CAPRI

CONSORTIUM FOR ASPHALT PAVEMENT RESEARCH AND IMPLEMENTATION

CAPRI Project

Field Evaluation of Long-Term Cracking Performance of High RAM Asphalt Mixtures

capriasphalt.us

Field Evaluation of Long-Term Cracking Performance of High RAM Asphalt Mixtures

By

Tyler Jones
Amy Epps Martin
Edith Arámbula Mercado
Texas A&M Transportation Institute
College Station, Texas

Elie Hajj
University of Nevada, Reno
Reno, Nevada

AUGUST 2026

ACKNOWLEDGEMENTS

The author would like to thank the Consortium for Asphalt Pavement Research and Implementation (CAPRI) for sponsoring this important research. Special thanks to TxDOT, WisDOT, DelDOT, Washoe County-Nevada, Rieth-Riley Construction Co.

Table of Contents

List of Figures

List of Tables

1.0 Introduction

The use of recycled asphalt materials (RAM), including reclaimed asphalt pavement (RAP) and recycled asphalt shingles (RAS), in newly constructed asphalt pavements is now ubiquitous with average RAP content increasing from 15.6% by weight of mixture in 2009 to 22.9% in 2024. This asphalt technology results in annual economic and environmental benefits with total savings of $4.7 billion in materials costs and 72.5 million cubic yards of landfill space [1]. Because of the significant financial and environmental benefits that increase with RAM content, it is imperative to understand how to engineer its use for balanced performance with respect to key pavement distresses, such as cracking, rutting, and moisture susceptibility.

The use of RAM as an aged component material inherently results in asphalt mixtures that are stiffer and more brittle than virgin mixtures without RAM. This causes the mixture to be more resistant to rutting but also less resistant to cracking. To increase RAM content in mixtures, strategies to mitigate cracking performance issues are usually needed. Recycling agents provide one mitigation strategy that can be used to improve the field performance of high RAM mixtures [2].

Two issues arise with the more widespread use of recycling agents in high RAM mixtures. First, there are limited studies documenting the long-term field performance of high RAM mixtures. Table 1 includes a list of individual department of transportation (DOT) studies (both completed and ongoing) that monitored field performance of a limited number of high RAM asphalt mixtures (3-12) after different in-situ aging periods [3, 4, 5, 6, 7, 8, 9, 10].

TABLE 1 DOT FIELD MONITORING STUDIES OF HIGH RAM MIXTURES

State	Age
OK (Ongoing)	0.5
OR	1
ND (Ongoing)	2
NE	2-3
CA	2.5-3
WI	3-4
IL	3-5
VA (Ongoing)	11-12 & 4-6

The second issue is the lack of a standardized mix design approach for high RAM mixtures. AASHTO PP 127 *Development of Balanced and Durable Asphalt Mixtures with High Recycled Asphalt Materials Contents* provides guidance on component material selection and proportioning, binder blend evaluation, and mixture performance evaluation [11]. These procedures were developed through a series of national research and implementation efforts sponsored by the National Cooperative Highway Research Program (NCHRP) in Projects 09-58 [12], 20-44(24) [13], and 09-65 [14].

However, much of the original performance data, particularly from NCHRP 09-58, was based on early-life pavement observations. Under this Consortium for Asphalt Pavement Research and Implementation (CAPRI) project, the re-evaluation of these same field sections after 8-10 years in-service provides a unique opportunity to strengthen confidence in the existing criteria prescribed in AASHTO PP 127 and, whenever warranted, propose refinements.
In addition, this project offers insights into the long-term field performance of high RAM asphalt mixtures with recycling agents. This chapter summarizes the NCHRP Project 09-58 and AASHTO PP 127 that form the basis for this CAPRI project *Field Evaluation of Long-Term Cracking Performance of High RAM Asphalt Mixtures*.

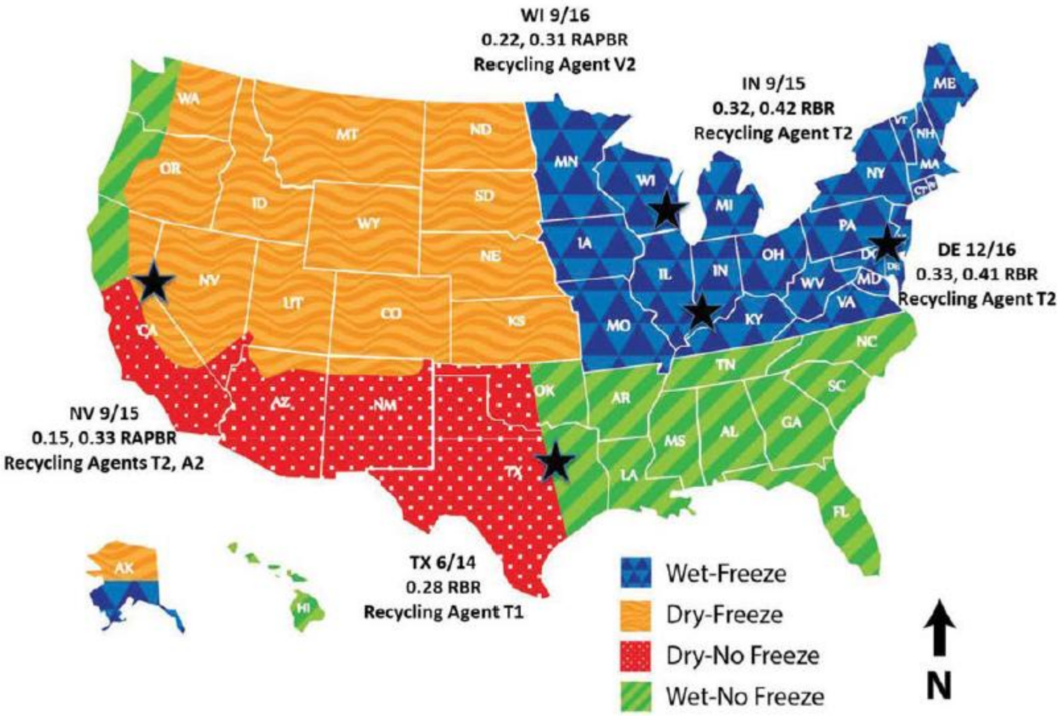
1.1 NCHRP Project 09-58

NCHRP Project 09-58 *The Effects of Recycling Agents on Asphalt Mixtures with High RAS and RAP Binder Ratios* was an almost five-year study (2014-2019) that explored the effects of recycling agents on asphalt mixtures that contained high RAM contents, including RAP and RAS. The study included five field projects located in Texas (TX), Nevada (NV), Indiana (IN), Wisconsin (WI), and Delaware (DE) as shown in Figure 1. Binder contents ranged from 4.5% to 5.9%, and recycled binder ratios (RBRs) ranged from 0 to 0.42 across 18 field sections with RBR defined in Equation 1.

$$RBR = \frac{(Pb_{RAP} * P_{RAP}) + (Pb_{RAS} * P_{RAS})}{100 * Pb_{total}} \quad (1)$$

where:
Pb_{RAP} = binder content of the RAP,
P_{RAP} = percentage of RAP by weight of mixture,
Pb_{RAS} = binder content of RAS,
P_{RAS} = percentage of RAS by weight of mixture, and
Pb_{total} = binder content of the combined mixture.

It is important to note that recycling agent dosages in these field sections were primarily based on supplier recommendations and were not explicitly optimized using performance-based criteria, as later described in AASHTO PP 127.



Notes: T = tall oil, A = aromatic extract, V = vegetable oil

FIGURE 1 MAP OF NCHRP 09-58 FIELD PROJECTS

Table 2 provides a summary of the field projects, including number of field sections, construction date, layer thickness, monitoring time periods, traffic parameters, and cumulative degree days (CDD) that normalizes the diverse climates. CDD at the current field performance monitoring session was defined as the number of degree days above freezing since construction based on average daily temperature at the nearest climate station.

$$CDD = \sum(MDD) \quad (2)$$

$$MDD = (T - 32) * D \quad (3)$$

Where:
MDD = Monthly Degree Days
T = Monthly Average Maximum Temperature
D = Number of days in the month

These average daily temperature data were accessed from the Weather Underground website [15]. Traffic was characterized by cumulative millions of Equivalent Single Axle Loads (mESALs) at the current field performance monitoring session and percent trucks. These traffic data were estimated as follows:

$$ESAL = AADT * T * DD * LD * TF * 365 \quad (4)$$

where:
AADT = Average Annual Daily Traffic = total number of vehicles per day
T = percentage of truck traffic
DD = directional distribution factor = proportion of traffic in design direction
LD = lane distribution factor = proportion of traffic in design lane
TF = average truck factor = ESALs for truck type = 1.7

Traffic data was not available for the NV and IN field projects; however, both sites correspond to low-volume local roads. The available traffic data was obtained from the respective state departments of transportation (DOT) for each field project.

TABLE 2 FIELD PROJECTS SUMMARY

Field Project	# Field Sections	Constr. (m/yr)	Layer/ Core Thick. (mm)	Monitoring (yrs)	Traffic (mESALs)	Truck (%)	CDD (k)
TX	3	06/14	50	0, 1, 2, 3, 10.5	3.43	14.8	178
NV	5	09/15	75	0, 1, 2, 9.8	N/A	N/A	133
IN	3	09/15	38	0, 1, 2, 9.6	N/A	N/A	100
WI	4	09/16	50	0, 1, 8.6	1.70	8.2	79
DE	3	12/16	50	0, 1, 8.3	0.015	1.0	102

N/A=data was not available at the time of this project

Figure 2 shows the CDD progression for each field project during the period of NCHRP 09-58, with field performance monitoring denoted by the black diamonds. In the cold, northern Freeze climates (in IN, NV, WI, and DE), CDD slows and plateaus during the winter when temperatures are near freezing; while the southern No Freeze climate in TX shows CDD continuing to accumulate at a more steady pace (i.e., without plateaus).

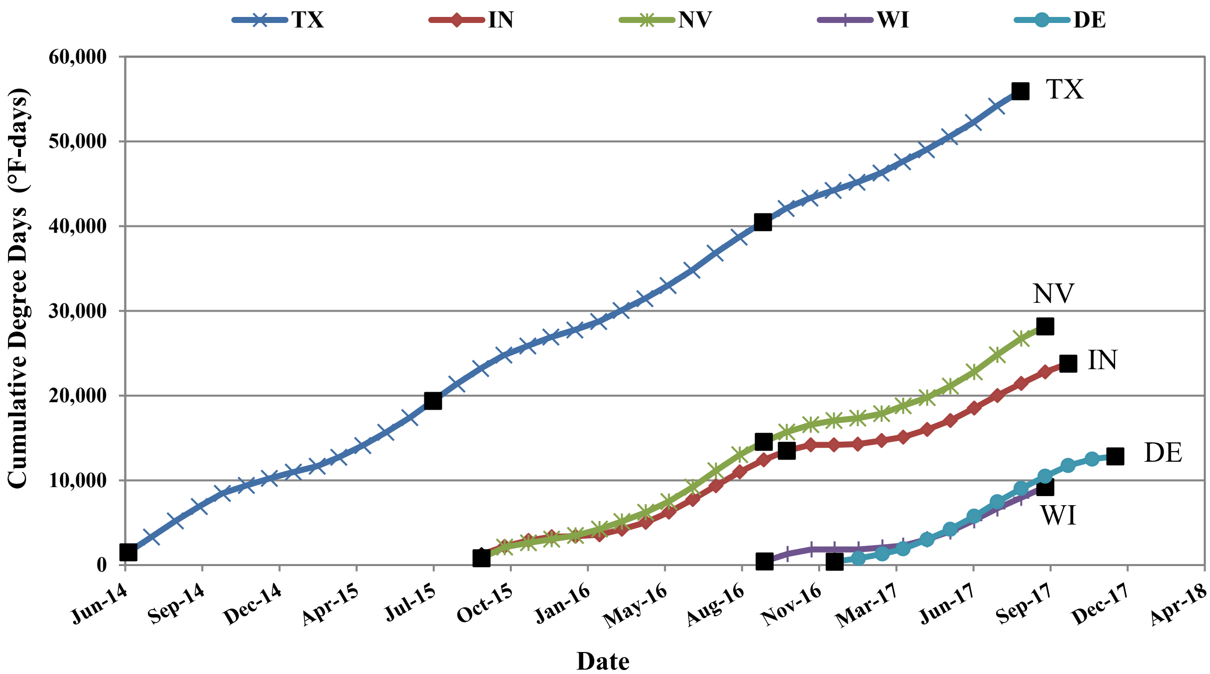


FIGURE 2 CDD DURING NCHRP 09-58

The field sections in Table 2 are presented by date of construction, but in this report the order of the results follows the extremes of field performance with the NV field project exhibiting the best performance and the IN field project showing the worst performance. Accordingly, the WI, TX, and DE field projects are also presented in order of general field performance from best to worst.

1.1.1 Nevada Field Project

The NV field project consisted of five field sections: a virgin section, a DOT control section at typical RBR, a recycled section, a rejuvenated section with a tall oil T2 recycling agent, and a rejuvenated section with an aromatic extract A2 recycling agent. Table 3 provides the mixture characteristics for these NV field sections that all contain the same polymer-modified asphalt (PMA) PG 64-28P and only RAP.

TABLE 3 NV ASPHALT MIXTURES

Field Section	RAM	RBR	Binder Content	PG Grade	Additive
Virgin	0	0	5.4%	64-28P	0
DOT Control	15% RAP	0.15 RAP	5.0%	64-28P	0
Recycled	33% RAP	0.33 RAP	4.6%	64-28P	0
Rejuvenated T2	33% RAP	0.33 RAP	4.5%	64-28P	2.0% T2
Rejuvenated A2	33% RAP	0.33 RAP	4.6%	64-28P	2.0% A2

1.1.2 Indiana Field Project

The IN field project consisted of three field sections: a virgin section, a DOT control section at typical RBR, and a rejuvenated section with a tall oil T2 recycling agent. As a cracking mitigation strategy for high RBR mixtures, the DOT control and rejuvenated sections incorporated a softer PG binder than that for the virgin section. Table 4 provides the mixture characteristics for these IN field sections with both the control and rejuvenated sections containing RAS.

TABLE 4 IN ASPHALT MIXTURES

Field Section	RAM	RBR	Binder Content	PG Grade	Additive
Virgin	0	0	5.9%	64-22	0
DOT Control (Soft)	28% RAP + 2% RAS	0.32 = (0.25 RAP + 0.07 RAS)	5.8%	58-28	0
Rejuvenated (Soft)	16% RAP + 8% RAS	0.42 = (0.14 RAP + 0.28 RAS)	5.8%	58-28	3.5% T2

1.1.3 Wisconsin Field Project

The WI field project consisted of four field sections: a DOT control section at typical RBR, a recycled section, a recycled section with a softer PG binder, and a rejuvenated section with a vegetable oil V2 recycling agent. Table 5 provides the mixture characteristics for these WI field sections that all contain only RAP.

TABLE 5 WI ASPHALT MIXTURES

Field Section	RAM	RBR	Binder Content	PG Grade	Additive
DOT Control	27% RAP	0.22 RAP	5.6%	58-28	0
Recycled	36% RAP	0.31 RAP	5.4%	58-28	0
Recycled (Soft)	36% RAP	0.31 RAP	5.4%	52-34	0
Rejuvenated	36% RAP	0.31 RAP	5.4%	58-28	1.2% V2

1.1.4 Texas Field Project

The TX field project consisted of three field sections: a virgin section, a DOT control section at typical RBR, and a rejuvenated section with a tall oil T1 recycling agent. As a cracking mitigation strategy for high RBR mixtures, the DOT control and rejuvenated sections incorporated a softer PG binder than that for the virgin section. Table 6 provides the mixture characteristics for these TX field sections with both the control and rejuvenated sections containing RAS. The DOT control mixture also included a dose of warm mix asphalt (WMA) additive.

TABLE 6 TX ASPHALT MIXTURES

Field Section	RAM	RBR	Binder Content	PG Grade	Additive
Virgin	0	0	4.9%	70-22P	0
DOT Control (Soft) WMA	10% RAP + 5% RAS	0.28 = (0.1 RAP + 0.18 RAS)	4.9%	64-22	0.5% WMA
Rejuvenated (Soft)	10% RAP + 5% RAS	0.28 = (0.1 RAP + 0.18 RAS)	4.9%	64-22	2.7% T1

1.1.5 Delaware Field Project

The DE field project consists of three sections: a DOT control section at typical RBR with WMA additive, a rejuvenated section with a tall oil T2 recycling agent, and a rejuvenated section with T2 recycling agent and WMA additive. Table 7 provides the mixture characteristics for these DE field sections that all contain the same PG 64-28 unmodified binder and RAS. The dose of the recycling agent is also low in both rejuvenated sections.

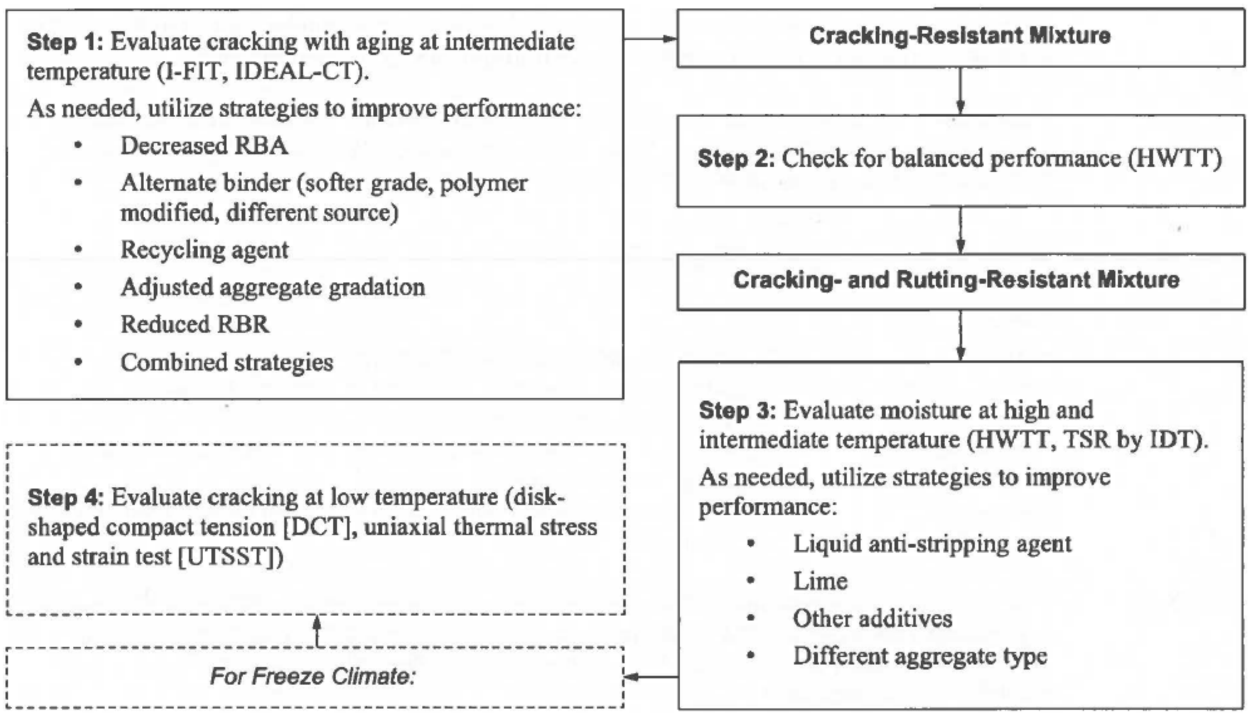
TABLE 7 DE ASPHALT MIXTURES

Field Section	RAM	RBR	Binder Content	PG Grade	Additive
DOT Control WMA	20% RAP + 4% RAS	0.34 = (0.17 RAP + 0.17 RAS)	5.4%	64-28	0.4% WMA
Rejuvenated	29% RAP + 4% RAS	0.41 = (0.24 RAP + 0.17 RAS)	5.4%	64-28	0.8% T2
Rejuvenated WMA	29% RAP + 4% RAS	0.41 = (0.24 RAP + 0.17 RAS)	5.4%	64-28	0.8% T2 + 0.25% WMA

1.2 AASHTO PP 127

AASHTO PP 127 was drafted based on the suite of tools developed in NCHRP 09-58. NCHRP Project 20-44(24) *Pilot Test of Proposed Standard Practice for Recycling Agents in Asphalt Mixtures Incorporating RAP and RAS* was a two-year implementation study (2019-2021) that applied these tools to two additional field projects in California (CA) and DE with eight total field sections. This follow-on project produced a revised draft AASHTO Standard Practice that condensed the recycled binder availability (RBA) cracking mitigation strategy, added a note suggesting correction for air voids and thickness are needed when evaluating cores, revised the recycling agent dose selection method, and suggested that only one cracking mitigation strategy be used in order to avoid rutting and moisture failures.

The final development of AASHTO PP 127 was completed in NCHRP Project 09-65 *Capturing Durability of High Recycled Binder Ratio (RBR) Asphalt Mixtures* which was a three-year study (2021-2024) that expanded the mixture performance evaluation tools and thresholds to include moisture susceptibility and aging in two climate regions (Freeze and No Freeze). A stepwise approach for designing robust, durable high RBR mixtures was also added as shown in Figure 3 [11]. Table 8 provides the complete set of mixture evaluation tools contained in AASHTO PP 127.



Notes: I-FIT = Illinois Flexibility Index Test; IDEAL-CT = Indirect Tensile Cracking Test; HWTT = Hamburg Wheel-Track Test; TSR = Tensile Strength Ratio; DCT = Disk-Shaped Compact Tension Test; UTSST = Uniaxial Thermal Stress and Strain Test.

FIGURE 3 STEPWISE APPROACH FOR DESIGN OF DURABLE HIGH-RBR MIXTURES FROM AASHTO PP 127

TABLE 8 AASHTO PP 127 MIXTURE EVALUATION TOOLS

Test	Parameter	Suggested Performance Threshold: Freeze Climate	Suggested Performance Threshold: No Freeze Climate	Potential Durability Issue
High-Temperature, Short-Term Aging				
T 324	$N_{12,5}$	≥10,000 for PG 58-XX at 45˚C	≥10,000 for PG 58-XX at 50˚C	Rutting
		≥15,000 for PG 64-XX at 45˚C	≥15,000 for PG 64-XX at 50˚C	
		≥20,000 for PG 70-XX at 45˚C	≥20,000 for PG 70-XX at 50˚C	
T 324	SIP	≥9,000 at 45˚C	≥9,000 at 50˚C	Moisture
	SN	≥2,000 at 45˚C	≥2,000 at 50˚C	Moisture
Intermediate-Temperature, Track with Aging				
T 324, TP 132	$G\text{-}R_m$	≤8,000 MPa after STOA		Cracking
		≤19,000 MPa after LTOA-compacted		
Intermediate-Temperature, Short-Term Aging				
T 393	FI	≥7	≥7	Cracking
T 283	TSR	≥80% with freeze-thaw	≥80% without freeze-thaw	Moisture
Intermediate-Temperature, Short- plus Long-Term Aging				
ASTM D8225	CT_{Index}	≥30 after LTOA	≥20 after LTOA	Cracking
Low-Temperature, Short- and/or Long-Term Aging				
T 419	S_m and m-$value_m$	≤Romero(2016) threshold on m-$value_m$ after STOA	-	Cracking
ASTM D8303	CRI_{ENV}	≥17 after LTOA-compacted	-	Cracking
ASTM D7313	G_f	≥400 after STOA	-	Cracking

Notes:–denotes not applicable; STOA = short-term oven aging (2 hours at 135˚C); LTOA = long-term oven aging of loose mix (Freeze: 6 hours at 135˚C, No Freeze: 8 hours at 135˚C); LTOA-compacted = long-term oven aging of compacted specimens (5 days at 85˚C).

1.2 Objective and Scope

The objective of this CAPRI project *Field Evaluation of Long-Term Cracking Performance of High RAM Mixtures* is to strengthen confidence in the binder and mixture performance thresholds found in AASHTO PP 127 and propose revisions as needed. Key tasks included monitoring the long-term (8-10 years) field performance of the 18 NCHRP 09-58 field sections with high RAM contents and comparing to laboratory evaluation results gathered from NCHRP Project 09-58 and measured on cores collected after 8-10 years in-service.

2.0 Experiment

This section describes the field activities, laboratory evaluation, and comparative analysis tools utilized in this project.

2.1 Field Activities

This section describes the distress evaluation procedure used for the current field performance monitoring session, the core procurement process, and the laboratory evaluation methods utilized for these cores.

2.1.1 Distress Evaluation

The Long-Term Pavement Performance (LTPP) distress identification manual was used as a guide for distress identification on each of the field sections [16]. Although rutting was recorded, only transverse and longitudinal cracking were used to assess the field performance based on the scope of the project and relevant later-life distress. Combined (transverse and longitudinal) percent lane cracking was determined from crack maps as follows, after discounting joint cracking:

$$\%\,Lane\ Cracking = \frac{(Cracking_{longitudinal} + Cracking_{transverse}) \times 1ft}{Lane\ Width \times Evaluation\ Area\ Length} \times 100 \tag{5}$$

MAP-21 thresholds of good, fair, and poor were initially adopted to categorize combined percent lane cracking [17]. Cracking less than 5% was defined as good. Cracking greater than 5% but less than 20% was defined as fair, and cracking greater than 20% was defined as poor. To further delineate field performance, two additional thresholds at 3% and 10% cracking were adopted. From 3% to 5% cracking was relabeled as adequate to capture sections with some cracking, and between 10% and 20% cracking was redefined as marginal, recognizing the originally proposed poor threshold of 10% in MAP-21. Table 9 presents these thresholds and associated color coding that is used throughout this report. The thresholds also include a designated letter used in the presentation of results subsequently.

TABLE 9 PERCENT LANE CRACKING THRESHOLDS

% Lane Cracking	Threshold
0	Good(G)
3	Adequate(A)
5	Fair(F)
10	Marginal(M)
20	Poor(P)

2.1.2 Core Procurement

Twelve cores were taken per field section. Cores were procured from the shoulder of the road where accessible to be consistent with those taken during NCHRP 09-58. If the shoulder was not wide enough for core procurement, cores were taken from between the wheel paths. Additional cores were taken over existing cracks and indicated only mixture performance-related top-down surface cracking and reflection cracking was present.

Cores were six inches in diameter and when possible, cores were drilled down to the bottom of the target asphalt pavement layer. In some cases, full depth cores were taken because pavement layers could not be separated. Cores were then marked and sent back to the laboratory. Cores were stored in a low temperature chamber to limit aging.

Once the cores were brought to the laboratory, they were measured for thickness, diameter, and air voids per ASTM D3203. Air voids statistics for each section are provided in Table 10. After measurements were completed, cores were divided to be tested for intermediate- and low-temperature cracking resistance. Cores in the same field section with similar air voids were grouped together and evaluated by the same test.

