

Validation Techniques for Setting BMD Test Criteria



Case Study: *MaineDOT Benchmarking*



Objectives

Maine Department of Transportation (MaineDOT) is proactively advancing performance-based mix design (Balanced Mix Design, BMD) to improve rutting and cracking performance by aligning lab results with observed field outcomes across Maine's varying climates and traffic spectra. This case study highlights the MaineDOT benchmarking efforts to assess existing mixes with candidate tests and establish initial criteria. *Their goal was to select appropriate tests that were reasonably fast, repeatable, and matched perceived field performance.*

Benefit

MaineDOT initiated benchmarking in 2015 (expanded in 2019) to build a broad, representative baseline of performance for field-produced mixes across binder grades, nominal maximum aggregate sizes (NMAS), reclaimed asphalt pavement (RAP) contents, and traffic levels, supporting a transition toward BMD.

Background

The benchmarking covers all mixture designs and randomly selected designs over several years, ensuring broad representation without targeting specific variables.

Methodology

MaineDOT's benchmarking methodology was systematic, focusing on field-produced asphalt mixtures to establish baseline performance data.

- Performance Tests Evaluated:** [Hamburg Wheel Tracking Test](#) (HWTT, AASHTO T 324), [Indirect Tensile Asphalt Cracking Test](#) (IDEAL-CT, ASTM D8225), [Direct Tension Cyclic Fatigue Test](#) using the Asphalt Mixture Performance Tester (AMPT), [Stress Sweep Rutting](#) (SSR) test using the AMPT, [High Temperature Indirect Tension](#) (HT-IDT) strength (HT-ITS), and the [Rapid Shear Rutting Test](#) (IDEAL-RT, ASTM D8360).
- Sampling/Handling and Data Collection:** MaineDOT collected samples from plant production, reheating for testing without further aging to reflect field conditions. MaineDOT created Excel spreadsheets for individual tests and combined results. The dataset was extensive, including 1,500 HWTT, 3,200 IDEAL-CT, 1,600 IDEAL-RT, 350 HT-IDT strength, and 2,000 AMPT tests (cyclic fatigue and SSR) for over 100 mixtures.
- Analysis and Factors Considered:** binder grade (e.g., PG 64S-28, PG 64E-28, PG 70E-28), binder source and modification (polymer-modified vs. unmodified), mix type, NMAS, volumetric properties, aggregate source and RAP content (typically 0–20%, with higher RAP reducing cracking resistance), traffic levels, categorized by Average Annual Daily Traffic (AADT: <5,000, 5,000–10,000, >10,000).
- Participation in Round Robin Testing:** In collaboration with the Federal Highway Administration (FHWA) and the National Center for Asphalt Technology (NCAT), round-robin work quantified inter-lab reproducibility and compared production/compaction scenarios.

Results of the Study (To Date)

The benchmarking yielded significant findings, summarized in the following table:

Test	Key Metrics	Findings
HWTT	SIP, RRI, NRRRI, Rut Depth at 20,000 passes	Mixtures with rut depth < 12.5 mm and Stripping Inflection Point (SIP) > 15,000 passes showed good resistance. Normalized Rutting Resistance Index (NRRRI) >1 indicated compliance. Preliminary field correlation confirmed, especially for high-traffic mixes. Implementation decision: Adopt for rutting/moisture screening; criteria to be refined with local performance comparisons.
IDEAL-CT	CT _{Index} Fracture Energy (J/m ²)	Polymer-modified binders (e.g., PG 64E-28) improved CT _{Index} , with finer NMAS (9.5 mm) showing higher values. Higher RAP reduced values, mitigated by modified binders. Average CT _{Index} : 154 (AADT < 5,000), 128 (5,000–10,000), 125 (>10,000, excluding outliers). Implementation decision: Adopt for cracking screening with provisional lower bounds; continue to tune by climate subzones and project performance.
Cyclic Fatigue (AMPT)	Apparent Damage Capacity (Sapp)	Modified mixtures had Sapp 8-24, suitable for standard traffic (<10 million ESALs). ~20% unmodified mixtures fell below Sapp 8, indicating lower cracking resistance. No clear correlation with IDEAL-CT. Implementation decision: Not for routine use (due to test duration/cost) and unclear value for screening.
SSR (AMPT)	Rutting Resistance	Variability influenced by NMAS and binder PG, with ANOVA suggesting significant factors. Time-consuming, taking a week for results. Implementation decision: Pilot only / Not for routine use due to duration.
IDEAL-RT	RT _{Index} , Shear Strength	Added in later phases, 1,600 results showed rapid assessment potential for rutting, enhancing quality assurance (QA). Implementation decision: Adopt as a rapid QA/production check; develop compatibility rules with HWTT outcomes.

Recommendations

- **Climate and Traffic Considerations:** Maine’s moderate climate may slow asphalt aging compared to southern states, affecting rutting development rates. Proposed adjustments to the IDEAL-CT test temperature (e.g., 22°C coastal, 19°C inland) based on regional low-temperature PG grades (e.g., -28°C coastal, -34°C inland) were suggested for future research.

