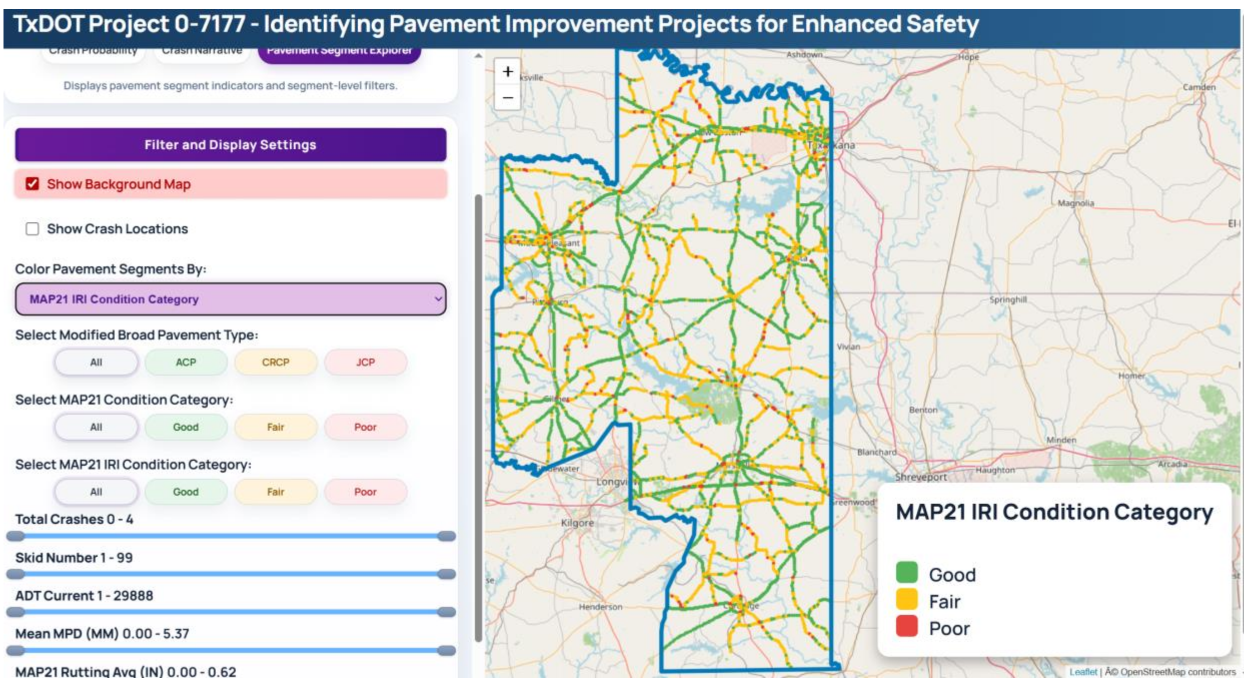



**Figure 30. Pavement Sections in TxDOT's Atlanta District by MAP-21 IRI Condition Category.**

## Slider Filters

The tool includes multiple range-based sliders that allow users to filter pavement section displays based on features such as total crashes, SN, average daily traffic (ADT), mean MPD, and MAP-

21 rutting condition categories. Users can define minimum and maximum thresholds to dynamically display segments that fall within the selected ranges.

The Total Crashes filter is provided as a dual-handle range slider spanning 0 to 4 crashes. Users can select minimum and maximum crash thresholds to display pavement sections whose total number of crashes during the study period falls within the chosen range. For example, Figure 31 shows pavement sections in TxDOT's Atlanta District color-coded based on total crash counts.

The SN filter is also provided as a dual-handle range slider spanning SN values from 1 to 99. Users can define minimum and maximum SN thresholds to display pavement sections whose measured friction levels fall within the selected range. For example, Figure 32 shows pavement sections in TxDOT's Atlanta District with SN values ≥ 40 color-coded based on pavement type.

The Mean MPD filter allows users to examine pavement sections across a range of macrotexture conditions. For example, Figure 33 shows pavement sections with Mean MPD ≥ 1 mm and SN ≥ 30, color-coded by crash count.

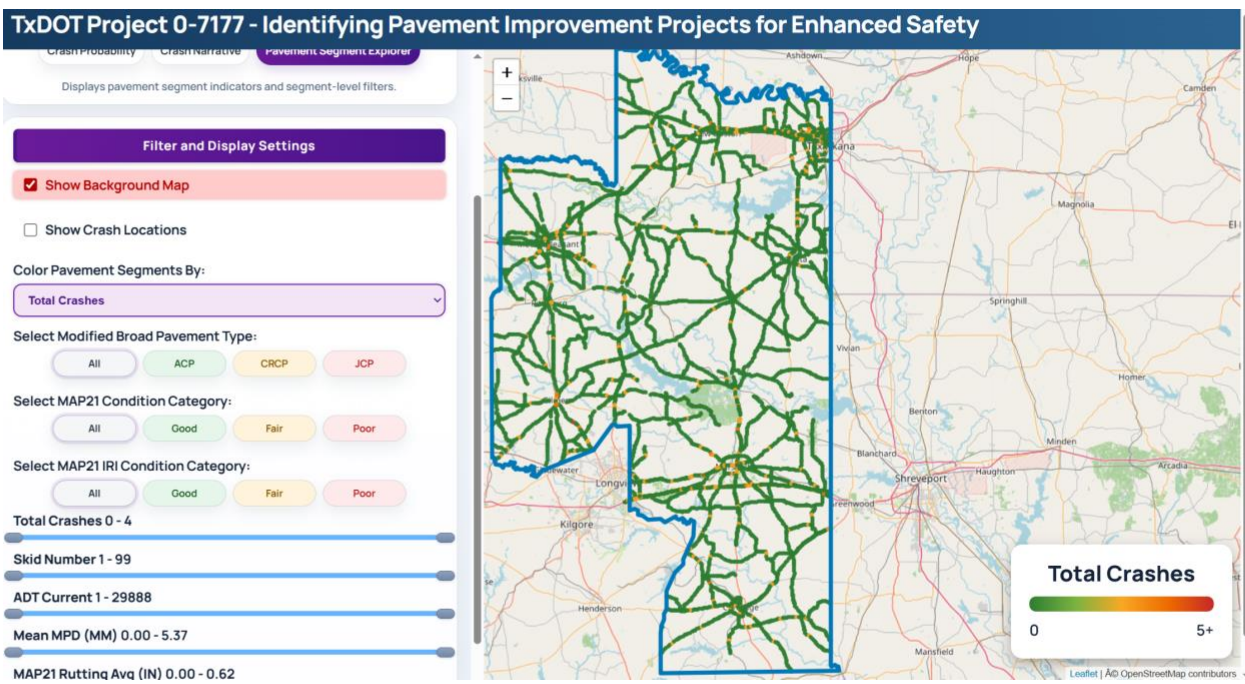



**Figure 31. Pavement Sections in TxDOT's Atlanta District Color-Coded Based on Total Number of Crashes in FY 2021–2023.**

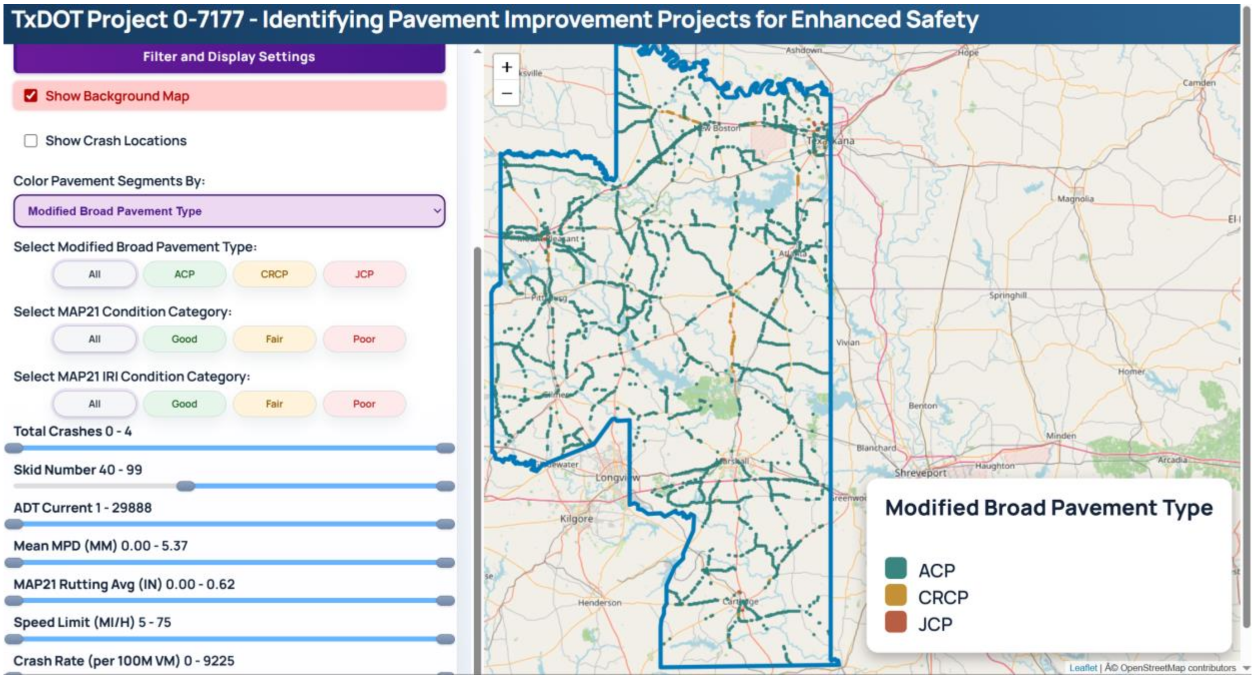



**Figure 32. Pavement Sections in TxDOT's Atlanta District with SN ≥ 40, Color-Coded by Pavement Type.**

The Current ADT filter allows users to examine pavement sections across a wide range of traffic volumes. For example, Figure 33 displays segments in TxDOT's Atlanta District with current ADT values ≤ 15,000 vpd, color-coded by crash risk category.