TABLE 10 AIR VOIDS OF CORES

Field Section	Average AV (%)	AV COV (%)
NV Virgin	3.5	15.4
NV DOT Control	3.5	18.7
NV Recycled	3.5	17.6
NV Rejuvenated T2	3.5	36.2
NV Rejuvenated A2	3.1	11.8
IN Virgin	11.7	5.3
IN DOT Control (Soft)	8.8	8.7
IN Rejuvenated (Soft)	8.0	10.1
WI DOT Control	5.9	5.3
WI Recycled	8.2	4.5
WI Recycled (Soft)	7.2	4.7
WI Rejuvenated	6.0	8.3
TX Virgin	8.1	4.2
TX DOT Control (Soft) WMA	11.0	4.4
TX Rejuvenated (Soft)	11.9	4.3
DE DOT Control WMA	6.3	7.1
DE Rejuvenated	5.8	17.7
DE Rejuvenated WMA	6.2	7.9

2.2 Laboratory Evaluation

Laboratory evaluation focused on cracking resistance to compare to long-term (8-10 years) field performance. Intermediate-temperature cracking resistance was evaluated through cracking tolerance index (CT_{Index}) by the Indirect Tensile Cracking Test (IDT-CT), more commonly called the Ideal Cracking Test (IDEAL-CT), per ASTM D8225 at 25˚C and flexibility index (FI) by the Illinois Flexibility Index Test (I-FIT) per AASHTO T 393 at 25˚C. CT_{Index} measures resistance to both crack initiation and propagation, while FI of a notched specimen measures primarily resistance to crack propagation. Low-temperature cracking resistance was evaluated through cracking resistance index due to environmental conditions (CRI_{Env}) by the Uniaxial Thermal Stress and Strain Test (UTSST) per ASTM D8303 and low-temperature cracking index (LCI) by the modified low-temperature semi-circular bending (LT-SCB) test per modified AASHTO T 394 proposed by the Airfield Asphalt Pavement Technology Program (AAPTP) BMD cracking study. Modifications included line-load displacement control for 0.15 mm/min loading rate on a 50 mm thick specimen with a 15 mm notch. LCI was defined as the ratio of the critical strain energy to the peak and the pre-peak secant slope between 25% and 75% of peak load. CRI_{Env} measures resistance to both crack initiation and propagation, while LCI of a notched specimen primarily measures resistance to crack propagation. Both low temperature cracking tests were conducted at -18 ˚C (NV, WI, DE) and -12 ˚C (IN, TX), which were 10 ˚C warmer than the low-temperature PG grades required by the corresponding climates.

2.2.1 Intermediate-Temperature Cracking

FI data from NCHRP Project 09-58 included results from laboratory mixed-laboratory compacted (LMLC) and reheated plant mixed-laboratory compacted (RPMLC) specimens and cores taken at construction. Available results for LMLC specimens included those after short-term oven aging (4hrs@135C per previous version of AASHTO R 30 (2014)) and after both short- and long-term aging (5d@85C for compacted specimens per AASHTO R 30 (2014)). Available results for RPMLC specimens included those with only reheating and after long-term aging (5d@85C for compacted specimens per previous version of AASHTO R 30 (2014)). FI and CT_{Index} were measured for cores procured at the current field performance monitoring session to tie to NCHRP Project 09-58 testing and add IDEAL-CT as it is adopted or considered for use in BMD systems across the U.S.

2.2.2 Low Temperature-Cracking

CRI_{Env} data from NCHRP Project 09-58 was more limited across field sections but did include results from LMLC and RPMLC specimens and cores taken at construction with some of the same aging conditions as for FI results. While not included in AASHTO PP 127 that is focused on mix design, a threshold of 38 was established based on early field performance for RPMLC with LTOA in NCHRP Project 09-58.

CRI_{Env} and LCI were measured for cores procured at the current field performance monitoring session to tie to NCHRP Project 09-58 testing and add the modified LT-SCB test per AASHTO T 394 as it is proposed for use in BMD systems based on the Airfield Asphalt Pavement Technology Program (AAPTP) cracking effort. Modifications include line load displacement control, 0.15mm/min loading rate, and 50mm specimen thickness.

2.2.3 Binder Testing

After completion of intermediate-temperature cracking tests on the field cores procured at the current field performance monitoring session, the binder was extracted and recovered. An automated extraction system per ASTM D8159 was used followed by recovery through a rotary evaporator per ASTM D5404. The traditional Glover-Rowe (G-R) parameter at 15°C and 0.005 rad/sec was calculated based on measurements from the Dynamic Shear Rheometer (DSR) per AASHTO T 315. The available data from NCHRP Project 09-58 was from laboratory-prepared binder blends and not extracted and recovered from cores collected at construction. The Poker Chip Ductility (PCD) test was also conducted to determine percent ductility per AASHTO TP 150 at 25˚C with a threshold of 150% [18].

2.3 Comparative Analysis Tools

This section presents the comparative analysis tools used to determine if field sections exhibit similar or different laboratory performance and explore the relationship between laboratory and field performance.

2.3.1 Statistical Comparisons – Laboratory Results

For each field project, results from each laboratory test were assumed to be normally distributed and subjected to Levene’s test to check the assumption of equal variances. If there were no statistical differences in variance across field sections, an analysis of variance (ANOVA) and Tukey’s honestly significant difference (HSD) comparisons were conducted to determine which field sections were statistically different at a 0.05 significance level. If the Levene test detected statistically different variances among field sections, Games-Howell comparisons were used to determine statistical differences. UTSST results only included two replicates per section and thus could not be analyzed for statistical differences.

2.3.2 Field versus Laboratory

Since each field project was monitored at a different CDD, a consistent threshold of 70,000 was adopted as a critical evaluation point also used at the National Center for Asphalt Technology (NCAT) test track [**Error! Reference source not found.**]. The field performance in terms of percent lane cracking for each field section was determined at 70,000 CDD through linear interpolation. This field performance was then used to compare with laboratory performance results of LMLC and RPMLC specimens gathered from NCHRP Project 09-58. Field performance at the current CDD was used to compare with laboratory results from aged field cores procured at the current performance monitoring session.
Because the IDEAL-CT was not conducted in NCHRP Project 09-58, an estimated CT_{Index} at construction was determined through a correlation between CT_{Index} and FI measured on the aged field cores shown in Figure 4.

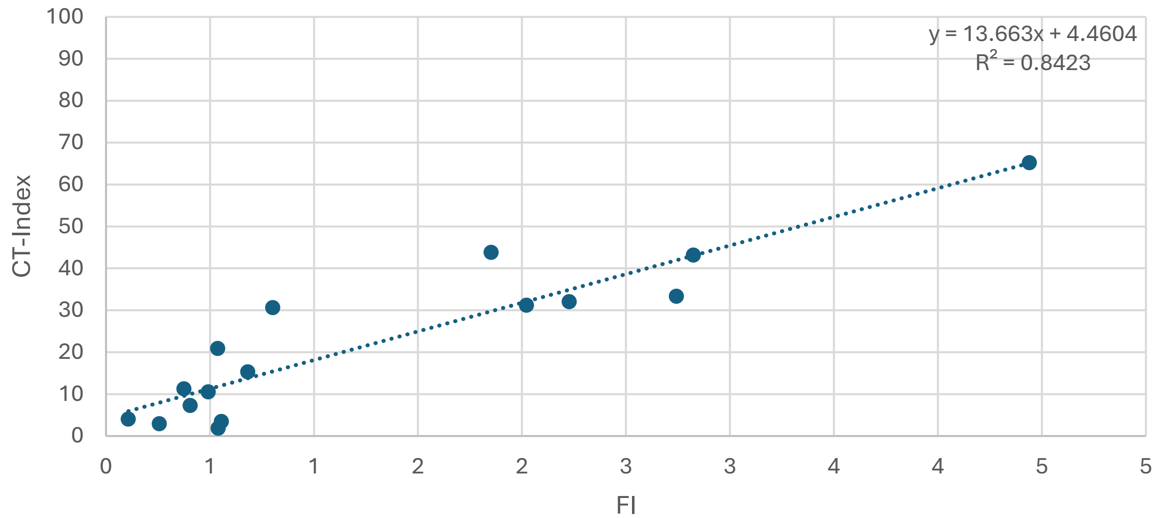


FIGURE 4 CT$_{\text{INDEX}}$ AND FI CORRELATION FROM AGED FIELD CORE RESULTS.

After the CT$_{\text{Index}}$ at construction was estimated, aging in the field was compared to aging in the laboratory. Field core data from NCHRP Project 09-58 was compared to the current aged field core results. The current aged field core results were divided by those estimated for the construction cores to evaluate how well the mixtures retained their cracking performance after 8-10 years in-service. From the laboratory results, the LTOA results were divided by the STOA results to obtain a similar cracking resistance retention ratio. These ratios were then compared to assess the laboratory aging protocols with respect to in-service aging in the field.

2.3.3 Confusion Analysis

A confusion analysis was also utilized to tie field performance to laboratory results from NCHRP Project 09-58. Table 11 provides the framework for the confusion matrix. Laboratory results are classified as passing or failing based on the laboratory thresholds contained in AASHTO PP 127 shown in Table 8 or from other efforts as previously noted. Field performance is also classified as passing or failing based on the expanded MAP-21 thresholds shown in Table 9 with failing defined as marginal or poor performance. Thus, when the laboratory results and field performance are compared, there are four outcomes for each comparison: passing both laboratory and field, failing both laboratory and field, failing in the laboratory and passing in the field, and passing in the laboratory and failing in the field. The corresponding colors for each of these scenarios are also shown in Table 11. When the results are either both passing (P-P) or both failing (F-F) in both the laboratory and the field, the results are correlated. When both results pass their thresholds, it is considered a true positive (TP) outcome. When both results fail to reach their thresholds, it is considered a true negative (TN) outcome. When the laboratory fails and the field passes (F-P) or the laboratory passes and the field fails (P-F), the results are uncorrelated. When the results pass in the laboratory but fail in the field, it is considered a false positive (FP) outcome. When the results fail in the laboratory and pass in the field, it is considered a false negative (FN) outcome. Of the two uncorrelated results, the FP outcome is of more concern than the FN outcome that is more conservative. A limitation of this analysis is that differences in traffic are not accounted for, but cracking was limited to top-down cracking and reflection cracking in the surface mixture. Thus, the assumption that the pavement structure was designed for the various traffic levels was upheld.

TABLE 11 CONFUSION MATRIX FRAMEWORK

Performance Classification		Laboratory Performance Threshold	
		Pass	Fail
Field Performance Threshold	Pass	Pass–Pass (P-P), True Positive (TP)	Fail–Pass (F-P), False Negative (FN)
	Fail	Pass–Fail (P-F), False Positive (FP)	Fail–Fail (F-F), True Negative (TN)

Once the confusion matrix outcomes are determined, those results were further analyzed through accuracy, precision, sensitivity, specificity, and balanced accuracy parameters based on equations 6, 7, 8, 9 [20], and 10 [21].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

$$\text{Sensitivity (True Positive Rate)} = \frac{TP}{TP+FN} \quad (8)$$

$$\text{Specificity (True Negative Rate)} = \frac{TN}{TN+FP} \quad (9)$$

$$\text{Balanced Accuracy} = \frac{Sensitivity+Specificity}{2} \quad (10)$$

Accuracy represents the overall proportion of correct agreement between laboratory and field performance. Precision reflects the reliability of a laboratory “pass” in predicting acceptable field performance. Sensitivity indicates the ability of the laboratory criteria to correctly identify mixtures that perform well in the field, while specificity reflects the ability to correctly identify mixtures that perform poorly in the field. Balanced accuracy, defined as the average of sensitivity and specificity, provides a more representative measure of performance classification for small or unbalanced datasets.

3.0 Results and Analysis

Figure 5 shows the CDD at the current field performance monitoring session for each field section with all sections beyond the 70,000 CDD critical age. This threshold of 70,000 CDD reflects approximately 4-8 years of in service field aging for this project. The results of the field performance evaluation were compiled and can be seen in Figure 6 presents the overall field performance results for each field section with respect to percent lane cracking. The DE field sections exhibit poor performance, and the rejuvenated IN section reflects the worst cracking performance. The results are color coded to reflect the categories defined in Table 9.

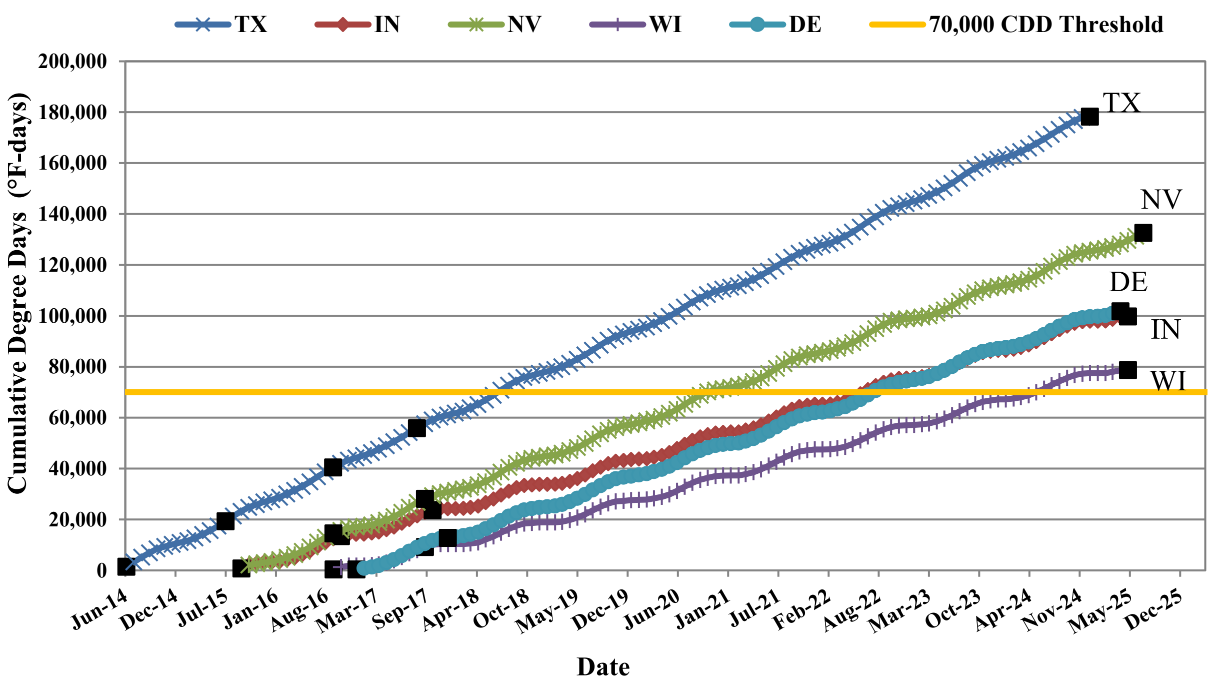
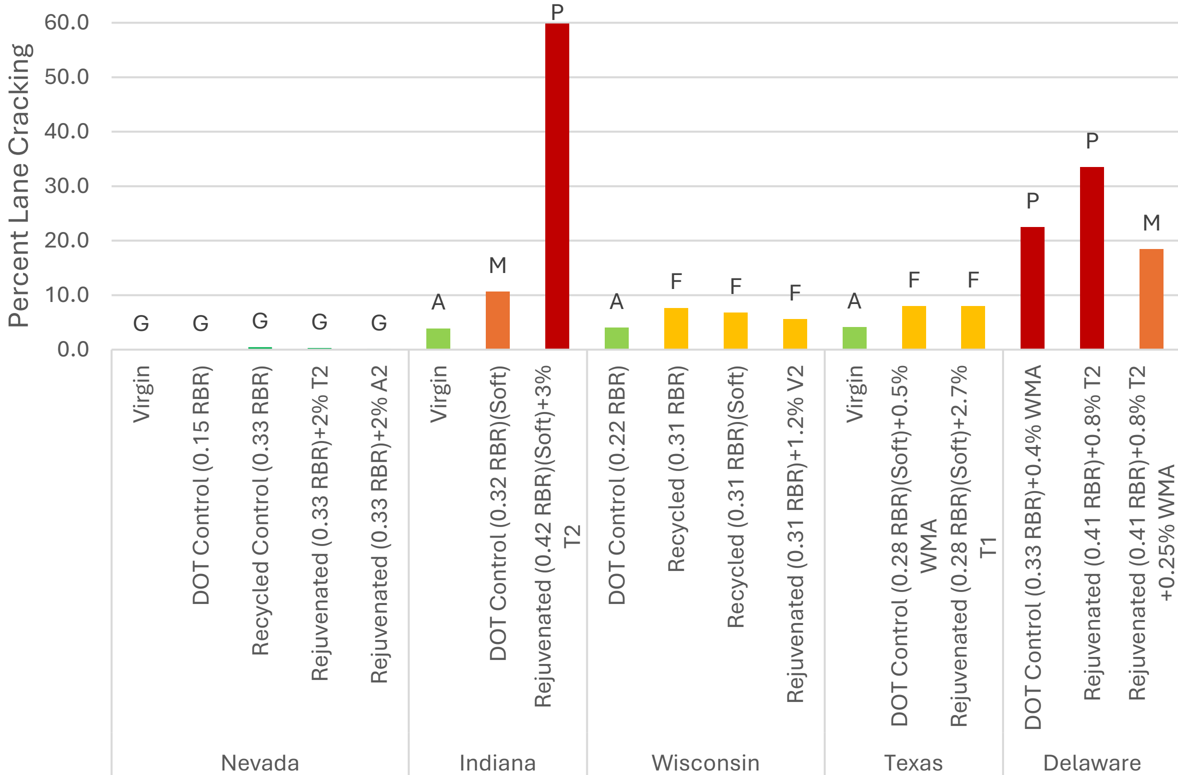


FIGURE 5 CURRENT CDD



Notes: Dark green = Good (<3% Lane Cracking); light green = Adequate (<5% Lane Cracking); Yellow = Fair (<10% Lane Cracking); Orange = Marginal (<20% Lane Cracking); Red = Poor (≥20% Lane Cracking)

FIGURE 6 FIELD CRACKING PERFORMANCE

Among the cracking that was observed in the field sections, the split between transverse cracking and longitudinal cracking is shown in Figure 7. The NV and WI field sections exhibited an overwhelming majority of transverse cracking, while the DE field sections exhibited mostly longitudinal cracking. The TX and IN field projects show a mixture of transverse and longitudinal cracking.

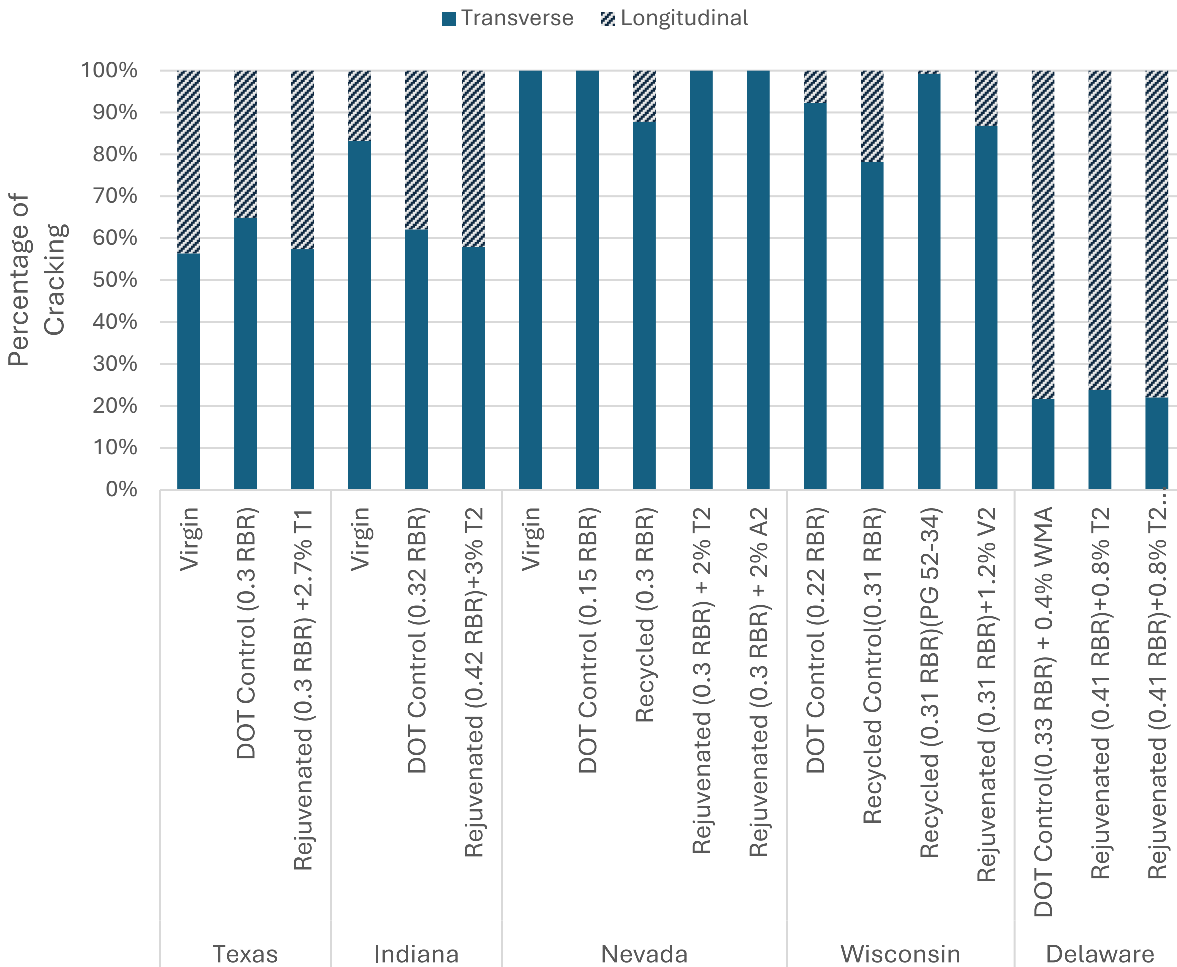
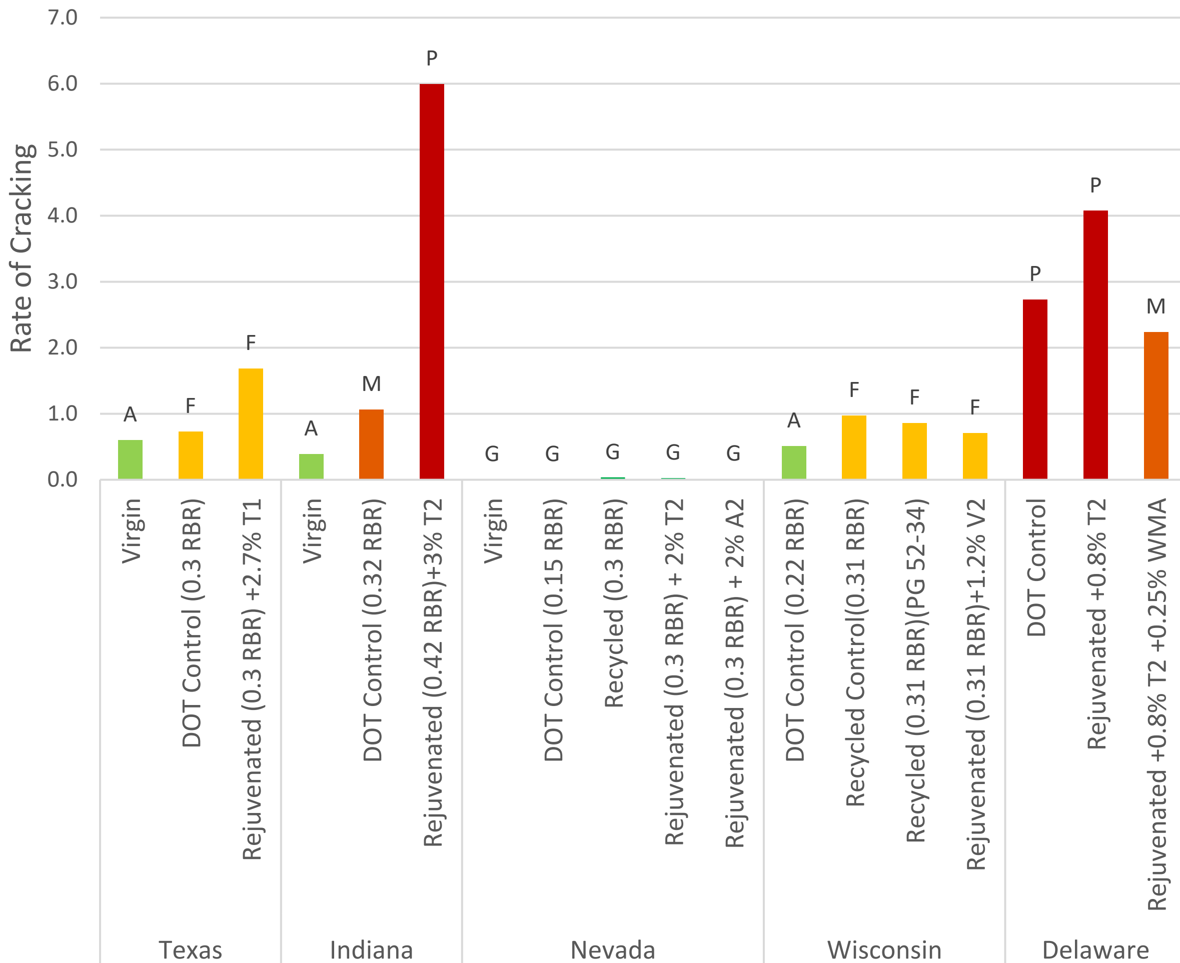


FIGURE 7 TYPES OF FIELD CRACKING

Figure 8 provides another overall assessment of the evolution of field cracking by showing the lane cracking rate at 70,000 CDD, defined in Equation 9 per 10,000 CDD. Based on a comparison across all 18 field sections, a candidate threshold of 1.5% could serve to identify field sections requiring maintenance or treatment which separates a majority of the poor and marginal performing sections from the others.

$$Average\ Rate\ of\ Cracking = \frac{Percent\ Lane\ Cracking}{10{,}000\ CDD} \quad (9)$$

The sections from the DE field project as well as the rejuvenated section from the IN field project showed high rates of cracking. The rejuvenated section from the TX field project also initially showed high rates of cracking, but the placement of a chip seal prior to the current field performance monitoring session reduced the rate of distress accumulation.



Notes: Dark green = Good (<3% Lane Cracking); light green = Adequate (<5% Lane Cracking); Yellow = Fair (<10% Lane Cracking); Orange = Marginal (<20% Lane Cracking); Red = Poor (≥20% Lane Cracking)

FIGURE 8 AVERAGE RATE OF % LANE CRACKING AT 70,000 CDD

Within each field project the traffic was consistent (Table 1), but air voids varied across the sections as shown in Table 10.

3.1 Nevada

Of all the 18 field sections in the CAPRI project, the NV field sections showed the best field performance as shown in Figure 9 with a maximum y-axis of 3.5% lane cracking. At the time of the field evaluation in April 2025 at 9.8 years in service, all the NV field sections exhibited percent lane cracking values below 0.5%. At 70,000 CDD, the NV field sections all had less than 0.25% lane cracking and were categorized as good.

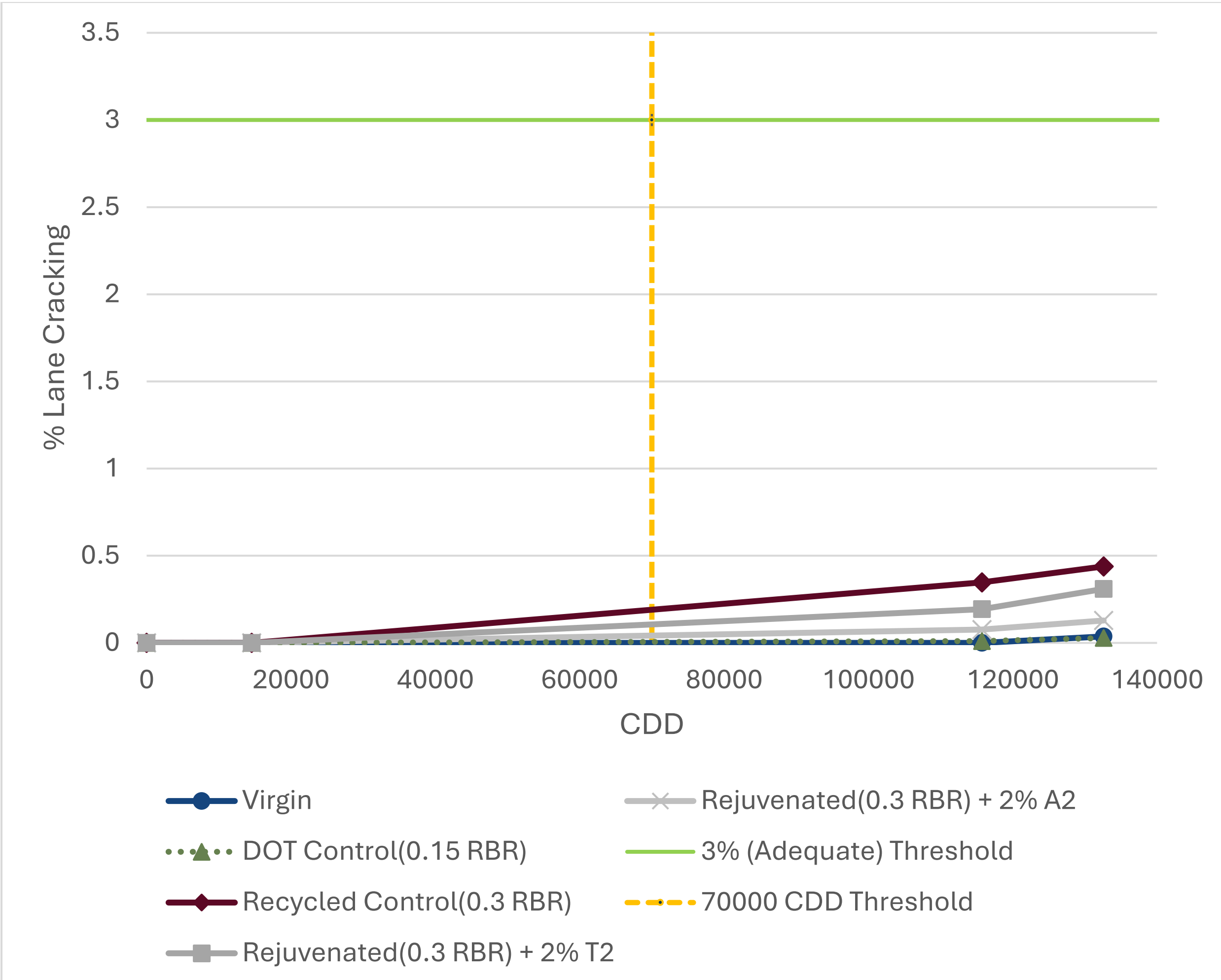


FIGURE 9 NV FIELD PERFORMANCE

3.1.1 Field (70,000 CDD) versus Laboratory Comparison

The data available from NCHRP Project 09-58 is shown in Table 12 in green and with a check mark.

