Using Multi-Objective Optimization to Enhance Calibration of Performance Models in the *Mechanistic–Empirical Pavement Design Guide*

PUBLICATION NO. FHWA-HRT-17-104

JUNE 2018

U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

LTPP™

FOREWORD

This report documents research that applied Long-Term Pavement Performance Data to develop an improved approach to calibrating the American Association of State Highway and Transportation Officials' (AASHTO) AASHTOWare® Pavement ME Design performance models.[1] Whereas the current AASHTO guidelines used in the Pavement ME Design software for calibration of the performance prediction models to local conditions (e.g., materials, traffic, and climate) relies on single-objective minimization of bias and standard error (STE), this report investigates the use of multi-objective optimization to enhance the calibration of the performance models.

The multi-objective optimization results in a final pool of tradeoff solutions where none of the viable sets of calibration factors are prematurely eliminated. This report also demonstrates the application of engineering judgment and qualitative criteria to select reasonable calibration coefficients from the final pool of solutions that result from the multi-objective optimization. More reasonable calibration factors result in a more justifiable pavement design when considering multiple aspects of pavement performance. This investigation revealed that simply evaluating the bias and STE is not adequate for a comprehensive evaluation of performance prediction models. This report is intended for pavement engineers and State transportation departments.

Cheryl Allen Richter, Ph.D., P.E.
Director, Office of Infrastructure
Research and Development

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA-HRT-17-104	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Using Multi-Objective Optimization to Enhance Calibration of Performance Models in the *Mechanistic–Empirical Pavement Design Guide*		5. Report Date June 2018 6. Performing Organization Code
7. Author(s) Nima Kargah-Ostadi, Jose Rafael Menendez, and Mina Zhou		8. Performing Organization Report No.
9. Performing Organization Name and Address Fugro Consultants, Inc. 8613 Cross Park Drive Austin, TX 78754		10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTFH61-14-C-00025
12. Sponsoring Agency Name and Address Federal Highway Administration Office of Research, Development, and Technology Turner-Fairbank Highway Research Center 6300 Georgetown Pike McLean, VA 22101-2296		13. Type of Report and Period Covered Final report; July 2014–September 2016 14. Sponsoring Agency Code HRDI-30
15. Supplementary Notes The FHWA Contracting Officer's Representative was Deborah Walker (HRDI-30).		
16. Abstract This research study devised two scenarios for application of multi-objective optimization to enhance calibration of performance models in the American Association of State Highway and Transportation Officials (AASHTO) AASHTOWare® Pavement ME Design software.(1) In the primary scenario, mean and standard deviation of prediction error are simultaneously minimized to increase accuracy and precision at the same time. In the second scenario, model prediction error on data from Federal Highway Administration's Long-Term Pavement Performance test sections and error on available accelerated pavement testing data are treated as independent objective functions to be minimized simultaneously. The multi-objective optimization results in a final pool of tradeoff solutions, where none of the viable sets of calibration factors are eliminated prematurely. Exploring the final front results in more reasonable calibration coefficients that could not be identified using single-objective approaches. This report demonstrates the application of engineering judgment and qualitative criteria to select reasonable calibration coefficients from the final pool of solutions that result from the multi-objective optimization. More reasonable calibration factors result in a more justifiable pavement design considering multiple aspects of pavement performance. This investigation revealed that simply evaluating the bias and standard error is not adequate for a comprehensive evaluation of performance prediction models.		
17. Key Words *Mechanistic–Empirical Pavement Design Guide* (MEPDG), AASHTOWare® Pavement ME Design software, multi-objective optimization, calibration, validation, pavement performance models, evolutionary algorithms		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. http://www.ntis.gov

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	152	

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

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AADTT	average annual daily truck traffic
AAE	average absolute error
AASHTO	American Association of State Highway and Transportation Officials
AC	asphalt concrete
ANN	Artificial Neural Network
ANNACAP	Artificial Neural Networks for Asphalt Concrete Dynamic Modulus Prediction
APADS	Asphalt Pavement Analysis and Design System
APT	accelerated pavement testing
ARB	asphalt rubber binder
ATB	asphalt-treated base
AVC	Automatic Vehicle Classification
BAA	Broad Agency Announcement
BSG	bulk specific gravity
DASR	dominant aggregate size range
DOT	department of transportation
EA	evolutionary algorithm
EICM	Enhanced Integrated Climatic Model
ES	evolution strategy
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration
FWD	falling weight deflectometer
GA	genetic algorithm
GB	granular base
GPS	General Pavement Studies
GRG	generalized reduced gradient
GSA	global sensitivity analysis
HCD	historical climate data
HMA	hot-mix asphalt
HVS	Heavy Vehicle Simulator
IDE	integrated development environment
LTPP	Long-Term Pavement Performance
MEPDG	*Guide for Mechanistic–Empirical Design of New and Rehabilitated Pavement Structures*
MERRA	Modern-Era Retrospective Analysis for Research and Applications
MOEA	multi-objective evolutionary algorithm
NCHRP	National Cooperative Highway Research Program
NSGA	nondominated sorted genetic algorithm
NSI	Normalized Sensitivity Index
OAT	one-at-a-time
PG	performance grade
PLUG	Pavement Loading User Guide
PMA	polymer-modified asphalt

PMS	pavement management system
RAP	recycled asphalt pavement
RMSE	root-mean-squared error
RSM	response surface model
SG	specific gravity
SDR	Standard Data Release
SPS	Specific Pavement Studies
SSE	sum of squared errors
STE	standard error
TRF	traffic
VBA	Visual Basic for Applications
WIM	weigh in motion
XML	Extensible Markup Language

EXECUTIVE SUMMARY

The American Association of State Highway and Transportation Officials (AASHTO) has published guidelines to calibrate the pavement performance prediction models used in the AASHTOWare® Pavement Mechanistic–Empirical Design software to local materials–traffic–climate conditions.[1] The AASHTO recommended calibration process consists of several steps to increase accuracy and precision of the calibrated models through single-objective minimization of bias and standard deviation of error. This research project investigated the application of multi-objective optimization to enhance the calibration of the performance models within the *Guide for Mechanistic–Empirical Design of New and Rehabilitated Pavement Structures* (MEPDG).[2]

Using a multi-objective optimization approach enables researchers to escape preconception, avoid excessive concentration on only one aspect of the problem, and combine multiple sources of information in an objective manner. This research project devised two scenarios for application of multi-objective optimization to enhance calibration of MEPDG performance models. In the primary multi-objective scenario, mean and standard deviation of prediction error are simultaneously minimized to increase accuracy and precision at the same time. In this manner, the information from a single calibration run is fully implemented, and an additional round of computationally intensive calibration is avoided. In the second scenario, model prediction error on data from Federal Highway Administration's Long-Term Pavement Performance test sections and error on available accelerated pavement testing (APT) data are treated as independent objective functions to be minimized simultaneously. As a result, the multiple sources of data with different materials–traffic–climate conditions and disparate measurement protocols are combined in an objective manner.

Using this multi-objective calibration approach results in a final pool of tradeoff solutions. This way, none of the viable sets of calibration factors are eliminated prematurely, and all of the nondominated solutions are included in the final tradeoff front. Exploring the final front might reveal unknown aspects of this calibration problem and result in more reasonable calibration coefficients that could not be identified using single-objective approaches. This study demonstrates the application of engineering judgment and qualitative criteria to select reasonable calibration coefficients from the final pool of solutions that result from the multi-objective optimization. More reasonable calibration factors result in a more justifiable pavement design considering multiple aspects of pavement performance.

Although there was no fundamental way to prove whether there was a theoretical conflict between the selected objective functions, the shape of the final nondominated front indicated that the selected objective functions conflicted with one another, and therefore, the application of a multi-objective optimization approach was justified. In the first multi-objective calibration scenario, the simultaneous minimization of bias and standard error (STE) resulted in calibrated models that had higher precision (lower STE) and higher generalization capability (lower difference in bias between calibration and validation data) compared to the single-objective calibration. While this scenario was more successful in the calibration of rutting models for overlaid pavements on Florida Specific Pavement Studies (SPS)-5 data, it did not result in desirable accuracy levels for rutting models on new pavements using Florida SPS-1 data. The

results of the second multi-objective scenario demonstrated that incorporation of the disparate source of performance data (Florida Department of Transportation APT data) as a separate objective function has significantly improved the prediction accuracy, precision, and generalization capability of the calibrated rutting model on SPS-1 data.

The qualitative comparison of the calibrated models showed that using the multi-objective approach has resulted in predicted rutting distributions that are more similar in flatness (kurtosis) to the measured rutting distributions. However, the same was not true about skewness. The low goodness-of-fit indicator for scatterplots of predicted versus measured rutting in the case of all calibration approaches reveals that the MEPDG rutting models have an inherent lack of precision that might not be addressed with the calibration process. This is perhaps because the variability in pavement materials has not been captured in these models. The final selected calibration factors for rutting in unbound pavement layers (base and subgrade) were more reasonable in the multi-objective approach, compared to insignificant values achieved through single-objective calibration. Once again, this possibility of applying engineering judgment demonstrates the value of the multi-objective calibration in providing a final pool of solutions from which to choose.

To combine the quantitative and qualitative success metrics of a performance prediction model, the measured and predicted rutting deterioration trends were examined. While the two-objective calibration on SPS-5 data had significantly improved the prediction of rutting deterioration rates compared to the single-objective calibration, the multi-objective calibration results on SPS-1 did not exhibit the same quality. This investigation revealed that simply evaluating the bias and STE is not adequate for a comprehensive evaluation of performance prediction models. Therefore, it is recommended that the comprehensive comparison framework presented in this study be used when selecting suitable performance prediction models.

CHAPTER 1. INTRODUCTION

BACKGROUND

The *Guide for Mechanistic–Empirical Design of New and Rehabilitated Pavement Structures* (MEPDG) was developed under the National Cooperative Highway Research Program (NCHRP) Project 1-37A.[2] This pavement design methodology is based on various performance prediction models that relate mechanistically calculated pavement responses to empirically measured field performance. The performance models in the current American Association of State Highway and Transportation Officials (AASHTO) AASHTOWare® Pavement ME Design software were calibrated based on Federal Highway Administration's (FHWA's) Long-Term Pavement Performance (LTPP) data from all North American regions with various material characteristics, traffic patterns, and climatic conditions.[1]

Numerous State agencies have adopted the AASHTO recommended calibration procedure developed in NCHRP Project 1-40B for their State or regional conditions.[3,4] This calibration approach minimizes the sum of squared errors (SSE) between measured and predicted pavement performance. Different States have reported calibration results with varying levels of success depending on the performance models. In this single-objective optimization approach to calibration, the SSE and, if needed, the standard deviation of error, is minimized in two separate steps.

Using a multi-objective optimization approach enables researchers to escape preconception, avoid excessive concentration on only one aspect of the problem, and combine multiple sources of information in an objective manner. In this study, application of a multi-objective optimization approach to enhance calibration of the AASHTOWare® Pavement ME Design software performance models is investigated. To demonstrate the multi-objective approach, a subset of the LTPP data from Florida is selected for calibration of MEPDG permanent deformation (rutting) prediction models for new and rehabilitated flexible pavements. In addition to the LTPP data, some accelerated pavement testing (APT) data were received from the Florida Department of Transportation (FDOT) State Materials Office. This project was awarded by FHWA as a result of a proposal in response to the FHWA Broad Agency Announcement (BAA) on analysis of LTPP data.

RESEARCH OBJECTIVES AND OUTCOMES

The proposed research is aimed at calibration and validation of pavement performance prediction models, which was considered as an objective (research area d.i) in the BAA. Employing LTPP data, this research effort will support strategic objective number 5, "Development of Pavement Response and Performance Models Applicable to Pavement Design and Performance Prediction," of the current LTPP Strategic Plan for LTPP Data Analysis. Through application of alternative computational tools, calibration of pavement performance models is enhanced. The following enhancements are anticipated as a result of this research effort:

- Increase utilization of the available information in each iteration of the computationally intensive calibration process by simultaneously increasing accuracy and precision (minimizing mean and standard deviation of error) of the performance models.

- Combine multiple disparate sources of data into the calibration process in an objective manner.

- Provide engineering judgment and qualitative criteria to select reasonable calibration coefficients from the final pool of solutions that result from a multi-objective optimization.

Using this multi-objective calibration approach, multiple sources of information are incorporated in an objective manner, resulting in a final set of tradeoff solutions. This way, none of the viable sets of calibration factors are eliminated prematurely, and all of the nondominated solutions are included in the final tradeoff front. Exploring the final front might reveal unknown aspects of this calibration problem and result in more reasonable calibration coefficients that could not be identified using single-objective approaches.

If successful, a superior calibration of the ME software performance models is realized using multi-objective optimization compared to conventional single-objective methods. Of particular value, this study demonstrates how State and local agencies can adopt this multi-objective approach to determine more reasonable calibration factors for their pavement networks at all age and distress levels. This contribution can result in more economical and justifiable pavement design considering multiple aspects of pavement performance.

REPORT ORGANIZATION

Following an overview of the corresponding literature in chapter 2, this report discusses the extraction of relevant LTPP and APT data and generating the ME software input files in chapter 3. Chapter 4 explains the programming approach to single-objective (according to AASHTO guidelines) and multi-objective calibration. Chapter 5 presents the calibration results and discusses a comprehensive comparison of the final performance models calibrated through the single-objective and multi-objective approaches. Finally, chapter 6 concludes with insights from the discussion of results and provides some recommendations for future implementation of this multi-objective calibration approach and further research to enhance the methodology.

CHAPTER 2. REVIEW OF LITERATURE ON CALIBRATING THE MECHANISTIC–EMPIRICAL PAVEMENT PERFORMANCE MODELS

INTRODUCTION

The first step in the proposed research study was a comprehensive literature review regarding calibration of the pavement performance models in the AASHTOWare® Pavement ME Design software. The literature review also included multi-objective model calibration studies in research areas other than pavement engineering. The major objective of this literature review was to identify important sources of information for model calibration and to formulate corresponding objective functions. The review also enabled researchers to base their selection of range and precision of possible calibration factors on previous calibration studies.

Results of previous calibration efforts indicated little to no problem in calibration of thermal (transverse) cracking and smoothness prediction models. Therefore, the existing single-objective calibration procedure seems to be sufficient for these two models.(3) There seem to be difficulties in calibration of the longitudinal cracking model associated with the lack of fit of the global model.(5) These difficulties require reconsideration of the formulation for the longitudinal cracking model, and therefore, this model is not considered for this research project.

The permanent deformation model has been reported to consistently overpredict measured pavement rutting. On the other hand, the fatigue cracking model has been reported to underpredict actual pavement distress in most studies. A more sophisticated calibration procedure could address these consistent deviations of model predictions from measured pavement performance. This research project was therefore originally focused on local calibration of prediction models for rutting and fatigue cracking in flexible pavements. Hence, the following literature review includes information regarding both models. However, due to limitations in the resources and schedule of this project, only the permanent deformation models were considered to demonstrate the proof of concept for multi-objective calibration.

MECHANISTIC–EMPIRICAL PAVEMENT PERFORMANCE MODELS

The amount of total permanent deformation in flexible pavements is calculated as the sum of plastic deformations in each of the hot-mix asphalt (HMA), base, and subgrade layers. The model for predicting rutting (permanent deformation) in HMA layers (inches)[1] of flexible pavements has the form of equation 1:(1)

$$\Delta_{p(HMA)} = h_{HMA}\varepsilon_{p(HMA)} = h_{HMA}\varepsilon_{r(HMA)}\beta_{r1}k_Z 10^{k_1}T^{k_2\beta_{r2}}N^{k_3\beta_{r3}} \quad (1)$$

Where:

Δ_p = predicted rutting (inches).
h_{HMA} = thickness of the HMA layer (inches).
ε_p = plastic strain in the layer (inch/inch).

[1]For consistency with how measurements are recorded in the LTPP database, all layer thickness measurements are presented in inches in this report. These measurements can be converted to centimeters: 1 inch = 2.54 cm.

ε_r = resilient (recoverable) strain in the layer (inch/inch).
T = layer temperature.
N = number of load repetitions.
k_1, k_2, k_3 = global field calibration parameters (from NCHRP 1-40D Recalibration, k_1 = –3.35412, k_2 = 1.5606, k_3 = 0.4791).
β_{r1}, β_{r2}, β_{r3} = local or mixture field calibration factors; these factors were all set to 1.0 for the global calibration.
k_z = depth confinement factor, which is calculated through equation 2:

$$k_Z = (C_1 + C_2 D) \times 0.328196^D \quad (2)$$

Where:
D = depth below the surface
C_1 and C_2 = coefficients to calculate the depth confinement factor; these coefficients are calculated according to equations 3 and 4:

$$C_1 = -0.1039 \times {h_{HMA}}^2 + 2.4868 \times h_{HMA} - 17.324 \quad (3)$$

$$C_2 = 0.0172 \times {h_{HMA}}^2 - 1.7331 \times h_{HMA} + 27.4280 \quad (4)$$

The model for predicting rutting in unbound (base, subbase, and subgrade soil) layers (inches) of flexible pavements has the form of equation 5:[1]

$$\Delta_{p(soil)} = h_{soil}\,\varepsilon_{p(soil)} = h_{soil}\,\beta_{s1} k_1 \varepsilon_v \left(\frac{\varepsilon_0}{\varepsilon_r}\right) e^{-\left(\frac{\rho}{N}\right)^{\beta}} \quad (5)$$

Where:
h_{soil} = thickness of the unbound layer/sublayer (inches).
k_1 = global calibration coefficient; k_1 = 2.03 for granular materials, and k_1 = 1.35 for fine-grained materials.
β_{s1} = local calibration factor; this factor was set to 1.0 for the global calibration; it is also called β_{GB} for unbound base layers and β_{SG} for subgrade layers.
ε_v = average vertical resilient or elastic strain in the layer (inch/inch) calculated by the structural response model.
ε_0 = strain intercept (inch/inch) determined from laboratory repeated-load permanent deformation tests.
ε_r = resilient (recoverable) strain (inch/inch) imposed in laboratory test to obtain material properties.

$\frac{\varepsilon_0}{\varepsilon_r}$ = strain ratio that is calculated using equation 6:

$$log\left(\frac{\varepsilon_0}{\varepsilon_r}\right) = \frac{\left(e^{(\rho)^{\beta}} . a_1 M_r^{b_1}\right) + \left(e^{(\rho/10^9)^{\beta}} . a_9 M_r^{b_9}\right)}{2} \quad (6)$$

β and ρ are material properties that are calculated according to equations 7 and 8:

$$\log(\beta) = -0.61119 - 0.017638(W_c) \tag{7}$$

$$\rho = 10^9 \left(\frac{C_0}{(1-(10^9)^{\beta})}\right)^{\frac{1}{\beta}} \tag{8}$$

Where W_c is water content (%) that is calculated using equation 9:

$$W_c = 51.712 \left[\left(\frac{M_r}{2555}\right)^{\frac{1}{0.64}}\right]^{-0.3586.GWT^{0.1192}} \tag{9}$$

Where:

GWT = depth of ground water table (ft).

C_0 = factor depending on the material resilient modulus and is calculated through equation 10:

$$C_0 = Ln\left(\frac{a_1 M_r^{b1}}{a_9 M_r^{b9}}\right) = 0.0075 \tag{10}$$

Where:

M_r = resilient modulus of the unbound layer/sublayer (psi).

a_1, a_9, b_1, b_9 = regression constants; a_1 = 0.15, a_9 = 20.0, b_1 = 0.0, and b_9 = 0.0.

The prediction model for fatigue (bottom–up or alligator) cracking (percent of total lane area) in flexible pavements has the form of equation 11:[1]

$$FC_{Bottom\ -up} = \left(\frac{6000}{1+e^{\left(C_1 \times C_1^* + C_2 \times C_2^* Log\left(100 \times DI_{bottom\ -up}\right)\right)}}\right) \times \left(\frac{1}{60}\right) \tag{11}$$

Where:

$FC_{Bottom–up}$ = bottom–up alligator cracking.

C_1^* and C_2^* are coefficients that can be calculated using equations 12 and 13:

$$C_1^* = -2 \times C_2^* \tag{12}$$

$$C_2^* = -2.40874 - 39.748(1+h_{HMA})^{-2.856} \tag{13}$$

Where

h_{HMA} = total HMA thickness.

$DI_{bottom–up}$ = damage index that is calculated using equation 14:

$$DI_{bottom\ -up} = \sum (\Delta DI)_{j,m,l,p,T} = \sum \left(\frac{N}{N_{f-HMA}}\right)_{j,m,l,p,T} \tag{14}$$

Where:

ΔDI = incremental damage index.
N = actual number of axle load applications within a specific period.
j = axle load interval.
m = axle load type (single, tandem, tridem, quad, or special axle configuration).
l = truck type using the truck classification groups included in the MEPDG.
p = month.
T = median temperature for the five temperature intervals used to subdivide each month.
N_{f-HMA} = allowable number of axle load applications for a flexible pavement to fatigue cracking, and it is calculated using equation 15:

$$N_{f-HMA} = k_{f1}(C)(C_h)\beta_{f1}(\varepsilon_t)^{k_{f2}\beta_{f2}}(E)^{k_{f3}\beta_{f3}} \quad (15)$$

Where:

ε_t = tensile strain at the critical location.
E = dynamic modulus measured in compression.
k_{f1}, k_{f2}, k_{f3} = global field calibration parameters (from NCHRP 1-40D Recalibration, k_{f1} = 0.007566, k_{f2} = –3.9492, k_{f3} = –1.281).[6]
β_{f1}, β_{f2}, β_{f3} = local or mixture field calibration factors; these factors were all set to 1.0 for the global calibration.
C = constant depending on mix properties and calculated using equations 16 and 17:

$$C = 10^{M} \quad (16)$$

$$M = 4.84\left(\frac{V_{be}}{V_a + V_{be}} - 0.69\right) \quad (17)$$

Where:

V_a = air voids at the time the roadway is opened to traffic (%).
V_{be} = effective asphalt content by volume of the mix placed on the roadway (%).
C_h = thickness correction term, and it is calculated using equation 18:

$$C_h = \frac{1}{0.000398 + \frac{0.003602}{1 + e^{(11.02 - 3.49 \times h_{HMA})}}} \quad (18)$$

The local calibration procedure is aimed at determining the calibration factors that minimize the difference between measured and predicted pavement performance. This process includes reducing bias through minimization of average prediction error and lessening error variation through reduction of the standard deviation of error. Table 1 lists the calibration factors or coefficients that need to be determined in the model calibration process for rutting and fatigue cracking in flexible pavements.

Table 1. Calibration factors in prediction models for rutting and fatigue cracking in flexible pavements.[3]

Performance Model	Calibration Objective: Reduce Bias	Calibration Objective: Reduce STE
Permanent deformation	k_1, β_{r1}, β_{GB}, and/or β_{SG}	k_2, k_3 and β_{r2}, β_{r3}
Fatigue cracking	C_2 or β_{f1}	β_{f2}, β_{f3} and C_1

STE = standard error.

Based on the NCHRP Project 1-40B, the corresponding calibration factors in table 1 were found to be contributing to bias and standard error (STE).[4] The current single-objective calibration procedure determines the calibration factors in two steps corresponding to "eliminating" bias and reducing STE, respectively.[3] However, the multi-objective calibration approach in this research project will involve the determination of optimum values for all calibration factors to reduce bias and STE at the same time.

INPUT VARIABLES

A very important task in calibration and implementation of AASHTOWare® Pavement ME Design software is selection of accurate values for input variables. Three main categories of pavement structure, climate, and traffic variables require ample efforts to determine corresponding values for every design project. By the same token, many State agencies have sponsored research efforts to characterize local pavement materials, determine local climatic data, and classify local traffic patterns. In fact, several State agencies have developed databases or software that specify corresponding values for each input variable to be used in the implementation of AASHTOWare®.[7,8]

The majority of the States have used LTPP data in combination with their State pavement management system (PMS) database to develop their MEPDG calibration database.[9,10] Differences in distress identification protocols between LTPP and State PMS surveys are a source of concern regarding the combination of these data sources to be used in model calibration efforts. Some States have addressed this issue by interpreting their distress data according to the *Distress Identification Manual for the Long-Term Pavement Performance Program* and using the transformed data.[11]

There are three levels of data precision (hierarchical input levels) for MEPDG input variables. Level 1 input values are site-specific data based on laboratory or field measurements that are the most accurate values. Level 2 values are derived based on correlations with other locally measured parameters or available historical data that were not necessarily measured at the specific site. Level 3 data are the default values that were established based on national averages, correlations, or both. Depending on the sensitivity of the predicted output to each input variable, it is important to use level 1 data when available.

Regarding asphalt material characterization in the MEPDG performance models, the most important (influential) input variable is the dynamic modulus of HMA. FHWA has developed software based on Artificial Neural Network (ANN) models to populate the LTPP database with dynamic modulus data.[12] Several State departments of transportation (DOTs) have also conducted HMA material characterization studies to determine asphalt binder and mixture

properties to be used as level 1 (agency-specific) input values in AASHTOWare® Pavement ME Design software.[13] One of the key efforts in these studies was an evaluation of the Witczak model for calculation of dynamic modulus. Most of these studies found the Witczak model to produce reasonable predictions for dynamic modulus of HMA with conventional binders and mixtures. However, further modifications were required for binders with higher performance grades (PGs) and nonconventional mixtures, such as high recycled asphalt pavement (RAP) content, stone-matrix asphalt, cold-recycled asphalt, and warm-mix asphalt mixtures.

Most of the studies on characterization of unbound materials in flexible pavements have focused on determining resilient modulus values for typical granular aggregate base materials and local subgrade soils.[13] Several studies have also developed a resilient modulus prediction model based on soil parameters. In addition, falling weight deflectometer (FWD) and other nondestructive test results have been implemented to determine the resilient modulus values. The LTPP database contains repeated load resilient modulus test results, and FWD measured deflections that could be utilized in this regard.

The LTPP database contains extensive climatic data either measured at LTPP sites or estimated from adjacent weather stations. The impact of climatic and environmental parameters on material properties of unbound pavement layers is captured using the Enhanced Integrated Climatic Model (EICM) in MEPDG. Several studies have evaluated the predictions of EICM with test data.[13] Change in resilient modulus values due to seasonal variations and behavior of unsaturated soils is another topic currently under research in this area.

Traffic data inputs for the AASHTOWare® Pavement ME Design software have been calculated and are accessible for LTPP sites. LTPP data have been utilized to establish level 3 traffic inputs for the MEPDG. Several States have developed agency-specific traffic data and axle load spectra.[13] Some have also developed customized software to calculate MEPDG traffic inputs from weigh-in-motion (WIM) data.

SENSITIVITY ANALYSIS OF MEPDG PERFORMANCE MODELS

Sensitivity analysis of performance prediction models is a qualitative assessment that can be implemented for multiple purposes, such as the following:

- For evaluation of the appropriate range of input variables and model parameters. The sensitive range is determined as the range within which a change in variables or parameters will result in a significant change in model output. Only the values within this sensitive range are used for model calibration and subsequent performance predictions.

- For a qualitative assessment of model function. The reasonableness of model behavior in terms of its response to increasing or decreasing input variables is evaluated against engineering principles. Model behavior can be used in the comparison of different prediction models.

This research project will include a sensitivity analysis on the final calibrated models for the second purpose. The majority of sensitivity analyses conducted on MEPDG performance models in the literature correspond to the first purpose and have been carried out before calibration. In

this project, the results of the previous studies will be utilized to determine the suitable range of input variables and calibration factors. The following are two types of sensitivity analyses on MEPDG performance models in the literature:

- Sensitivity analysis of model predictions to changes in input variables.
- Sensitivity analysis of model output to variations in calibration factors.

Sensitivity to Input Variables

The most comprehensive sensitivity analysis of MEPDG performance models to changes in input variables was carried out in the NCHRP Project 01-47, and the results provide valuable information regarding range and precision of input values to be considered for calibration of each model.[14] The adopted sensitivity metric was a Normalized Sensitivity Index (NSI), which represents percent change in predicted performance from its design limit value, normalized to a percentage change in an input variable.

This study comprised extensive one-at-a-time (OAT) sensitivity analyses in addition to comprehensive global sensitivity analysis (GSA). In contrast to the OAT analyses, the GSA varied all design inputs simultaneously across the entire problem domain. General agreements between OAT and GSA rankings of sensitivity to various input variables suggest that there were no significant interactions among design inputs. Therefore, the OAT analyses, which are computationally less demanding, could be adequate for sensitivity analysis of MEPDG performance models.

Multivariate linear regression and ANNs were utilized to fit response surface models (RSMs) to the GSA results, allowing for evaluation of sensitivities to design input variables. The ANN resulted in more accurate and robust representations of the compound relations between input design variables and output performance values. Based on frequency distributions and summary statistics generated using the ANN RSM, a "mean plus/minus two standard deviations" ($\mu \pm 2\sigma$) normalized sensitivity metric ($NSI_{\mu\pm2\sigma}$) was derived, which incorporates the mean sensitivity and the variability of the sensitivity across the problem domain. This metric was used to develop the following sensitivity categories:

- Hypersensitive—$NSI_{\mu\pm2\sigma} > 5$.
- Very Sensitive—$1 < NSI_{\mu\pm2\sigma} < 5$.
- Sensitive—$0.1 < NSI_{\mu\pm2\sigma} < 1$.
- Nonsensitive—$NSI_{\mu\pm2\sigma} < 0.1$.

The hypersensitive, very sensitive, and sensitive design inputs for rutting and fatigue cracking models are listed in table 2. As indicated in this table, the performance predictions are most sensitive to the dynamic modulus (E^*) of HMA layers. Poisson's ratio and thickness of the HMA layer and the surface shortwave absorptivity are also important input variables to which these models have shown high sensitivity.

The extreme sensitivity of performance models to the lower and upper shelves of HMA dynamic modulus master curve (alpha and delta parameters) is a questionable behavior. Nevertheless, this calls for careful characterization of dynamic modulus using mix-specific laboratory

measurements. In addition, accurate representation is required for thickness and Poisson's ratio values. The most challenging insight from this sensitivity analysis is that the performance models are very sensitive to several uncertain variables, such as the surface shortwave absorptivity for HMA, thermal conductivity, and heat capacity of stabilized bases, that cannot be readily measured.

Table 2. Sensitive design inputs for rutting and fatigue cracking models.[14] $NSI_{\mu\pm2\sigma}$ values are given in parentheses.

Distress	Input Category	Hypersensitive	Very Sensitive	Sensitive
Fatigue cracking	HMA properties	*E** alpha (–15.9) *E** delta (–13.2) Thickness (–7.5)	Air voids (+3.4) Effective binder volume (–2.9) Surface shortwave absorptivity (+1.3) Poisson's ratio (–1.0)	Unit weight (+1.0) Heat capacity (–0.6) High-temperature PG (–0.5) Thermal conductivity (–0.4)
Fatigue cracking	Base properties	—	Resilient modulus (–2.7) Thickness (–1.0)	Poisson's ratio (+0.9)
Fatigue cracking	Subgrade properties	—	Resilient modulus (–3.4)	Liquid limit (–0.8) Percent passing no. 200 (–0.7) Poisson's ratio (–0.6) Groundwater depth (–0.2) Plasticity index (+0.1)
Fatigue cracking	Other properties	—	Traffic volume (+3.9)	Operating speed (–0.8)
AC rutting	HMA properties	*E** alpha (–24.4) *E** delta (–24.4)	Surface shortwave absorptivity (+4.6) Poisson's ratio (–4.3) Thickness (–4.2)	Unit weight (–0.9) Heat capacity (–0.8) High-temperature PG (–0.7) Low-temperature PG (+0.2) Thermal conductivity (+0.2)
AC rutting	Base properties	—	—	Thickness (+0.2) Poisson's ratio (–0.2) Resilient modulus (+0.1)
AC rutting	Subgrade properties	—	—	Percent passing no. 200 (–0.1) Liquid limit (–0.1)
AC rutting	Other properties	—	Traffic volume (+1.9) Operating speed (–1.1)	
Total rutting	HMA properties	*E** alpha (–9.0) *E** delta (–9.0)	Surface shortwave absorptivity (+1.7) Thickness (–1.6) Poisson's ratio (–1.5)	Unit weight (–0.3) Heat capacity (–0.3) High-temperature PG (–0.2)
Total rutting	Base properties	—	—	Resilient modulus (–0.2)
Total rutting	Subgrade properties	—	—	Resilient modulus (–0.3) Percent passing no. 200 (–0.1)
Total rutting	Other properties	—	—	Traffic volume (+0.7) Operating speed (–0.4)

—No input variable is in this sensitivity category for this performance model; AC = asphalt concrete; E* = dynamic modulus of the HMA layer.