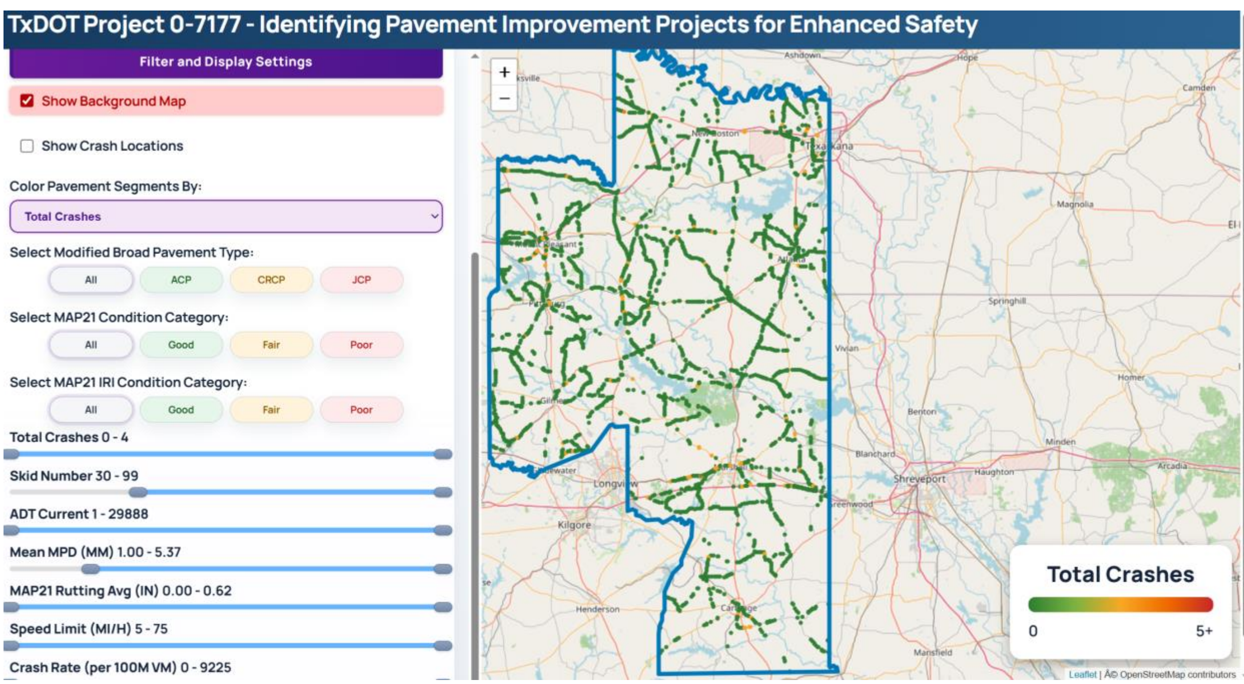



**Figure 33. Pavement Sections in TxDOT's Atlanta District with Mean MPD ≥ 1 mm and SN ≥ 30, Color-Coded by Crash Count.**

# CHAPTER 7: CONCLUSIONS AND RECOMMENDATIONS

## CONCLUSIONS

This report presented crash susceptibility models, an evaluation of pavement-related factors contributing to hydroplaning crashes, and a GIS-based tool for assessing pavement susceptibility to crashes and identifying candidate pavement safety improvement projects. The study used a large dataset from across Texas that included data for 18,843 on-system roadbed-miles and 24,873 crashes that occurred on them during a 3-year period (FY 2021–2023).

To examine the relationships between crash susceptibility (i.e., the probability of crash occurrence) and contributing factors, researchers employed logistic regression. The factors considered included pavement condition indicators, as well as traffic, road geometrics, and environmental factors. Table 14 summarizes the variables included in the crash susceptibility models for ACP with ST, ACP without ST, CRCP, and JCP.

**Table 14. Factors Affecting Crash Susceptibility for Different Pavement Types.**

| Factor | ACP with ST | ACP without ST | CRCP | JCP |
|---|---|---|---|---|
| SN | ↓ | ↓ | NS | ↓ |
| IRI (inches/mile) | ↑ | ↑ | ↑ | NS |
| MPD (inches) | NS | ↓ | ↓ | NS |
| Faulting (inches) | NA | NA | NA | ↑ |
| Pothole present (no = 0/yes = 1) | ↑ | NS | NA | NA |
| AADT (vpd) | ↑ | ↑ | ↑ | ↑ |
| Percentage trucks in AADT (%) | ↓ | ↓ | ↓ | ↓ |
| Downhill (downhill = 1/uphill = 0) | NS | ↑ | NS | NS |
| Absolute longitudinal slope (%) | NS | NS | ↓ | NS |
| Deflection angle (degrees) | ↑ | ↑ | NS | NS |
| Precipitation normal (inches) | ↑ | ↑ | ↑ | ↑ |

↑=positive correlation; ↓=negative correlation.
AADT = annual average daily traffic; ACP = asphalt concrete pavement; CRCP = continuously reinforced concrete pavement; IRI = International Roughness Index; JCP = jointed concrete pavement; MPD = mean profile depth; NA = not applicable; NS = not significant; SN = skid number; ST = surface treatment.

Regarding pavement conditions, pavement roughness (measured by IRI) and skid resistance (measured by SN) were significant in most cases. MPD had a negative correlation with crashes for CRCP and ACP without ST. Pothole presence had a positive correlation with crashes only for ACP with ST. Other pavement surface distresses such as rut depth, raveling, flushing, and failures for ACP and punchout for CRCP had no statistically significant effect on crash susceptibility. Thus, the results of this study indicated that pavement maintenance and rehabilitation treatments that reduce roughness, improve skid resistance, increase macrotexture depth, and fix potholes are likely to reduce crash occurrence likelihood.

The observed significance of traffic, roadway geometry, and environmental factors was consistent with findings from previous studies.

As part of the hydroplaning analysis, researchers considered inputs from 627 hydroplaning crashes, identified using LLM analysis of the crash narratives. Researchers estimated WFT at the time and location of these crashes using four established models. Key findings revealed that most hydroplaning crashes (90 percent) occurred when the mean WFT on the road was 0.05 inches or less—0.05 inches is significantly lower than the commonly cited threshold of 0.095 inches. Nearly three-quarters (73 percent) of hydroplaning crashes occurred during moderate to heavy rainfall; the remaining crashes occurred during light rain. These results highlighted the need for hydroplaning predictive models that account for a wide range of rainfall conditions, including light-intensity rain.

Pavement surface properties were also critical in determining the likelihood of hydroplaning-related crashes. Hydroplaning crash frequencies decreased sharply for MPD values $> 0.030$ inches, while segments with MPD values $\leq 0.020$ inches had increased numbers of hydroplaning crashes. Similarly, high-friction surfaces (SN $> 40$) experienced significantly fewer crashes in both wet- and dry-pavement conditions, whereas wet-pavement crashes clustered on low-friction sections (SN $\leq 30$).

An online GIS tool was developed to identify and highlight potential traffic safety concerns by analyzing crash probability, pavement conditions, and related roadway characteristics. It provides insight into possible contributing factors (e.g., low skid resistance, poor surface condition), which enables engineers to understand where and why elevated crash risk may exist.

## RECOMMENDATIONS

Based on the results of this study regarding the impact of pavement-related factors on crash occurrence, the research team offers the following recommendations to further support TxDOT in identifying candidate sites for pavement safety improvements:

- The crash susceptibility models indicated that proactive pavement safety improvement efforts should prioritize maintaining adequately high levels of surface friction, macrotexture, and smoothness. Such efforts are likely to reduce crash risk under both dry- and wet-pavement conditions.
- With all independent variables set to their mean values, the probability of crash occurrence in any given year was 3.1 percent for ACP with ST, 12.9 percent for ACP without ST, 26.3 percent for CRCP, and 28.5 percent for JCP. These values could serve as a reference (or benchmark) for assessing crash susceptibility for each pavement type individually. These reference values should not, however, be used to compare the safety of different pavement types because the influencing factors (e.g., traffic volume and rainfall intensity) vary across pavement types.

- Because data on actual roadway cross slopes were unavailable, researchers estimated and used the as-designed cross slopes in this study. The routine collection of actual cross slope field measurements by TxDOT could improve the accuracy of crash susceptibility models and WFT estimates in future efforts.
- Researchers recommend that the crash susceptibility models and GIS tool be implemented on a trial basis in at least one TxDOT district. Based on this pilot implementation, enhancements can then be made, and guidelines can be developed to support other TxDOT districts interested in adopting the models and GIS tool.
- Future work could include the following:
    - Integrate the crash susceptibility models into TxDOT's pavement management decision trees and 4-year pavement management planning process.
    - Validate and calibrate the new models and thresholds for different climatic regions and districts using data from FY 2024 onward.
    - Extend the GIS tool to incorporate treatment recommendation modules, in which data-driven models and guidelines suggest appropriate interventions such as surface treatments, friction improvement strategies, drainage enhancements, or geometric modifications based on identified crash risk factors and pavement conditions.

# REFERENCES


Agrawal, S. K., Meyer, W. E., & Henry, J. J. (1977). *Measurement of hydroplaning potential* (Report No. FHWA-PA-72-6). Pennsylvania Transportation Institute, Pennsylvania State University.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Al-Masaeid, H. R. (1997). Impact of pavement condition on rural road accidents. *Canadian Journal of Civil Engineering*, *24*(4), 523–531.

