

Validation Techniques for Setting BMD Test Criteria



Foreword

Balanced Mix Design (BMD) validation means that laboratory test results have a strong, defensible relationship with actual field performance. Establishing validated thresholds is critical—if criteria are too loose, premature pavement failures or reduced performance may occur; if too stringent, viable mixtures may be eliminated, driving up costs. Clear, data-driven validation provides the balance needed for durable and cost-effective pavements.

The development and validation of test criteria for asphalt mixtures are critical to the successful implementation of BMD, ensuring long-term pavement durability, cost-effectiveness, and sustainability. Strategies such as benchmarking, open road test sections, modeling, heavy vehicle simulation (HVS), accelerated loading facilities (ALF), test tracks, pilot projects, and forensic studies each offer unique strengths and limitations. No single strategy fully addresses the diverse requirements posed by varying pavement types, traffic conditions, and environmental factors. Agencies must, therefore, strike a balance between rigorous evaluation and practical considerations, including time constraints and available resources. Integrating multiple Strategies can enhance confidence in the validation of outcomes while reducing risk.

This document provides concise guidance on available validation Strategies, highlights their advantages and challenges, and provides a framework for implementation. Supported by real-world case studies from state departments of transportation (DOTs) and other relevant agencies, this resource provides practical insights into successful validation strategies for BMD performance tests. Regardless of the strategy, three principles apply: (1) No single strategy is sufficient—multiple strategies should be combined, (2) Validation takes time, so efforts should start early, and (3) Criteria must encompass both acceptable and unacceptable performance ranges.

Validation Strategies

The strategies are presented in order of increasing cost and complexity, from initial screening tools to large-scale field efforts. All strategies involve testing mixtures with selected or proposed [laboratory tests](#), such as the [Hamburg Wheel Tracking Test](#) (HWTT, AASHTO T 324) to assess rutting resistance or the [Indirect Tensile Asphalt Cracking Test](#) (IDEAL-CT, ASTM D8225) to assess cracking resistance.



Benchmarking involves testing existing mixtures with selected or proposed tests to establish baseline metrics and evaluate whether tests can differentiate among typical agency mixes. Beyond serving as a cost-effective initial screening tool, it is critical for building stakeholder buy-in and industry awareness of BMD goals. Agencies like MaineDOT use cumulative distribution curves

during benchmarking to evaluate how an entire population of test results performs relative to proposed thresholds, thereby quantifying the proportion of current mixes that meet or exceed target values. While benchmarking provides an efficient, low-risk starting point, its initial thresholds lack field validation and must be refined through other strategies.

Modeling with Laboratory Data leverages laboratory tests and mechanistic-empirical (ME) models to predict pavement performance based on mechanical properties. This strategy serves as a "reasonableness check" for candidate test bounds derived from benchmarking, confirming that the proposed criteria align with predicted long-term distress for specific climates and traffic conditions.

For agencies like Virginia DOT, mechanistic modeling provides "ME anchors" that build confidence in rapid index tests by linking them to predicted field performance. However, its accuracy depends on high-quality input data and may oversimplify complex field interactions.



Heavy Vehicle Simulation (HVS)/Accelerated Loading Facilities (ALF) employ mobile or stationary equipment to apply controlled, repetitive loading to pavement sections, simulating years of traffic in a matter of months. This strategy accelerates pavement distress and bridges the gap between

laboratory and field conditions. However, it is expensive and does not fully capture the diverse field conditions or aging effects. It's best for evaluating mix performance under high-traffic conditions, validating specific failure modes, and testing experimental or innovative mixes in controlled settings.

Closed Test Tracks involve dedicated facilities where full-scale pavement sections are subjected to realistic yet managed traffic loading. They enable simultaneous monitoring of multiple designs and are ideal for comparing BMD mixtures with traditional volumetric designs in high-traffic or heavy-load applications. The Texas DOT and Oklahoma DOT case studies at the NCAT Test Track demonstrated that BMD can significantly extend the lifespan of overlays compared to traditional designs. Similar to HVS, these tracks accelerate traffic loading to shorten validation timeframes but do not fully reflect the long-term aging effects of a 20-year service life.



Open Road Test Sections entail constructing full-scale pavement sections on actual roadways to evaluate performance under real, yet consistent, traffic and environmental conditions over time. They provide valuable real-world data reflecting local conditions and long-term behavior, but are time-intensive, costly, and subject to variability in traffic, climate, and potential variations in support conditions. This strategy is best for validating mix designs for high-traffic or critical roadways. (See: [CAPRI Report](#))

Pilot Projects implement new BMD mixtures on a small scale in real-world settings to assess performance before widespread adoption. They offer practical data under actual conditions and reduce risk, but require extended monitoring and are often subject to variability in traffic, climate, and support conditions. Pilot projects are ideal for testing mixes, validating performance in specific climates, and building stakeholder confidence. Some pilot projects may involve substantial tonnages, depending on the scope.



Forensic Studies involve replicating mixtures of known field performance in the lab, both good and bad, to assess criteria. The availability of similar materials, binders, and aggregates can limit the ability to replicate original conditions, and results may be influenced by variability in construction practices or environmental factors. Forensic studies are best suited for investigating premature pavement failures.



Each Strategy offers unique advantages, but no single strategy fully addresses the diverse needs of asphalt mix validation (see Table 1). Benchmarking and modeling are cost-effective for initial screening and risk reduction, while open road test sections, test tracks, HVS/ALF, along with pilot projects, provide more robust data at higher costs and time commitments. The choice of Strategies depends on factors such as mix type, traffic conditions, environmental factors, agency resources, and project timelines. Combining multiple Strategies enhances confidence in performance criteria while balancing practical constraints.

