VTRC Virginia Transportation Research Council

We Bring Innovation to Transportation

Impact of Production Variability on Balanced Mix Designs in Virginia

http://www.virginiadot.org/vtrc/main/online_reports/pdf/23-r20.pdf

BENJAMIN F. BOWERS, Ph.D., P.E.
Assistant Professor, Department of Civil and Environmental Engineering, Auburn University

TIANA LYNN
Graduate Student, Department of Civil and Environmental Engineering, Auburn University

FAN YIN, Ph.D., P.E.
Assistant Director and Associate Research Professor, National Center for Asphalt Technology

NATHAN MOORE, P.E.
Lab Research Engineer, National Center for Asphalt Technology

STACEY DIEFENDERFER, Ph.D., P.E.
Associate Principal Research Scientist, Virginia Transportation Research Council

ILKER BOZ, Ph.D.
Research Scientist, Virginia Transportation Research Council

Final Report VTRC 23-R20

VIRGINIA TRANSPORTATION RESEARCH COUNCIL
530 Edgemont Road, Charlottesville, VA 22903-2454

vtrc.virginiadot.org

Standard Title Page - Report on Federally Funded Project

1. Report No.: FHWA/VTRC 23-R20	2. Government Accession No.:	3. Recipient's Catalog No.:
4. Title and Subtitle: Impact of Production Variability on Balanced Mix Designs in Virginia		5. Report Date: June 2023 6. Performing Organization Code:
7. Author(s): Benjamin Bowers, Ph.D., P.E.; Tiana Lynn; Fan Yin, Ph.D., P.E.; Nathan Moore, P.E.; Stacey Diefenderfer, Ph.D., P.E.; Ilker Boz, Ph.D.		8. Performing Organization Report No.: VTRC 23-R20
9. Performing Organization and Address: Auburn University National Center for Asphalt Technology 277 Technology Parkway Auburn, AL 36830 Virginia Transportation Research Council 530 Edgemont Road Charlottesville, VA 22903		10. Work Unit No. (TRAIS): 11. Contract or Grant No.: 116425
12. Sponsoring Agencies' Name and Address: Virginia Department of Transportation 1401 E. Broad Street Richmond, VA 23219 Federal Highway Administration 400 North 8th Street, Room 750 Richmond, VA 23219-4825		13. Type of Report and Period Covered: Final Contract 14. Sponsoring Agency Code:
15. Supplementary Notes: This is an SPR-B report.		

16. Abstract:

Balanced Mix Design (BMD) promotes durable pavements by ensuring resistance to multiple modes of distress through laboratory performance testing. BMD can also address material variability concerns present in volumetric design and encourage sustainability by allowing more recycled materials and innovative technologies. An example of addressing material variability concerns includes changes in aggregate specific gravity, which can change the volumetric properties through production. If performance testing is assessed through production, performance is verified despite inherent aggregate variability. Benchmarking studies and pilot projects have been conducted to select performance tests and threshold limits and evaluate the feasibility of implementing BMD in production. However, more information is needed regarding how plant variability can influence performance test results, specifically for Cantabro Mass Loss test, Asphalt Pavement Analyzer Rut Depth test (APA), and Indirect Tensile Strength Cracking test (IDT-CT), to fully implement performance testing into production.

The objective of this study was to evaluate the influence of production variability on performance test results to determine if mixtures balanced during design could become unbalanced during production. Fourteen currently accepted mixture designs, 10 BMD mixtures and 4 Superpave mixtures (to serve as a comparison between current standards and BMD designs), were recreated in the laboratory. Additionally, the mixes were adjusted to produce coarse and fine gradations and high and low binder contents according to currently accepted tolerance limits to simulate production variability. Two interaction mixtures assessing changes in gradation and binder content were evaluated as well as two critically aged mixtures.

Based on performance results, it was determined that mixtures originally balanced could become unbalanced due to production variability. Mixtures showed excellent rutting resistance but could be susceptible to durability and cracking issues as the gradation and binder content change. The Cantabro, APA, and IDT-CT were sensitive to changes in binder content. High binder content improved Cantabro and IDT-CT results but reduced the mixtures resistance to rutting, though this increase generally did not cause results to fail the threshold limit. Gradation was not found to be significant for IDT-CT, however, a coarse gradation negatively influenced Cantabro results, and a fine gradation hindered APA results. When evaluating the interaction between changes in gradation and binder content, performance test results were sometimes further negatively influenced. Complexities in results could be a performance test response to changes in volumetric parameters.

Based on the findings, this study recommends further refinement of the BMD specifications to ensure mixtures stay balanced through production despite variability. VTRC should also continue to assess the influence of volumetric parameters on performance test results since the interaction mixtures showed greater influences than other variations or different trends all together. Benefits of this study include furthering implementation of performance-based design and acceptance and informing VDOT regarding potential challenges with plant variations and their impact on volumetrics and performance results.

17 Key Words: Balanced Mix Design; Asphalt Mixtures; IDT-CT; APA; Cantabro Mass Loss, Variability, Plant Production		18. Distribution Statement: No restrictions. This document is available to the public through NTIS, Springfield, VA 22161.	
19. Security Classif. (of this report): Unclassified	20. Security Classif. (of this page): Unclassified	21. No. of Pages: 82	22. Price:

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

FINAL REPORT

IMPACT OF PRODUCTION VARIABILITY ON BALANCED MIX DESIGNS IN VIRGINIA

Benjamin F. Bowers, Ph.D., P.E.
Assistant Professor
Department of Civil and Environmental Engineering
Auburn University

Tiana Lynn
Graduate Student
Department of Civil and Environmental Engineering
Auburn University

Fan Yin, Ph.D., P.E.
Assistant Director and Associate Research Professor
National Center for Asphalt Technology

Nathan Moore, P.E.
Lab Research Engineer
National Center for Asphalt Technology

Stacey Diefenderfer, Ph.D., P.E.
Associate Principal Research Scientist
Virginia Transportation Research Council

Ilker Boz, Ph.D.
Research Scientist
Virginia Transportation Research Council

VTRC Project Manager
Stacey Diefenderfer, Ph.D., P.E., Virginia Transportation Research Council

In Cooperation with the U.S. Department of Transportation
Federal Highway Administration

Virginia Transportation Research Council
(A partnership of the Virginia Department of Transportation
and the University of Virginia since 1948.)

Charlottesville, Virginia

June 2023
VTRC 23-R20

DISCLAIMER

The project that is the subject of this report was done under contract for the Virginia Department of Transportation, Virginia Transportation Research Council. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Virginia Department of Transportation, the Commonwealth Transportation Board, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation. Any inclusion of manufacturer names, trade names, or trademarks is for identification purposes only and is not to be considered an endorsement.

Each contract report is peer reviewed and accepted for publication by staff of the Virginia Transportation Research Council with expertise in related technical areas. Final editing and proofreading of the report are performed by the contractor.

ABSTRACT

Balanced Mix Design (BMD) promotes durable pavements by ensuring resistance to multiple modes of distress through laboratory performance testing. BMD can also address material variability concerns present in volumetric design and encourage sustainability by allowing more recycled materials and innovative technologies. An example of addressing material variability concerns includes changes in aggregate specific gravity, which can change the volumetric properties through production. If performance testing is assessed through production, performance is verified despite inherent aggregate variability. Benchmarking studies and pilot projects have been conducted to select performance tests and threshold limits and evaluate the feasibility of implementing BMD in production. However, more information is needed regarding how plant variability can influence performance test results, specifically for Cantabro Mass Loss test, Asphalt Pavement Analyzer Rut Depth test (APA), and Indirect Tensile Strength Cracking test (IDT-CT), to fully implement performance testing into production.

The objective of this study was to evaluate the influence of production variability on performance test results to determine if mixtures balanced during design could become unbalanced during production. Fourteen currently accepted mixture designs, 10 BMD mixtures and 4 Superpave mixtures (to serve as a comparison between current standards and BMD designs), were recreated in the laboratory. Additionally, the mixes were adjusted to produce coarse and fine gradations and high and low binder contents according to currently accepted tolerance limits to simulate production variability. Two interaction mixtures assessing changes in gradation and binder content were evaluated as well as two critically aged mixtures.

Based on performance results, it was determined that mixtures originally balanced could become unbalanced due to production variability. Mixtures showed excellent rutting resistance but could be susceptible to durability and cracking issues as the gradation and binder content change. The Cantabro, APA, and IDT-CT were sensitive to changes in binder content. High binder content improved Cantabro and IDT-CT results but reduced the mixtures resistance to rutting, though this increase generally did not cause results to fail the threshold limit. Gradation was not found to be significant for IDT-CT, however, a coarse gradation negatively influenced Cantabro results, and a fine gradation hindered APA results. When evaluating the interaction between changes in gradation and binder content, performance test results were sometimes further negatively influenced. Complexities in results could be a performance test response to changes in volumetric parameters.

Based on the findings, this study recommends further refinement of the BMD specifications to ensure mixtures stay balanced through production despite variability. VTRC should also continue to assess the influence of volumetric parameters on performance test results since the interaction mixtures showed greater influences than other variations or different trends all together. Benefits of this study include furthering implementation of performance-based design and acceptance and informing VDOT regarding potential challenges with plant variations and their impact on volumetrics and performance results.

FINAL REPORT

IMPACT OF PRODUCTION VARIABILITY ON BALANCED MIX DESIGNS IN VIRGINIA

Benjamin F. Bowers, Ph.D., P.E.
Assistant Professor
Department of Civil and Environmental Engineering
Auburn University

Tiana Lynn
Graduate Student
Department of Civil and Environmental Engineering
Auburn University

Fan Yin, Ph.D., P.E.
Assistant Director and Associate Research Professor
National Center for Asphalt Technology

Nathan Moore
Lab Research Engineer
National Center for Asphalt Technology

Stacey Diefenderfer, Ph.D., P.E.
Associate Principal Research Scientist
Virginia Transportation Research Council

Ilker Boz, Ph.D.
Research Scientist
Virginia Transportation Research Council

INTRODUCTION

Background

The goal of asphalt mix design is to provide a durable mixture that meets the economic, social, and environmental needs of the traveling public. Asphalt mixture design approaches have evolved as technological advancements and the understanding of mixture performance and performance prediction have grown. Previous asphalt mixture design methods (Marshall followed by Superpave) were effective in mitigating at least one mode of distress, but often left the mixture susceptible to other distresses. This often resulted in premature maintenance and intervention, and thus additional economic investment. To mitigate the shortcomings of the Marshall mix design system, the Virginia Department of Transportation (VDOT) began implementing the Superpave mixture design system in 1997, and achieved full implementation by 2002 (Diefenderfer et al., 2021a). Although performance testing was originally part of the Superpave mix design plan for various traffic levels, it was not implemented due to complex equipment requirements and specialized training; therefore, design and acceptance are based on volumetric properties (Asphalt Institute, 2014; Boz et al., 2022). This system was successful in reducing rutting within the mixture but, in many cases, resulted in mixtures that were prone to cracking and durability issues. There are a number of other concerns with designing and accepting asphalt mixtures solely based on volumetric properties, as done in the Superpave design methodology:

1. Aggregate specific gravity and mineralogy can change throughout production. Slight variations in these properties can produce considerable differences within the design volumetrics.

2. Recycled binder availability for blending is unknown, which can influence the quantity of total active binder in the mixture. In addition, the true grade of the recycled binder could be variable throughout the stockpile. Blending availability and variability can potentially compromise the durability and longevity of the mix.

3. The effectiveness and interaction of innovative technologies are not quantified or assessed in the volumetric process (Diefenderfer et al., 2021b).

In the early 2000s, there was growing interest in increasing the reclaimed asphalt pavement (RAP) content in new pavements to reduce costs and environmental burden, and by 2007 VDOT increased the maximum allowed RAP content to 30% for surface mixtures (Diefenderfer et al., 2021a). Increasing RAP content also increases the amount of brittle, aged binder within the mixtures, potentially leading to mixtures that are more susceptible to cracking and durability issues. Coupled with the inherent cracking and durability issues associated with Superpave, additional brittle and aged binder caused concern for the longevity of flexible pavements.

Balanced Mix Design (BMD) addresses these concerns by evaluating performance during design and production through various performance tests to resist multiple modes of distress. Using performance testing in design and production ensures acceptable performance is achieved and maintained throughout the construction process despite unknown RAP binder contribution and potential changes in aggregate properties and allows for the use of innovative technologies to increase performance and lifespan.

Based on these benefits, VDOT has had growing interest in BMD, establishing an implementation framework, conducting various research studies, and developing special provisions. The BMD framework includes:

- Selecting laboratory tests to assess common modes of distress and selecting threshold limits for those tests (Diefenderfer and Bowers, 2019; Bowers et al., 2022)
- Constructing and observing pilot projects
- Equipment acquisition
- Training personnel

- Refining specifications through research
- Initial implementation (Diefenderfer et al., 2021a).

Diefenderfer and Bowers (2019) and Bowers et al. (2022) conducted studies evaluating various performance tests based on the results of approved mixtures to select the laboratory performance tests for VDOT's BMD specification. These tests were to assess common modes of distress for Virginia, namely durability, rutting, and cracking, and were selected based on correlations to known mix properties, simple testing, equipment availability, testing efficiency, and repeatability. From this study, the Cantabro Mass Loss test to assess durability, the Asphalt Pavement Analyzer (APA) Rut Depth test to quantify rutting, and the Indirect Tensile Asphalt Cracking Test (IDT-CT) to evaluate cracking were selected. In addition, threshold limits for each of the tests were selected and are listed in 1. These tests and threshold limits were then verified through a study conducted by Diefenderfer et al. (2021a).

Table 1. Selected Performance Tests and Threshold Limits

Performance Test	Threshold Limit
Cantabro Mass Loss Test	Mass Loss (ML) $\leq$ 7.5%
Asphalt Pavement Analyzer (APA) Rut Depth Test	Rut depth $\leq$ 8 mm
Indirect Tensile Asphalt Cracking Test (IDT-CT)	$CT_{index} \geq 70$

CT = cracking tolerance.

Following test and threshold selection, two special provision specifications were developed (*Special Provision for Dense-Graded Surface Mixtures Designed Using Performance Criteria* and *Special Provision for High Reclaimed Asphalt Pavement (RAP) Content Surface Mixtures Designed Using Performance Criteria*), and in 2019 two field trial studies assessing nine mixtures were evaluated. Different materials and innovative technologies were evaluated in this study based on the Cantabro, APA, and IDT-CT tests on reheated and non-reheated samples. It was concluded from this study that additional research was needed to evaluate the influence of mixture properties on performance test results. The same is true for high RAP contents (40%) and innovative technologies, as some trial work was very successful. However, more work is needed to determine the conditions and requirements where that could be expected to be consistently produced by industry and prove to be a durable, performing pavement in the field over time.

Prior to full implementation, it is necessary to develop an effective quality assurance (QA) plan to ensure that performance evaluated in the laboratory is replicated in the field. However, to develop this plan, the influence of production variability on performance, as defined by the laboratory performance testing criteria, must first be determined to assess whether currently acceptable variability in production can cause mixtures to become unbalanced. This study aims to assess production variability with respect to performance to inform decisions regarding refining the specifications for full implementation of BMD and completing the steps outlined in the original framework.

PURPOSE AND SCOPE

This study aims to inform VDOT and its contractors regarding the influence of production variability on the laboratory performance of mixtures according to the Cantabro, APA, and IDT-CT tests. Results of this research will aid in decision-making regarding specification refinement and next steps toward full implementation, including QA, for BMD.

The objective of this project is to evaluate the impact of production variability on performance test results to determine if production variability can cause balanced mixtures in the design phase to become unbalanced during production. Further, it is important to determine if these differences produce statistically different results from the originally balanced design mixture. To evaluate these objectives, fourteen approved mixtures were reproduced in the laboratory according to the Job Mix Formula (JMF) to serve as controls. Controls were then compared to variations in design, which included coarse and fine gradations at the optimum asphalt content and increased and decreased asphalt contents at the design gradation. Two mixtures were selected to assess the interaction of changing both gradation and asphalt content by combining upper and lower asphalt contents with coarse and fine gradations. In addition, two mixtures were selected to assess critical aging of the mixture and to simulate in-service mixture performance.

METHODS

To achieve the objectives in this study, the following tasks were completed:

1. Conduct a literature review to capture the state of the practice.
2. Identify mixture designs and materials for evaluation.
3. Select tolerance limits; simulate plant and material variability.
4. Establish coarse and fine gradations based on the tolerance limits.
5. Process materials and fabricate samples for performance testing.
6. Evaluate performance based on ML, APA, and IDT-CT.
7. Perform statistical analysis to determine if mixtures balanced in design could become unbalanced in production.

Literature Review

A literature review was conducted to understand the state of the practice with respect to BMD and production tolerances leveraging the Auburn University library and subsequent journal subscriptions. Key databases were searched, such as the Transportation Research Information Database (TRID), Sage Premier, and the American Society of Civil Engineers (ASCE) journal databases, to name a few. The literature was synthesized and used to inform the study.

Mixture Designs and Materials

Fourteen approved mixture designs, listed in Table 2, were evaluated in this study. Four were designed under the Superpave specification according to Section 211 of the *Road and Bridge Specification* to serve as a comparison to the 10 that were designed under the BMD special provisions (*Special Provision for Dense-Graded Surface Mixtures Designed Using Performance Criteria* or the *Special Provision for High Reclaimed Asphalt Pavement (RAP) Content Surface Mixtures Designed Using Performance Criteria, revised version 3/18/2020)* (VDOT, 2016). The Superpave mixtures were limited to a RAP content of 30%, whereas the BMD mixtures had RAP contents of 35-40%. Mixtures evaluated had a nominal maximum aggregate size (NMAS) of 9.5 mm or 12.5 mm. Since this study was conducted in conjunction with field trials, materials (aggregate, RAP, and binder) were sampled during construction by VTRC staff or asphalt producer staff so that the materials would be similar between studies and results could be compared. Aggregate stockpile gradations were provided by the contractor and used unless apparent discrepancies arose during blending.

Table 2. Mix Designs

Producer	Design Method	NMAS	RAP%	Binder	Additive
A	Volumetric Design	9.5	30	PG 64S-22	-
	Balanced Mix Design	9.5	40	PG 58-28	-
	Balanced Mix Design	9.5	40	PG 64S-22	Recycling Agent
B	Volumetric Design	12.5	30	PG 64S-22	-
	Balanced Mix Design	12.5	40	PG 58-28	-
	Balanced Mix Design	12.5	40	PG 64S-22	Recycling Agent
C	Balanced Mix Design	12.5	35	PG 58-28	-
D	Volumetric Design	12.5	30	PG 64S-22	-
	Balanced Mix Design	12.5	40	PG 58-28	-
E	Balanced Mix Design	12.5	35	PG 58-28	Recycling Agent
	Balanced Mix Design	12.5	35	PG 58-28	Softening Oil+ Fibers
F	Volumetric Design	9.5	30	PG 64S-22	-
	Balanced Mix Design	9.5	40	PG 58-28	-
	Balanced Mix Design	9.5	40	PG 64S-22	Recycling Agent

NMAS = nominal maximum aggregate size; RAP = reclaimed asphalt pavement; PG = performance grade.

Selecting Tolerance Limits

This study is predicated on simulating production variability based on currently accepted variability according to the specification. Acceptable quality in production is defined as the mean result falling within the range listed in Table 3, with respect to the JMF, corresponding to the number of tests conducted. These tolerances were used to determine if mixtures deviating from the JMF but still falling within acceptable production limits could produce mixtures that did not meet laboratory performance test criteria. From this table, the tolerance limits were selected by choosing the corresponding number of tests. To mirror production practices, the range of number of tests to select was narrowed down to between 2 and 8. Table 3 is used for material acceptance in which the mean of 8 tests, typically from a 4,000-ton lot, must meet the tolerance limits. However, if a test fails the specification, it is resampled and tested again. If the subsequent test does not meet the specification, corrective action can be pursued, thus reducing the possible tolerances to those that correspond to 2-8 tests.

In a similar study conducted for MassDOT, Mogawer et al. (2019) used the following tolerance limits to evaluate the influence of production variability on performance test results: ±6% for No.4 sieve, ±5% for No. 8 sieve, ±3% for No. 16 sieve, ±1% for No. 200 sieve, and ±0.3% for asphalt content (AC). These values are similar to values in Table 3 between 2 and 4 tests. From an internal survey, producers stated that for NMAS 12.5 mm mixtures, typical variability for the No.4 sieve was ±4% from the JMF, which is similar to acceptable tolerances for four tests from Table 3. Based on previous studies and the internal survey, tolerance limits corresponding to four tests were selected to simulate production variability for this study.

Table 3: Process Tolerances from Road and Bridge Specification, Section 211.08, Table II-15 (VDOT, 2016)

Process Tolerance

Tolerance on Each Laboratory Sieve and Asphalt Content: Percent Plus and Minus												
No. Tests	Top Size[1]	1 1/2”	1”	3/4”	1/2”	3/8”	No. 4	No. 8	No. 30	No. 50	No. 200	A.C.
1	0.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	6.0	5.0	2.0	.60
2	0.0	5.7	5.7	5.7	5.7	5.7	5.7	5.7	4.3	3.6	1.4	0.43
3	0.0	4.4	4.4	4.4	4.4	4.4	4.4	4.4	3.3	2.8	1.1	0.33
4	0.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	3.0	2.5	1.0	0.30
5	0.0	3.6	3.6	3.6	3.6	3.6	3.6	3.6	2.7	2.2	0.9	0.27
6	0.0	3.3	3.3	3.3	3.3	3.3	3.3	3.3	2.4	2.0	0.8	0.24
7	0.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	2.3	1.9	0.8	0.23
8	0.0	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.1	1.8	0.7	0.21
12	0.0	2.3	2.3	2.3	2.3	2.3	2.3	2.3	1.7	1.4	0.6	0.17

[1]Defined as the sieve that has 100% passing as defined in Table II-13.

Establishing Coarse and Fine Gradations

To establish the coarse and fine gradation targets, the tolerance limits from Table 3 were added and subtracted from the JMF. Stockpile percentages were then optimized, as shown in Table 4, to produce a gradation closest to the target gradation without allowing any of the sieves to fall outside of tolerance. This method typically resulted in one or two sieves that deviated from the JMF to the maximum limit while others were relatively close to the limit, as shown in the gradation variation example in Table 5 and Figure 1. During production, variability is likely to occur for some sieves but is not as likely to vary to the extreme limit for every sieve. Therefore, this method was pursued to mirror production practices as well as optimize sample preparation procedures. Since stockpile quantities were altered in using this method, combined bulk specific gravity (Gsb) was adjusted accordingly.

Table 4: Example of Optimized Stockpile Percentages to Produce Coarse and Fine Gradations, Producer E 35R PG 58-28 with Softening Oil and Fibers

	Stockpile Percentages		
Stockpile	**JMF**	**Coarse**	**Fine**
78's B	11.0%	19.9%	10.5%
8's PE	14.8%	10.0%	11.0%
17's B	5.0%	10.0%	5.0%
17's PE	33.1%	25.0%	36.5%
Baghouse	1.1%	0.1%	2.0%
RAP	35.0%	35.0%	35.0%

JMF = job mix formula; RAP = reclaimed asphalt pavement.

Table 5: Example Gradation Variations, Producer E 35R PG 58-28 with Softening Oil and Fibers

Sieve	JMF	-TL	Coarse	+TL	Fine
3/4"	100		100.0		100.0
1/2"	98		97.7		98.3
3/8"	90	86.6	88.3	94.6	91.4
# 4	61	58.6	59.2	66.6	66.4
# 8	42	39.2	40.3	47.2	46.3
# 16			28.5		33.4
# 30	25	20.4	21.1	26.4	25.4
# 50		14.6	15.0	19.6	18.7
#100			8.9		11.7
#200	5.7	4.7	4.69	6.7	6.52

JMF = job mix formula; -TL = target gradation, where the tolerance limit was subtracted from the JMF; coarse = the coarse gradation produced by changing stockpile quantities in an effort to match the -TL; +TL = target gradation, where the tolerance limit was added to the JMF; fine = the fine gradation produced by changing stockpile quantities in an effort to match the +TL.