Anderson, D. A., Huebner, R. S., Reed, J. R., Warner, J. C., & Henry, J. J. (1998). *Improved surface drainage of pavements.* Pennsylvania Transportation Institute, Pennsylvania State University.

Anastasopoulos, P. C., Mannering, F. L., Shankar, V. N., & Haddock, J. E. (2012). A study of factors affecting highway accident rates using the random-parameters tobit model. *Accident Analysis & Prevention*, *45*, 628–633.

Arteaga, C., & Park, J. (2025). A large language model framework to uncover underreporting in traffic crashes. *Journal of Safety Research*, *92*, 1–13. https://doi.org/10.1016/j.jsr.2024.11.009

Booth, J. G., Casella, G., Friedl, H., & Hobert, J. P. (2003). Negative binomial loglinear mixed models. *Statistical Modelling*, *3*, 179–191. http://journals.sagepub.com/doi/pdf/10.1191/1471082X03st058oa

Buddhavarapu, P., Banerjee, A., & Prozzi, J. A. (2013). Influence of pavement condition on horizontal curve safety. *Accident Analysis & Prevention*, *52*, 9–18.

Buss, A., Guriguis, M., Claypool, B., Gransberg, D., & Williams, R. C. (2016). *Chip seal design and specifications* (Report No. FHWA-OR-RD-17-03). Oregon Department of Transportation, Research.

Cafiso, S., Montella, A., D'Agostino, C., Mauriello, F., & Galante, F. (2021). Crash modification functions for pavement surface condition and geometric design indicators. *Accident Analysis & Prevention*, *149*. https://doi.org/10.1016/J.AAP.2020.105887

Chan, C. Y., Huang, B., Yan, X., & Richards, S. (2010). Investigating effects of asphalt pavement conditions on traffic accidents in Tennessee based on the pavement management system (PMS). *Journal of Advanced Transportation*, *44*(3), 150–161. https://doi.org/10.1002/ATR.129

Chen, S., Saeed, T. U., & Labi, S. (2017). Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach. *Analytic Methods in Accident Research*, *16*, 75–89.

Das, S., Dutta, A., Dey, K., Jalayer, M., & Mudgal, A. (2020). Vehicle involvements in hydroplaning crashes: Applying interpretable machine learning. *Transportation Research Interdisciplinary Perspectives*, *6*. https://doi.org/10.1016/j.trip.2020.100176

Federal Highway Administration. (2023). *Pavement Friction for Road Safety: Primer on Friction Measurement and Management Methods* (Report No. FHWA-SA-23-007). U.S. Department of Transportation.

Gallaway, B. M., Schiller, R. E. Jr., & Rose, J. G. (1971). *The effects of rainfall intensity, pavement cross slope, surface texture, and drainage length on pavement water depths* (Report No. 138-5). Texas A&M Transportation Institute, Texas A&M University.

Gallaway, B. M., Ivey, D. L., Hayes, G., Ledbetter, W. B., Olson, R. M., Woods, D. L., & Schiller, R. E. Jr. (1979). *Pavement and geometric design criteria for minimizing hydroplaning* (Report No. FHWA-RD-79-31). Texas A&M Transportation Institute, Texas A&M University.

Ghanbari, A. (2017). *Impact of pavement condition indicators on road safety in New Brunswick*. [Master's thesis, University of New Brunswick].

Graves, S. J., Rochowiak, D., & Anderson, M. (2005). *Mining and analysis of traffic safety and roadway condition data* (Report No. 04310). University Transportation Center for Alabama, University of Alabama.

Gengenbach, W. (1967). *The effect of wet pavement on the performance of automobile tires* (Translated from German for limited distribution by Calspan). Universität Karlsruhe.

Grigorev, A., Saleh, K., Ou, Y., & Mihăiţă, A.-S. (2025). Enhancing traffic incident management with large language models: A hybrid machine learning approach for severity classification. *International Journal of Intelligent Transportation Systems Research*, *23*(1), 259–280. https://doi.org/10.1007/s13177-024-00448-7

Gunaratne, M., Lu, Q., Yang, J., Metz, J., Jayasooriya, W., Yassin, M., & Amarasiri, S. (2012). *Hydroplaning on multi-lane facilities* (Report No. BDK84-977-14). Florida Department of Transportation.

Hall, J. W., Smith, K. L., & Littleton, P. (2009). *Texturing of concrete pavements* (Report No. 634). National Cooperative Highway Research Program. https://doi.org/10.17226/14318

Horne, W. B. (1968). Tire hydroplaning and its effects on tire traction. *Highway Research Record*, *214*, 24–33.

Huang, X., Feng, Y., & Zhang, Z. (2024). ChatGPT-based method for generating automobile accident reports. *Proceedings of the 4th International Conference on Electronic Information Engineering and Computer Science,* 1174–1177. https://doi.org/10.1109/EIECS63941.2024.10800246

Henry, J. J. (2000). *Evaluation of pavement friction characteristics* (Synthesis of Highway Practice No. 291). National Cooperative Highway Research Program, Transportation Research Board, National Research Council.

Huebner, R. S., Reed, J. R., & Henry, J. J. (1986). Criteria for predicting hydroplaning potential. *Transportation Research Record*, *1100*, 1–7.

Hranac, R., Sterzin, E., Krechmer, D., Rakha, H., & Farzaneh, M. (2006). *Empirical studies on traffic flow in inclement weather* (Report No. FHWA-HOP-07-073). Federal Highway Administration, U.S. Department of Transportation.

Hussein, N., Hassan, R., & Fahey, M. T. (2021). Effect of pavement condition and geometrics at signalized intersections on casualty crashes. *Journal of Safety Research*, *76*, 276–288. https://doi.org/10.1016/J.JSR.2020.12.021

Jayasooriya, W., & Gunaratne, M. (2014). Evaluation of widely used hydroplaning risk prediction methods using Florida's past crash data. *Transportation Research Record*, *2457*(1), 140–150. https://doi.org/10.3141/2457-15

Kassu, A., & Anderson, M. (2019). Analysis of severe and non-severe traffic crashes on wet and dry highways. *Transportation Research Interdisciplinary Perspectives*, *2*. https://doi.org/10.1016/J.TRIP.2019.100043

Kutela, B., Combs, T., John Mwekh'iga, R., & Langa, N. (2022). Insights into the long-term effects of COVID-19 responses on transportation facilities. *Transportation Research Part D: Transport and Environment*, *111*. https://doi.org/10.1016/J.TRD.2022.103463

Lee, H. S., & Ayyala, D. (2020). *Enhanced hydroplaning prediction tool*. Applied Research Associates and Florida Department of Transportation.

Lee, J., Nam, B., & Abdel-Aty, M. (2015). Effects of pavement surface conditions on traffic crash severity. *Journal of Transportation Engineering*, *141*(10).

Li, Z., Ci, Y., Chen, C., Zhang, G., Wu, Q., Qian, Z., Prevedouros, P. D., & Ma, D. T. (2019). Investigation of driver injury severities in rural single-vehicle crashes under rain conditions using mixed logit and latent class models. *Accident Analysis & Prevention*, *124*, 219–229. https://doi.org/10.1016/J.AAP.2018.12.020

Li, Y., & Huang, J. (2014). Safety impact of pavement conditions. *Transportation Research Record*, *2455*(1), 77–88.

Li, Y., Liu, C., & Ding, L. (2013). Impact of pavement conditions on crash severity. *Accident Analysis & Prevention, 59*, 399–406. https://doi.org/10.1016/j.aap.2013.06.028

Li, J. Q., Liu, W., Yang, X., Lu, P., & Wang, K. C. (2022). Statistical safety performance models considering pavement and roadway characteristics. *Journal of Advanced Transportation*. https://doi.org/10.1155/2022/5871601

Luo, W., Wang, K. C. P., & Li, L. (2019). Field test validation of water film depth (WFD) prediction models for pavement surface drainage. *International Journal of Pavement Engineering, 20*(10), 1170–1181.

Ma, Z., Shao, C., Yue, H., & Ma, S. (2009). Analysis of the logistic model for accident severity on urban road environment. *Proceedings of the IEEE Intelligent Vehicles Symposium*, 983–987. https://doi.org/10.1109/IVS.2009.5164414

Mayora, J. M. P. & Piña, R. J. (2009). An assessment of the skid resistance effect on traffic safety under wet-pavement conditions. *Accident Analysis & Prevention*, *41*(4), 881–886. https://doi.org/10.1016/j.aap.2009.05.004

Merritt, D. K., Lyon, C. A., & Persaud, B. N. (2015). *Evaluation of pavement safety performance* (Report No. FHWA-HRT-14-065). Federal Highway Administration, U.S. Department of Transportation.

Mumtarin, M., Chowdhury, M. S., & Wood, J. S. (2023). *Large language models in analyzing crash narratives: A comparative study of ChatGPT, Bard, and GPT-4.* [Preprint]. ArXiv. https://doi.org/10.48550/arXiv.2308.13563

National Academies of Sciences, Engineering, and Medicine. (2021). *Guidance to predict and mitigate dynamic hydroplaning on roadways*. The National Academies Press. https://doi.org/10.17226/26287

National Oceanic and Atmospheric Administration (NOAA). (2023). *Precipitation frequency data server*. https://hdsc.nws.noaa.gov/pfds/

Norton, E. C., & Dowd, B. E. (2018). Log odds and the interpretation of logit models. *Health Services Research*, *53*(2), 859–878. https://doi.org/10.1111/1475-6773.12712

Nowakowska, M. (2010). Logistic models in crash severity classification based on road characteristics. *Transportation Research Record,* 2148, 16–26. https://doi.org/10.3141/2148-03

Pourhassan, A., Gheni, A. A., & El Gawady, M. A. (2022). Water film depth prediction model for highly textured pavement surface drainage. *Transportation Research Record,* 2676(2), 100–117.