Sensitivity to Calibration Factors

Li et al. (2009) introduced another kind of sensitivity analysis, which is used to determine range and precision of calibration factors.[15] This study on calibration of MEPDG flexible pavement models for Washington DOT examines sensitivity of distress output to the change in each calibration factor. This sensitivity is represented by a metric called elasticity, which was calculated as in equation 19:[15]

$$E_{distress}^{C_i} = \frac{\partial(distress)/distress}{\partial(C_i)/C_i} \quad (19)$$

Where:

$E_{distress}^{C_i}$ = the elasticity of calibration factor C_i for the associated distress condition.
$\partial(distress)$ = change in distress.
distress = initial distress.

$E_{distress}^{C_i}$ is calculated as the ratio of normalized change in predicted distress divided by the normalized change in calibration factor. A positive value means that the predicted distress increases as the calibration factor increases, and a negative value implies that the predicted distress decreases as the calibration factor increases. Based on typical pavement structure, traffic, and climatic data in the Washington DOT PMS database, table 3 indicates elasticity values for calibration factors in MEPDG rutting and fatigue cracking models.[15]

Table 3. Elasticity of MEPDG calibration factors in rutting and fatigue cracking models for Washington State DOT flexible pavements.[15]

Distress	Calibration Factor	Elasticity	Related Input Variables
Fatigue cracking	β_{f1}	–3.3	Effective binder content, air voids, AC thickness
Fatigue cracking	β_{f2}	–40	Tensile strain
Fatigue cracking	β_{f3}	20	Material stiffness
Fatigue cracking	C_1	1	AC thickness
Fatigue cracking	C_2	0	Fatigue damage, AC thickness
Fatigue cracking	C_3	≈0	No related variable
Rutting	β_{r1}	0.6	Layer thickness, layer resilient strain
Rutting	β_{r2}	20.6	Temperature
Rutting	β_{r3}	8.9	Number of load repetitions

AC = asphalt concrete.

The higher absolute values of elasticity for β_{f2}, β_{f3}, β_{r2}, and β_{r3} indicate that model predictions are more sensitive to these calibration factors. As a result, successful calibration requires a higher degree of precision for these factors compared to the others in the optimization procedure. It

should be noted that increasing the precision of calibration factors requires higher computational cost of the optimization procedure. Therefore, the selected precision for each factor should be commensurate with its corresponding elasticity.

It should also be noted that the elasticity metric needs to be identified according to the local pavement structure, climate, and traffic data. In another study on calibration of AASHTOWare® Pavement ME Design software for Iowa, Ceylan et al. used a similar sensitivity metric to calculate the change in performance prediction caused by change in calibration factors.[16]

STATE CALIBRATIONS OF MEPDG PERFORMANCE MODELS

Global calibration and validation of MEPDG performance models were completed using a subset of LTPP data based on national averages.[2] Ever since, numerous State DOTs have been in the process of calibrating these models to their own regional materials–traffic–climate conditions. Two important studies of NCHRP 9-30 and NCHRP 1-40B have provided guidelines in this regard.[17,4] The NCHRP Synthesis 457 provides a comprehensive report on the pavement design practices and MEPDG implementation status in various States across the country.[10] This report also includes agency implementation challenges and details case examples of the MEPDG implementation process in three States.

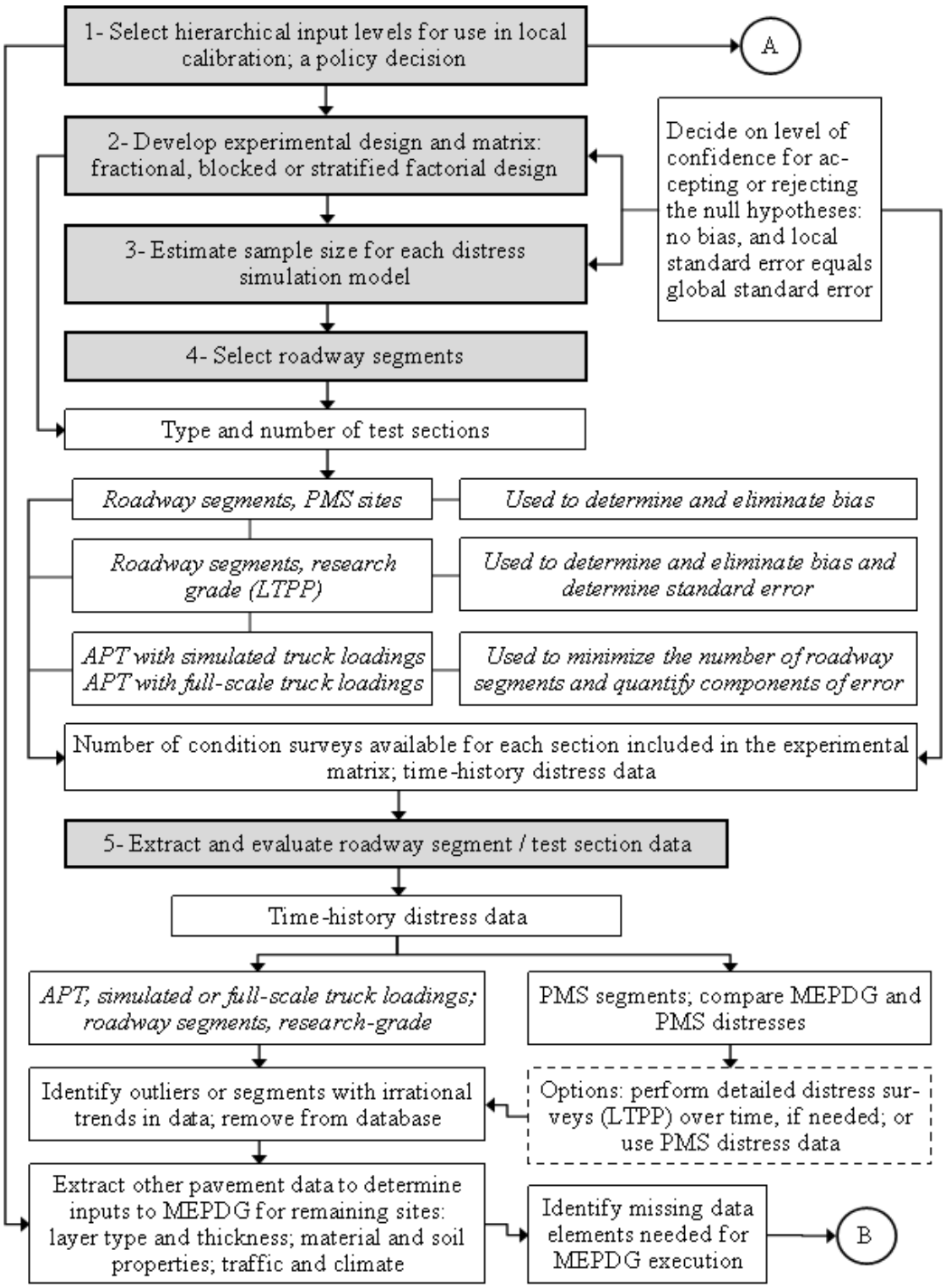
NCHRP 1-40B provides the following 11-step procedure for verification, calibration, and validation of the MEPDG models for local conditions, which has been adopted by AASHTO:[3]

1. Select hierarchical input level.
2. Develop experimental plan and sampling template.
3. Estimate sample size.
4. Select roadway segments.
5. Evaluate project and distress data.
6. Conduct field testing and forensic investigation.
7. Assess local bias.
8. Eliminate local bias.
9. Assess STE of the estimate.
10. Reduce STE of the estimate.
11. Interpret the results.

Statistical significance testing is recommended at various steps to determine if the models need further calibration. At the seventh step, the significance of the bias (the average difference between predicted and measured performance) is tested. If there is a significant bias in prediction of pavement performance measures, the first round of calibration is conducted at the eighth step to eliminate bias. For example, during this step for the rutting models, the SSE is minimized by adjusting the β_{r1}, β_{GB}, and β_{SG} calibration factors.

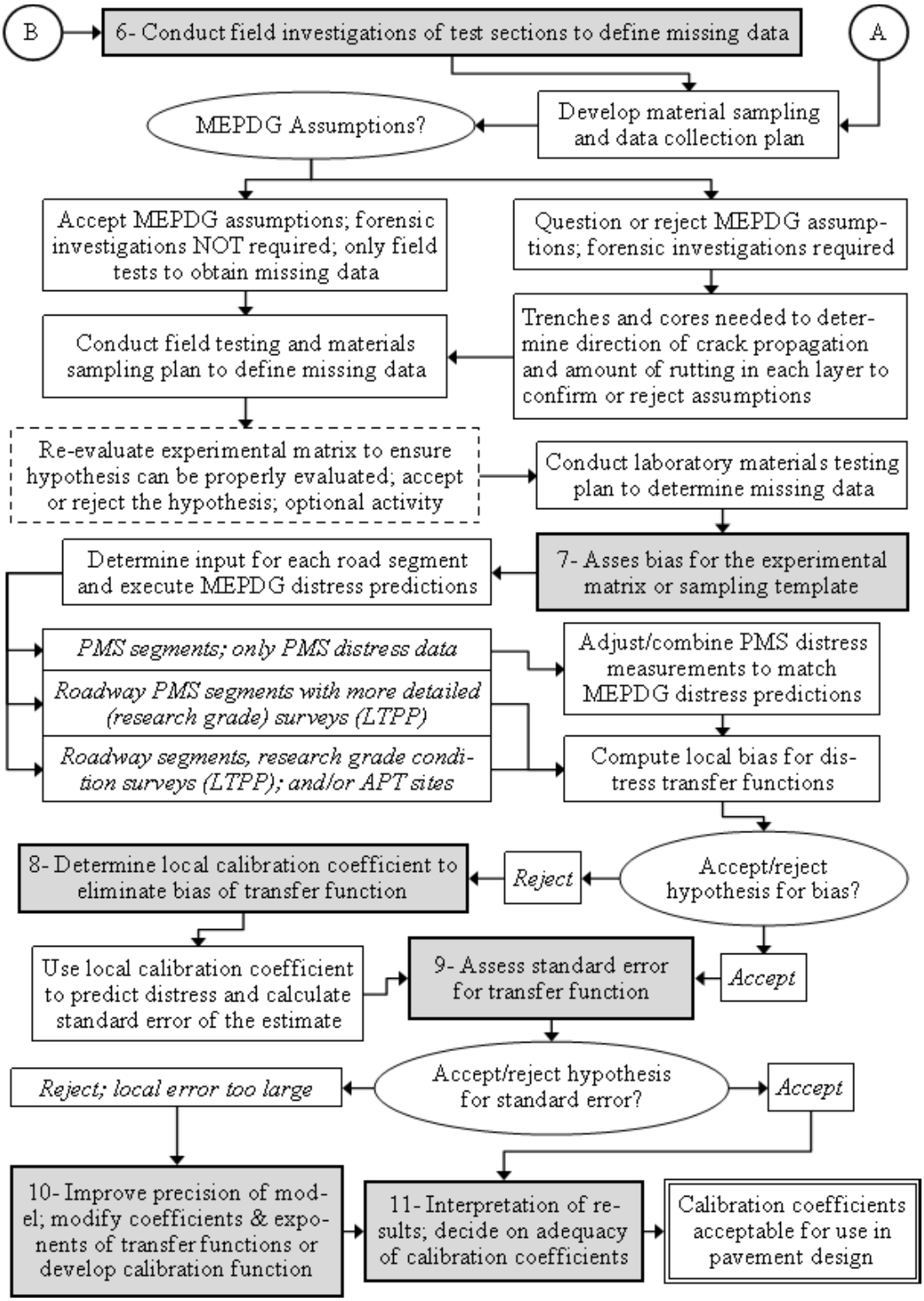
At the ninth step, the STE (standard deviation of error among the calibration dataset) is evaluated by comparing it to the STE from the national global calibration. If there is a significant STE, the second round of calibration at the 10th step tries to reduce the STE by adjusting the β_{r2} and β_{r3} calibration factors. A final validation step checks for the reasonableness of performance

predictions. The flowcharts depicted in figure 1 and figure 2 demonstrate this calibration process.[3]



Reprinted from *Guide for the Local Calibration of the Mechanistic–Empirical Pavement Design Guide*, 2010, by the American Association of State Highway and Transportation Officials, Washington, DC. Used by permission.

Figure 1. Flowchart. The AASHTO recommended procedure for local calibration of MEPDG performance models, steps 1 through 5.[3]



Reprinted from *Guide for the Local Calibration of the Mechanistic–Empirical Pavement Design Guide*, 2010, by the American Association of State Highway and Transportation Officials, Washington, DC. Used by permission.

Figure 2. Flowchart. The AASHTO recommended procedure for local calibration of MEPDG performance models, steps 6 through 11.[3]

The most recent literature review on calibration of MEPDG in different States was conducted as part of a study for Georgia DOT.[13] Results of this literature review and other similar studies have been compiled to provide a summary of the State calibration efforts in table 4 and a list of reported calibration factors for flexible pavement fatigue cracking and rutting models in table 5.

Most of the past calibration studies suggest that the MEPDG rutting prediction models overpredict rutting in unbound pavement layers. However, the rutting predicted in asphalt concrete (AC) layers seems to be easily calibrated. The fatigue (alligator) cracking model seems to underpredict actual pavement distress and has high variation in the predicted values. There seems to be little to no problem in calibration of transverse cracking and smoothness prediction models. There seems to be no specific trend for the flexible pavement longitudinal cracking model, and none of the studies reported a successful calibration of it. The MEPDG longitudinal cracking model is not considered in the scope of this study because several past studies have expressed concern on the lack of fit of this model.[18] Difficulty in differentiating longitudinal cracks in the wheelpath from alligator cracking patterns might have contributed to errors in measured longitudinal cracking values.

Differences among various distress identification protocols (e.g., LTPP versus State PMS) and the subjective nature of identifying distress type and severity have been noted as sources of measurement error that cause significant challenges in calibration of mechanistic models to field-measured performance data.[19,20]

Table 4. Major State efforts for calibration of MEPDG performance models.

Study	Scope	Major Findings
NCHRP 1-37A[2]	National calibration of MEPDG models	National calibration of MEPDG models
NCHRP 9-30[17]	Calibration of flexible pavement performance models for structural and mix design	Procedures for adjusting global coefficients according to lab data
NCHRP 1-40A[21]	Independent review of the MEPDG	Rutting is overpredicted in unbound pavement layers.
NCHRP 1-40B[4]	11-step recommended calibration procedure	11-step recommended calibration procedure
NCHRP 1-40D[6]	National recalibration of MEPDG models	National recalibration of MEPDG models
Von Quintus and Moulthrop 2007[5]	Calibration of MEPDG flexible pavement performance models for Montana	Lack of fit for the longitudinal flexible pavement cracking model
Kang et al. 2007[7]	Midwest regional pavement performance database for MEPDG calibration	Database creation is very labor intensive and unreliable.
Von Quintus 2008[18]	Overview of selected studies on local calibration of MEPDG	Summary of flexible pavement local calibration factors from national and local calibrations
Muthadi and Kim 2008[22]	Calibration of MEPDG flexible pavement performance models for North Carolina	Calibration factors for rutting and fatigue cracking models. MEPDG models underpredict fatigue cracking.
Banerjee et al. 2009[23]	Calibration of MEPDG flexible pavement performance models for Texas	Regional and local calibration factors for rutting
Li et al. 2009[15]	Calibration of MEPDG flexible pavement performance models for Washington	The important calibration factors were identified according to the sensitivity of the models to them.
Titus-Glover and Mallela 2009[24]	Calibration of MEPDG performance models for Ohio	Calibration of MEPDG performance models for Ohio
Souliman et al. 2010[25]	Calibration of MEPDG flexible pavement performance models for Arizona	Calibration of MEPDG flexible pavement performance models for Arizona
Hoegh et al. 2010[26]	Calibration of MEPDG rutting models for Minnesota	Modified rutting model based on MnROAD data
Hall et al. 2011[27]	Calibration of MEPDG flexible pavement performance models for Arkansas	Variation in predicted fatigue cracking remains high and is not improved by calibration.
Williams and Shaidur 2013[28]	Calibration of MEPDG performance models for Oregon	Calibration of MEPDG performance models for Oregon
Ceylan et al. 2013[16]	Calibration of MEPDG performance models for Iowa	Nationally calibrated rutting model provides acceptable predictions for Iowa.
Mallela et al. 2013[29]	Calibration of MEPDG performance models for Colorado	Calibration of MEPDG performance models for Colorado

MnROAD = Minnesota Department of Transportation pavement test track.

Table 5. Local calibration factors for MEPDG fatigue cracking and rutting prediction models.

Performance Model	HMA Fatigue	HMA Fatigue	HMA Fatigue	Bottom–Up Cracking	Bottom–Up Cracking	HMA Rutting	HMA Rutting	HMA Rutting	Base Rutting	Subgrade Rutting
Coefficient	β_{f1}	β_{f2}	β_{f3}	C_1	C_2	β_{r1}	β_{r2}	β_{r3}	β_{GB}	β_{SG}
National	1	1	1	1	1	1	1	1	1	1
AR	1	1	1	0.688	0.294	1.2	1	0.8	1	0.5
AZ*	0.729	0.8	0.8	0.732	0.732	3.63	1.1	0.7	0.111	1.38
CO^	130.367	1	1.2178	0.07	2.35	1.34	1	1	0.4	0.84
IA	1	1	1	1	1	1	1.15	1	0	0
MO	1	1	1	1	1	1.07	1	1	0.01	0.4375
MT	13.21	1	1.25	1	1	7	1.13	0.7	1	0.3
NC*	1.41	–2.82	–6.67	0.4372	0.15049	1.0175	1	1	1.5803	1.10491
OH	1	1	1	1	1	0.51	1	1	0.32	0.33
OR	1	1	1	0.56	0.225	1.48	1	0.9	0	0
UT	1	1	1	1	1	0.56	1	1	0.604	0.4
WA*	0.96	0.97	1.03	1.071	1	1.05	1.109	1.1		0
WI*	1	1.2	1.5	1	1	1.0157	1	1	0.01	0.5731
WY^	1	1	1	0.4951	1.469	1.0896	1	1	0.9475	0.6897
Midwest	1	1.2	1.5	1	1	1	1	1	1	1
Average”	2.1190	0.6682	0.4009	0.8488	0.7638	1.7757	1.0445	0.9273	0.4039	0.4569
Range”	0.729 to 13.21	–2.82 to 1.2	–6.67 to 1.5	0.4372 to 1.071	0.15049 to 1.469	0.51 to 7	1 to 1.15	0.7 to 1.1	0.0 to 1.5803	0.0 to 1.38
COV (%)”	174	174	588	27	47	108	6	15	139	97

*Calibration factors reported by Von Quintus et al. (2013) were different from the ones found in this literature search (references in table 4).[5]
^These values are not final.
”These statistics exclude CO and WY values.
COV = coefficient of variation.

Table 5 shows significant variance among the States in terms of the β_{f1}, β_{f2}, β_{f3}, β_{r1}, β_{GB}, and β_{SG} calibration factors as indicated by their corresponding high coefficients of variation. Therefore, it is important that the optimum coefficients be determined for these calibration factors to ensure compliance to local pavement performance. In addition, C_1 and C_2 also show some variation among different calibration efforts. The number of calibration factors determined to be equal to 1 (1.0), which are the global calibration values, are more for the fatigue cracking model compared to the permanent deformation model. This could be interpreted as a superior global model having been developed for fatigue cracking compared to rutting.

OTHER MEPDG CALIBRATION EFFORTS

The measurement error in the performance data records is known to be greatly undermining precision of calibrated MEPDG models.[18] Therefore, Hall et al. suggested a new output format for the performance models to predict ranges of distress instead of an exact value.[27]

To account for the effect of maintenance or rehabilitation activities, Li et al. suggested developing piecewise performance models for Washington State.[30] Pavement serviceable life was divided into three time periods of early age, rehabilitation, and overdistressed situations. They used regression to develop models for each time period.

In addition to the national research studies conducted to determine the global calibration factors for permanent deformation model, some States have conducted their own laboratory tests in this regard.[31] For example, Jadoun and Kim used results of the triaxial repeated load permanent deformation test to determine the global *k* factors for 12 different HMA mixtures.[32]

The majority of these studies used exhaustive search methods such as the generalized reduced gradient (GRG) method to minimize SSE between measured and predicted performance. These methods are local optimization techniques that are dependent on seed values and typically get stuck at a local minimum of error. Jadoun and Kim compared a genetic algorithm (GA) to the GRG method for calibration of rutting and fatigue cracking models for North Carolina.[32] They demonstrated that the GA method provides a more global minimum of SSE compared to the GRG method in predicting rutting. However, this superior optimization does not result in a reasonable match between predicted and measured fatigue cracking.

It should be noted that the applied GA code is highly sensitive to the control parameters used to manipulate the evolutionary process of optimization. Therefore, there might be variants of this GA code that perform better, and the best set of control parameters needs to be determined for each optimization problem. Several evolution strategies (ESs) have been developed in the evolutionary computation literature that evolve and adapt control parameters along with optimization solutions and with respect to the objective function space. Application of these ESs would result in a more robust optimization.

MULTI-OBJECTIVE CALIBRATION STUDIES

All of the MEPDG calibration studies focus on minimization of a single-objective function (SSE) for all distress severity levels and all pavement ages in the considered network. Incorporating multiple sources of information might reveal unknown aspects of this calibration problem and result in more reasonable calibration coefficients. Multi-objective evolutionary

algorithms (MOEAs) are derivative-free, global optimization heuristics that provide a set of tradeoff solutions independent of seed values.[33]

MOEAs have been used in pavement management studies to optimize the allocation of resources to various treatment alternatives considering multiple criteria.[34–36] They have also been vastly implemented in water resources research to design long-term groundwater monitoring schemes and to calibrate hydrologic models.[37,38] The multi-criteria framework provided by this kind of calibration has enabled recognition and handling of errors and uncertainties and detection of prominent behavioral solutions with acceptable tradeoffs in hydrologic modeling efforts within the past decade.[39]

INSIGHTS AND OBSERVATIONS FROM THE LITERATURE REVIEW

The following are the key observations drawn from this literature review:

- Different data sources are being incorporated in the MEPDG calibration process, but attention should be paid to the differences in performance measurement protocols.

- The sensitivity of performance models to the HMA dynamic modulus, thickness, and Poisson's ratio calls for careful characterization of these values. MEPDG performance models are very sensitive to several uncertain variables, such as the surface shortwave absorptivity for HMA, thermal conductivity, and heat capacity of stabilized bases, that cannot be readily measured.

- Model predictions are more sensitive to some local calibration factors compared to others. Therefore, the selected precision for each factor should be commensurate with its corresponding elasticity.

- MEPDG rutting and fatigue cracking models have been reported to overpredict and underpredict actual pavement distresses, respectively. The local values calculated for the calibration factors, β_{f1}, β_{f2}, and β_{f3} for fatigue cracking model and β_{r1}, β_{GB}, and β_{SG} for rutting model, seem to be significantly different among various reviewed calibration efforts.

- Differences among various distress identification protocols (e.g., LTPP versus State PMS) and the subjective nature of identifying distress type and severity have been noted as sources of measurement error that cause significant challenges in calibration of mechanistic models to field-measured performance data.

- Due to the challenge posed by distress measurement errors, some researchers have proposed conducting model calibration using ranges of distress instead of exact values. Furthermore, to account for different pavement behavior during its various life stages, it has been suggested that model calibration be carried out separately across different periods of pavement service life.

- Global heuristic optimization methods, such as evolutionary algorithms (EAs), could possibly identify a more optimum set of calibration coefficients compared to the local exhaustive search methods.

- The multi-criteria calibration framework provided by MOEAs has enabled recognition and handling of errors and uncertainties and detection of prominent behavioral solutions with acceptable tradeoffs in hydrologic modeling efforts.

IMPACT ON RESEARCH APPROACH

Based on the findings of this literature review, the following considerations corresponding to the above observations were recommended for the research approach:

- This study should be performed using LTPP data for flexible pavements within a specific region comprising one or more States with similar climatic and subgrade conditions. In addition to LTPP data for the selected region, utilization of another source of data, such as State PMS or APT data, should be considered.

- Careful characterization of the HMA dynamic modulus, thickness, and Poisson's ratio is necessary for the success of this project. In this regard, the results of ANNs for Asphalt Concrete Dynamic Modulus Prediction (ANNACAP) software for LTPP test sections should be implemented. This software could be used for non-LTPP data sources when applicable.

- The selected precision for each factor in the optimization procedure should be commensurate with the corresponding sensitivity of the performance model to that calibration factor. This is an important consideration because the precision of these unknown variables directly relates to the computational cost of the optimization problem.

- This research project will focus on local calibration of prediction models for rutting on new and overlaid flexible pavements.

- The multi-objective calibration approach could incorporate the different data characteristics (performance measurement protocols) of different data sources in an objective manner.

- Using a multi-objective calibration approach and by simultaneously minimizing the error in predicting pavement performance from disparate data sources, the calibration coefficients that provide a tradeoff among pavement behavior during different experiments will be determined.

- In this study, MOEAs will be implemented. These global optimization heuristics have good global search ability, are less dependent on seed values (techniques such as restarting have been shown to significantly decrease dependence on seed values), and do not require the mathematical formula (to find the derivative) of the objective functions.[37]

- Using MOEA, multiple sources of information can be incorporated in an objective manner, resulting in a final set of tradeoff solutions. This way, none of the possible sets of calibration factors will be eliminated prematurely, and all of the nondominated solutions will be included in the final tradeoff front. Exploring the final front might reveal unknown aspects of this calibration problem and result in more reasonable calibration coefficients that could not be identified using single-objective approaches.

Several scenarios can be devised for multi-objective formulation of calibration, all of which could overcome cognitive challenges and add to the knowledge of this problem. More than one set of multiple objectives will be considered to explore new aspects of the calibration problem. The idea is to optimize multiple objectives simultaneously. The following are the proposed sets of objectives up to this stage of the study:

- Statistical outcomes (increasing accuracy and precision simultaneously).
    i. Minimize average error (bias).
    ii. Minimize error standard deviation.
- Data sources (an objective approach to incorporate different sources of data).
    i. Minimize error on LTPP data.
    ii. Minimize error on APT data.

In the primary multi-objective scenario, mean and standard deviation of prediction error are simultaneously minimized to reduce the bias and STE at the same time. In this manner, the information from a single calibration run is fully implemented, and an additional round of computationally intensive calibration is avoided.

In the second multi-objective scenario for calibration of MEPDG performance models, the error in predicting the performance of pavements within different performance data sources will be used as separate objective functions to be minimized simultaneously. In addition to LTPP test sections, data from State PMS or APT facilities in the same region can be considered for this scenario. This scenario comprises an objective approach to incorporate different sources of data. Finally, a combination of two or more of the above scenarios could also be considered for the multi-objective calibration approach.

CHAPTER 3. PREPARATION OF MEPDG INPUTS FROM LTPP DATA

INTRODUCTION

To demonstrate the multi-objective approach, a subset of the LTPP data from a specific region was selected for calibration of MEPDG permanent deformation (rutting) prediction models for new and rehabilitated flexible pavements.

This chapter discusses the extraction of relevant LTPP data, calculation of the necessary parameters, and selection of other necessary values for generating the AASHTOWare® Pavement ME Design software input files. Following the discussion of the region selection process according to data availability, the LTPP data extraction and calculation of the required general, performance, traffic, climate, and structure/materials values are discussed.

LTPP DATA AVAILABILITY

The following steps demonstrate the process used to narrow the search for a suitable set of LTPP sections to collect calibration data. The LTPP InfoPave™ website has been very useful in filtering the relevant test sections and identifying the availability of the required data.[40]

1. The 1,746 LTPP test sections with AC surface were selected.

2. The relevant LTPP flexible pavement experiments were identified to be the following 1,019 sections:

    i. General Pavement Studies (GPS)-1.
    ii. GPS-2.
    iii. GPS-6.
    iv. Specific Pavement Studies (SPS)-1.
    v. SPS-5.
    vi. SPS-8.
    vii. SPS-9N, SPS-9O.

3. The LTPP sections with at least one rutting measurement were selected: 1,014 sections.

4. Considering coarse- and fine-grained subgrade soils and the four LTPP climatic regions, table 6 provides the number of available sections in each category.

Table 6. Available number of test sections for each LTPP climatic region and subgrade type.

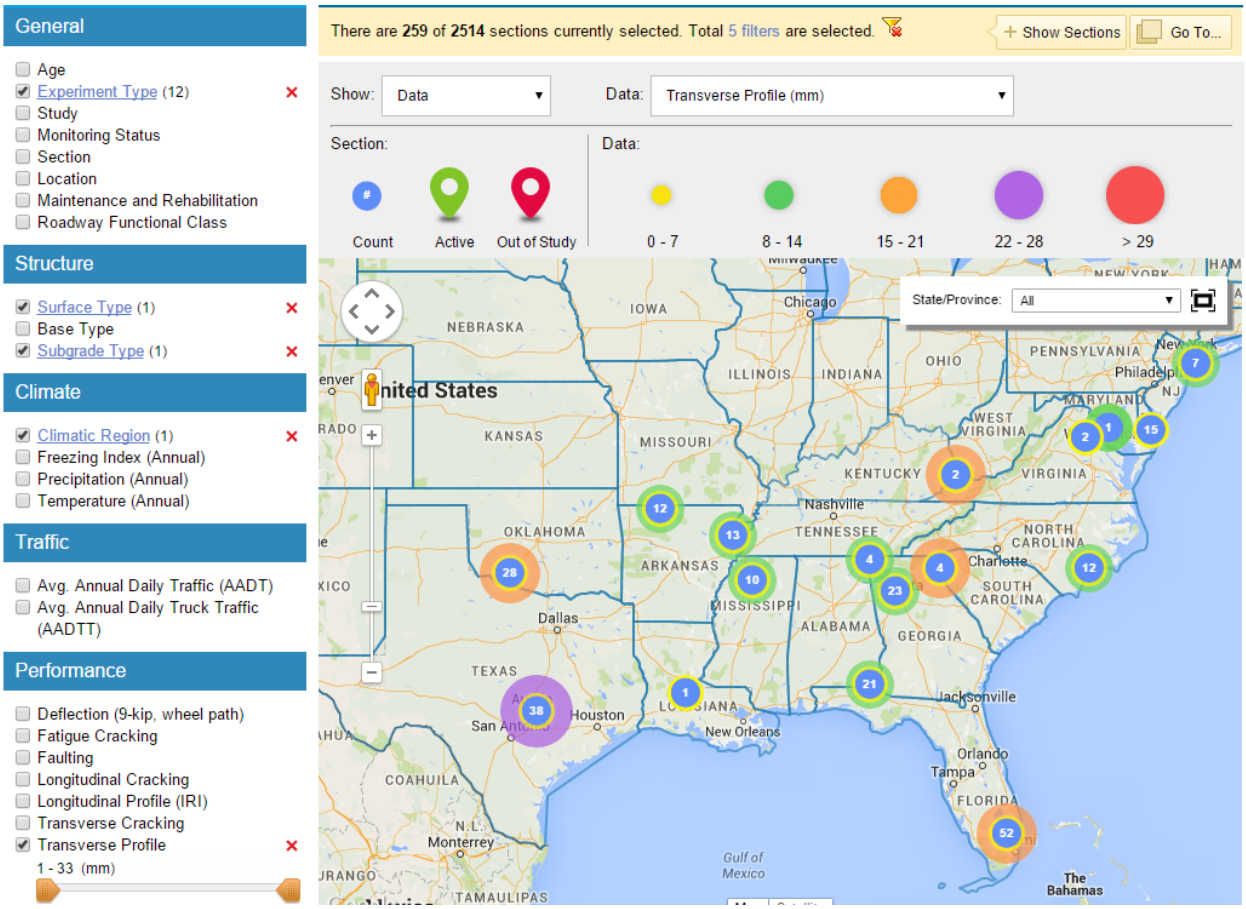
Coarse-Grained Subgrade	Sections	Fine-Grained Subgrade	Sections
Dry, freeze	100	Dry, freeze	39
Dry, no freeze	122	Dry, no freeze	28
Wet, freeze	143	Wet, freeze	132
Wet, no freeze	259	Wet, no freeze	191

There are more options among sections on coarse-grained subgrade soils (624 test sections) than fine-grained subgrade soils (390) and within the wet climatic regions (725) compared to dry regions (289). Based on the number of test sections in table 6, the search was narrowed to sections in the wet, no freeze climatic region.