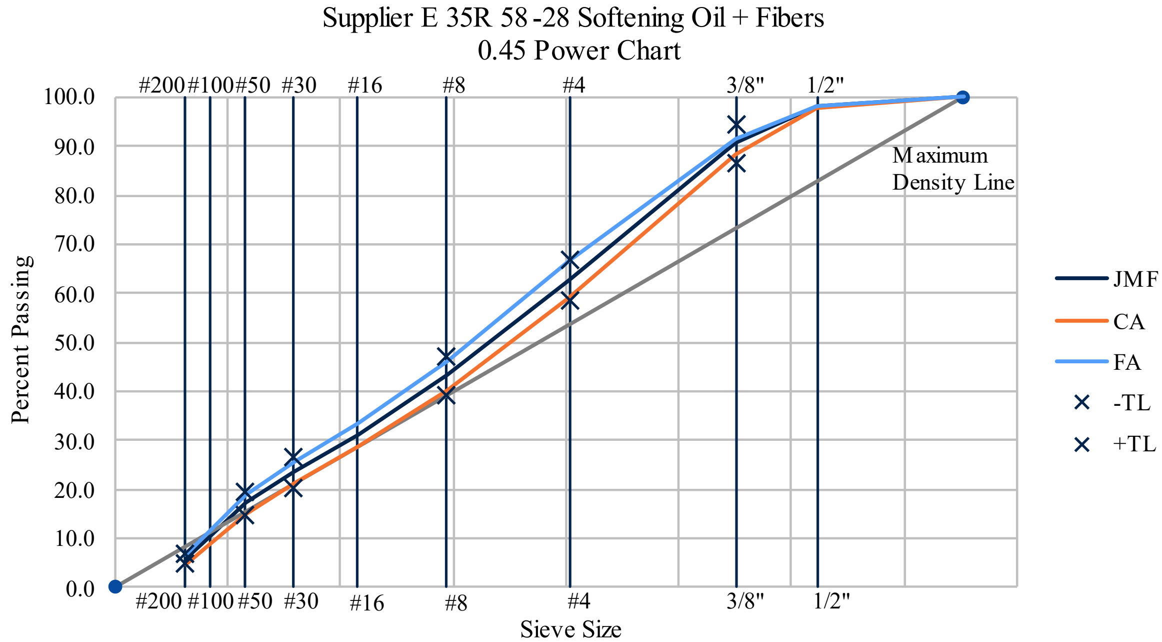

Figure 1. Example Gradation Variations Power 45 Chart, Producer E 35R PG 58-28 with Softening Oil and Fibers. JMF = job mix formula; CA = coarse adjusted gradation; FA – fine adjusted gradation; TL = target limit.

Material Processing and Specimen Preparation

Prior to specimen fabrication and testing, RAP and aggregate materials were dried by air drying and oven drying, respectively, and processed by fractionating. RAP asphalt content (AC) and gradation were determined according to AASHTO T 308, A*sphalt Binder Content by Ignition Oven method*, and AASHTO T 27, *Sieve Analysis for Fine and Coarse Aggregate*. This was conducted to confirm RAP AC from the design to ensure accurate total asphalt content was produced in the laboratory. Sieve analysis, in accordance with AASHTO T 27, was also determined for as-received RAP, before evaluating the binder content as per AASHTO T 308, to determine the sieve to fractionate RAP material to reduce aggregate variability in the laboratory. The sieve chosen for fractionation corresponded to 50% passing for the unburned (black rock) RAP gradation.

Samples were batched according to the respective gradation (JMF, coarse, and fine), heated, and combined with asphalt binder, recycling agents (RAs), or fibers according to the JMF or the manufacturer's instructions. Production practices were simulated in the laboratory by using a bucket mixer and standard conditioning times: 2 hours for Cantabro and APA and 4 hours for IDT-CT. Cantabro samples were compacted to N_{design} (50 gyrations) at a height of 115±5 mm with a gyratory compactor, while APA and IDT-CT were compacted to 7.0% air voids and a height of 75 mm and 62 mm, respectively. Air voids were verified for all test samples according to AASHTO T-166, *Standard Method of Test for Bulk Specific Gravity (Gmb) of Compacted Asphalt Mixtures Using Saturated Surface-Dry Specimens*.

Material Processing Challenges and Solutions

Several challenges were encountered and addressed in processing materials and preparing for sample fabrication. Most challenges involved RAP and aggregate variability, as well as reproducing contractor methods. Solutions were formulated within the research teams at NCAT and VTRC with insight from industry and agency partners when appropriate. Challenges relating to each supplier and subsequent solutions are summarized below:

- Supplier A: The unburned RAP gradations, or black rock gradations, evaluated in the laboratory were finer for some sieves as compared to the contractor's historica data for burned gradations. Since the RAP aggregate is coated with asphalt, the unburned gradation should be coarser than burned gradation data. To address variability in the RAP for this supplier, fractionated RAP was reblended to produce a gradation similar to the JMF.

- Supplier B: Following the ignition oven burn, it was determined that the mean asphalt content was 1.1% higher than the RAP AC listed in the design documents. The virgin binder content was reduced to produce a total binder content consistent with the JMF. No other blending or adjustments were required.

- Supplier C: Half of the RAP supplied was processed which had been crushed and screened, while the other half was millings which had not undergone any processing post milling from the project site. To replicate contractor practices, the millings were blended with the processed RAP; however, to reduce variability within laboratory procedures, the material was screened over a 1-inch sieve. Due to limited processing of the RAP, considerable variability was observed between gradation samples. However, the mean percentage passing for each sieve was similar to design data from the contractor, and therefore, no additional processing or blending was conducted. In addition, the RAP AC was 1% lower than the design so virgin binder was added to be consistent with the total binder content listed in the JMF.

- Supplier D: The JMF gradation could not be produced with the data provided by the contractor, and aggregate gradations were completed to troubleshoot the discrepancy. After correspondence with the contractor, it was determined that additional aggregate material would need to be ordered. The RAP gradation presented some variability but resulted in only a 0.2% difference in binder content compared to the design.

- Supplier E: RAP gradation presented low variability for this supplier and was consistent with the design data. The RAP AC was 0.7% different from the design, and the virgin binder content was adjusted accordingly. Since RAP for this supplier did not present considerable variability as with Supplier A, interactions between gradation and binder content were assessed for this supplier instead of Supplier A. However, on account of this change, more material had to be sampled. Although the RAP was found to be consistent with the design, the aggregate blend produced based on contractor stockpile data did not result in a consistent blend with the JMF. After assessing aggregate gradations and contractor correspondence, it was determined that the contractor produces their final design through the plant rather than the laboratory. Thus, the JMF includes plant-related aggregate breakdown, which is not accounted for in laboratory stockpile batching percentages. To reflect this aggregate breakdown in the laboratory, additional baghouse fines were added. However, the breakdown at the plant was considerable and the mixture would have required approximately 4% baghouse fines. Excessive amounts of baghouse fines can begin to distort the rest of the gradation. To address this issue, a finer source of baghouse fines was substituted.

- Supplier F: RAP gradations showed consistency between samples, and RAP AC was only -0.2% from design and was adjusted accordingly.

For several suppliers, aggregate and RAP variability was inherent such that gradations and asphalt content deviated from design. Based on this observation, it can be inferred that other material properties may have changed as well, such as aggregate bulk specific gravity, which influences other volumetric parameters. Challenges encountered in this project bring to light the potential material variability between design and construction and the need for performance indicators beyond volumetrics for both design and quality control. In addition, these observations illuminate the need for more control over RAP stockpiles especially since the RAP content is a considerable portion of the mixture at 30-40%.

Performance Tests

Cantabro Mass Loss

The Cantabro Mass Loss test was conducted in accordance with AASHTO TP 108, *Standard Method of Test for Abrasion Loss of Asphalt Mixture Specimens*, for a minimum of 3 replicates where samples are subject to 300 revolutions in a Los Angeles abrasion machine without steel charges. Prior to testing, samples were air dried after determining air voids and allowed to rest at room temperature. A total of 245 samples were tested for this study. A ML ≤ 7.5% is considered a passing result.

Asphalt Pavement Analyzer Rut Depth

The APA test was conducted in accordance with AASHTO T 340, *Standard Method of Test for Determining Rutting Susceptibility of Hot Mix Asphalt (HMA) Using the Asphalt Pavement Analyzer (APA)*, on four cylindrical specimens set in two tracks. The tracks were subjected to 8,000 passes at 6 ℃. Samples were allowed to air dry after determining air voids and then were conditioned in the APA environmental chamber for 3 hours, to simulate VTRC conditioning procedures, before testing. A total of 312 specimens were tested making up 156 tracks. A rut depth ≤ 8 mm is considered a passing result.

Indirect Tensile Cracking Test

The IDT-CT test was conducted in accordance with ASTM D8225-19, *Standard Test Method for Determination of Cracking Tolerance Index of Asphalt Mixture Using the Indirect Tensile Cracking Test at Intermediate Temperature*, using a hydraulic load frame and TestQuip software. Samples were conditioned in an environmental chamber at 25℃ for 2 hours after being air dried due to determining air voids by SSD method. At least 5 samples were tested for each set of tests for a total of 519 samples tested. A $CT_{index} \geq 70$ is considered a passing result.

Analysis

Before data analysis, the data was screened for outliers according to ASTM E178, *Dealing with Outlying Observations*. If a data point was identified as an outlier, the test set standard deviation was compared to an industry-accepted value. If the standard deviation was greater than that value, then the outlying point was rejected; however, if the standard deviation was less, the point was included in the data set. No samples were rejected from the Cantabro and APA data set. However, a few samples were identified as outliers for the IDT-CT data according to ASTM E178 with a 90% confidence level. The CT_{index} standard deviation from ASTM D8225 single lab repeatability is 13.5. For samples that are identified as outliers but belong to a set that has a standard deviation of 13.5 or less, the data was kept for analysis. If they are outliers and belong to a data set that exceeds 13.5, the data point was rejected. A total of 7 samples were rejected from the data, which reduced the number of samples for IDT-CT from 526 to 519. It should be noted that while ASTM D8225 states 13.5 as the standard deviation for single lab repeatability, there are a number of studies recently completed or ongoing that evaluate this threshold and future work should reflect the state-of-the-practice (Boz et al., 2022; Habbouche et al., 2021; Habbouche et al., 2022).

Data analysis was first evaluated by comparing laboratory reproduced JMF results to each mixture variation to quantify the impact of production variability on performance test results by assessing the percent change from design according to Equation 1.

$$Percent\ Change\ from\ Design = \frac{Variation\ of\ JMF - JMF}{JMF} * 100 \qquad \text{[Eq. 1]}$$

Following preliminary analysis, statistical analysis was conducted. One-way Analysis of Variance (ANOVA) test was evaluated in Minitab© for mass loss, rut depth, and CT_{index} to determine if at least one of the means of the variations was statistically different. A confidence interval of 95% was used ($\alpha = 0.05$) where the null hypothesis, which assumes all the variation means are equal, is rejected if the p-value is less than the α-value of 0.05. If the null hypothesis was rejected and the means were found to be significant, Dunnett's test was evaluated to determine if the mean of each respective variation was statistically different from the design, O-J. These statistical tests assume normality and equal variance for the data (Devore and Farnum, 2005). Normality and equal variance were evaluated at a 95% confidence interval. The majority of designs (70%-100% depending on the test) confirmed equal variance such that this assumption was accepted.

RESULTS AND DISCUSSION

Literature Review

BMD was officia y defined b the Balanced Mix Design Task Force as "asphalt mi design using performance tests on appropriately conditioned specimens that address multiple modes of distress taking into consideration mix aging, traffic, climate and location within the pavement structure" (West et al., 2018; Wang et al.2023). The National Asphalt Pavement Association (NAPA), in conjunction with the National Center for Asphalt Technology (NCAT), has been maintaining a record of BMD implementation efforts across the U.S. Currently, this database shows that the majority of states are pursuing BMD to at least some degree. For example, some states are in the pre-implementation phase, while many others have implemented BMD specifications (NAPA, 2023). Through this implementation process, states have conducted benchmarking studies to select performance tests and the respective threshold limits as well as assess the current state of their designs, filling the knowledge gap within this area.

Following benchmarking, pilot projects have been conducted to evaluate implementation of the selected performance tests within the QC process and to determine if production data deviates from design data. Based on the Virginia Transportation Research Council (VTRC) 2019 field trials report, a conclusion was made that results from performance testing indicated variability although gradation and asphalt content were similar, and that additional information was needed to evaluate the relationship between changes in gradation and asphalt content and performance tests results (Diefenderfer et al., 2021b). A similar conclusion was made by Ling and Buchanan (2022) when evaluating Vermont plant produced mixtures. This study observed that performance results were generally higher than design values. The authors advised that factors during production that can influence performance results be considered and understood when applying performance testing during production, including aging condition, production variability, and material variability (Ling and Buchanan, 2022).

Assessing performance testing during production is imperative to pursuing full implementation of BMD and has been evaluated in many states by benchmarking mixtures and pilot projects. However, there is limited knowledge regarding how plant variability can influence these results. Mogawer et al. (2019) evaluated the influence of gradation, binder content, and binder source on balanced mix designs to determine if originally balanced designs could become unbalanced due to production variability or acceptable plant practices such as changing binder source. A balanced design according to the Volumetric Design with Performance Verification method was established, and gradation, asphalt content, and asphalt source were then varied according to Massachusetts Department of Transportation (MassDOT) tolerances. Variations produced upper and lower limits for asphalt content for both the coarse and fine gradations, which was also evaluated for each of the asphalt sources. Performance was evaluated according to the Hamburg Wheel-tracking Devise (HWTD), Illinois Flexibility Index Test (I-FIT), IDT-CT, and Bending Beam Rheometer (BBR).

Based on this study, none of the mixes were susceptible to rutting, indicating that current tolerances produce rut resistant mixes. The Flexibility Index (FI) for the IFIT test showed significant changes with respect to gradation but did not suggest that one gradation (coarse or fine) was superior to the other, while asphalt content was also found to be significant and produced a correlation of increasing FI with increasing binder content. IDT-CT did not show variability with respect to gradation, but significant differences were observed when asphalt content and asphalt source were varied. Based on this study, mixtures balanced during design could become unbalanced during production as a result of production variability. The main cause for mixtures to become unbalanced was related to failing durability or cracking criteria due to low binder content (Mogawer et al., 2019).

Similarly, Austerman et al. (2018) evaluated the effects of production variations, namely asphalt content, asphalt source, and percent passing the No. 200 sieve, on BMD. Two mixtures were designed volumetrically, and performance verified according to the HWTD and the IFIT tests. MassDOT specifications were used in the volumetric design as well as for determining the thresholds for test tolerances. When variations in asphalt content were evaluated, the HWTD and IFIT results indicated that both mixes stayed balanced except for one mix that fell outside of tolerance for the stripping inflection point (SIP). For the IFIT data, the FI increased when the amount of asphalt binder increased, as expected. Variations of percent passing the No. 200 sieve also resulted in balanced mixes; however, the +1% passing the No. 200 sieve mixes did experience more rutting than mixes with less dust. A statistical difference was observed when changing the asphalt binder source. One asphalt binder source consistently exhibited lower values of fracture energy (FE) compared to the other. Although variations were significant, performance results met threshold criteria such that the mixes stayed balanced. However, this study showed that due to the variability observed it is possible to produce a mixture that is unbalanced based on the tolerances established by state agencies (Austerman et al., 2018).

Previous studies concluded that mixtures originally balanced could become unbalanced during production by failing to meet at least one of the performance criteria due to changes in gradation, asphalt content, and asphalt source. Transitioning to full implementation and using performance testing in production, it will be critical to understand the implications of material and plant variability, specifically related to native Virginia materials and the Virginia BMD specifications.

Impact of Production Variability on Performance Test Results

Each of the 14 mixtures were reproduced in the laboratory and evaluated for performance according to the Cantabro, APA, and IDT-CT tests. Volumetric and gradation summaries for each design are included in Appendix A and performance test results are included in Appendix B. Table 6 details the notation convention where the binder content is listed first, followed by the gradation. For example, O-J is the JMF reproduced in the laboratory with the optimum asphalt content (OAC) at the design gradation. In addition, mixtures are designated by producer (A, B, C, etc.), RA content followed b the etter ‘R’ (5R for 35% RAP), performance grade (PG), and recycling agent (RA) or additive.

Table 6. Mixture Variations Legend

Design Variation	Abbreviation
Optimum Asphalt Content (OAC)	O
Low Asphalt Content (-0.3% OAC)	L
High Asphalt Content (+0.3% OAC)	H
Design Gradation from the JMF	J
Coarse Gradation	C
Fine Gradation	F

Figure 2 summarizes the cracking and rutting resistance of the mixtures and their respective variations to illustrate the general performance tendencies of standard Superpave and BMD mixtures. Variations resistant to cracking and rutting plot in the upper left corner where the CT_{index} is greater than 70 and the rut depth is less than 8mm.

The mixtures show excellent resistance to rutting despite gradation and binder content variations. The only test to fail to meet rutting criteria was the interaction of high binder content and a fine gradation. Austerman et al. (2018) and Mogawer et al. (2019) noted that balanced mixtures were sensitive to high binder contents and high dust contents, which was defined as +1% passing the No.200 sieve. This is consistent with the fine gradation in this study. Based on these observations from the literature, the H-F variation produces a worst-case scenario.

With respect to the cracking criteria, mixtures spanned the cracking resistant and cracking susceptible regions indicating that mixtures may become unbalanced when production variability is considered. Further, the figure shows general clustering of the data based on binder content where low binder content produces lower CT indices whereas higher binder content produces higher CT indices. Inadequate cracking resistance did not correspond to rutting susceptibility, thus none of the variations were susceptible to both cracking and rutting.

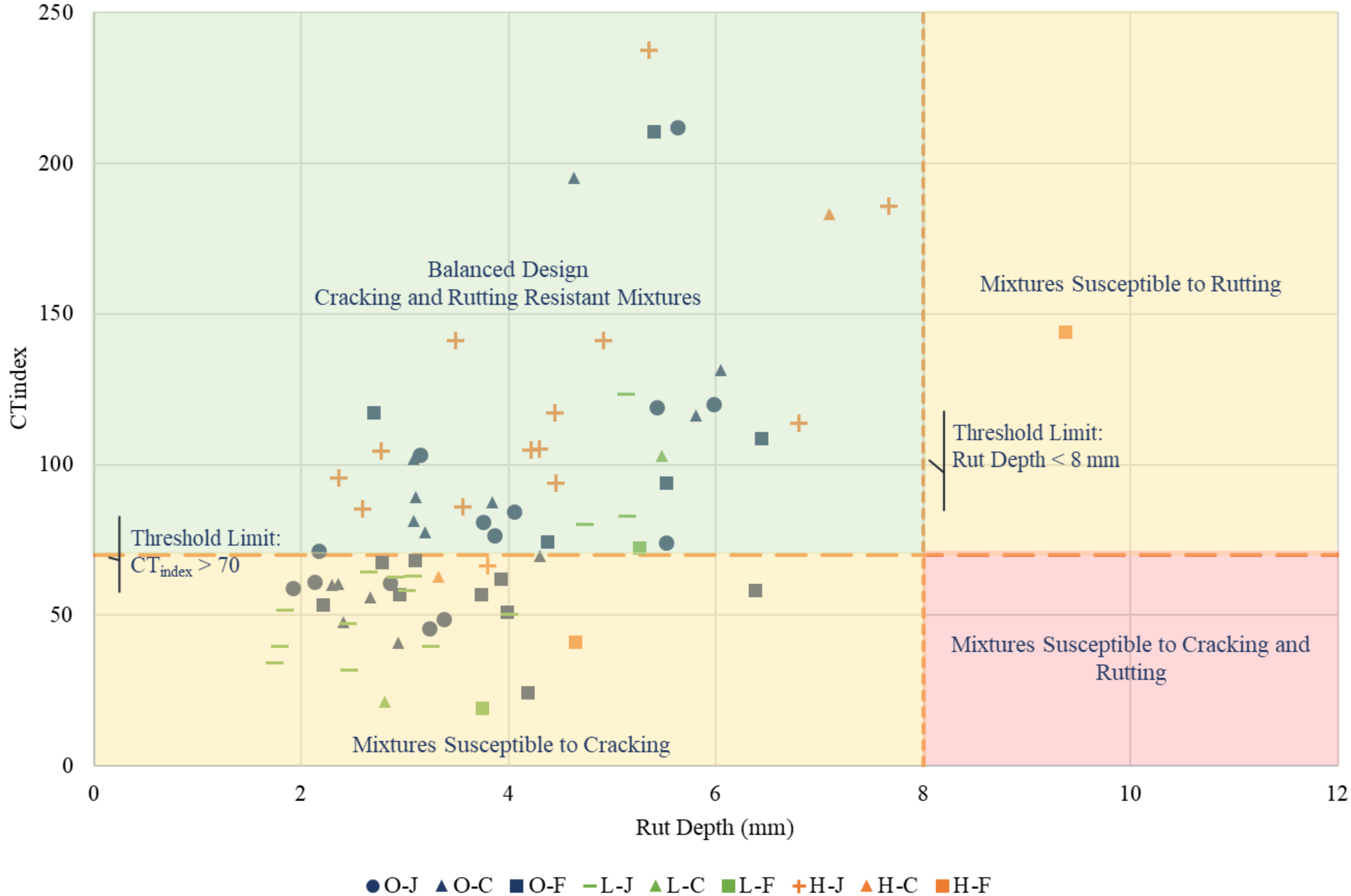

Figure 2. Comparison of Mean Cracking and Rutting Performance for Each Mixture Variation. CT = cracking tolerance; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Cantabro Mass Loss

Figure 3 details the mean Cantabro Mass Loss (ML) for each variation of each mixture. Two mixtures met the threshold limit despite the variations, while three mixtures did not meet the threshold limit for any of the variations. One of these designs was the Superpave design and was not originally intended to meet the threshold criteria. However, the other two were BMD designs. These mixtures fall under Producer A, which had considerable RAP variability as previously discussed. Although the two 40R designs for Producer A met the durability criteria in the design from the producer, the O-J for these designs did not meet specification requirements when mixed and compacted using the raw materials sampled during production. This suggests that material variability, specifically for the RAP, could cause mixtures to become unbalanced.

The Producer A 30R mixture only meets the specification for the high binder content which falls just below the threshold limit, as shown in Figure 4. This design has less aged RAP binder compared to the 40R designs as well as high binder content which resulted in the highest volume of effective binder (Vbe) compared to the other variations. Cox et al. (2017) found that increasing Vbe increased durability by decreasing ML. However, when air voids were held constant and Vbe was further increased, ML also increased. This was a result of the voids in the mineral aggregate (VMA) increasing and causing the gradation to become unstable. The authors also noted that high RAP contents estimate a higher Vbe because it is difficult to properly estimate the absorbed asphalt in the RAP that is not available to contribute to the effective binder. For the 30R design, the less RAP but higher Vbe are likely the cause of the passing result.

The remaining 9 mixtures displayed a blend of results passing and failing threshold requirements. For Producer B, each of the designs exhibited both passing and failing results as shown in Figure 5. The 30R design was not originally intended to meet the specification but showed passing results for the high binder content. Originally designed to meet the durability criteria, the 40R designs did not meet the threshold limit for O-J. Results were slightly higher but similar to the threshold limit, and these deviations are likely a result of between-lab variability.