Roe, P. G., Webster, D. C., & West, G. (1998). *The relationship between the surface texture of roads and accidents* (Report No. 367). Transport Research Laboratory.

Ross, N. F., & Russam, K. (1968). The depth of rain water on road surfaces (Report No. LR 236). Road Research Laboratory, Ministry of Transport.

Roy, U., Farid, A. & Ksaibati, K. (2023). Effects of pavement friction and geometry on traffic crash frequencies: A case study in Wyoming. *International Journal of Pavement Research and Technology*, *16*(6), 1468–1481. https://doi.org/10.1007/s42947-022-00208-4

Start, M. R., Kim, J., & Berg, W. D. (1998). Potential safety cost-effectiveness of treating rutted pavements. *Transportation Research Record*, *1629*(1), 208–213. https://doi.org/10.3141/1629-23

Tehrani, S., Cowe Falls, L. & Mesher, D. (2017). Effects of pavement condition on roadway safety in the province of Alberta. *Journal of Transportation Safety & Security*, *9*(3), 259–272. https://doi.org/10.1080/19439962.2016.1194352

Texas Department of Transportation (TxDOT). (2019). *Hydraulic design manual*.

Texas Department of Transportation (TxDOT). (2024). *Roadway design manual*.

Vinayakamurthy, M., Mamlouk, M., Underwood, S., & Kaloush, K. (2017). *Effect of pavement condition on accident rate*. National Transportation Center, University of Maryland.

Wang, K. C. P. & Li Q. (2017). *Safety analysis opportunities using pavement surface characterization based on 3D laser imaging* (Report No. FHWA-SA-17-046). Federal Highway Administration, U.S. Department of Transportation.

# APPENDIX

## TOOL DEVELOPMENT TECHNOLOGIES AND DATA

This appendix documents the web-based GIS tool for supporting pavement safety diagnostics and identification of candidate pavement safety improvement projects. The tool is intended to bring together the key outputs of the project into one interactive environment so that TxDOT users can review crash susceptibility, pavement condition, crash history, and narrative-derived safety evidence at the segment level.

This appendix focuses on two topics: how the tool was developed from a web technology standpoint, and what datasets are used by the application.

### Web-Based Tool Architecture

The current prototype follows a lightweight web GIS architecture. The application is served through a password-protected web entry page, after which the user enters a single-page map interface. Most of the analytical display logic occurs on the client side, while preprocessed project data are loaded as GIS-ready files.

The architecture can be summarized as follows:

1. A PHP-based access layer controls entry to the map.
2. An HTML/CSS/JavaScript front end renders the user interface.
3. JavaScript map logic loads and filters GIS-ready files for display.
4. Pavement segment, boundary, narrative, and crash datasets are stored as GeoJSON or CSV.
5. Browser dynamically renders layers, filters, popups, and map symbology for the selected district and analysis mode.

This architecture provides web deployment, browser-based access, and compatibility with a district-level screening workflow.

### Web Technologies Used in the Tool

Table 15 lists the principal web technologies used in the current implementation and explains their roles in the application.

**Table 15. Primary Web Technologies Used in the Current GIS Tool Implementation.**

| Technology | Type | Role in the Tool |
|---|---|---|
| PHP | Server-side scripting | Controls password-protected access to the tool and manages the login-to-map page transition. |
| HTML | Front-end structure | Defines the layout of the login page, sidebar, map container, popup containers, and modal windows. |
| CSS | Front-end styling | Controls the visual layout of the application, including the sidebar, buttons, filters, map loader, popup appearance, and overall dashboard styling. |
| JavaScript | Client-side logic | Handles district switching, analysis selection, filter interaction, data loading, layer rendering, popup generation, and state persistence. |
| Leaflet | JavaScript mapping library | Provides the interactive GIS map, layer management, popups, legends, district boundaries, and base-map rendering. |
| noUiSlider | JavaScript UI library | Implements the interactive slider controls used for crash probability, pavement filters, and narrative thresholds. |
| Papa Parse | JavaScript parsing library | Loads and parses the crash CSV file used for crash-point display and crash-level popups. |
| OpenStreetMap tiles | Base map service | Supplies the background reference map used for roadway and geographic context. |
| Google Street View link integration | External web service linkage | Allows the user to open a Street View location from crash or pavement segment popup windows for visual review. |
| GeoJSON and CSV | Data exchange formats | Store the GIS-ready pavement, district boundary, narrative, and crash datasets consumed by the web application. |

## Project Data Used in the Tool

The current tool uses four primary project data groups:

- District pavement segment layers.
- Crash dataset for crash locations and crash-level details.
- District boundary layer.
- Narrative-derived pavement segment layers.

These data groups support the three analysis modes of the tool: Crash Probability Estimation, Crash Narrative, and Pavement Segment Explorer. The following subsections describe each dataset in detail.

### District Pavement Segment Layers

The district pavement segment layers are the core GIS layers used by the tool. Each district is stored as a separate GeoJSON file, and each feature represents a pavement segment with associated pavement, crash, traffic, geometric, and model-output attributes. These layers are used in both the Crash Probability Estimation and Pavement Segment Explorer modes.

Table 16 presents the district pavement segment dataset at both the file level and the attribute-group level. The table shows the file structure, geometry type, primary use of the layer in the tool, and the main groups of pavement, crash, traffic, geometric, environmental, and model-output fields carried by each segment.

**Table 16. District Pavement Segment Dataset Structure and Key Attribute Groups.**

| Component | Representative Fields/Structure | Description |
|---|---|---|
| Current file structure | One GeoJSON file per district, for example pavement/19—ATLANTA.geojson | District-based file organization used for map loading and district filtering. |
| Geometry type | Line-based pavement segment features | Each feature represents one pavement section/segment. |
| Primary role in tool | Crash Probability Estimation and Pavement Segment Explorer | Main network layer used for crash-probability display and condition-based exploration. |
| Segment identification | SEGMENT_KEY, RESPONSIBLE DISTRICT | Unique segment reference and district assignment used throughout the tool. |
| Crash history | TOTAL_CRASHES, CRASH_RATE | Number of crashes associated with the segment and derived crash-rate measure used for screening. |
| Condition scoring | CONDITION_SCORE, DISTRESS_SCORE | Composite pavement condition indicators available for segment review and filtering. |
| Surface friction and texture | SKID_NUMBER, MEAN_MPD (MM) | Friction and macrotexture measures relevant to wet-weather and pavement safety screening. |
| Smoothness and rutting | MAP21 IRI AVG (IN/MI), MAP21 RUTTING AVG (IN) | Roughness and rutting measures used in the explorer and interpretation of pavement condition. |
| Categorical condition fields | MAP21 CONDITION CATEGORY, MAP21 IRI CONDITION CATEGORY | Preclassified condition categories used in filtering and map display. |
| Pavement type | BROAD_PAVEMENT_TYPE, MODIFIED_BROAD_PAVEMENT_TYPE, DETAILED PVMNT TYPE ROAD LIFE | Pavement family and detailed pavement classification used in filtering and popup interpretation. |
| Traffic exposure | ADT_CURRENT, TRUCK_ADT_PERCENTAGE | Exposure-related traffic information used in both model interpretation and map exploration. |
| Geometry | SEGMENT_LENGTH (MI), LONGITUDINAL_GRADE (%), DEFLECTION_ANGLE (DEGREES) | Roadway geometry variables relevant to crash susceptibility and hydroplaning-related screening. |
| Environmental context | ANN_PRCP_NORMAL (IN) | Climate context used in the project modeling framework. |

| Component | Representative Fields/Structure | Description |
|---|---|---|
| Model outputs | CRASH_PROB_NEW_PERCENT, OVERALL_NEW_20 | Segment-level crash probability estimate and comparison context used in popup interpretation. |

**Crash Dataset**

The crash dataset is stored as a CSV file (integrated_crash.csv). Each row represents a crash record associated with a segment and includes crash location, timing, weather and surface conditions, severity, pavement attributes, roadway attributes, and the original investigator narrative. This dataset supports the crash-point overlay and crash-level popup windows.

Table 17 presents the crash dataset at both the file level and the information-group level. The table shows the file structure, the record unit used in the crash file, the role of the dataset in the web tool, and the main groups of spatial, temporal, severity, roadway, pavement, weather, and narrative-related fields contained in the crash records.