The InfoPave™ map in figure 3 indicates the distribution of the 259 sections within the wet, no freeze climate and on coarse subgrades.[40] Most of these test sections are in Florida, Texas, Oklahoma, Georgia, and Alabama. The highest recorded amount of measured rut depth is between 22 and 28 mm (as indicated on the map).

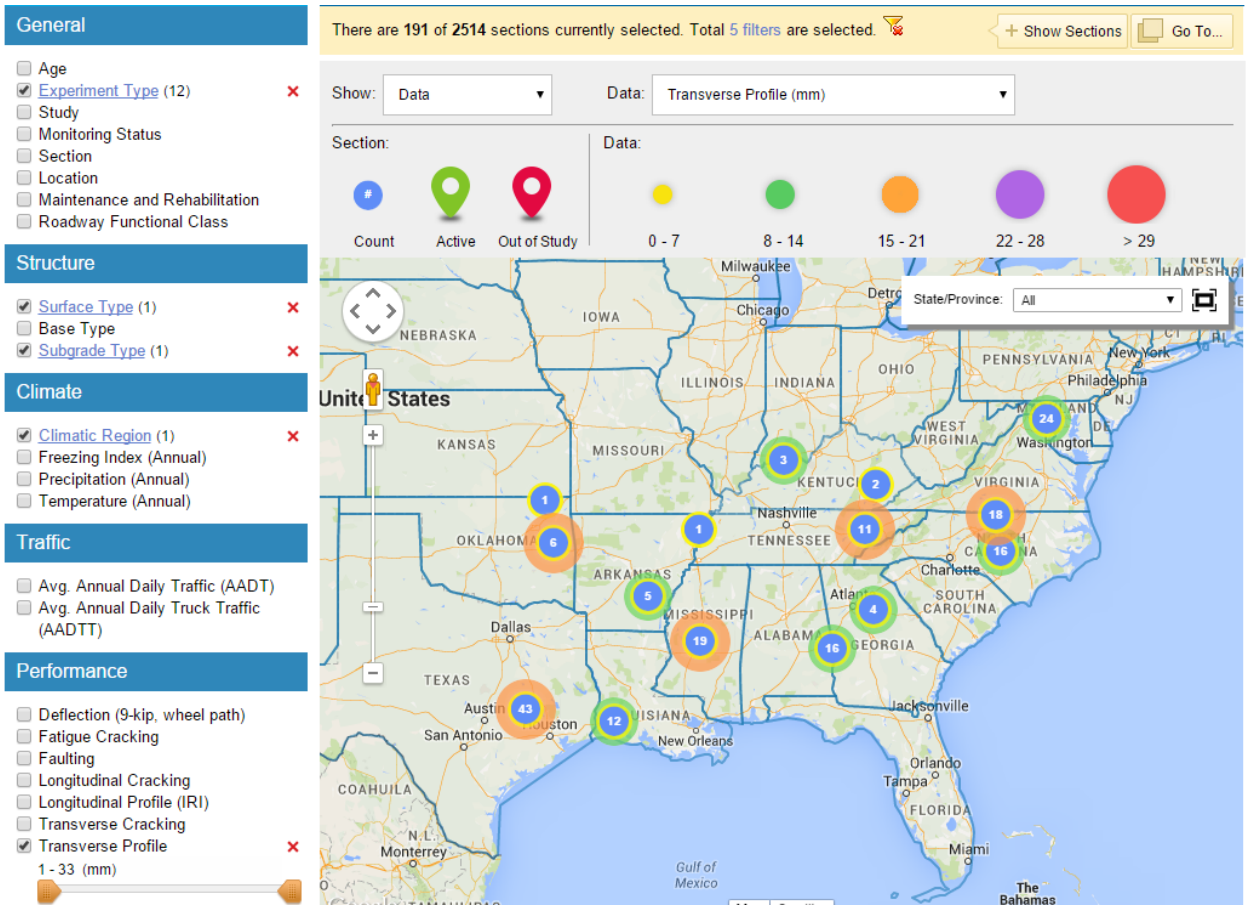
The InfoPave™ map in figure 4 shows the distribution of the 191 sections within the wet, no freeze climate and on fine subgrades.[40] The majority of these test sections are located in Texas, Maryland, Mississippi, and Virginia, in that order. The highest recorded amount of measured rut depth is between 15 and 21 mm (as indicated on the map).

From the InfoPave™ maps, it seems Florida has the highest number (52) of eligible test sections to be used for this study.[40] In addition to the results of this preliminary evaluation being in favor of selecting the Florida region for this project, FDOT has offered the flexible pavement data collected at their APT facility to be used in this study. The eligible LTPP test sections in Florida and the corresponding general information are listed in table 7. The LTPP Standard Data Release (SDR) number 29.0 was used for identifying the available data for this study.[40] The *LTPP Information Management System (IMS) User Guide* was used to identify the relevant sources (tables and fields) of data.[41]



Source: FHWA.

Figure 3. Screenshot. Location of sections within the wet, no freeze climate and on coarse subgrades from InfoPave™.



Source: FHWA.

Figure 4. Screenshot. Location of sections within the wet, no freeze climate and on fine subgrades from InfoPave™.

Table 7. General information on the 52 flexible test sections on coarse subgrade soils in Florida.

#	STATE_CODE	SHRP_ID	Highway	Direction	EXPERIMENT_TYPE	CONST_DATE	OL_DATE	END_DATE	NEW_OR_OVERLAY	BASE TYPE
1	12	0101	US-27	South	SPS-1	3/8/1995	N/A	7/13/2012	New	GB
2	12	0102	US-27	South	SPS-1	3/7/1995	N/A	7/13/2012	New	GB
3	12	0103	US-27	South	SPS-1	3/7/1995	N/A	7/13/2012	New	ATB
4	12	0104	US-27	South	SPS-1	3/7/1995	N/A	7/13/2012	New	ATB
5	12	0105	US-27	South	SPS-1	3/7/1995	N/A	7/13/2012	New	GB
6	12	0106	US-27	South	SPS-1	11/20/1995	N/A	7/13/2012	New	GB
7	12	0107	US-27	South	SPS-1	5/1/1995	N/A	7/13/2012	New	GB
8	12	0108	US-27	South	SPS-1	5/2/1995	N/A	7/13/2012	New	GB
9	12	0109	US-27	South	SPS-1	3/7/1995	N/A	7/13/2012	New	GB
10	12	0110	US-27	South	SPS-1	11/20/1995	N/A	7/13/2012	New	ATB
11	12	0111	US-27	South	SPS-1	11/20/1995	N/A	7/13/2012	New	ATB

#	STATE_CODE	SHRP_ID	Highway	Direction	EXPERIMENT_TYPE	CONST_DATE	OL_DATE	END_DATE	NEW_OR_OVERLAY	BASE TYPE
12	12	0112	US-27	South	SPS-1	11/28/1995	N/A	7/13/2012	New	ATB
13	12	0161	US-27	South	SPS-1	3/3/1995	N/A	7/13/2012	New	GB
14	12	0502	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
15	12	0503	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
16	12	0504	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
17	12	0505	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
18	12	0506	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
19	12	0507	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
20	12	0508	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
21	12	0509	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
22	12	0561	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
23	12	0562	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
24	12	0563	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
25	12	0564	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
26	12	0565	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
27	12	0566	US-1	South	SPS-5 to GPS-6S	4/1/1971	4/18/1995	7/24/2014	Overlay	GB
28	12	0901	I-10	East	SPS-9O	6/1/1963	7/23/1996	5/15/2008	Overlay	GB
29	12	0902	I-10	East	SPS-9O	6/1/1963	7/24/1996	5/15/2008	Overlay	GB
30	12	0903	I-10	East	SPS-9O	6/1/1963	7/23/1996	5/15/2008	Overlay	GB
31	12	0959	I-10	East	SPS-9O	6/1/1963	7/24/1996	5/15/2008	Overlay	GB
32	12	1030	US-1	South	GPS-1 to 6S	4/1/1971	N/A	7/24/2014	N/A	GB
33	12	1060	SH-878	West	GPS-1	10/1/1979	N/A	3/7/2003	N/A	GB
34	12	1370	SH-407	North	GPS-1 to 6S	6/1/1973	8/15/2000	4/24/2015	Overlay	GB
35	12	3995	I-95	North	GPS-1	12/1/1975	N/A	4/17/1997	N/A	GB
36	12	3996	US-19	North	GPS-1	4/1/1974	N/A	6/1/1998	N/A	GB
37	12	3997	US-17	South	GPS-1 to 6S	6/1/1974	2/7/1995	7/13/1999	Overlay	GB
38	12	4096	SH-20	West	GPS-2 to GPS-6C	5/1/1974	2/21/2003	N/A	Overlay	ATB
39	12	4097	I-10	East	GPS-2	1/1/1986	N/A	1/15/2005	N/A	CTB
40	12	4099	SH-884	West	GPS-1	6/1/1976	N/A	6/29/1992	N/A	GB
41	12	4100	SH-85	North	GPS-2 to GPS-6S	8/1/1976	8/29/2002	8/12/2012	Overlay	ATB
42	12	4101	SH-528	East	GPS-1 to 6B	5/1/1967	7/31/1991	8/28/1996	Overlay	GB
43	12	4103	SH-836	West	GPS-1	6/1/1982	N/A	6/22/2000	N/A	GB
44	12	4105	SH-9A	North	GPS-1	12/1/1984	N/A	6/3/1993	N/A	GB
45	12	4106	I-95	North	GPS-1 to 6S	8/1/1987	11/15/2003	11/24/2009	Overlay	GB
46	12	4107	SH-70	West	GPS-1	10/1/1983	N/A	5/4/1998	N/A	GB
47	12	4108	SH-30	West	GPS-2	6/1/1986	N/A	10/25/1996	N/A	ATB
48	12	4135	US-27	North	GPS-1 to 6B	2/1/1971	2/15/1992	6/12/2006	Overlay	GB
49	12	4136	US-27	North	GPS-1 to 6B	2/1/1971	2/15/1992	6/12/2006	Overlay	GB
50	12	4137	US-27	North	GPS-1 to 6B	12/1/1970	2/15/1992	6/12/2006	Overlay	GB
51	12	4154	SH-442	East	GPS-1	6/1/1970	N/A	3/31/1998	N/A	GB
52	12	9054	SH-200	West	GPS-1	10/1/1974	N/A	9/26/1997	N/A	GB

SHRP_ID = Strategic Highway Research Program identification number; CONST_DATE = construction date; OL_DATE = overlay date; N/A = no adequate data; GB = granular base; ATB = asphalt-treated base; CTB = cement-treated base.

The field named NEW_OR_OVERLAY in table 7 shows which test sections (13 sections total) were found most suitable for calibration of rutting prediction models for new pavements and which sections (27 sections total) were deemed more appropriate for calibration of rutting models for overlaid pavements. These selections were made based on the availability of monitoring data after original pavement construction or after overlay construction. In 12 test sections marked as N/A, there were not adequate data on maintenance and rehabilitation history before the sections were assigned to the LTPP program and monitored. Therefore, using the data from these 12 test sections is not recommended for the calibration process.

These test sections have adequate climatic data to be used. The AASHTOWare® Pavement ME Design software uses historical climate data (HCD) files, and each one of the LTPP sections do have a corresponding climate station in that database, which can be downloaded from the http://me-design.com/MEDesign/ClimaticData.html Web location. Recently, the LTPP program has also made Modern-Era Retrospective Analysis for Research and Applications (MERRA) data available since the SDR 29.0 and through the InfoPave™ website.(40) In this research study, the original HCD files downloaded from the AASHTOWare® website were used as inputs. The next steps are exploring the amount of available traffic and structure data.

Table 8 shows the LTPP source tables and amount of available traffic data for the 52 test sections that met the selection criteria in Florida. The LTPP traffic (TRF) module contains tables that start with the TRF_MEPDG_ prefix. These tables contain traffic data estimated for input into the MEPDG software based on test sites that have more than 210 d of monitored data per year. However, for other site-years where the total number of accepted monitoring data was less than 210 d in that year, traffic data can still be found in the tables starting with the TRF_MONITOR_ prefix. As table 8 demonstrates, the majority of the test sections have some (for some years) traffic data available to be used in this study. The details of input data extraction and calculations are explained later in this chapter.

Table 9 shows the amount of available LTPP data for the most important structural factors. It was assumed that the material properties for one SPS section could be applied to other sections with the same material source within the same SPS site. The LTPP data field PROJECT_LAYER_NO in the tables starting with the TST prefix (the testing module, which contains the results of materials testing conducted under the LTPP program) was used to decide whether test results from one test section could be applied to other test sections within an SPS project site. It is noted that, due to construction variability, this might not be an accurate estimation, but it will be more representative of the field materials when compared to the MEPDG default values. The details of input data extraction and calculations are explained later in this chapter.

Table 8. Source and availability of traffic data for the selected 52 flexible sections in Florida.

Data Item	LTPP Source Tables	Number of Sections With Available Data
AADTT	TRF_MEPDG_AADTT_LTPP_LN	49
AADTT	TRF_MONITOR_LTPP_LN	52
AADTT	TRF_HIST_EST_ESAL	15
AADTT	TRF_MON_EST_ESAL	52
Axle load distributions	TRF_MEPDG_AX_DIST_ANL	31
Axle load distributions	TRF_MONITOR_AXLE_DISTRIB	52
Number of axles per truck	Estimated based on data from TRF_MONITOR_LTPP_LN	52
Vehicle class distributions	TRF_MEPDG_VEH_CLASS_DIST	49
Vehicle class distributions	TRF_MONITOR_LTPP_LN	52
Monthly adjustment factors	TRF_MEPDG_MONTH_ADJ_FACTR	49
Hourly adjustment factors	TRF_MEPDG_HOURLY_DISTRIB	28
Traffic growth factor	Estimated based on data from TRF_MONITOR_LTPP_LN	52
Traffic growth function	Not available, assumed compound growth	N/A
LTPP lane, direction, and lane width	INV_GENERAL	39
LTPP lane, direction, and lane width	SPS_ID	13
Lane and directional distribution	Not directly available, both assumed to be 1.0 for calibration purposes	N/A
Operational speed	SECTION_GENERAL, not available for Florida sections	N/A
Tire pressure	Not available, MEPDG defaults	N/A
Axle configuration, wheelbase, and wheel location	Not available, MEPDG defaults	N/A
Truck wander	Not available, MEPDG defaults	N/A

AADTT = average annual daily truck traffic; N/A = no adequate data.

Table 9. Source and availability of structure data for the selected 52 flexible sections in Florida.

Data Item	LTPP Source Tables	Sections
Thickness of all layers	TST_L05B	52
Poisson's ratio for all layers	Not available	MEPDG defaults
AC dynamic modulus	TST_AC07	34
AC dynamic modulus	TST_ESTAR estimated values	40
AC air voids	For SPS-9: TST_SP02	4
AC air voids	Other than SPS-9: TST_AC02 + TST_AC03	37
AC effective binder volume	For SPS-9: TST_SP02	4
AC effective binder volume	Other than SPS-9: TST_AC03 + TST_AC04 + TST_AG01 + TST_AG02	23 to 37
AC shortwave absorptivity	Not available	MEPDG defaults
AC PG grading	LTPP Bind Online	43
AC heat capacity	Not available	MEPDG defaults
AC thermal conductivity	Not available	MEPDG defaults
Base resilient modulus	TST_UG07_SS07_WKSHT_SUM	38
Subgrade resilient modulus	TST_UG07_SS07_WKSHT_SUM	38
Subgrade percent passing no. 200	TST_SS01_UG01_UG02	39
Subgrade Atterberg limits	TST_UG04_SS03	39

In this section, it was determined that there are adequate structure–climate–traffic data in the LTPP database for some of the selected test sections to be used in calibration of rutting prediction models for the new and rehabilitated flexible pavements. However, the final data element is the performance monitoring data. Table shows the availability of the rutting measurements for the selected LTPP test sections in Florida. This table shows the rutting data only for the 40 test sections that were assigned to calibration of new (13 sections) or rehabilitated (27 sections) performance models in table 7.

Table 10. Availability of rutting data for the selected LTPP flexible pavements in Florida.

#	STATE_ CODE	SHRP_ ID	EXPERIMENT_ TYPE	NEW_OR_ OVERLAY	BASE_ TYPE	FIRST_RUT_ DATE	LAST_RUT_ DATE	NUMBER_RUT_ MEASUREMENT
1	12	0101	SPS-1	New	GB	2/9/2000	3/29/2011	12
2	12	0102	SPS-1	New	GB	2/9/2000	3/29/2011	12
3	12	0103	SPS-1	New	ATB	2/9/2000	3/29/2011	12
4	12	0104	SPS-1	New	ATB	2/9/2000	3/30/2011	12
5	12	0105	SPS-1	New	GB	2/9/2000	3/29/2011	12
6	12	0106	SPS-1	New	GB	2/9/2000	4/4/2011	12
7	12	0107	SPS-1	New	GB	2/9/2000	4/4/2011	12
8	12	0108	SPS-1	New	GB	2/9/2000	4/4/2011	12
9	12	0109	SPS-1	New	GB	2/9/2000	3/30/2011	12
10	12	0110	SPS-1	New	ATB	2/9/2000	4/4/2011	12
11	12	0111	SPS-1	New	ATB	2/9/2000	3/30/2011	12
12	12	0112	SPS-1	New	ATB	2/9/2000	3/30/2011	12
13	12	0161	SPS-1	New	GB	2/9/2000	11/3/2006	10
14	12	0502	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/1/2013	16
15	12	0503	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/1/2013	16
16	12	0504	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	4/1/2011	15
17	12	0505	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/3/2013	16
18	12	0506	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/2/2013	16
19	12	0507	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/2/2013	16
20	12	0508	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/1/2013	16
21	12	0509	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/1/2013	16
22	12	0561	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/1/2013	14
23	12	0562	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/2/2013	14
24	12	0563	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/2/2013	14
25	12	0564	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/3/2013	14
26	12	0565	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	10/1/2013	14
27	12	0566	SPS-5 to GPS-6S	Overlay	GB	1/21/1996	1/29/2014	15
28	12	0901	SPS-9O	Overlay	GB	7/25/1996	10/11/2006	8
29	12	0902	SPS-9O	Overlay	GB	7/25/1996	10/11/2006	8
30	12	0903	SPS-9O	Overlay	GB	7/25/1996	10/11/2006	8
31	12	0959	SPS-9O	Overlay	GB	7/25/1996	1/24/2004	7
32	12	1370	GPS-1 to 6S	Overlay	GB	10/29/2001	10/10/2013	7
33	12	3997	GPS-1 to 6S	Overlay	GB	1/25/1996	3/1/1999	2
34	12	4096	GPS-2 to GPS-6C	Overlay	ATB	10/23/2003	3/25/2014	5
35	12	4100	GPS-2 to GPS-6S	Overlay	ATB	10/2/2002	2/13/2012	4
36	12	4101	GPS-1 to 6B	Overlay	GB	4/14/1992	1/22/1996	3
37	12	4106	GPS-1 to 6S	Overlay	GB	1/20/2005	5/11/2009	3
38	12	4135	GPS-1 to 6B	Overlay	GB	3/10/1994	5/5/2004	9
39	12	4136	GPS-1 to 6B	Overlay	GB	3/10/1994	5/5/2004	9
40	12	4137	GPS-1 to 6B	Overlay	GB	3/10/1994	6/9/2006	9

Average measured rutting values across the entire length of each test section can be extracted from the LTPP table MON_T_PROF_INDEX_SECTION, and the individual rutting values measured at 50-ft intervals can be extracted from the LTPP table MON_T_PROF_INDEX_ POINT.

GENERATION OF MEPDG INPUT VARIABLES BASED ON LTPP DATA

In the previous section, the availability of LTPP data in the selected Florida region was explored. This section explains the details of LTPP data extraction and the required calculations and assumptions to generate inputs for the AASHTOWare® Pavement ME Design software version 2.2.[(1)] The required MEPDG data have been classified into the following groups: project information, performance criteria, traffic data, climate data, pavement structure and materials, and permanent deformation (rutting) data. The LTPP SPS-1 and SPS-5 sites in Florida were selected for demonstration of the novel approach in this study for calibration of the rutting models for new and overlaid pavements, respectively.

Project Information

The software has an interface to enter general information for every design project (table 11). For calibration purposes, the significant information is the design life. The default value is 20 yr, but it is important to enter a design life that encompasses the available performance data to be able to calibrate the models to those data. As table 7 and table demonstrate, the latest rutting measurement date compared to the construction date of the original or overlay surface was on SPS-5 sections, and the surface age at the latest measurement date was 19 yr. Therefore, in this study, the default design life of 20 yr was found to be adequate for calibration.

Table 11. General project information.

Data Item	LTPP Source Tables	Data Field/Value
Section ID	EXPERIMENT_SECTION	SHRP_ID
Project location	EXPERIMENT_SECTION	STATE_CODE
Design type	New pavement or overlay per table 10	N/A
Pavement type	Flexible pavement	N/A
Design life (years)	20 yr	N/A
Base construction year/month	EXPERIMENT_SECTION	CN_ASSIGN_DATE
Pavement construction year/month	EXPERIMENT_SECTION	CN_ASSIGN_DATE
Traffic opening year/month	EXPERIMENT_SECTION	ASSIGN_DATE

Performance Criteria

The software has an interface to enter some performance criteria to be met by the specific pavement design (table 12). However, this information is not significant for the purpose of calibrating the performance models. Therefore, the default values were used for this study.

Table 12. Performance criteria.

Data Item	LTPP Source Tables	Data Field/Value
Initial IRI (m/km)	MEPDG defaults	1
Terminal IRI (m/km) limit	MEPDG defaults	2.7
Terminal IRI (m/km) reliability	MEPDG defaults	90
AC top-down fatigue cracking (m/km) limit	MEPDG defaults	378.8
AC top-down fatigue cracking (m/km) reliability	MEPDG defaults	90
AC bottom–up fatigue cracking (%) limit	MEPDG defaults	25
AC bottom–up fatigue cracking (%) reliability	MEPDG defaults	90
AC thermal cracking (m/km) limit	MEPDG defaults	189.4
AC thermal cracking (m/km) reliability	MEPDG defaults	90
Permanent deformation—total pavement (mm) limit	MEPDG defaults	19
Permanent deformation—total pavement (mm) reliability	MEPDG defaults	90
Permanent deformation—AC-only (mm) limit	MEPDG defaults	6
Permanent deformation—AC-only (mm) reliability	MEPDG defaults	90

IRI = International Roughness Index.

Traffic Data

Table 8 lists the general availability of the required traffic data within the LTPP database. In this section, the details of traffic input data extraction and calculation are explained. Table 13 lists some of the data sources used for traffic inputs. The LTPP TRF module contains tables that start with the TRF_MEPDG_ prefix. These tables contain traffic data estimated for input into the MEPDG software based on test sites that have more than 210 d of monitored data per year. However, for other site-years where the total number of accepted monitoring data was less than 210 d in that year, traffic data can still be found in the tables starting with the TRF_MONITOR_ prefix.

Table 13. Traffic input data sources and default values.

Data Item	LTPP Source Tables	Data Field/Default Value
Two-way AADTT	TRF_MONITOR_LTPP_LN	TRUCKS_LTPP_LN
Number of lanes	Not applicable	1
% of trucks in design direction	Not applicable	100
% of trucks in design lane	Not applicable	100
Operational speed (kph)	SECTION_GENERAL Not available for Florida sections	SPEED_LIMIT Used 100 kph based on posted limit.
Traffic capacity cap	Not enforced	N/A
Average axle width (m)	MEPDG defaults	2.59
Dual tire spacing (mm)	MEPDG defaults	305
Tire pressure (single tire) (kPa)	MEPDG defaults	827.4
Tandem axle spacing (m)	MEPDG defaults	1.31
Tridem axle spacing (m)	MEPDG defaults	1.25
Quad axle spacing (m)	MEPDG defaults	1.25
Mean wheel location (mm)	MEPDG defaults	460
Traffic wander standard deviation (mm)	MEPDG defaults	254
Design lane width (m)	MEPDG defaults	3.6576
Average spacing of short axles (m)	MEPDG defaults	3.66

Data Item	LTPP Source Tables	Data Field/Default Value
Average spacing of medium axles (m)	MEPDG defaults	4.57
Average spacing of long axles (m)	MEPDG defaults	5.49
Percent trucks with short axles	MEPDG defaults	33
Percent trucks with medium axles	MEPDG defaults	33
Percent trucks with long axles	MEPDG defaults	34
Vehicle class distribution (%)	TRF_MEPDG_VEH_CLASS_DIST Or TRF_MONITOR_LTPP_LN	Multiple fields in these tables are used.
Growth rate (%) by vehicle class	Estimated based on data from TRF_MONITOR_LTPP_LN	TRUCKS_LTPP_LN
Growth function	Compound growth was assumed.	Compound growth was assumed.
Monthly adjustment factors by vehicle class	TRF_MEPDG_MONTH_ADJ_FACTR	Multiple fields in these tables are used.
Hourly adjustment factors	MEPDG defaults were used.	MEPDG defaults were used.
Axles per truck for each vehicle class and axle group	Estimated based on data from TRF_MONITOR_LTPP_LN	Multiple fields in these tables are used.
Axle load distribution for every axle group	TRF_MEPDG_AX_DIST_ANL Or TRF_MONITOR_AXLE_DISTRIB	Multiple fields in these tables are used.

The traffic inputs in the MEPDG software comprise two main interfaces. The first interface includes basic traffic information, axle configuration, lateral wander, vehicle class distribution and growth, monthly adjustment, and axles per truck; the second interface includes the axle load distributions for single, tandem, tridem, and quad axles.

A software routine called LTPP *Pavement Loading User Guide* (PLUG) had previously been developed based on Microsoft® Access to populate the input data required for axle load distributions in the second interface.[(42)] However, the LTPP PLUG does not populate other traffic data required for the first interface in the MEPDG software. Therefore, during the current project, a series of Visual Basic for Applications (VBA) macros were developed in an Microsoft® Excel platform to extract, calculate, and create Extensible Markup Language (XML) data files to be imported into the MEPDG software for the first traffic inputs interface. In this manner, all traffic input data were imported into the MEPDG software using the generated XML files from the VBA macro codes created in this project and from the LTPP PLUG software.

Within the basic traffic information, the initial two-way average annual daily truck traffic (AADTT) is the most significant information. In this study, the LTPP table TRF_MONITOR_LTPP_LN, which is based on Automatic Vehicle Classification (AVC) and WIM equipment, was used as the source of AADTT. The data field TRUCKS_LTPP_LN in this table gives an estimate of the annual number of trucks in each class based on AVC and WIM information. The sum of these values for every year gives an estimate of AADTT in the LTPP lane for that year. Since the objective of this study is calibration of performance models based on performance measurements within LTPP test sections, these AADTT estimates were used along with the assumption that all of the roadway traffic has passed through the LTPP lane. In other words, a 100-percent value was used for directional and lane adjustment factors.

In MEPDG, the flexible pavement response model only requires the load spectrum, tire contact pressure distributions, and areas of contact for traffic characterization (see page 3.3.42 of the final report for NCHRP Project 1-37A).[2] However, for the slab cracking prediction model, the axle configuration, traffic wander, and wheelbase are considered critical factors (see page KK-9 and KK-10 of appendix KK of the final report for NCHRP Project 1-37A).[2] Since the axle configuration, lateral wander, and wheelbase information are only used for analysis and design of jointed concrete pavements, the default software values were used for this study.

The vehicle class distribution is extracted from the LTPP table TRF_MEPDG_VEH_CLASS_DIST where available (for any site-year where more than 210 d of monitored data have been recorded) and estimated from the LTPP table TRF_MONITOR_LTPP_LN otherwise. To generate MEPDG input values, the annual class distributions were averaged among all years of available data.

Data from the TRUCKS_LTPP_LN field in the TRF_MONITOR_LTPP_LN table were used to estimate a traffic growth factor by each vehicle class. For every vehicle class in each site, the available data were used with linear interpolation to fill in the gaps between the final available year and the initial year (for which AADTT was input) and create a continuous series of truck counts. Then the initial counts and the final cumulative counts were used with an assumption of compound growth to estimate the growth factor *r*. In equation 20, T_f is the cumulative truck count for every class at the final year of available data, T_i is the truck count at the initial year, and Y is the difference in the number of years between the initial year and the final year of available data. Equation 20 was recursively solved using a VBA macro to estimate *r*, or growth factor, for every vehicle class:

$$T_f = T_i\left(\frac{(1+r)^Y - 1}{r}\right) \tag{20}$$

Where:

T_f = cumulative traffic at the end of the period.
T_i = traffic at the beginning of the period.
r = growth factor.
Y = period in years.

For monthly adjustment factors, data from the LTPP table TRF_MEPDG_MONTH_ADJ_FACTR were used. The LTPP TRF_MEPDG_HOURLY_DISTRIB has hourly adjustment factors only for the LTPP sites that were in the traffic pooled fund study. Since hourly distributions are more important in analysis of jointed concrete pavements, the MEPDG default values were used for hourly distributions in this study.

For every vehicle class, based on the estimated count of axles in each axle group (single, tandem, tridem, or quad) within the TRF_MONITOR_LTPP_LN table, the number of axles per truck is calculated. This number is then averaged among all years with available data to generate the MEPDG inputs.

For each LTPP site-year with adequate data (more than 210 d of monitored data per year), the TRF_MEPDG_AX_DIST_ANL table was used to generate the MEPDG required inputs. For other site-years, data from the TRF_MONITOR_AXLE_DISTRIB table were converted to generate the corresponding MEPDG inputs. All of these data were reformatted to be input into the LTPP PLUG tables. Then the PLUG software was used to generate the required XML files for each site.

Climate Data

The AASHTOWare® Pavement ME Design software has two options to gather the climatic data—a single weather station or a virtual weather station. The appropriate option is selected based on the proximity of the section to a certain weather station. Existing weather stations (the data for which can be downloaded at http://me-design.com/MEDesign/ClimaticData.html) are listed by location and can be searched by latitude and longitude. The latitude and longitude data are available in table SECTION_COORDINATES of the LTPP database. Table 14 lists the data source and the corresponding fields.

Table 14. Climate information.

Data Item	LTPP Source Tables	Data Field/Value
Longitude (degrees, minutes)	SECTION_COORDINATES	LONGITUDE
Latitude (degrees, minutes)	SECTION_COORDINATES	LATITUDE
Elevation (ft)	AWS_LOCATION	ELEVATION
Depth of water table (ft)	N/A	5 m (+1)
Climate station	MEPDG	N/A

Note: (+1) a default value of 5 m was set for the depth of the water table, since this information is not available from the LTPP database.

The LTPP InfoPave™ website has a tool for extracting the National Aeronautics and Space Administration MERRA climatic data in the form of HCD files for each LTPP test section.[40] This tool also provides the station information in the form of a station.dat file. These files can be copied to the climatic data folder, which has been designated by the user for the AASHTOWare® software.[1] This way, the user will also have the option to use MERRA climatic data. For the current project, only the existing weather station data were used.