TABLE 12 NV LABORATORY DATA AVAILABILITY

		FI w/ LMLC		FI w/ RPMLC		CRI_{Env} w/ LMLC		CRI_{Env} w/ RPMLC	
		STOA	LTOA	No LTOA	LTOA	STOA	LTOA	No LTOA	LTOA
Nevada	Virgin	-	-	✓	✓	-	-	-	✓
	DOT Control (0.15 RBR)	✓	✓	✓	✓	✓	✓	-	✓
	Recycled Control(0.3 RBR)	-	-	✓	✓	✓	✓	-	✓
	Rejuvenated (0.3 RBR) + 2% T2	✓	✓	✓	✓	✓	✓	-	✓
	Rejuvenated (0.3 RBR) + 2% A2	✓	✓	✓	✓	✓	✓	-	✓

Figure 10 and Figure 11 present the average FI results for LMLC and RPMLC specimens at available aging conditions, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD or Games-Howell at a 0.05 significance level. Letters in uppercase were classified by Tukey's HSD, and letters in lower case were classified by Games-Howell. The current threshold in AASHTO PP 127 for FI of LMLC specimens after STOA of 7 is also provided. From the results for the LMLC specimens, the DOT control section met this threshold and performed better than the rejuvenated sections. The results also reflected the estimated field performance ranking at 70,000 CDD between the DOT control and rejuvenated sections. Based on the results for the RPMLC specimens, the virgin section performed better than the rest of the sections, which also reflected the ranking in the field.

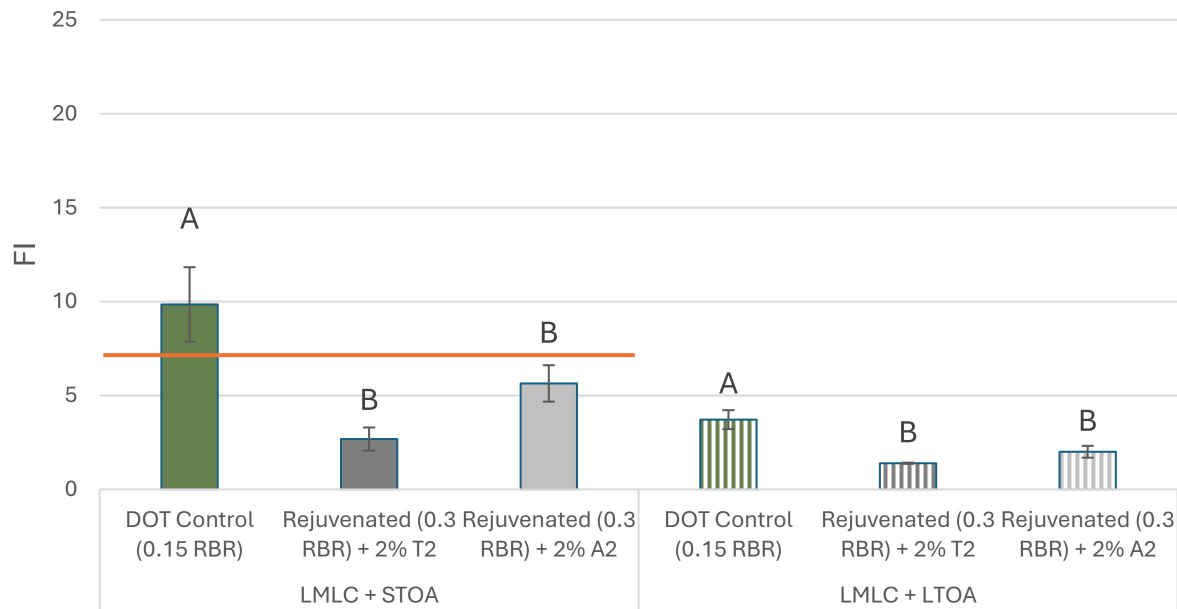


FIGURE 10 NV FI RESULTS FOR LMLC SPECIMENS

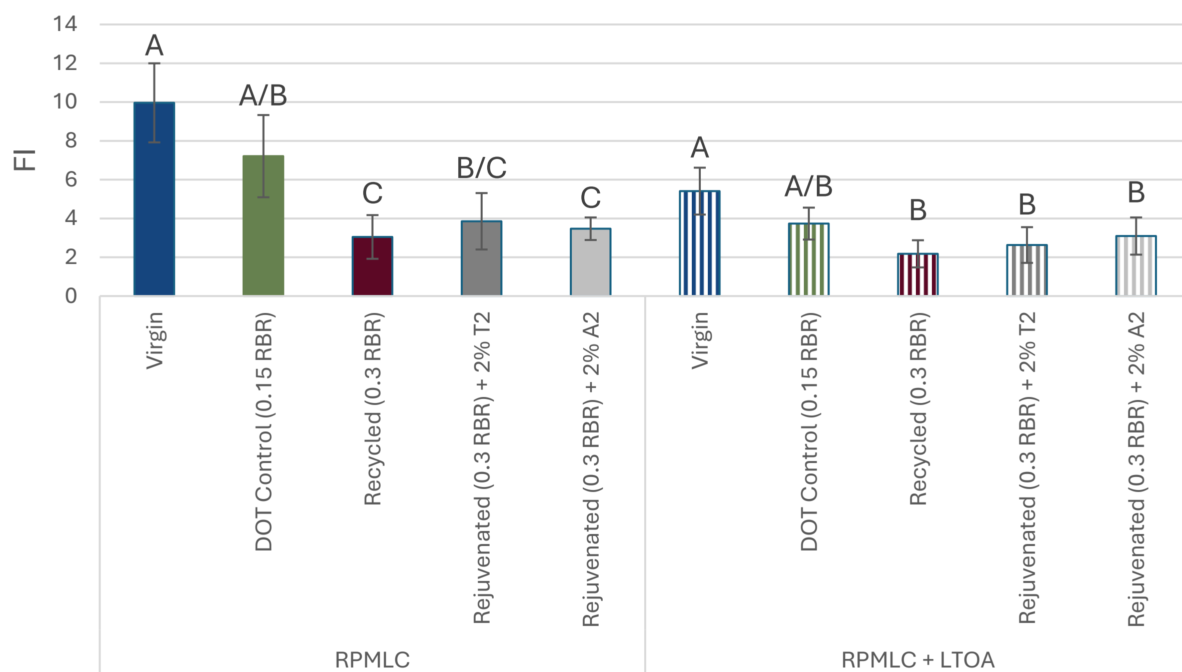


FIGURE 11 NV FI RESULTS FOR RPMLC SPECIMENS

Figure 12 and Figure 13 present the average CRI_{Env} results for LMLC and RPMLC specimens at available aging conditions, respectively, with error bars indicating plus or minus one standard deviation from the average value. The current thresholds for CRI_{Env} of LMLC specimens after LTOA of 17 (AASHTO PP 127) and of RPMLC specimens after LTOA of 38

(NCHRP 09-58) are also provided. All the sections met both thresholds except the recycled section for RPMLC specimens. As with the FI results, the virgin section performed better than the rest of the sections based on the RPMLC results which reflected field performance rankings. LMLC results were more variable.

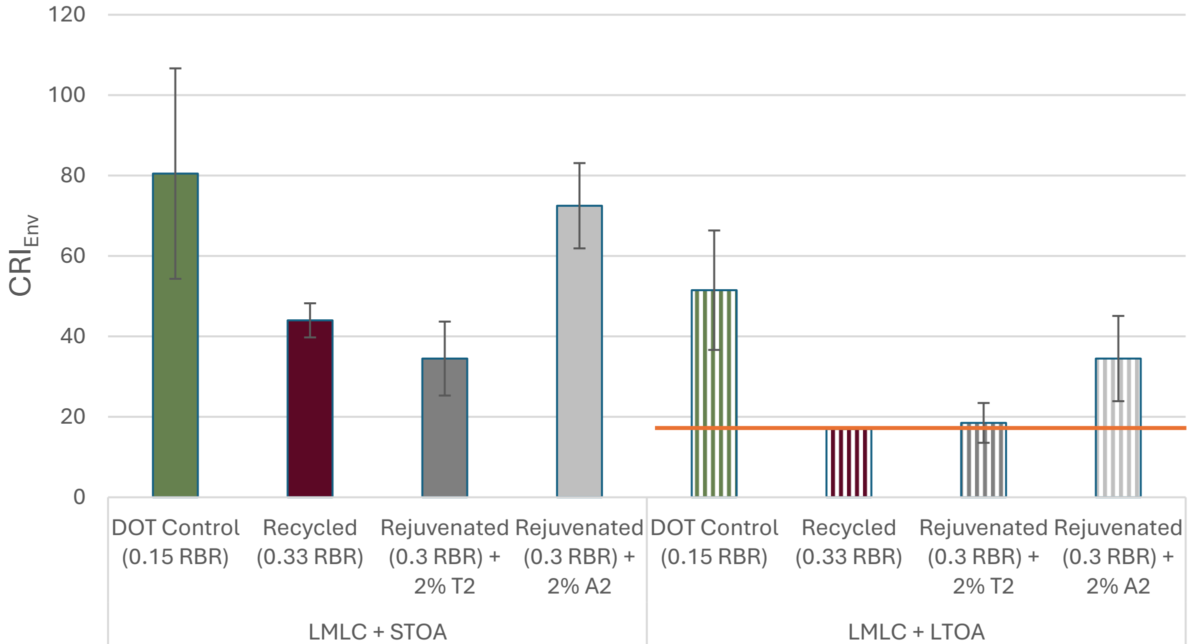


FIGURE 12 NV CRI_{ENV} RESULTS FOR LMLC SPECIMENS

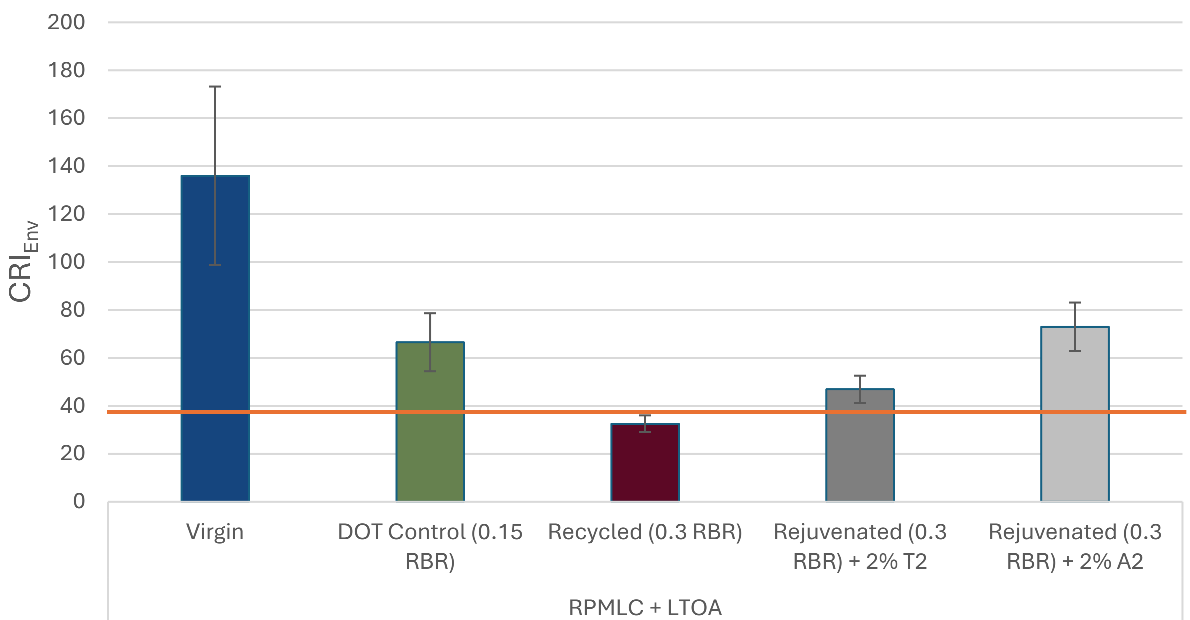


FIGURE 13 NV CRI_{ENV} RESULTS FOR RPMLC SPECIMENS

3.1.2 Field (Current) versus Aged Core Results

Because the laboratory test methods require specified specimen dimensions and air void levels (7%), thresholds established in AASHTO PP 127 are not applicable for cores. In addition, cores may reflect different aging levels compared to laboratory-prepared specimens, further limiting the direct use of the established criteria.

Figure 14 and Figure 15 present the average FI and CT_{Index} results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. For both intermediate-temperature cracking resistance parameters, the virgin and DOT control cores performed better than the recycled and rejuvenated cores at construction and with long-term field aging in service.

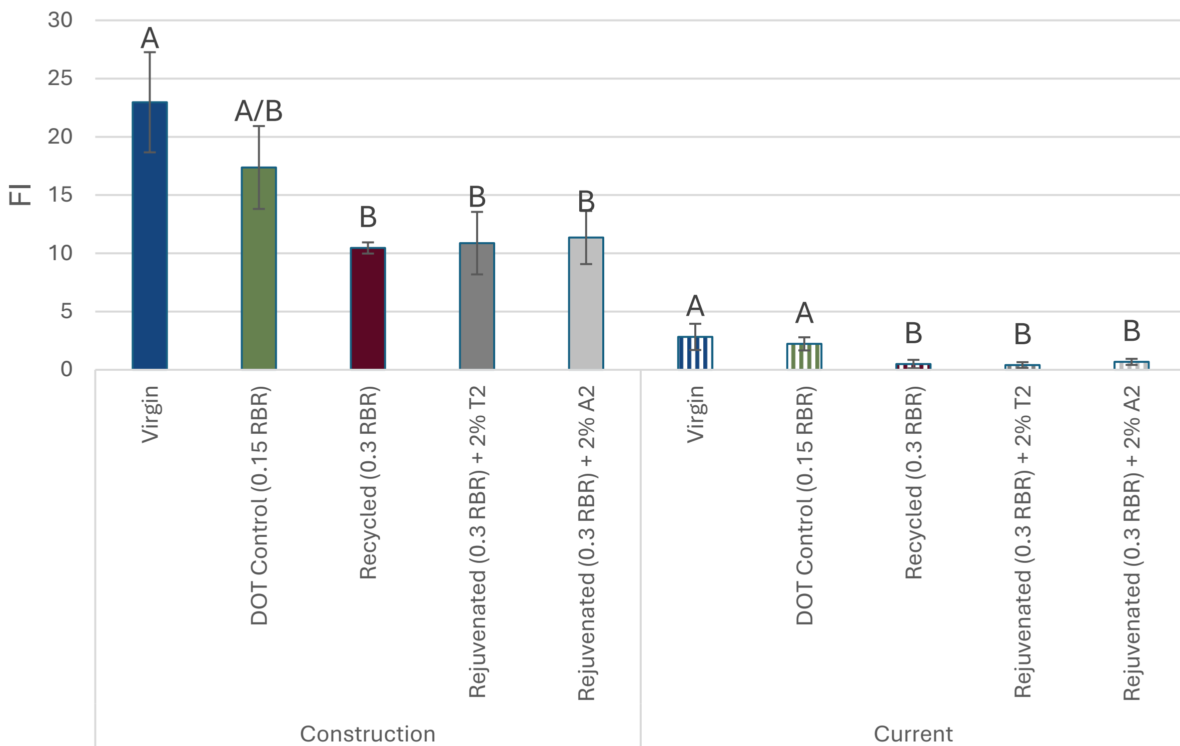


FIGURE 14 NV FI RESULTS FOR AGED FIELD CORES

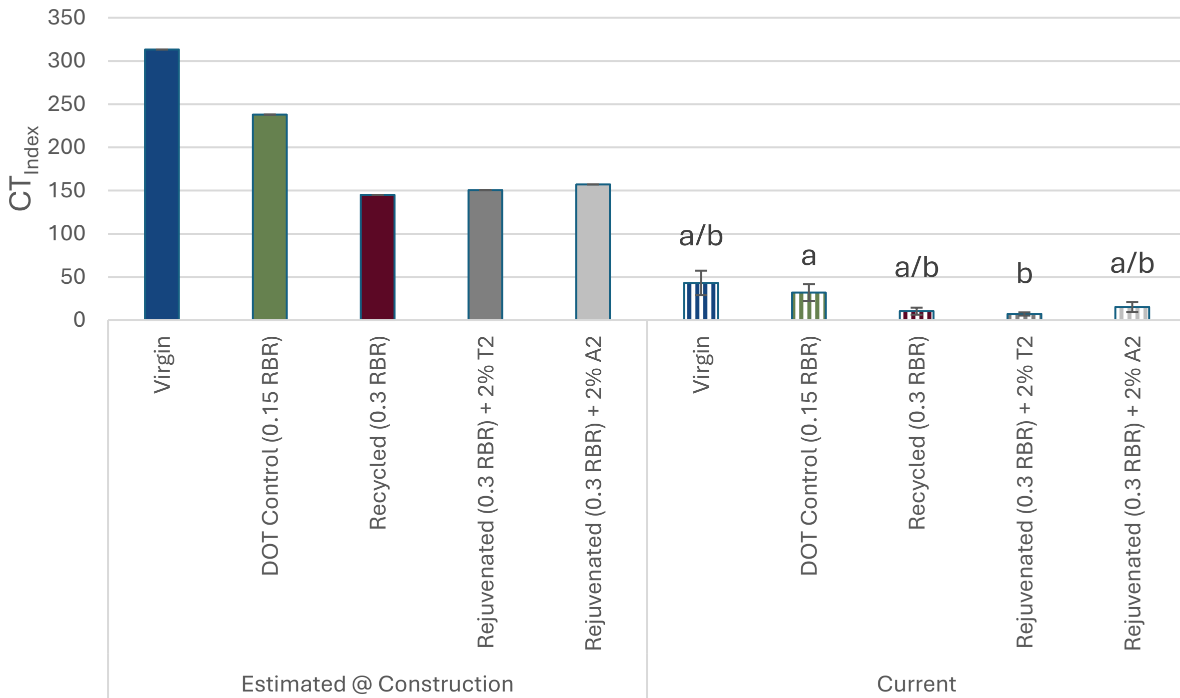


FIGURE 15 NV CT_{INDEX} RESULTS FOR AGED FIELD CORES

Figure 16 and Figure 17 present the average LCI and CRI_{Env} results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. The CRI_{Env} results show some differentiation with the virgin core performing the best at construction. LCI core results show no statistical difference between sections.

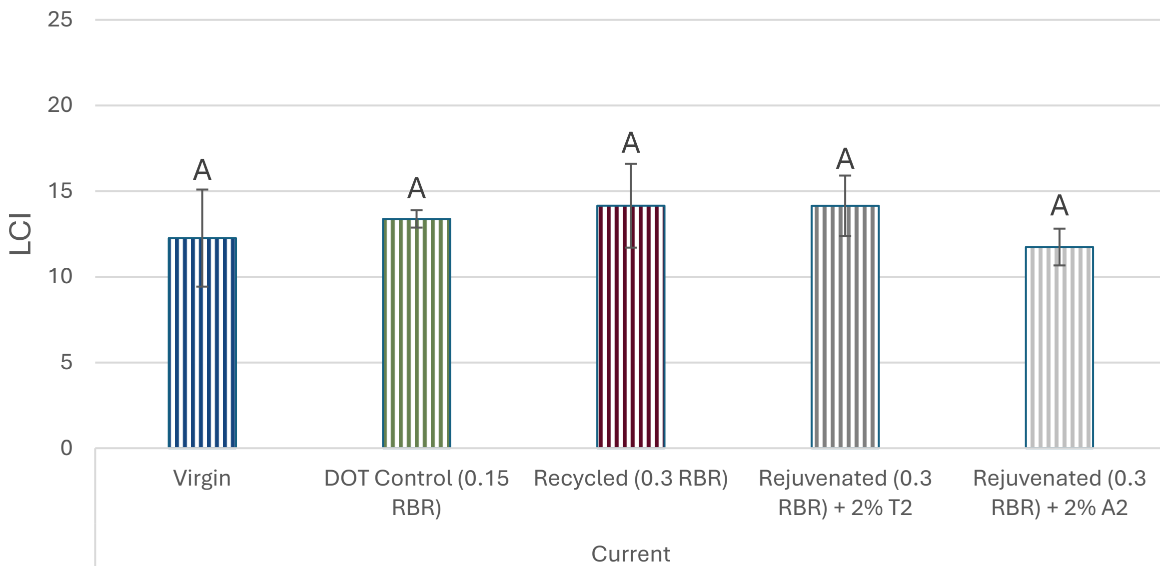


FIGURE 16 NV LCI RESULTS FOR AGED FIELD CORES

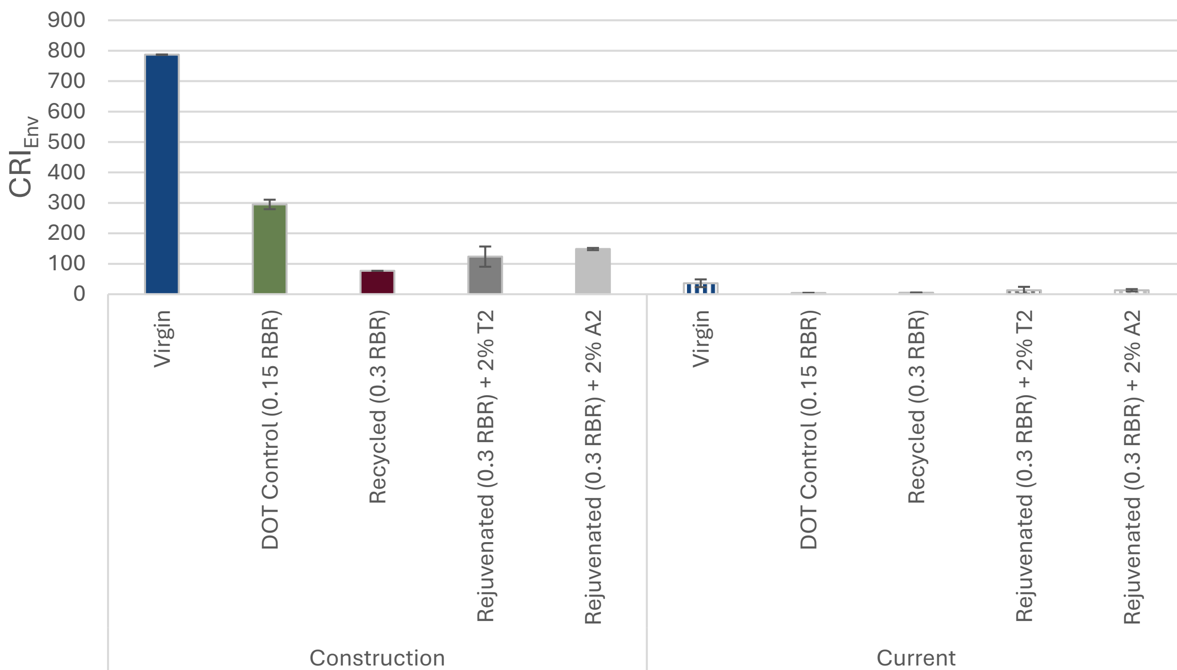


FIGURE 17 NV CRI_{ENV} RESULTS ADJUSTED TO -22°C FOR AGED FIELD CORES

3.2 Indiana

Figure 18 with a maximum y-axis of 60% lane cracking shows that the IN field sections exhibited clear differences in field performance. The rejuvenated section had the worst cracking in the IN field project. At 70,000 CDD, the virgin section was categorized as good, the DOT control section was classified as fair, and the rejuvenated section was assessed as poor. At the time of the field evaluation in May 2025 at 9.6 years in service, the performance of the virgin section had worsened to the adequate category, the DOT control section moved to the marginal classification, and the rejuvenated section remained in the poor category.

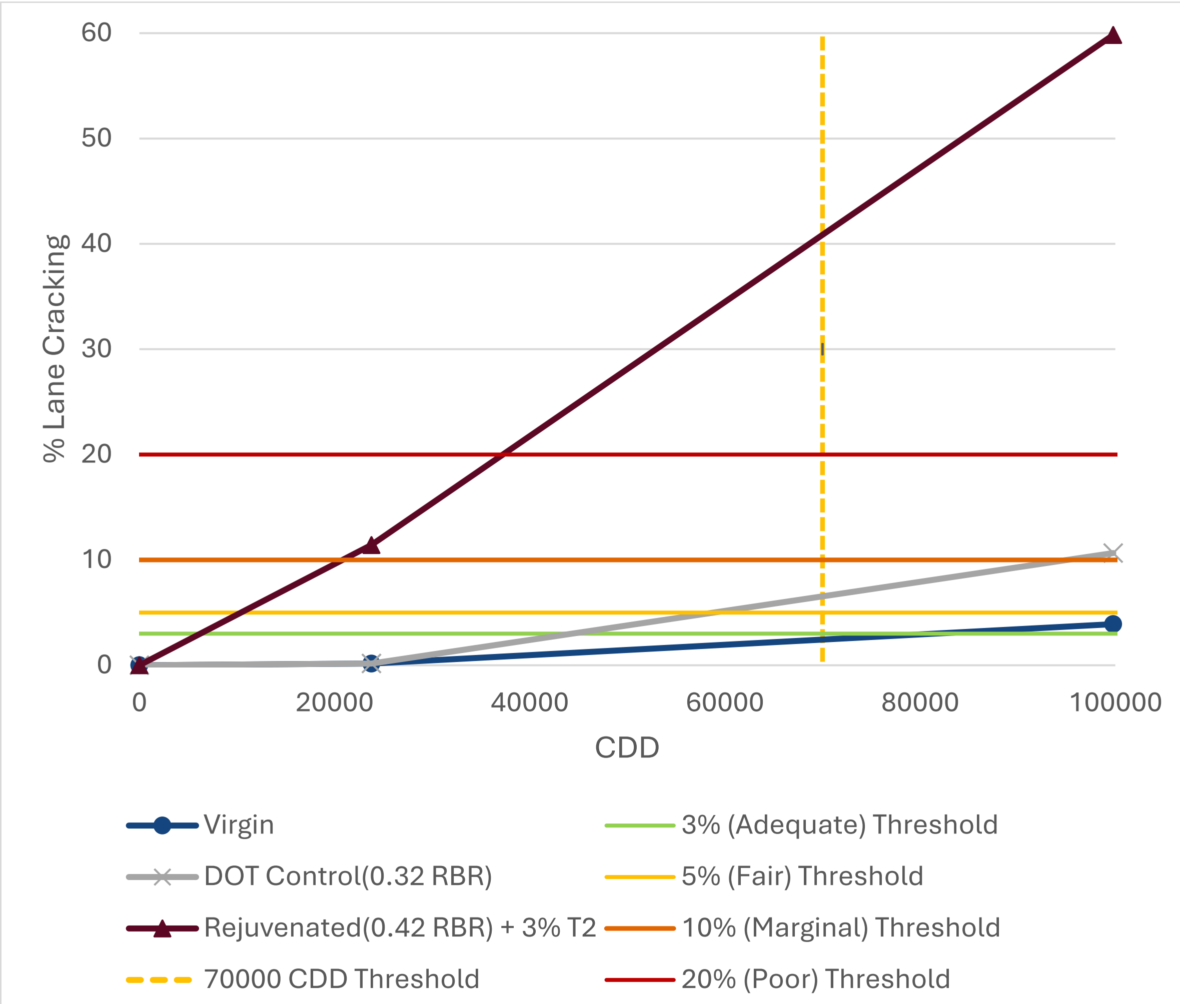


FIGURE 18 IN FIELD PERFORMANCE

3.2.1 Field (70,000 CDD) versus Laboratory Comparison

The data available from NCHRP Project 09-58 is shown in Table 13 in green and with a check mark. The unavailability of FI data for RPMLC specimens for the rejuvenated section is likely due to embrittlement of the high RBR with high RAS mixture.