Challenges and Opportunities

- **Buy-in (DOT & industry):** DOT leadership expressed caution around pilot costs and schedule risk; industry management highlighted training needs to integrate BMD testing into QC plans. Focused communication and hands-on training can reduce perceived risk.
- **Time/cost of advanced tests:** AMPT cyclic fatigue and SSR require extended durations (≈1 to 2 weeks), impacting decision timelines. MaineDOT is prioritizing rapid tests (IDEAL-CT/IDEAL-RT/HWTT) for screening.
- **Validation Across Scenarios:** Correlating lab results with field performance under differing conditions remains a priority; ongoing collaboration with FHWA’s Mobile Asphalt Technology Center supports this need.

Conclusions and Recommendations

Benchmarking provided a cost-effective foundation for evaluating multiple performance tests, generating a strong evidence base, and a shortlist of implementable tests with provisional criteria. Because benchmarking uses existing mixtures and routine data collection, it provides agencies with an efficient, low-risk approach to identify reasonable starting thresholds before investing in more resource-intensive validation methods. These benchmark-derived criteria serve as initial guideposts and should be refined as test results from specific projects are compared with actual field performance, accounting for climate, traffic, materials, and local conditions.

References

- ["2023 Northeast Peer Exchange on Balanced Mix Design \(BMD\) Outcomes Summary,"](#) FHWA, March 2023.

- "[Framework of Asphalt Balanced Mix Design \(BMD\) for New England Transportation Agencies](#)," Mogawer et al., 2020.
- "[Balanced Mix Design Benchmarking of Field-Produced Asphalt Mixtures in Maine, U.S.](#)," Veeraragavan et al., 2021.
- [NAPA BMD Resource Guide](#)

Excerpt: Cumulative Distribution IDEAL-CT Test Results

Background: Role of Cumulative Distributions in Assessing Test Results

A cumulative distribution provides a clear way to evaluate how an entire population of test results performs relative to potential specification limits or benchmark thresholds. Instead of relying only on averages or ranges, a cumulative distribution displays the percentile at which certain performance levels occur, making it easier to see how mixtures compare, where performance begins to diverge between categories, and what proportion of mixtures meet or exceed a target value. This approach is especially useful in benchmarking, where agencies seek to understand typical performance trends, identify outliers, and quantify the benefits of factors such as binder grade or aggregate size.

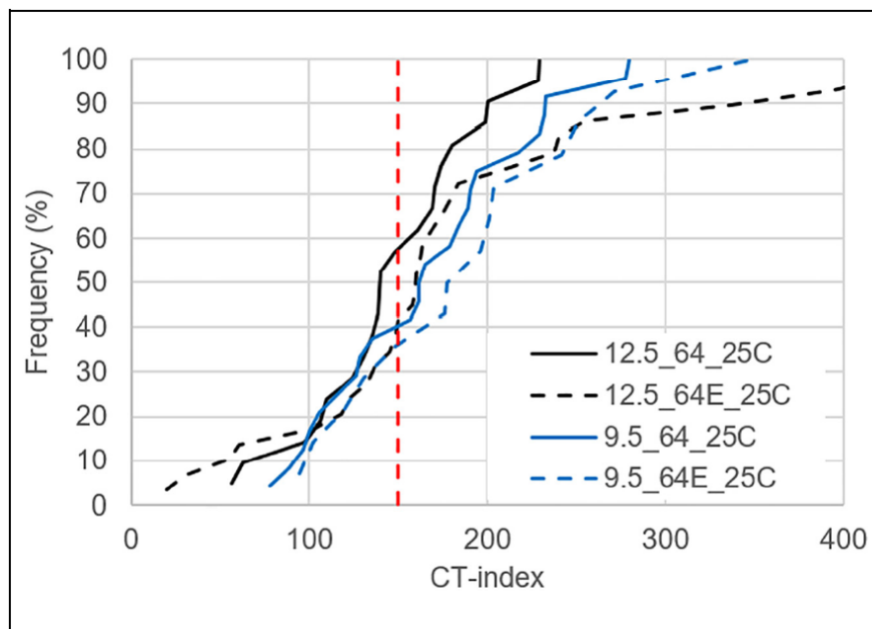


Figure 6. Cumulative distributions of indirect tensile cracking test (IDEAL-CT) test results.

Figure 6 presents cumulative distribution curves of IDEAL-CT cracking tolerance index (CT_{Index}) results, separated by binder grade (unmodified PG 64-28 vs. modified PG 64E-28) and nominal maximum aggregate size (12.5 mm vs. 9.5 mm).

In Figure 6, the cumulative curves were close to each other between the 20th and 40th percentiles or to a CT_{Index} value of approximately 150. After that point, the modified asphalt mixtures showed much higher CT_{Index} values than the unmodified asphalt mixtures. The effect of NMAS on the IDEAL-CT results may be attributed to an increase in binder content (P_b) for a finer mixture. Figure 6 also shows the preliminary minimum CT_{Index} requirement of 150 recommended in a previous study by Mogawer and Bennert for asphalt mixtures in Maine. (Veeraragavan, 2021)

Agency and Research Entities