Pavement Structure and Materials Data

Pavement layers are identified in LTPP with a unique number (LAYER_NO). Layer number 1 is always assigned to the lowest layer (subgrade) in the pavement structure, and additional layers above it are indicated with progressively larger layer numbers. In addition to LAYER_NO, which is specific to each test section, for the SPS projects, which have more than one test section per site, a layer identifier, PROJECT_LAYER_CODE, is available. Pavement layers with the same material properties from different test sections along the entire site are identified using this field code. Even though the sequence of a layer might be different in different test sections (different LAYER_NO), a similar PROJECT_LAYER_CODE indicates that those layers from different test sections were constructed at the same time and using similar materials.

Since not every section has test results for each layer, an expansion data process was applied to populate the layer structure properties. Under this process, the test results from one section are expanded to the other section when the PROJECT_LAYER_CODE in both sections is the same for a given layer. This process is applied to the entire site for the Florida SPS-1 and SPS-5 test sections.

Before applying the described procedure, sections need to be ordered according to the construction sequence (SECTION_START and SECTION_END fields in the SPS_PROJECT_STATIONS table). Then blank fields are filled with data available from the closest section having the same PROJECT_LAYER_CODE. The description of the layer was also taken into consideration for the expansion of subgrade test results.

Layer Thickness and Type of Material

Layer thickness and type of material are extracted from the LTPP table TST_L05B (SECTION_LAYER_STRUCTURE). Table 15 summarizes the data required for input into MEPDG software and the corresponding LTPP tables and fields.

Table 15. Layer thickness and type of material.

Data Item	LTPP Source Tables	Data Field/Value
Thickness (mm)	TST_L05B, SECTION_LAYER_STRUCTURE	REPR_THICKNESS
Construction number	TST_L05B, SECTION_LAYER_STRUCTURE	CONSTRUCTION_NO
Layer number	TST_L05B, SECTION_LAYER_STRUCTURE	LAYER_NO
Type of material	TST_L05B, SECTION_LAYER_STRUCTURE	MATL_CODE
Material that is similar among SPS sections on one site	SECTION_LAYER_STRUCTURE	PROJECT_LAYER_CODE

New Asphalt Concrete Layer

Wherever data were available, input level 1 was considered for the AC layer. Under this input level, binder properties, mixture volumetric properties, dynamic modulus, and creep compliance need to be provided according to laboratory test results. Mixture volumetric properties calculations were made using available LTPP test results prior to introducing the data into the software. Dynamic modulus master curve is calculated internally in the software when it is provided with the dynamic modulus for a combination of laboratory test results. Individual values for creep compliance are introduced, and the software calculates the master curve.

Mixture Volumetric Information

The mixture volumetric information was calculated with the LTPP data and applying the weight–volume relationships for asphalt mixtures.[43] Relevant values for the volumetric calculation were taken from LTPP TST tables as described in table 16.

Table 16. Mixture volumetric data.

Data Item	LTPP Source Tables	Data Field/Value
Unit weight (kg/m^3)	TST_AC02	BSG 1,000
Effective binder content (%)	TST_AC04	ASPHALT_CONTENT_MEAN (P_b = asphalt, percent by total weight of mixture)
Effective binder content (%)	TST_AC03	MAX_SPEC_GRAVITY (G_{mm} = maximum SG of paving mixture)
Effective binder content (%)	TST_AG01	BSG_OF_COARSE_AGG (G_{sb} = BSG of aggregate)
Effective binder content (%)	TST_AG02	BSG_OF_FINE_AGG (G_{sb} = BSG of aggregate)
Air voids (%)	TST_AC02	BSG (G_{mb} = BSG of compacted mixture)
Air voids (%)	TST_AC03	MAX_SPEC_GRAVITY (G_{mm} = maximum SG of paving mixture)

BSG = bulk specific gravity.

The following relationships were applied to calculate the effective binder content (equation 21) and the air voids (equation 26).

$$P_{be} = P_b - \left(\frac{P_{ba}}{100}\right)(1 - P_b) \quad (21)$$

Where:

P_{be} = effective binder content (%).

P_b = asphalt, percent by total weight of mixture.

P_{ba} = absorbed asphalt, percent by weight of aggregate, calculated using equation 22:

$$P_{ba} = 100\left(\frac{G_{se} - G_{sb}}{G_{sb}G_{se}}\right)G_b \quad (22)$$

Where G_{se} is the effective specific gravity (SG) of aggregate, calculated using equation 23:

$$G_{se} = \frac{1 - P_b}{\frac{1}{G_{mm}} - \frac{P_b}{G_b}} \quad (23)$$

Where G_{sb} is the bulk specific gravity (BSG) of aggregate, calculated using equation 24:

$$G_{sb} = \frac{P_1 + P_2 + \cdots .. + P_n}{\frac{P_1}{G_1} + \frac{P_2}{G_2} + \cdots . + \frac{P_n}{G_n}} \quad (24)$$

Where:

P_i = percentages by weight of aggregates.

G_i = BSG of aggregates.

G_b = asphalt SG, considered 1.01.(44)

G_{mm} = maximum SG of paving mixture, calculated using equation 25:

$$G_{mm} = \frac{100}{\frac{P_s}{G_{se}} + \frac{P_b}{G_b}} \quad (25)$$

Then the air voids are calculated with equation 26:

$$V_a = \left(1 - \frac{G_{mb}}{G_{mm}}\right).100 \quad (26)$$

Where:

V_a = air voids (%).

G_{mb} = BSG of compacted mixture.

Binder and Asphalt Concrete Properties

Binder properties are available in the LTPP database, except for the softening point (table 17). This item is obtained indirectly by applying the correlation between temperature and viscosity,

considering that the softening point value is associated uniquely with a viscosity of 13,000 poise.[45] Viscosity–temperature parameters correspond to MEPDG recommended values for an asphalt cement grade 40-50 (A = 10.5254 and VTS = –3.5047). Applying equation 27, the softening point temperature for the specified viscosity is 47 ºC.

$$\log\log\eta = A + VTS\ log\ T_r \tag{27}$$

Where:

η = viscosity, centipoise (cP).
T_r = reference temperature, R.
A = regression intercept.
VTS = regression slope or viscosity temperature susceptibility.

Table 17. Binder properties.

Data Item	LTPP Source Tables	Data Field/Value
Poisson's ratio	MEPDG defaults	MEPDG defaults
Softening point (ºC) at 1,300 pascal-sec	MEPDG defaults	47 ºC
Absolute viscosity (pascal-sec) at 60 ºC	TST_AE_02	ABSOLUTE_VISC_140_F
Kinematic viscosity (centistokes) at 135 ºC	TST_AE_02	KINEMATIC_VISC_275_F
SG at 25 ºC	TST_AE_03	SPECIFIC_GRAVITY
Penetration at temperature 25 ºC	TST_AE_02	PENETRATION_77_F

AC dynamic modulus master curve is calculated internally in the MEPDG. This calculation is done based on the individual dynamic modulus values entered for a set of frequency and temperature combinations. The general expression for the dynamic modulus master curve is in equation 28:[1]

$$log(E^*) = \delta + \frac{\alpha}{1 + e^{\beta+\gamma\left[log\,(t)-c\left(log\,(\eta)-log\,(\eta_{T_r})\right)\right]}} \tag{28}$$

Where:

t = time of loading at a given temperature of interest.
δ, α = fitting parameters; for a given set of data, δ represents the minimum value of E^*, and $\delta + \alpha$ represents the maximum value of E^*.
β, γ = parameters describing the shape of the sigmodal function.
c = fitting parameter.
η_{T_r} = viscosity at reference temperature.

It should be noted that the dynamic modulus values available within the LTPP TST_ESTAR_ tables have been calculated using an ANN model developed under a previous LTPP data analysis project. This ANN model uses actual laboratory-tested resilient modulus values available within the LTPP TST_AC07_ tables. Considering that calculated dynamic modulus values based on a relationship to the resilient modulus test results are used here, by MEPDG definition, this is level 2 input. However, the only option within the software to enter calculated dynamic modulus values was to use the level 1 input option. If the level 2 input

option is selected, the software requires material properties other than the resilient modulus to predict the dynamic modulus based on the Witczak model.

Table 18 shows the LTPP tables and corresponding fields where the required input values regarding binder and mixture properties have been stored. In using the TST_ESTAR_ tables, the ESTAR_LINK corresponding to PREDICTIVE_MODEL number 1 was used, as the first predictive ANN model is the one estimating dynamic modulus based on resilient modulus lab testing results.

Table 18. Mixture properties.

Data Item	LTPP Source Tables	Data Field/Value
Poisson's ratio	MEPDG defaults	MEPDG defaults
Temperature levels for dynamic modulus	TST_ESTAR_MODULUS (TST_ESTAR_MASTER.PREDICTIVE_ MODEL = 1) which has been populated based on TST_AC07_V2_MR_SUM values	TEMPERATURE
Frequency levels for dynamic modulus	TST_ESTAR_MODULUS (TST_ESTAR_MASTER.PREDICTIVE_ MODEL = 1) which has been populated based on TST_AC07_V2_MR_SUM values	FREQUENCY
Dynamic modulus	TST_ESTAR_MODULUS (TST_ESTAR_MASTER.PREDICTIVE_ MODEL = 1) which has been populated based on TST_AC07_V2_MR_SUM values	ESTAR
Creep compliance (1/GPa)	MEPDG defaults	MEPDG defaults
Thermal conductivity (watt/meter-kelvin)	MEPDG defaults	MEPDG defaults
Heat capacity (joule/kilogram-kelvin)	MEPDG defaults	MEPDG defaults
Thermal contraction	MEPDG defaults	MEPDG defaults

Existing Asphalt Concrete Properties

SPS-5 sections are experiments with HMA overlays. For this type of pavement structures, the MEPDG requires the backcalculated elastic modulus or the results of a structural adequacy evaluation of the existing pavement. In all the sections, the backcalculated elastic modulus was obtained from the LTPP database. During the history of the LTPP program, two data analysis studies were conducted to backcalculate all of the FWD data. The most recent study was the LTPP Determination of In-Place Elastic Layer Modulus: Backcalculation Methodology and Procedure.(46) The latest backcalculated elastic modulus data became available since the SDR 29.0 and can be downloaded from the InfoPave™ website.(40) The backcalculated elastic modulus along with the frequency and temperature are required in the MEPDG software for the overlay analysis. Table 19 shows the LTPP data tables and fields where the backcalculated moduli and the corresponding information can be found.

Table 19. LTPP data tables and fields for backcalculated moduli.

Data Item	LTPP Source Tables	Data Field/Value
Backcalculated modulus averaged for each FWD pass	BAKCAL_MODULUS_SECTION_LAYER	AVG_MODULUS
Backcalculated layer	BAKCAL_MODULUS_SECTION_LAYER	BC_LAYER_NO
Corresponding LTPP layer number	BAKCAL_LAYER_LINK	LAYER_NO
FWD pass number	BAKCAL_MODULUS_SECTION_LAYER	FWD_PASS
Test date	BAKCAL_PASS	TEST_DATE
Temperature	BAKCAL_BASIN	SURFACE_TEMP
Frequency	N/A	15 Hz
Gradation percent passing 3/4-inch sieve	TST_AG04	THREE_FOURTHS_PASSING
Gradation percent passing 3/8-inch sieve	TST_AG04	THREE_EIGHTHS_PASSING
Gradation percent passing Nº4 sieve	TST_AG04	NO_4_PASSING
Gradation percent passing Nº200 sieve	TST_AG04	NO_200_PASSING

For the testing frequency, several studies have recommended a value of 1/2*t* where *t* is the period of the FWD load pulse and can be extracted from the FWD time history data that are available in the LTPP Ancillary Information Management System and can be downloaded through InfoPave™.[44,40] Based on an observation of several of the FWD time histories, it was decided that a value of 16 Hz corresponding to $t = 0.03$ s was suitable to use in this project. This value also coincides with the findings from several past studies.[48–50]

Asphalt-Treated Base and Permeable Asphalt-Treated Base

Some of the SPS-1 sections have asphalt-treated bases (ATBs) and permeable asphalt-treated bases. Those materials behave similarly to HMA concrete. So, LTPP source tables and fields are the same as the ones described for HMA.

Additional Asphalt Concrete Layer Properties

The additional AC layer properties required by the software are listed in table 20.

Table 20. Additional AC layer properties.

Data Item	LTPP Source Tables	Data Field/Value
Surface shortwave absorptivity	MEPDG defaults	0.85
Is endurance limit applied?	MEPDG defaults	False
Endurance limit (microstrain)	MEPDG defaults	100
Layer interface	MEPDG defaults	Full friction interface

Unbound and Subgrade Materials

Unbound materials response is characterized by the resilient modulus calculated from the LTPP materials testing module data (TST_UG07_SS07_* tables). The representative resilient modulus was calculated according to the guidelines provided in the NCHRP Project 1-28A study, which found that the summary resilient modulus should be reported using equation 29

and calculated for the following stress states: σ_3 = 5 psi and σ_1 = 15 psi for aggregate base/subbase and σ_3 = 2 psi and σ_1 = 6 psi for subgrade soils.[51] The calculated resilient modulus values are listed in table 39 of appendix A.

$$M_r = k_1 P_a \left[\frac{\theta}{P_a}\right]^{k_2} \left[\frac{\tau_{oct}}{P_a} + 1\right]^{k_3} \quad (29)$$

Where:

k_1, k_2, k_3 = regression constants.
P_a = atmospheric pressure equal to 14.7 psi.
θ = bulk stress, calculated using equation 30:

$$\theta = \sigma_1 + 2\sigma_3 \quad (30)$$

Where τ_{oct} is the octahedral shear stress, calculated using equation 31:

$$\tau_{oct} = \frac{\sqrt{2}}{3}(\sigma_1 - \sigma_3) \quad (31)$$

Where σ_1 and σ_3 are the principal stresses.

Table 21 shows the LTPP source tables and fields for the required input data for unbound aggregate and subgrade soils materials properties.

Table 21. LTPP data sources for unbound materials properties.

Data Item	LTPP Source Tables	Data Field/Value
Layer thickness (inches)	TST_L05B	REPR_THICKNESS
Poisson's ratio	N/A	MEPDG defaults based on AASHTO soil classification
Coefficient of lateral earth pressure (Ko)	N/A	MEPDG defaults based on AASHTO soil classification
Resilient modulus (level 2)	N/A	Equation 29
Average applied max axial stress	TST_UG07_SS07_WKSHT_ SUM	APPLIED_MAX_AXIAL_STRESS_ AVG (σ_1)
Confining pressure	TST_UG07_SS07_WKSHT_ SUM	CON_PRESSURE (σ_3)
Average resilient modulus	TST_UG07_SS07_WKSHT_ SUM	RES_MOD_AVG (M_r)
Average applied cyclic stress	TST_UG07_SS07_WKSHT_ SUM	APPLIED_CYCLIC_STRESS_AVG (S_{cyclic})
Average resilient strain	TST_UG07_SS07_WKSHT_ SUM	RES_STRAIN_AVG (ε_r)
Percent passing 0.020 mm	TST_SS02_UG03	HYDRO_02
Percent passing # 200	TST_SS01_UG01_UG02	NO_200_PASSING
Percent passing # 80	TST_SS01_UG01_UG02	NO_80_PASSING
Percent passing # 40	TST_SS01_UG01_UG02	NO_40_PASSING
Percent passing # 10	TST_SS01_UG01_UG02	NO_10_PASSING
Percent passing # 4	TST_SS01_UG01_UG02	NO_4_PASSING
Percent passing 3/8"	TST_SS01_UG01_UG02	THREE_EIGHTHS_PASSING
Percent passing 1/2"	TST_SS01_UG01_UG02	ONE_HALF_PASSING
Percent passing 3/4"	TST_SS01_UG01_UG02	THREE_FOURTHS_PASSING
Percent passing 1"	TST_SS01_UG01_UG02	ONE_PASSING
Percent passing 1 1/2"	TST_SS01_UG01_UG02	ONE_AND_HALF_PASSING
Percent passing 2"	TST_SS01_UG01_UG02	TWO_PASSING
Percent passing 3"	TST_SS01_UG01_UG02	THREE_PASSING
Liquid Limit	TST_UG01_SS03	LIQUID_LIMIT
Plasticity Index	TST_UG01_SS03	PLASTICITY_INDEX
AASHTO soil classification	TST_AG04	NO_10_PASSING
AASHTO soil classification	TST_SS01_UG01_UG02	NO_4_PASSING and NO_200_PASSING
AASHTO soil classification	TST_UG01_SS03	PLASTIC_LIMIT and PLASTICITY_INDEX
Maximum dry unit weight (pcf)	N/A	MEPDG defaults based on AASHTO soil classification
Saturated hydraulic conductivity (m/hr)	N/A	MEPDG defaults based on AASHTO soil classification
Specify gravity of soils (Gs)	N/A	MEPDG defaults based on AASHTO soil classification
Optimum gravimetric water content (%)	N/A	MEPDG defaults based on AASHTO soil classification
Soil water characteristic curve parameter (af, bf, cf, hr)	N/A	MEPDG defaults based on AASHTO soil classification

N/A = no adequate data.

The AASHTOWare® Pavement ME Design software does not allow an ATB layer directly on top of the subgrade. Therefore, the subgrade was split into two layers with the same materials, each one with half the subgrade thickness.

For the Florida SPS-1 site, the pavement structure is supported by a compacted limerock embankment. The construction report indicated a hard material underneath the embankment that, for the modeling purposes, was considered as bedrock with the properties noted in table 22.

Table 22. Bedrock material properties.

Data Item	LTPP Source Tables	Data Field/Value
Layer thickness	MEPDG defaults	Semi-infinite
Unit weight	MEPDG defaults	2,240
Poisson's ratio	MEPDG defaults	0.15
Elastic modulus	MEPDG defaults	5,171

Backcalculated values of resilient modulus are applied when laboratory results are not available. These backcalculated values need to be adjusted to laboratory conditions to use in ME design. The adjustment to laboratory condition is done internally by the software according to the selected C-value that is listed in table 23.[(52)]

Table 23. C-values to convert the backcalculated layer modulus values to an equivalent resilient modulus measured in laboratory.

Layer Type	Location	C-value of M_r/E_{FWD} Ratio
Aggregate base/subbase	Between a stabilized and HMA layer	1.43
Aggregate base/subbase	Below a PCC layer	1.32
Aggregate base/subbase	Below an HMA layer	0.62
Subgrade/embankment	Below a stabilized subgrade/embankment	0.75
Subgrade/embankment	Below an HMA or PCC layer	0.52
Subgrade/embankment	Below an unbound aggregate base	0.35

PCC = portland cement concrete.

Pavement Permanent Deformation

During each LTPP manual distress survey, the transverse profile of the pavement sections is measured at every 50 ft, which results in rutting measurements for both wheelpaths at 11 test locations across the length of each test section. Point-by-point rutting measurements and calculated average section rutting values are stored in the LTPP database. Table 24 shows the LTPP data source for rutting measurements.

Table 24. LTPP data source for rutting measurements (wire reference method).

Data Item	LTPP Source Tables	Data Field/Value
Measurement date	MON_T_PROF_INDEX_SECTION MON_T_PROF_INDEX_POINT	SURVEY_DATE
Average section left wheelpath rutting	MON_T_PROF_INDEX_SECTION	LLH_DEPTH_WIRE_REF_MEAN
Average section right wheelpath rutting	MON_T_PROF_INDEX_SECTION	RLH_DEPTH_WIRE_REF_MEAN
Point-by-point left wheelpath rutting	MON_T_PROF_INDEX_POINT	LLH_DEPTH_WIRE_REF
Point-by-point right wheelpath rutting	MON_T_PROF_INDEX_POINT	RLH_DEPTH_WIRE_REF

Rutting measurements used to be conducted using a 1.8-m straightedge and a reference wire. Later, the LTPP program adopted the Face Dipstick device, which measures the transverse profile elevations at every foot along the width of the lane. Even after using the Dipstick, the rutting values have been recorded according to a simulated straightedge and the wire reference methods. There has been no concrete evidence as to which method produces a more repeatable or representative rutting measurement. In this study, the values recorded according to the wire reference method have been used as measured rutting values to be used in the calibration of permanent deformation models. In the wire reference method, the maximum displacement between the reference wire line and pavement surface is calculated in the left- and right-lane halves. Reference wire is placed at profile end points and connects peaks, which protrude above horizontal datum end points, with straight lines. Displacement is computed perpendicular to horizontal datum between end points.

Pavement permanent deformation is the result of incremental deformation in each layer of the pavement. MEPDG calculates the incremental deformation for each subseason at the mid-depth of each sublayer within the pavement system. Each layer contributes to the total permanent deformation according to its material properties, climate, and load conditions. The rutting measurements in the LTPP database are for the total pavement structure, and trenching measurements are not available to calibrate the permanent deformation models for each layer independently. Therefore, the calibration factors for the following models need to be adjusted in a way that minimizes the difference between the LTPP measured rutting and the total pavement rutting calculated using equations 32 and 33.

$$RD = \sum_{i=1}^{n} \varepsilon_p^i \,.\, h^i \tag{32}$$

$$\Delta RD = \Delta_{p(HMA)} + \Delta_{p(base)} + \Delta_{p(subbase)} + \Delta_{p(soil)} \tag{33}$$

Where:

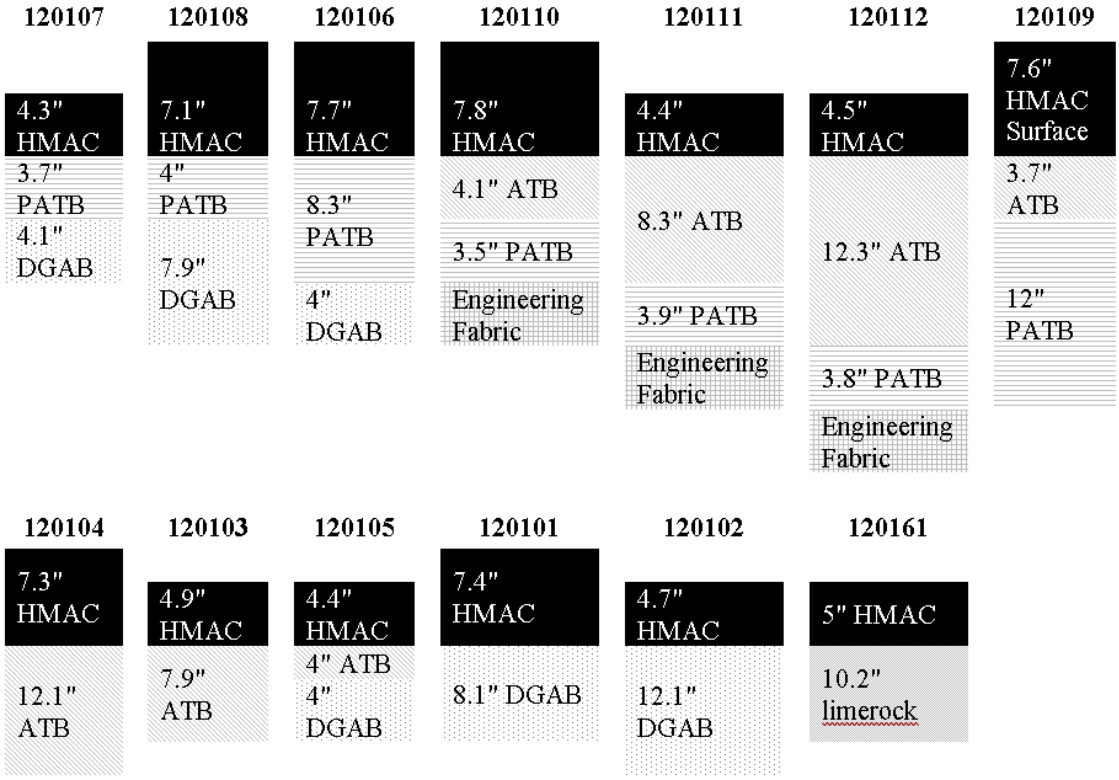
RD = pavement permanent deformation.
ε_p^i = total plastic strain in sublayer i.
h^i = thickness of sublayer i.
n = number of sublayers.
$\Delta_{p(HMA)}$ = accumulated permanent or plastic vertical deformation in the HMA layers/sublayers, calculated using equation 1 (inches).
$\Delta_{p(base)}$, $\Delta_{p(subbase)}$, $\Delta_{p(soil)}$ = permanent or plastic deformation for the unbound layers/sublayers, calculated using equation 5 (inches).

ASSEMBELED CALIBRATION DATASETS

This section of the report presents general information on the different LTPP and non-LTPP datasets that have been used in this project. Data from 13 Florida LTPP SPS-1 test sections and 11 FDOT APT sections were used in calibrating the permanent deformation model for new pavements. Data from 15 Florida LTPP SPS-5 sections were used in calibrating the permanent deformation models for overlaid pavements.

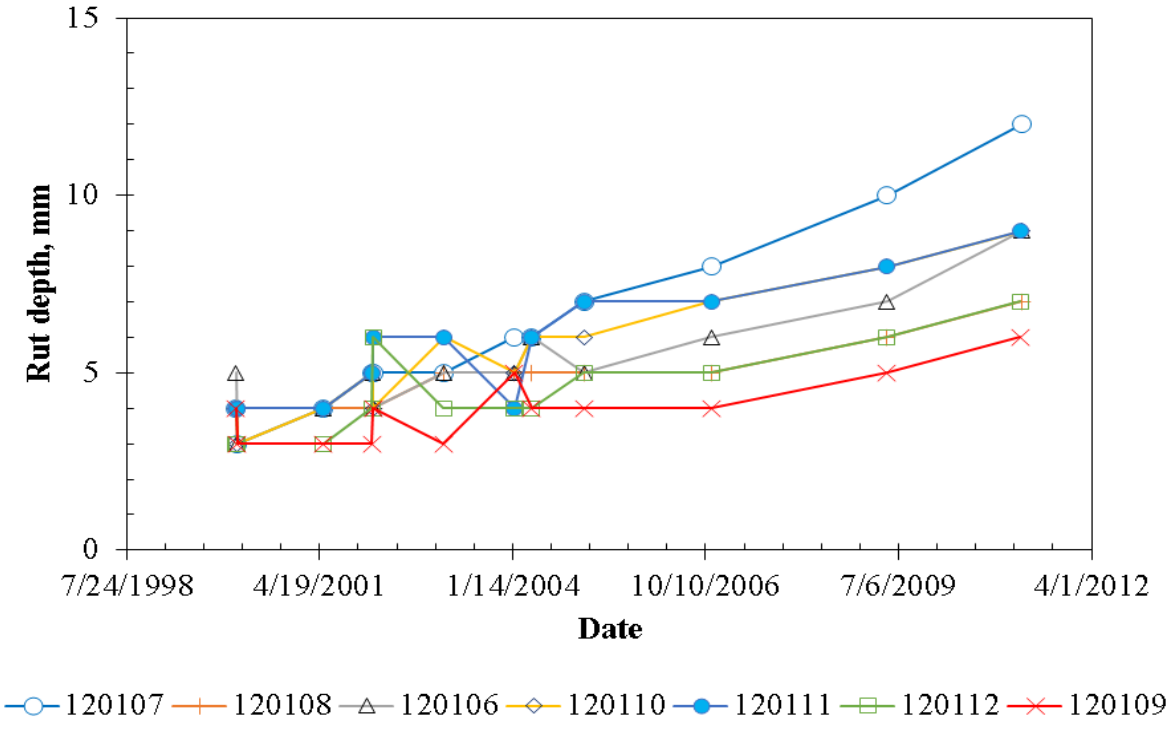
SPS-1 Sections

Figure 5 shows the pavement structure in the Florida SPS-1 test sections in the same order that they exist onsite. Figure 6 and figure 7 show the trend in the average section rutting values measured on Florida SPS-1 test sections from 2000 to 2011. As it can be seen, despite the increasing trend in rutting values with time, the trends for different test sections are not parallel to each other, and they cross at several points. This indicates the inherent variability in the involved parameters and the measurement methods. Also, the averaging of rutting values from different locations on each test section obscures the real trends.



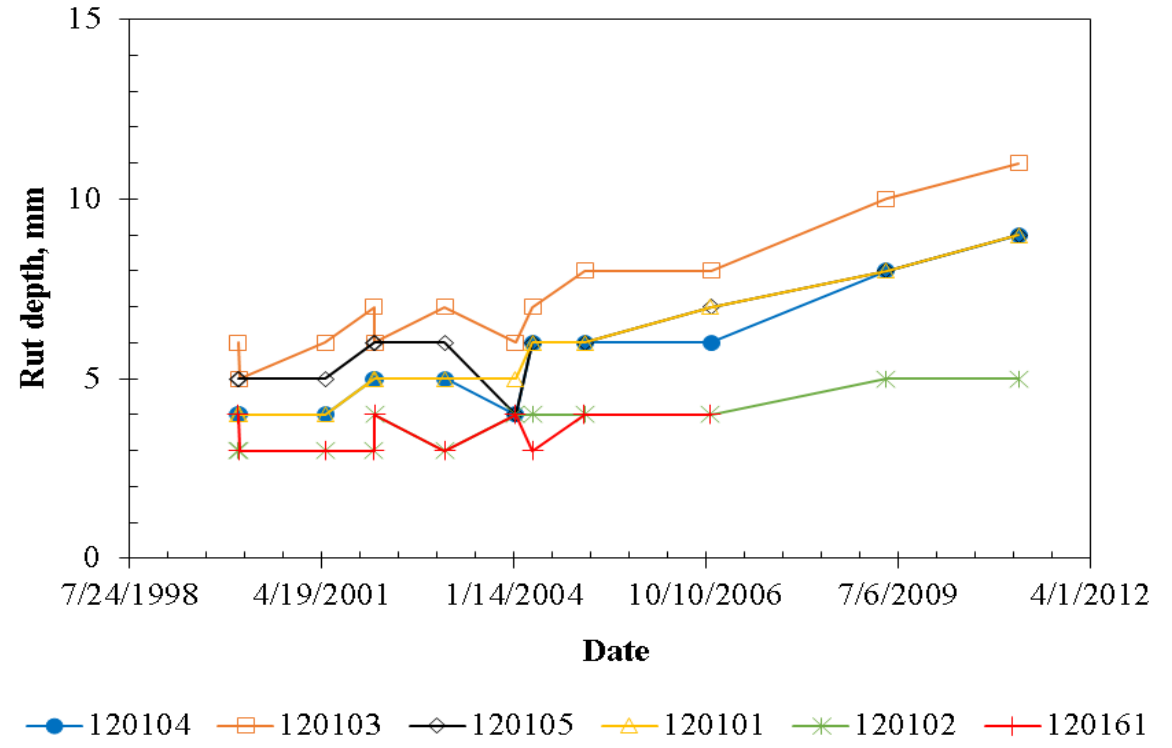
Source: FHWA.
HMAC = hot-mix asphalt concrete; DGAB = dense-graded aggregate base.

Figure 5. Illustration. Pavement structure in Florida SPS-1 test sections.



Source: FHWA.

Figure 6. Chart. Average rutting measurements on SPS-1 test sections 120107 to 120109.



Source: FHWA.