TABLE 13 IN LABORATORY DATA AVAILABILITY

State	Section	FI w/ LMLC		FI w/ RPMLC		CRI_{Env} w/ LMLC		CRI_{Env} w/ RPMLC	
		STOA	LTOA	STOA	LTOA	STOA	LTOA	STOA	LTOA
Indiana	Virgin	✓	✓	✓	✓	-	-	✓	✓
	DOT Control (0.32 RBR) (Soft)	✓	✓	✓	✓	-	-	✓	✓
	Rejuvenated (0.42 RBR) (Soft)+3% T2	✓	✓	-	-	-	-	✓	✓

Figure 19 and Figure 20 present the average FI results for LMLC and RPMLC specimens at available aging conditions, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. The current threshold in AASHTO PP 127 for FI of LMLC specimens after STOA of 7 is also provided. From the results for the LMLC specimens, only the virgin section met this threshold and performed better than the rejuvenated section. These results reflected the estimated field performance ranking at 70,000 CDD.

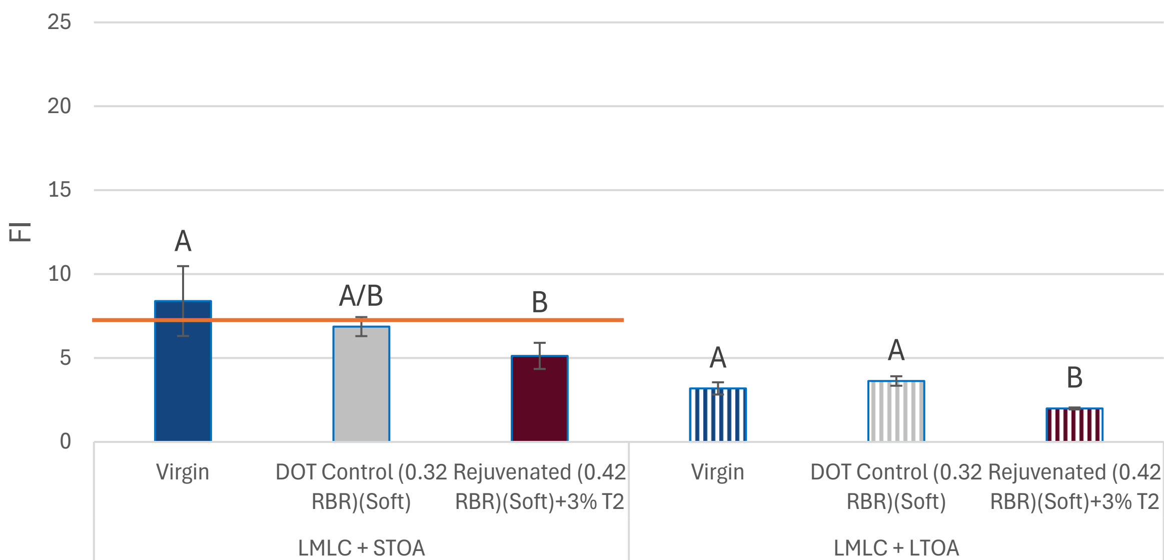


FIGURE 19 IN FI RESULTS FOR LMLC SPECIMENS

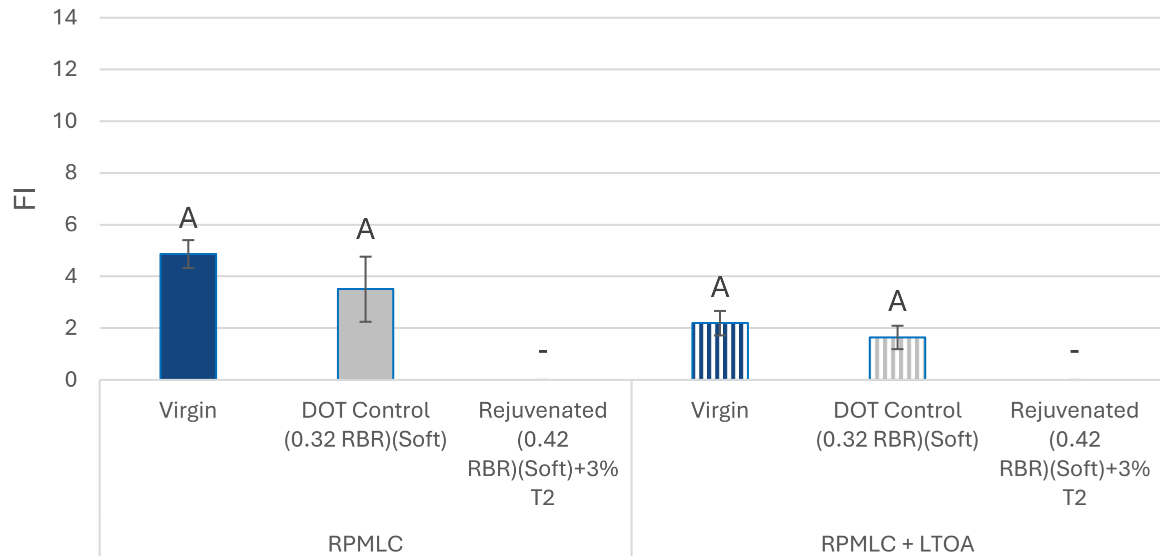


FIGURE 20 IN FI RESULTS FOR RPMLC SPECIMENS

Figure 21 presents the average CRI_{Env} results for RPMLC specimens at available aging conditions with error bars indicating plus or minus one standard deviation from the average value. The current threshold for CRI_{Env} of RPMLC specimens after LTOA of 38 (NCHRP 09-58) is also provided. None of the sections met this threshold. Based on these RPMLC results, the virgin and DOT control sections performed better than the rejuvenated section for both aging conditions, which also reflected field performance rankings.

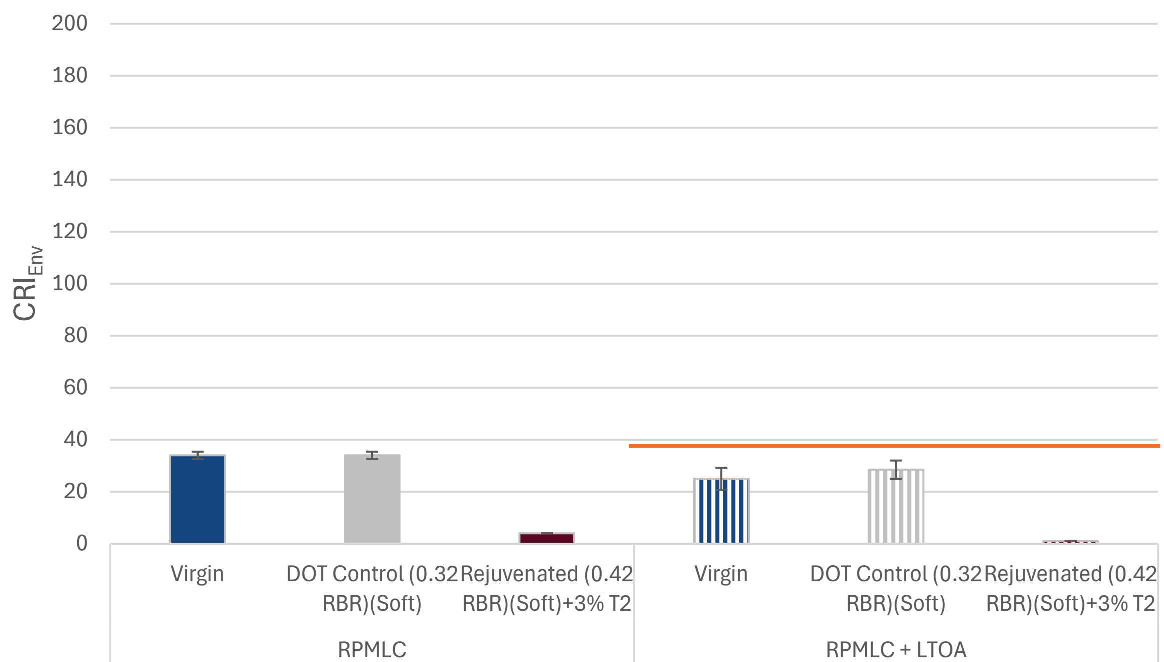


FIGURE 21 IN CRI_{ENV} RESULTS FOR RPMLC SPECIMENS

3.2.2 Field (Current) versus Aged Core Results

Because the laboratory test methods require specified specimen dimensions and air void levels (7%), thresholds established in AASHTO PP 127 are not applicable for cores. Figure 22 and Figure 23 present the average FI and CT_{Index} results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD or Games-Howell at a 0.05 significance level. Letters in uppercase were classified by Tukey's HSD, and letters in lowercase were classified by Games-Howell. For both intermediate-temperature cracking resistance parameters, the virgin section cores exhibit improved performance compared to the DOT control and rejuvenated section cores at construction and with long-term field aging in service. In agreement with field performance rankings, the FI also statistically differentiated the DOT control core from the rejuvenated core at construction, but not after long-term field aging.

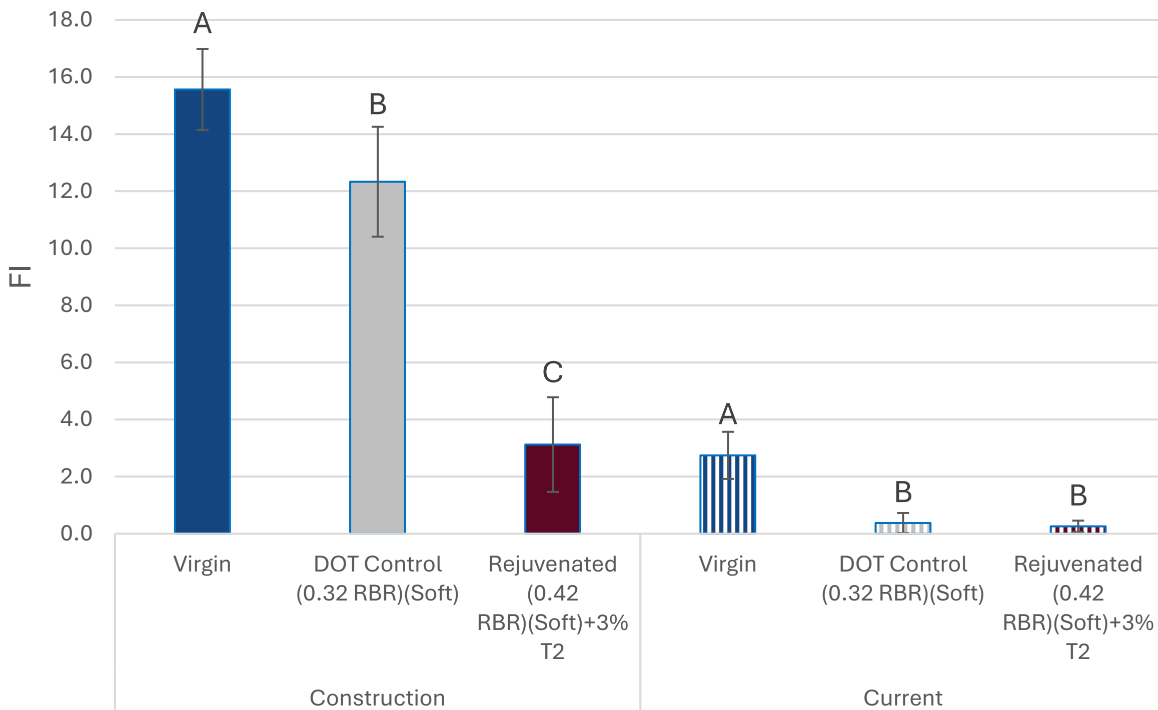


FIGURE 22 IN FI RESULTS FOR AGED FIELD CORES

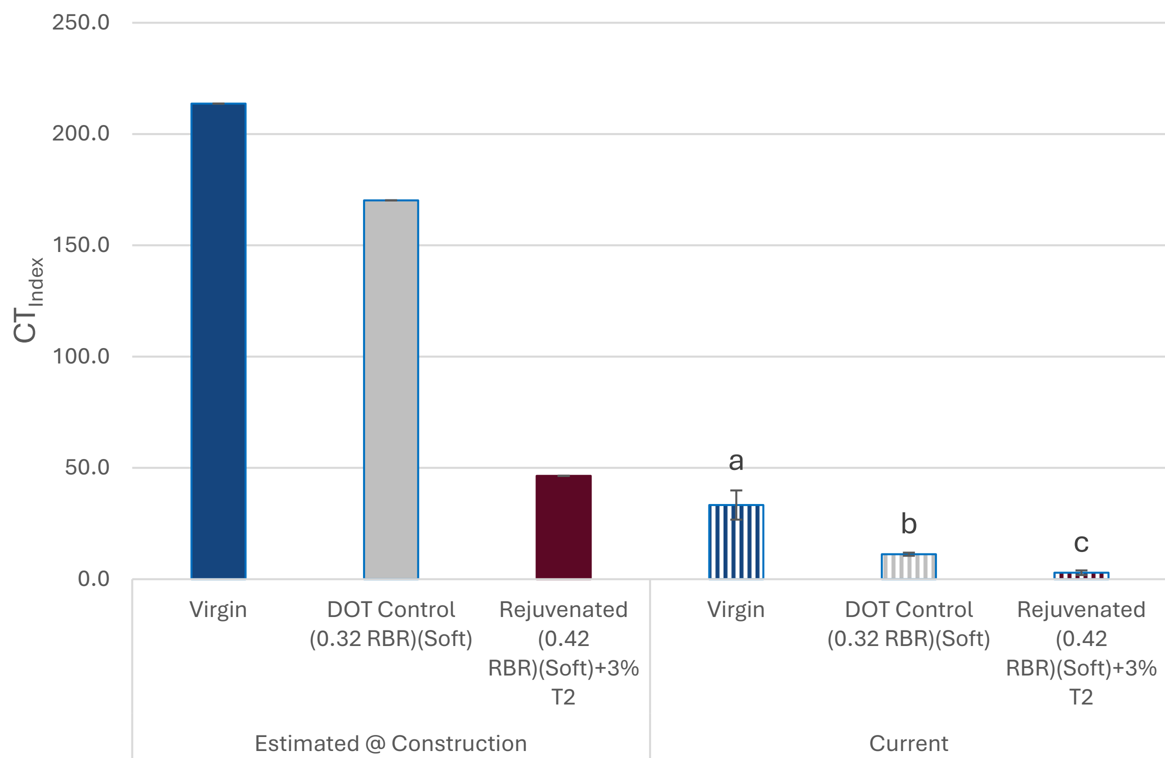


FIGURE 23 IN CT_{INDEX} RESULTS FOR AGED FIELD CORES

Figure 24 presents the average LCI results for field cores at available time periods with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. CRI_{Env} could not be measured on the IN field cores because they were too thin. LCI core results show no statistical difference between sections.

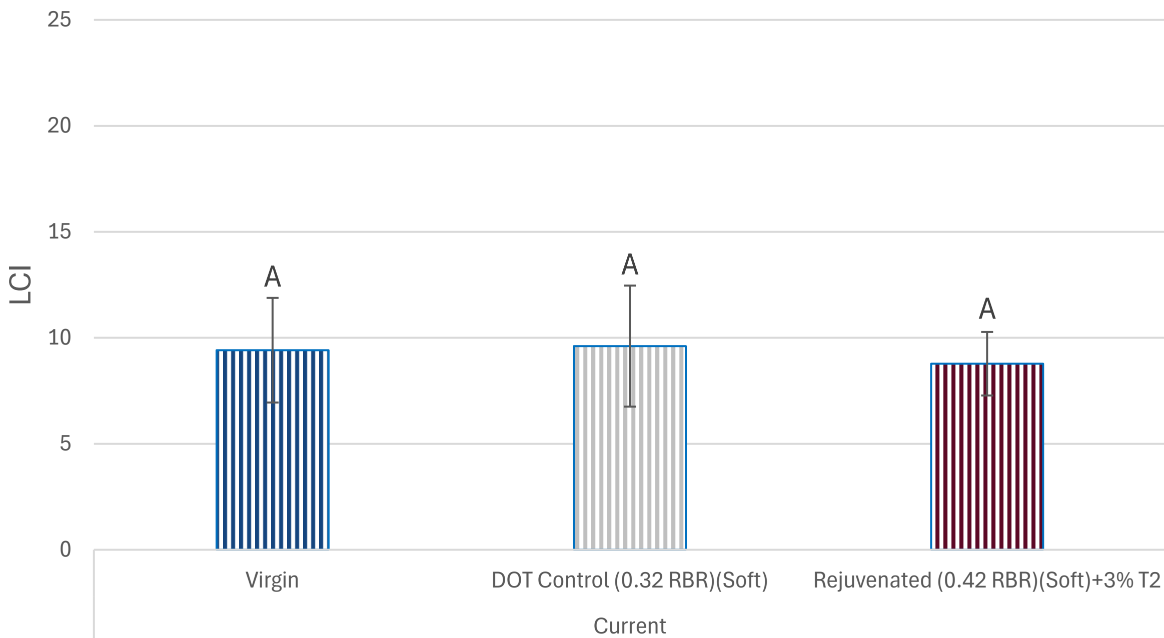


FIGURE 24 IN LCI RESULTS FOR AGED FIELD CORES

3.3 Wisconsin

Figure 25 with a maximum y-axis of 12% lane cracking shows that the WI field sections consistently exhibited some cracking. At 70,000 CDD, the performance of the DOT control and rejuvenated sections was categorized as adequate, but the rejuvenated section was approaching the fair category. At this critical age, both recycled sections were classified as fair. At the time of the field evaluation in May 2025 at 8.6 years in service, the performance of the DOT control section remained adequate while that of all other sections was classified as fair.

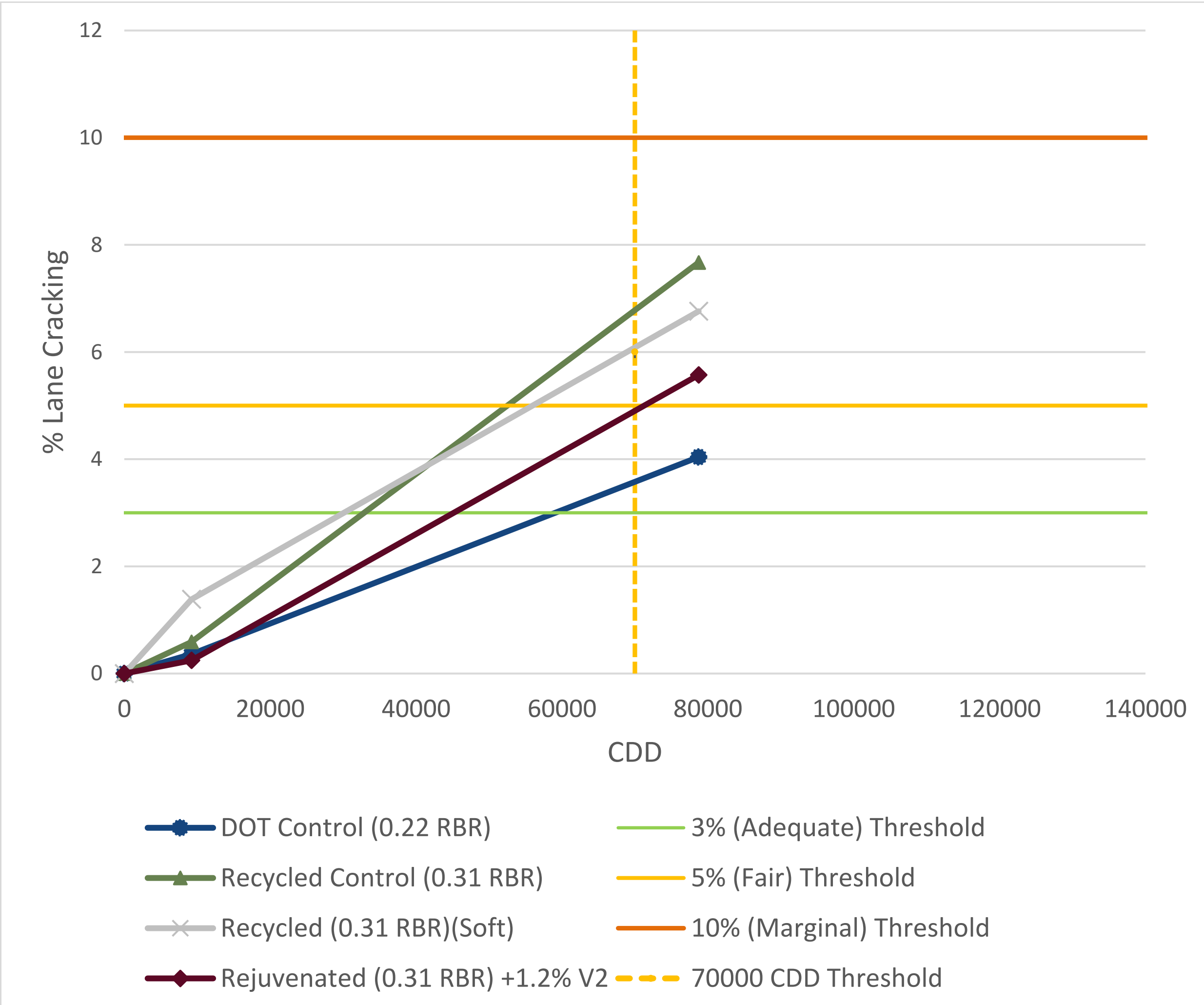


FIGURE 25 WI FIELD PERFORMANCE

3.3.1 Field (70,000 CDD) versus Laboratory Comparison

The complete set of data available from NCHRP Project 09-58 is shown in Table 14 in green and with a check mark.

Table 14 WI Laboratory Data Availability

State	Section	FI w/ LMLC		FI w/ RPMLC		CRI_{Env} w/ LMLC		CRI_{Env} w/ RPMLC	
		STOA	LTOA	STOA	LTOA	STOA	LTOA	STOA	LTOA
Wisconsin	DOT Control (0.22 RBR)	✓	✓	✓	✓	✓	✓	✓	✓
	Recycled Control(0.31 RBR)	✓	✓	✓	✓	✓	✓	✓	✓
	Recycled (0.31 RBR) (Soft)	✓	✓	✓	✓	✓	✓	✓	✓
	Rejuvenated (0.31 RBR) +1.2% V2	✓	✓	✓	✓	✓	✓	✓	✓

Figure 26 and Figure 27 present the average FI results for LMLC and RPMLC specimens at available aging conditions, respectively, with error bars indicating plus or minus one standard deviation and different letters indicating statistical differences by Tukey's HSD or Games-Howell at a 0.05 significance level. Letters in uppercase were classified by Tukey's HSD, and letters in lower case were classified by Games-Howell. The current threshold in AASHTO PP 127 for FI of LMLC specimens after STOA of 7 is also provided. From the results for the LMLC specimens, all sections met this threshold. The relative results for the recycled section across specimen types and aging conditions also reflected the estimated field performance ranking at 70,000 CDD.

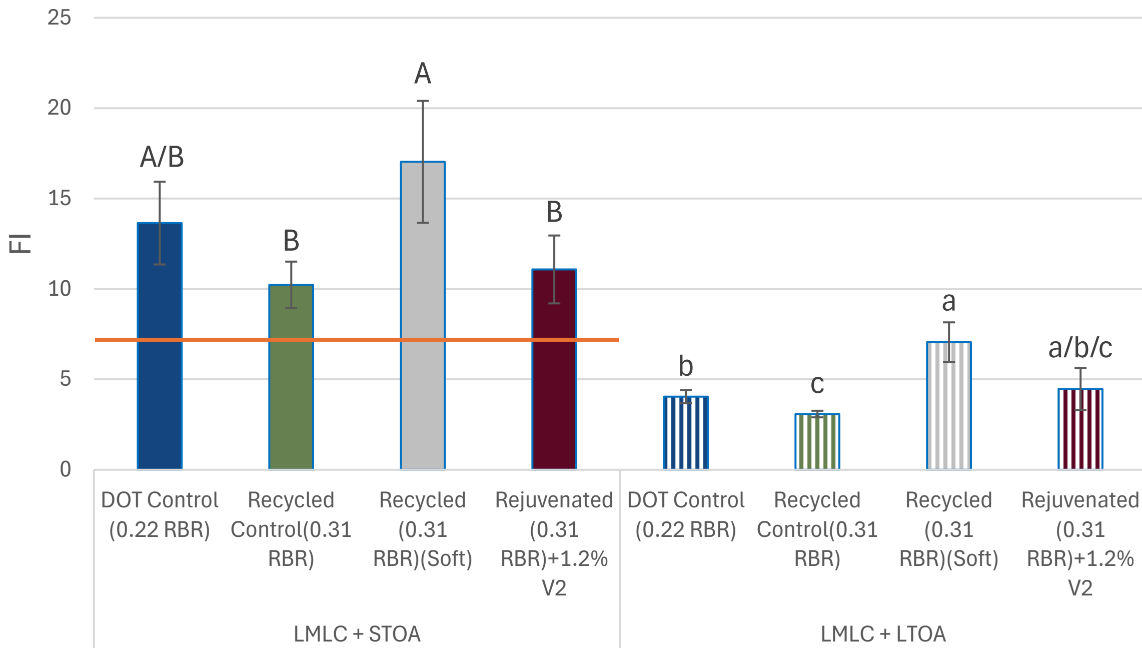


FIGURE 26 WI FI RESULTS FOR LMLC SPECIMENS

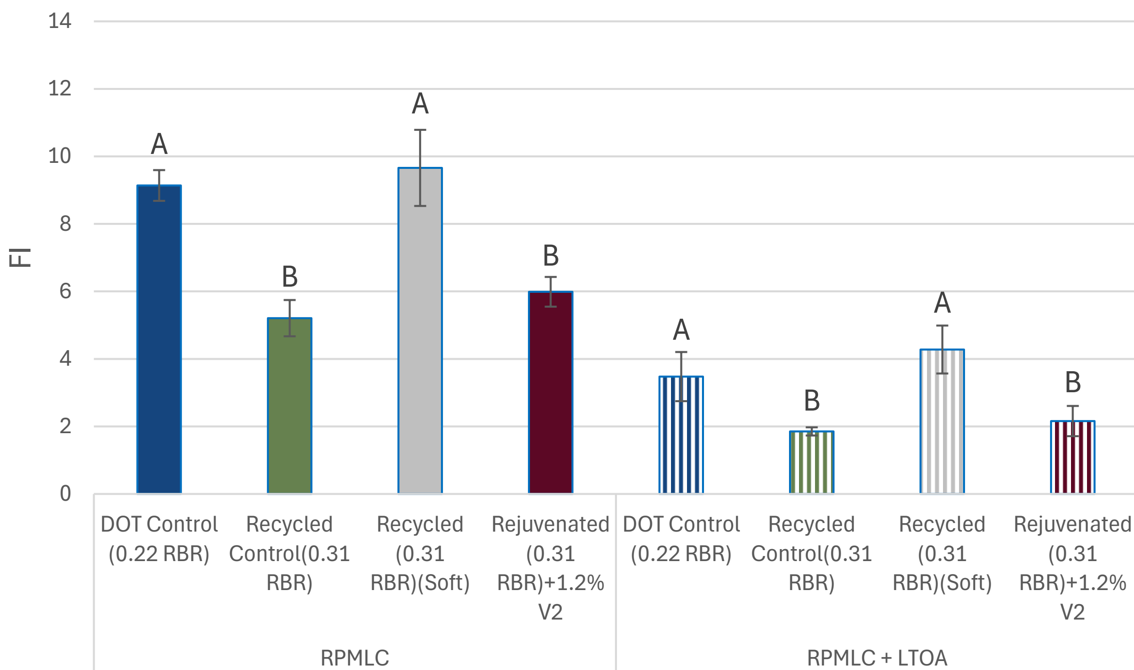


FIGURE 27 WI FI RESULTS FOR RPMLC SPECIMENS

Figure 28 and Figure 29 present the average CRI_{Env} results for LMLC and RPMLC specimens at available aging conditions, respectively, with error bars indicating plus or minus one standard deviation from the average value. The current thresholds for CRI_{Env} of LMLC specimens after LTOA of 17 (AASHTO PP 127) and of RPMLC specimens after LTOA of 38 (NCHRP 09-58) are also provided. For the LMLC results, the DOT control and recycled sections met the threshold. For the RPMLC results, only the recycled section with a softer binder met the threshold, highlighting a successful use of this cracking mitigation strategy. The only statistical differentiation was shown for the RPMLC specimens after LTOA, but the laboratory performance of the recycled section with a softer binder is not reflected in the field performance rankings.