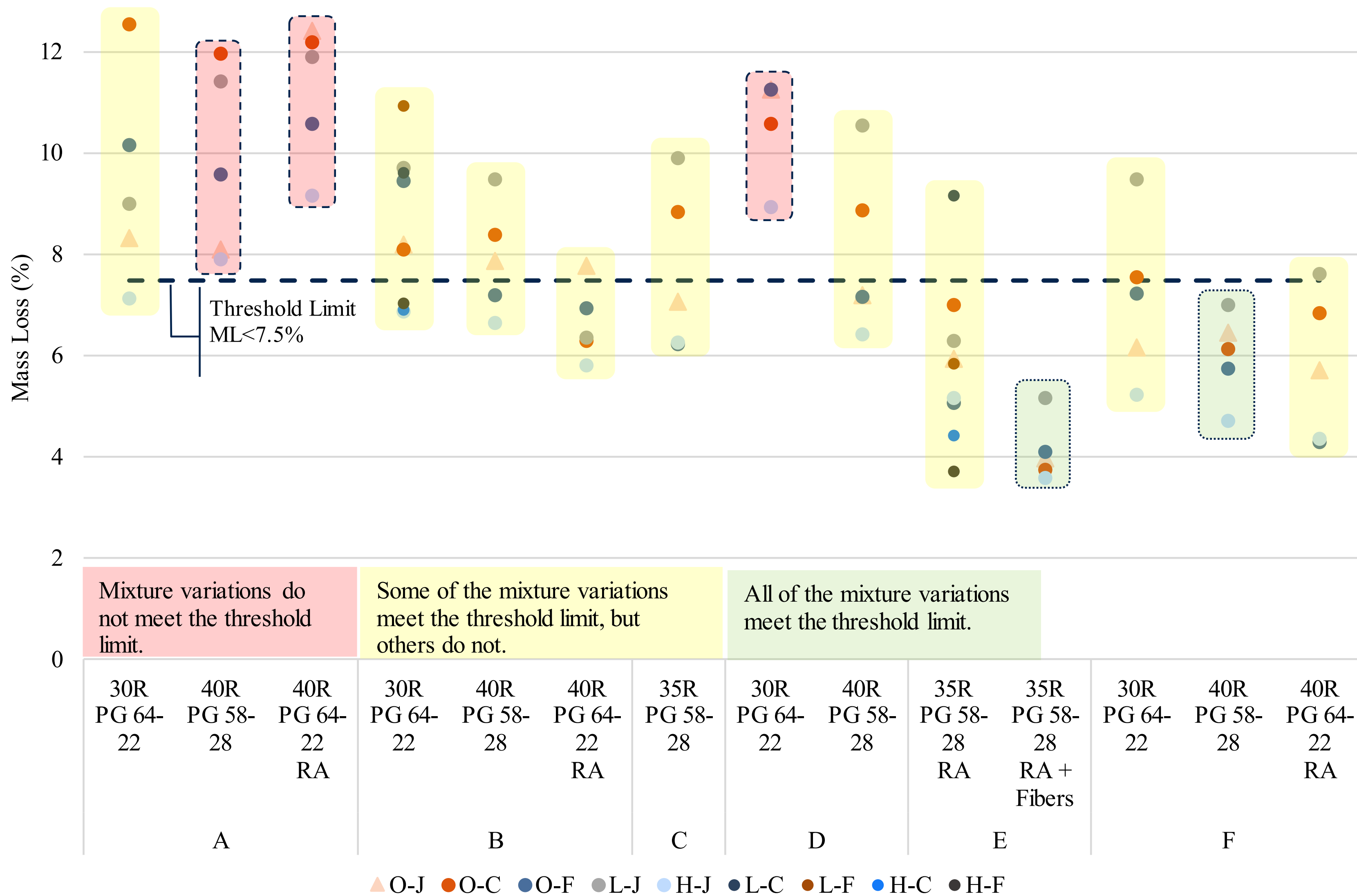

Figure 3. Cantabro Mass Loss Data. PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

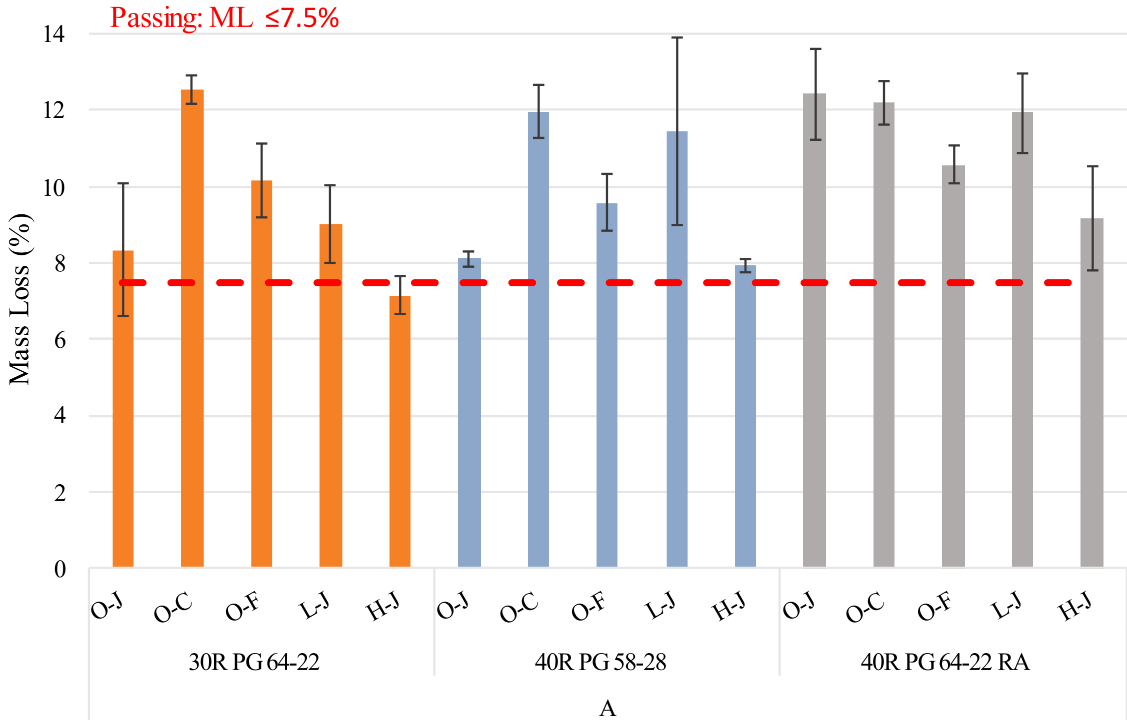

Figure 4. Producer A Mean Mass Loss Results. Error Bars = ± 1 standard deviation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

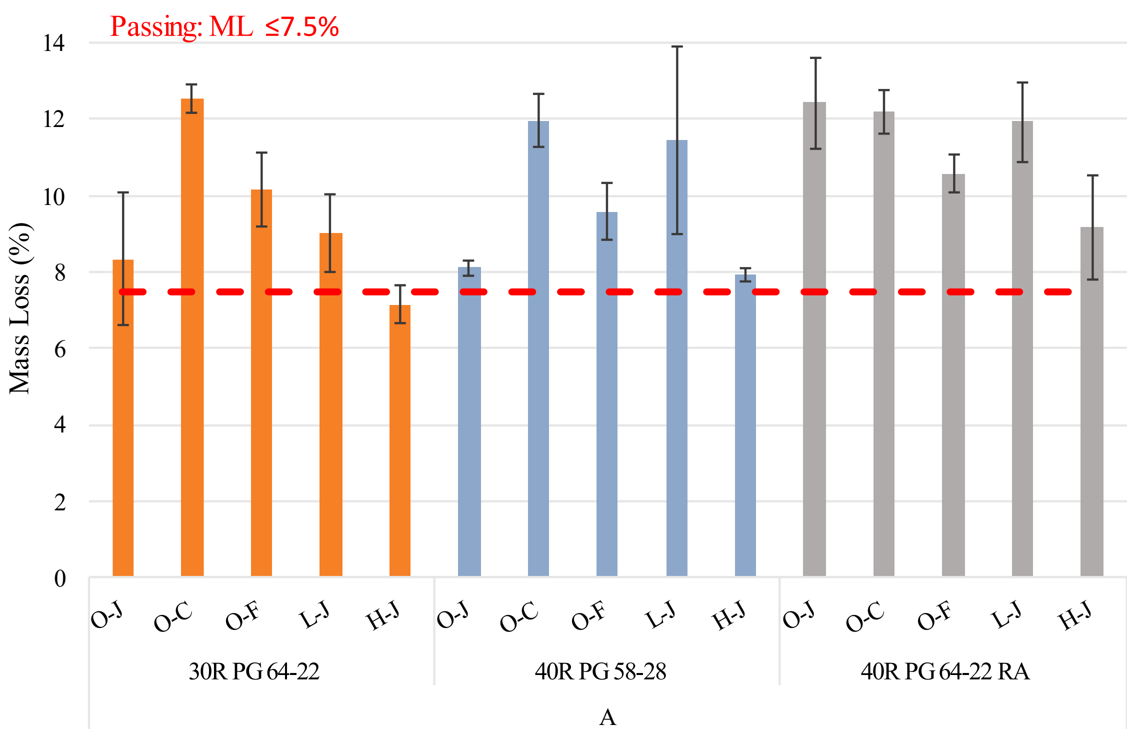

Figure 5. Producer B Mean Mass Loss Results. Error Bars = ± 1 standard deviation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Producers C, D, and F show similar and expected trends where O-J falls below the threshold limit, indicating passing values, but as the mixtures vary according to gradation and binder content, they begin to fall outside of the threshold limit, failing specification criteria. (See Appendix B and Appendix D for results.) One mixture for Producer E met requirements for each variation, while the other mixture for Producer E generally met the criteria except for one interaction variation, L-C, as shown in Figure 6.

Cox et al. (2017) found that lower Vbe, which in many cases is a result of lower binder, decreases durability, and air voids also significantly affect durability with respect to the Cantabro test. A correlation between ML and air voids was also observed in this study through linear regression of the data as illustrated in Figures 7 and 8, which provide an assessment for all of the data as well as for each producer. The overall trend results in an R^2 of 0.42. While some producers resulted in a stronger regression, others did not, namely Producer A. For Producer E, the L-C variation for the 35R PG 58-28 RA mixture resulted in the highest air voids, as shown in Table 7. The failing result is likely a factor of both lower Vbe and higher air voids. Aside from extreme cases, this mixture likely would have stayed balanced with respect to Cantabro since the other variations met the requirements. It is important to note that in current VDOT performance plus volumetric verification procedures the O-J mixtures would fail volumetrics, thus making the mix design unacceptable despite passing the Cantabro test requirement.

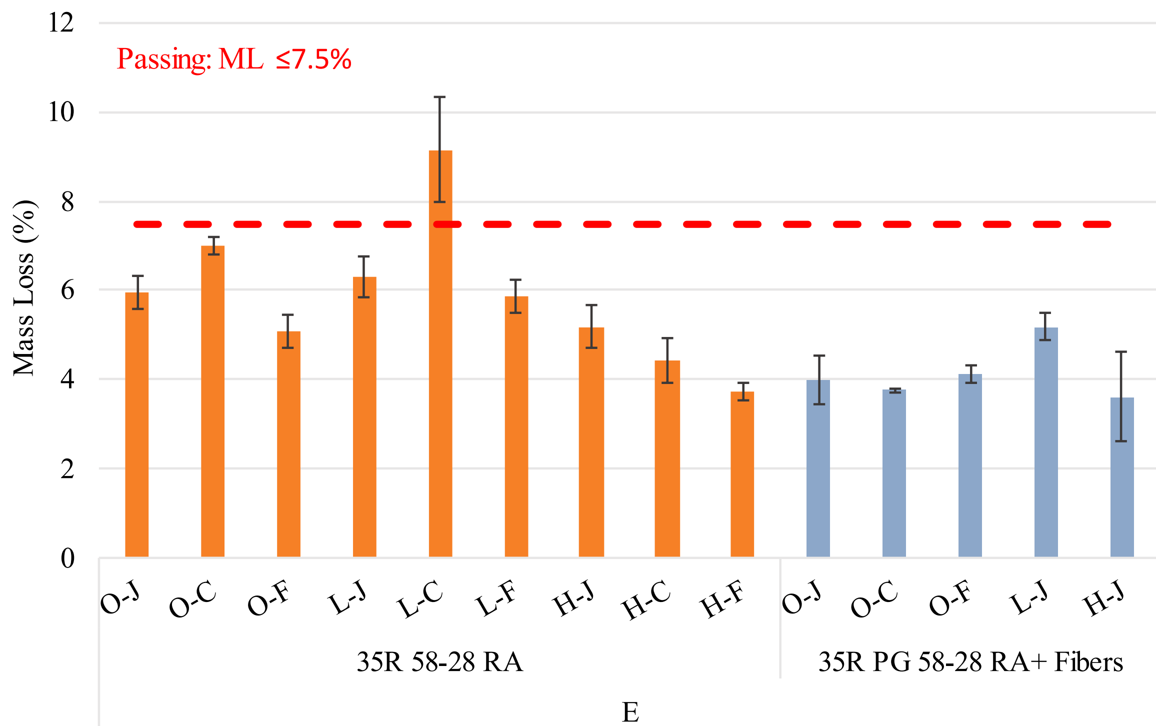

Figure 6. Producer E Mean Mass Loss Results. Error Bars = ± 1 standard deviation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

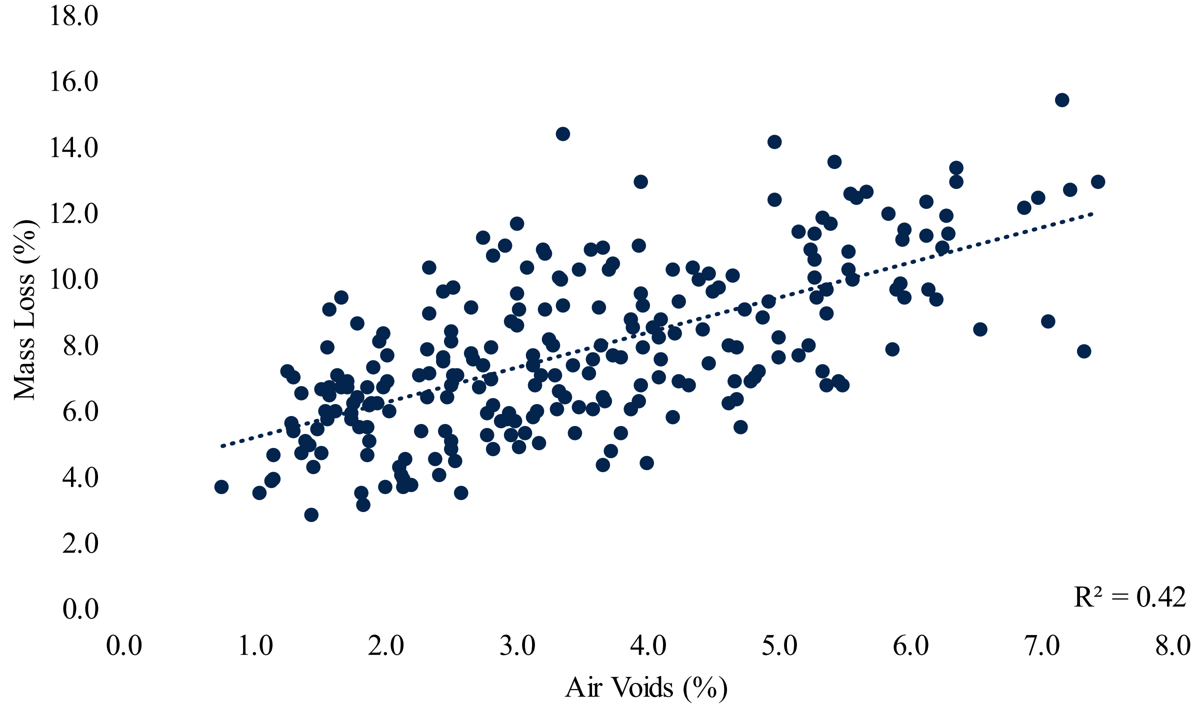

Figure 7. Linear Regression Correlation between Mass Loss Results and Laboratory Compacted Specimens.

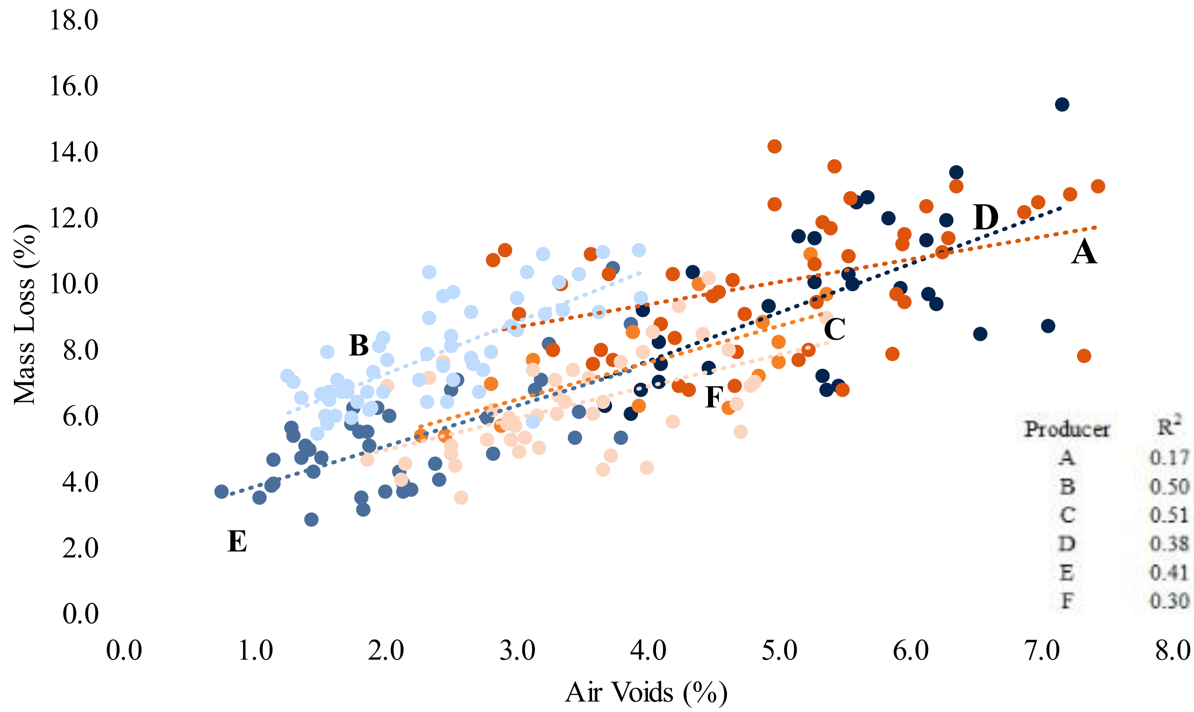

Figure 8. Linear Regression Correlation between Mass Loss Results and Laboratory Compacted Specimen Air Voids by Producer.

Table 7. Mean Air Voids for Cantabro Samples

Producer	Mixture Design	O-J	O-C	O-F	L-J	L-C	L-F	H-J	H-C	H-F
A	30R PG 64-22	5.4	7.1	3.2	6.4			4.6		
	40R PG 58-28	4.7	6.3	4.3	5.2			3.6		
	40R PG 64-22 RA	5.4	5.9	3.2	5.7			4.3		
B	30R PG 64-22	2.6	2.6	2.4	3.3	3.5	2.9	1.7	2.0	1.9
	40R PG 58-28	2.8	2.9	1.6	3.4			1.7		
	40R PG 64-22 RA	2.2	1.6	1.3	2.8			1.6		
C	35R PG 58-28	4.6	4.7	3.3	4.8			2.8		
D	30R PG 64-22	5.6	5.6	5.1	6.4			4.3		
	40R PG 58-28	4.5	6.6	5.1	6.0			3.8		
E	35R PG 58-28 RA	1.9	2.7	1.3	3.1	3.6	1.8	1.5	1.3	1.0
	35R PG 58-28 RA + Fibers	2.1	2.1	2.2	3.4			1.6		
F	30R PG 64-22	3.4	4.1	2.3	4.7			3.0		
	40R PG 58-28	3.2	4.6	3.3	3.9			2.3		
	40R PG 64-22 RA	2.9	3.2	3.3	4.4			2.5		
	Count	**0**	**6**	**0**	**8**	**2**	**0**	**0**	**0**	**0**

Red circles indicate high air voids (within the top 5% for the mixture). Green circles indicate low air voids (lowest 5% of air voids for the mixture). Yellow circles are values in between. PG = performance grade; RA = recycling agent. O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Table 8 summarizes the analysis results for Cantabro mass loss. From Equation 1, negative values for percent change indicate that the test results for the variation were less than the O-J values, which is characteristic of more durable mixtures. However, positive values for percent change indicate that the change in gradation or binder content resulted in a greater mass loss than O-J, hindering durability. Percent change was averaged across the 14 mixtures to generalize trends, and individual values for each design and variation are included in Appendix B. The average percent change for each variation was listed for Producer B 30R and Producer E 35R in Table 8 to evaluate the influence of interaction variations on performance results. The results were not averaged to further evaluate if consistent trends emerged between designs or if results were producer specific.

Table 8. Analysis Results Summary for Cantabro Mass Loss

Mixture Variation	Avg. % Change for 14 Designs	B 30R PG 64-22	E 35R PG 58-28 RA	No. of Mixtures Statistically Significant
O-C	12%	-1%	18%	5
O-F	-1%	15%	-15%	0
L-J	22%	18%	6%	6
H-J	-16%	-16%	-13%	2
L-F	-	33%	-2%	1
L-C	-	17%	54%	1
H-C	-	-16%	-26%	1
H-F	-	-14%	-38%	1

Negative ("-") for percent chan e indicates a reduction in mass loss compared to the design or an increase in durability. Positive values indicate an increase in mass loss compared to the design and indicate a decrease in durability. PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Low binder content negatively influenced durability, increasing ML 22% on average, while high binder content improved durability reducing ML by an average of 16%. With respect to gradation, the coarse gradation reduced durability by increasing the ML by 12%, where the fine gradation showed only a minor improvement on durability but largely showed consistent results to O-J. Higher air voids have been found to reduced durability, and this study found that often the coarse gradation and low binder content had higher air voids supporting findings in the literature (Cox et al., 2017).

When evaluating the interaction mixtures, trends appear to be mixture specific and may suggest more intricate dependencies upon volumetric parameters such as Vbe or dust-to-binder ratio (F/A). For B 30R, the greatest mass loss corresponded to L-F which reduced durability by increasing ML 33%. L-F had a lower Vbe and also had the highest F/A ratio. L-C also had lower Vbe but the F/A ratio was not as high as L-F since the coarse gradation resulted in 2% less dust than the fine gradation. As a result, L-C increased ML 17%, which is considerably less than L-F (33%).

Based on the literature, it could be expected that L-C would result in the greatest ML since it has the lowest binder content and generally the highest air voids. While this was not the case for Producer B 30R, this tendency was observed for Producer E 35R where ML increased 54%. For this mixture, L-F slightly improved durability but was not statistically different from the design. High binder content improved durability for both coarse and fine gradations for both mixtures. The greatest reduction in mass loss among these combinations was for H-F and suggests that the high dust content coupled with the high binder could act as a mastic extender benefitting durability.

An ANOVA was conducted to assess the statistical similarities and differences between means, along with Dunnett's test if the variations mean(s) were found to e statistically different. P-values for the statistical analysis for ML are listed in Table 9. The number of mixtures that produced a significant result were counted for each of the respective variations and summarized in Table 8. For binder content, low binder was significant for 6 of the 14 mixtures (43%), whereas high binder was only significant for 2. Similar trends observed for percent change for gradation were also observed for statistical significance. The fine gradation did not produce a considerable change from O-J values and was also not statistically significant. The coarse gradation had a considerable percent change and was also statistically significant for 5 mixtures. With respect to the interaction mixtures, each variation was found to be statistically significant for one of the two designs.

Cantabro results were sensitive to changes in binder content causing mass loss to improve as binder content was increased and worsen as binder content was decreased. Based on the literature, the Cantabro test is sensitive to changes in Vbe and air voids, which are largely a function of the VMA. From the interaction mixtures, it appears that the influence of these parameters can dictate the influence on ML results causing some variations to show improved results while other mixtures show a reduction in durability.

Table 9: P-values for ANOVA and Dunnett’s Test for Cantabro Results

Producer	Design	ANOVA p-value	Dunnett's Method p-value							
			O-C	O-F	L-J	H-J	L-C	L-F	H-C	H-F
A	30R PG 64-22	0.001	0.001	0.181	0.868	0.514	-	-	-	-
	40R PG 58-28	0.005	0.009	0.409	0.022	0.999	-	-	-	-
	40R PG 64-22 RA	0.012	0.995	0.139	0.927	0.008	-	-	-	-
B	30R PG 64-22	0.000	1.000	0.266	0.137	0.224	0.171	0.009	0.255	0.340
	40R PG 58-28	0.111	-	-	-	-	-	-	-	-
	40R PG 64-22 RA	0.008	0.019	0.207	0.024	0.003	-	-	-	-
C	35R PG 58-28	0.002	0.109	0.657	0.011	0.694	-	-	-	-
D	30R PG 64-22	0.003	0.830	1.000	0.039	0.064	-	-	-	-
	40R PG 58-28	0.000	0.029	1.000	0.000	0.426	-	-	-	-
E	35R PG 58-28 RA	0.000	0.139	0.269	0.955	0.396	0.000	1.000	0.016	0.001
	35R PG 58-28 Softening Oil + Fibers	0.032	0.958	0.993	0.059	0.817	-	-	-	-
F	30R PG 64-22	0.000	0.032	0.103	0.000	0.155	-	-	-	-
	40R PG 58-28	0.084	-	-	-	-	-	-	-	-
	40R PG 64-22 RA	0.002	0.333	0.181	0.056	0.208	-	-	-	-

Blank ce ls indicate Dunnett’s test was not conducted ecause the ANOVA found means were not statistically different. Mixtures including 5 variations had 9 degrees of freedom. Producer B 30R PG 64-22 and Producer E 35R PG 58-28 RA were interaction mixtures which included 9 variations and therefore had 17 degrees of freedom. PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

APA Rut Depth

As previously mentioned, APA results showed excellent resistance to rutting despite changes in gradation and binder content. Figure 9 and Figure 10 provide the APA results, and data can be found in Appendix B and Appendix E. The only test that did not meet threshold requirements was the interaction H-F for one mixture (35R PG58-28 RA from producer E). Other studies have found that both high binder and high dust contents negatively influence rutting resistance (Austerman et al., 2018; Mogawer et al., 2019). While the second interaction mix (30R PG64-22 from producer B) had passing results, the H-F variation also had the most rutting.

Between mixture designs, the balanced designs showed similar results to the 30R Superpave counterpart. A balanced design for Producer B was the only mixture to show generally higher rut depths for 40R PG 58-28 compared to the two other designs. Figure 9 shows a consistent spread in data that encapsulates a central point or the variations cluster around the extremes. The higher binder content is most often the highest rutting, whereas the lower binder content typically results in the least rutting. Fine mixes tend to rut more, whereas coarse mixes tend to rut less. This indicates that there is a connection between rut depth and volumetrics, though in nearly all cases the mix design stays balanced regardless of changes to binder content or gradation. However, data for Producer E seems to cluster together except for one point that is considerably higher than the rest, which is shown in Figure 10. Most of the variations appear to oscillate near the same value with one variation, either H-F or H-J, which is considerably higher. This further suggests a relationship between rut depth and volumetric parameters that cause rut depth to increase as the interaction between gradation and binder content changes.