**Table 17. Crash Dataset Structure and Key Information Groups.**

| Component | Representative Fields/Structure | Description |
|---|---|---|
| Current file structure | Single CSV file: integrated_crash.csv | One crash record per row. |
| Primary role in tool | Crash location overlay and crash-level popups | Crash-point layer and crash-level evidence source used in the map interface. |
| Spatial reference within tool | Latitude, Longitude | Coordinates used to display crash points. |
| Linkage field | Segment_Key | Links crashes to pavement segments. |
| Time coverage | FY2021-FY2023 | Analysis period represented in the current crash file. |
| Narrative support | Investigator_Narrative and theme labels | Supports crash-level narrative display and narrative-based screening. |
| Crash identifiers | Crash_ID, Segment_Key, Fiscal_Year | Unique crash identity, linked segment, and analysis period. |
| Spatial location | Latitude, Longitude, COUNTY, RESPONSIBLE DISTRICT | Crash point location and administrative context used in the map and popup. |
| Time and temporal context | Crash_Date, Crash_Time, Day_of_Week, Day, Night | Date and time information used for popup display and temporal interpretation. |
| Weather and surface condition | Wthr_Cond_ID, Surf_Cond_ID, Wet, Dry, Precipitation (in/h) | Weather and pavement-surface condition context associated with the crash. |
| Weather-station linkage | Wthr_Station_ID, Nearest_Station_Dist (mi) | Reference to the weather station and station distance used in precipitation integration. |
| Crash severity | Crash_Sev_ID, K, A, B, C, O | Severity coding and severity distribution fields. |
| Roadway condition and geometry | Road_Algn_ID, SPD_LIMIT (MI/H), ADT_CURRENT, NUM_LANES, LANE_WIDTH (FT), | Roadway and operational characteristics associated with the crash record. |

| Component | Representative Fields/Structure | Description |
|---|---|---|
| | LONGITUDINAL_GRADE (MI/MI), DEFLECTION_ANGLE (DEGREES), CROSS_SLOPE (%) | |
| Pavement condition and distress | SKID_NUMBER, DISTRESS_SCORE, CONDITION_SCORE, MAP21 CONDITION CATEGORY, MAP21 IRI AVG (IN/MI), MEAN_MPD (MM), MAP21 RUTTING AVG (IN), POTHOLE QTY | Pavement-related fields used to interpret whether crash patterns may be related to pavement condition. |
| Pavement type | BROAD_PAVEMENT_TYPE, MODIFIED_BROAD_PAVEMENT_TYPE, DETAILED PVMNT TYPE ROAD LIFE | Surface and structural pavement type information associated with the crash-linked segment. |
| Narrative-derived labels | Potholes, Rough Pavement, Hydroplaning, Wet Pavement, Loss of Control at Curves, Curve Mentioned, Other | Crash theme labels used to support the narrative-analysis component of the tool. |
| Narrative text | Investigator_Narrative | Original narrative text used for crash-level popup display and narrative-analysis interpretation. |

**District Boundary Dataset**

The district boundary layer provides district polygons for visual context. This layer supports district selection by showing the selected TxDOT district boundary on the map.

Table 18 summarizes the district boundary dataset. The table shows the file structure, geometry type, record unit, role of the boundary layer in the tool, and the key district identification fields used for map display.

**Table 18. TxDOT District Boundary Dataset Used in the GIS Tool.**

| Item | Description |
|---|---|
| Current file structure | Single GeoJSON file: txdot_district_boundaries.geojson. |
| Geometry type | Polygon district boundaries. |
| Primary record unit | One TxDOT district polygon. |
| Primary role in tool | Displays the selected district boundary as a spatial frame of reference. |
| Key fields | DIST_NBR, DIST_NM, TXDOT_DIST_NBR, TXDOT_DIST_NM, Label. |

**Narrative-Derived Segment Layers**

The narrative-derived segment layers are stored as separate GeoJSON files for the four narrative themes currently implemented in the tool:

- Hydro_new.geojson.
- Rough_new.geojson.
- Loc_new.geojson.
- Potholes_new.geojson.

Each layer stores pavement segments associated with a narrative-derived crash theme and includes fields needed for popup interpretation. Table 19 presents the narrative-derived segment layers at both the file level and the attribute-group level. The table shows the file structure, geometry type, role of the layers in the Crash Narrative mode, and the main groups of theme-count, pavement-context, district-comparison, and crash-probability fields carried by these features.

**Table 19. Narrative-Derived Segment Layer Structure and Key Attribute Groups.**

| Component | Representative Fields/Structure | Description |
|---|---|---|
| Current file structure | Four GeoJSON files, one per narrative theme | Theme-specific files used in the crash narrative analysis mode. |
| Geometry type | Line-based pavement segment features | Each feature represents one pavement section/segment. |
| Primary role in tool | Hydroplaning, rough pavement, potholes, and loss of control at curves | Displays narrative-derived segment themes in the map interface. |
| Segment identity | SEGMENT_KEY, RESPONSIBLE DISTRICT | Segment reference and district assignment. |
| Narrative theme intensity | Hydroplaning_new, RoughPavement_new, LossControl_new, Potholes_new | Count or theme-specific indicator used by the narrative threshold controls. |
| Pavement safety context | SKID_NUMBER, MEAN_MPD (MM) | Key pavement surface indicators used in the narrative popup. |
| District comparison context | DISTRICT_AVG_SKID, DISTRICT_AVG_MPD | District-average comparison values used to interpret the selected segment. |
| Crash probability context | CRASH_PROB_NEW_PERCENT, OVERALL_NEW_20 | Segment crash-probability context used in the narrative popup. |

## TOOL FUNCTIONS

Section 2 documents the user-facing functions of the GIS tool. The current interface is organized around three analysis modes:

1. Crash Probability Estimation.
2. Crash Narrative.
3. Pavement Segment Explorer.

The discussion below follows the normal user workflow, starting at login and then moving through each analysis mode.

### Login and Access Control

The tool begins with a password-protected login page. Figure 34 shows the initial entry screen that the user sees before accessing the map. This access-control step is part of the current prototype deployment and is implemented through the PHP front end described earlier in Section 1.

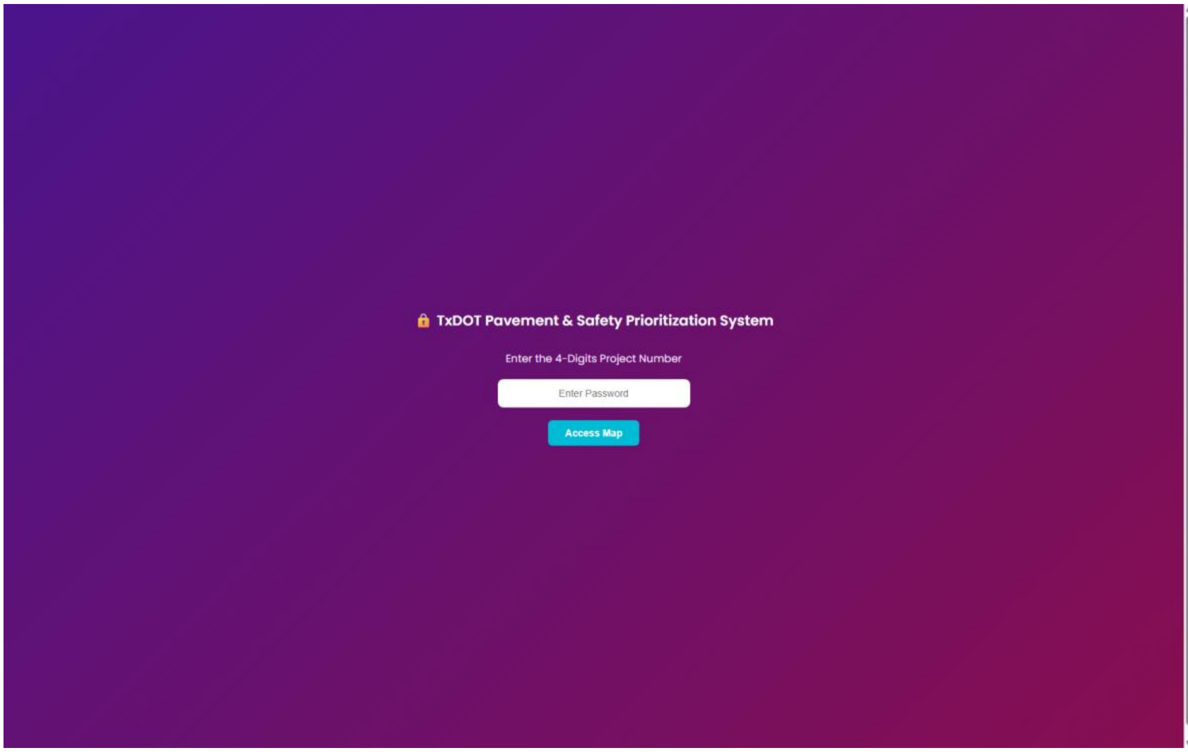



**Figure 34. Password-Protected Entry Page for the GIS Tool.**

After entering the project password, the user is redirected to the interactive map interface.

### Common Interface Functions

The common functions shared across the tool are listed below:

- TxDOT district selection, which limits the displayed network and associated map layers to the selected district.
- Analysis-type selection, which switches the map among Crash Probability Estimation, Crash Narrative, and Pavement Segment Explorer.
- Background map toggle, which turns the base map on or off for display.
- Crash location toggle, which turns the crash-point overlay on or off.
- District-specific display of pavement segments and district boundary, which renders the pavement segment layer and the selected TxDOT district boundary for the active district.

Figure 35 shows the overall map interface after login. The layout shown in the figure is shared across the three analysis modes: district selection at the top, analysis selection beneath it, filter and display settings in the main control panel, and the map view on the right.