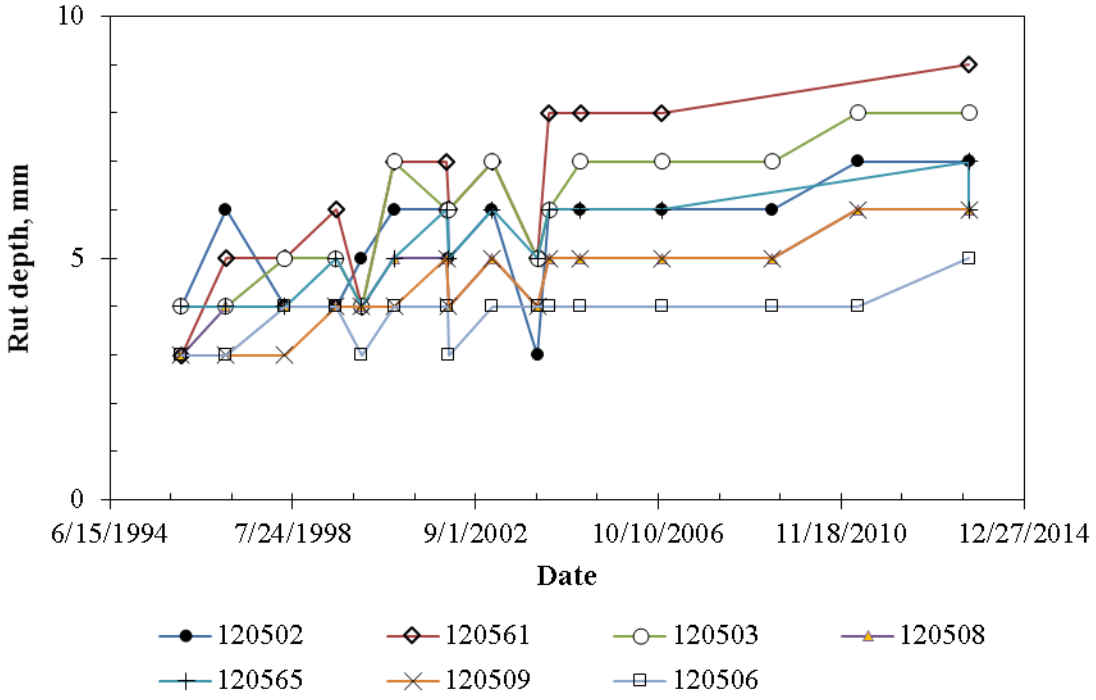
Figure 7. Chart. Average rutting measurements on SPS-1 test sections 120104 to 120161.

Figure 8 shows the pavement structure in the Florida SPS-5 test sections in the same order that they exist on the site. Figure 9 and figure 10 show the trend in the average section rutting values measured on Florida SPS-5 test sections from 1995 to 2013.

SPS-5 Sections

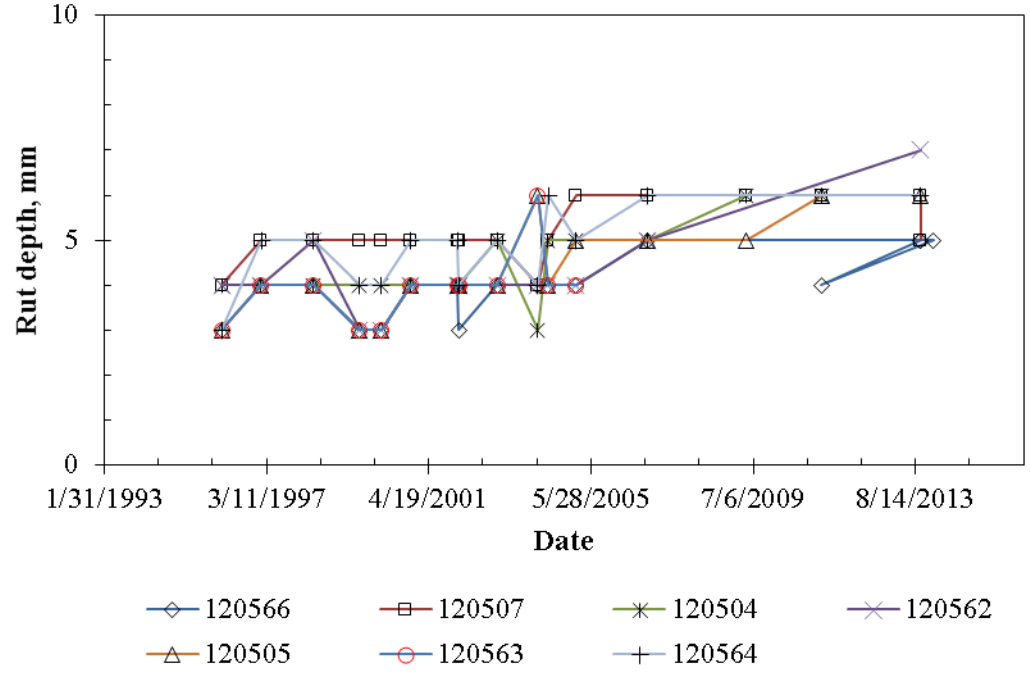
Source: FHWA.

Figure 8. Illustration. Pavement structure in Florida SPS-5 test sections.



Source: FHWA.

Figure 9. Chart. Average rutting measurements on SPS-5 test sections.



Source: FHWA.

Figure 10. Chart. Average rutting measurements on SPS-5 test sections.

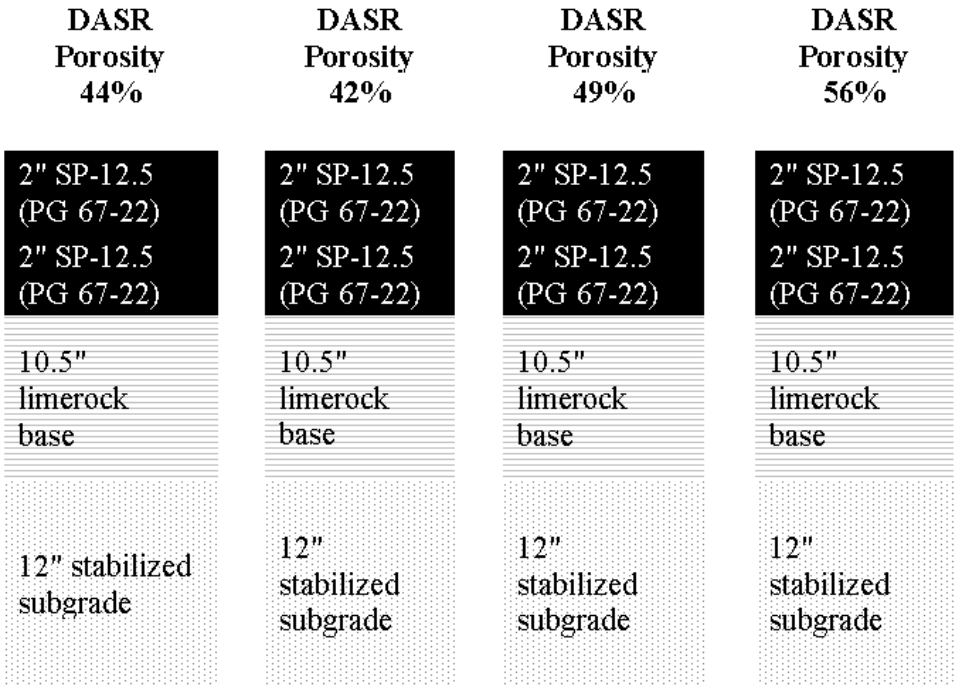
FDOT APT Data

FDOT has an APT facility with a heavy vehicle simulator (HVS). FDOT has used this facility for conducting several AC pavement experiments, of which two were selected in this project to provide information for the rutting calibration process.

Dominant Aggregate Size Range Gradation Model Experiment

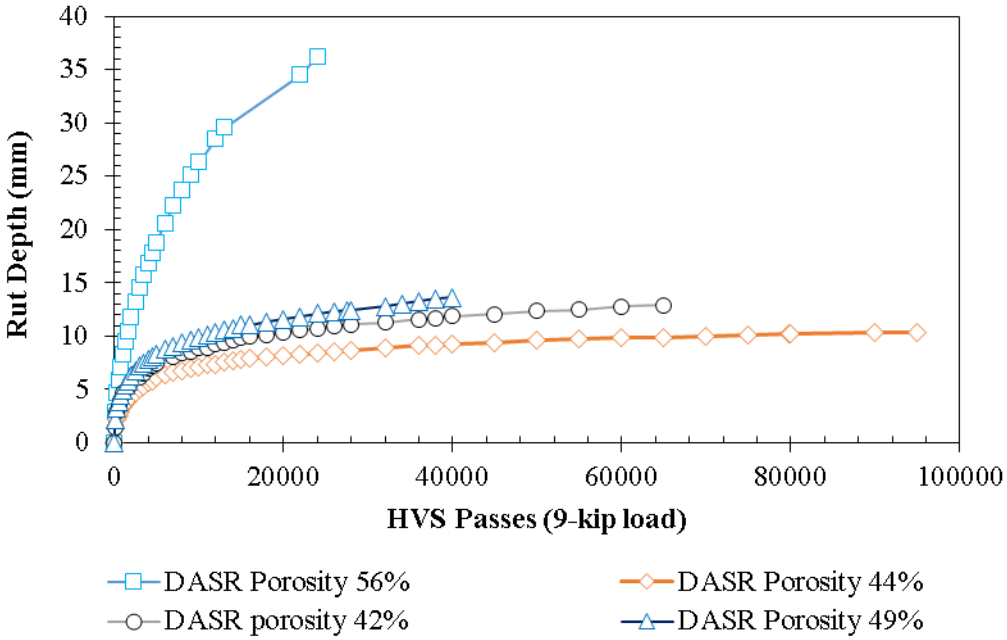
FDOT established an accelerated pavement experiment to test various aggregate gradations to resist rutting. The approach taken by FDOT is known as the dominant aggregate size range (DASR) gradation model. Four sections were built to be tested under the HVS. Test sections were trafficked at 50 ºC using a 455-mm wide-base single tire inflated to 390 kPa and loaded to 40 kN.[53]

Each test section consisted of two layers of HMA of 2 inches. Those layers were placed over a 20.5-inch limerock base and a 12-inch limerock stabilized subgrade. The DASR porosity ranged from 42 to 56 percent, and mixes with lower DASR porosity exhibited less rutting. Figure 11 presents a schematic of the pavement structure in the four test sections, and figure 12 shows the rutting measurements.



Source: FHWA.
SP-12.5 = Superpave with nominal maximum aggregate size of 12.5 inches.

Figure 11. Illustration. FDOT DASR project sections.[54]



Source: FHWA.

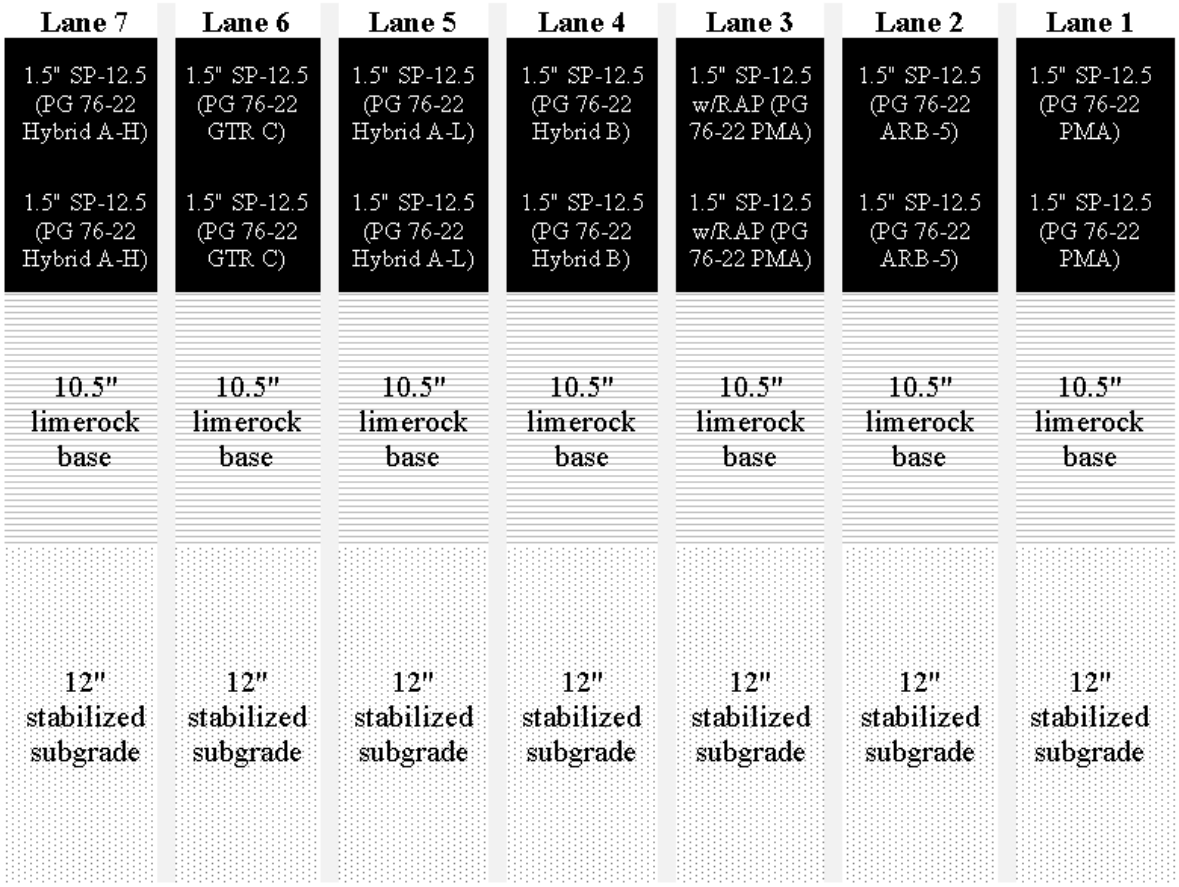
Figure 12. Chart. Rutting for the four sections tested under FDOT DASR project.

Asphalt Rubber Binder Experiment

FDOT built a set of seven APT sections as a part of a study to compare HMA mixtures constructed using a polymer-modified asphalt (PMA) and an asphalt rubber binder (ARB). The objective of this study was to find a way to make ARB handle and perform similarly to PG 76-22, Florida's "gold standard" binder. The experiment included three PMA sections and four ARB sections. Testing lanes were milled and resurfaced, approximately 1 inch of existing asphalt remained in place after milling, and then each lane was resurfaced with two 1.5-inch layers of a 12.5-inch nominal maximum aggregate size fine-graded Superpave mixture. The asphalt mixtures of each lane were the same except for the binder type.[(54)]

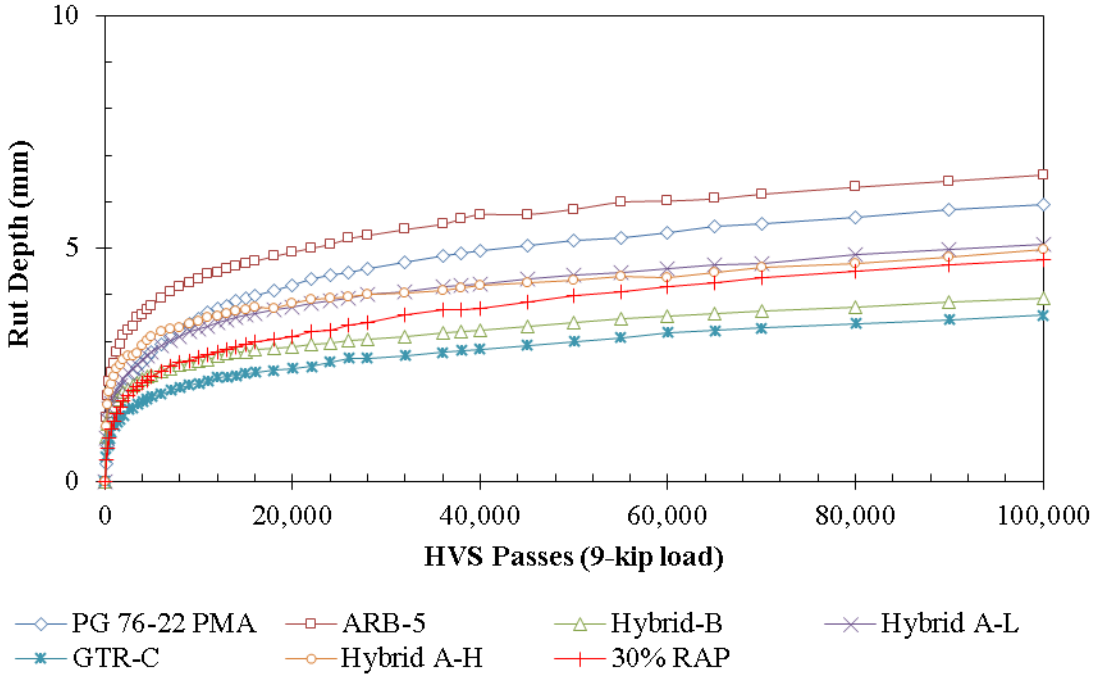
Figure 13 presents a schematic of the pavement structure in the four test sections, and figure 14 shows the rutting measurements.

Accelerated loading was performed using FDOT's HVS with a super single tire loaded to 9 kip and inflated to 110 psi. A wheel wander of 4 inches was used and the test temperature maintained at 50 ºC. All the mixtures showed good rutting performance. All mixtures with a PG 76-22 (ARB) exhibited slightly better rutting resistance than the control mix.



Source: FHWA.
GTR = ground tire rubber.

Figure 13. Illustration. FDOT ARB project sections.



Source: FHWA.

Figure 14. Chart. Rutting for the seven sections tested under FDOT ARB project.

CHAPTER 4. PROGRAMMING METHODOLOGY

INTRODUCTION

In the proposed approach for the execution of this research project, task 4 was anticipated for all the programming activities that were needed to implement a multi-objective calibration of the MEPDG performance models. An overview of the general programming activities is described in the following sections of this report. First, the approach to implement the AASHTO recommended calibration (single-objective) is presented. Second, the multi-objective evolutionary optimization framework is outlined. Third, the methodology for calculation of the predicted performance using MEPDG is explained. And finally, the multiple considered objective functions and their corresponding calculations are clarified.

PROGRAMMING SINGLE-OBJECTIVE CALIBRATION

The plan for task 3 of this research project was to use the available AASHTOWare® Pavement ME Design software to conduct a conventional single-objective calibration according to the AASHTO guidelines, the results of which would serve as a comparison baseline. In this manner, this study will compare the multi-objective approach against a single-objective approach currently being used by the State agencies.

NCHRP 1-40B provides an 11-step procedure for verification, calibration, and validation of the MEPDG models for local conditions, which has been adopted by AASHTO.[4,3] At the seventh step, the significance of the bias (the average difference between predicted and measured performance) is tested. If there is a significant bias in prediction of pavement performance measures, the first round of calibration is conducted at the eighth step to eliminate bias. During this step, the SSE is minimized by adjusting the β_{r1}, β_{GB}, and β_{SG} calibration factors.

At the ninth step, the STE (standard deviation of error among the calibration dataset) is evaluated by comparing it to the STE from the national global calibration. If there is a significant STE, the second round of calibration at the 10th step tries to reduce the STE by adjusting the β_{r2} and β_{r3} calibration factors. A final validation step checks for the reasonableness of performance predictions.

The AASHTO guidelines provide regression equations for calculating the STE for each sublayer (AC, granular layers, and fine-grained layers). These equations were initially developed, not based on trenching data, but based on a "systematic approach" of assigning a fraction of the total measured rutting to each layer. During the NCHRP Project 09-30A, some trenching experiments were conducted.[31] However, the STE equations for each layer have not actually been updated accordingly. Only the STE regression equation for the total pavement rutting was updated following that study. In the absence of trenching data, only the STE in calculating the total pavement rutting is compared to the STE from the global calibration, which was about 0.1 inch.

Global heuristic optimization methods such as EAs could possibly identify a more optimum set of calibration coefficients compared to the local exhaustive search methods. That is why in this project an EA is used for each step of the required single-objective optimization to calibrate the rutting models for new and overlaid pavements.

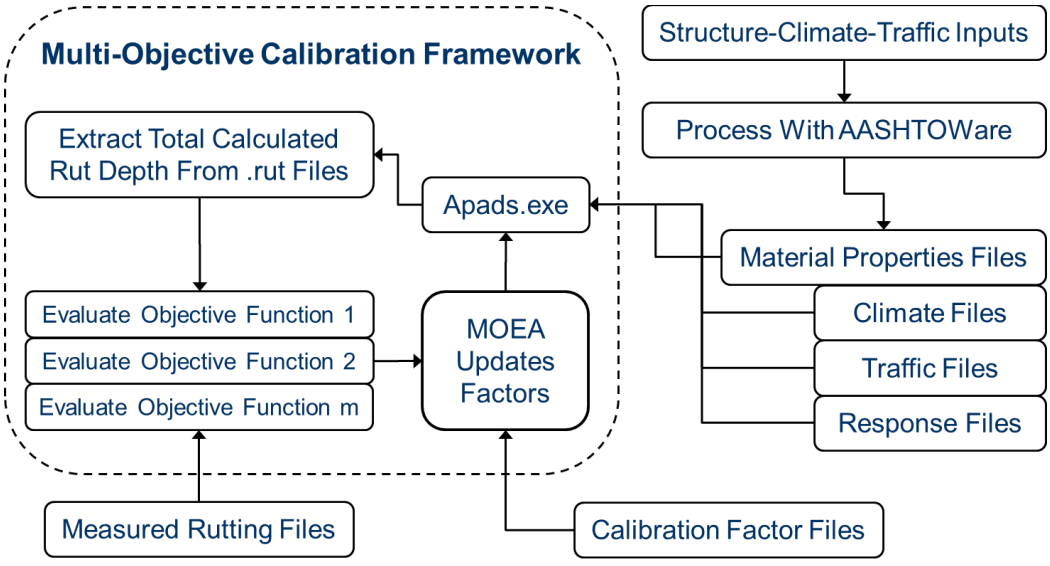
The single-objective calibration process is used in calibrating rutting models for new AC pavements using LTPP SPS-1 Florida site data and overlaid AC pavements using LTPP SPS-5 Florida site data.

PROGRAMMING MULTI-OBJECTIVE OPTIMIZATION FRAMEWORK

In this research project, MOEAs have been implemented to optimize the multiple objective functions involved. Using a multi-objective optimization approach allows escaping preconception, avoiding excessive concentration on only one aspect of the problem, and combining multiple sources of information in an objective manner.

MOEA Framework (moeaframework.org) is a free and open-source Java framework for multi-objective optimization using a variety of EAs, including GAs and ESs. In this object-oriented framework, an instance of the "abstract problem" class needs to be created, in which the calculation process for the multiple objective functions is implemented. Then an instance of the "problem execution" class is created, where the abstract multi-objective optimization problem is solved using a selected "algorithm."

At the first round of programming for this project, a Java code was developed to integrate the performance model engine of the AASHTOWare® Pavement ME Design software and the multi-objective EA into a calibration software as indicated in figure 15. After creating all the input data files for all the calibration test sections, the ME Design software was executed once to generate the materials properties, climate, traffic, and pavement response files. The framework runs the Asphalt Pavement Analysis and Design System (APADS) software using these files and the updated calibration factors at each iteration to calculate the MEPDG rutting values.(1) A comparison of these calculated values to LTPP measured values is used to evaluate the multiple objective functions (error functions) for each of the calibration factor sets. The MOEA then updates the solution population members (calibration factor sets) according to the results of objective function evaluations.



Source: FHWA.

Figure 15. Flowchart. Multi-objective calibration framework.

If improving final optimization solutions in terms of one objective degrades them in terms of other objectives, there is a conflict between the objectives; a clear example is the cost–performance tradeoff. When there is no conflict between multiple objectives, they can be combined using an engineered weighing scheme, and the problem would change into a single-objective optimization. The general goal of any multi-objective optimization is to identify the Pareto-optimal tradeoffs between multiple objectives. It is called a Pareto-optimal tradeoff because the best compromise among the multiple conflicting objectives is sought, and it was defined by Vilfredo Pareto (1896).[55] A Pareto-optimal solution has at least one better objective value compared to any other feasible solution in the decision space, while performing as well or worse in all remaining objectives. A solution X1 is called dominated if and only if another feasible solution X2 performs better than X1 in terms of at least one objective and as well as X1 in terms of others. A set of Pareto-optimal solutions is called the nondominated or Pareto-optimal front.

EAs have good global search ability, are less dependent on seed values, and do not require the mathematical formula to take the derivative of the objective functions.[56] In addition, EAs seem more suitable for multi-objective optimization because they are population based (they find several members of the Pareto-optimal front in a single run of the algorithm instead of having to perform a series of separate runs) and are less sensitive to the shape and discontinuities in the Pareto-optimal front.[33] A multi-objective GA called epsilon-dominance nondominated sorted GA (NSGA)II (ε-NSGAII) is implemented for multi-objective calibration.[57] ε-NSGAII has performed better in terms of effectiveness, efficiency, and reliability compared to other evolutionary multi-objective algorithms on complicated water resources applications.[57,58] It is also demonstrated to be easier to implement than traditional multi-objective EAs due to its simplified parameterization, adaptive population sizing, and automatic termination.[57]

The ε-NSGAII was designed based on an earlier successful second-generation multi-objective GA, the NSGAII, while the first-generation algorithm was called NSGA.[59,60] NSGA implemented the nondomination sorting approach recommended by Goldberg.[61] Fitness is assigned based on the level of domination (number of solutions that dominate the solution being evaluated) and similarity of solutions (i.e., fitness sharing). NSGA's front-based fitness assignment ensures that the solutions are found along the full extent of the Pareto front. NSGAII improved upon NSGA by implementing a more efficient nondomination sorting scheme, eliminating the need to specify the sharing parameter, and adding elitism and crowded tournament selection.[59] Following real-parameter simulated binary crossover (mating) and polynomial mutation, the *N* best individuals are selected from the combination pool of parent and child populations, preserving the elite population members. The two-step crowded tournament selection favors the individuals with lower rank (dominated by fewer other solutions), and if two solutions share the same rank, the solution with larger crowding distance is preferred (crowding distance is the largest cuboid surrounding the solution, in which no other solutions are present). NSGAII's extensive problem-specific parameter calibration was minimized in ε-NSGAII using epsilon-dominance archiving, adaptive population sizing, and automatic termination.[57]

Taking advantage of the epsilon-dominance concept, the user can specify the precision to quantify each objective. In the first step, the search space of the problem is divided into grid blocks, each block having a width of epsilon based on the specified precision. According to the developers of the algorithm, "Larger ε values result in a coarser grid (and ultimately fewer

solutions) while smaller values produce a finer grid."[57] Epsilon can be viewed as publishable precision or error tolerance determined to avoid wasting computational resources on unjustifiably precise results.[62] If there are multiple solutions in a grid block, only the solution closest to the lower left-hand corner of the block is saved (assuming minimization of all objectives). Step 2 comprises nondomination sorting based on the grid blocks. In this step, each grid block is used as a reference, and any other solution grids to the top and right side of the reference grid will be eliminated.[57] Step 3 is called "thinning of solutions," where dominated solutions are eliminated. As a result, a more even search of the objective space is encouraged.

The ε-NSGAII starts with exploiting small populations to "precondition" search, and then it automatically adapts population size according to problem difficulty and explores further areas of the solution space.[57] This is carried out through a series of "connected runs." An offline archive is used to store epsilon-nondominated solutions found after each generation, which are subsequently used to direct the search in the next run. While the search is directed using previously evolved solutions (25 percent elitism among runs), adding new random solutions (75 percent) encourages the exploration of additional regions of the search space.[57] If the number and quality of nondominated solutions does not increase above a minimum threshold (delta parameter) between two successive runs, the algorithm is automatically terminated across all populations. The initial use of smaller population sizes, the elimination of random seed analysis, and the elimination of trial-and-error application runs to determine search parameters are all contributing to lower computational cost.

The probability of mating and mutation are assumed to be 1 and 1/*n*, respectively, with *n* being the number of unknown parameters to calculate (number of calibration factors in this case). The initial population size is assumed 10, and the maximum number of generations per run is 250. The search is terminated after 10,000 function evaluations unless terminated due to the delta parameter of 10 percent beforehand.[58]

MEPDG PERFORMANCE PREDICTION FOR FUNCTION EVALUATION

After executing the preliminary runs of the developed program, it was found that each evaluation using the APADS software takes about 4 min to complete. This is because APADS conducts both the pavement analysis to calculate responses and then the calculation of pavement performance. However, pavement responses do not change by changing the calibration factors in consecutive iterations, and there is no need to conduct pavement analysis at every iteration. There was no option to calculate only the pavement performance using the responses. While the APADS software was launched simultaneously for all the LTPP test sections used in the calibration database, each optimization iteration was taking about 30 min to evaluate all the solution population members (calibration factor sets) on all the calibration data (using a computer with Intel Core i7-4600 2.1 GHz and 8 GB of RAM). Since the complete calibration could take hundreds of iterations, it was not feasible to continue using the APADS software in this framework. Therefore, it was decided to simulate the Pavement ME software to replace the APADS routine in this framework.

As explained above, due to the long computational time of the APADS analysis software, it was decided to use the MEPDG equations to calculate pavement performance based on the generated materials properties, climate, traffic, and pavement response data files. The calculation of total

pavement deformation using these equations was embedded in the code for objective function evaluation, which involves calculation of the difference between measured and calculated rutting. The following sections explain how these calculations were conducted to simulate the performance predictions in the Pavement ME software.

Total Pavement Permanent Deformation

The total rutting is the summation of each layer permanent deformation. Pavement ME software calculates the monthly incremental pavement total deformation for the design life. The expression to calculate total pavement deformation when the pavement structure is composed of HMA layer(s) and granular base (GB) layer(s), on top of a fine-grained subgrade, is in equation 34:

$$RD = RD_{HMA} + RD_{GB} + RD_{SG} = \varepsilon_{p(HMA)} h_{HMA} + \delta_{a(GB)}(N) + \delta_{a(SG)}(N) \quad (34)$$

Where:

RD = total pavement permanent deformation.
RD_{HMA} = HMA layer permanent deformation.
RD_{GB} = GB permanent deformation.
RD_{SG} = fine-grained subgrade permanent deformation.
h_{HMA} = thickness of the HMA (inches).
$\varepsilon_{p(HMA)}$ = permanent strain of the HMA layer.
$\delta_{a(GB)}(N)$ = permanent or plastic deformation for granular material.
$\delta_{a(SG)}(N)$ = permanent or plastic deformation for fine-grained materials.

The initial round of rutting calculation with global calibration factors (local factors set to 1.0) was done using the Pavement ME software. Subsequent rutting calculation within the optimization framework was done using a simulated approach based on the MEPDG equations to save computational time. This process can be done efficiently, since pavement temperature, number of axle load repetitions, and elastic strain do not need to be recalculated for each new set of local calibration factors. The AASHTOWare® output files that provide the data for the global rutting calculation are described in table 25.

Table 25. AASHTOWare® Pavement ME Design software data files.

Data Item	Software File	Data Field/Value
Layer thickness	input.tmp	Material thickness
Water content	input.tmp	Initial water content
Local and global calibration factors	calibrationfactor.dat	HMA rutting (columns 4, 5, 6), subgrade rutting (columns 1, 3)
Cumulative traffic per month	TruckGrowth.csv	Column 2
Noncumulative number of trucks in each class per month	TruckGrowthByClass.csv	For each class
Hourly temperature	thermal.tmp	*Column#* = *Round*($0.3681h_{ac}$ + 2.2857, 0)
Vertical strain in each subseason under each axle type and at different horizontal locations under the load	_VertStrain.txt	Strain
Rutting per layer	rut.tmp	AC1, GB2, SG3, Total

Most of the data files generated by the Pavement ME software include data at a monthly frequency, except for the thermal.tmp file, which includes hourly pavement layer temperature values. It appears that the APADS software is calculating rutting values on a submonthly basis. Each "subseason" is one-fifth of the month, which includes one-fifth of the total monthly traffic and an average temperature for a 20-percent percentile of a normal distribution generated based on hourly temperature data in that month. A normal distribution histogram is generated based on hourly temperature data in each month, and the area under the histogram is divided into five identical slices. Each slice indicates one subseason, and the centroid of each slice is used as the average temperature for that subseason. The software then cumulates the calculated rut depth values for the five subseasons in the month to calculate the monthly rutting. The _VertStrain file includes vertical strain in each subseason, under each axle type, and at different horizontal locations under the load. However, it is not clear which depth (layer) of the pavement structure these strain values are calculated for. Since the available data files do not contain adequate details on the intermediate pavement response data, it was decided to simulate the software output instead of using the documented process in the *AASHTO Manual of Practice*.[51]

Note: The simulation process explained here onward is not the same procedure that the AASHTOWare® Pavement ME Design software is using to calculate rut depth (according to the *AASHTO Manual of Practice*). The following sections explain the simulation methodology for calculating rutting in asphalt bound and unbound pavement layers.