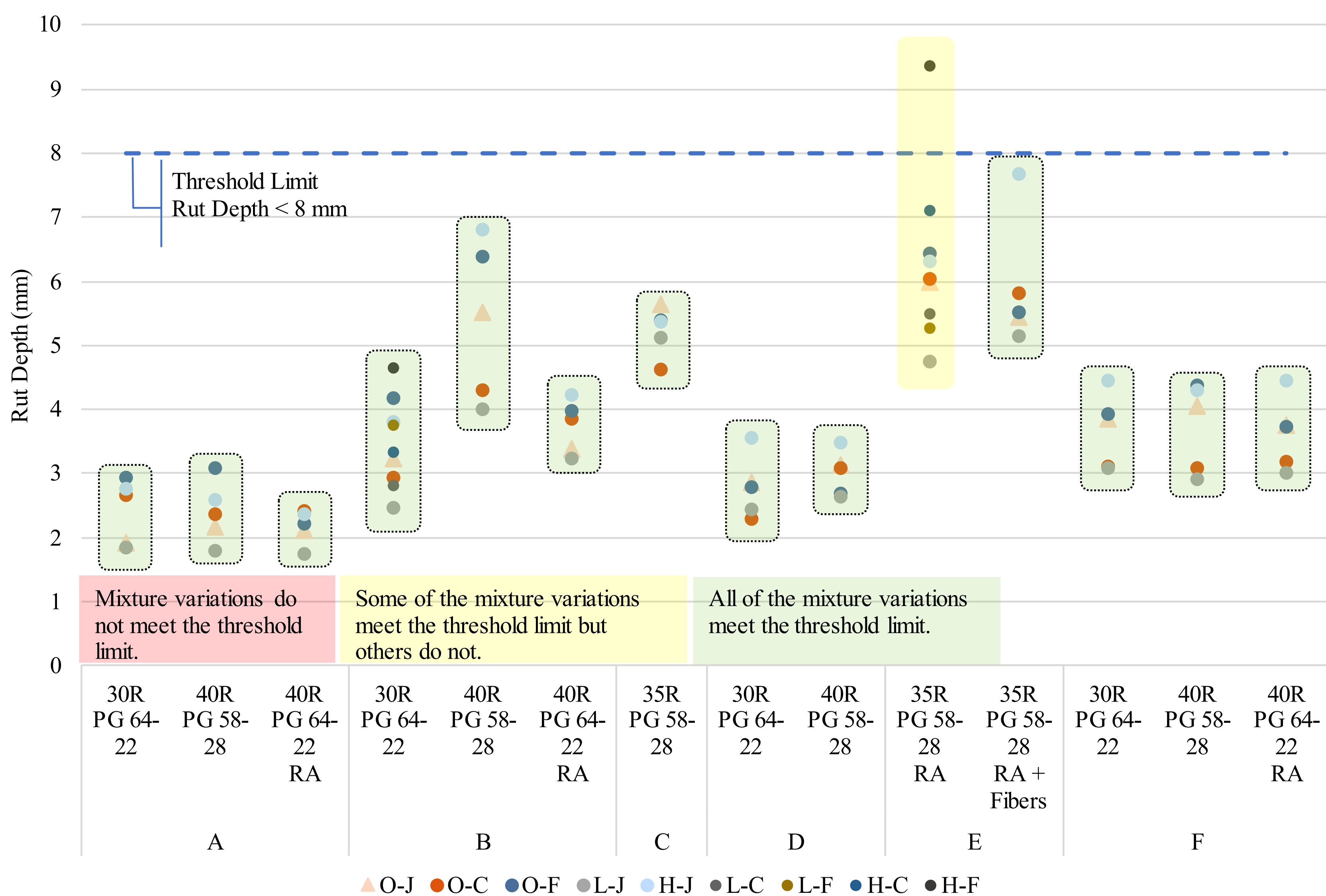

Figure 9. APA Rut Depth Data. PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

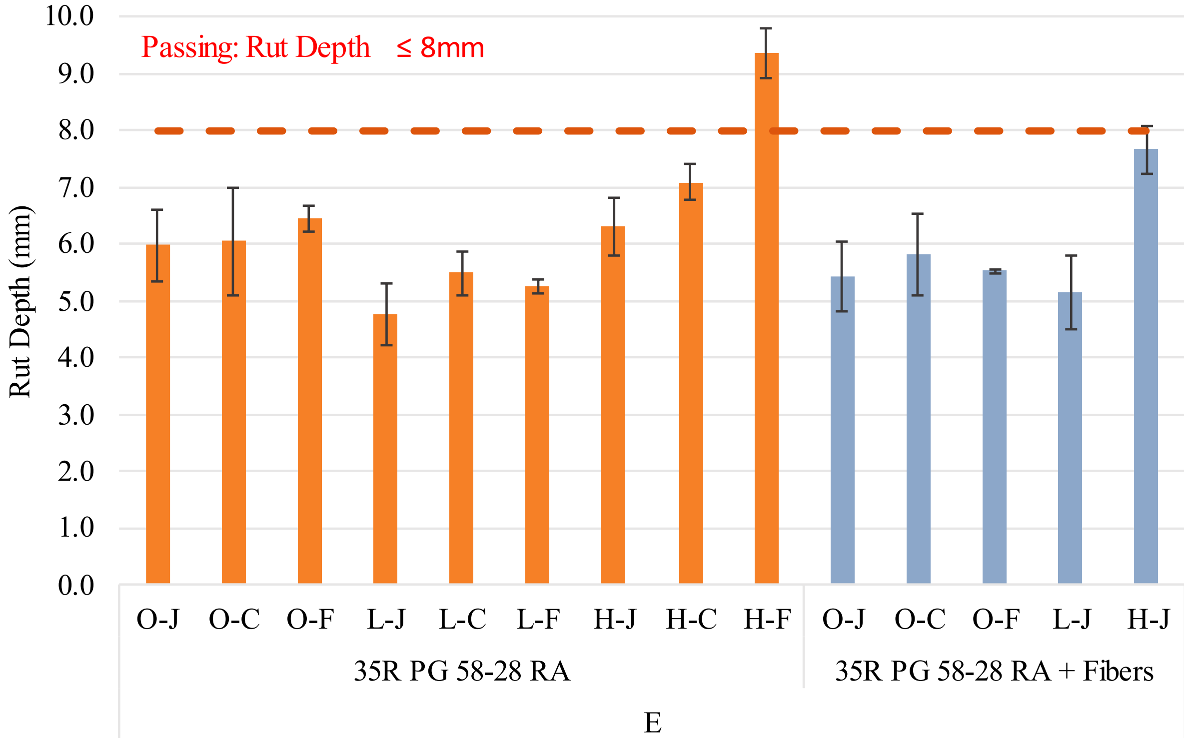

Figure 10. Average Rut Depth for Producer E. Error Bars = ± 1 standard deviation. PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Table 10 includes percent change and statistical analysis results for APA. Like the Cantabro test, an increase in percent change indicates an increase in rut depth compared to the design, resulting in a mixture that is more rut susceptible. Negative values indicate a rut depth less than the O-J indicating a more rut resistant mixture variation. Percent change was calculated according to Equation 1 and found that binder content influenced results the most. On average, rut depth decreased 16% for the low variation but increased 18% for the high variation. The fine gradation also had a negative influence on rutting resistance by increasing rut depths by 11% compared to the design. However, the coarse gradation improved rutting resistance by 3%.

When evaluating the interaction mixtures, the H-F variation increased rut depths by 43% and 57% for Producer B and Producer E, respectively. By combining both factors that have been observed to influence rutting resistance, high binder and high dust content, considerably higher rut depths occurred. Low binder content improved rutting resistance, and coupled with the coarse gradation, the change in rut depth is similar to the L-J (-16%) and O-C (-3%) variations. However, when evaluating the percent change with respect to the L-F variation conflicting results are observed between the two producers. For Producer B, L-F hinders rutting resistance, whereas for Producer E it increases rutting resistance by resulting in a rut depth less than the design. This suggests that there may be volumetric parameters influencing results. A low binder content would typically have a considerable effect on rut resistance, but when the gradation changes that influence is reduced or changed all together.

Table 10. Analysis Results Summary for APA Rut Depth

Mixture Variation	Avg. % Change for 14 Designs	B 30R PG 64-22	E 35R PG 58-28 RA	No. of Mixtures Statistically Significant
O-C	-3%	-9%	1%	0
O-F	11%	29%	8%	1
L-J	-16%	-24%	-21%	0
H-J	18%	17%	6%	0
L-F	-	16%	-12%	0
L-C	-	-13%	-8%	0
H-C	-	3%	19%	0
H-F	-	43%	57%	2

Negative ("-") for percent chan e indicates a reduction in rut depth compared to the design, or an increase in rutting resistance. Positive values indicate an increase in rut depth compared to the design and indicate a decrease in rutting resistance. PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Table 10 includes the statistical analysis results for the ANOVA and Dunnett's tests. If the ANOVA test failed to reject the null hypothesis, the Dunnett's test was not conducted. Generally, the mixtures did not show statistical significance for the variations likely due to limited sample size. Production practices were replicated in this study where only two tracks were tested for each test set. Based on limited replicates and the inherent testing variability, the statistical tests are likely not sensitive or robust enough to detect changes in gradation and binder content. While this is the case, the tests did determine statistical significance for H-F, which had resulted in the greatest deviations from design. Based on this observation, the percent change, or other methods for quantifying differences in the data, may be more effective in detecting deviations from design due to the low replicate size.

Table 11. P-values for ANOVA and Dunnett's Test for APA Results

Producer	Design	ANOVA p-value	Dunnett's Method p-value O-C	O-F	L-J	H-J	L-C	L-F	H-C	H-F
A	30R PG 64-22	0.027	0.107	0.035	0.995	0.07	-	-	-	-
	40R PG 58-28	0.071	-	-	-	-	-	-	-	-
	40R PG 64-22 RA	0.549	-	-	-	-	-	-	-	-
B	30R PG 64-22	0.004	0.942	0.134	0.265	0.545	0.779	0.643	1.000	0.021
	40R PG 58-28	0.044	0.386	0.633	0.245	0.347	-	-	-	-
	40RPG 64-22 RA	0.678	-	-	-	-	-	-	-	-
C	35RPG 58-28	0.780	-	-	-	-	-	-	-	-
D	30R PG 64-22	0.617	-	-	-	-	-	-	-	-
	40R PG 58-28	0.636	-	-	-	-	-	-	-	-
E	35R PG 58-28 RA	0.002	1.000	0.967	0.337	0.995	0.952	0.796	0.420	0.003
	35R PG 58-28 Softening Oil + Fibers	0.064	-	-	-	-	-	-	-	-
F	30R PG 64-22	0.034	0.178	0.999	0.165	0.337	-	-	-	-
	40R PG 58-28	0.024	0.118	0.786	0.066	0.89	-	-	-	-
	40R PG 64-22 RA	0.175	-	-	-	-	-	-	-	-

Mixtures including 5 variations had 9 degrees of freedom. Producer B 30R PG 64-22 and Producer E 35R PG 58-28 RA were interaction mixtures that included 9 variations and therefore had 17 degrees of freedom. PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

IDT-CT

Figure 11 illustrates the IDT-CT data for each of the variations of the mixtures in this study. Only one mixture did not meet the threshold limit for each of the variations assessed as shown in Figure 12. However, this mixture was a 30R Superpave design that was not originally designed to meet BMD criteria. The other three Superpave designs produced some variations that did meet the threshold limit. Despite varying both gradation and binder content, three mixtures met the threshold limit in all cases. The remaining 10 mixtures, a combination of Superpave and balanced designs, produced results that met the specification for some variations and some that did not.

When evaluating data based on producer, the balanced designs yielded similar CT_{index} results indicating that designs using a RA appear to be equivalent to using a softer binder grade (PG 58-28). CT_{index} results are detailed further in Appendix B and Appendix F

.

When designs were recreated in the lab (O-J), Producer A and Producer B fell below the threshold limit despite being balanced designs. An example of this is shown in Figure 12. These designs were the RA modified mixtures, and the material variability, specifically for the RAP, may be influencing the performance of the mixture. The purpose and goal of RAs is to reduce the influence of aging of the RAP binder regarding cracking resistance and durability. If variability in the RAP causes an increase in RAP AC compared to the design, the produced mixture will have more aged binder, and the RA dosage determined in design may not be sufficient in reviving the RAP. RAP AC for Producer B was 1.1% higher than the design as stated in the 'Material Processing Cha enges' section. In contrast, Producer resulted in passing results for each of the O-J variations including the RA modified mixture as shown in Figure 14. For Producer F, RAP AC was 0.2% less than the design. This indicates that more RA was supplied to the produced mixture than in the design. Material variability was accounted for in the laboratory from a quantitative perspective, but designs were not redesigned with respect to additives and RAs. The material variability for the RAP could be reducing the positive influence of RAs on CT_{index} results.

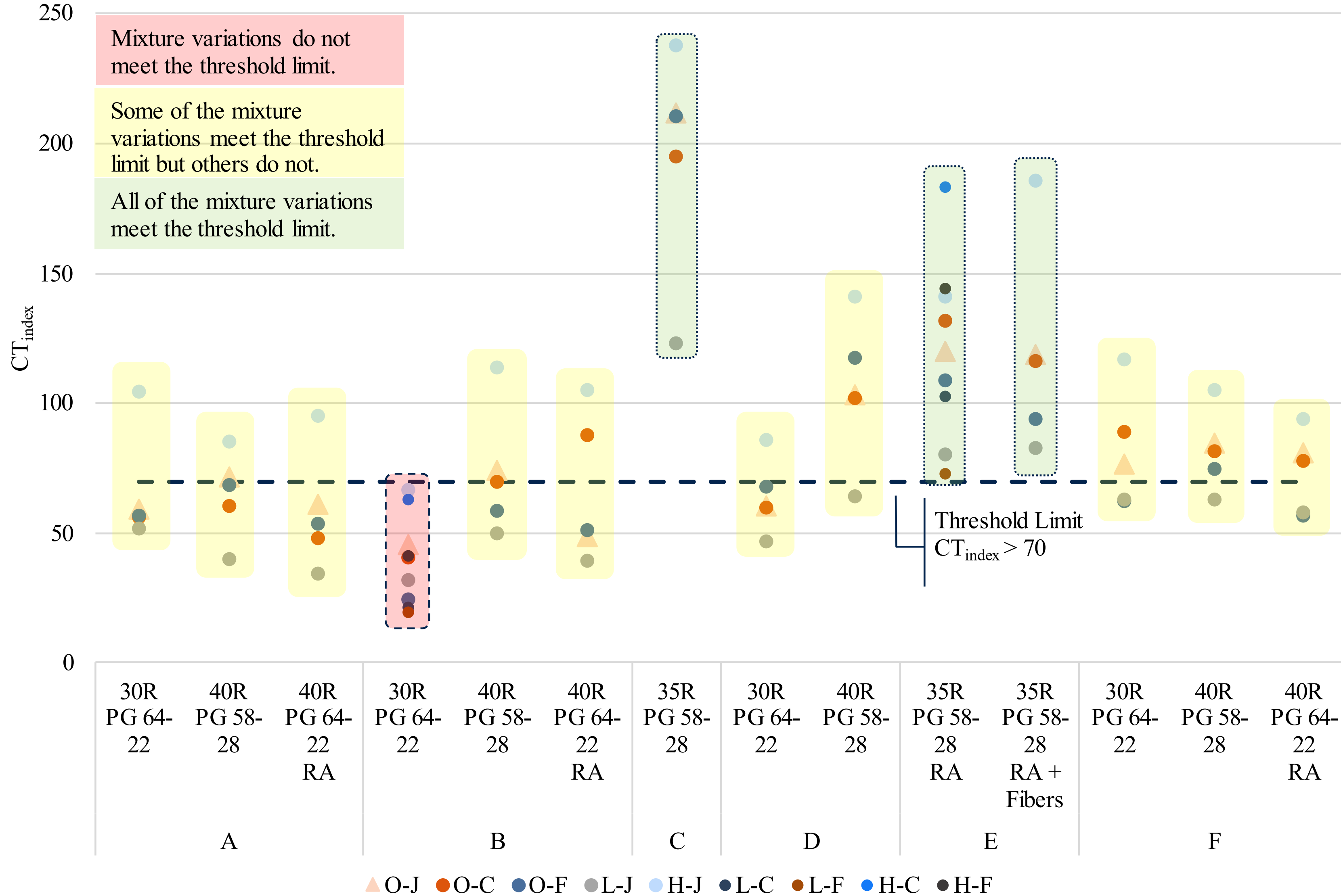

Figure 11. IDT-CT Data. CT = cracking tolerance; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Figure 12 depicts IDT-CT results for Producer B, which compares a Superpave mixture, evaluating the interaction of changes in gradation and binder content, as well as two balanced mixtures. For the interaction mixture, clustering of the data with respect to binder content was observed. Results descended from high to low binder where high binder content showed the best cracking resistance. Based on gradation, a general trend appeared where CT_{index} descends between O-J, O-C, and O-F which is also mirrored in the 40R PG 58-28 design. However, this relationship is not observed for the 40R PG 64-22 RA design. For this design, the coarse gradation results exceed the fine gradation. The trend of coarse gradation improving results more than the fine gradation was also observed for Producer E as shown in Figure 13. The IDT-CT may not be sensitive to changes in gradation, aligning with the findings of Mogawer et al. (2019), which could cause general trends in results with respect to gradation to change. Results may also be influenced by the combination of gradation and binder content within the volumetric parameters rather than strictly the gradation, or by using rejuvenators.

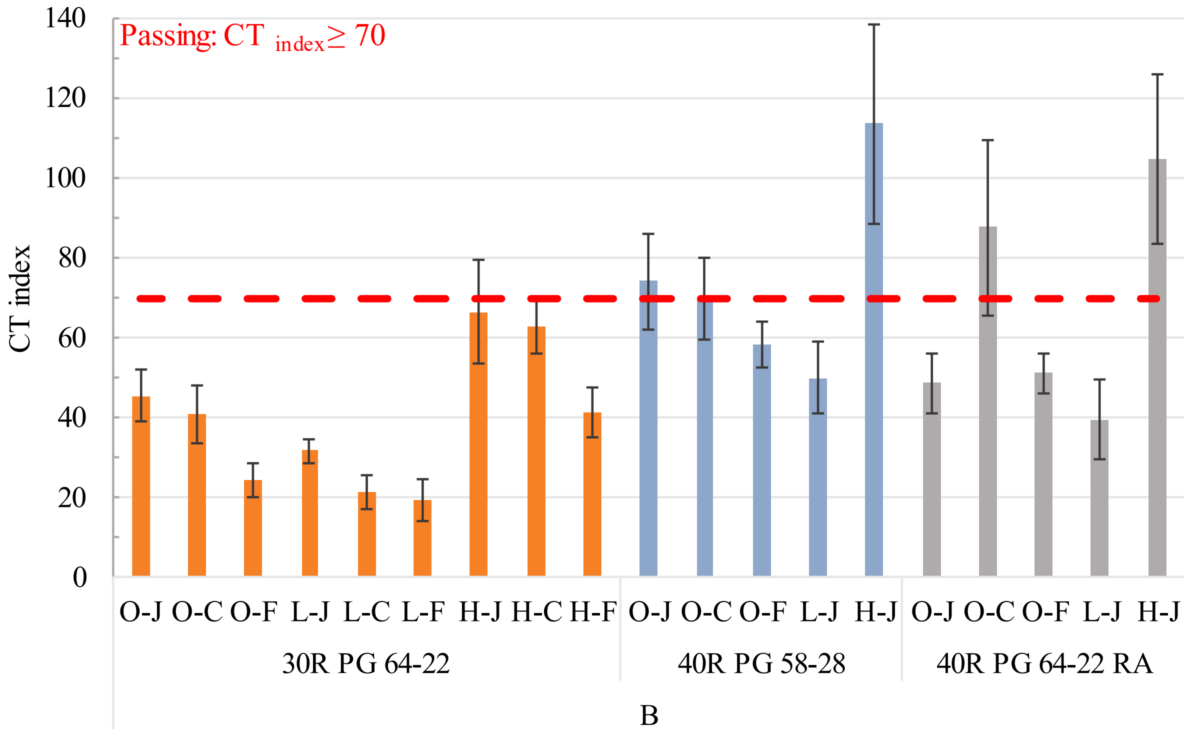

Figure 12. Producer B Mean CT_{index} Results. Error Bars = ± 1 standard deviation; CT = cracking tolerance; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

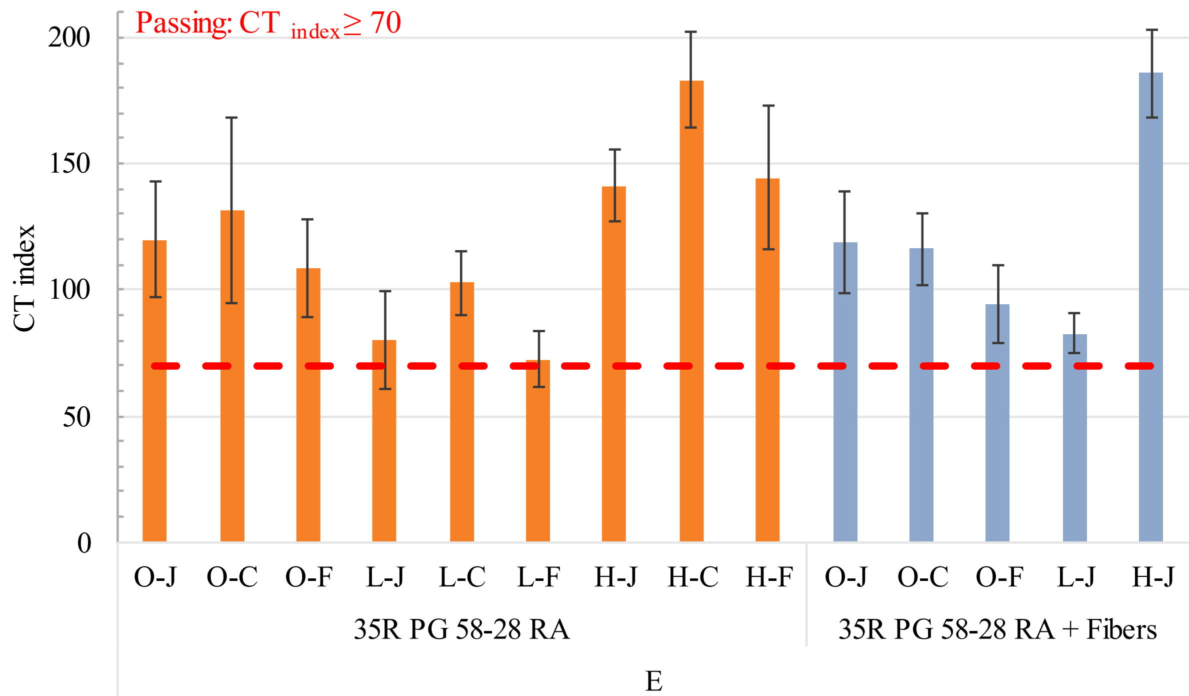

Figure 13. Producer E Mean CT_{index} Results. Error Bars = ± 1 standard deviation; CT = cracking tolerance; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

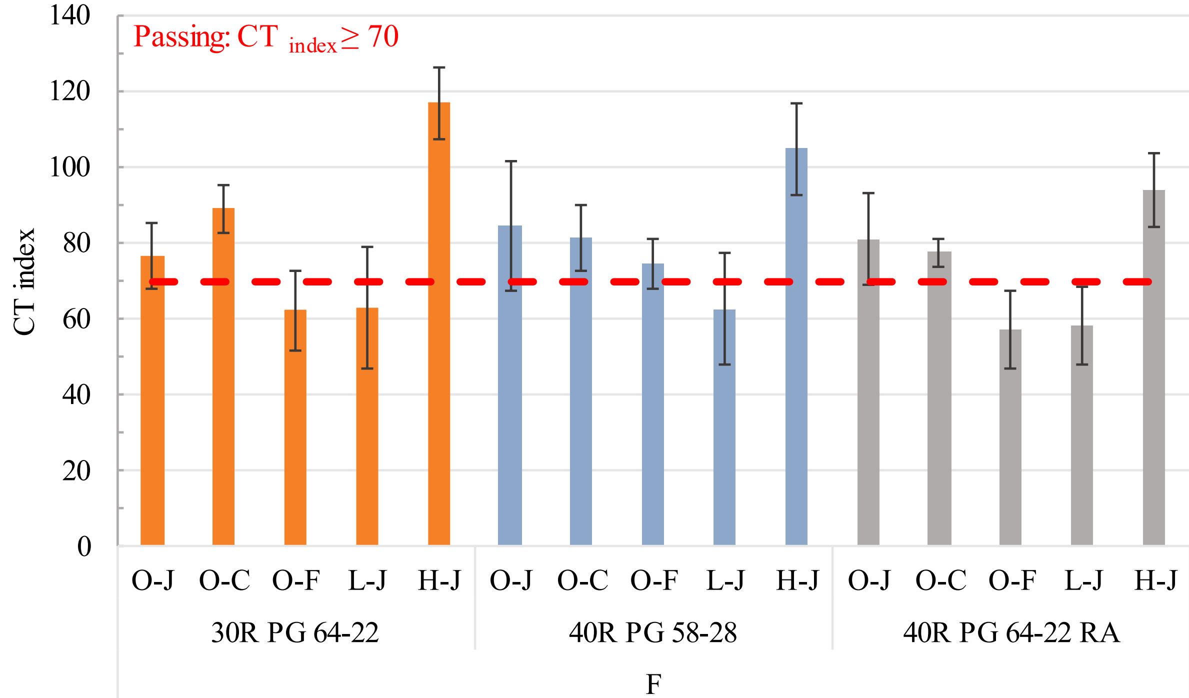

Figure 14. Producer F Mean CT_{index} Results. Error Bars = ± 1 standard deviation; CT = cracking tolerance; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Table 12 summarizes the IDT-CT analysis results, and Table 13 provides p-values for the ANOV and Dunnett's test. Percent change was calculated according to Equation 1. Positive values indicate that the variation had a higher CT_{index} compared to O-J indicating that those variations are more resistant to cracking compared to the original design. Negative values for average percent change indicate that the CT_{index} of the variation was lower than the respective design and is more susceptible to cracking. Changes in binder content show the greatest sensitivity on CT_{index} where high binder content improved cracking resistance by 45%, on average, but the low binder content hindered cracking resistance by 30%. Gradation was not found to influence results as considerably as binder content. On average, the coarse gradation improved results by 2%, where the fine gradation lowered CT_{index} by 11%.