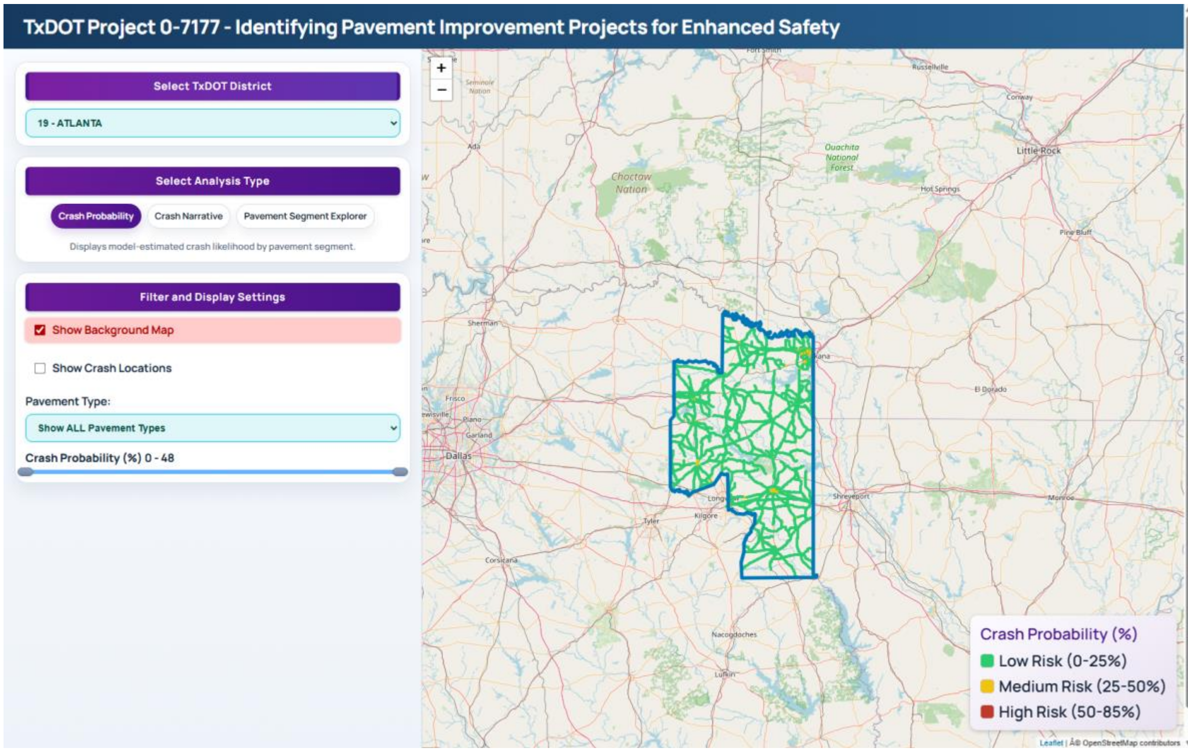



**Figure 35. Overall Map Interface Showing District Selection, Analysis Selection, Filter Controls, and District-Level Segment Display in the Crash Probability Estimation Mode.**

## Crash Probability Estimation

The Crash Probability Estimation mode is the model-based screening view of the tool. In this analysis mode, pavement segments are displayed according to their estimated probability of crash occurrence.

### *District-Level Crash Probability Display*

As shown in Figure 35, the first function of this mode is district-level visualization of the crash susceptibility outputs. Once a district is selected, the pavement sections in that district are displayed and colored by crash probability category. The display presents segments by relative crash-probability level within the selected district.

### *Crash Probability Filters*

Within this mode, the user can further refine the display using crash-probability-specific controls, including pavement type and crash probability range. These controls restrict the displayed network to selected pavement families and selected crash-probability ranges.

*Section-Level Popup Window*

The Crash Probability Estimation mode also includes a detailed section-level popup. Figure 36 shows an example popup. This popup summarizes the selected segment, including its segment key, segment length, total crashes, crash probability, pavement type information, and model input variables. The popup also includes a Street View link so the user can visually inspect the roadway context.

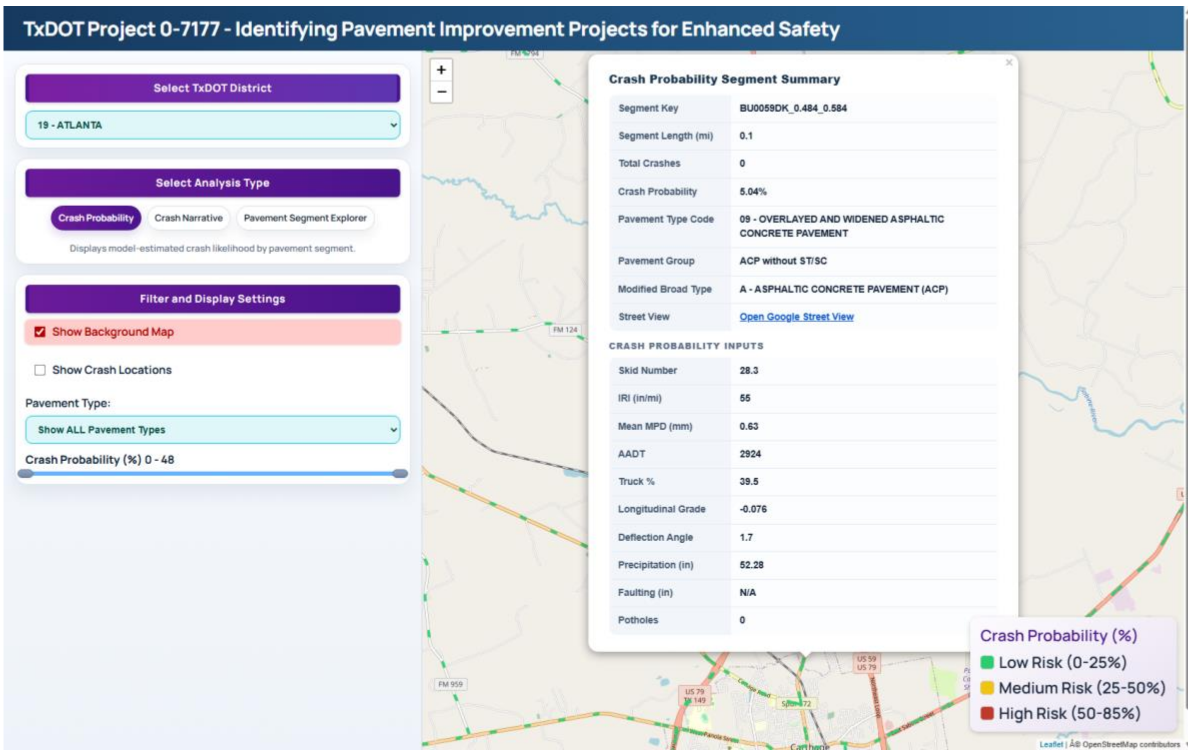



**Figure 36. Section-Level Popup in the Crash Probability Estimation Mode.**

The popup presents the segment attributes associated with the selected crash-probability result, including the displayed model output and supporting explanatory variables.

*Google Street View Review*

The popup-level Street View function is shown in Figure 37. When the user selects the Street View link from a crash or pavement-segment popup, the tool opens an in-application Street View modal. The modal presents image-based roadway context at the selected crash or segment location.

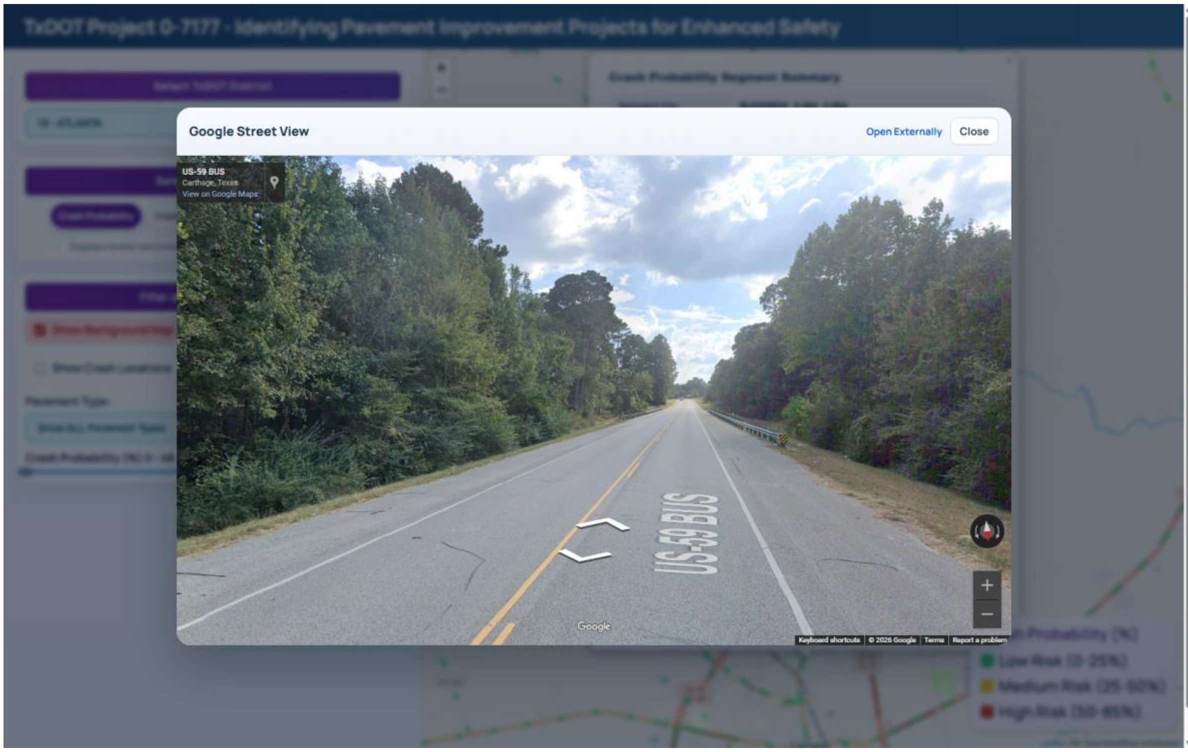



**Figure 37. In-Application Google Street View Window Launched from a Popup.**

Figure 37 shows the link between mapped crash and pavement records and image-based site context within the same interface.

*Crash Location Overlay*

Another function of this mode is the crash-location overlay. When the user enables Show Crash Locations, crash points are added to the map. Figure 38 shows this function. The overlay displays crash-point locations together with the crash-probability segment layer.