Simulating Permanent Deformation in Asphalt Concrete Layers

For better accuracy in simulating the ME software calculations, the permanent deformation in the AC layers was assumed to be calculated on an hourly basis instead of a subseason basis. The general shape of the rutting prediction model was used to conduct this simulation, assuming rutting and traffic to be cumulative values (note that in the original MEPDG equation, this is only true if the temperature was constant in this period, as it is in laboratory tests). The hourly temperature in each pavement layer is available in the thermal.tmp file. The available cumulative monthly traffic and calculated global rutting prediction were transformed linearly to cumulative hourly values (which is again an assumption used for this simulation and not necessarily true). After this transformation, the hourly strain is backcalculated with equation 35 for the global model with local calibration factors equal to 1. This is because the available monthly maximum vertical strain data (_VertStrain.txt) for each axle type (single, tandem, tridem, quad) could not be attributed to a specific pavement layer and were not readily transformed to hourly values. As noted previously, an hourly period is used to achieve better accuracy in simulating the software results.

$$\varepsilon_{r_j} = \frac{RD_{g,j}}{k_Z 10^{k_1} h_{HMA} T_j^{k_2} N_j^{k_3}} \quad (35)$$

Where:

ε_{r_j} = backcalculated resilient or elastic strain for hour *j*.

$RD_{g,j}$ = HMA permanent deformation calculated using MEPDG software for global calibration factors (*g* stands for global); the total amount of rutting accumulated up to hour *j* in each month *i* is calculated (assuming linear increase in rutting within each month) using equation 36:

$$RD_{g,j} = RD_{g,i-1} + \left(\frac{RD_{g,i} - RD_{g,i-1}}{24d_i}\right) H_j \quad (36)$$

Where:

i = month.

j = hour.

d_i = total number of days in the month *i*.

H_j = number of hours from the beginning of each month up to hour *j*.

For *D* in equation 2, D_m is used, which is the depth below the surface for each HMA layer *m* calculated using equation 37:

$$D_m = \frac{h_{HMA,m}}{2} + \sum_{i=1}^{m-1} h_{HMA,i} \quad (37)$$

Where:

m = HMA layer number from top to bottom.

k_1, k_2, k_3 = global field calibration parameters.[6]

k_1 = –3.35412, k_2 = 1.5606, k_3 = 0.4791.

h_{HMA} = thickness of the HMA sublayer/layer (inches).

T_j = layer temperature for hour *j* from thermal.tmp file.

N_j = number of axle load repetitions up to hour *j* (assuming linear increase in traffic within each month) calculated using equation 38:

$$N_j = N_{i-1} + \left(\frac{N_i - N_{i-1}}{24d_i}\right) H_j \quad (38)$$

The subsequent monthly rutting calculations for each month *i* and any different set of calibration factors can be done using the above backcalculated hourly strain values in equation 39:

$$RD_{HMA,i} = MAX_{j=1}^{24d_i} \left(\beta_{r1}\, k_Z 10^{k_1}\, h_{HMA}\, T_j^{k_2\beta_{r2}} \varepsilon_{r_j} N_j^{k_3\beta_{r3}}\right) \quad (39)$$

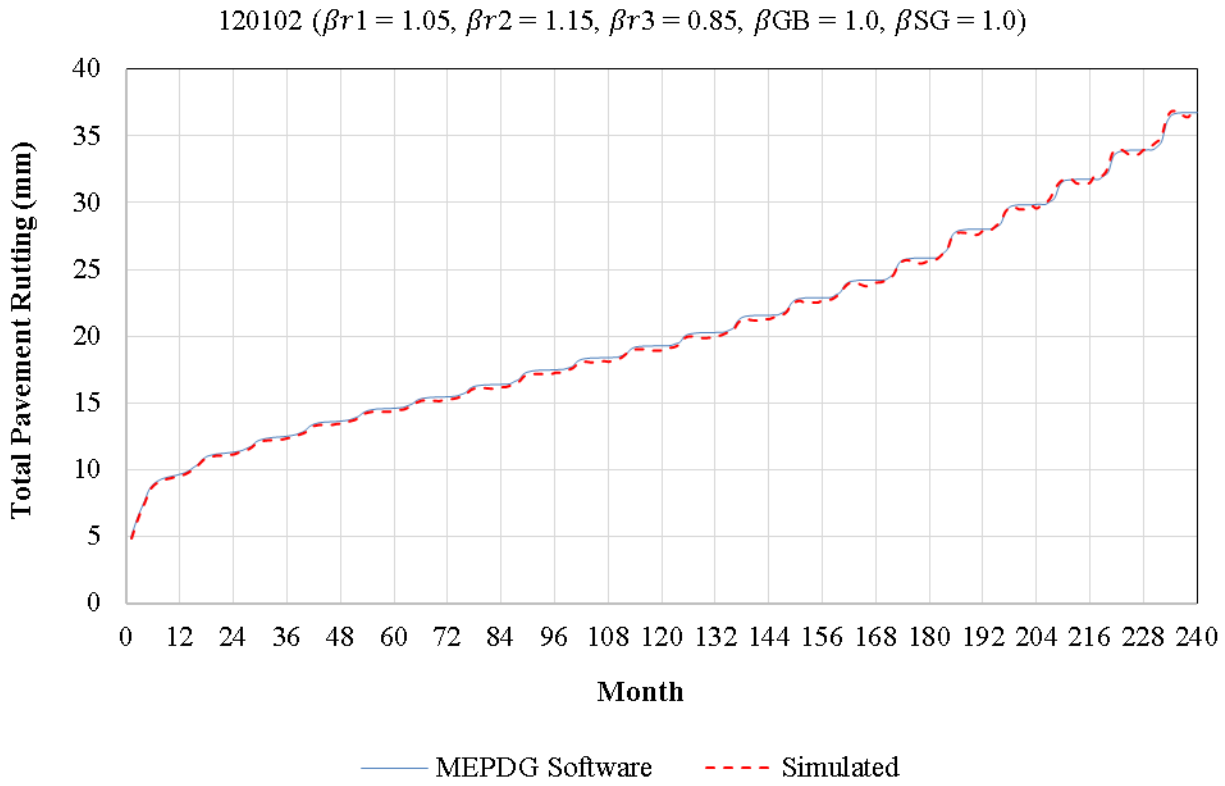
Where:

$RD_{HMA,i}$ = total HMA permanent deformation accumulated up to month *i*.

$\beta_{r1}, \beta_{r2}, \beta_{r3}$ = local or mixture field calibration factors.

This approach was tested through a comparison of the Pavement ME software results with various pavement structures and different local calibration factors. Examples for the comparison of this simulated calculation to the ME software can be found in appendix C. It should be noted that the same approach was used for ATB layers. Figure 16 shows an example comparison of the simulated total pavement rutting to the Pavement ME software output. As evident in this figure

and similar figures in appendix C, the simulated calculation and the software output follow a similar trend. While there are some intermediate decrements in rutting values simulated within each month, the total accumulated rutting at the end of each month is very close to the software output. The reason for the intermediate decrements (which do not comply with the theory of rutting accumulation) is the assumptions made for the simulation, which may not be inherently correct. One of those assumptions is that traffic and rutting values are increasing linearly within each month, which might not be true. As explained before, the actual subseason pavement response data for each layer are not provided by the AASHTOWare® software, and therefore, an exact calculation (according to the MEPDG equations) could not be conducted for this project.



Source: FHWA.

Figure 16. Chart. Example comparison of the simulated calculation to Pavement ME software output (on SPS-1 test section 120102).

Simulating Permanent Deformation in Unbound Materials

A similar approach was applied for the coarse- and fine-grained unbound layers. The relationship between the calculated elastic strain and the laboratory resilient strain is backcalculated for the global model using equation 40:

$$\varepsilon_{p,j} = \varepsilon_{v,j}\left(\frac{\varepsilon_{0,\mathrm{j}}}{\varepsilon_{r,j}}\right) = \frac{\delta_{g,j}}{k_1\, h_{ub}\, \left| e^{-\left(\frac{\rho}{N_j}\right)^{\beta}} \right|} \tag{40}$$

Where:

ε_p = permanent vertical strain.

ε_v = vertical resilient or elastic strain in the layer calculated by the structural response model.

$\delta_{g,j}$ = permanent or plastic deformation for granular and fine-grained materials calculated by the Pavement ME software for the global calibration factors; the total amount of rutting accumulated up to hour j in each month i is calculated using equation 41:

$$\delta_{g,j} = \delta_{a_{i-1}} + \left(\frac{\delta_{a_i} - \delta_{a_{i-1}}}{24d}\right) H_j \tag{41}$$

Where:

k_1 = global calibration coefficient.

k_1 = 2.03 for granular materials.

k_1 = 1.35 for fine-grained materials.

h_{ub} = thickness of the unbound layer/sublayer (inches).

N_j = number of axle load repetitions up to hour j.

Knowing the initial relationship among permanent and resilient strains and then rutting in each unbounded layer with any set of local calibration factors is calculated using equation 42:

$$\delta_{ub,j} = \beta_{s1} k_1 h_{ub}\ \varepsilon_{p,j}\ \left| e^{-\left(\frac{\rho}{N_j}\right)^{\beta}} \right| \tag{42}$$

Where:

$\delta_{ub,,j}$ = permanent or plastic deformation for granular and fine-grained materials accumulated up to hour j.

β_{s1} = local calibration factor.

CALCULATION OF THE MULTIPLE OBJECTIVE FUNCTIONS

In this section, the concept and equations for calculation of each objective function have been outlined. These objective functions quantify the multiple sources of information that could be combined to enhance model calibration. These multiple objective functions will be optimized simultaneously to determine the calibration coefficients that provide a tradeoff among the various objectives.

Several scenarios can be devised for multi-objective formulation of calibration, all of which could overcome cognitive challenges and add to our knowledge of this problem. More than one set of multiple objectives will be considered to explore new aspects of the calibration problem.

The following are the identified approaches and the multiple objective functions that need to be simultaneously minimized for each scenario:

- Scenario 1: Optimizing major statistical outcomes.
  i. SSE, which represents bias (average error) that is an estimate of model accuracy.
  ii. Standard deviation of error (or STE), which represents variation in error that is an estimate of model precision.
- Scenario 2: Combining different sources of data.
  i. SSE on LTPP data.
  ii. STE on LTPP data.
  iii. SSE on FDOT APT data.
  iv. STE on FDOT APT data.

In the first alternative scenario, mean and standard deviation of prediction error are simultaneously minimized to reduce the bias and STE (increase model accuracy and precision) at the same time. In this manner, the information from a single calibration run is fully implemented, and an additional round of computationally intensive calibration is avoided. This two-objective optimization will be separately executed in calibrating rutting models for new AC pavements using LTPP SPS-1 Florida site data and overlaid AC pavements using LTPP SPS-5 Florida site data.

In the second scenario, the SSE and STE in predicting permanent deformation of pavements within different performance data sources will be used as separate objective functions to be minimized simultaneously. This four-objective optimization will be used in calibrating rutting models for new AC pavements using the LTPP SPS-1 Florida site data and the FDOT APT data at the same time. This scenario will comprise an objective approach to incorporate different sources of data. This scenario was not applied to overlaid pavements because the available APT data were only for new pavement structures.

As explained before, when there is no conflict between multiple objectives, they can be combined using an engineered weighing scheme, and the problem would change into a single-objective optimization. However, in the case of the selected objective functions in this study, the existence of a conflict cannot be proven theoretically. In the case of the second scenario, the sources of data for multiple objective functions are different, and the errors on LTPP and APT data may or may not be in conflict. In the case of the first scenario, equations 43 and 44 are the expanded expressions for calculating SSE and STE, which are to be minimized.

$$SSE = \sum_{i=1}^{n} (Rp_i - Rm_i)^2 = \sum_{i=1}^{n} \left({Rp_i}^2 + {Rm_i}^2 - 2Rp_i Rm_i \right) \tag{43}$$

$$STE = \left(\frac{\sum_{i=1}^{n} \left((Rp_i - Rm_i) - \overline{(Rp - Rm)} \right)^2}{n-1} \right)^{\frac{1}{2}} \tag{44}$$

Where:

n = the number of rutting data records.

Rp_i = the predicted (for a specific pavement structure at a specific time) rutting record i.

Rm_i = the corresponding measured rutting record i.

Equation 45 is an expanded expression for STE from equation 44.

$$STE = \left(\frac{\sum_{i=1}^{n} \left((Rp_i - Rm_i)^2 + \overline{(Rp - Rm)}^2 - 2(Rp_i - Rm_i)\overline{(Rp - Rm)} \right)}{n - 1} \right)^{\frac{1}{2}} \quad (45)$$

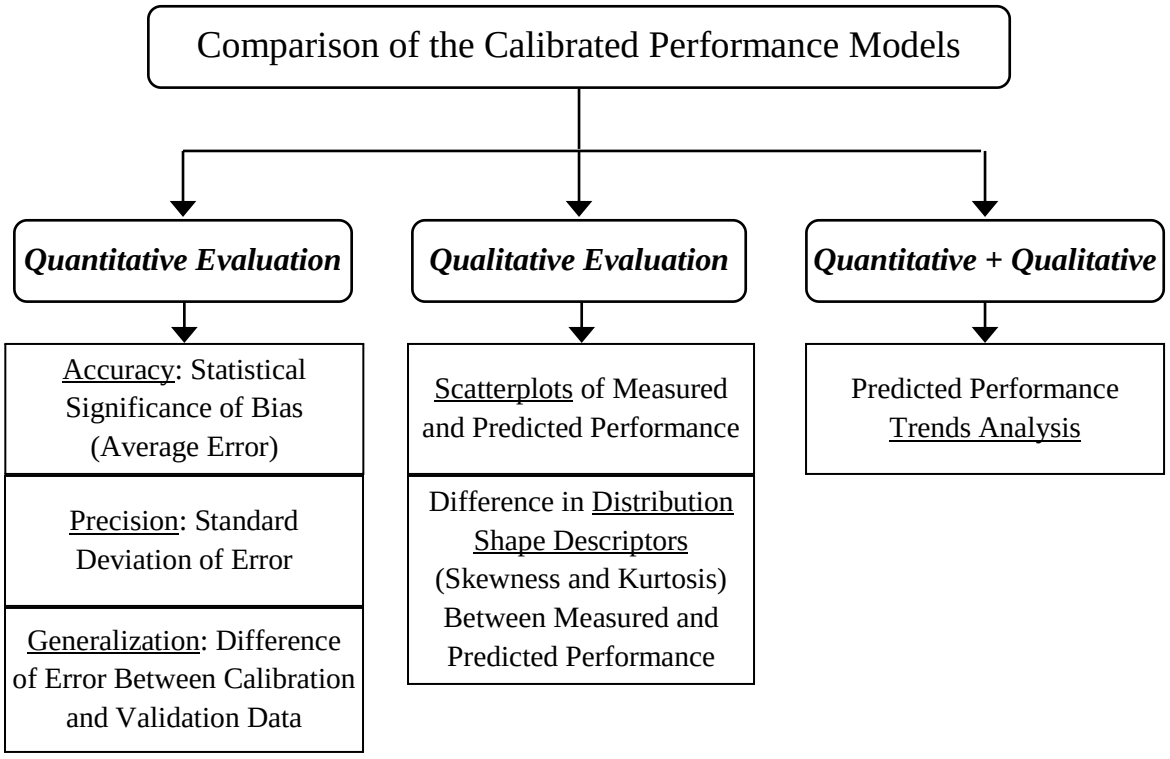
For different sets of calibration factors, the SSE and STE (from equations 43 and 45) could be increased or decreased either in the same or the opposite direction of each other. A set of calibration factors that results in the minimum SSE does not necessarily guarantee a minimum STE. This means that a calibrated model that exhibits higher accuracy (lower SSE) might not necessarily have higher precision (lower STE) as well. As it cannot be proven whether the selected objective functions are in conflict, they cannot be combined with the goal of simplifying the problem into a single-objective optimization. Therefore, a multi-objective optimization approach was used.

CHAPTER 5. COMPARISON OF MULTI-OBJECTIVE TO SINGLE-OBJECTIVE CALIBRATION RESULTS

INTRODUCTION

A comprehensive framework was devised for comparison of the multi-objective calibration results to the AASHTO recommended single-objective calibration. This comparison framework is based on quantitative and qualitative evaluation of the calibrated models.[63] Figure 17 shows the different steps within the comparison framework.

The measured rutting databases for both LTPP and APT were randomly divided into 80 percent for calibration and 20 percent for validation. Following the calibration of the models using the calibration data, the calibrated prediction models were tested using the validation dataset.



Source: FHWA.

Figure 17. Flowchart. Framework for comparison of the calibrated performance models.

For quantitative evaluation, accuracy, precision, and generalization capability of the models are assessed using the calibration and validation datasets. To evaluate the accuracy of the multi-objective approach compared to the conventional single-objective calibration, statistical significance testing is required to determine whether the bias between measured and predicted performance values is statistically significant before and after each calibration process.

The bias (average error) is an estimate of the model accuracy, and the STE represents the precision of the performance model. According to the AASHTO recommended calibration

guidelines, the STE of the calibrated model is compared to the STE from the national global calibration to determine its significance.

Once the models are calibrated, they are validated on a validation dataset to check for the reasonableness of performance predictions. This step also provides a measure of the generalization capability (repeatability) of the calibrated models. The model that generalizes better will be the model that matches measured pavement performance equally well on both the validation and the calibration datasets.

For qualitative evaluation of the calibrated models, scatterplots of measured versus predicted performance on the calibration and validation datasets were used. Another qualitative assessment could be a sensitivity analysis of the calibrated models to the changes in input variables, which is a substantial task and beyond the scope of the current project. The reasonableness of this model behavior can be compared between the single-objective and multi-objective calibration methods. An additional qualitative evaluation was conducted by comparing the shape of the distribution of measured versus predicted pavement performance. Statistical distribution shape descriptors such as nonparametric skewness and kurtosis can be used for this purpose.

Examining the predicted rate of change in performance indices compared to measured deterioration trends, a combination of quantitative and qualitative evaluations was implemented to compare different calibrated models.

This comparison of the multi-objective approach to the conventional single-objective method is demonstrated in calibrating rutting models for new AC pavements using LTPP SPS-1 Florida site data, new AC pavements using FDOT APT data, and overlaid AC pavements using LTPP SPS-5 Florida site data.

SINGLE-OBJECTIVE CALIBRATION RESULTS

A single-objective calibration according to the AASHTO guidelines was conducted, the results of which serve as a comparison baseline. The AASHTO recommended approach includes an 11-step procedure for "verification," "calibration," and "validation" of the MEPDG models for local conditions. The verification involves an examination of accuracy and precision of the global (nationally calibrated) model on the local dataset.

Table 26 and table 27 list the important statistics regarding this verification of the rutting models for new pavements using Florida SPS-1 site data and overlaid pavements using Florida SPS-5 site data, respectively. As expected, the results show a significant positive bias and a high standard deviation of error because the global model was calibrated based on all LTPP data from across North America. The next steps will demonstrate the AASHTO recommended procedure for "eliminating" the bias and potentially reducing the STE.

Table 26. "Verification" of the global rutting model for new pavements on Florida SPS-1.

Statistic	Calibration Data	Validation Data	Combined Data
Data records	128	26	154
SSE	11,135.13 mm^2 (17.26 $inch^2$)	2,396.75 mm^2 (3.72 $inch^2$)	13,531.88 mm^2 (20.97 $inch^2$)
RMSE	9.33 mm (0.367 inch)	9.60 mm (0.379 inch)	9.37 mm (0.370 inch)
Bias	+8.05 mm (0.317 inch)	+8.34 mm (0.328 inch)	+8.10 mm (0.319 inch)
p value (paired *t*-test) for bias	5.11E – 39 < 0.05; significant bias	1.28E – 08 < 0.05; significant bias	5.51E – 47 < 0.05; significant bias
STE	4.73 mm (0.186 inch)	4.85 mm (0.191 inch)	4.73 mm (0.186 inch)
Generalization capability	N/A	N/A	96.4%
R^2 (goodness of fit)	0.0391	0.123	0.0487

RMSE = root-mean-squared error; generalization capability = 100 – normalized difference in bias between the calibration and validation datasets. N/A = not applicable.

Table 27. "Verification" of the global rutting model for overlaid pavements on Florida SPS-5.

Statistic	Calibration Data	Validation Data	Combined Data
Data records	189	39	228
SSE	62,207.04 mm^2 (96.42 $inch^2$)	11,607.67 mm^2 (17.99 $inch^2$)	73,814.71 mm^2 (114.41 $inch^2$)
RMSE	18.14 mm (0.714 inch)	17.25 mm (0.680 inch)	17.99 mm (0.709 inch)
Bias	+16.39 mm (0.645 inch)	+15.71 mm (0.618 inch)	+16.27 mm (0.640 inch)
p value (paired *t*-test) for bias	7.40E – 71 < 0.05; significant bias	1.18E– 15 < 0.05; significant bias	9.78E– 86 < 0.05; significant bias
STE	7.80 mm (0.311 inch)	7.22 mm (0.284 inch)	7.70 mm (0.304 inch)
Generalization capability	N/A	N/A	95.85%
R^2 (goodness of fit)	0.0609	0.0018	0.0504

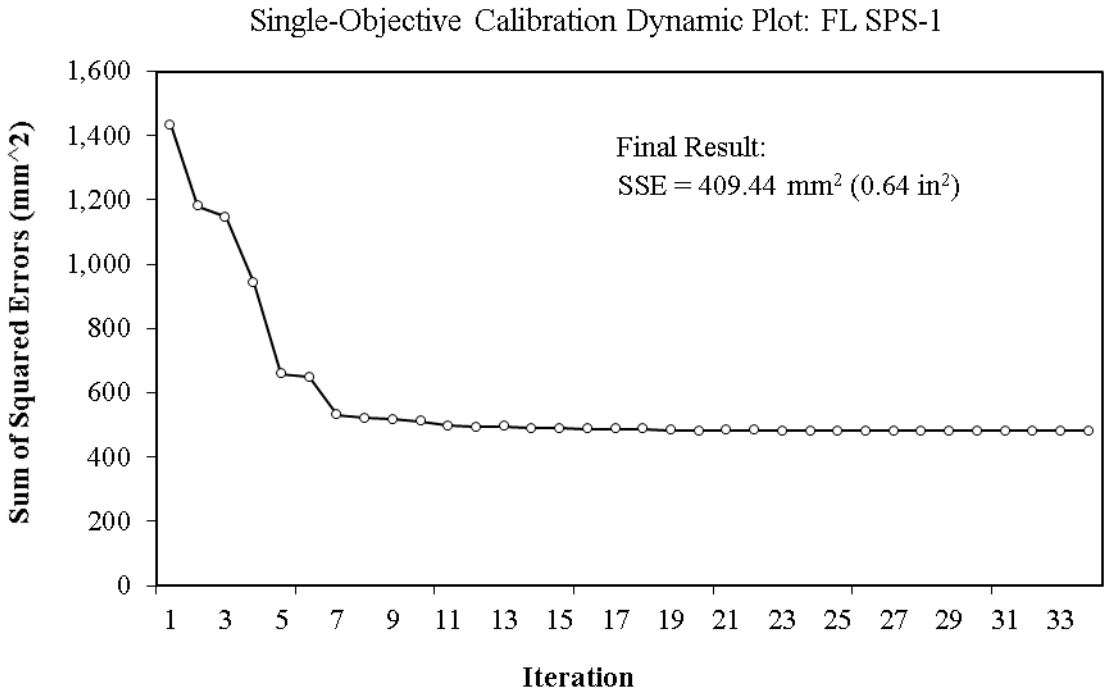
N/A = not applicable; RMSE = root-mean-squared error.

At the seventh step of the 11-step AASHTO recommended calibration procedure, the significance of the bias (the average difference between predicted and measured performance) is tested. If there is a significant bias in prediction of pavement performance measures, the first round of calibration is conducted at the eighth step to eliminate bias. During this step, the SSE is minimized by adjusting the β_{r1}, β_{GB}, and β_{SG} calibration factors.

At the ninth step, the STE (standard deviation of error among the calibration dataset) is evaluated by comparing it to the STE from the national global calibration, which was about 0.1 inch. If there is a significant STE, the second round of calibration at the tenth step tries to reduce the STE by adjusting the β_{r2} and β_{r3} calibration factors. A final validation step checks for the reasonableness of performance predictions on the validation dataset that has not been used for model calibration.

Global heuristic optimization methods such as EAs could possibly identify a more optimum set of calibration coefficients compared to the local exhaustive search methods. That is why in this project GA was used for single-objective optimization in calibrating rutting models for new AC pavements using the LTPP SPS-1 Florida site data and overlaid AC pavements using the LTPP SPS-5 Florida site data.

Figure 18 shows a dynamic plot of SSE as the optimization iterations pass by for single-objective calibration on Florida SPS-1 data. The optimization is stopped when the SSE for the best member of the population does not change more than 1 mm^2 for at least 10 consecutive generations (in this case, after 34 iterations).



Source: FHWA.

Figure 18. Chart. Dynamic plot of SSE in single-objective optimization on Florida SPS-1 data.

Table 28 lists the important statistics regarding these single-objective calibration results of rutting models for new pavements on Florida SPS-1. The final results of the single-objective minimization of SSE demonstrate a *p* value (0.096) greater than 0.05 for a paired *t*-test of measured versus predicted total rut depth on calibration data. Therefore, there is not enough evidence to reject the null hypothesis of the bias being insignificant. It seems that the continuation of this optimization process will not significantly improve the results (figure 18). On the other hand, the *p* value is much higher (0.47) for the validation dataset, and therefore the bias seems to be even less significant on the validation dataset, which was not used for model calibration. The negative bias indicates that the calibrated model underpredicts rut depth values on average by 0.25 mm, which is insignificant compared to the accuracy of manual rut depth measurements that is 1 mm in the LTPP program.

There seems to be a lack of precision of the calibrated model demonstrated by the high amount of scatter in figure 19. However, since the overall STE is lower than the national global calibration results (0.068 inch < 0.129 inch), according to AASHTO calibration guidelines, there is no need for the second round of optimization to reduce STE. Therefore, the following are the selected calibration factors:

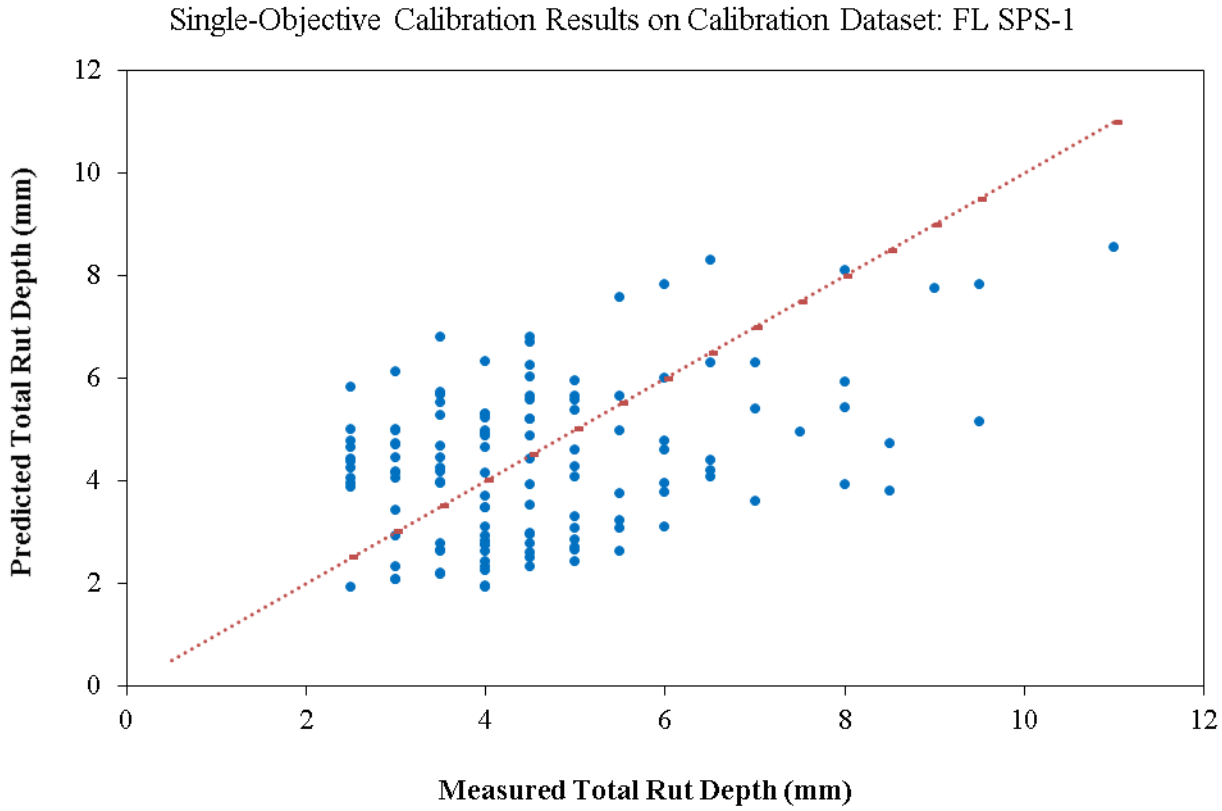
β_{r1} = 0.522, β_{r2} = 1.0, β_{r3} = 1.0, β_{GB} = 0.011, and β_{SG} = 0.171.

Table 28. Single-objective calibration results of rutting models for new pavements on Florida SPS-1.

Statistic	Calibration Data	Validation Data	Combined Data
Data records	128	26	154
SSE	409.44 mm^2 (0.64 inch^2)	59.70 mm^2 (0.09 inch^2)	469.14 mm^2 (0.73 inch^2)
RMSE	1.79 mm (0.071 inch)	1.52 mm (0.060 inch)	1.75 mm (0.069 inch)
Bias	–0.26 mm (–0.010 inch)	–0.20 mm (–0.008 inch)	–0.25 mm (–0.010 inch)
p value (paired *t*-test) for bias	0.096 > 0.05; insignificant bias	0.47 > 0.05; insignificant bias	0.072 > 0.05; insignificant bias
STE	1.78 mm (0.070 inch)	1.53 mm (0.060 inch)	1.73 mm (0.068 inch)
Generalization capability	N/A	N/A	70%
R^2 (goodness of fit)	0.1519	0.2553	0.165

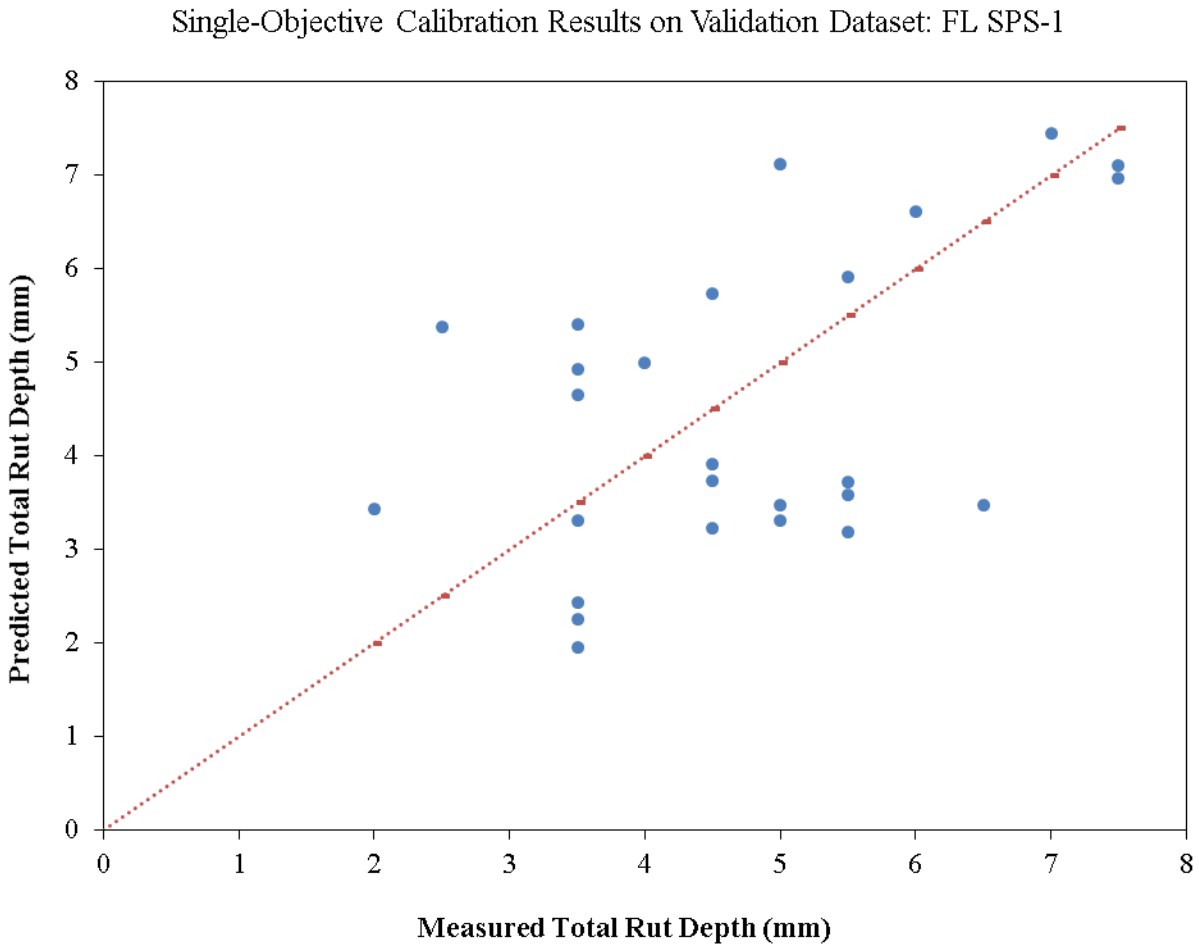
N/A = not applicable; RMSE = root-mean-squared error.