Conversely, the interaction mixtures incorporating changes in binder content as well as gradation seem to have an increased effect on results. L-J and O-F both negatively influenced CT_{index} decreasing it by 30% and 11%, respectively. The combination of these two conditions yielded a CT_{index} decrease of 57% and 40% for Producer B and Producer E, respectively. Similarly, high binder content improves cracking resistance, and on average increases the CT_{index} by 45%. However, combined with a fine gradation, H-F, opposite trends are observed based on producer. For Producer E, H-F improves results, while it hinders cracking resistance for Producer B. In contrast, the H-C variation improves results for both producers, and Producer E shows a considerable increase in CT_{index} of 53%.

Based on the results of the interaction mixtures, there may be underlying influences of volumetrics on CT_{index}, such as effective binder content. A relationship between effective binder content and CT_{index} could explain the trends observed between H-F and H-C variations. H-C has a higher effective binder content, due to less fines within the structure, which has been known to improve cracking resistance.

Table 12. Analysis Results Summary for CT_{index}

Mixture Variation	Avg. % Change for 14 Designs	B 30R PG 64-22	E 35R PG 58-28 RA	No. of Mixtures Statistically Significant
O-C	2%	-10%	10%	1
O-F	-11%	-47%	-9%	2
L-J	-30%	-30%	-33%	9
H-J	45%	46%	18%	10
L-F	-	-57%	-40%	2
L-C	-	-53%	-14%	1
H-C	-	38%	53%	2
H-F	-	-10%	20%	0

Negative percent change indicates a reduction in CT_{index} compared to the design, or a decrease in cracking resistance. Positive values indicate an increase in CT_{index} compared to the design and indicate an increase in cracking resistance. "-" com ination not e aluated; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

Table 13. P-values for ANOVA and Dunnett’s Test for IDT-CT Results

Producer	Design	ANOVA p-value	Dunnett's Method p-value							
			O-C	O-F	L-J	H-J	L-C	L-F	H-C	H-F
A	30R PG 64-22	0.000	0.961	0.989	0.617	0.000	-	-	-	-
	40R PG 58-28	0.000	0.244	0.963	0.000	0.123	-	-	-	-
	40R PG 64-22 RA	0.000	0.278	0.735	0.007	0.001	-	-	-	-
B	30R PG 64-22	0.000	0.734	0.000	0.005	0.000	0.000	0.000	0.000	0.798
	40R PG 58-28	0.000	0.963	0.283	0.051	0.001	-	-	-	-
	40R PG 64-22 RA	0.000	0.000	0.996	0.714	0.000	-	-	-	-
C	35R PG 58-28	0.000	0.749	1.000	0.000	0.375	-	-	-	-
D	30R PG 64-22	0.000	1.000	0.660	0.138	0.002	-	-	-	-
	40R PG 58-28	0.000	0.999	0.111	0.000	0.000	-	-	-	-
E	35R PG 58-28 RA	0.000	0.926	0.896	0.007	0.323	0.757	0.001	0.000	0.178
	35R PG 58-28 Softening Oil + Fibers	0.000	0.996	0.051	0.005	0.000	-	-	-	-
F	30R PG 64-22	0.000	0.107	0.075	0.061	0.000	-	-	-	-
	40R PG 58-28	0.000	0.966	0.385	0.009	0.021	-	-	-	-
	40R PG 64-22 RA	0.000	0.931	0.000	0.001	0.062	-	-	-	-

Mixtures including 5 variations had 9 degrees of freedom. Producer B 30R PG 64-22 and Producer E 35R PG 58-28 RA were interaction mixtures that included 9 variations and therefore had 17 degrees of freedom. “-“ com ination not evaluated; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

The variations for gradation and binder content yield considerable differences for percent change as well as statistical significance. Binder content was found to significantly influence CT_{index} results for the majority of mixtures. H-J was statistically significant for 10 mixture designs according to Dunnett's test, and low binder content was significant for 9. O-C and O-F were only significant for 1 and 2 mixtures, respectively, and reflects many of the conclusions made based on the percent change discussion as well as in the study conducted by Mogawer et al. (2019). For the interaction mixtures, L-F and H-C were statistically significant for each mixture evaluated, while L-C was significant for one, and H-F did not significantly influence results. Similar trends emerged from the percent change and the statistical significance analysis, concluding that changes in binder content is most significant to CT_{index} results. The fine gradation could have a minor influence on results, but the interaction between gradation and binder content suggests that changes in volumetric properties may be influencing CT_{index} results.

IDT-CT for Critically Aged Specimens

Figure 15 and Figure 16 display CT_{index} results for loose-mix, long-term oven aged (8 hours at 1 5°C) samples compared to the short-term aged (4 hours at the compaction temperature) counterparts. The critically aged condition produces results that are roughly half the CT_{index} of samples that are short-term aged, as detailed in Table 14. From Figure 15, H-J appears to have the greatest rate of reduction with respect to CT_{index}, which is confirmed in Table 14 since it produces the greatest percent difference of 62% as calculated according to Equation 2. Aging significantly reduces CT_{index} indicated by p-values less than 0.05 (Table 14) which was determined by a two-sample unrelated means t-test. Critical aging accelerates the embrittlement of the binder resulting in significantly lower CT_{index} values which was also determined by Chen et al. (Chen et al., 2020).

$$\textbf{Percent Difference} = \frac{\text{CTindex}_{\text{short-term aging}} - \text{CTindex}_{\text{long-term aging}}}{\text{CTindex}_{\text{short-term aging}}} * 100 \qquad \textbf{[Eq.2]}$$

Table 14. Critical Aging (135°C for 8hrs) Results Compared to Short-term Aged Results

Producer	Variation	Aging, 135°C for 8hrs	Short-term aging	Percent Difference, %	t-test, p-value
B	O-J	33.8	74.2	54%	0.000
	O-C	41.0	69.9	41%	0.000
	O-F	33.3	58.5	43%	0.001
	L-J	21.5	50.1	57%	0.000
	H-J	43.2	113.7	62%	0.000
D	O-J	51.2	103.2	50%	0.000

O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

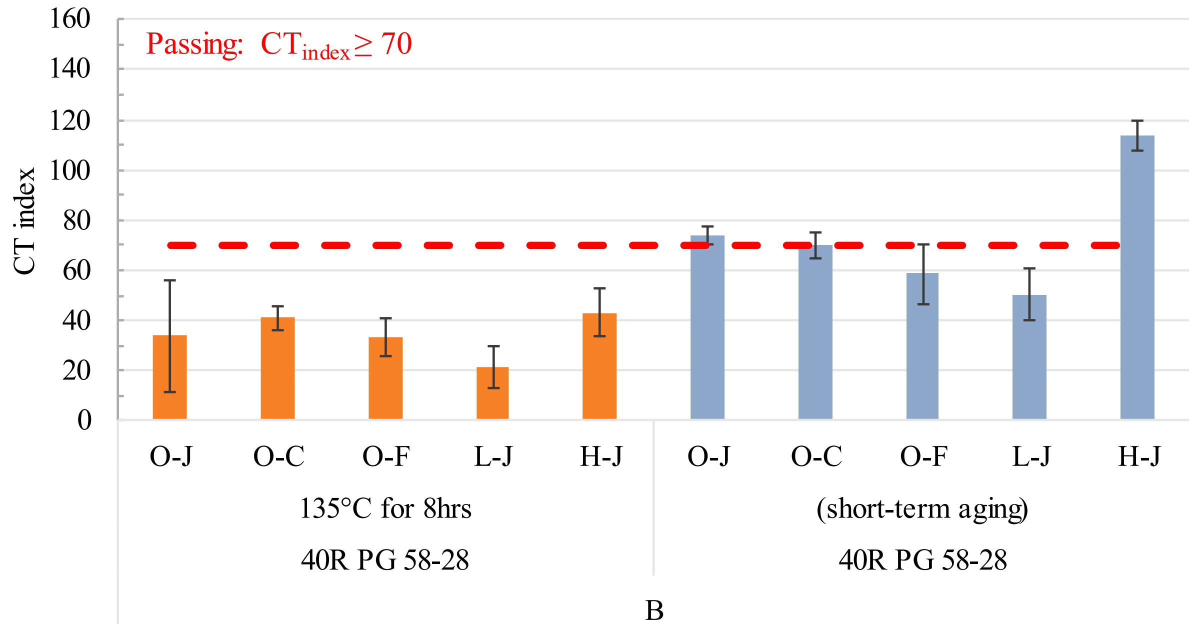

Figure 15. Producer B CT$_{index}$ Results Comparing Loose-mix Oven Aging for 8 Hours at 135°C to 4 Hour Short-term Oven Aging at the Compaction Temperature. Error Bars = ± 1 standard deviation; CT = cracking tolerance; PG = performance grade; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

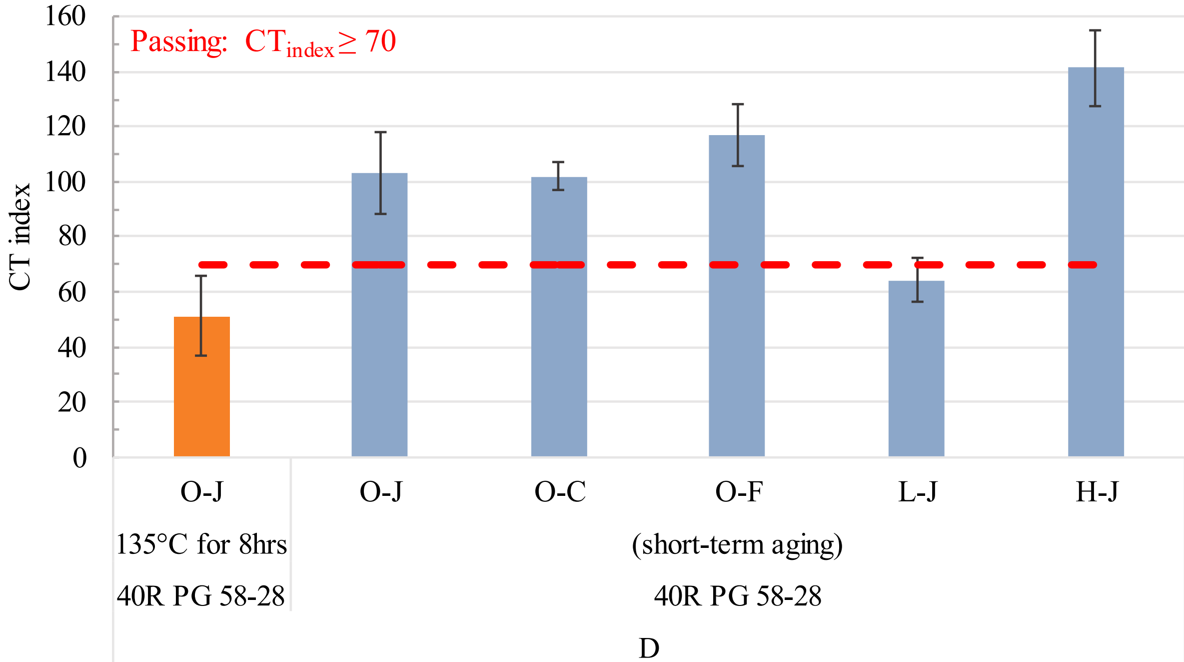

Figure 16. Producer D CT_{index} Results Comparing Loose-mix Oven Aging for 8 Hours at 135°C to 4 Hour Short-term Oven Aging at the Compaction Temperature. Error Bars = ± 1 standard deviation; CT = cracking tolerance; PG = performance grade; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

The relationship between the aged and unaged condition was assessed to evaluate the consistency in aging based on production variability shown in Figure 17. Data for Producer B and Producer D was plotted with the unaged CT_{index} on the y-axis and the aged condition on the x-axis. The data produced a linear trend with an R^2 of 0.69 indicating a good correlation between results. This suggests that aging is not influencing some variations more than others. If this was the case, the correlation would be more skewed resulting in a lower R^2 since the rate of change in CT_{index} would be greater for some variations than others. This is a limited data set with only two mixtures evaluating critical aging for O-J, and one mixture evaluating the influence of production variability on aging. These observations should be confirmed with a broader data set encompassing more mixtures to determine if these observations are mixture specific.

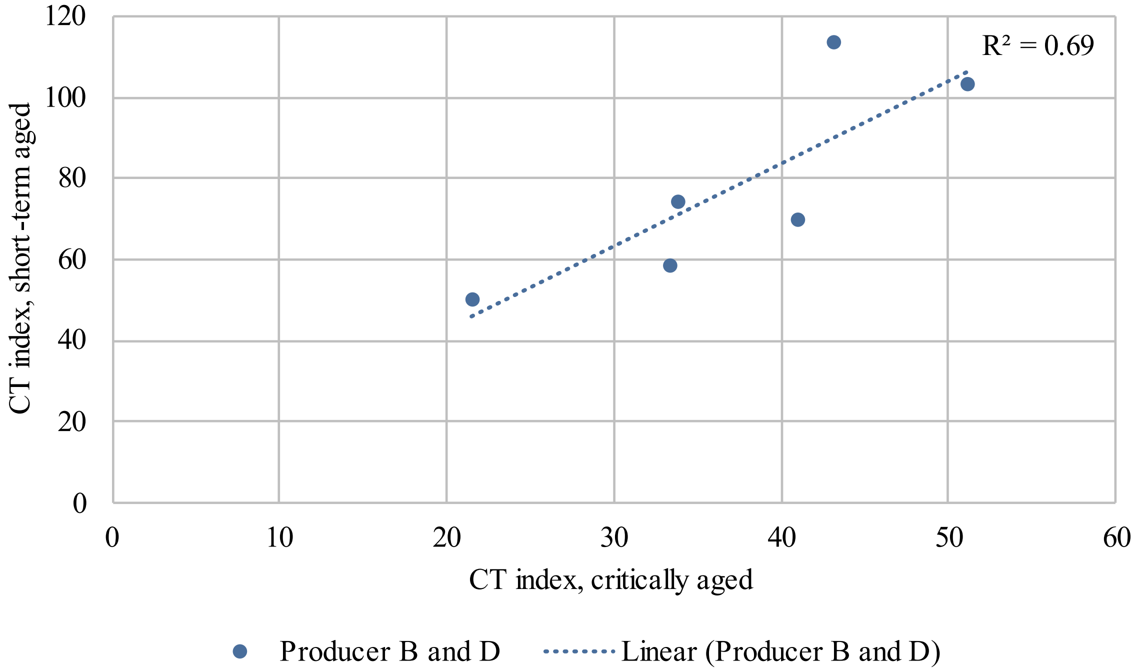

Figure 17. Linear Correlation Between Long-term Aged and Short-term Aged Conditions for CT_{index}. CT = cracking tolerance.

Balancing Durability, Cracking, and Rutting Resistance

Previous design methods have been effective in addressing one key mode of pavement distress, either cracking or rutting, leaving mixtures exposed and susceptible to other distresses. BMD requires that mixtures are resistant to multiple modes of distress, and specifically for Virginia, must meet Cantabro, APA, and IDT-CT test criteria. While most mixtures and variations of those mixtures were resistant to rutting, several mixtures showed a susceptibility to cracking or durability issues as the gradation or binder content varied. Table 15 summarizes whether mixtures met the threshold limit for each variation based on the performance tests. Two mixtures met Cantabro criteria despite variations in gradation and binder content, and three mixtures met IDT-CT requirements. However, only one mixture of the 14 evaluated remained balanced by meeting the threshold requirements for each performance test. This indicates that mixtures can become unbalanced due to production variability and tend to be susceptible to durability issues and cracking.

Table 15. Summary of Mixtures Meeting Performance Threshold Limits for each Variation; Yes = All Variations Met the Threshold Limit, No = Some or None of the Variations Met the Threshold Limit

Producer	Design	Design Method	All Variations Met Threshold Limit? Y/N		
			Cantabro	APA	IDT-CT
A	30R PG 64-22	Volumetric Design	No	Yes	No
	40R PG 58-28	Balanced Mix Design	No	Yes	No
	40R PG 64-22 RA	Balanced Mix Design	No	Yes	No
B	30R PG 64-22	Volumetric Design	No	Yes	No
	40R PG 58-28	Balanced Mix Design	No	Yes	No
	40R PG 64-22 RA	Balanced Mix Design	No	Yes	No
C	35R PG 58-28	Balanced Mix Design	No	Yes	Yes
D	30R PG 64-22	Volumetric Design	No	Yes	No
	40R PG 58-28	Balanced Mix Design	No	Yes	No
E	35R PG 58-28 RA	Balanced Mix Design	No	No	Yes
	35R PG 58-28 Softening Oil + Fibers	Balanced Mix Design	Yes	Yes	Yes
F	30R PG 64-22	Volumetric Design	No	Yes	No
	40R PG 58-28	Balanced Mix Design	Yes	Yes	No
	40R PG 64-22 RA	Balanced Mix Design	No	Yes	No

PG = performance grade; RA = recycling agent; APA = Asphalt Pavement Analyzer Rut Test; IDT-CT = Indirect Tensile Strength test

Impact of Aggregate (Gsb) Variability on Volumetric Parameters

The objective of this study was to evaluate the influence of production variability, mainly the influence of changes in gradation and binder content within acceptable deviations based on production tolerance limits, on performance test results. In establishing the coarse and fine variations and processing the materials for sample preparation, material variability was observed and indicated that the gradation and RAP properties, gradation and RAP binder content, could change etween design and production as detailed in the 'Material ha enges' section. It is reasonable to assume that if gradation presents variability between design and production, other aggregate properties could be changing as well, such as the bulk specific gravity of the aggregate (Gsb). The Gsb value is pivotal for volumetric calculations and influences several parameters such as voids in the mineral aggregate (VMA), percent of absorbed binder (Pba), percent of effective binder (Pbe), and dust-to-binder ratio (F/A) which are detailed in Equations 3-6 (Asphalt Institute, 2014).

The Pba calculation is based on both the effective specific gravity (Gse) and the bulk specific gravity (Gsb). Gse is a function of the maximum theoretical specific gravity (Gmm), percent stone (Ps), percent binder (Pb), and specific gravity of the binder (Gb). For laboratory mixed samples, the amount of aggregate and asphalt is controlled, and the Gmm is tested frequently, which suggests that variability for Gse could be reasonably low. However, Gsb reflects natural changes in aggregate mineralogy, which produces variability and is also not verified as frequently, suggesting that this parameter could have more inherent variability.

$$VMA, \% = 100 - \frac{G_{mb} * P_s}{G_{sb}} \quad \text{[Eq. 3]}$$

$$P_{ba}, \% = 100 * \frac{G_{se} - G_{sb}}{G_{se} * G_{sb}} * G_b \quad \text{[Eq. 4]}$$

$$P_{be}, \% = P_b - \frac{P_{ba}}{100} * P_s \quad \text{[Eq. 5]}$$

$$\frac{F}{A}\ Ratio = \frac{P_{0.075}}{P_{be}} \quad \text{[Eq. 6]}$$

where

VMA = voids in the mineral aggregate,
Gmb = bulk specific gravity of the mix,
Ps = percent of aggregate,
Gsb = bulk specific gravity of the aggregate,
Pba = percent of absorbed binder,
Gse = effective specific gravity,
Gb = specific gravity of the binder,
Pbe = percent of effective binder,
Pb = total binder percentage,
F/A ratio = dust to binder ratio, and
$P_{0.075}$ = percent passing the No. 200 sieve.

To assess the influence of changes in Gsb on volumetric parameters, Gsb data was collected from the production year for each producer and compared to the design data. For Producer D and E, the producer provided gravities from 2021 production year which was compared to the design, and for the remaining producers 2020 production data was compared to the design gravities. To limit the number of variables in this assessment, data corresponding to the O-J variation was used so the differences reflected in volumetric parameters are mainly influenced by changes in the aggregate (Gsb). O-J reflects material variability between design and production, but this variation does not simulate production variability since it was merely the recreation of the design in the laboratory and does not account for changes in gradation or binder content.

Table 16 lists the Gsb, VMA, Pba, Pbe, and F/A values for both design and production for each mixture. The absolute value of Gsb values produced a range of 0.003 to 0.039. Negative values indicate that the design Gsb was lower than production data and is included to describe the direction of change while the magnitude describes the degree of change in Gsb values. Deviations in Gsb values of 0.013 or less, which is also the multilaboratory precision for one standard deviation as per AASHTO T 85, *Standard Method of Test for Specific Gravity and Absorption of Coarse Aggregate*, impacted the volumetric parameters less (e.g., <0.4% for VMA). It is also important to note that the acceptable multilaboratory precision for Gsb for two results is 0.038, which is just under the highest values (0.039) recorded. However, the VMA differences were much higher as well.

Producers A, C, and D resulted in a range less than or equal to 0.013 which resulted in a range of 0.1% to 0.4% for VMA. Whereas, Producers B, E, and F had a Gsb range up to 0.039 and resulted in a difference in VMA between design and production of up to 1.1%. This indicates that material variability between design and production can change the VMA by over a percent. For these calculations, all other variables were constant where calculations were made for each sample with its unique Gmb and the only difference between values in Table 11 are the result of changing the Gsb.

For Pba, the influence of changes in Gsb are considerable. Pba calculations are dependent on the relationship between Gse and Gsb. Gse should have a greater magnitude, compared to Gsb, since the volume is based on the volume of aggregate and water permeable voids that remain after the voids have filled with asphalt. Whereas the Gsb has a larger volume since it accounts for the volume of aggregate and the water permeable voids which results in a lower specific gravity. However, for some producers Pba resulted in a negative value indicating that the Gsb was greater than Gse based on Equations 3-6. Since Gse is based on parameters that are verified frequently, this deviation is most likely a result of differences in Gsb. Figure 18 illustrates changes in Pba between design and construction depicting considerable deviations between the two and the tendency for some values to result in negative values.

Table 16. Sensitivity of Volumetric Calculations to Changes in Bulk Specific gravity (Gsb); All Calculations Based on O-J Variation

Producer	Design	Design Gsb	Prod. Gsb	Range	Design Avg. of VMA, %	Prod. Avg. of VMA, %	Range	Design Avg. of Pba, %	Prod. Avg. of Pba, %	Range	Design Avg. of Pbe, %	Prod. Avg. of Pbe, %	Range	Design Avg. F/A	Prod. Avg. F/A	Range
A	30R PG 64-22	2.908	2.915	-0.007	18.2	18.4	-0.2	0.41	0.32	0.09	5.21	5.30	-0.08	1.05	1.03	0.02
	40R PG 58-28	2.910	2.913	-0.003	17.5	17.7	-0.2	0.31	0.25	0.06	5.21	5.27	-0.06	1.17	1.15	0.01
	40R PG 64-22 RA	2.910	2.913	-0.003	18.2	18.4	-0.1	0.29	0.23	0.06	5.22	5.28	-0.06	1.16	1.15	0.01
B	30R PG 64-22	2.902	2.872	0.030	14.5	13.6	0.9	0.65	1.02	-0.37	4.68	4.33	0.35	1.19	1.29	-0.10
	40R PG 58-28	2.931	2.902	0.029	16.1	15.3	0.8	0.28	0.63	-0.35	5.34	5.00	0.33	1.08	1.15	-0.07
	40R PG 64-22 RA	2.931	2.902	0.029	15.9	15.0	0.8	0.21	0.57	-0.35	5.40	5.06	0.33	1.07	1.14	-0.07
C	35R PG 58-28	2.689	2.685	0.004	19.1	18.9	0.1	-0.23	-0.18	-0.06	6.42	6.37	0.05	0.83	0.84	-0.01
D	30R PG 64-22	2.705	2.718	-0.013	17.7	18.1	-0.4	0.69	0.50	0.19	5.25	5.43	-0.18	1.15	1.11	0.04
	40R PG 58-28	2.710	2.721	-0.011	16.8	17.0	-0.3	0.59	0.47	0.12	5.24	5.36	-0.11	0.99	0.97	0.02
E	35R PG 58-28 RA	2.701	2.677	0.024	16.2	15.5	0.7	0.10	0.43	-0.34	6.11	5.80	0.32	0.93	0.98	-0.05
	35R PG 58-28 Softening Oil + Fibers	2.711	2.687	0.025	16.5	15.7	0.8	0.06	0.41	-0.35	6.14	5.82	0.33	0.92	0.97	-0.05
F	30R PG 64-22	2.908	2.938	-0.029	16.2	17.1	-0.8	0.38	0.02	0.36	5.14	5.48	-0.34	1.12	1.05	0.07
	40R PG 58-28	2.907	2.946	-0.039	16.5	17.7	-1.1	0.24	-0.23	0.47	5.37	5.82	-0.45	1.09	1.01	0.08
	40R PG 64-22 RA	2.907	2.946	-0.039	16.3	17.4	-1.1	0.23	-0.25	0.47	5.39	5.83	-0.45	1.09	1.01	0.08

Design values = values from the design submittal document to produce the Job Mix Formula (JMF); Production values (Prod.) = values from the contractor from around the time of construction; Range = the different between design values and production values; Gsb = bulk specific gravity; VMA = voids in the mineral aggregate; Pba = percent of absorbed binder; Pbe = percent of effective binder; F/A = dust-to-binder ratio.