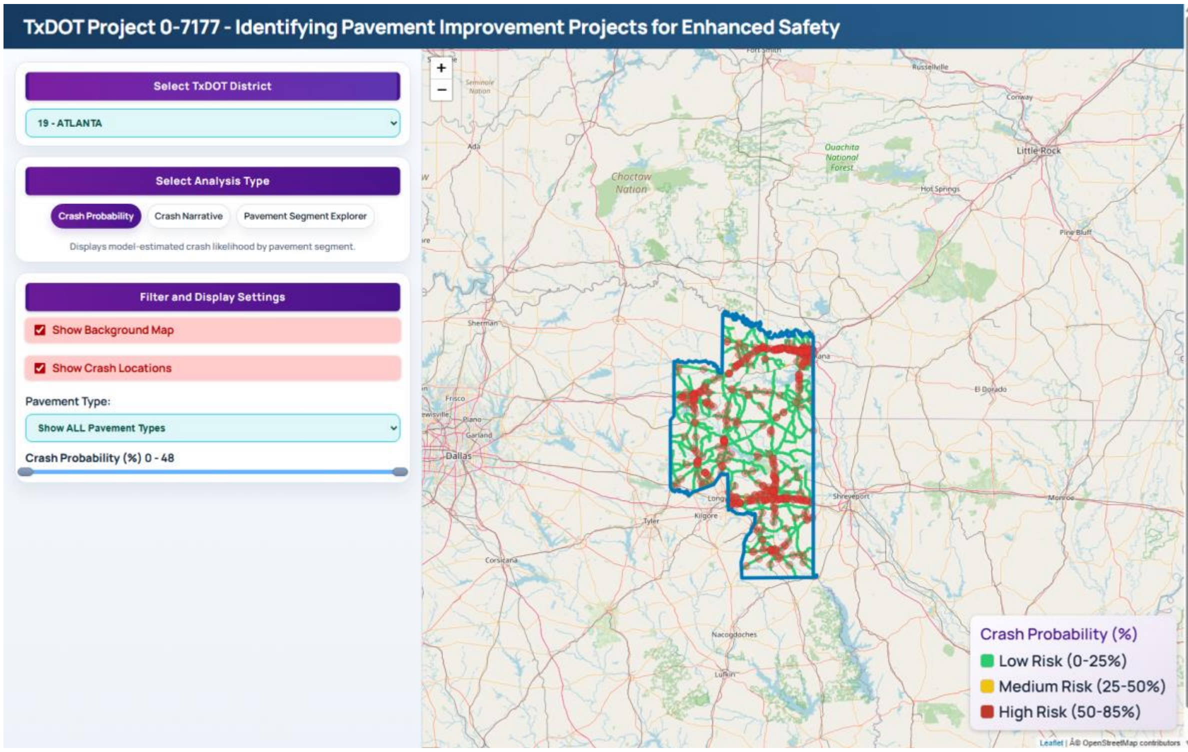



**Figure 38. Crash Location Overlay in the Crash Probability Estimation Mode.**

This overlay supports comparison between observed crash locations and the mapped crash-probability results.

### Crash Narrative

The Crash Narrative mode provides a second type of evidence layer. Instead of displaying only model outputs, it displays pavement sections associated with themes identified from investigator narratives.

*Narrative Theme Selection*

The current implementation includes four narrative themes:

- Hydroplaning.
- Rough Pavement.
- Loss of Control at Curves.
- Potholes

The user can switch among these themes and apply the threshold control associated with the selected theme. Figure 39 shows the hydroplaning view. This figure demonstrates how the interface displays a narrative-derived layer together with threshold filtering.

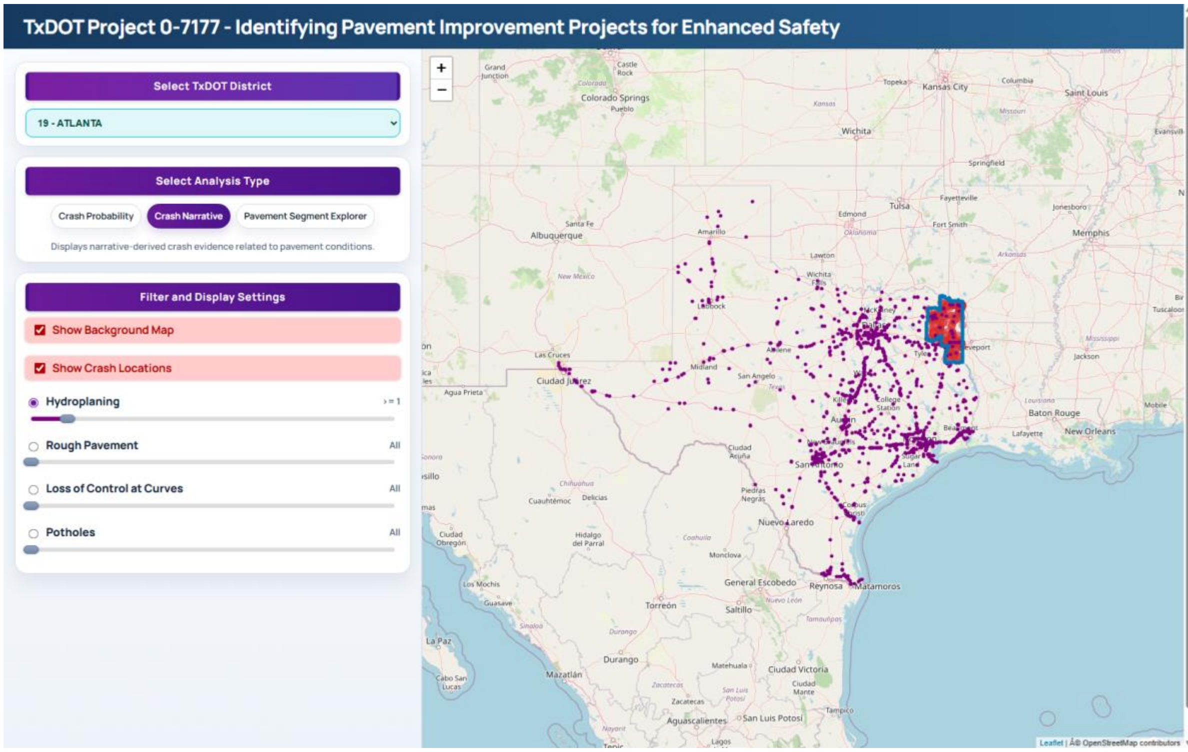



**Figure 39. Crash Narrative Mode Showing Hydroplaning-Related Pavement Sections and the Hydroplaning Threshold Control.**

This mode presents narrative-derived evidence as a separate segment layer alongside the other mapped project outputs.

*Narrative Popup Window*

The narrative mode also includes a detailed popup for the selected segment. Figure 40 shows an example of this popup. It presents segment-level narrative summary information, crash probability context, possible contributing factors, and a recommended safety-improvement direction.

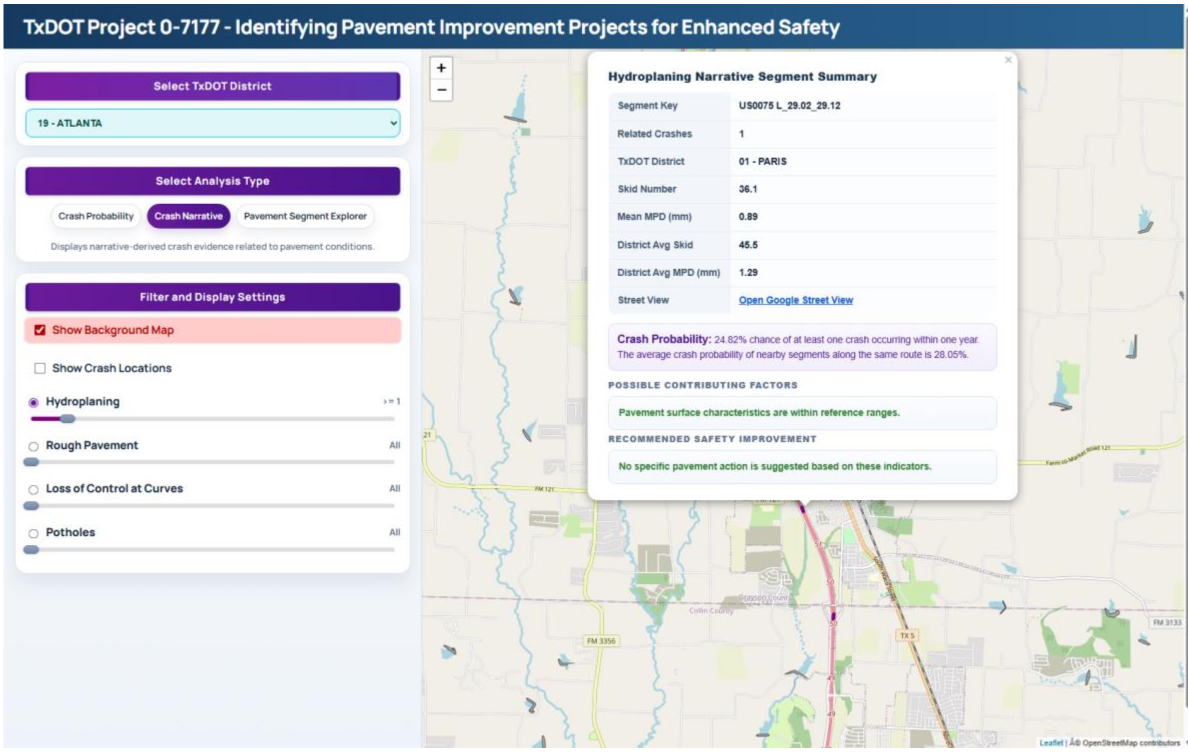



**Figure 40. Narrative Popup for a Hydroplaning-Related Segment in the Crash Narrative Mode.**

The popup contains a summary table and three interpretive sections. The summary table reports the selected segment, the number of related crashes for the active narrative theme, the TxDOT district, and the pavement-surface context for that segment.