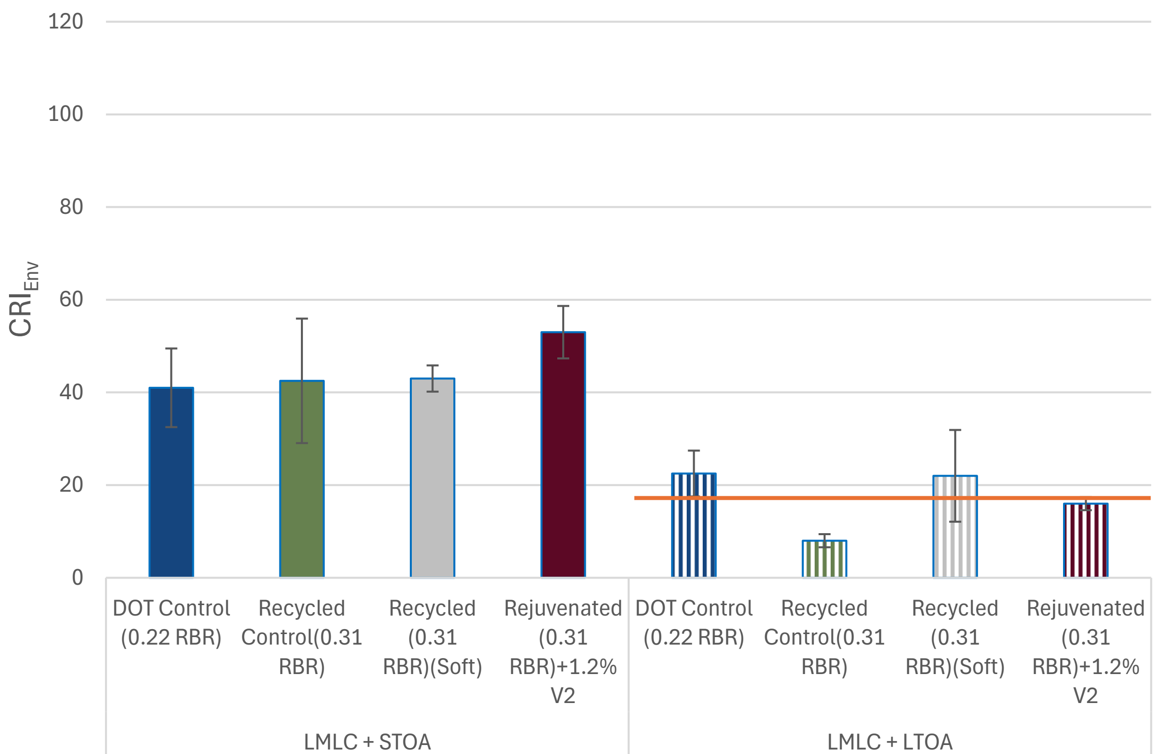


FIGURE 28 WI CRI_{ENV} RESULTS FOR LMLC SPECIMENS

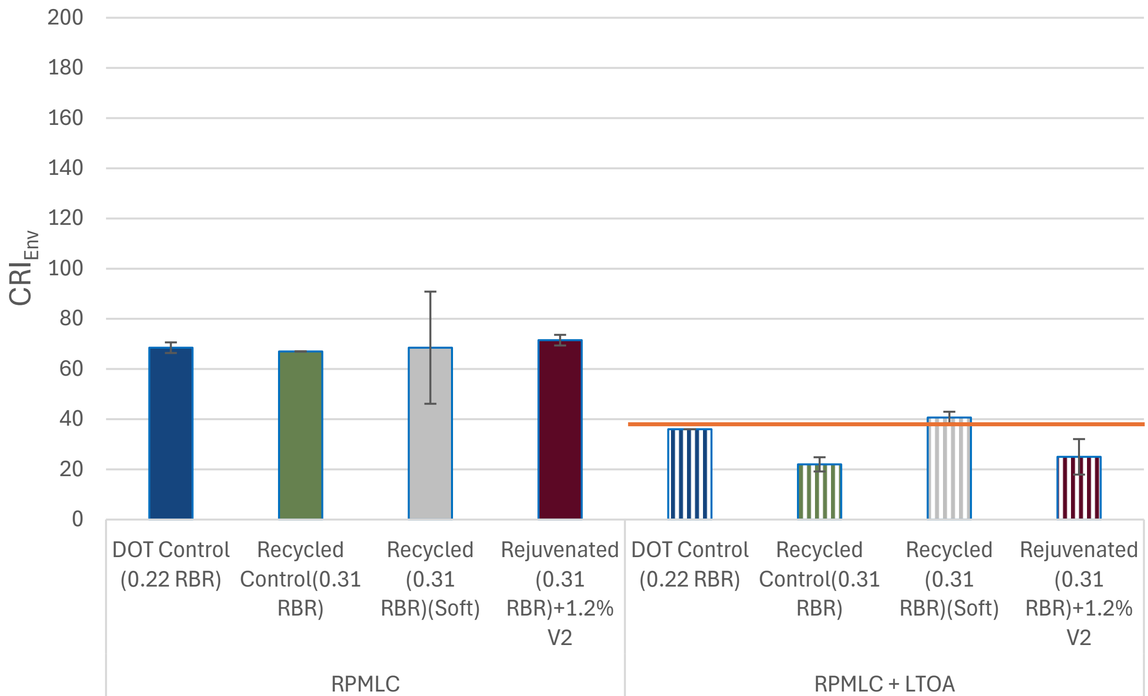


FIGURE 29 WI CRI_{ENV} RESULTS FOR RPMLC SPECIMENS

3.3.2 Field (Current) versus Aged Core Results

Since the laboratory test methods require specified specimen dimensions and air void levels (7%), thresholds established in AASHTO PP 127 are not applicable for cores. Figure 30 and Figure 31 present the average FI and CT_{Index} results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD or Games-Howell at a 0.05 significance level. Letters in uppercase were classified by Tukey's HSD, and letters in lower case were classified by Games-Howell. For both intermediate-temperature cracking resistance parameters, the recycled core with a softer binder exhibited improved performance at construction and with long-term field aging in service, although this is not reflected in the field performance rankings which may be due to differing air voids.

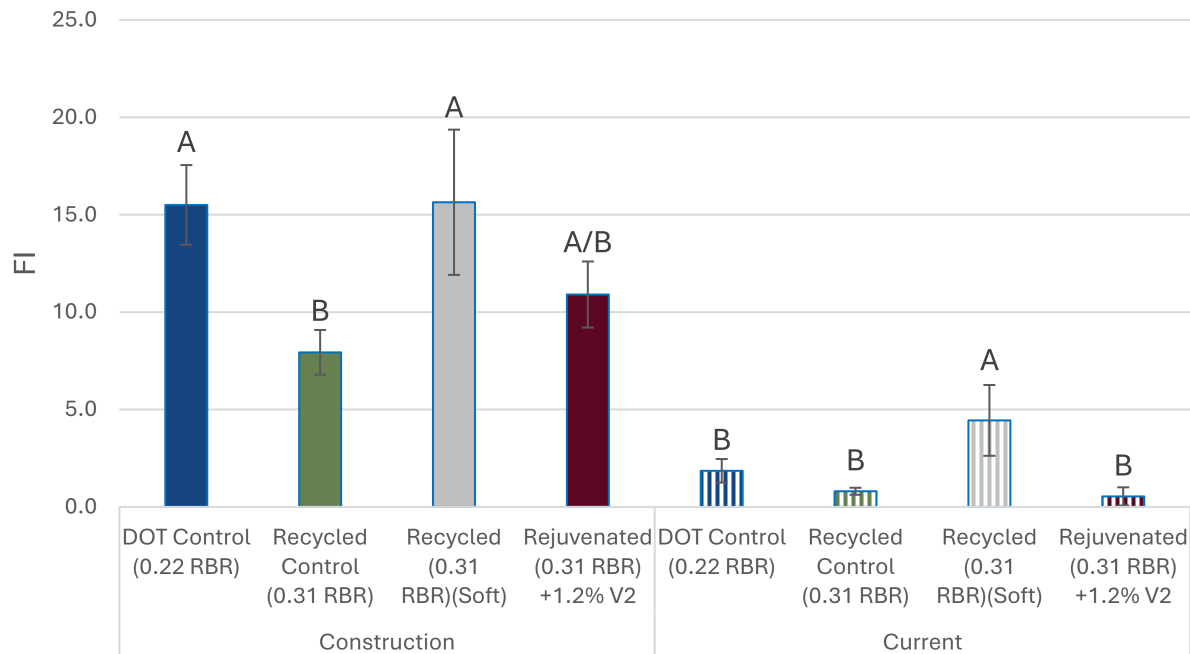


FIGURE 30 WI FI RESULTS FOR AGED FIELD CORES

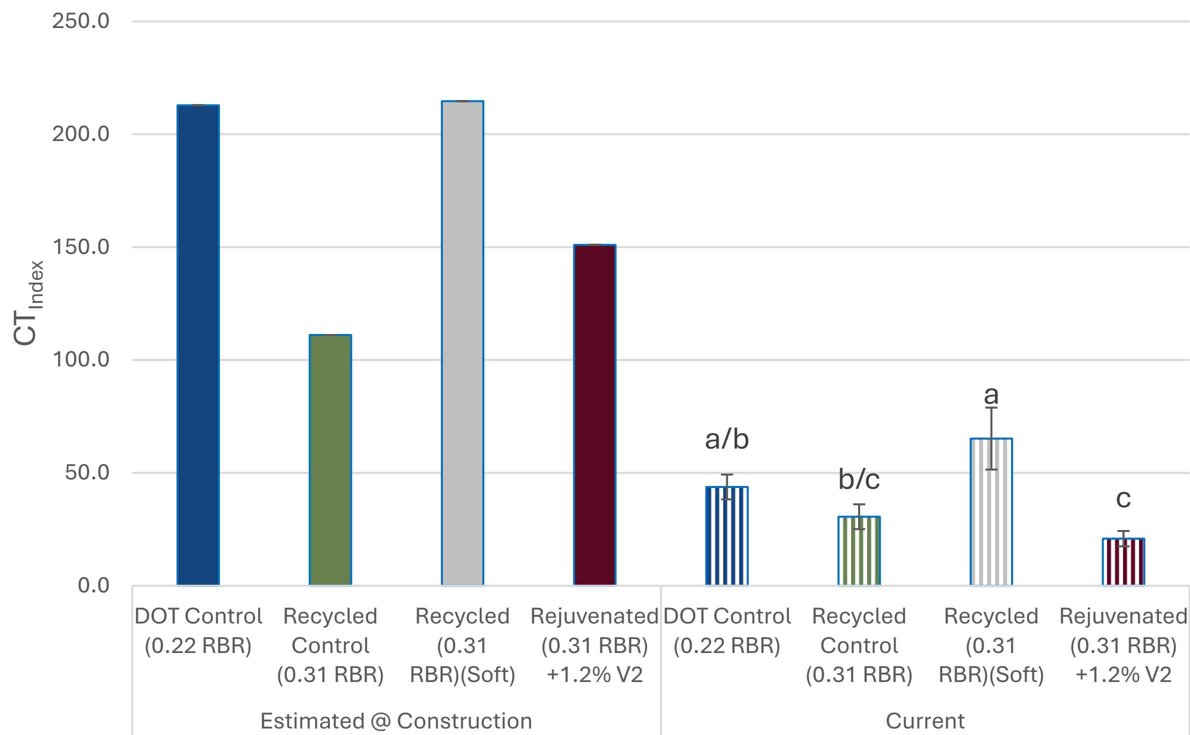


FIGURE 31 WI CT_{INDEX} RESULTS FOR AGED FIELD CORES

Figure 32 and Figure 33 present the average LCI and CRI_{Env} results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. The CRI_{Env} results again showed the recycled core outperforming the other cores, although this result is not reflected in the field performance ranking. LCI core results show no statistical difference between sections.

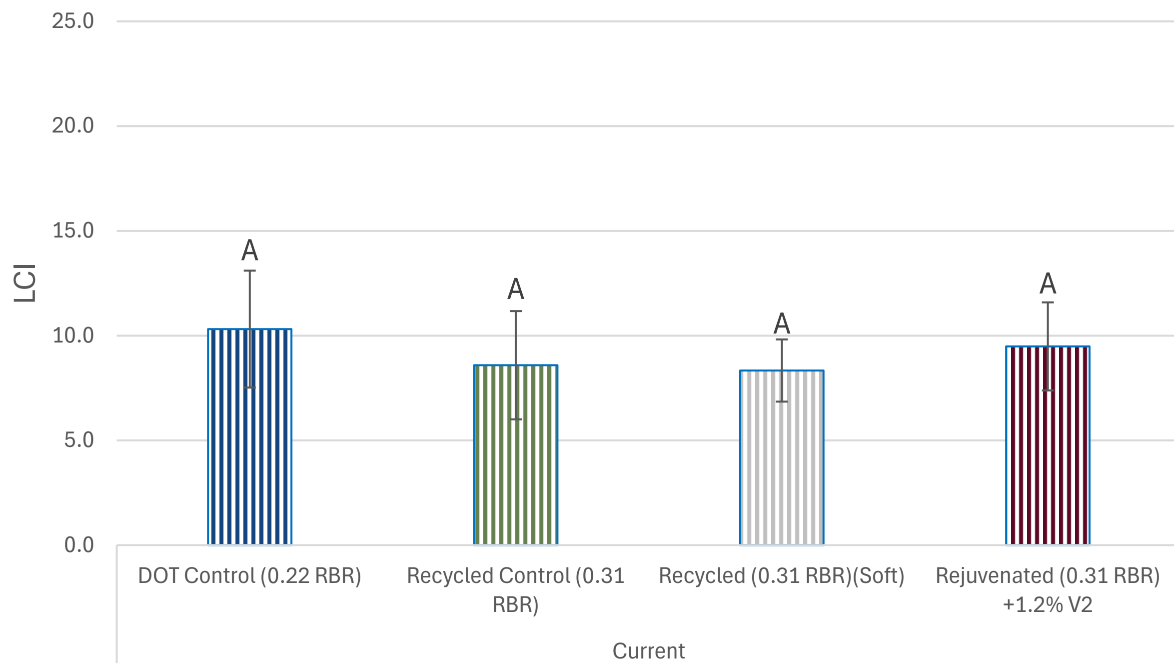


FIGURE 32 WI LCI RESULTS FOR AGED FIELD CORES

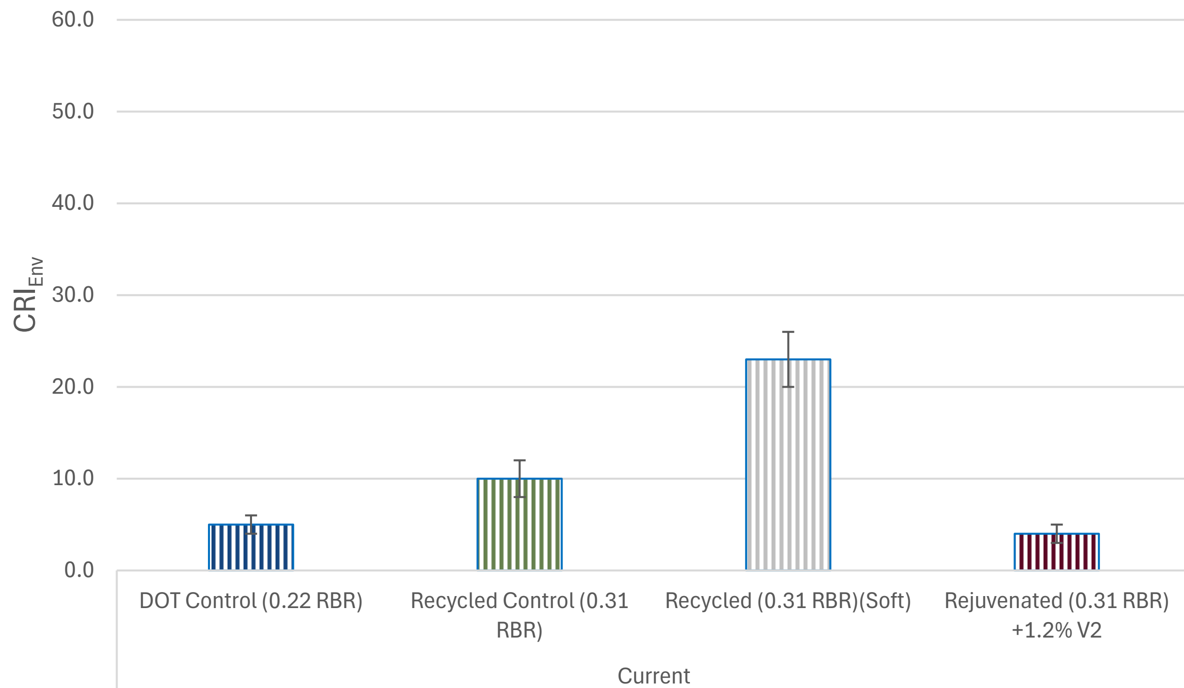


FIGURE 33 WI CRI_{ENV} RESULTS ADJUSTED TO -22C FOR AGED FIELD CORES

3.4 Texas

Figure 34 with a maximum y-axis of 13% lane cracking shows that the TX field sections were treated with a chip seal at approximately 101,000 CDD. Before this treatment, the rejuvenated section exhibited a high rate of cracking. At 70,000 CDD, the performance of the virgin section was categorized as adequate, the DOT control section as fair, and the rejuvenated section as marginal. At the time of the field evaluation in December 2024 at 10.5 years in service and after the chip seal was placed, the performance of the virgin section remained adequate, and the DOT control and rejuvenated sections were both categorized as fair.

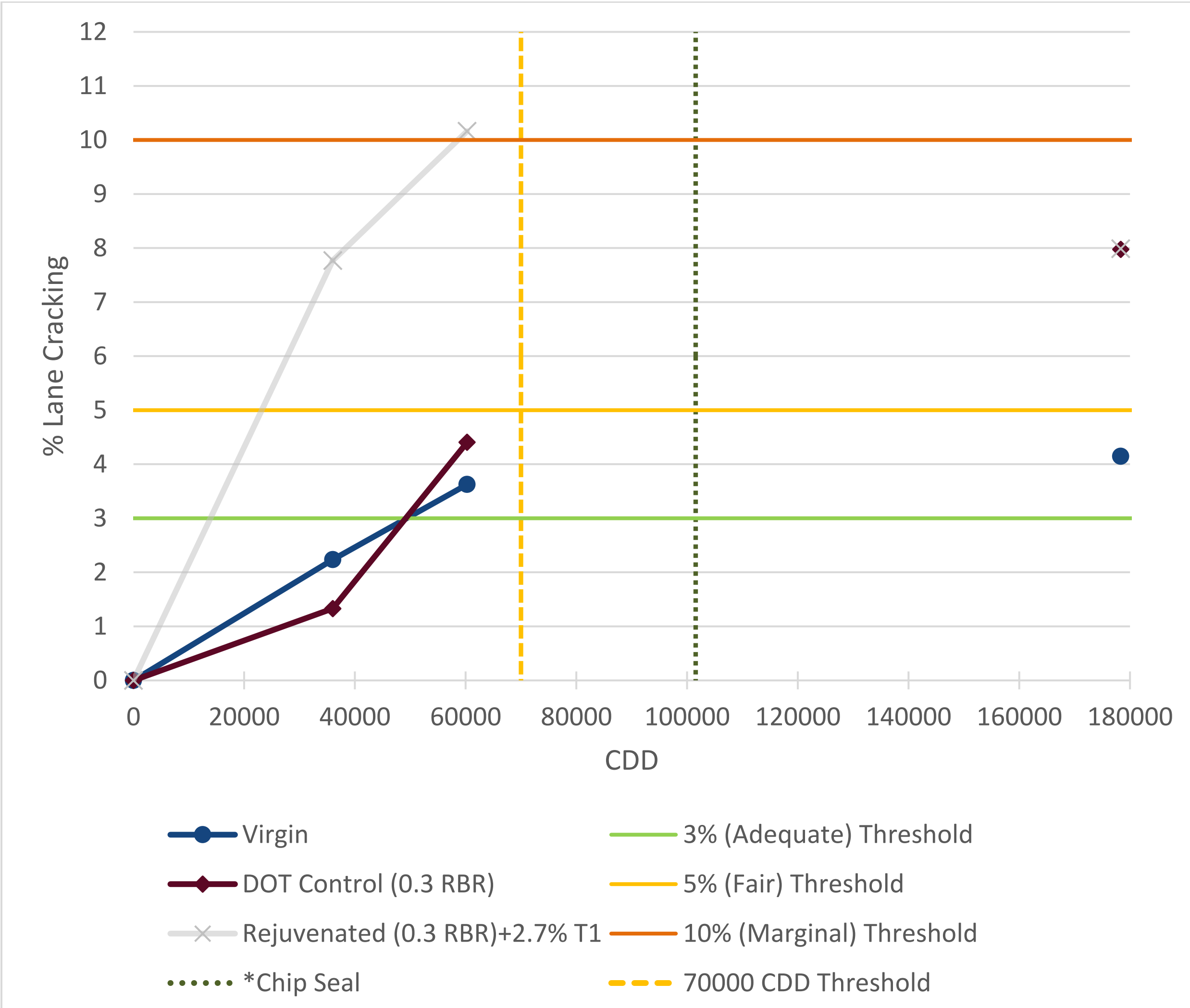


FIGURE 34 TX FIELD PERFORMANCE

3.4.1 Field (70,000 CDD) versus Laboratory Comparison

The data available from NCHRP Project 09-58 is shown in Table 15 in green and with a check mark. The unavailability of FI data for RPMLC specimens is likely due to embrittlement of the high RBR with high RAS mixtures.

TABLE 15 TX LABORATORY DATA AVAILABILITY

State	Section	FI w/ LMLC		FI w/ RPMLC		CRI_{Env} w/ LMLC		CRI_{Env} w/ RPMLC	
		STOA	LTOA	STOA	LTOA	STOA	LTOA	STOA	LTOA
Texas	Virgin	✓	✓	✓	-	-	-	✓	✓
	DOT Control (0.3 RBR) (Soft)	✓	✓	✓	-	-	-	✓	✓
	Rejuvenated (0.3 RBR) (Soft) +2.7% T1	✓	✓	-	-	-	-	✓	✓

Figure 35 and Figure 36 present the average FI results for LMLC and RPMLC specimens at available aging conditions, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. The current threshold in AASHTO PP 127 for FI of LMLC specimens after STOA of 7 is also provided. From the results for the LMLC specimens, only the virgin section met this threshold and exhibited statistically better cracking performance after STOA which is also reflected for just the virgin section in the estimated field performance ranking at 70,000 CDD.

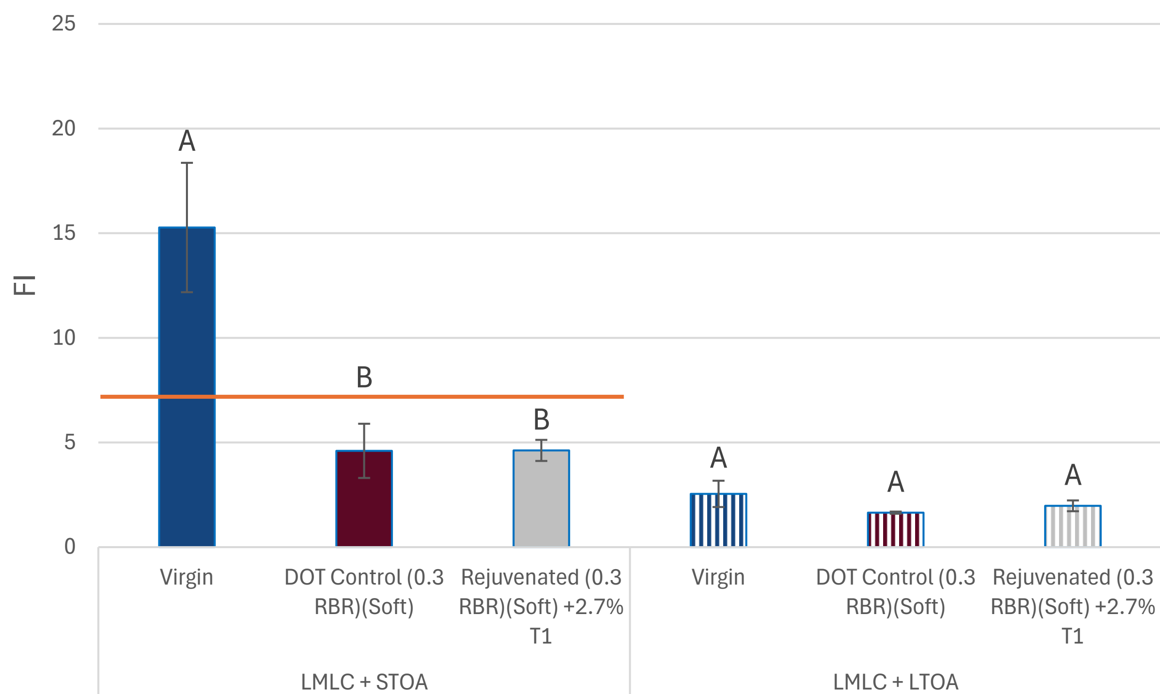


FIGURE 35 TX FI RESULTS FOR LMLC SPECIMENS

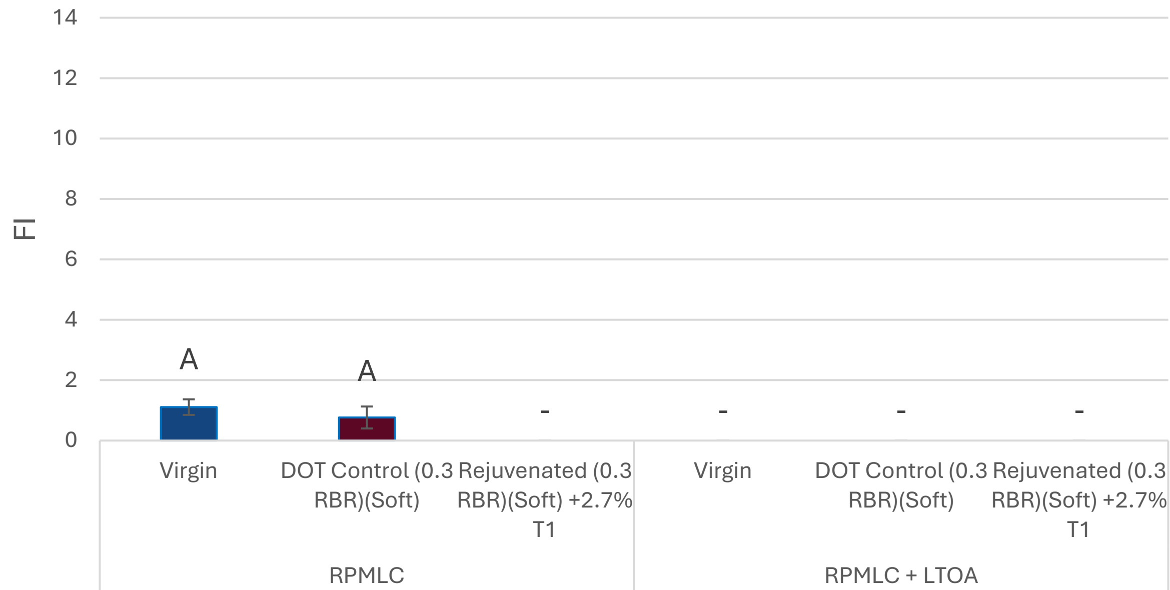


FIGURE 36 TX FI RESULTS FOR RPMLC SPECIMENS

Figure 37 presents the average CRI_{Env} results for RPMLC specimens at available aging conditions with error bars indicating plus or minus one standard deviation. The current threshold for CRI_{Env} of RPMLC specimens after LTOA of 38 (NCHRP 09-58) is also provided. None of the sections met this threshold, but these results were reflected in the field performance rankings.