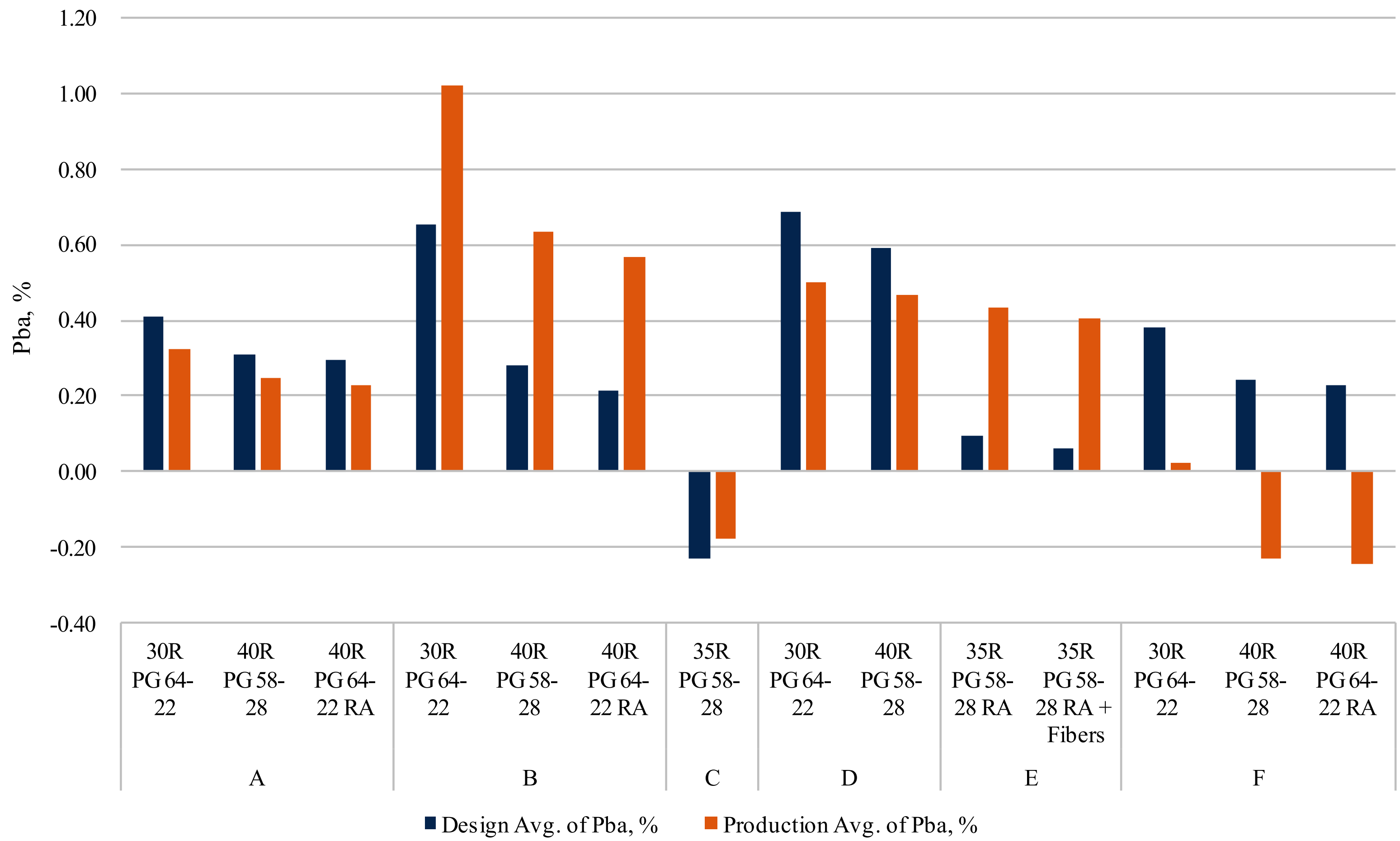

Figure 18: Changes in Pba Between Design and Production Due to Differences in Gsb, Some of Which Resulted in Negative Values Which Theoretically are Not Possible and are Most Likely Related to Changes in Gsb. Pba = percent of absorbed binder; PG = performance grade; RA = recycling agent.

Pbe and F/A are a function of Pba, and the influence of Gsb is translated into the parameters through the impact of Pba. F/A did not result in considerable deviations like Pbe with the greatest range being 0.1. For Pbe, the greatest range was 0.45%, which is considerable since the total Pbe is typically between 5-6% for this data.

This analysis highlights the potential for aggregate variability within a fairly short time and how that variability influences other volumetric parameters. Caution should be taken when designing and accepting asphalt mixtures with respect to volumetric parameters alone since they can be influenced considerably by changes in Gsb. For volumetric design and acceptance, Gsb should be evaluated frequently, and designs updated regularly. One of the key reasons for transitioning to BMD was the concern of aggregate variability and its influence on volumetric parameters used for material acceptance, as summarized in the introduction. Results of this analysis confirm these concerns and further support the need for performance-based testing not only in design but also in quality control for product acceptance.

SUMMARY OF FINDINGS

Material Variability

- During material processing, it was determined that materials had inherent variability between design and sampling, causing aggregate or RAP gradations to change or RAP asphalt content to change. Adjustments were made to ensure that designs were similar to the JMF gradation and total binder content.
- Updated aggregate gravities were supplied by the producers to assess the influence of Gsb values on volumetric parameters. When other material inputs were held constant and Gsb was varied from design to production values, VMA increased as much as 1.1%. Pba could produce unreasonable (negative) values, which is a concern since it is required to calculate other volumetric parameters, such as Pbe and F/A ratio.

Influence of Production Variability on Performance Test Results

- Only one mixture stayed balanced and met the threshold limits for each respective performance test. Ten mixtures were originally designed to meet the BMD specifications indicating that mixtures can become unbalanced because of production variability.

Cantabro Durability Results

- When mixtures were varied in the laboratory according to the tolerance limits, three mixtures did not meet the threshold limit (ML≤7.5%) for any of the variations. The majorit of mixtures resulted in some passing and some failing variations indicating that mixtures designed to meet the threshold limit could fail the criteria due to production variability. Only two mixtures met the requirements for each of the variations assessed.

- Binder content influenced ML results where high binder content improved durability and low binder content hindered durability, which was also statistically significant for six designs.

- Coarse gradation reduced durability and was significant for five designs, while the fine gradation did not significantly influence results.

- Based on the interaction variations, some mixtures may be sensitive to changes in gradation and binder content more than other mixtures, which may be influenced by the interaction between volumetric parameters such as VMA, VTM, or Vbe.

APA Rutting Results

- APA results showed excellent rutting resistance where 13 of the 14 mixtures met the rutting criteria for all the variations evaluated. Only one mixture variation, H-F, failed to meet the specification requirements.

- Low binder content positively influenced rut depths improving rutting resistance, and high binder content increased rut depths. The fine gradation also resulted in increasing rut depth, while the coarse gradation only slightly improved results.

- When evaluating the interaction mixtures, the H-F mixture produced the greatest rut depths. High binder and a fine gradation negatively influence results, and the combination of the two produced an even greater impact. For L-F, differing results were observed between producers where L-F hindered rutting resistance for Producer B but improved rutting resistance for Producer E. Alternate trends based on producer may indicate a sensitivity to volumetric parameters.

IDT-CT Results

- Three mixtures met the IDT-CT criteria despite variations in gradation and binder content. The only mixture not to meet the requirements for all the variations was a Superpave design and not originally designed to meet BMD criteria. The remaining 10 mixtures produced variations that passed while others failed the threshold limit.

- Binder content was significant for CTindex results improving results for the high binder content and hindering results for the low binder content.

- Generally, gradation did not significantly influence results since only 1 or 2 mixtures produced statistically significant results. The fine gradation negatively influenced results and coarse gradation only slightly improved CTindex, although both observations were not considered to be statistically significant.

- For the interaction mixtures, some variations show conflicting results between producers. H-F for Producer E increased CTindex but showed a decrease for Producer B. However, L-F had a significant, negative effect on cracking resistance for both producers. Although gradation at the optimum binder content was not statistically significant, when combined with the low binder content, results were confounded, reducing CTindex more than when only one variable changes. This indicates that there may be an influence of volumetric parameters.

- Critical aging mixtures significantly reduced CTindex. Production variability does not appear to influence aging such that the rate of reduction for CTindex between the unaged and aged samples were similar for each variation.

CONCLUSIONS

- *Inherent material variability and variability related to control of RAP stockpiles can considerably change volumetric properties.* This has been a main concern for the Superpave design method, and one of the reasons to transition to BMD. Analysis from this study confirmed these concerns despite the controlled experimental nature, and further supports the need for performance-based testing during design and production to ensure that quality asphalt mixtures are being produced even as volumetric properties shift and change during production.

- *Interactions between changing gradation and binder content appear to have a more complex influence on performance test results and may indicate an influence of volumetric parameters.* As gradation and binder content changes, key volumetric properties such as VMA, Vbe, VTM, and F/A ratio change which may cause a different response for performance tests.

- *Mixtures balanced during design can become unbalanced during production.* When gradation and binder content were allowed to vary within the acceptable tolerance limits, some variations for some mixtures failed to meet the performance threshold limits, typically for cracking and durability requirements.

RECOMMENDATIONS

1. *VDOT's Materials Division should continue with the implementation of BMD.* This is especially pertinent considering the inherent variability of materials and the effect on volumetric properties traditionally used to measure mixture quality and acceptance.

2. *VDOT's Materials Division and VTRC should evaluate ways to encourage mixtures to stay balanced through design and production.* Two options that could be pursued include, but are not limited to:

   — Adjust threshold limits for design so that the potentially worst performing variations in production should meet current BMD specification threshold limits. This option is less conservative and may cause balanced mixtures, designed with satisfactory performance expectations, to be accepted in a state that no longer provides the expected field performance.

— Inform contractors of the potential impacts of acceptable production variations so that they can adjust mix designs appropriately to insure passing results despite plant and material variability. This will help ensure mixture performance as understood by the current state of BMD and the established thresholds, while potentially leading to mixes that are somewhat overdesigned to ensure they stay balanced during production.

3. *VTRC should continue to assess how volumetric parameters influence performance test results.* The interaction mixtures may have a greater influence on results compared to variations that only vary one parameter (either gradation or binder content).

4. *VTRC should continue to assess the effect of laboratory aging on performance to simulate how mixtures will perform in-service.*

IMPLEMENTATION AND BENEFITS

Researchers and the technical review panel (listed in the Acknowledgments) for the project collaborate to craft a plan to implement the study recommendations and to determine the benefits of doing so. This is to ensure that the implementation plan is developed and approved with the participation and support of those involved with VDOT operations. The implementation plan and the accompanying benefits are provided here.

Implementation

Regarding Recommendation 1, VDOTs Materials Division and VTRC should work together to continue exploring BMD and the impacts on production variability. This will further inform specification development, and support full implementation of the specification, with the goal of enhanced mixture quality and performance.

Regarding Recommendation 2, the VDOT Materials Division in conjunction with VTRC (and/or Auburn University and NCAT) should develop technical briefs, presentations, and/or webinars describing the findings of this report to inform mix designers of considerations during the design of BMD mixtures to ensure that they stay balanced during production.

Regarding Recommendation 3, VTRC should consider volumetric parameters and their influence on performance test results in ongoing and future BMD research projects. VTRC can implement this in the next phase of the ongoing BMD pilot project research.

Regarding Recommendation 4, VTRC Project No. 122013, Developing Long-term Aging Protocols for Cracking Performance Evaluation of Asphalt Mixtures in Virginia, is ongoing. The objective of this research is twofold: (1) to develop practical long-term aging protocols for asphalt SMs with A and D designations that can be used in mix design verification and production of asphalt mixtures for quality control and acceptance purposes; and (2) to develop preliminary performance criteria for CT index for the to-be-developed long-term aging protocols. Both objectives will provide necessary refinement and improvements to performance testing used in Virginia's BMD initiative. The outcomes of this effort are expected to be available in November 2024.

Benefits

Regarding Recommendation 1, VDOT will move toward a performance-based asphalt mixture design and acceptance approach, reducing the potential impact of material and plant variation impacts on volumetrics and improving the long-term performance of the asphalt mixtures.

*Regarding Recommendation 2***,** VDOT will be adequately positioned to manage potential challenges with regard to plant variation and its impact on results by educating contractors to ensure that produced mixtures pass current and future BMD requirements.

Regarding Recommendation 3, a fundamental understanding of how volumetrics change with plant and material variation during production and impact on BMD performance tests will further support the transition to BMD.

Regarding Recommendation 4, connecting laboratory aging to long term performance will further strengthen the confidence VDOT has in BMD tests and will allow for better planning and maintenance scheduling due to reduced risk.

ACKNOWLEDGMENTS

The authors thank the Virginia Transportation Research Council (VTRC) and the Federal Highway Administration (FHWA) for their support of this research. The authors are grateful to the following individuals who served on the technical review panel for this study: Robert Crandol, P.E. (Project Champion, Assistant State Maintenance Engineer, Maintenance Division, VDOT), Angela Beyke, P.E. (Assistant State Materials Engineer, Materials Division, VDOT), Kwame Adu-Gyamfi, Ph.D., P.E. (District Materials Engineer, Fredericksburg District, VDOT), Tommy Schinkel, P.E. (District Materials Engineer, Richmond District, VDOT), and David Shiells, P.E. (District Materials Engineer, Northern Virginia District, VDOT), as well as former member Sungho Kim, Ph.D., P.E. (formerly of VDOT). The authors thank Troy Deeds, Donnie Dodds, and Jennifer Samuels of VTRC, John Burch, Ben Earl, and Zack Robb, formerly of VTRC, and Trace Fontana, Elizabeth Turochy, Elizabeth Lowman, Benjamin Prowell, Kevin Ambrose, Tina Taylor, and Vicki Adams of Auburn University and the National Center for Asphalt Technology, for their efforts.

REFERENCES

Asphalt Institute. *MS-2: Asphalt Mix Design Methods*. Asphalt Institute, Lexington, KY, 2014.

Austerman, A.J., Mogawer, W.S., Stuart, K.D. Influence of Production Considerations on Balanced Mix Designs. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2672, Issue 28, 2018, pp. 426-437.

Bowers, B.F., Diefenderfer, S.D., Moore, N. and Lynn, T. Balanced Mix Design and Benchmarking: A Case Study in Establishing Performance Test Thresholds. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, Issue 12, 2022

Boz, I., Habbouche, J., Diefenderfer, S. and Bilgic, Y., 2022. Precision Estimates and Statements for Performance Indices from the Indirect Tensile Cracking Test at Intermediate Temperature. *Transportation Research Record*, Vol. 2676(5), pp.225-241.

Chen, C., Yin, F., Andriescu, A., Moraes, R., Mensching, D., Tran, N., Taylor, A., West, R. Preliminary Validation of the Critical Aging Protocol for NCAT Top-down Cracking Experiment. *Journal of the Association of Asphalt Paving Technologists.* Vol. 89, 2020.

Cox, B.C., Smith, B.T., Howard, I.L., James, R.S. State of Knowledge for Cantabro Testing of Dense Graded Asphalt. *Journal of Materials in Civil Engineering*, Vol. 29, Issue 10, 2017.

Devore, J.L. and Farnum, N.R. *Applied Statistics for Engineers and Scientists*. Belmont, CA, Brooks, Cole, 2005.

Diefenderfer, S.D. and Bowers, B.F. Initial Approach to Performance (Balanced) Mix Design: The Virginia Experience. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2673, Issue 2, 2019, pp. 335-345.

Diefenderfer, S.D., Boz, I., and Habbouche, J. *Balanced Mix Design for Surface Asphalt Mixtures: Phase I: Initial Roadmap Development and Specification Verification*. VTRC 21R-15. Virginia Transportation Research Council, Charlottesville, 2021a.

Diefenderfer, S. D., Boz, I., and Habbouche, J. *Balanced Mix Design for Surface Asphalt Mixtures: 2019 Field Trials*. VTRC 21R-21.Virginia Transportation Research Council, Charlottesville, 2021b.

Habbouche, J., Boz, I., Diefenderfer, S.D. and Bilgic, Y.K., 2021. *Round Robin Testing Program for the indirect tensile cracking test at intermediate temperature: Phase I* (No. FHWA/VTRC 22-R3). Virginia Transportation Research Council (VTRC).

Habbouche, J., Boz, I. and Diefenderfer, S.D., 2022. *Interlaboratory Study for the Indirect Tensile Cracking Test at Intermediate Temperature: Phase II* (No. FHWA/VTRC 23-R3). Virginia Transportation Research Council (VTRC).

Ling, C. and Buchanan, S. Implementation Considerations of Balanced Mix Design in Practice: Recent Experience in Vermont. *Road Materials and Pavement Design*, Vol. 24, Issue 3, 2022, pp. 836-849.

Mogawer, W. S., Stuart, K.D., Austerman, A.J., Zhou, F., Romero, P. Balanced Mix Design Sensitivity to Production Tolerance Limits and Binder Source. *Journal of the Association of Asphalt Paving Technologists*, Vol. 88, 2019.

NAPA. Balanced Mix Design Resource Guide: Implementation Efforts. National Asphalt Pavement Association, 2023 https://www.asphaltpavement.org/expertise/engineering/resources/bmd-resource-guide/implementation-efforts. Accessed January 16, 2023.

VDOT. *Road and Bridge Specification*. Virginia Department of Transportation, 2016.

Wang, Y. D., Liu, J., Liu, J. Integrating Quality Assurance in Balanced Mix Designs for Durable Asphalt Mixtures: State-of-the-Art Literature Review. *Journal of Transportation Engineering, Part B: Pavements*, Vol 149, Issue 1, 2023.

West, R., Rodezno, C., Leiva, F., Yin, F. *Development of a Framework for Balanced Mix Design*. Project NCHRP 20-07/Task 406, Transportation Research Board, Washington, DC, 2018.

APPENDIX A

GRADATION AND VOLUMETRIC PROPERTIES FOR EACH DESIGN BY GRADATION TYPE (JMF, COARSE, FINE)

Table A1. Producer A Gradation and Volumetric Properties for Each Design and Each Gradation Type (JMF, Coarse, Fine)

Producer A

	Mix ID		30R PG 64-22	**Mix ID**		40R PG 58-28	**Mix ID**		40R PG 64-22 RA
	Mixture Type		SM-9.5A	**Mixture Type**		SM-9.5A	**Mixture Type**		SM-9.5A
	NMAS, mm		9.5	**NMAS, mm**		9.5	**NMAS, mm**		9.5
	Virgin Binder Grade		PG 64-22	**Virgin Binder Grade**		PG 58-28	**Virgin Binder Grade**		PG 64-22
	RAP Content, %		30%	**RAP Content, %**		40%	**RAP Content, %**		40%
	Recycling Agent		No	**Recycling Agent**		No	**Recycling Agent**		Yes
	Design Asphalt Content, %		5.60	**Design Asphalt Content, %**		5.50	**Design Asphalt Content, %**		5.50
	JMF	**Coarse**	**Fine**	**JMF**	**Coarse**	**Fine**	**JMF**	**Coarse**	**Fine**
VTM, %	5.4	7.1	3.2	4.7	6.3	4.3	5.4	5.9	3.2
Gsb	2.915	2.915	2.916	2.913	2.914	2.913	2.913	2.914	2.913
Gse	2.942	2.942	2.934	2.934	2.939	2.936	2.932	2.940	2.935
Gmm	2.665	2.665	2.659	2.663	2.667	2.665	2.662	2.668	2.664
Gmb	2.521	2.476	2.574	2.538	2.499	2.551	2.517	2.511	2.578
VMA, %	18.4	19.8	16.7	17.7	19.0	17.2	18.4	18.6	16.4
Vbe	13.0	12.7	13.5	13.0	12.7	13.0	12.9	12.7	13.1
VFA, %	70.6	64.2	80.8	73.4	66.8	75.3	70.4	68.3	80.2
Pbe, %	5.3	5.3	5.4	5.3	5.2	5.2	5.3	5.2	5.2
FA Ratio	1.0	0.8	1.2	1.2	1.0	1.3	1.1	1.0	1.3
Gradation (Percent Passing)									
1/2 in (12.5 mm)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
3/8 in (9.5 mm)	96.5	96.2	96.9	96.5	96.3	96.8	96.5	96.3	96.8
No. 4 (4.75 mm)	59.6	55.6	63.6	59.0	56.9	62.7	59.0	56.9	62.7
No. 8 (2.36 mm)	37.3	33.7	41.3	36.8	34.8	40.1	36.8	34.8	40.1
No. 16 (1.18 mm)	25.8	23.0	28.9	26.0	24.3	28.4	26.0	24.3	28.4
No. 30 (0.6 mm)	18.4	16.3	20.5	18.8	17.4	20.5	18.8	17.4	20.5
No. 50 (0.3 mm)	12.5	10.9	14.0	13.1	11.9	14.3	13.1	11.9	14.3
No. 100 (0.15 mm)	8.1	6.8	9.4	8.8	7.7	9.8	8.8	7.7	9.8
No. 200 (0.075 mm)	5.47	4.46	6.50	6.07	5.20	6.79	6.07	5.20	6.79

PG = performance grade; NMAS = nominal maximum aggregate size; RAP = reclaimed asphalt pavement; JMF = job mix formula; VTM = voids total mixture; Gsb = bulk (dry) specific gravity; Gse = effective specific gravity; Gmm = maximum theoretical specific gravity; Gmb = bulk specific gravity; VMA = voids in the mineral aggregate; Vbe = volume of effective binder; VFA = voids filled with asphalt; Pbe = percent of effective binder; FA ratio = dust-to-binder ratio.

Table A2. Producer B Gradation and Volumetric Properties for Each Design and Each Gradation Type (JMF, Coarse, Fine)

Producer B

	Mix ID		30R PG 64-22			40R PG 58-28	**Mix ID**		40R PG 64-22 RA
	Mixture Type		SM-12.5A	**Mixture Type**		SM-12.5A	**Mixture Type**		SM-12.5A
	NMAS, mm		12.5	**NMAS, mm**		12.5	**NMAS, mm**		12.5
	Virgin Binder Grade		PG 64-22	**Virgin Binder Grade**		PG 58-28	**Virgin Binder Grade**		PG 64-22
	RAP Content, %		30%	**RAP Content, %**		40%	**RAP Content, %**		40%
	Recycling Agent		No	**Recycling Agent**		No	**Recycling Agent**		Yes
	Design Asphalt Content, %		5.3	**Design Asphalt Content, %**		5.6	**Design Asphalt Content, %**		5.6
	JMF	**Coarse**	**Fine**	**JMF**	**Coarse**	**Fine**	**JMF**	**Coarse**	**Fine**
VTM, %	2.6	2.6	2.4	2.8	2.9	1.6	2.2	1.6	1.3
Gsb	2.872	2.881	2.868	2.902	2.905	2.855	2.902	2.905	2.855
Gse	2.957	2.966	2.950	2.952	2.951	2.938	2.949	2.941	2.927
Gmm	2.690	2.697	2.685	2.673	2.672	2.662	2.670	2.665	2.653
Gmb	2.620	2.627	2.620	2.597	2.595	2.620	2.612	2.623	2.618
VMA, %	13.6	13.6	13.5	15.5	15.7	13.4	15.0	14.8	13.4
Vbe	11.0	11.1	11.1	12.7	12.8	11.8	12.8	13.2	12.1
VFA, %	80.9	81.1	82.0	81.8	81.7	88.1	85.5	89.4	90.3
Pbe, %	4.3	4.3	4.3	5.0	5.1	4.6	5.1	5.2	4.8
FA Ratio	1.3	1.1	1.4	1.1	0.9	1.4	1.1	0.9	1.4
Gradation Percent Passing									
1/2 in (12.5 mm)	97.1	95.9	98.2	97.9	96.9	99.0	97.9	96.9	99.0
3/8 in (9.5 mm)	88.3	85.2	91.5	89.9	86.9	92.9	89.9	86.9	92.9
No. 4 (4.75 mm)	59.9	56.6	63.7	58.1	54.0	61.4	58.1	54.0	61.4
No. 8 (2.36 mm)	40.9	37.7	42.9	38.0	35.2	41.0	38.0	35.2	41.0
No. 16 (1.18 mm)	29.9	26.4	31.1	26.4	24.2	29.7	26.4	24.2	29.7
No. 30 (0.6 mm)	21.7	18.7	22.3	18.9	17.2	21.6	18.9	17.2	21.6
No. 50 (0.3 mm)	15.4	13.1	16.0	13.4	11.9	15.7	13.4	11.9	15.7
No. 100 (0.15 mm)	9.1	8.0	9.4	8.9	7.8	10.0	8.9	7.8	10.0
No. 200 (0.075 mm)	5.51	4.97	5.92	5.76	4.77	6.63	5.76	4.77	6.63

PG = performance grade; NMAS = nominal maximum aggregate size; RAP = reclaimed asphalt pavement; JMF = job mix formula; VTM = voids total mixture; Gsb = bulk (dry) specific gravity; Gse = effective specific gravity; Gmm = maximum theoretical specific gravity; Gmb = bulk specific gravity; VMA = voids in the mineral aggregate; Vbe = volume of effective binder; VFA = voids filled with asphalt; Pbe = percent of effective binder; FA ratio = dust-to-binder ratio.