The Crash Probability section reports the annual crash probability for the selected segment together with a comparison value for nearby segments along the same route. This gives the user a route-level comparison rather than showing only the segment-specific value.

The Possible Contributing Factors section presents screening-level statements based on the pavement indicators shown in the popup. In the current implementation, these statements describe whether the displayed friction- and texture-related indicators are notably lower than reference values or whether the pavement surface characteristics are within reference ranges.

The Recommended Safety Improvement section provides corresponding screening-level actions linked to those factor statements. These statements are intended to support diagnostic review of the selected segment and do not replace engineering evaluation.

### Pavement Segment Explorer

The Pavement Segment Explorer mode is the condition- and attribute-based exploratory mode of the tool. It displays pavement indicators and crash-history variables across the district network and supports filtering based on combinations of those variables.

*Color-By Indicator Display*

In this mode, the user may color the pavement network by multiple indicators, including pavement type, MAP-21 condition category, MAP-21 IRI condition category, total crashes, skid number, ADT, MPD, rutting, speed limit, and crash rate. Figure 41 shows the pavement network color-coded by MAP-21 condition category.

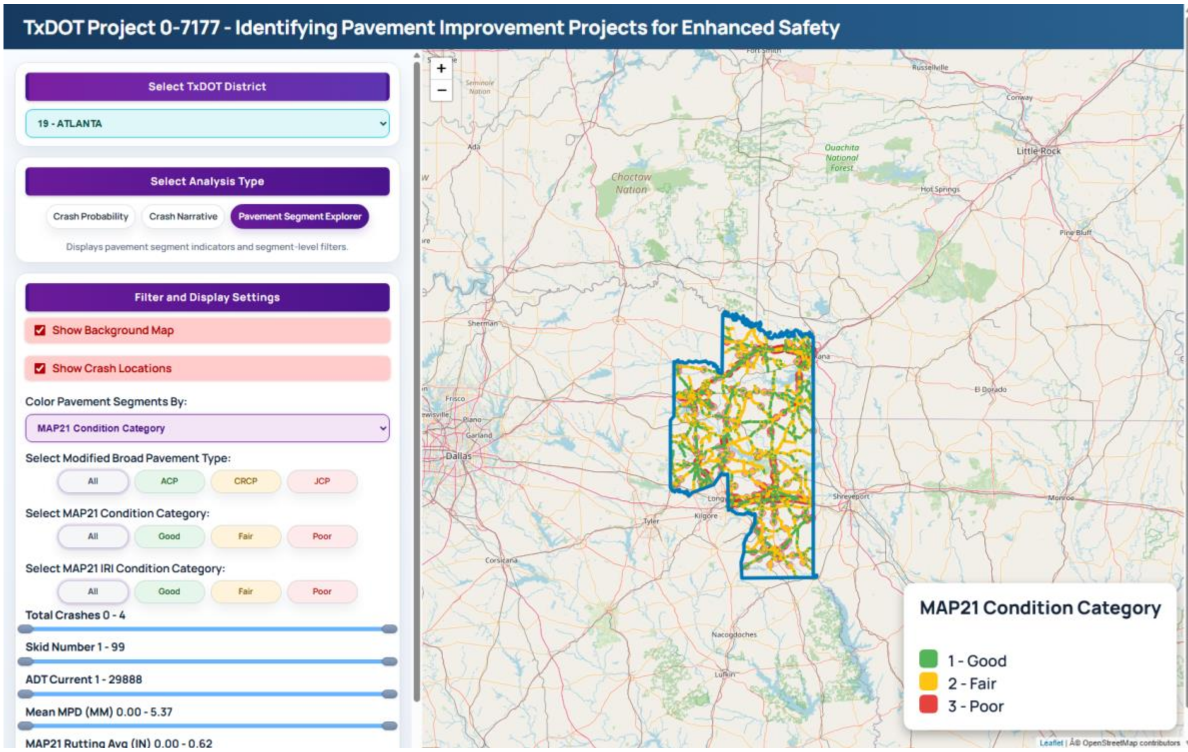



**Figure 41. Pavement Segment Explorer Mode with Segments Color-Coded by MAP-21 Condition Category.**

Figure 41 shows one example of indicator-based display in this mode using MAP-21 condition categories.

*Multi-Criteria Segment Filtering*

One core function of the Pavement Segment Explorer mode is its multi-criteria filtering capability. The user can apply condition-based and crash-history-based filters to narrow the displayed set of segments. Figure 42 shows one example in which the network is color-coded by total crashes and filtered to a specific crash-count range.

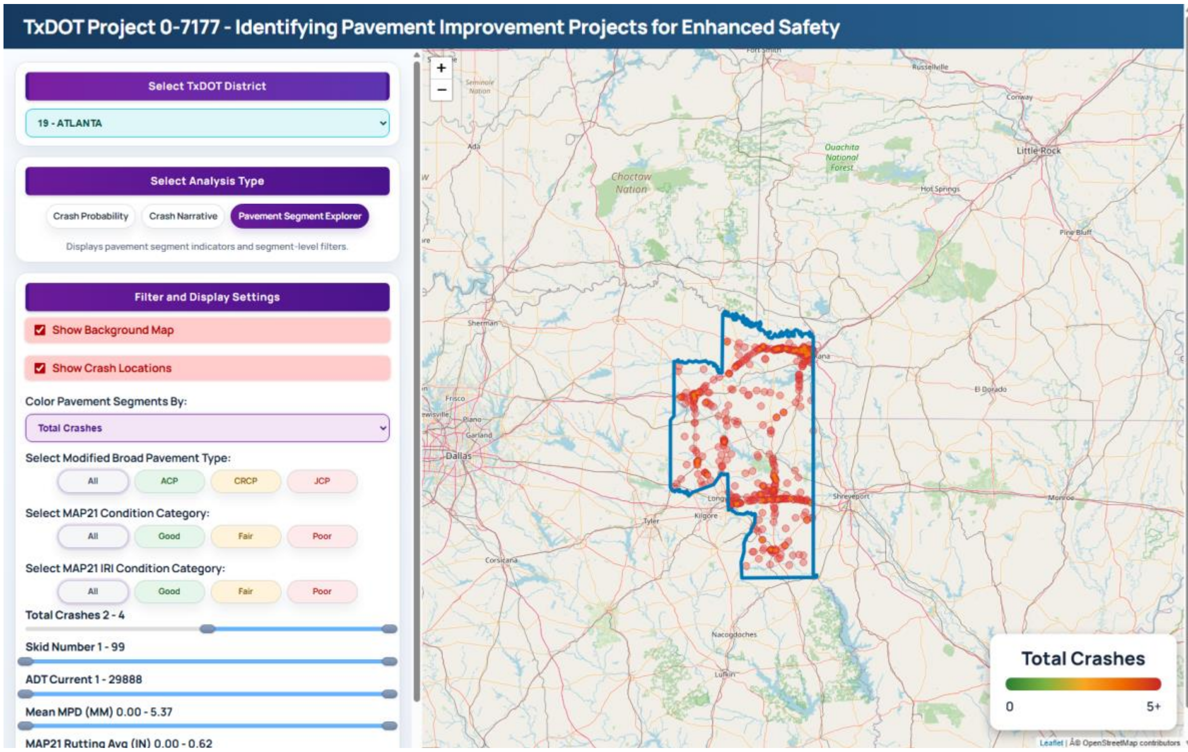



**Figure 42. Pavement Segment Explorer Mode with the Total Crashes Filter Applied.**

This filtering function displays only the segments that satisfy the selected condition, pavement-type, and crash-history criteria.

### Functional Summary

The three analysis modes serve different but complementary purposes:

- Crash Probability Estimation supports model-based screening of crash susceptibility.
- Crash Narrative supports evidence-based screening using narrative-derived crash themes.
- Pavement Segment Explorer supports direct exploration of pavement indicators and multi-criteria filtering.

This three-part structure organizes the tool around modeled risk, observed crashes, pavement condition, and narrative-derived evidence within one interface for candidate-section screening.

## FUTURE WORK

The current tool provides a screening framework that can be extended in two main directions: expansion toward treatment recommendation and future verification and calibration using newly available TxDOT data.

### Maintenance and Treatment Recommendation Logic

One future enhancement is expansion of the tool from a screening platform into a treatment-support platform. In future versions, the tool could be configured to associate combinations of pavement condition, crash history, crash probability, and narrative evidence with candidate maintenance or rehabilitation strategies. This would allow the tool to move beyond identification of candidate sections and provide preliminary guidance on potential treatment categories such as friction improvement, preservation treatment, localized repair, resurfacing, or rehabilitation.

This type of enhancement would keep the existing diagnostic role of the tool while adding a direct connection between the screening results and potential pavement treatment decisions. The treatment suggestions would remain subject to engineering review, field validation, and project-level judgment.

### Verification and Calibration with New TxDOT Data

A second future enhancement is use of newly available TxDOT data to verify and calibrate the current analytical models and screening logic. As additional years of crash data, pavement condition data, traffic data, and other relevant roadway data become available, these data can be used to evaluate whether the current relationships identified in the tool remain stable and whether the current thresholds and model outputs continue to perform as expected.

This future work would support model verification, recalibration of prediction relationships where needed, and periodic updating of the screening logic used in the tool. In this way, the web tool could be maintained as an evolving decision-support application rather than a fixed prototype tied only to the current project dataset.