Proposed Validation Framework

In 2023-24, the Federal Highway Administration (FHWA) hosted seven regional BMD peer exchanges (link to [TechBrief](#)). The primary motivation for State DOTs to move to BMD is that **volumetrics do not always yield optimal performance**. The Proposed Validation Framework demonstrates how the various validation strategies can work together (see Figure 1). It is organized into four phases, ensuring harmony between cost, time, and robust validation. This enables DOTs to confidently adopt BMD criteria tailored to specific traffic, environmental, and material conditions.



In the **assessment** phase, State DOTs evaluate current pavement performance and benchmark existing mixtures using proposed or selected BMD tests to establish baseline metrics. This step identifies gaps in current volumetric-based designs and informs the selection of appropriate tests. *“Assessment” = “Selecting the Right Tests”*

Setting initial criteria phase leverages literature reviews, benchmarking data from mixtures with known performance, and can also include mechanistic-empirical modeling to establish initial performance thresholds.





These thresholds are subsequently refined in the **initial validation** phase, where strategies such as open-road test sections, HVS/ALF, closed test tracks, pilot projects, and forensic studies are combined to confirm the criteria under real-world or controlled conditions.

The last phase is **monitoring and refining**. The BMD peer exchanges also identify a key management gap – “cost-benefit analysis.” Long-term monitoring is key to quantifying the benefits of BMD and to refining criteria in the future.

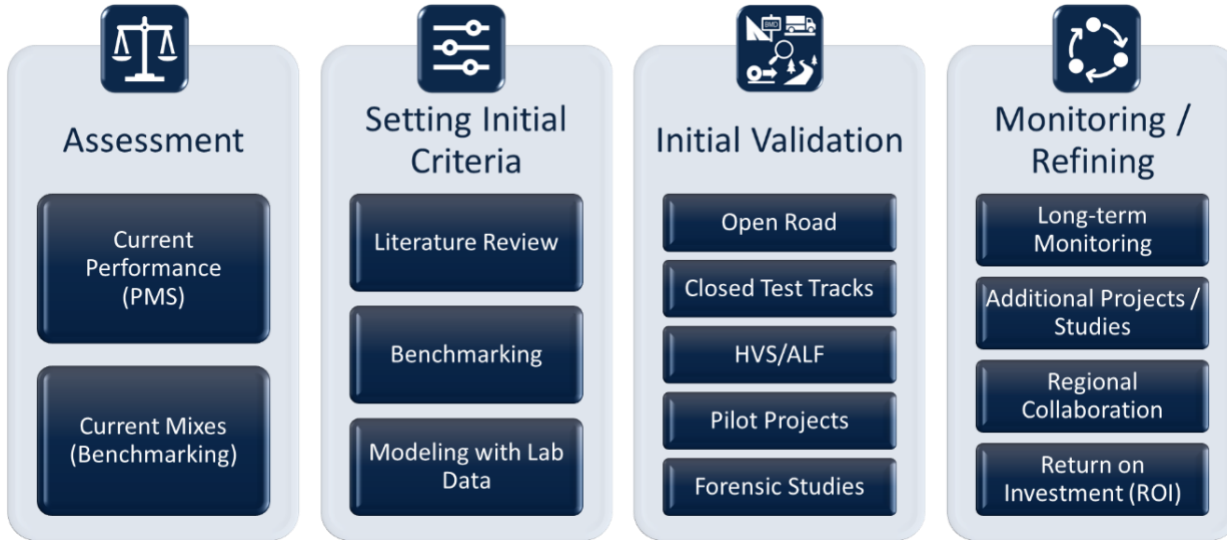


Figure 1. Proposed Validation Framework.

Table 1. Advantages, Disadvantages, and Limitations of Validation Strategies.

| | Bench- marking | Modeling w/ Lab | HVS/ALF | Closed Test Tracks | Open-Road Test Section | Pilot Project | Forensic Study |
|----------------------------------|-------------------|--------------------|---------|-----------------------|---------------------------|------------------|-------------------|
| Advantages | | | | | | | |
| Real-world Traffic | | | | | ✓ | ✓ | ✓ |
| Controlled Loading | | | ✓ | ✓ | | | |
| Real-world Environment | | | | | ✓ | ✓ | ✓ |
| Controlled Environment | | | ✓ | ✓ | | | |
| Long-Term Data Collection | | | | | ✓ | ✓ | |
| Relative Effectiveness | ★ | ★★ | ★★ | ★★★★ | ★★★★★ | ★★★★ | ★★ |
| Accelerated Testing | | | ✓ | ✓ | | | |
| Stakeholder Buy-In | 🔑 | | | | 🔑 | 🔑 | |
| Comprehensive Data | | | ✓ | ✓ | ✓ | | |
| Production Reheat Assessment | | | ✓ | ✓ | ✓ | ✓ | |
| Assess Innovative Materials | | | ✓ | ✓ | ✓ | | ✓ |
| Relative Cost | \$ | \$ | \$\$ | \$\$ | \$\$\$ | \$\$\$ | \$ |
| Disadvantages/Limitations | | | | | | | |
| Slow Data Accumulation | | | | | X | X | |
| Limited Control | X | | X | | X | X | X |
| Spatial Variability | | | | | X | X | X |
| Limited Rep. of Real-World | | X | X | X | | | |
| Limited Flexibility | X | X | | X | | | X |
| Complexity and Cost | | | | | | | |
| Granularity of Data | X | X | | | | X | X |
| Data Accuracy | | X | | | | | X |
| Tie to Field Performance | X | ? | X | | | | |
| Material Replication Challenges | | | | | | | X |

Case Studies of Validation Strategies: [Click Here.](#)

Current CAPRI Members



Join CAPRI-2: <https://pooledfund.org/Details/Solicitation/1642>