Table A3. Producer C Gradation and Volumetric Properties for Each Design and Each Gradation Type (JMF, Coarse, Fine)

Producer C

	Mix ID		35R PG 58-28
	Mixture Type		SM-12.5A
	NMAS, mm		12.5
	Virgin Binder Grade		PG 58-28
	RAP Content, %		35%
	Recycling Agent		No
	Design Asphalt Content, %		6.20
	JMF	**Coarse**	**Fine**
VTM, %	4.6	4.7	3.3
Gsb	2.685	2.687	2.681
Gse	2.672	2.675	2.676
Gmm	2.432	2.434	2.435
Gmb	2.320	2.318	2.354
VMA, %	18.9	19.1	17.6
Vbe	14.3	14.3	14.3
VFA, %	75.8	75.2	81.5
Pbe, %	6.4	6.4	6.3
FA Ratio	0.8	0.7	1.0
Gradation Percent Passing			
1/2 in (12.5 mm)	95.7	94.8	97.4
3/8 in (9.5 mm)	84.8	82.5	88.8
No. 4 (4.75 mm)	49.5	45.5	53.0
No. 8 (2.36 mm)	33.7	30.4	36.1
No. 16 (1.18 mm)	24.3	22.0	25.9
No. 30 (0.6 mm)	16.8	14.8	18.4
No. 50 (0.3 mm)	10.1	8.4	11.8
No. 100 (0.15 mm)	6.7	5.3	8.1
No. 200 (0.075 mm)	4.63	3.63	5.65

PG = performance grade; NMAS = nominal maximum aggregate size; RAP = reclaimed asphalt pavement; JMF = job mix formula; VTM = voids total mixture; Gsb = bulk (dry) specific gravity; Gse = effective specific gravity; Gmm = maximum theoretical specific gravity; Gmb = bulk specific gravity; VMA = voids in the mineral aggregate; Vbe = volume of effective binder; VFA = voids filled with asphalt; Pbe = percent of effective binder; FA ratio = dust-to-binder ratio.

Table A4. Producer D Gradation and Volumetric Properties for Each Design and Each Gradation Type (JMF, Coarse, Fine)

Producer D						
	Mix ID		30R PG 64-22	**Mix ID**		40R PG 58-28
	Mixture Type		SM-12.5A	**Mixture Type**		SM-12.5A
	NMAS, mm		12.5	**NMAS, mm**		12.5
	Virgin Binder Grade		PG 64-22	**Virgin Binder Grade**		PG 58-28
	RAP Content, %		30%	**RAP Content, %**		40%
	Recycling Agent		No	**Recycling Agent**		No
	Design Asphalt Content, %		5.90	**Design Asphalt Content, %**		5.80
	JMF	**Coarse**	**Fine**	**JMF**	**Coarse**	**Fine**
VTM, %	5.6	5.6	5.1	4.5	6.6	5.1
Gsb	2.705	2.706	2.706	2.710	2.708	2.711
Gse	2.755	2.757	2.754	2.755	2.754	2.757
Gmm	2.507	2.509	2.507	2.511	2.510	2.513
Gmb	2.367	2.367	2.378	2.397	2.344	2.386
VMA, %	17.7	17.7	17.3	16.7	18.5	17.1
Vbe	12.1	12.0	12.2	12.1	11.9	12.0
VFA, %	68.3	68.1	70.5	72.8	64.2	70.4
Pbe, %	5.3	5.2	5.3	5.2	5.2	5.2
FA Ratio	1.1	1.0	1.3	1.0	0.8	1.2
Gradation Percent Passing						
1/2 in (12.5 mm)	94.4	93.7	95.1	96.0	94.8	96.7
3/8 in (9.5 mm)	85.6	83.9	87.3	87.6	85.0	89.1
No. 4 (4.75 mm)	57.8	54.0	62.0	54.5	50.1	56.9
No. 8 (2.36 mm)	36.6	33.7	39.8	34.6	31.1	36.8
No. 16 (1.18 mm)	24.4	22.1	26.6	23.3	20.6	25.1
No. 30 (0.6 mm)	17.1	15.3	18.8	16.6	14.3	18.1
No. 50 (0.3 mm)	12.0	10.6	13.4	11.7	9.8	13.1
No. 100 (0.15 mm)	8.4	7.2	9.6	7.8	6.4	8.8
No. 200 (0.075 mm)	6.03	5.00	6.97	5.20	4.20	5.98

PG = performance grade; NMAS = nominal maximum aggregate size; RAP = reclaimed asphalt pavement; JMF = job mix formula; VTM = voids total mixture; Gsb = bulk (dry) specific gravity; Gse = effective specific gravity; Gmm = maximum theoretical specific gravity; Gmb = bulk specific gravity; VMA = voids in the mineral aggregate; Vbe = volume of effective binder; VFA = voids filled with asphalt; Pbe = percent of effective binder; FA ratio = dust-to-binder ratio.

Table A5. Producer E Gradation and Volumetric Properties for Each Design and Each Gradation Type (JMF, Coarse, Fine)

Producer E

	Mix ID		35R PG 58-28 RA	**Mix ID**		35R PG 58-28 RA + Fibers
	Mixture Type		SM-12.5A	**Mixture Type**		SM-12.5A
	NMAS, mm		12.5	**NMAS, mm**		12.5
	Virgin Binder Grade		PG 58-28	**Virgin Binder Grade**		PG 58-28
	RAP Content, %		35%	**RAP Content, %**		35%
	Recycling Agent		Yes	**Recycling Agent**		Yes
	Design Asphalt Content, %		6.20	**Design Asphalt Content, %**		6.20
	JMF	**Coarse**	**Fine**	**JMF**	**Coarse**	**Fine**
VTM, %	1.9	2.7	1.3	2.1	2.1	2.2
Gsb	2.701	2.698	2.704	2.711	2.699	2.714
Gse	2.708	2.694	2.706	2.716	2.712	2.726
Gmm	2.459	2.449	2.458	2.465	2.463	2.474
Gmb	2.412	2.382	2.426	2.413	2.411	2.419
VMA, %	16.2	17.2	15.8	16.5	16.2	16.4
Vbe	14.3	14.5	14.5	14.4	14.1	14.2
VFA, %	88.1	84.1	91.9	87.2	87.0	86.5
Pbe, %	6.1	6.3	6.2	6.1	6.0	6.0
FA Ratio	0.9	0.8	1.1	0.9	0.8	1.1
Gradation Percent Passing						
1/2 in (12.5 mm)	98.0	97.6	98.3	98.2	97.7	98.3
3/8 in (9.5 mm)	90.2	88.1	91.4	90.6	88.3	91.4
No. 4 (4.75 mm)	62.0	59.3	65.0	62.6	59.2	66.4
No. 8 (2.36 mm)	42.5	40.3	45.0	43.2	40.3	46.3
No. 16 (1.18 mm)	31.3	29.4	33.4	31.0	28.5	33.4
No. 30 (0.6 mm)	23.7	22.0	25.5	23.4	21.1	25.4
No. 50 (0.3 mm)	16.6	15.1	18.2	17.1	15.0	18.7
No. 100 (0.15 mm)	9.8	8.7	11.0	10.5	8.9	11.7
No. 200 (0.075 mm)	5.68	4.71	6.68	5.67	4.69	6.52

PG = performance grade; NMAS = nominal maximum aggregate size; RAP = reclaimed asphalt pavement; JMF = job mix formula; VTM = voids total mixture; Gsb = bulk (dry) specific gravity; Gse = effective specific gravity; Gmm = maximum theoretical specific gravity; Gmb = bulk specific gravity; VMA = voids in the mineral aggregate; Vbe = volume of effective binder; VFA = voids filled with asphalt; Pbe = percent of effective binder; FA ratio = dust-to-binder ratio.

Table A6. Producer F Gradation and Volumetric Properties for Each Design and Each Gradation Type (JMF, Coarse, Fine)

Producer F

	Mix ID		30R PG 64-22	**Mix ID**		40R PG 58-28	**Mix ID**		40R PG 64-22 RA
	Mixture Type		SM-9.5A	**Mixture Type**		SM-9.5A	**Mixture Type**		SM-9.5A
	NMAS, mm		9.5	**NMAS, mm**		9.5	**NMAS, mm**		9.5
	Virgin Binder Grade		PG 64-22	**Virgin Binder Grade**		PG 58-28	**Virgin Binder Grade**		PG 64-22
	RAP Content, %		30%	**RAP Content, %**		40%	**RAP Content, %**		40%
	Recycling Agent		No	**Recycling Agent**		No	**Recycling Agent**		Yes
	Design Asphalt Content, %		5.50	**Design Asphalt Content, %**		5.60	**Design Asphalt Content, %**		5.60
	JMF	**Coarse**	**Fine**	**JMF**	**Coarse**	**Fine**	**JMF**	**Coarse**	**Fine**
VTM, %	3.4	4.1	2.3	3.2	4.6	3.3	2.9	3.2	3.3
Gsb	2.938	2.938	2.938	2.946	2.945	2.947	2.946	2.945	2.947
Gse	2.940	2.943	2.942	2.927	2.932	2.931	2.926	2.933	2.930
Gmm	2.668	2.671	2.671	2.656	2.660	2.659	2.653	2.659	2.657
Gmb	2.579	2.561	2.610	2.570	2.538	2.571	2.577	2.574	2.570
VMA, %	17.1	17.6	16.0	17.7	18.7	17.7	17.4	17.5	17.7
Vbe	13.7	13.5	13.8	14.4	14.1	14.3	14.6	14.3	14.4
VFA, %	80.3	76.7	85.9	81.8	75.5	81.2	83.5	81.8	81.7
Pbe, %	5.5	5.4	5.4	5.8	5.7	5.8	5.8	5.7	5.8
FA Ratio	1.1	0.9	1.2	1.0	0.9	1.2	1.0	0.9	1.2
Gradation Percent Passing									
1/2 in (12.5 mm)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
3/8 in (9.5 mm)	92.1	91.4	92.9	94.2	93.5	94.8	94.2	93.5	94.8
No. 4 (4.75 mm)	57.1	53.9	60.6	64.8	61.4	67.5	64.8	61.4	67.5
No. 8 (2.36 mm)	38.8	35.7	42.2	43.7	40.6	46.3	43.7	40.6	46.3
No. 16 (1.18 mm)	27.5	25.0	30.1	30.0	27.7	32.0	30.0	27.7	32.0
No. 30 (0.6 mm)	19.3	17.4	21.3	21.0	19.3	22.7	21.0	19.3	22.7
No. 50 (0.3 mm)	13.1	11.6	14.6	14.1	12.7	15.5	14.1	12.7	15.5
No. 100 (0.15 mm)	8.7	7.5	9.9	8.9	7.8	10.1	8.9	7.8	10.1
No. 200 (0.075 mm)	5.78	4.79	6.73	5.86	4.95	6.94	5.86	4.95	6.94

PG = performance grade; NMAS = nominal maximum aggregate size; RAP = reclaimed asphalt pavement; JMF = job mix formula; VTM = voids total mixture; Gsb = bulk (dry) specific gravity; Gse = effective specific gravity; Gmm = maximum theoretical specific gravity; Gmb = bulk specific gravity; VMA = voids in the mineral aggregate; Vbe = volume of effective binder; VFA = voids filled with asphalt; Pbe = percent of effective binder; FA ratio = dust-to-binder ratio.

APPENDIX B

PERFORMANCE DATA FOR CANTABRO, APA, AND IDT-CT

Table B1. Performance Data and Percent Change between the Design and Mixture Variations (see Equation 1)

Producer	Design	AC%, Gradation Type	Mass Loss (ML)			Rut Depth			CTindex		
			Avg. (%)	CV (%)	Change (%)	Avg.	CV (%)	Change (%)	Avg.	CV (%)	Change (%)
A	30R PG 64-22	O-J	8.3	21%		1.9	11%		59.1	21%	
		O-C	12.6	3%	50%	2.7	0%	39%	55.9	16%	-5%
		O-F	10.2	10%	22%	2.9	5%	54%	56.9	12%	-4%
		L-J	9.0	11%	8%	1.8	26%	-4%	51.6	19%	-13%
		H-J	7.2	7%	-14%	2.8	9%	44%	104.6	14%	77%
	40R PG 58-28	O-J	8.1	3%		2.2	5%		71.4	14%	
		O-C	12.0	6%	48%	2.4	11%	9%	60.5	15%	-15%
		O-F	9.6	8%	18%	3.1	15%	43%	68.4	15%	-4%
		L-J	11.4	22%	41%	1.8	17%	-17%	39.7	18%	-44%
		H-J	7.9	2%	-2%	2.6	15%	20%	85.2	18%	19%
	40R PG 64-22 RA	O-J	12.4	10%		2.1	33%		61.1	16%	
		O-C	12.2	5%	-2%	2.4	15%	13%	47.9	12%	-22%
		O-F	10.6	5%	-15%	2.2	9%	4%	53.7	18%	-12%
		L-J	11.9	9%	-4%	1.7	8%	-18%	34.2	19%	-44%
		H-J	9.2	15%	-26%	2.4	17%	11%	95.4	25%	56%
B	30R PG 64-22	O-J	8.2	3%		3.2	5%		45.6	14%	
		O-C	8.1	12%	-1%	2.9	5%	-9%	40.8	17%	-10%
		O-F	9.5	4%	15%	4.2	10%	29%	24.4	17%	-47%
		L-J	9.7	9%	18%	2.5	0%	-24%	31.7	10%	-30%
		L-C	9.6	13%	17%	2.8	7%	-13%	21.4	20%	-53%
		L-F	10.9	5%	33%	3.7	6%	16%	19.4	26%	-57%
		H-J	6.9	7%	-16%	3.8	22%	17%	66.6	20%	46%
		H-C	6.9	11%	-16%	3.3	11%	3%	63.0	11%	38%
		H-F	7.0	14%	-14%	4.6	3%	43%	41.2	15%	-10%
	40R PG 58-28	O-J	7.9	14%		5.5	24%		74.2	16%	
		O-C	8.4	28%	6%	4.3	6%	-22%	69.9	15%	-6%
		O-F	7.2	9%	-9%	6.4	13%	15%	58.5	10%	-21%
		L-J	9.5	2%	20%	4.0	7%	-27%	50.1	18%	-33%
		H-J	6.7	2%	-15%	6.8	5%	23%	113.7	22%	53%
	40R PG 64-22 RA	O-J	7.8	8%		3.4	7%		48.7	15%	
		O-C	6.3	11%	-19%	3.8	25%	14%	87.7	25%	80%
		O-F	6.9	5%	-11%	4.0	17%	18%	51.2	10%	5%
		L-J	6.4	7%	-18%	3.2	7%	-4%	39.6	26%	-19%
		H-J	5.8	6%	-25%	4.2	27%	25%	104.8	20%	115%
C	35R PG 58-28	O-J	7.1	9%		5.6	13%		211.9	17%	
		O-C	8.9	9%	26%	4.6	6%	-18%	195.2	12%	-8%
		O-F	6.2	13%	-12%	5.4	20%	-4%	210.6	18%	-1%
		L-J	9.9	10%	41%	5.1	21%	-9%	123.3	14%	-42%

		H-J	6.3	20%	-11%	5.4	12%	-5%	237.7	15%	12%
D	30R PG 64-22	O-J	11.3	12%		2.9	20%		60.7	16%	
		O-C	10.6	7%	-6%	2.3	17%	-20%	60.0	28%	-1%
		O-F	11.3	7%	0%	2.8	41%	-3%	67.7	9%	12%
		L-J	13.9	11%	23%	2.4	26%	-15%	47.1	14%	-22%
		H-J	8.9	7%	-21%	3.6	31%	24%	85.9	14%	42%
	40R PG 58-28	O-J	7.2	5%		3.1	25%		103.2	14%	
		O-C	8.9	6%	24%	3.1	25%	-2%	102.0	5%	-1%
		O-F	7.2	5%	0%	2.7	26%	-14%	117.4	10%	14%
		L-J	10.6	11%	47%	2.6	10%	-16%	64.3	13%	-38%
		H-J	6.4	6%	-11%	3.5	0%	11%	141.3	10%	37%
E	35R PG 58-28 RA	O-J	6.0	6%		6.0	13%		120.0	19%	
		O-C	7.0	3%	18%	6.0	19%	1%	131.5	28%	10%
		O-F	5.1	7%	-15%	6.4	4%	8%	108.7	18%	-9%
		L-J	6.3	7%	6%	4.7	15%	-20%	80.0	24%	-33%
		L-C	9.2	13%	54%	5.5	9%	-8%	102.9	12%	-14%
		L-F	5.9	6%	-2%	5.3	3%	-12%	72.6	15%	-40%
		H-J	5.2	9%	-13%	4.9	14%	6%	141.3	10%	18%
		H-C	4.4	12%	-26%	7.1	6%	19%	183.1	10%	53%
		H-F	3.7	5%	-38%	9.4	6%	57%	144.3	20%	20%
	35R PG 58-28 Softening Oil + Fibers	O-J	4.0	14%		5.4	14%		119.0	17%	
		O-C	3.7	1%	-6%	5.8	15%	7%	116.3	12%	-2%
		O-F	4.1	5%	3%	5.5	1%	2%	94.2	16%	-21%
		L-J	5.2	6%	31%	5.1	16%	-5%	82.8	10%	-30%
		H-J	3.6	28%	-9%	7.7	7%	41%	185.7	10%	56%
F	30R PG 64-22	O-J	6.2	4%		3.9	7%		76.4	11%	
		O-C	7.6	12%	22%	3.1	7%	-20%	89.1	7%	17%
		O-F	7.2	5%	17%	3.9	9%	1%	62.2	17%	-19%
		L-J	9.5	7%	54%	3.1	12%	-20%	62.7	26%	-18%
		H-J	5.2	3%	-15%	4.4	9%	15%	117.1	8%	53%
	40R PG 58-28	O-J	6.5	16%		4.1	2%		84.6	20%	
		O-C	6.1	13%	-5%	3.1	15%	-24%	81.4	11%	-4%
		O-F	5.7	22%	-11%	4.4	13%	8%	74.6	9%	-12%
		L-J	7.0	12%	8%	2.9	5%	-28%	62.6	24%	-26%
		H-J	4.7	4%	-27%	4.3	7%	6%	105.1	12%	24%
	40R PG 64-22 RA	O-J	5.7	13%		3.8	17%		81.0	15%	
		O-C	6.8	15%	20%	3.2	12%	-15%	77.5	5%	-4%
		O-F	4.3	4%	-25%	3.7	15%	-1%	57.0	18%	-30%
		L-J	7.6	14%	33%	3.0	17%	-20%	58.1	18%	-28%
		H-J	4.4	18%	-23%	4.5	9%	19%	94.0	10%	16%

Red cells indicate test values that did not meet the threshold requirement. AC = asphalt content; CV = coefficient of variation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

APPENDIX C

PERFORMANCE DATA FOR IDT-CT FOR CRITICALLY AGED MIXTURES

Table C1. Critical Aging IDT-CT Data

Producer	Design	Aging Condition	AC%, Gradation Type	CT_{index}	
				Avg.	CV (%)
B	40R PG 58-28	95°C for 48hrs	O-C	65.8	34%
			O-F	43.3	11%
	40R PG 58-28	135°C for 8hrs	O-J	33.8	22%
			O-C	41.0	20%
			O-F	33.3	28%
			L-J	21.5	17%
			H-J	43.2	11%
	40R PG 58-28	(short-term aging)	O-J	74.2	16%
			O-C	69.9	15%
			O-F	58.5	10%
			L-J	50.1	18%
			H-J	113.7	22%
D	40R PG 58-28	135°C for 8hrs	O-J	51.2	29%
	40R PG 58-28	(short-term aging)	O-J	103.2	14%
			O-C	102.0	5%
			O-F	117.4	10%
			L-J	64.3	13%
			H-J	141.3	10%

AC = asphalt content; CT = cracking tolerance; CV = coefficient of variation; PG = performance grade; Short-term aging = 4 hours at the compaction temperature; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

APPENDIX D

CANTABRO MASS LOSS DATA BY PRODUCER

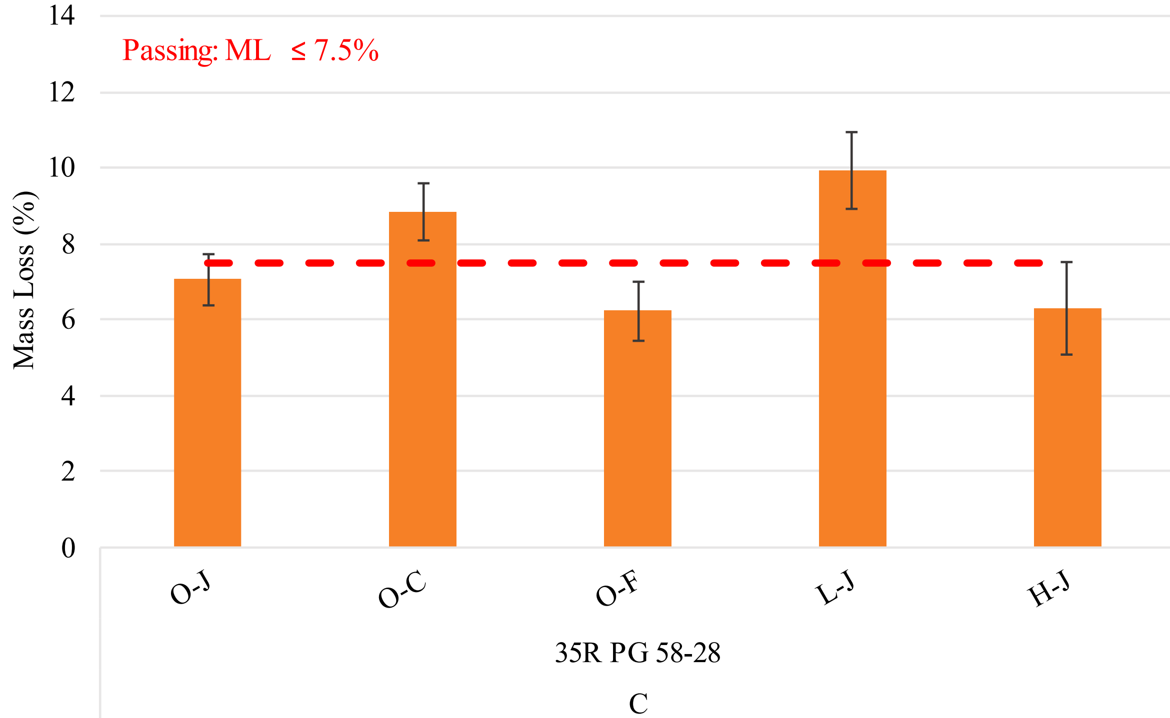

Figure D1. Producer C Mean Mass Loss Results. Error Bars = ± 1 standard deviation; PG = performance grade; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

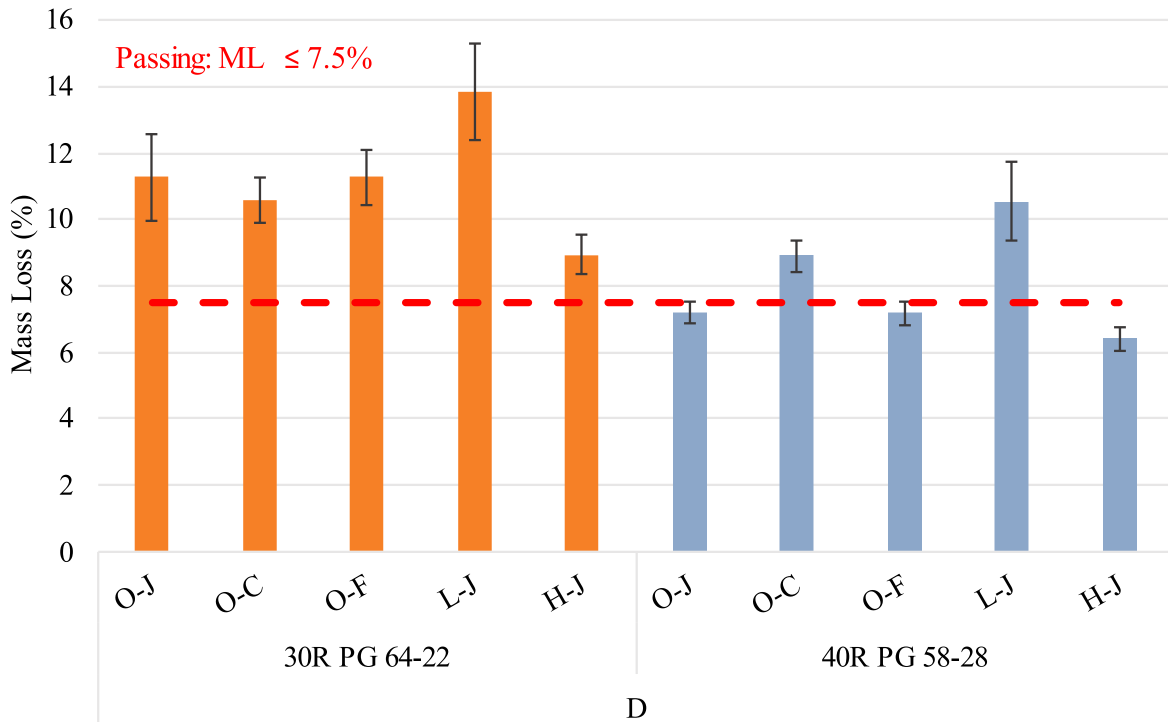

Figure D2. Producer D Mean Mass Loss Results. Error Bars = ± 1 standard deviation; PG = performance grade; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