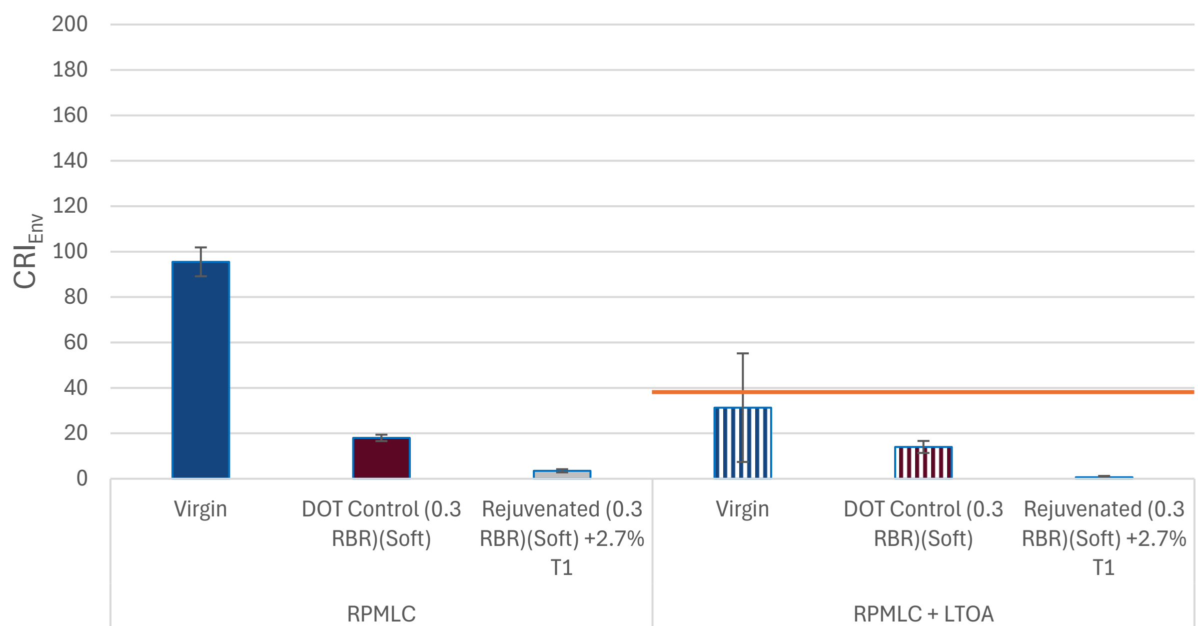


FIGURE 37 TX CRI_{ENV} WITH RPMLC

3.4.2 Field (Current) versus Aged Core Results

Because the laboratory test methods require specified specimen dimensions and air void levels (7%), thresholds established in AASHTO PP 127 are not applicable for cores. Figure 38 and Figure 39 present the average FI and CT_{Index} results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD or Games-Howell at a 0.05 significance level. Letters in uppercase were classified by Tukey's HSD, and letters in lower case were classified by Games-Howell. For both intermediate-temperature cracking resistance parameters, all cores exhibited poor performance with the virgin core exhibiting better results but not statistically improved compared to the DOT control and rejuvenated cores at construction but statistically different (improved for CT_{Index} but not for FI) with long-term field aging in service. These results are not in agreement with field performance rankings.

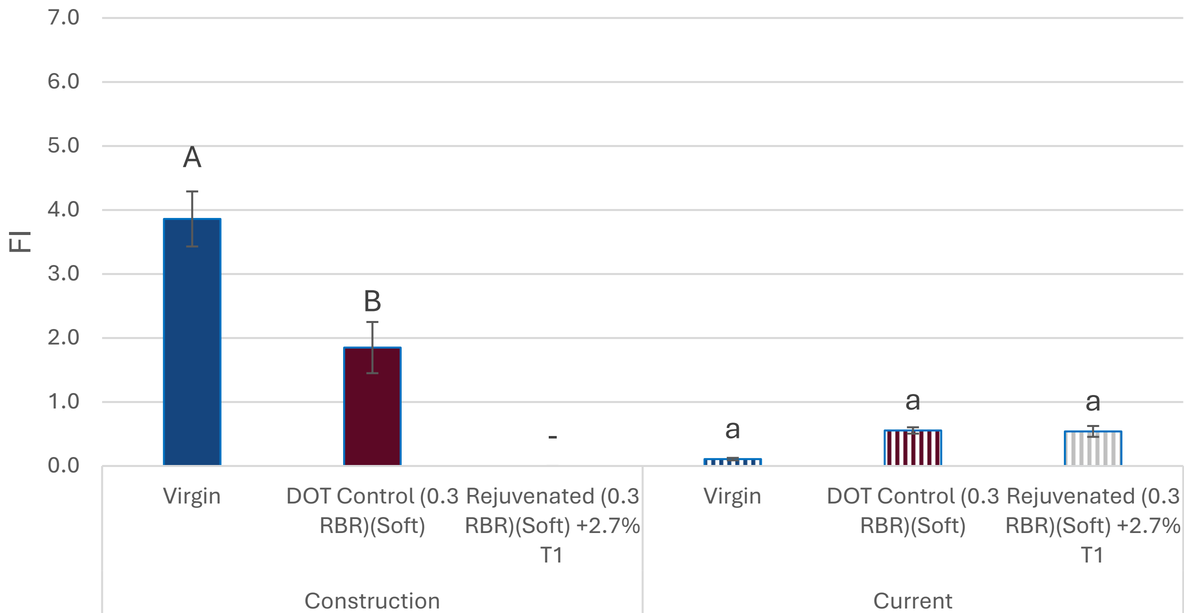


FIGURE 38 TX FI RESULTS FOR AGED FIELD CORES

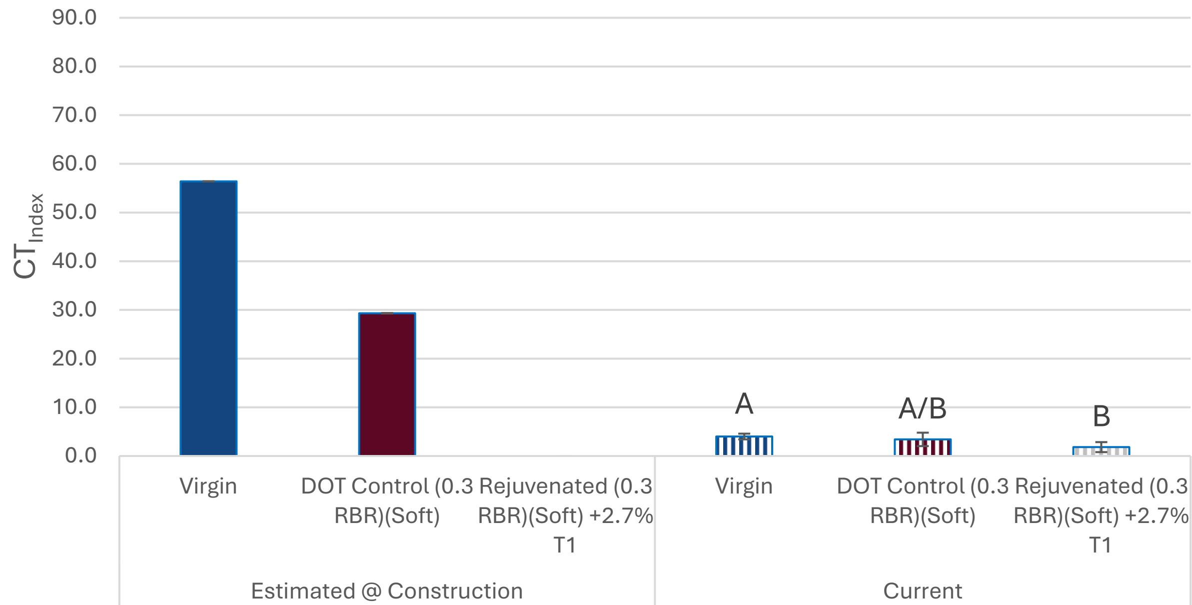


FIGURE 39 TX CT$_{INDEX}$ RESULTS FOR AGED FIELD CORES

Figure 40 and Figure 41 present the average LCI and CRI$_{Env}$ results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. Only one CRI$_{Env}$ result was available since the DOT control and rejuvenated cores were too brittle to be tested. LCI core results show no statistical difference between sections.

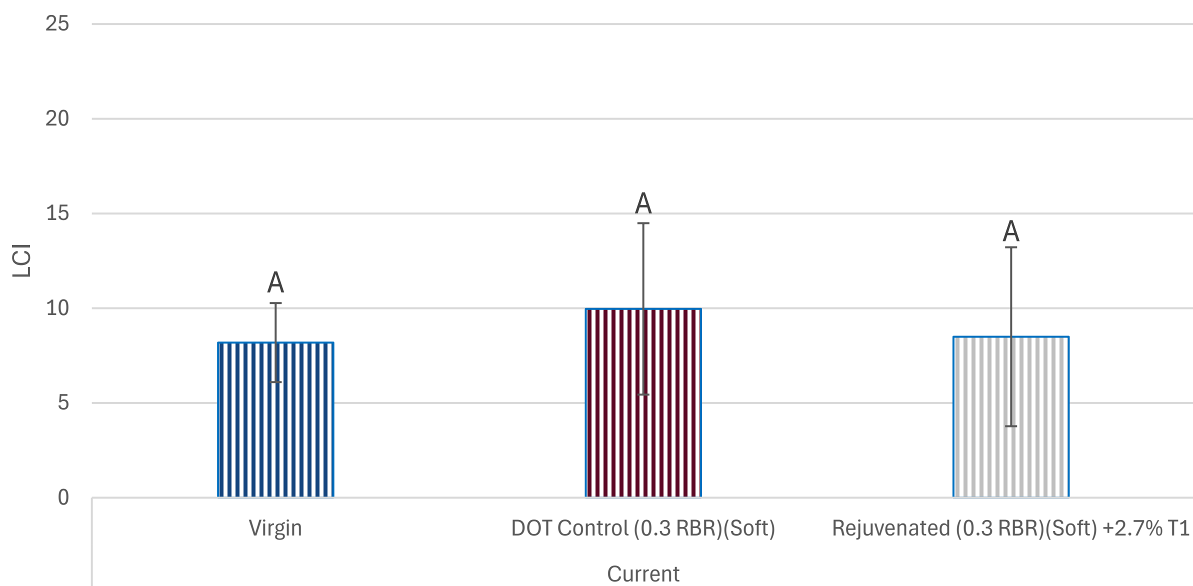


FIGURE 40 TX LCI RESULTS FOR AGED FIELD CORES

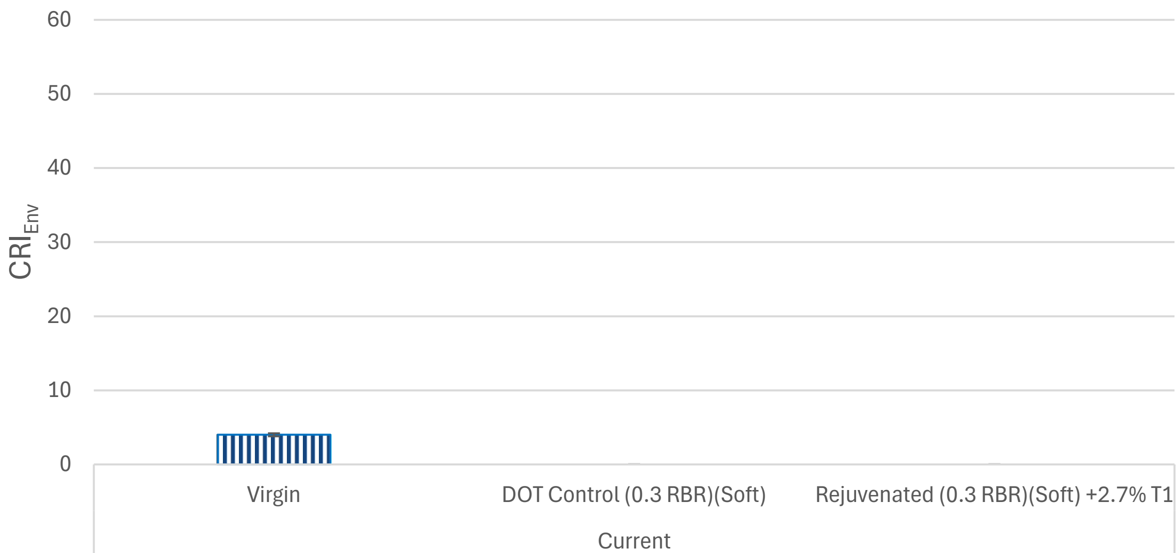


FIGURE 41 TX CRI_{ENV} RESULTS ADJUSTED TO -22C FOR AGED FIELD CORES

3.5 Delaware

Figure 42 with a maximum y-axis of 40% lane cracking shows that the DE field sections exhibited high rates of cracking and poor field performance. At 70,000 CDD, the performance of the DOT control and rejuvenated (WMA) sections was categorized as marginal, and the rejuvenated section was classified as poor. At the time of the field evaluation in April 2025 at 8.3 years in service, the performance of the rejuvenated (WMA) section remained marginal and the DOT control and rejuvenated sections were both categorized as poor.

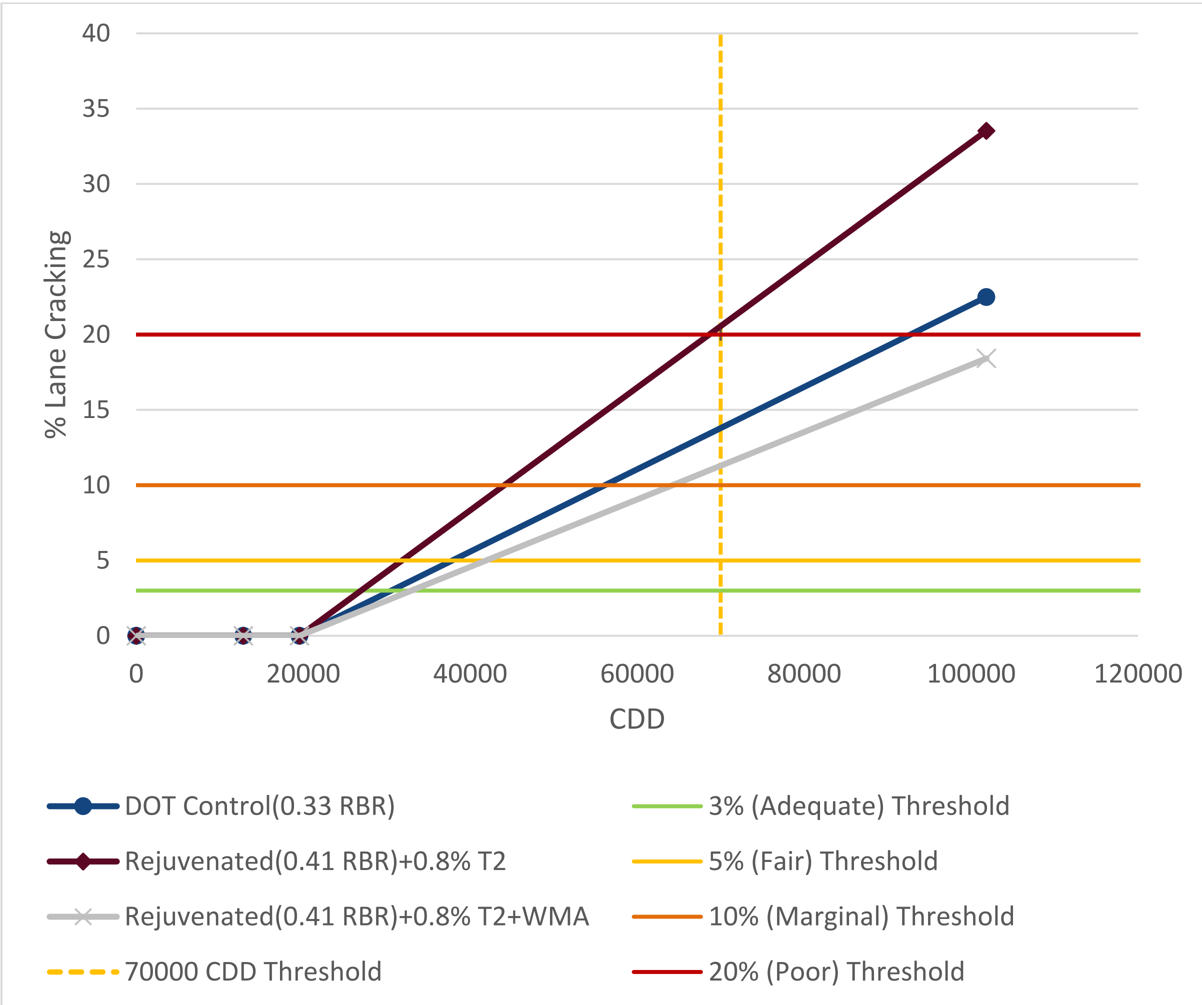


FIGURE 42 DE FIELD PERFORMANCE

3.5.1 Field (70,000 CDD) versus Laboratory Comparison

The relatively smaller set of data available from NCHRP Project 09-58 is shown in Table 16 in green and with a check mark. The DE field sections do not have as much data available as the sections were constructed later than the other field sections.

TABLE 16 DE LABORATORY DATA AVAILABILITY

State	Section	FI w/ LMLC		FI w/ RPMLC		CRI_{Env} w/ LMLC		CRI_{Env} w/ RPMLC	
		STOA	LTOA	STOA	LTOA	STOA	LTOA	STOA	LTOA
Delaware	DOT Control	✓	✓	-	-	-	-	-	-
	Rejuvenated +0.8% T2	✓	✓	-	-	-	-	-	-
	Rejuvenated +0.8% T2 +0.25% WMA	✓	✓	-	-	-	-	-	-

Figure 43 presents the average FI results for LMLC specimens at available aging conditions with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. The current threshold in AASHTO PP 127 for FI of LMLC specimens after STOA of 7 is also provided. None of the sections met this threshold, and all sections exhibited statistically equivalent performance which is reflected in the consistently poor field performance ranking at 70,000 CDD

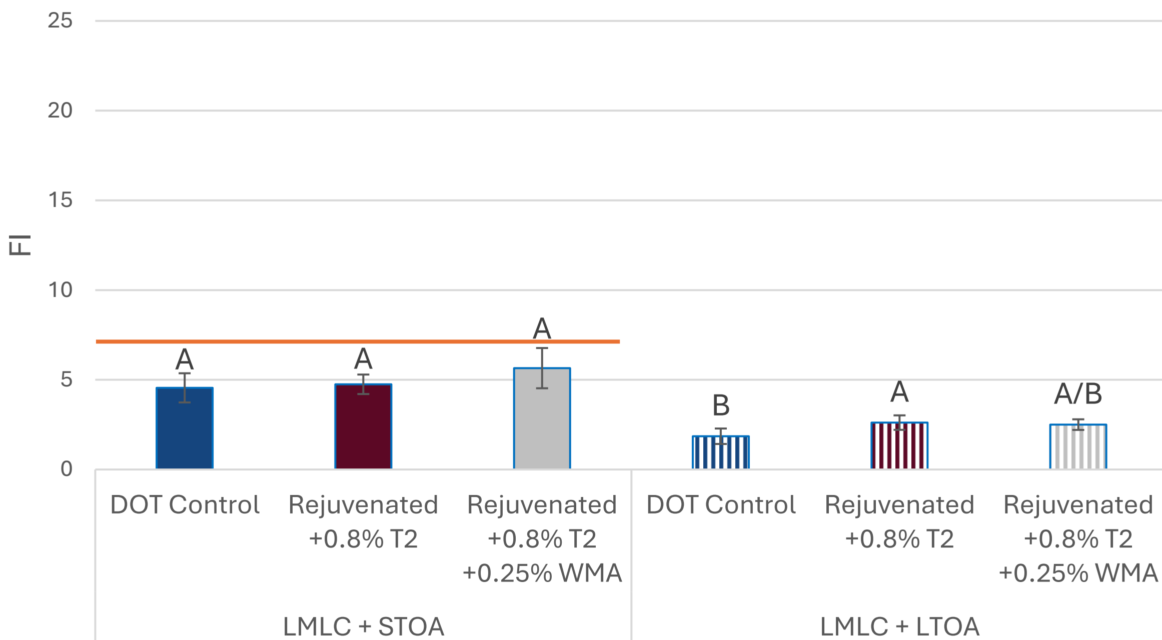


FIGURE 43 DE FI RESULTS FOR LMLC SPECIMENS

3.5.2 Field (Current) versus Aged Core Results

Because the laboratory test methods require specified specimen dimensions and air void levels (7%), thresholds established in AASHTO PP 127 are not applicable for cores.

Figure 44 and Figure 45 present the average FI and CT_{Index} results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. For FI results, all cores exhibited statistically equivalent performance. For the CT_{Index} results, unexpectedly high values were measured for the rejuvenated cores that do not reflect the poor field performance.

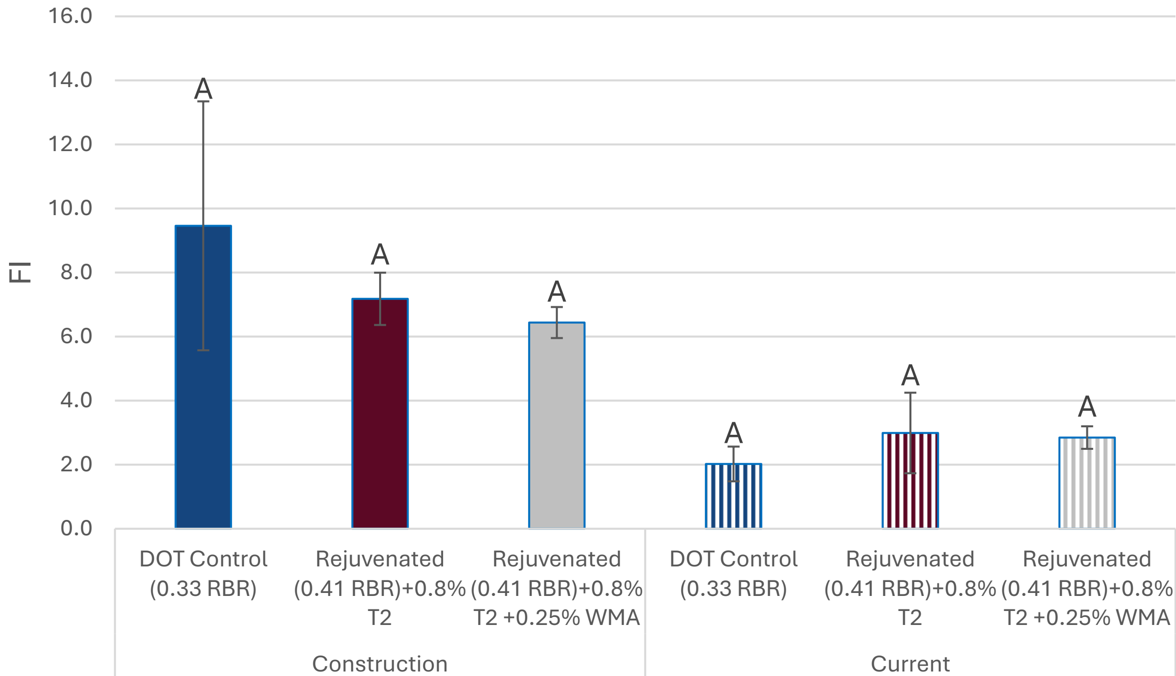


FIGURE 44 DE FI RESULTS FOR AGED FIELD CORES

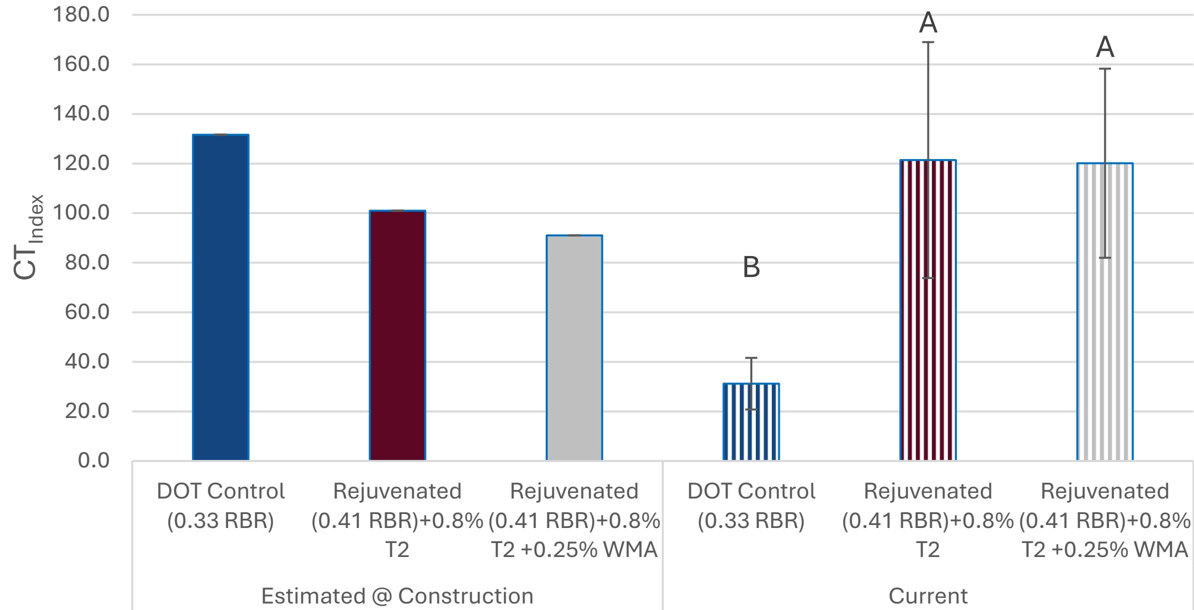


FIGURE 45 DE CT_{INDEX} RESULTS FOR AGED FIELD CORES

Figure 46 and Figure 47 present the average LCI and CRI_{Env} results for field cores at available time periods, respectively, with error bars indicating plus or minus one standard deviation from the average value and different letters indicating statistical differences by Tukey's HSD at a 0.05 significance level. The CRIEnv results were unexpectedly high for all cores, and thus they do not reflect the poor field performance. LCI core results show no statistical difference between sections.