Figure 19 and figure 20 show scatterplots of the measured versus predicted total rut depth on calibration (128 records) and validation (26 records) datasets, respectively, for Florida SPS-1. The measured rut depth data are discrete values at every 1 mm (the plots show 0.5 mm because the rut depth values are averaged between the left and right wheelpaths), while the predicted rut depth data are continuous values. Even though the bias and the STE are relatively low on both the calibration and validation datasets, there is a significant amount of scatter in these plots, and the goodness-of-fit indicator (R^2) is poor. This scatter could perhaps be due to two reasons. First and foremost, the calibrated model does not exhibit adequate precision. This could be tracked back to the lack of precision of the global model. Second, there is a high degree of variation (perhaps due to construction quality and environmental variability) in measured rut depths along each pavement section (and between the two wheelpaths) at each distress survey, while the model can only produce one value (for 50 percent reliability) for each pavement section at each specified time. This could add to the scatter observed in these plots.



Source: FHWA.

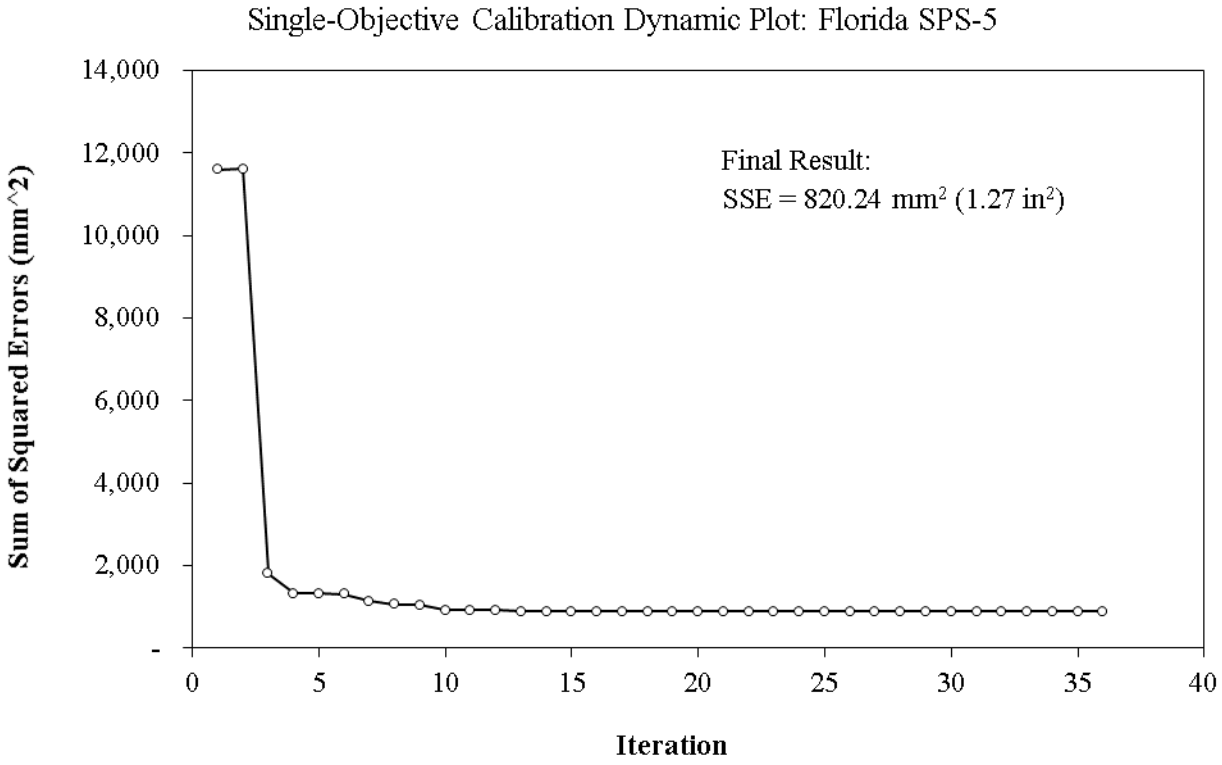
Figure 19. Scatterplot. Measured versus predicted single-objective calibration results of rutting models for new pavements on calibration dataset for Florida SPS-1.



Source: FHWA.

Figure 20. Scatterplot. Measured versus predicted single-objective calibration results of rutting models for new pavements on validation dataset for Florida SPS-1.

Figure 21 shows a dynamic plot of SSE as the optimization iterations pass by for single-objective calibration on Florida SPS-5 data. The optimization is stopped when the SSE for the best member of the population does not change more than 1 mm^2 for at least 10 consecutive generations (in this case, after 36 iterations).



Source: FHWA.

Figure 21. Chart. Dynamic plot of SSE in single-objective optimization on Florida SPS-5 data.

Table 29 lists the important statistics regarding these single-objective calibration results of rutting models for overlaid pavements on Florida SPS-5. The final results of the single-objective minimization of SSE demonstrate a low *p* value (0.0005) for a paired *t*-test of measured versus predicted total rut depth on calibration data. Therefore, the null hypothesis of the bias being insignificant has been rejected. However, it seems that the continuation of this optimization process will not significantly improve the results (figure 21). The *p* value is higher (0.013) for the validation dataset, but the bias is still significant on the validation dataset as well.

Table 29. Single-objective calibration results of rutting models for overlaid pavements on Florida SPS-5.

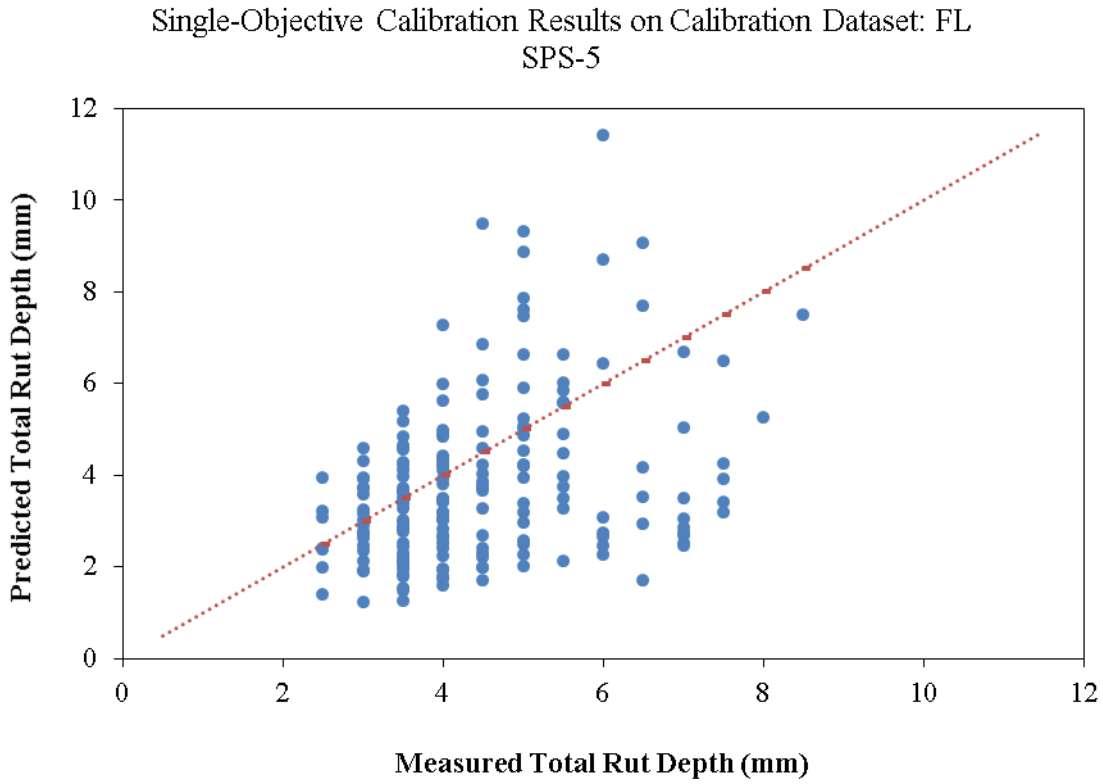
Statistic	Calibration Data	Validation Data	Combined Data
Data records	189	39	228
SSE	820.24 mm² (1.27 inch²)	128.37 mm² (0.20 inch²)	948.61 mm² (1.47 inch²)
RMSE	2.08 mm (0.082 inch)	1.81 mm (0.071 inch)	2.04 mm (0.080 inch)
Bias	–0.53 mm (–0.021 inch)	–0.72 mm (–0.028 inch)	–0.56 mm (–0.022 inch)
p value (paired *t*-test) for bias	0.0005 < 0.05; significant bias	0.013 < 0.05; significant bias	3.18E-05 < 0.05; significant bias
STE	2.02 mm (0.080 inch)	1.69 mm (0.066 inch)	1.97 mm (0.078 inch)
Generalization capability	N/A	N/A	64.15%
R^2 (goodness of fit)	0.1196	0.0213	0.1073

N/A = not applicable; RMSE = root-mean-squared error.

Since the overall STE is lower than the national global calibration results (0.078 inch < 0.129 inch), according to AASHTO calibration guidelines, there is no need for the second round of optimization to reduce STE. Therefore, the following are the selected calibration factors:

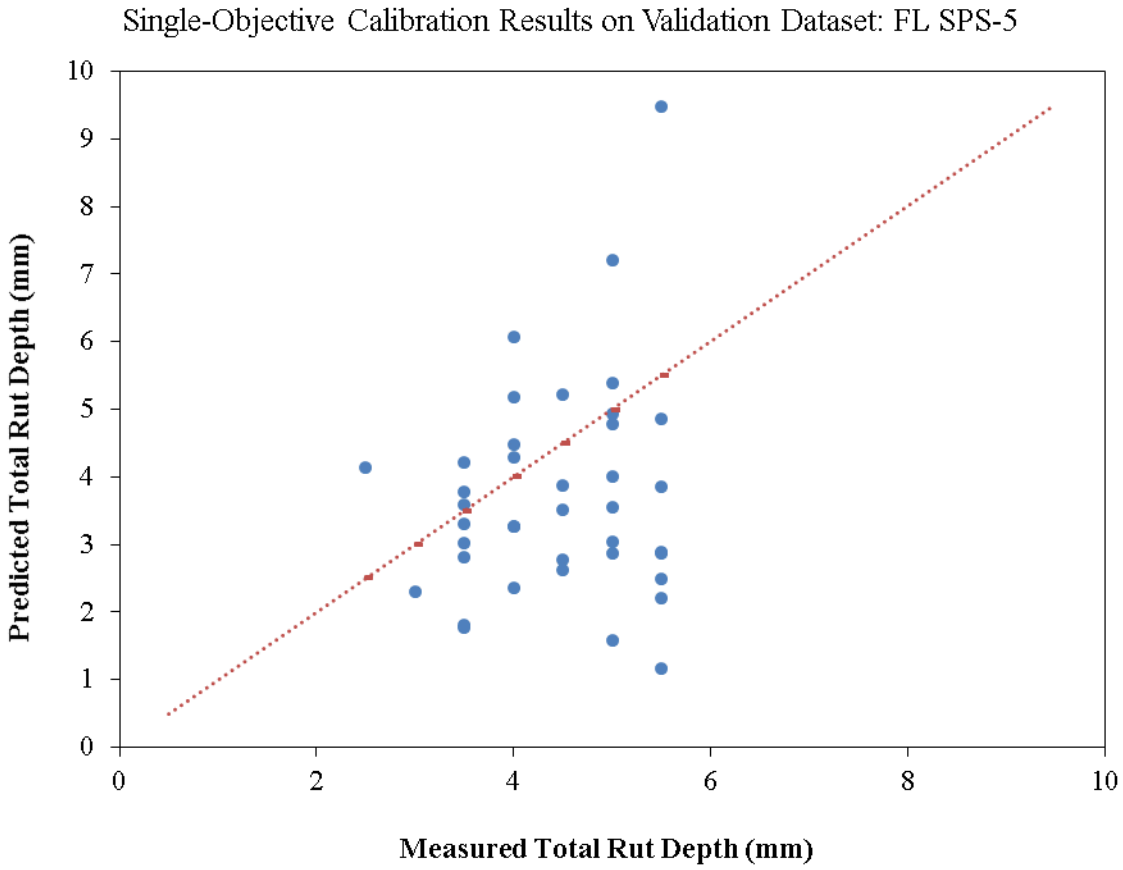
β_{r1} = 0.5004, β_{r2} = 1.0, β_{r3} = 1.0, β_{GB} = 0.0738, and β_{SG} = 0.1554.

Figure 22 and figure 23 show scatterplots of the measured versus predicted total rut depth on calibration (189 records) and validation (39 records) datasets, respectively, for Florida SPS-5. Similar to the results on SPS-1 data, there is a significant amount of scatter in these plots, and the goodness-of-fit indicator (R^2) is poor. The negative bias indicates that the calibrated model underpredicts rut depth values on average by 0.56 mm. While this bias is statistically significant, it is low compared to the accuracy of manual rut depth measurements, which is 1 mm in the LTPP program.



Source: FHWA.

Figure 22. Scatterplot. Measured versus predicted single-objective calibration results of rutting models for overlaid pavements on calibration dataset for Florida SPS-5.



Source: FHWA.

Figure 23. Scatterplot. Measured versus predicted single-objective calibration results of rutting models for overlaid pavements on validation dataset for Florida SPS-5.

MULTI-OBJECTIVE CALIBRATION RESULTS

As discussed in chapter 4, several scenarios can be devised for multi-objective formulation of calibration, all of which could overcome cognitive challenges and add to our knowledge of this problem. The following sections demonstrate the results of the two considered scenarios—the first for simultaneous utilization of multiple statistical data in the calibration process, and the second for objective incorporation of data from multiple disparate sources. Note that in the multi-objective calibration approaches, unlike the single-objective calibration, all of the involved calibration factors are evolved to determine the suitable factors.

The general goal of any multi-objective optimization is to identify the Pareto-optimal tradeoffs between multiple objectives. A Pareto-optimal tradeoff identifies the best compensation among the multiple conflicting objectives. The solutions presented on the Pareto-optimal front are nondominated solutions. A solution X1 is called dominated if, and only if, another feasible solution X2 performs better than X1 in terms of at least one objective and as well as X1 in terms of others. A set of nondominated solutions is called the nondominated or Pareto-optimal front. This nondominated solution set might contain information that advances knowledge of the problem at hand.

The final solution can be selected from the set using engineering judgment and/or other qualitative criteria. Statistical distribution shape descriptors (skewness and kurtosis) can be used in conjunction with engineering judgment to better match the distribution of measured and calculated rutting data on LTPP test sections. Difference in nonparametric skew (as calculated using equation 46) between the distributions of measured and calculated rutting values could be one of the criteria used to select the final solution among the nondominated pool of solutions. Skewness quantifies how symmetrical the distribution is.

$$NPS = \frac{\mu - \nu}{\sigma} \tag{46}$$

Where:

NPS = non-parametric skewness.
μ = mean.
ν = median.
σ = standard deviation.

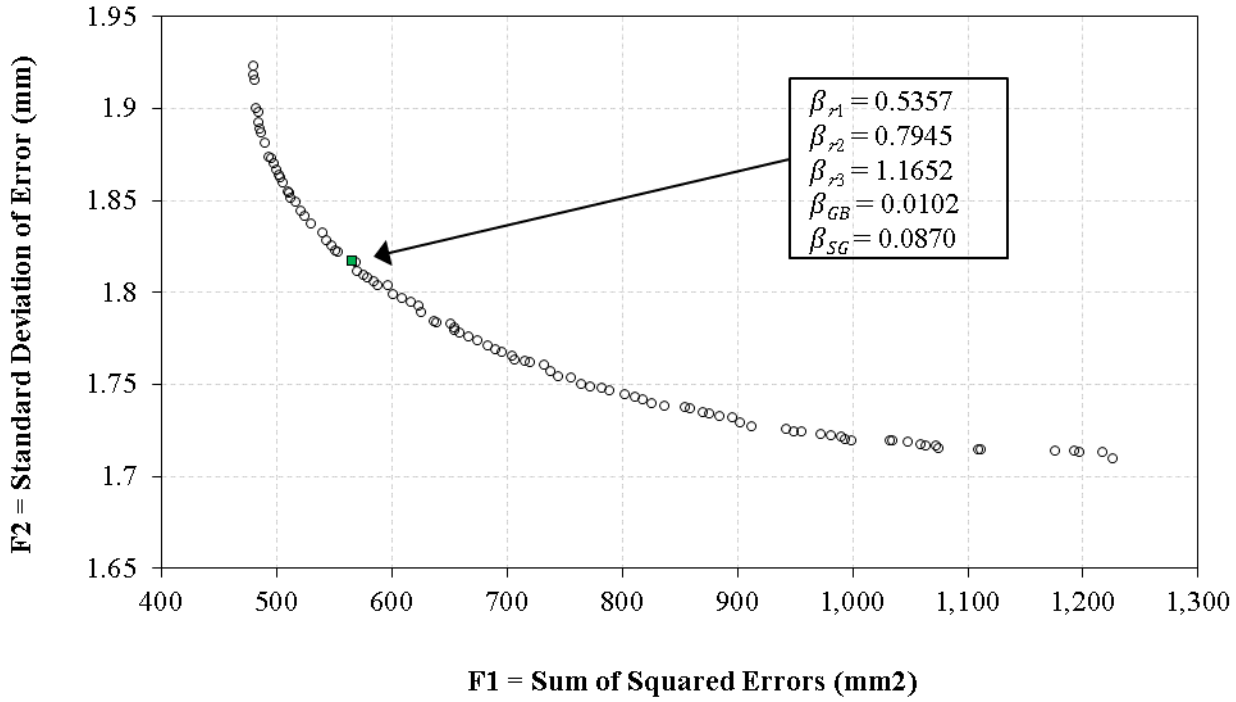
Another criterion is the difference in kurtosis (as calculated using equation 47) between the distributions of measured and calculated rutting values. Kurtosis quantifies whether the shape of the data distribution matches the Gaussian distribution. A flatter distribution has a negative kurtosis, and a distribution more peaked than a Gaussian distribution has a positive kurtosis.

$$\kappa = \frac{\sum_{i=1}^{n}(Y_i - \mu)^4}{n\sigma^4} \tag{47}$$

Scenario 1: Optimizing Major Statistical Outcomes (Two-Objective)

In the first alternative scenario, mean and standard deviation of prediction error were simultaneously minimized to reduce the bias and STE (increase model accuracy and precision) at the same time. In this manner, the information from a single calibration run was fully implemented, and an additional round of computationally intensive calibration was avoided.

Figure 24 presents the final Pareto-optimal front showing the nondominated solutions in terms of SSE and STE for the SPS-1 data. As explained in chapter 4, it cannot be proven whether the selected objective functions are in conflict. In other words, decreasing SSE might not necessarily decrease STE. As shown in this graph, the results of this specific optimization indicate that decreasing one objective might result in increasing the other. Since the two objectives of minimizing bias (increasing accuracy) and minimizing STE (increasing precision) seem to be conflicting objectives in this case, the application of a multi-objective optimization algorithm is justified. This means that a final calibrated model that exhibits higher accuracy might not necessarily have higher precision as well. In this specific case, the difference in STE is not significant among the different solutions.



Source: FHWA.

Figure 24. Scatterplot. The final nondominated solution set for two-objective calibration of rutting models for new pavements on Florida LTPP SPS-1 data.

All the solutions on this nondominated front are valid solutions in terms of the mathematical optimization problem at hand. This is because no solution is better than another solution in terms of all objective functions. A solution might perform better than another in terms of one objective function (e.g., higher accuracy), but it will be performing worse in terms of the other objective function (e.g., lower precision). However, qualitative criteria and engineering judgment could be practiced when selecting the most reasonable solution from this front.

Table 30 shows two of the candidate solutions on the final nondominated front for SPS-1 data (figure 24). These solutions have the minimum difference in skewness and kurtosis between the predicted and measured rutting values. Based on these results, the solution with minimum skewness difference seems to be the suitable solution, as its difference in kurtosis is not much higher than the solution with minimum kurtosis difference:

$\beta_{r1} = 0.54$, $\beta_{r2} = 0.79$, $\beta_{r3} = 1.16$, $\beta_{GB} = 0.01$, and $\beta_{SG} = 0.09$.

Table 30. Candidate solutions from the two-objective nondominated front for SPS-1, with minimum difference in skewness and kurtosis between predicted and measured distributions.

Candidate Solutions	β_{r1}	β_{r2}	β_{r3}	β_{GB}	β_{SG}	SSE (mm^2)	STE (mm)	Skewness Difference (%)	Kurtosis Difference (%)
Minimum difference in skewness	0.5357	0.7945	1.1652	0.0102	0.0870	494.19	1.66	43.27	34.48
Minimum difference in kurtosis	0.5224	0.8152	1.1548	0.0101	0.0118	617.18	1.79	55.40	33.18

It should be noted that this solution (highlighted in figure 24) also provides a more reasonable (farther from zero) calibration factor for the subgrade rutting, compared to the single-objective calibration results. The calibration factor for rutting in the base layer seems insignificant, but all the solutions on the nondominated front shared this issue. Since trench data were not available for this study, the models could not be calibrated accurately for the rutting in unbound layers, as they are often overwhelmed by bound layers with higher stiffness.

The final results in table 31 demonstrate a low *p* value for a paired t-test of measured versus predicted total rut depth on the calibration dataset. Therefore, the null hypothesis of the bias being insignificant has been rejected. The *p* value is higher for the validation dataset (0.001), but the bias is still significant on the validation dataset as well. The negative bias indicates that the calibrated model underpredicts rut depth values on average by 1.05 mm, which is at the same accuracy of manual rut depth measurements that is 1 mm in the LTPP program.

Table 31. Two-objective calibration results of rutting models for new pavements on Florida SPS-1.

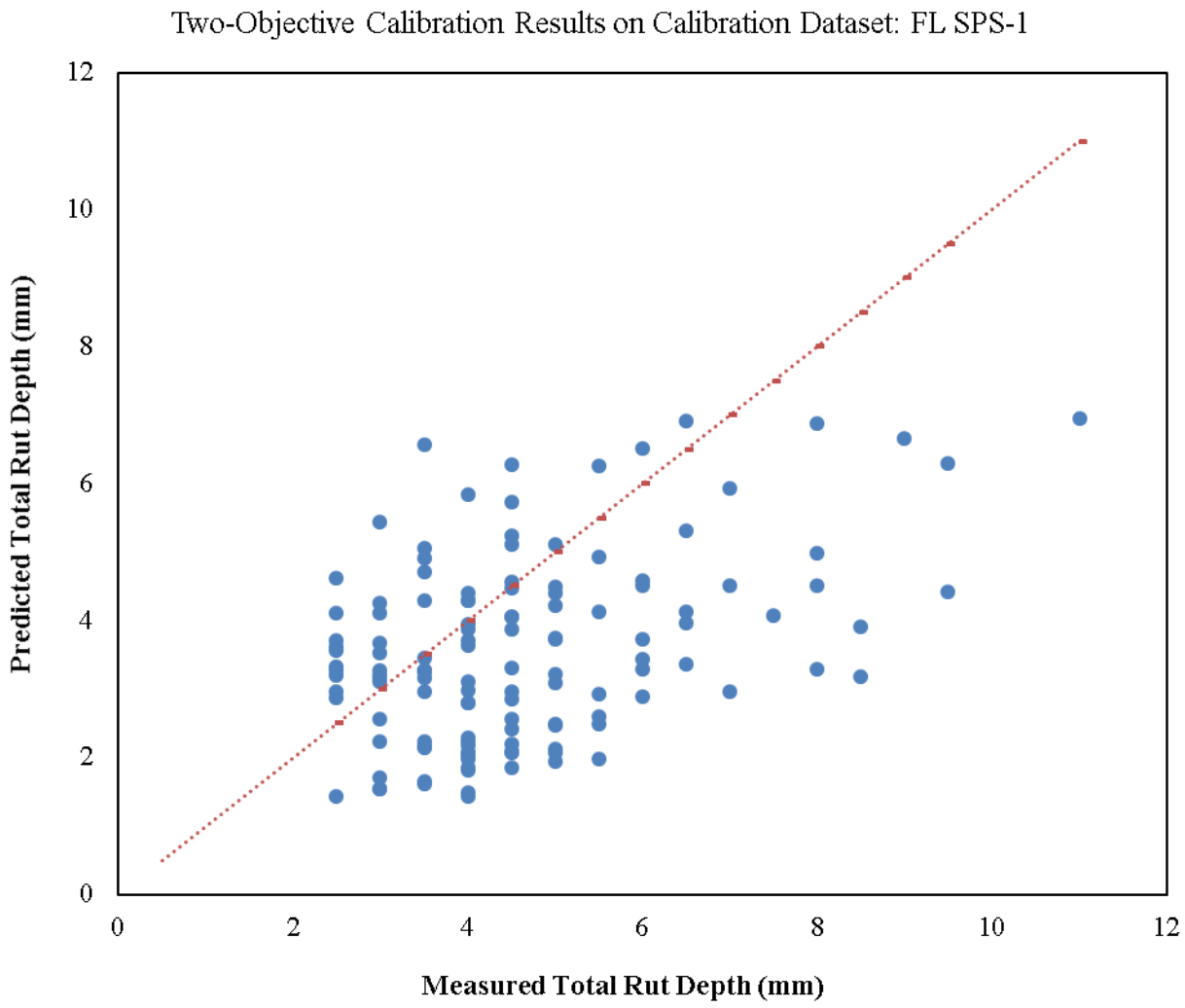
Statistic	Calibration Data	Validation Data	Combined Data
Data records	128	26	154
SSE	494.19 mm^2 (0.77 $inch^2$)	72.41 mm^2 (0.11 $inch^2$)	566.60 mm^2 |(0.88 $inch^2$)
RMSE	1.96 mm (0.077 inch)	1.67 mm (0.066 inch)	1.92 mm (0.076 inch)
Bias	–1.06 mm (–0.042 inch)	–1.01 mm (–0.040 inch)	–1.05 mm (–0.041 inch)
p value (paired *t*-test) for bias	4.56E-11 < 0.05; significant bias	0.001 < 0.05; significant bias	1.68E-13 < 0.05; significant bias
STE	1.66 mm (0.065 inch)	1.36 mm (0.054 inch)	1.61 mm (0.063 inch)
Generalization capability	N/A	N/A	95.28%
R^2 (goodness of fit)	0.1778	0.3067	0.194

N/A = not applicable; RMSE = root-mean-squared error.

In comparison to the single-objective calibration results, the two-objective calibration has resulted in lower accuracy (a higher bias), but higher precision (a lower STE), of the final model on both the calibration and validation datasets. This is because the standard deviation of error was also minimized simultaneously with the SSE. In addition to an increased precision, the other improvement is the generalization capability of the calibrated model. The model that was calibrated using two-objective optimization had more similar bias values on calibration and validation data, compared to the single-objective results.

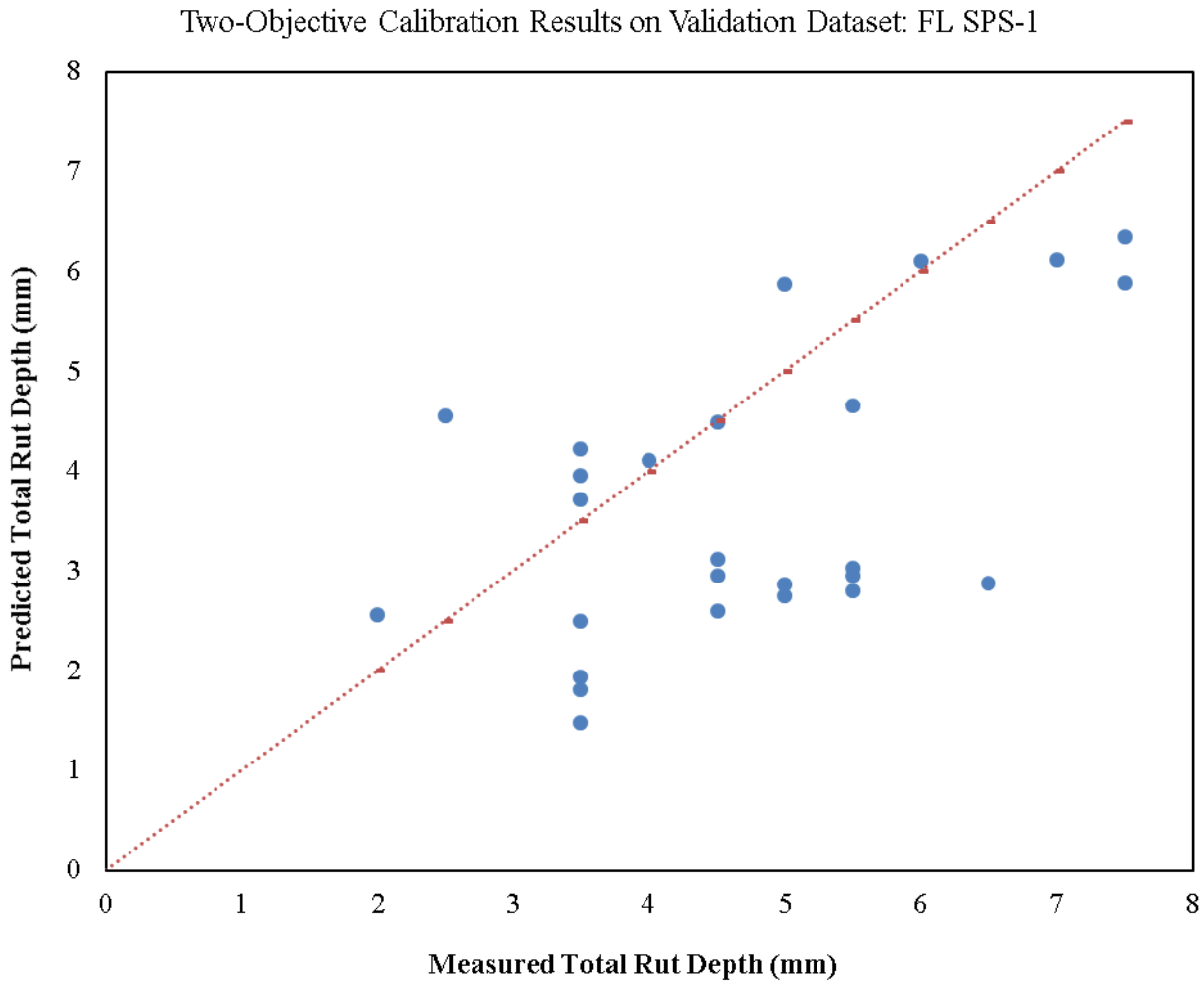
Figure 25 and figure 26 show scatterplots of the measured versus predicted total rut depth on calibration (128 records) and validation (26 records) datasets, respectively, for Florida SPS-1. Similar to the single-objective calibration results, there is a significant amount of scatter in these plots, and the goodness-of-fit indicator (R^2) is poor. As explained before, this scatter indicates the lack of precision of the rutting model. However, the overall STE is lower than the national global calibration results (0.063 inch < 0.129 inch).