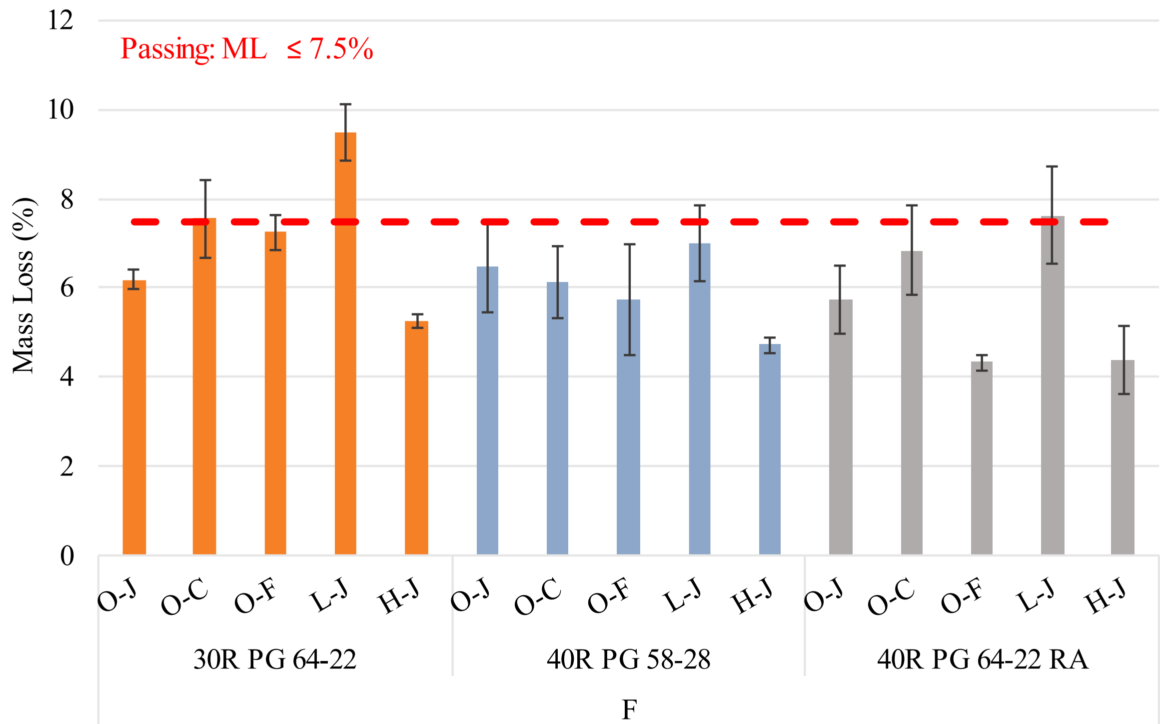

Figure D3. Producer F Mean Mass Loss Results. Error Bars = ± 1 standard deviation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

APPENDIX E

APA RUT DEPTH DATA BY PRODUCER

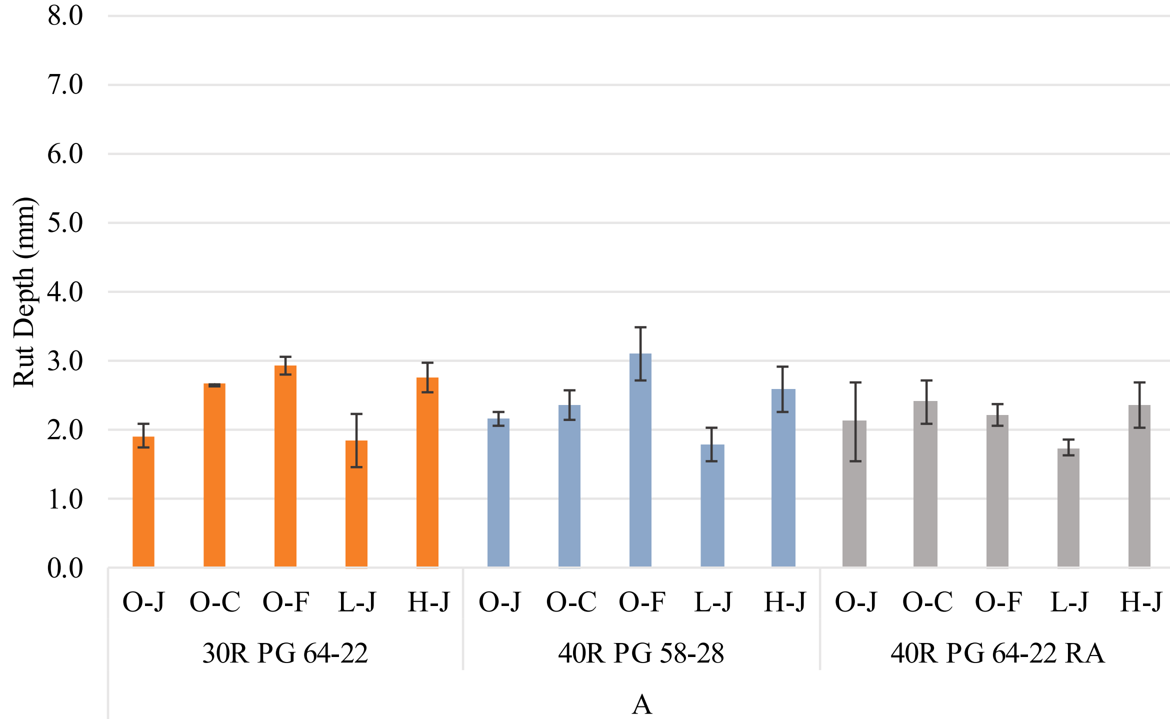

Figure E1. Producer A Mean Rut Depth Results. Error Bars = ± 1 standard deviation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

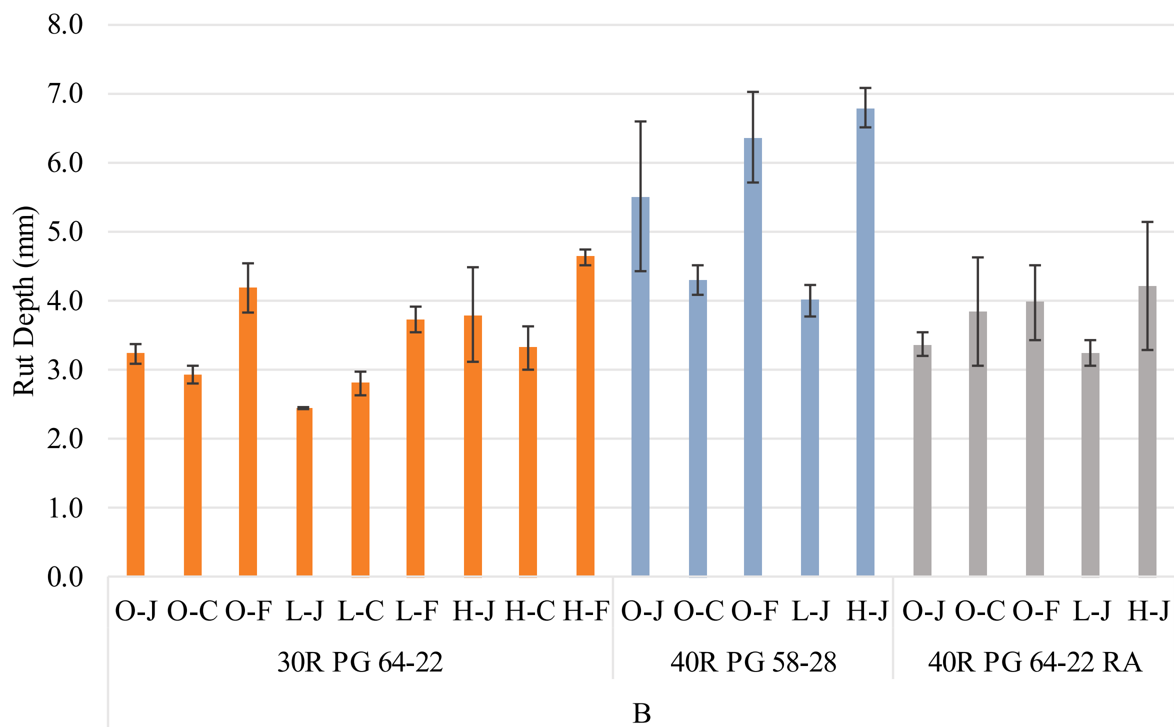

Figure E2. Producer B Mean Rut Depth Results Error Bars = ± 1 standard deviation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

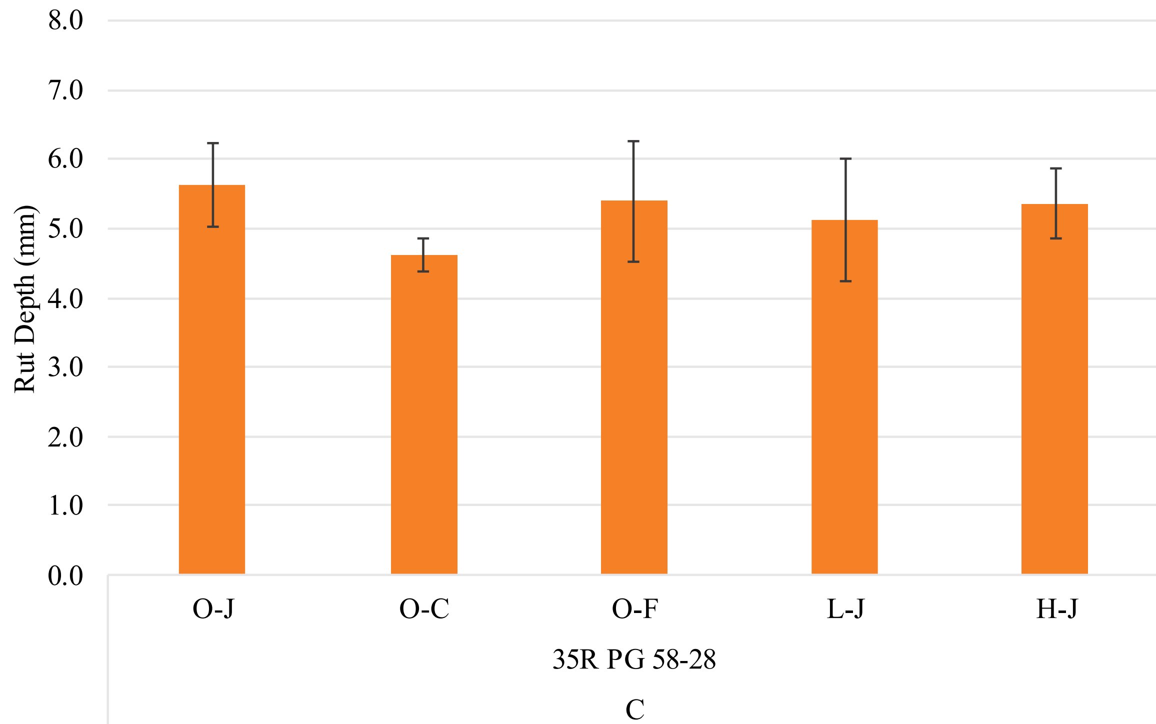

Figure E3. Producer C Mean Rut Depth Results. Error Bars = ± 1 standard deviation; PG = performance grade; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

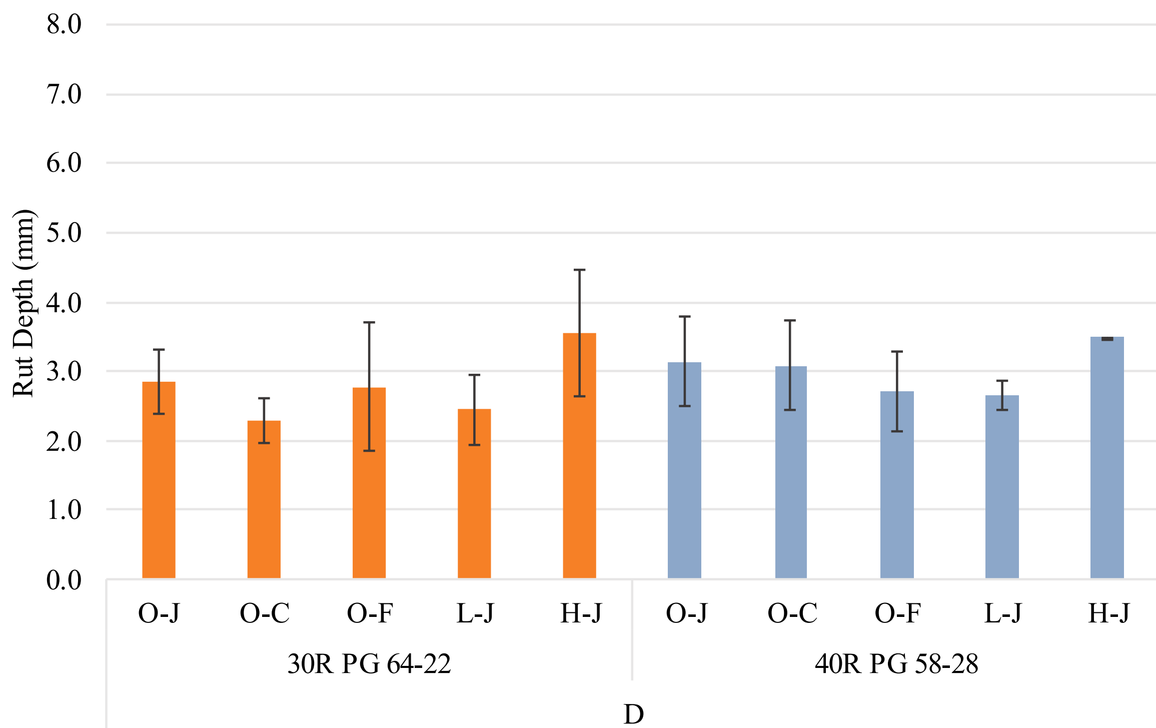

Figure E4. Producer D Mean Rut Depth Results. Error Bars = ± 1 standard deviation; PG = performance grade; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

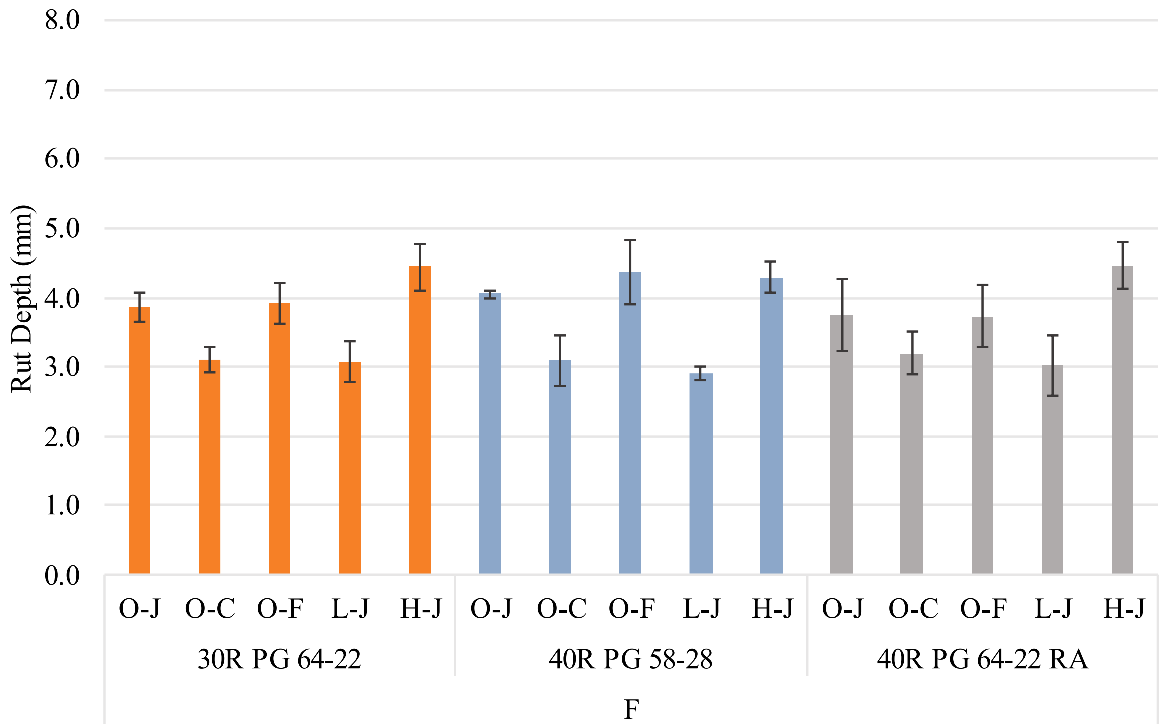

Figure E5. Producer F Mean Rut Depth Results. Error Bars = ± 1 standard deviation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

APPENDIX F

IDT-CT CT_{INDEX} DATA BY PRODUCER

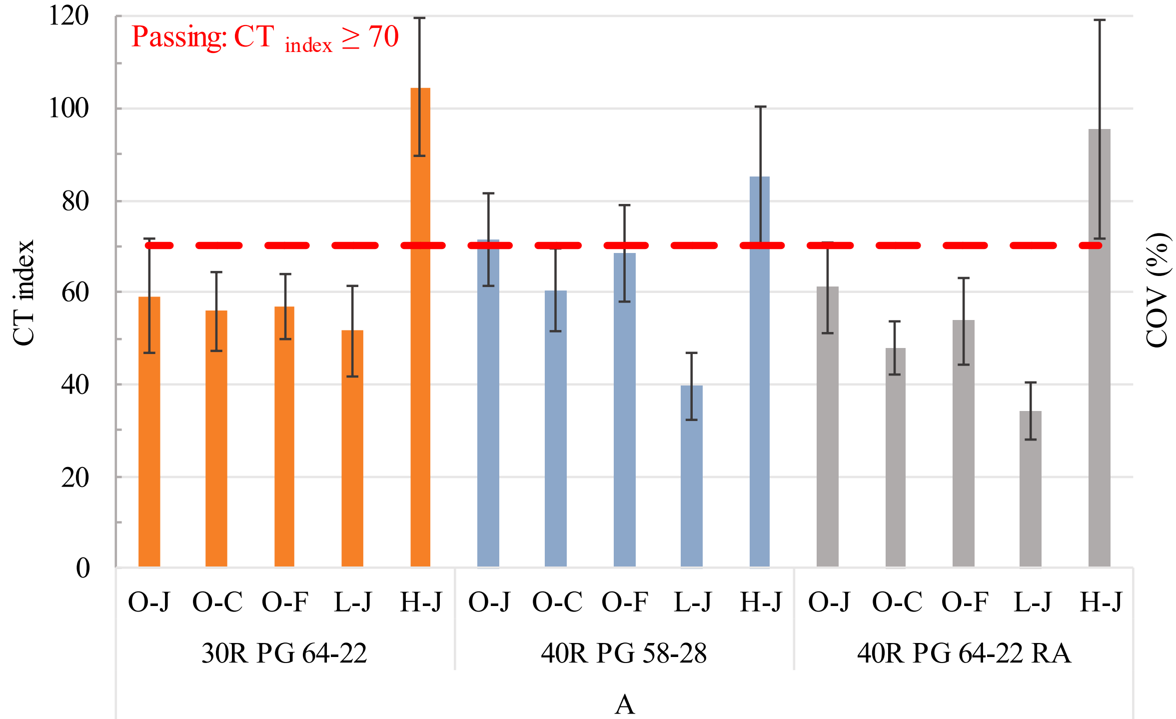

Figure F1. Producer A Mean CT$_{index}$ Results. Error Bars = ± 1 standard deviation; PG = performance grade; RA = recycling agent; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

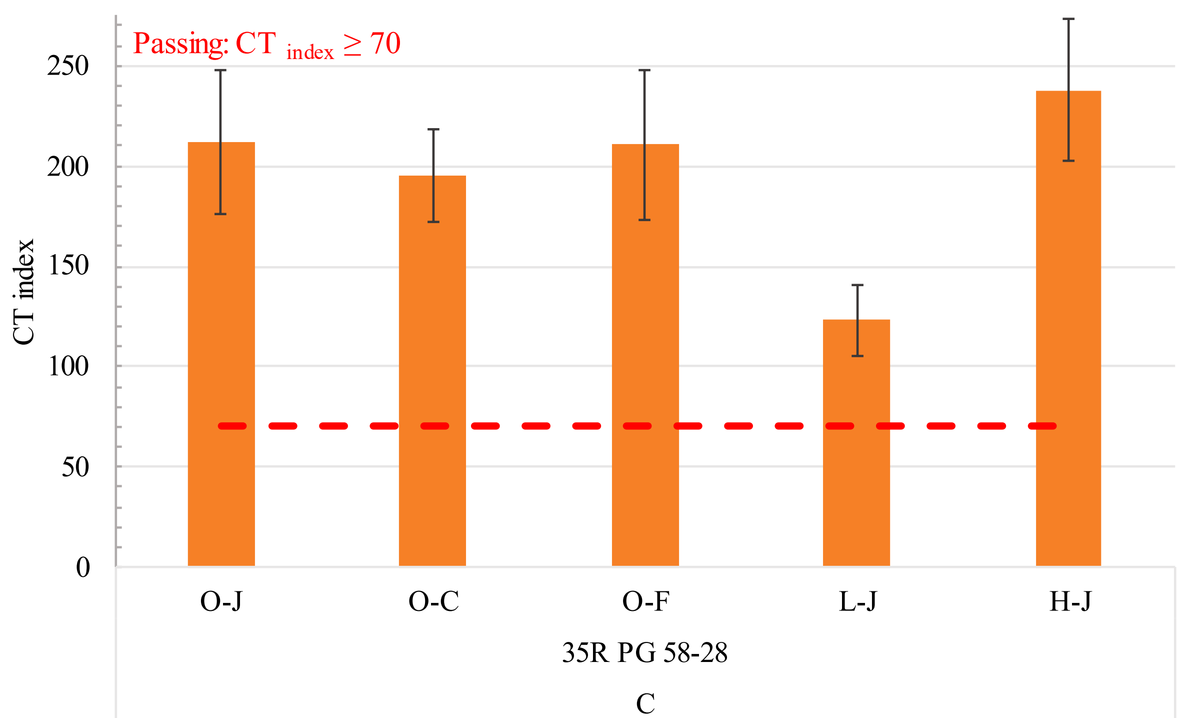

Figure F2. Producer C Mean CT$_{index}$ Results. Error Bars = ± 1 standard deviation; PG = performance grade; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.

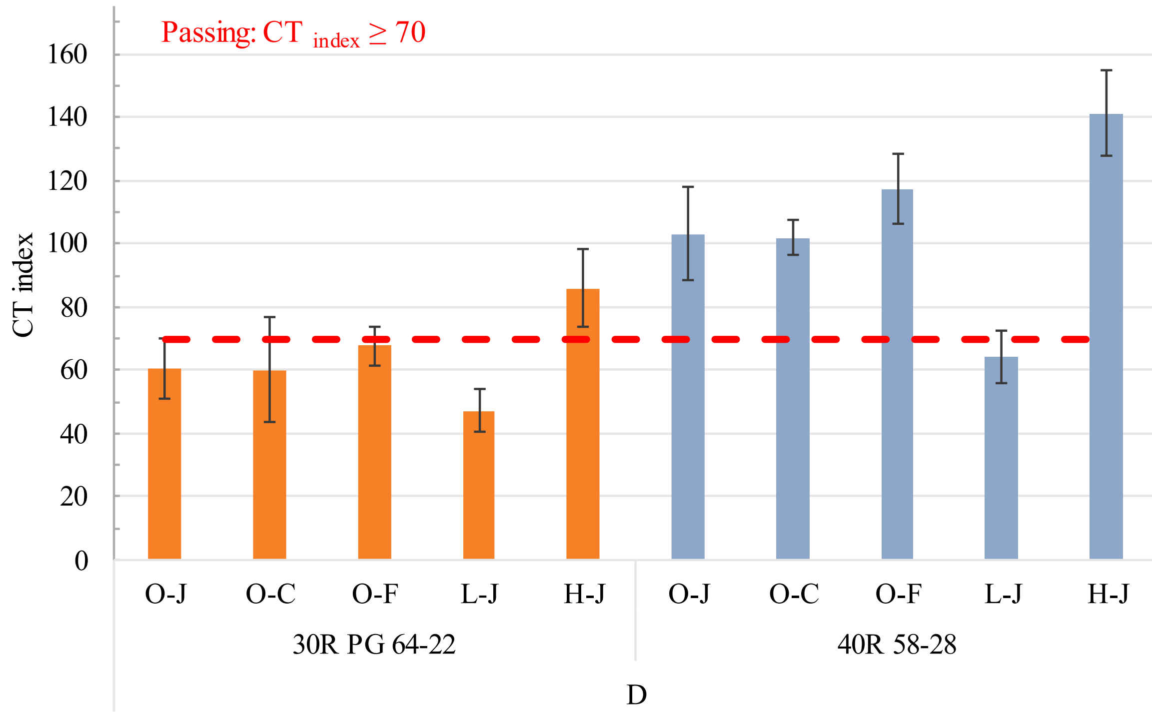

Figure F3. Producer D Mean CT_{index} Results. Error Bars = ± 1 standard deviation; PG = performance grade; O = optimum asphalt content; C = coarse gradation; F = fine gradation; L = low asphalt content; H = high asphalt content.