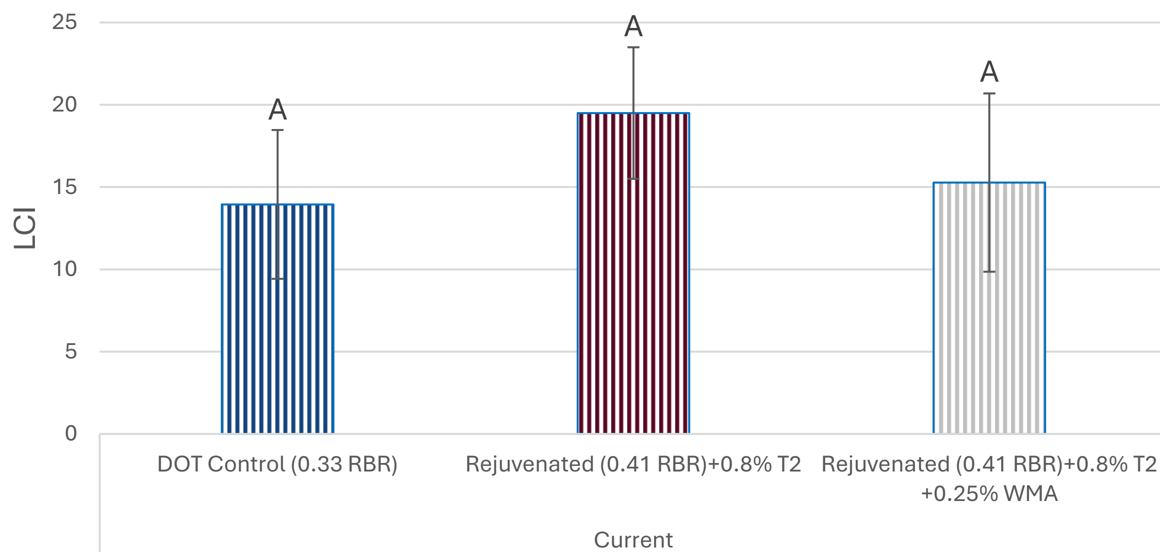


FIGURE 46 DE AGED FIELD CORE LCI RESULTS

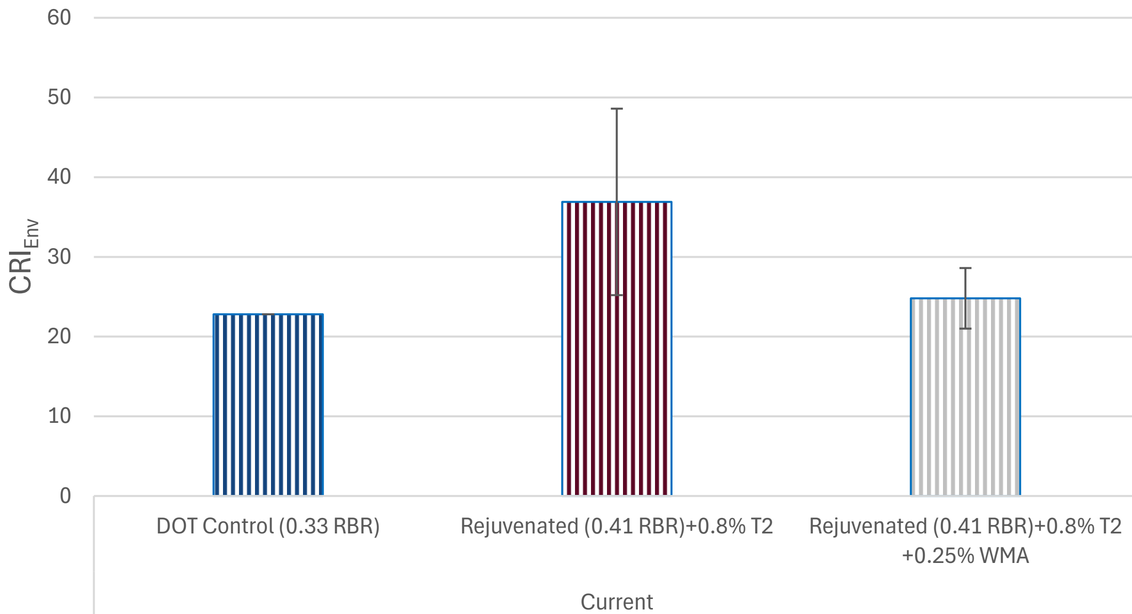


FIGURE 47 DE AGED FIELD CORE CRI$_{ENV}$(ADJ. -22°C) RESULTS

3.6 Overall Mixture Results

After field and laboratory evaluations were completed for each field project, further analysis of the results was conducted across all field projects. A confusion matrix analysis was completed for several thresholds in AASHTO PP 127 and beyond. After the confusion matrix analysis, an analysis of the laboratory aging protocols in comparison to field aging was conducted.

3.6.1 Confusion Matrix Analysis

To meet the primary objective of strengthening confidence in binder and mixture performance thresholds in AASHTO PP 127 and beyond and proposing revisions as needed, comparisons of laboratory results and field performance were analyzed to evaluate their correlation. As presented previously, correlated laboratory and field results either passed laboratory thresholds and exhibited less than 10% lane cracking (Pass-Pass) or failed laboratory thresholds and exhibited marginal or poor performance with more than 10% lane cracking (Fail-Fail). Uncorrelated results occurred when one either laboratory results passed the threshold with marginal or poor field performance or results failed the threshold with good, adequate, or fair field performance. After creation of the confusion matrix with these results, laboratory tests with low percentage correlation were identified as candidates for threshold adjustments to better align laboratory results with field performance. In addition, a laboratory pass paired with a field fail (Pass-Fail) was considered a more significant error than a laboratory fail paired with a field pass (Fail-Pass) that is more conservative. This issue was considered when adjusting performance thresholds. The results for the confusion matrix analysis are provided in Table 17. The confusion matrix analysis was not conducted on the following tests contained in AASHTO PP 127 (Table

8) as these were not included in the scope of this project: tensile strength ratio (TSR), Glover-Rowe for the mixture ($G\text{-}R_m$), fracture energy (G_f), flexural creep stiffness for the mixture (S_m), and rate of relaxation ($m\text{-value}_m$) for the mixture. The confusion matrix analysis was also not conducted for CRI_{Env} with LMLC and CT_{Index} due to lack of failure data and lack of data from NCHRP 09-58, respectively.

TABLE 17 CONFUSION MATRIX RESULTS

		Correlated		Uncorrelated	
Parameter	Threshold	TP (P-P)	TN (F-F)	FP (P-F)	FN (F-P)
PC Ductility	150%	6	6	0	6
FI w/ LMLC @ STOA	7	7	5	0	4
FI w/ LMLC @ LTOA	3	7	5	0	4
CRI_{Env} w/ RPMLC	38	5	2	0	8
CRI_{Env} w/ RPMLC	26	11	2	0	2

Figure 48 presents the confusion matrix for FI results for LMLC specimens after STOA. The FI threshold of 7 after STOA and a field performance threshold of 10% after 70,000 CDD correlated 75% of the time. This strong correlation provided confidence in this threshold for this aging condition, and revision was not proposed. The outside ring of the confusion chart is labeled "Corr" for results that correlate and "Uncorr" for results that do not correlate. The inner ring of the chart breaks down the confusion matrix results into true positive (TP), true negative (TN), false positive (FP), and false negative (FN) as defined previously.

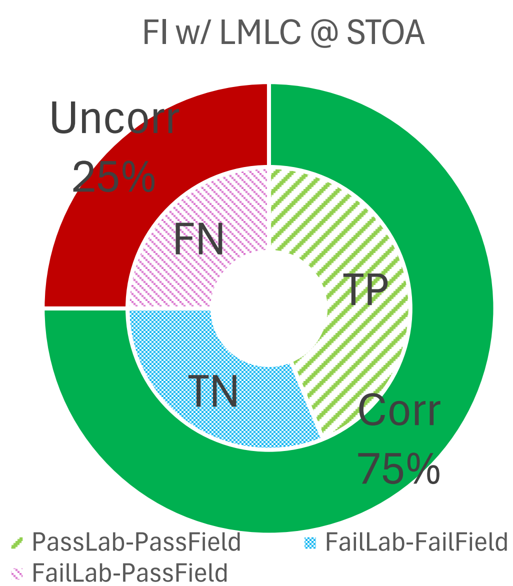


FIGURE 48 FI LMLC W/ STOA AND FIELD COMPARISON

Figure 49 presents the confusion matrix for FI results for LMLC specimens after LTOA with a proposed threshold of 3 with a strong correlation of 75%.

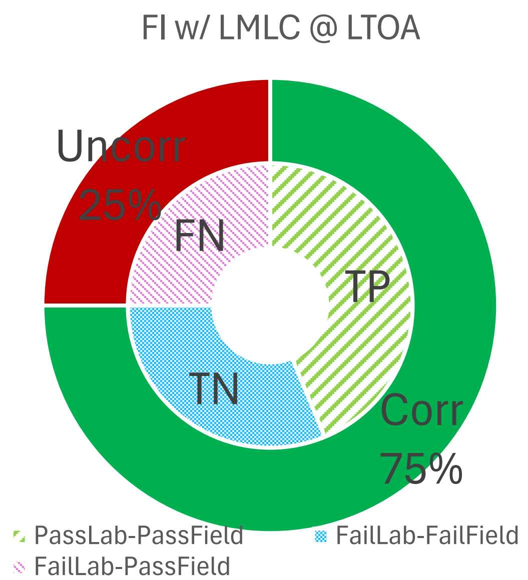


FIGURE 49 FI LMLC W/ LTOA PROPOSED THRESHOLD AND FIELD COMPARISON

Figure 50 presents the confusion matrix for CRI_{Env} results for RPMLC specimens after LTOA. The CRI_{Env} threshold of 38 for RPMLC specimens after LTOA and a field performance threshold of 10% after 70,000 CDD correlated 47% of the time. This is a weak correlation, so a new threshold of 26 was proposed. Figure 51 shows the revised confusion matrix with the new CRI_{Env} with RPMLC after LTOA threshold of 26 with correlation improving significantly to 87%.

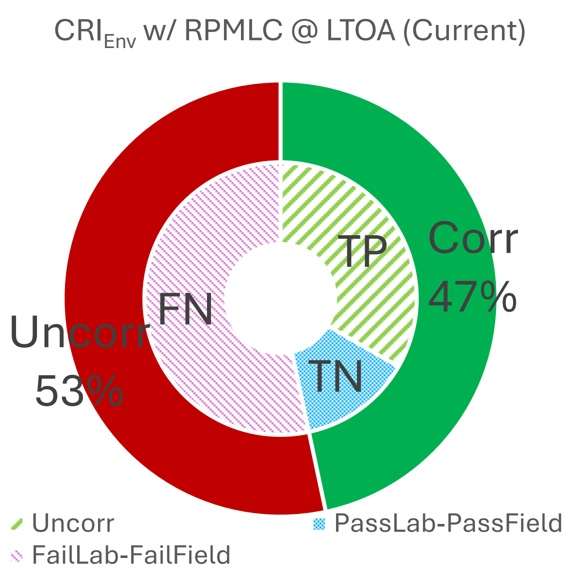


FIGURE 50 CRI_{ENV} RPMLC W/ LTOA THRESHOLD AND FIELD COMPARISON

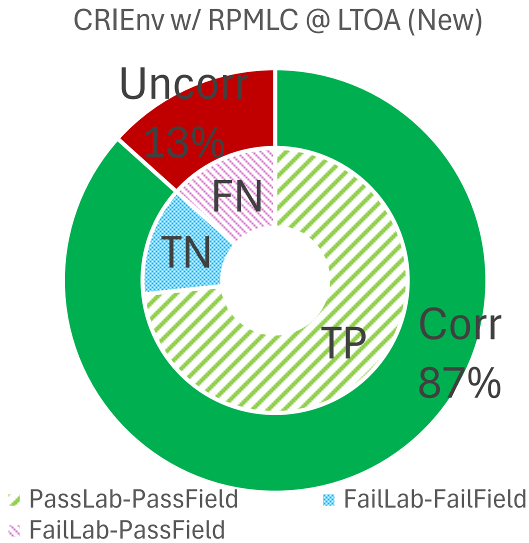


FIGURE 51 CRI$_{ENV}$ RPMLC W/ LTOA PROPOSED THRESHOLD AND FIELD COMPARISON

Further analysis of the confusion matrix results is presented in Table 18 that reflects the improved accuracy of the new proposed threshold for CRI$_{Env}$ with RPMLC after LTOA.

TABLE 18 CONFUSION MATRIX ANALYSIS

Parameter	Threshold	Accuracy	Precision	Sensitivity	Specificity	Balanced Acc.
PC Ductility	150%	0.667	1.000	0.500	1.000	0.750
FI w/ LMLC @ STOA	7	0.750	1.000	0.636	1.000	0.818
FI w/ LMLC @ LTOA	3	0.750	1.000	0.636	1.000	0.818
CRI$_{Env}$ w/ RPMLC @ LTOA	38	0.467	1.000	0.385	1.000	0.692
CRI$_{Env}$ w/ RPMLC @ LTOA	26	0.867	1.000	0.846	1.000	0.923

3.6.2 Cracking Resistance Ratio Analysis

The cracking resistance retention ratios for the laboratory mixtures and the field cores are shown in Figure 52. Data were not available for the DE field project for RPMLC specimens and for the TX field project for RPMLC specimens after LTOA. The CT$_{Index}$ cracking resistance retention ratio for the DE field project rejuvenated sections was also not included because of the unusually high CT$_{Index}$ for those sections that produced a negative ratio.

Based on Figure 52, the LMLC and RPMLC cracking resistance retention ratios were consistently higher than those of the field cores. Thus, the laboratory aging protocols did not sufficiently capture the performance retention of the field cores after 8-10 years. The aged laboratory specimens from the DE field project exhibited performance retention much closer to that of the field cores. This suggests that aging of compacted specimens for 5 days at 85˚C in the laboratory does not sufficiently age the asphalt mixture to simulate 8-10 years of aging in the field.

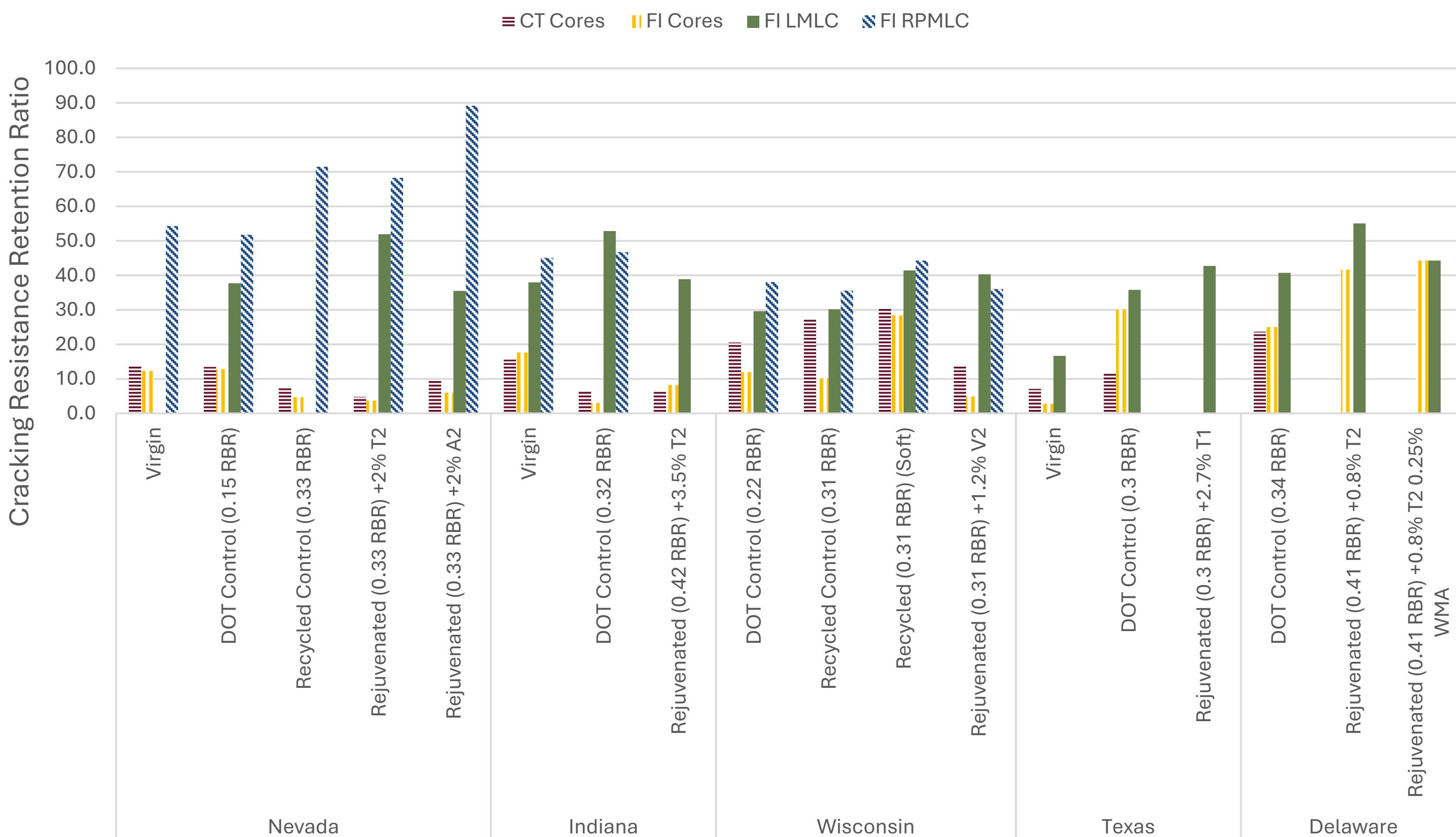
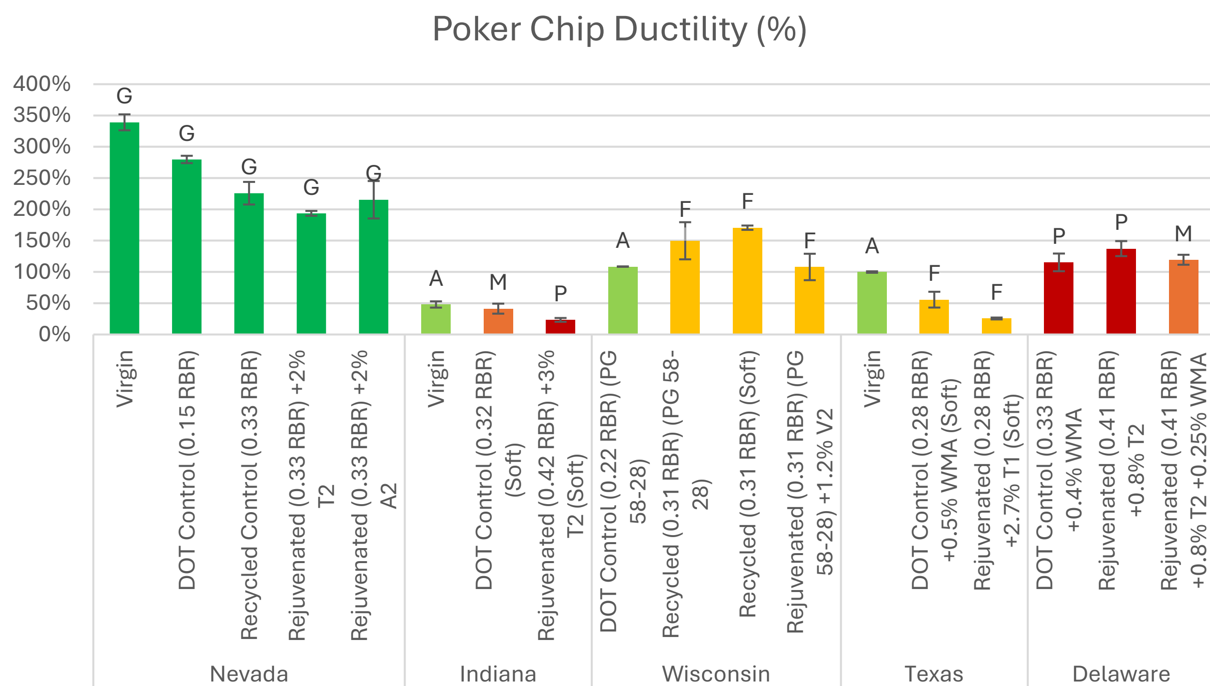


FIGURE 52 CRACKING RESISTANCE RETENTION RATIO

3.7 Overall Binder Results

Figure 53 presents the PCD results for the extracted and recovered binders along with the threshold of 150%. These results capture most of the sections with respect to their current field performance highlighted by color coded categories.



Notes: Dark green = Good (<3% Lane Cracking); light green = Adequate (<5% Lane Cracking); Yellow = Fair (<10% Lane Cracking); Orange = Marginal (<20% Lane Cracking); Red = Poor (≥20% Lane Cracking)

FIGURE 53 PCD RESULTS OF BINDERS RECOVERED FROM AGED FIELD CORES

Figure 54 provides the confusion matrix for PCD results of extracted and recovered binders from cores. The PCD threshold of 150% and a field performance threshold of 10% cracking correlated approximately 67% of the time which was considered adequate.

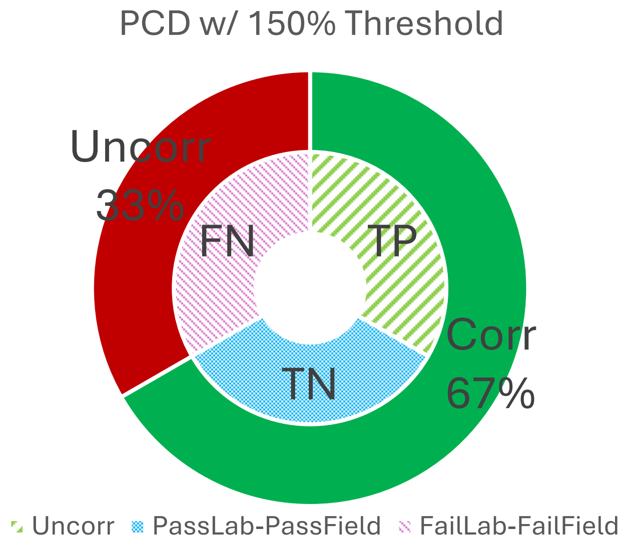


FIGURE 54 PCD THRESHOLD AND FIELD COMPARISON

Figure 55, Figure 56, Figure 57, Figure 58, and Figure 59 provide the G-R parameters for the binder extracted and recovered at the current field performance monitoring session. The thresholds of 180kPa and 600kPa are also provided for all binders except for those from the NV field sections since they are not applicable for PMA. The NV sections exhibited similar G-R values in Black space that reflected the similar field performance. The G-R values for the IN sections also reflect the separated field performance with cracking expected for all but the virgin section. The G-R values for the WI sections also reflected current field performance with respect to the lowest value for the recycled section without any high RBR cracking mitigation strategy. While the thresholds are not applicable for the TX virgin section with a PMA, the relative G-R results reflected the estimated field performance prior to placement of the chip seal. The DE sections exhibited similar G-R values, but their position in Black space below both thresholds did not reflect the poor field performance.

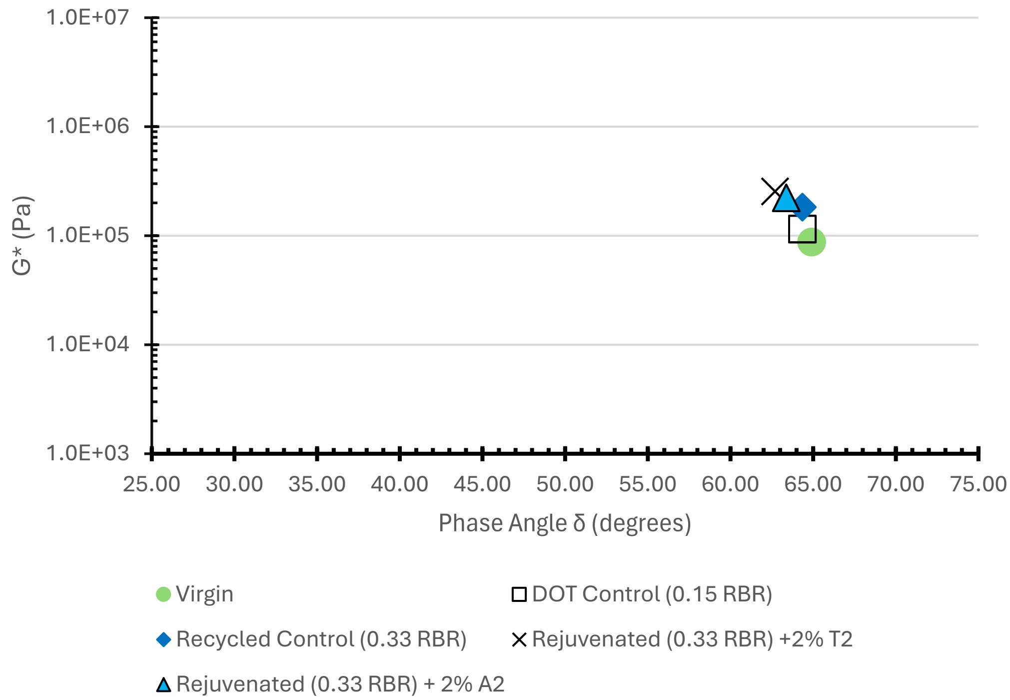


FIGURE 55 NV G-R RESULTS FOR BINDERS RECOVERED FROM AGED FIELD CORES

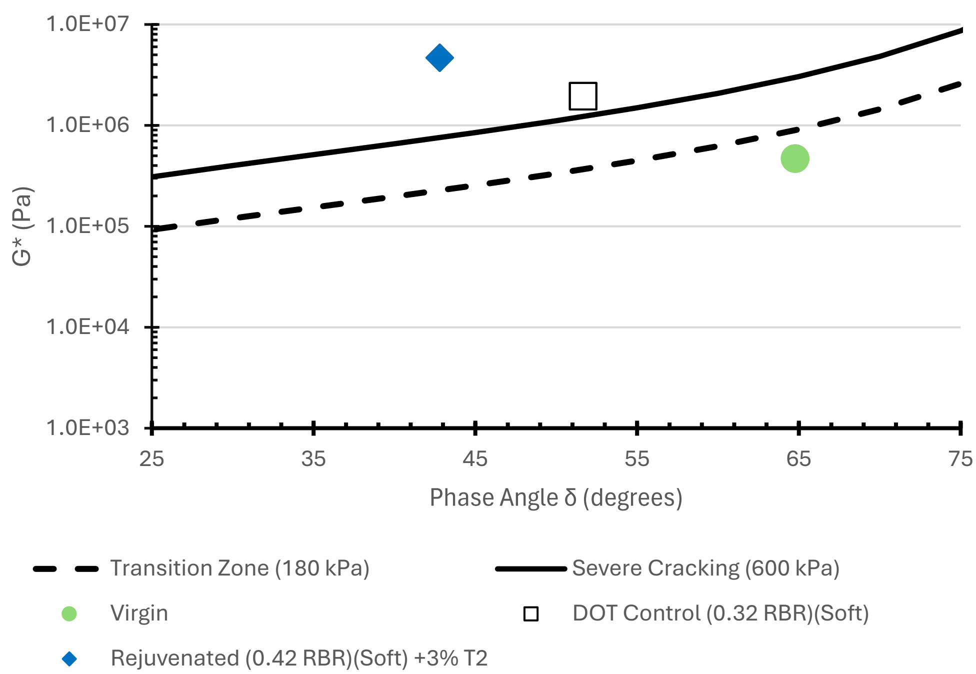


FIGURE 56 IN G-R RESULTS FOR BINDERS RECOVERED FROM AGED FIELD CORES

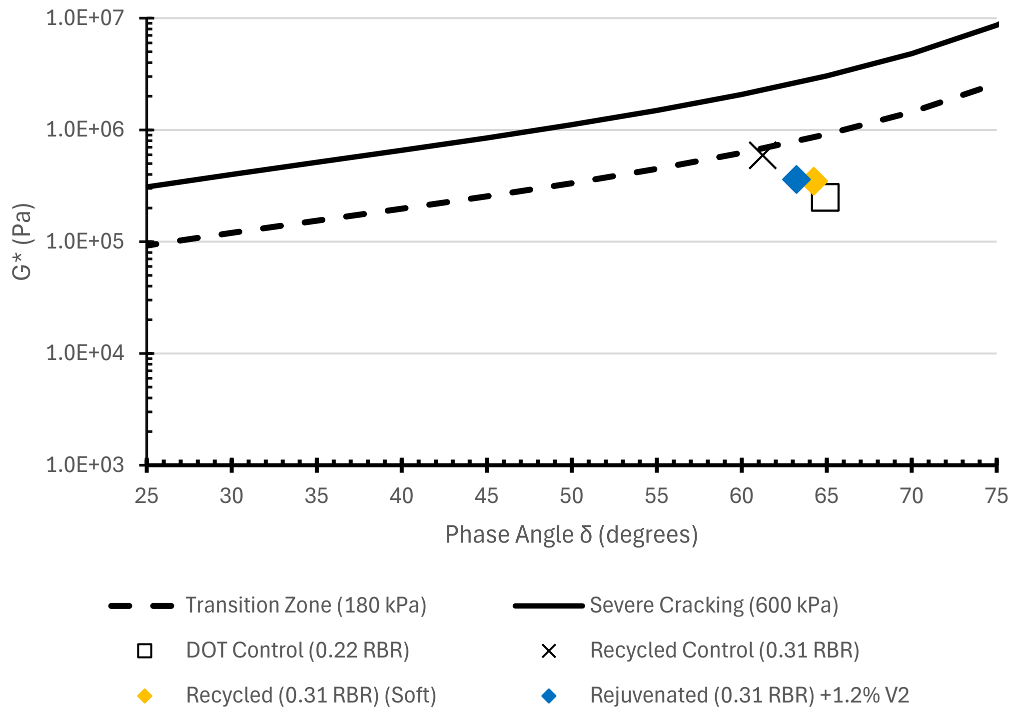


FIGURE 57 WI G-R RESULTS FOR BINDERS RECOVERED FROM AGED FIELD CORES

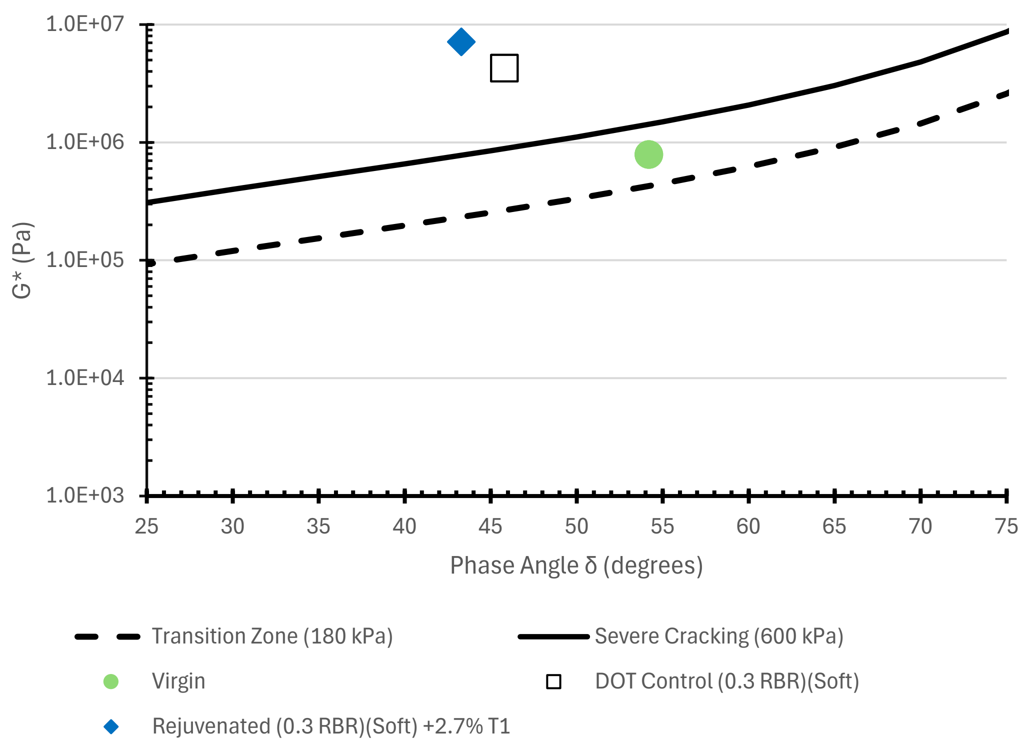


FIGURE 58 TX G-R RESULTS FOR BINDERS RECOVERED FROM AGED FIELD CORES

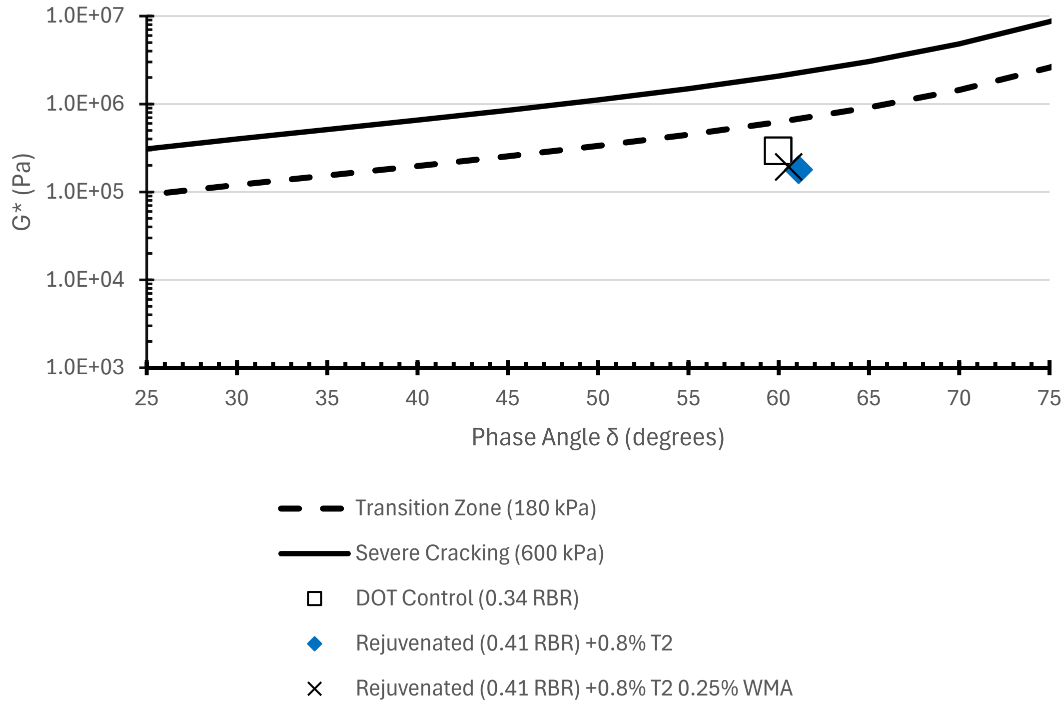


FIGURE 59 DE G-R RESULTS FOR BINDERS RECOVERED FROM AGED FIELD CORES

4.0 Conclusions

This section provides summaries of lessons learned from each field project, confusion analysis results and corresponding proposed revisions to AASHTO PP 127, and aging protocol comparison analysis. To conclude this report, overall findings and next steps are provided.

4.1 Field Project Summaries

Each of the field projects included different mixtures, climate conditions, traffic levels, and cracking type and severity. The assumption that the pavement structure was adequately designed for the traffic level was confirmed with only top-down surface cracking and reflections cracking present. Several field sections contained early generation tall oil recycling agents that are not conducive to healing and thus may have not prevented effectively cracking distress.

This project evaluated the relationship between laboratory results and field performance using two criteria: correlation and differentiation. Testing correlated with the field where laboratory results when judged with current thresholds accurately captured long-term field performance at 8-10 years in service, defined as 10% lane cracking (combined transverse and longitudinal) between fair and marginal categories. If laboratory results when judged with current thresholds did not directly correlate but still ranked mixtures the same as field performance, the laboratory testing was able to differentiate differences in performance. If there were no applicable thresholds, as for the aged field cores, but the laboratory test results generally agreed with the field performance, they were labeled as capturing the field performance.

Table 19 summarizes the key findings from the field projects, including project conditions, field performance at 70,000 CDD and at the current performance, and the main conclusions regarding laboratory-field correlation.

Overall, this summary shows clear differences in section performance across projects and highlights the effects of traffic level, climate, RAS/RAP content, and recycling agent effectiveness on cracking resistance and laboratory-field agreement.

TABLE 19 OVERALL SUMMARY OF FIELD PROJECTS

Project	Conditions / Key Features	70,000 CDD Performance	Current Performance	Main Takeaways
NV	Low-traffic local road in a cold climate; thermal cracking expected. All sections contained PMA.	All sections rated good.	All sections remained good with limited cracking after nearly 10 years.	The virgin section performed well across laboratory tests. The recycled section, despite higher RAP than the DOT control, did not meet originally established thresholds, but rejuvenated sections showed that recycling agents can be an effective high-RBR cracking mitigation strategy. Both PCD and G-R binder results aligned with field performance.
IN	Low-traffic local farm road; DOT control and rejuvenated sections contained 2% and 8% RAS, respectively.	Virgin: good; DOT control: fair; rejuvenated: poor.	Virgin: adequate; DOT control: marginal; rejuvenated: poor.	The high-RBR, high-RAS rejuvenated section showed a high cracking rate that was not mitigated by the early-generation tall oil recycling agent. Some laboratory measures reflected field rankings, including FI with STOA for LMLC specimens and CRI_{Env} for RPMLC specimens. G-R binder results also matched current performance.
WI	1.70 million ESALs and 8.2% trucks; no RAS or PMA; cold climate with expected thermal cracking.	DOT control and rejuvenated: adequate; recycled control and recycled: fair.	Only the DOT control remained adequate; the rejuvenated section worsened; both recycled sections remained fair.	Laboratory results for the recycled section generally reflected early rankings, but many tests overestimated aged field performance, particularly for the recycled section with the softer binder. G-R binder testing aligned with current field performance.
TX	Highest traffic level among field projects: 3.43 million ESALs and 14.8% trucks; virgin section contained PMA; DOT control and	Virgin: adequate; DOT control: fair; rejuvenated: marginal.	Virgin: adequate; DOT control: fair; rejuvenated: fair.	Sections containing RAS showed high cracking rates before chip sealing. Some laboratory results captured the adequate performance of the virgin section, including FI with STOA for LMLC specimens and CRI_{Env} for RPMLC specimens

Project	Conditions / Key Features	70,000 CDD Performance	Current Performance	Main Takeaways
	rejuvenated sections contained 5% RAS.			without LTOA. From field cores, CT_{Index} reflected current performance, and both PCD and G-R binder results aligned with field observations.
DE	Light traffic with 1.0% trucks; all sections contained 4% RAS and exhibited high cracking rates; rejuvenated sections used low doses of early-generation tall oil recycling agents.	DOT control and rejuvenated (WMA): marginal; rejuvenated: poor.	Only rejuvenated (WMA) remained marginal; DOT control and rejuvenated worsened to poor.	All sections showed high cracking susceptibility. FI with LMLC after both STOA and LTOA correlated with early field performance. For current performance, FI from field cores captured rankings, and PCD results for extracted binders were also consistent with field observations.

4.2 AASHTO PP 127 Revisions

Table 20 contains the possible revisions proposed for AASHTO PP 127. Thresholds that remain the same are located in the "Revised Threshold" column in bold. Thresholds that are being proposed are located in the same column and in italics. Thresholds that were not analyzed are indicated with "NA." All parameters in Table 20 are for with LMLC specimens for evaluation during mix design, and thus the proposed revision of the CRI_{Env} RPMLC threshold after LTOA is not included. However, the revision of the threshold for CRI_{Env} with LMLC after LTOA was proposed according to the same ratio as the revision to the RPMLC after LTOA threshold.

TABLE 20 AASHTO PP 127 REVISIONS

Parameter	Aging Condition	Current Threshold	Revised Threshold
$N_{12,5}$	-	≥10,000 for PG 58-XX at 45˚C ≥15,000 for PG 64-XX at 45˚C ≥20,000 for PG 70-XX at 45˚C	NA
SIP	-	≥9,000 at 45˚C	NA
SN	-	≥2,000 at 45˚C	NA
$G\text{-}R_m$	STOA	≤8,000 MPa	NA
	LTOA	≤19,000 MPa	
FI	STOA	≥7	**≥7**
FI	LTOA	-	*≥3*
TSR	-	≥80% without freeze-thaw	*NA*
TSR	-	≥80% with freeze-thaw	NA
CT_{Index} (No Freeze)	LTOA*	≥20	NA
CT_{Index} (Freeze)	LTOA**	≥30	NA
S_m and m-value$_m$	STOA	≤Romero (2016) threshold on m-value$_m$	NA
CRI_{ENV}	LTOA	≥17	≥12
G_f	STOA	≥400	NA

Notes: STOA for all specimens = 2hrs @ 135˚C, LTOA for compacted specimens = 5days @ 85˚C, Loose mix LTOA* specimens in non-freeze = 8hrs @135˚C (unmodified binders) & 5days @ 95˚C (PMA), Loose mix LTOA** specimens in freeze = 6hrs @135˚C (unmodified binders) & 3days @ 95˚C (PMA)

Based on the FI and CT correlation (Figure 4) shown in Figure 60 along with a 95% confidence interval, a CT value of 45 corresponded to the added FI threshold of 3. This provides confidence in the current threshold for CT in freeze climates after LTOA of 30 that falls within this interval.

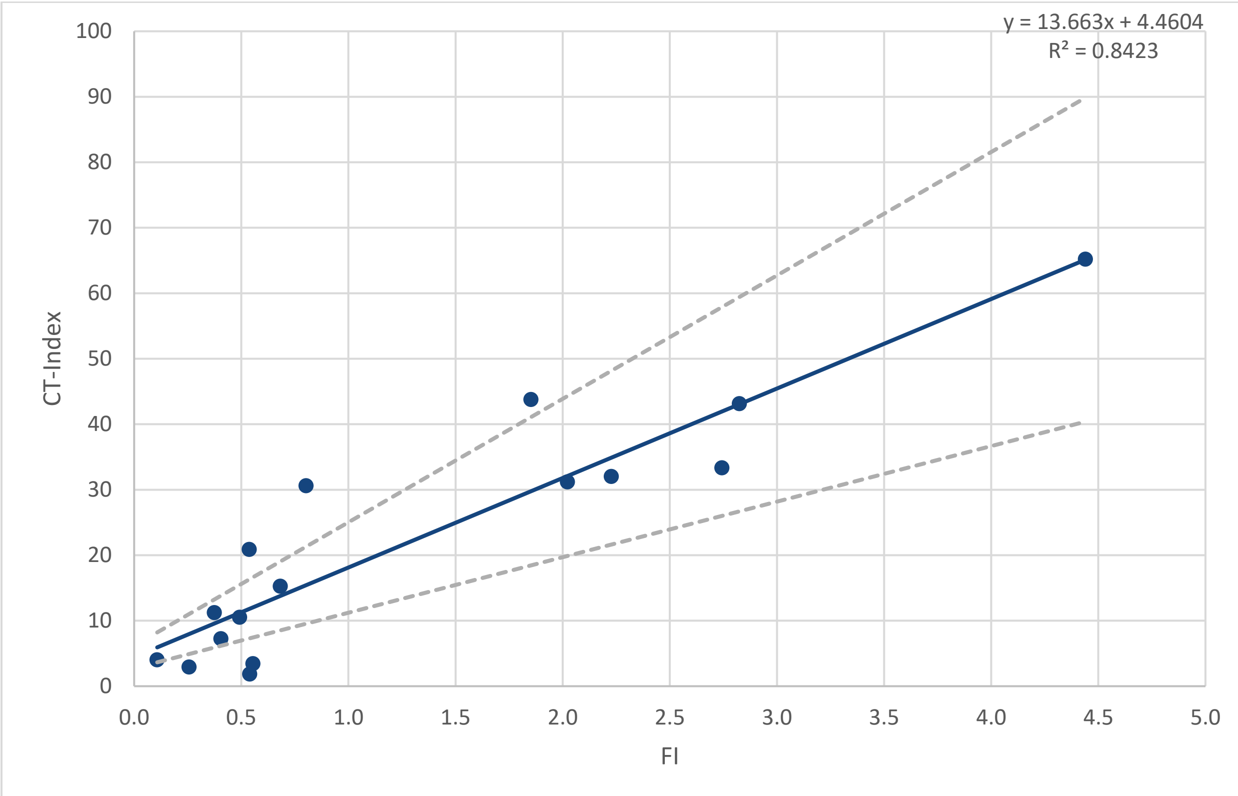


FIGURE 60 CT AND FI CORRELATION WITH 95% CONFIDENCE INTERVAL

4.3 Overall Findings

Based on the confusion analysis, there are several indicators that some of the laboratory thresholds contained in AASHTO PP-127 can be revised for better long-term field performance correlation. Potential adjustments after LTOA include reducing the threshold for CRI_{Env} with RPMLC specimens after LTOA to 26 and adding a threshold for FI with LMLC specimens after LTOA to 3. The current threshold for FI for LMLC specimens after STOA of 7 was adequate and required no adjustment.

The following overall findings emerged from this project:

- Field sections containing RAS exhibited marginal and poor cracking performance after 8-10 years in service. In the DE field project, low doses of recycling agents may also have been insufficient to offset the effects of RAS.
- Monitoring field performance at 70,000 CDD provides an adequate point for calibration and/or verification of laboratory cracking performance test thresholds using a confusion matrix analysis. Depending on the climate, 70,000 CDD will typically be reached within 4-8 years after construction.
- Chip seal maintenance treatments can reduce the rate of cracking based on field performance of the TX field project. In a field performance monitoring session at approximately 60,000 CDD, the rejuvenated section exhibited more than 10% lane cracking, while the virgin and DOT control sections had 3.6% and 4.4% lane cracking,

respectively. A chip seal was placed around 100,000 CDD. At the current field performance monitoring session at 178,321 CDD, the rejuvenated and DOT control sections both exhibited 8.0% lane cracking while the virgin section had 4.1% lane cracking. The placement of the chip seal prolonged the life of the sections and altered their cracking performance rating.

While sections with higher cracking rates naturally exhibited higher total percent lane cracking, monitoring the rate of crack propagation may be a more effective metric for timing of maintenance and rehabilitation treatments. By identifying sections with high rates of cracking above 1.5% per 10,000 CDD early in the pavement life, proactive maintenance and rehabilitation can be scheduled and implemented to extend the functional service life.

There were several lessons learned that can provide guidance for future studies focused on validation of mixture performance thresholds:

- For cracking resistance, field sections need to reach a climate-based age of at least 70,000 CDD which is typically achieved between 4-6 years.
- A minimum number of field sections is also required to provide sufficient confidence in the resulting conclusions.
- It is also important to identify the primary type(s) of pavement distress(es) challenging the DOT and the corresponding mixture performance test(s) that captures resistance to each distress.
- Consistent testing parameters and aging protocols are also essential for developing mix design systems tied to long-term field performance. For example, with relatively newer loose mixture aging protocols in AASHTO R 121, mixture performance thresholds will need to be verified.
- The chance of maintenance or rehabilitation for field sections increases with age, but lessons can still be learned after treatment.
- A range of traffic volumes should be utilized to provide robust conclusions.

4.4 Next Steps

Refinement of the AASHTO PP-127 performance thresholds should be proposed through AASHTO Committee on Materials and Pavements (COMP) in the next revision cycle. Results from this CAPRI project should also be shared with NCHRP 20-44(60) *Implementing a Training Program on the Use of High Recycled Binder Ratio Asphalt Mixtures* for inclusion in the training program under development for implementation of AASHTO PP 127. Expanding the confusion analysis with additional field sections where laboratory design results and long-term field performance results are available would further strengthen confidence in the national thresholds provided in AASHTO PP 127, especially for alternate loose mix LTOA protocols now available in AASHTO R 121.

Field validation of laboratory cracking performance thresholds through structured, annual monitoring of high-RAM mixtures over multiple years to 70,000 CDD (4-8 years depending on climate) is recommended for state or regional calibration. Standardizing the inspection protocols across all field sections is critical with analysis of long-term cracking performance. One of the objectives of a field performance monitoring program could be to evaluate the rate of % lane

cracking. These data could lead to the refinement of the 1.5% per 10,000 CDD, a threshold suggested by this CAPRI project for triggering pavement maintenance and rehabilitation.
As an example of impact, the NV DOT is currently working to adopt specifications to include provisions for high RAP mixtures using data measured in this project and the relationship developed between CT_{Index} and FI.

References

1. Williams, Brett A. and J. Richard Willis. 2025. "Asphalt Pavement Industry Survey on Recycled Materials and Warm-Mix Asphalt Usage: 2024." National Asphalt Pavement Association, Information Series 138, https://go.asphaltpavement.org/hubfs/IS138-2024_RAP-RAS-WMA_Survey_COMBINED_508.pdf?hsLang=en
2. Tran, Nam, Zhaoxing Xie, Grant Julian, Adam Taylor, Richard Willis, Mary Robbins, Shane Buchanan, 2016. "Effect of a Recycling Agent on the Performance of High-RAP and High-RAS Mixtures: Field and Lab Experiments." Journal of Materials in Civil Engineering, 29(1). https://ascelibrary.org/doi/10.1061/%28ASCE%29MT.1943-5533.0001697
3. Southern Plains Transportation Center (SPTC). 2024. "Improved Pavement Durability and Resiliency Using Balanced Mix Design with High RAP and Rejuvenator: A Path Toward Statewide Implementation." https://static1.squarespace.com/static/526c7a98e4b023d8f09390ed/t/68bf341b96a811509154dbdd/1757361179189/2-page-CY2-OU-LTU-12-Ali+-09-02-25_Final.pdf
4. Coleri, Erdem, Mayank Sukhija, Vipul Chitnis, Vikas Kumar, Servan Baran. 2026. "Implementation of Balanced Mix Design Methods in Oregon to Meet Long-Term Performance Goals." Oregon Department of Transportation. Report No. FHWA-OR-RD-26-10. https://www.oregon.gov/odot/Programs/ResearchDocuments/SPR852_BMD_Implementation.pdf
5. National Center for Asphalt Technology (NCAT). 2025. "Beyond Traditional Design: North Dakota validates balanced mix design to extend ND 14 pavement life." Asphalt Technology News. Spring 2025, Vol. 38, No. 1: 6-8. https://www.eng.auburn.edu/research/centers/ncat/research/newsletters/spring2025.pdf
6. Yazdipanah, Farzad, Mahdieh Khedmati, Robert Rea, Jamilla Teixeira, and Hamzeh F. Haghshenas. 2025. "Development of a Balance Mix Performance Framework for Assessing High-Recycled Asphalt Mixtures' Resistance: A Study Integrating Laboratory and Field Findings." Road Materials and Pavement Design 26 (11): 2836–55. https://doi.org/10.1080/14680629.2025.2465577
7. Guada, Irwin, John Harvey. 2025. "2025 Condition Survey of High RAP Pilot Projects on El Dorado 49, San Bernardino 215, and San Joaquin 26." UC Davis: University of California Pavement Research Center. Report No. UCPRC-TM-2025-04. https://doi.org/10.7922/G2862DVS
8. West, Randy, Rachel Cousins, Carolina Rodezno, Fan Yin, Nathan Moore, Matthew Sasser, Grant Julian. 2023. "Balanced Mixture Design Pilot and Test Sections." National Center for Asphalt Technology. WisDOT ID No. 0092-22-04. https://wisconsindot.gov/documents2/research/0092-22-04-final-report.pdf
9. Al-Qadi, Imad L., David L. Lippert, Shenghua Wu, Hasan Ozer, Greg Renshaw, Timothy R. Murphy, Abbas Butt, Satish Gundapuneni, James S. Trepanier, Joseph W. Vespa, Izak M. Said, Arturo F. Espinoza Luque, Fazal R. Safi. 2017. "Utilizing Lab Tests to Predict

Asphalt Concrete Overlay Performance." Illinois Center for Transportation. Report No. FHWA-ICT-17-020 https://doi.org/10.36501/0197-9191/17-026

10. Virginia Transportation Research Council (VTRC). 2025. "Field Performance of High RAP Mixtures." Last modified January 28, 2025. https://vtrc.virginia.gov/projects/all-projects/126826/
11. *Standard Practice for Balanced Mix Design (BMD)*. AASHTO PP 127-25, American Association of State Highway and Transportation Officials, 2020.
12. Epps Martin, Amy, Fawaz Kaseer, Edith Arámbula-Mercado, Akash Bajaj, Lorena Garcia Cucalon, Fan Yin, Arif Chowdhury, Jon Epps, Charles Glover, Elie Y. Hajj, Nathan Morian, Jo Sias Daniel, Mirkat Oshone, Reyhaneh Rahbar Rastegar, Chibuike Ogbo, and Gayle King. 2019. "The Effects of Recycling Agents on Asphalt Mixtures with High RAS and RAP Binder Ratios." NCHRP Report 927. National Cooperative Highway Research Program. https://doi.org/10.17226/25749
13. Leavitt, Aaron, Amy Epps Martin, Edith Arámbula-Mercado. 2023. "Pilot Test of Proposed Standard Practice for Use of Recycling Agents in Asphalt Mixtures with Recycled Asphalt Materials." National Cooperative Highway Research Program. NCHRP Project 20-44(24)
14. Epps Martin, Amy, Edith Arámbula-Mercado, Yong-Rak Kim, Juliana Montañez Gil, Santosh R. Kommidi, Nam Tran, Fan Yin, Raquel Moraes, Carolina Rodezno, Madhav Verma, Biswajit K. Bairgi, Gayle King. 2024. "Capturing Durability of High RBR Asphalt Mixtures." National Cooperative Highway Research Program. NCHRP Project 09-65. https://doi.org/10.17226/27991
15. https://www.wunderground.com/
16. Miller, John S., William Y. Bellinger. 2014. *Distress Identification Manual for the Long-Term Pavement Performance Program (Fifth Revised Edition*.) Report No. FHWA-HRT-13-092. Washington, DC: Federal Highway Administration, U.S. Department of Transportation. https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpp/13092/13092.pdf
17. Morian, Dennis A., Douglas Frith, Shelley Stoffels, Shervin Jahangirnejad. 2020. "Developing Guidelines for Cracking Assessment for Use in Vendor Selection Process for Pavement Crack Data Collection/Analysis Systems and/or Services." Report No. FHWA-RC-20-0005. Washington, DC: Federal Highway Administration, U.S. Department of Transportation. https://www.fhwa.dot.gov/pavement/management/pubs/fhwarc20005.pdf
18. Mohanraj, Kiran, Angelo Filonzi, Enad Mahmoud, Amit Bhasin. 2023. "Field Validation on the Use of Poker-Chip Test to Predict Cracking in Flexible Pavements." Journal of Transportation Engineering, Vol. 149(2), https://doi.org/10.1061/JPEODX.PVENG-1202
19. Chen, Chen, Fan Yin, Pamela Turner, Randy C. West, Nam Tran. 2018. "Selecting a Laboratory Loose Mix Aging Protocol for the NCAT Top-Down Cracking Experiment." Transportation Research Record: Journal of the Transportation Research Board. Vol. 2672(28). https://doi.org/10.1177/0361198118790639

20. Rainio, Oona., Jarmo Teuho, Riku Klén. 2024. "Evaluation Metrics and Statistical Tests for Machine Learning." Scientific reports. 14(1), 6086. https://doi.org/10.1038/s41598-024-56706-x
21. Zyl, Jean-Pierre van, Andries Petrus Engelbrecht, 2025, *Analysis of Classification Metric Behavior under Class Imbalance*, Egyptian Informatics Journal, Volume 31, 2025, 100711, https://doi.org/10.1016/j.eij.2025.100711