Overall, it seems that the two-objective optimization has increased the precision and generalization capability of the final calibrated model at the cost of decreasing accuracy. Since the changes in the results are not significant, conducting scenario 1 of the multi-objective optimization might not be worth the computational cost.



Source: FHWA.

Figure 25. Scatterplot. Measured versus predicted two-objective calibration results of rutting models for new pavements on calibration dataset for Florida SPS-1.

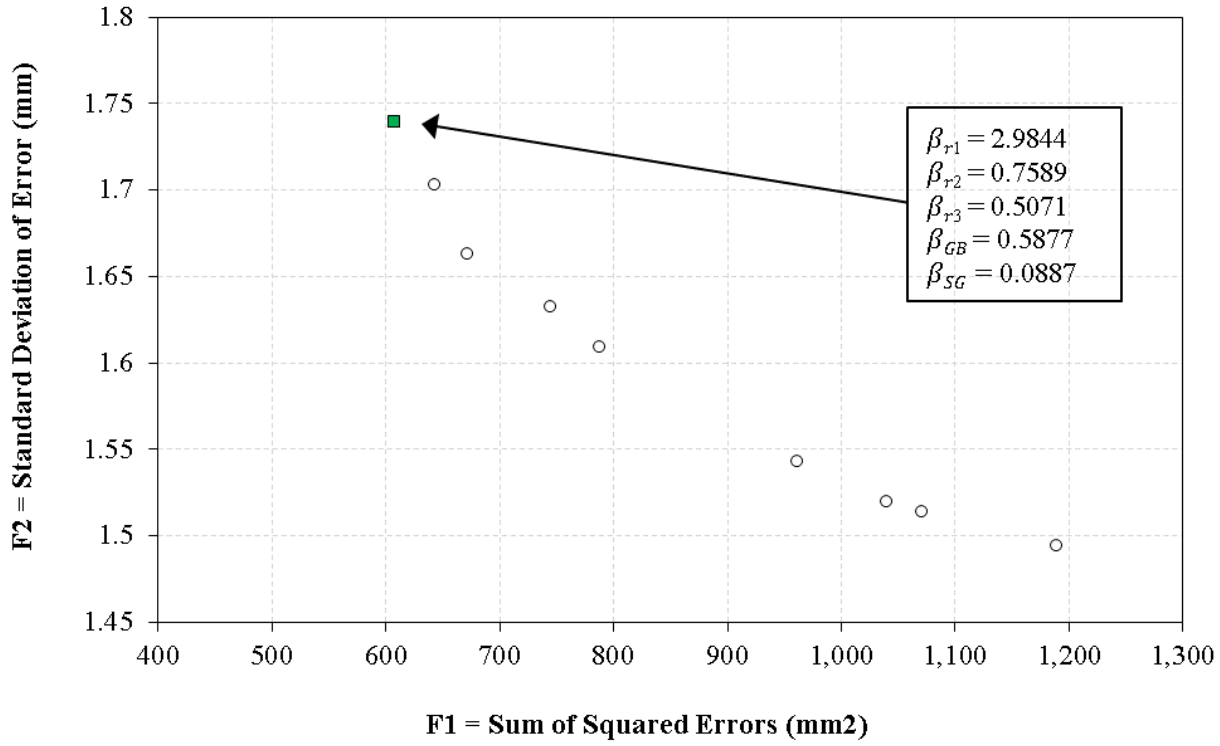


Source: FHWA.

Figure 26. Scatterplot. Measured versus predicted two-objective calibration results of rutting models for new pavements on validation dataset for Florida SPS-1.

From the two-objective Pareto-optimal front on SPS-1 (figure 24), it seems that using an epsilon parameter (MOEA precision factor) of 0.1 was too conservative and has produced too many solutions with similar objective function values. A higher epsilon value (equal to 1.0 instead of 0.1) was used for the two-objective calibration on SPS-5 data, and that has increased the speed and efficiency of the multi-objective optimization drastically (by about 70 percent). Figure 27 shows the final Pareto-optimal front showing the nondominated solutions in terms of SSE and STE for the SPS-5 data.

Again, the skewness and kurtosis of the predicted rutting distribution with every solution on the nondominated front were calculated and compared to the measured values (table 32). From these calculations, it is evident that one of the solutions on the final front produces rutting predictions that have the lowest difference in distribution skewness and kurtosis from the distribution of measured rut depth values. However, there is another similar solution (in bold font) that has resulted in the most reasonable calibration factor for the subgrade rutting. The final selected solution is also highlighted in figure 27.



Source: FHWA.

Figure 27. Scatterplot. The final nondominated solution set for two-objective calibration of rutting models for overlaid pavements on Florida LTPP SPS-5 data.

Table 32. Solutions from the two-objective nondominated front for SPS-5, with difference in skewness and kurtosis between predicted and measured data distributions.

Candidate Solutions	β_{r1}	β_{r2}	β_{r3}	β_{GB}	β_{SG}	SSE (mm^2)	STE (mm)	Skewness Difference (%)	Kurtosis Difference (%)
Other viable solution	1.8778	0.7588	0.5071	0.4438	0.0226	1,188.72	1.49	52.33	19.48
Other viable solution	2.1152	0.9602	0.5071	0.4405	0.0246	1,070.47	1.51	58.65	29.54
Other viable solution	2.1152	0.9602	0.5071	0.5877	0.0246	744.75	1.63	61.78	25.72
Other viable solution	2.1185	0.9939	0.5069	0.4370	0.0256	1,039.45	1.52	51.96	33.24
Other viable solution	2.1185	0.9939	0.5069	0.4746	0.0225	961.70	1.54	59.60	31.58
Minimum skewness and kurtosis difference	2.9712	0.7589	0.5071	0.5877	0.0253	787.72	1.61	51.89	19.28
Other viable solution	2.9838	0.7589	0.5071	0.5877	0.0540	671.14	1.66	56.40	22.83
Other viable solution	**2.9844**	**0.7589**	**0.5071**	**0.5877**	**0.0887**	**541.80**	**1.64**	**53.49**	**25.64**
Other viable solution	2.9844	0.9599	0.5071	0.5877	0.0500	643.39	1.70	56.25	31.35

The final selected solution seems to have reasonable calibration factors for base and subgrade rutting: β_{r1} = 2.98, β_{r2} = 0.76, β_{r3} = 0.51, β_{GB} = 0.59, and β_{SG} = 0.09. Table 33 shows the prediction results of this final solution on SPS-5 calibration and validation datasets. While the low bias value of –0.44 mm is much less than the LTPP measurement precision of 1 mm, the bias is statistically significant on the calibration dataset. However, the bias is statistically insignificant on the validation dataset that was not used in the calibration process.

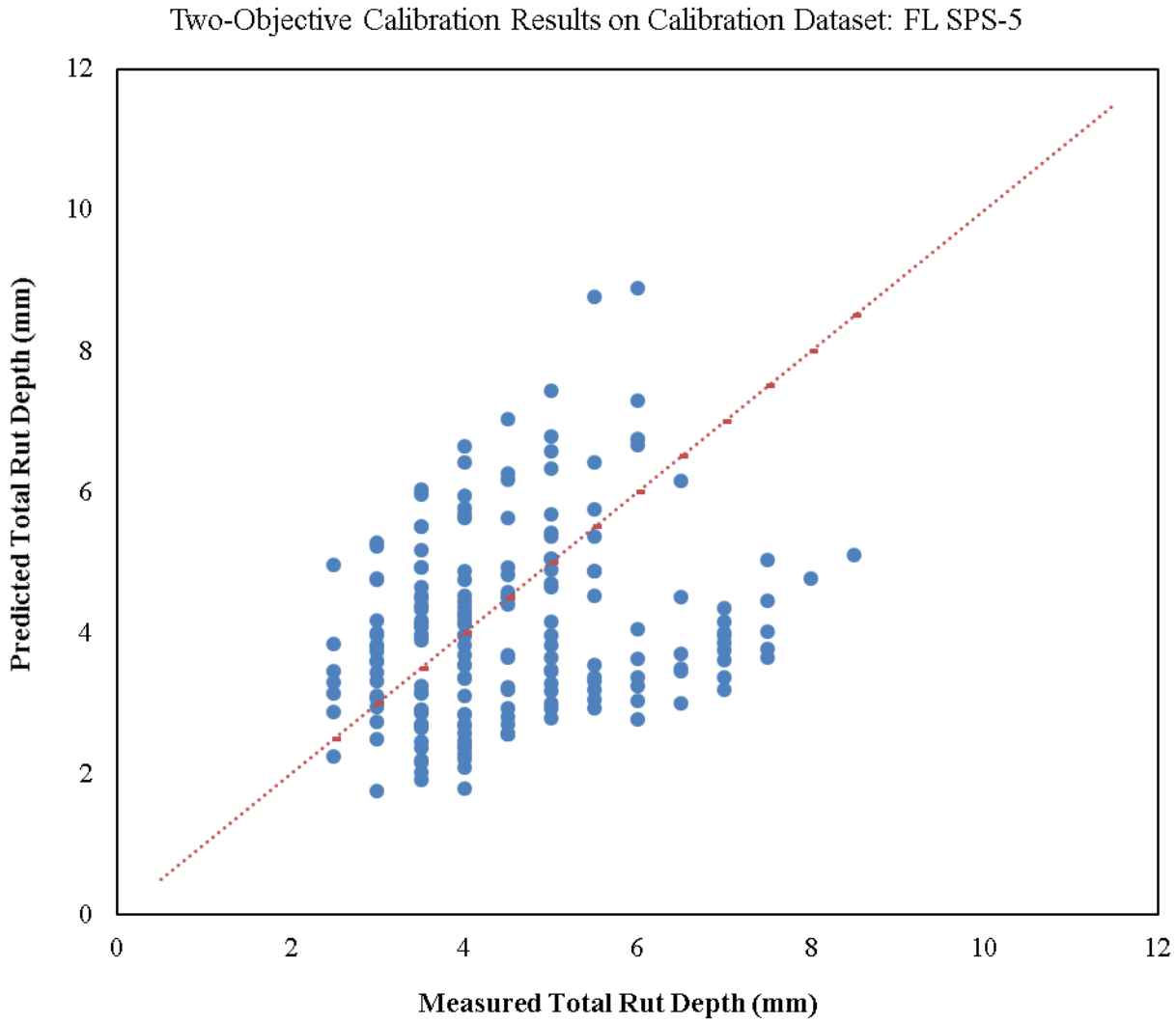
Table 33. Two-objective calibration results of rutting models for overlaid pavements on Florida SPS-5.

Statistic	Calibration Data	Validation Data	Combined Data
Data records	189	39	228
SSE	541.80 mm^2 (0.84 inch^2)	91.65 mm^2 (0.14 inch^2)	633.45 mm^2 (0.98 inch^2)
RMSE	1.69 mm (0.066 inch)	1.53 mm (0.060 inch)	1.67mm (0.066 inch)
Bias	–0.44 mm (–0.017 inch)	–0.47 mm (–0.018 inch)	–0.45 mm (–0.018 inch)
p value (paired t-test) for bias	0.0003 < 0.05; significant bias	0.051 > 0.05; insignificant bias	4.16E-05 < 0.05; significant bias
STE	1.64 mm (0.065 inch)	1.48 mm (0.058 inch)	1.61 mm (0.063 inch)
Generalization capability	N/A	N/A	93.18%
R^2 (goodness of fit)	0.0427	0.0013	0.0348

N/A = not applicable; RMSE = root-mean-squared error.

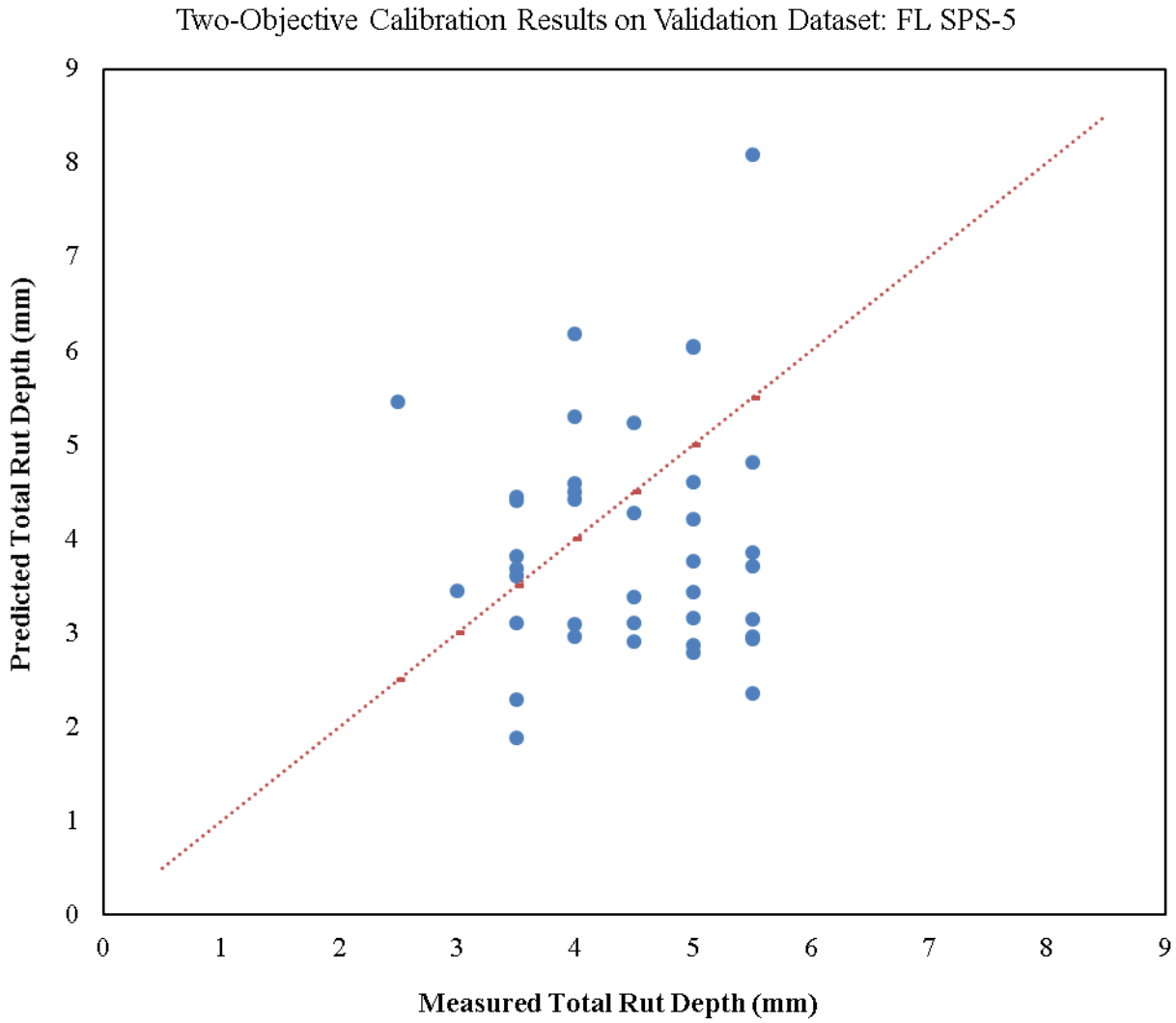
Figure 28 and figure 29 show scatterplots of the measured versus predicted total rut depth on calibration (189 records) and validation (39 records) datasets, respectively, for Florida SPS-5. Similar to the single-objective calibration, there is significant scatter in these plots. However, the standard deviation of error is lower for the model that was calibrated using the two-objective approach compared to the single-objective approach. The overall standard deviation of error is much lower than the STE from the national global calibration (0.063 < 0.129).

Compared to the single-objective calibration, the rutting model for overlaid pavements (SPS-5) that was calibrated using the two-objective approach shows both a higher accuracy and higher precision. In addition, the two-objective calibration has resulted in a model with much greater generalization capability.



Source: FHWA.

Figure 28. Scatterplot. Measured versus predicted two-objective calibration results of rutting models for overlaid pavements on calibration dataset for Florida SPS-5.



Source: FHWA.

Figure 29. Scatterplot. Measured versus predicted two-objective calibration results of rutting models for overlaid pavements on validation dataset for Florida SPS-5.

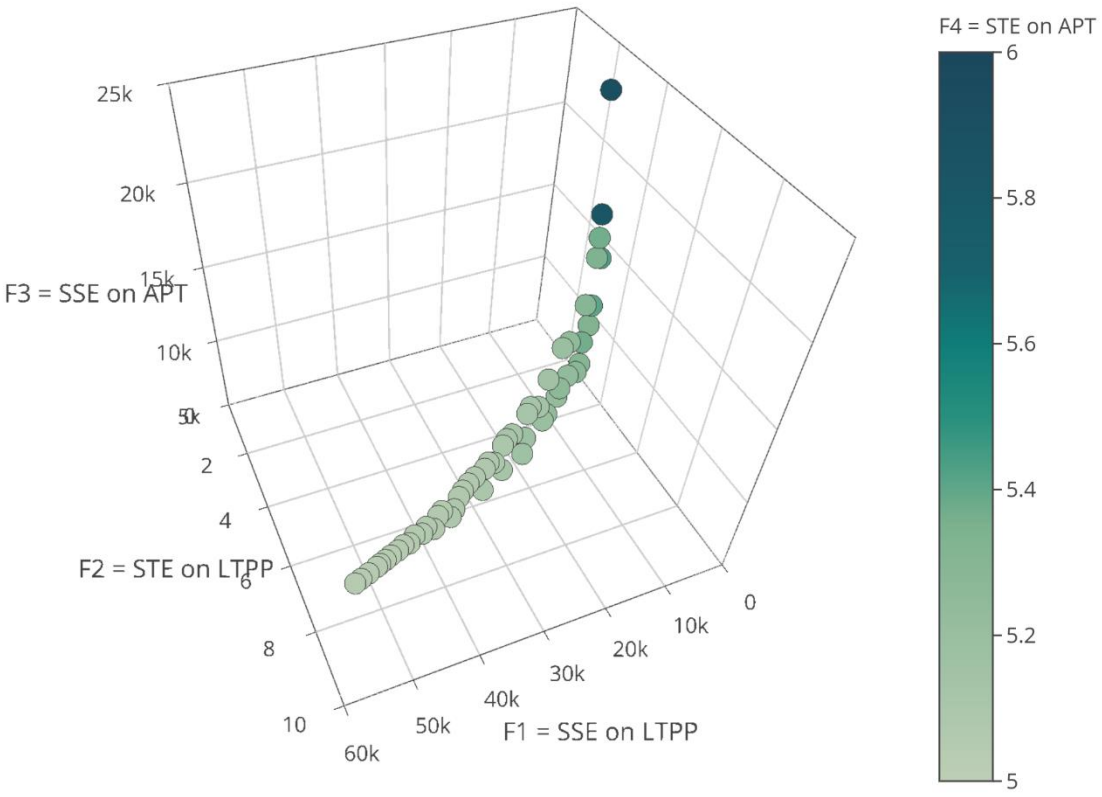
Scenario 2: Combining Different Sources of Data (Four-Objective)

In the second scenario, the SSE and STE in predicting permanent deformation of pavements within different performance data sources were used as separate objective functions to be minimized simultaneously. This four-objective optimization was used in calibrating rutting models for new AC pavements using the LTPP SPS-1 Florida site data and the FDOT APT data at the same time. In this manner, information from pavement performance near the end of its service life can be incorporated in the calibration process. This scenario comprised an objective approach to incorporate different sources of data. This scenario was not applied to overlaid pavements because the available APT data were only for new pavement structures.

Figure 30 shows the Pareto-optimal front for the four-objective calibration of the rutting models for new pavements using SPS-1 and APT data, simultaneously. In this figure, F1 and F2 are SSE

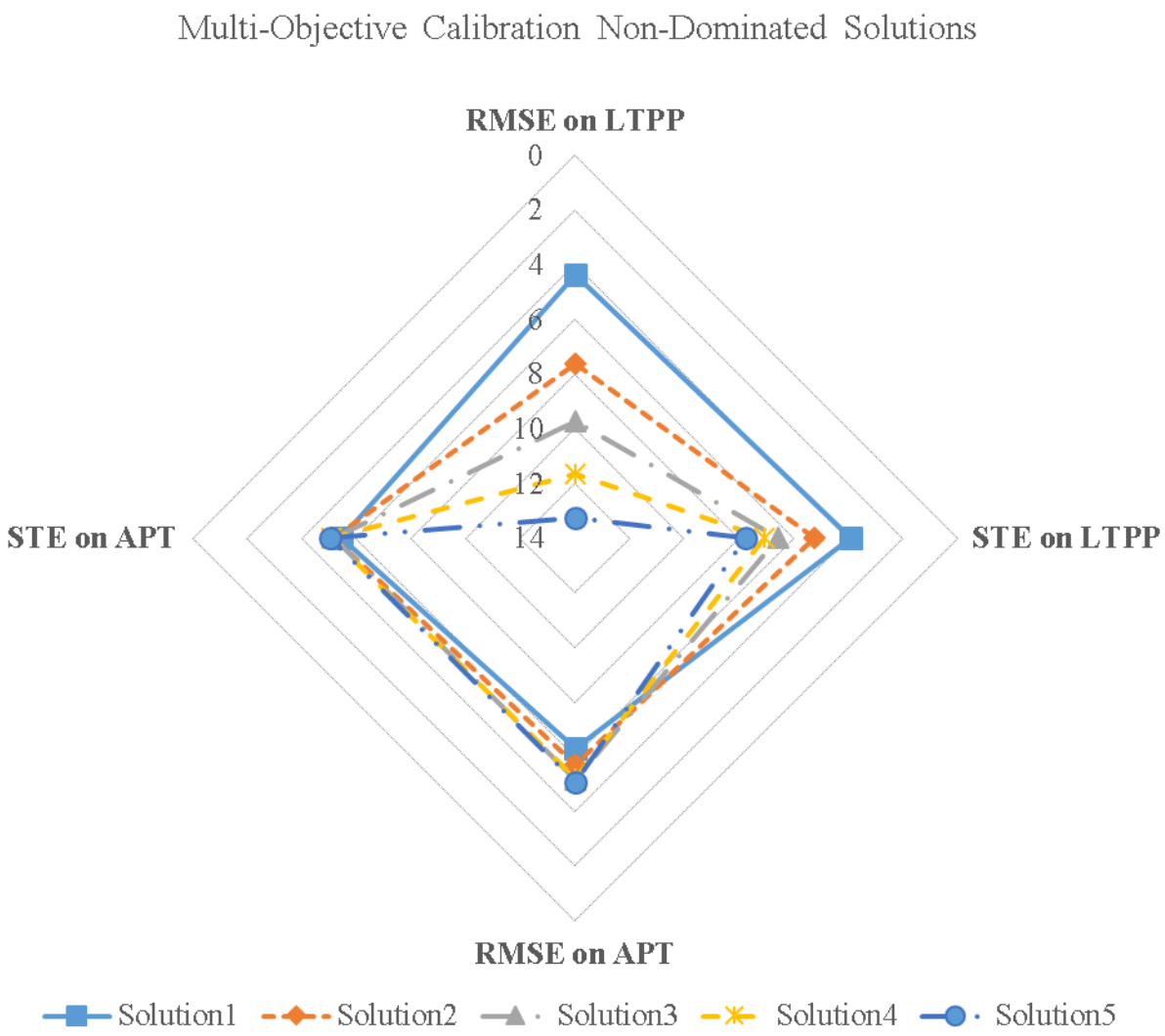
and STE on Florida LTPP SPS-1 data, and F3 and F4 are SSE and STE on FDOT APT data. While F1, F2, and F3 are indicated on the three dimensions, the values for F4 are indicated using a color chart, where lower values are closer to lighter gray. Figure 31 shows another approach to visualizing the final nondominated front for this four-objective optimization. In this figure, root-mean-squared error (RMSE) is shown instead of SSE so that the values of the average error objective functions (F1 and F3) are relatively in the same scale and comparable with standard deviation functions (F2 and F4). Unlike figure 30, where each solution (set of calibration factors) was indicated using a circle, in figure 31, each solution is represented by a line that crosses the objective function axes at different locations.

Figure 32 and figure 33 exhibit two-dimensional representations of the final Pareto-optimal front using pairwise comparison of SSE and STE, respectively. Each figure compares the average error or standard deviation of error on SPS-1 data to that on APT data. Note that some solutions on these fronts might seem suboptimal; however, that is because these figures are showing a two-dimensional shadow of the final four-dimensional nondominated front on one plane. Also note that there is a crowd of solutions on these fronts because an epsilon value of 0.1 was used; an epsilon value of 1.0 would result in a more reasonable density of solutions in the final front. Similar to the first multi-objective calibration scenario, these figures show a conflict between the F1 and F3 and F2 and F4 objective functions, which advocates for the application of multi-objective optimization. The advantage of using the multi-objective approach in this case is that solutions closer to the performance measured on SPS-1 sections can be preferred over other solutions closer to the performance observed in APT data. This is because the LTPP sections are actual in-service highway pavements exposed to actual climate and traffic, as opposed to APT sections, which have been exposed to simulated accelerated loading under a constant temperature.



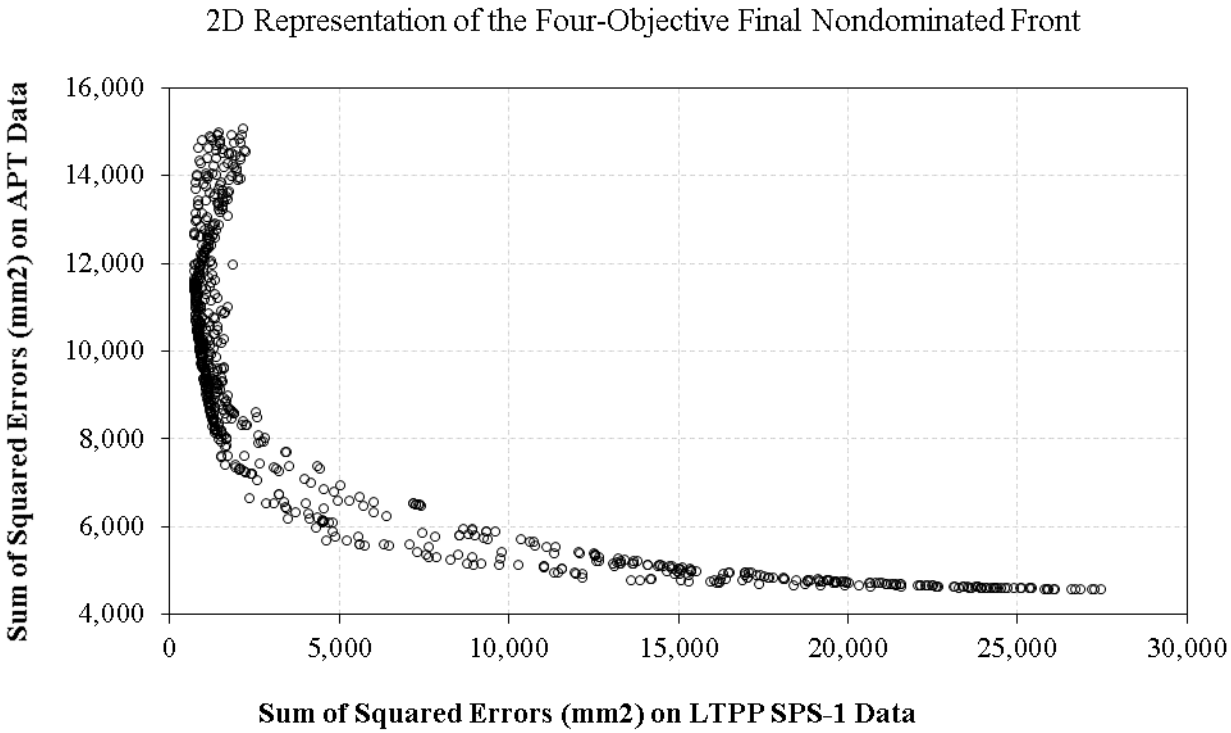
Source: FHWA.

Figure 30. Scatterplot. The final nondominated solution set for four-objective calibration of rutting models for new pavements: F1 and F2 are SSE and STE on Florida LTPP SPS-1 data, and F3 and F4 are SSE and STE on FDOT APT data.



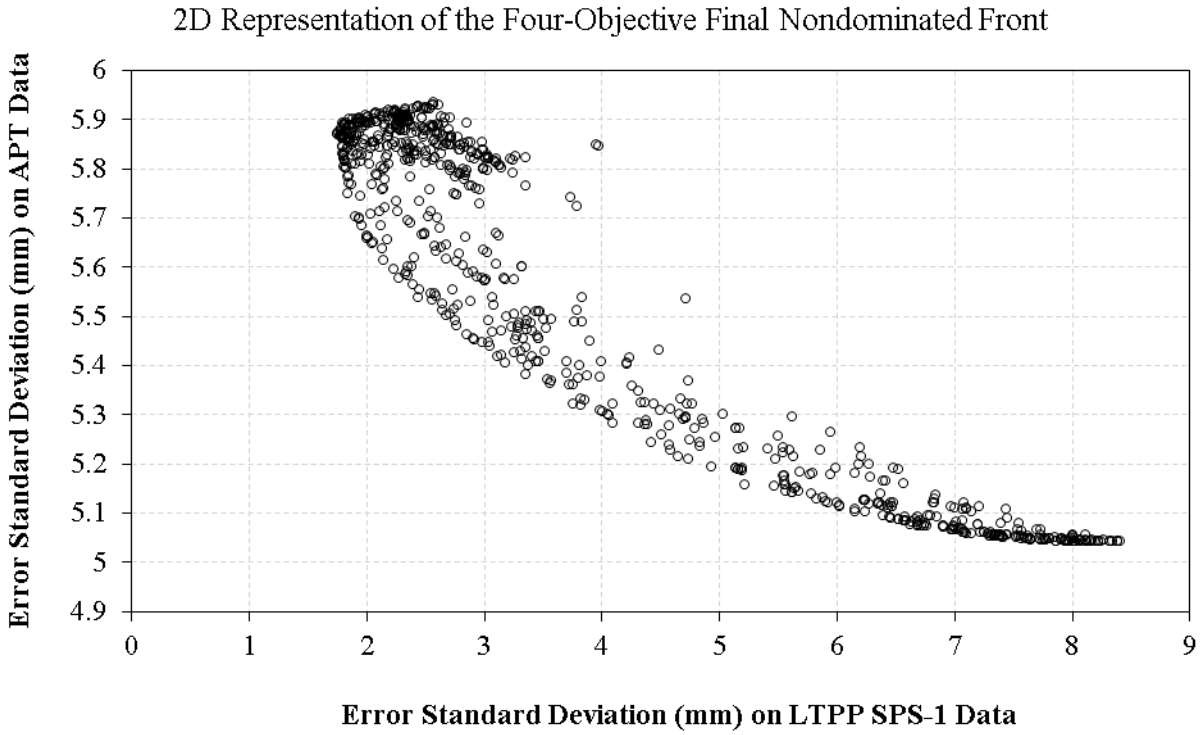
Source: FHWA.

Figure 31. Chart. The final nondominated solution set for four-objective calibration of rutting models for new pavements: RMSE and STE on Florida SPS-1 and FDOT APT data.



Source: FHWA.

Figure 32. Scatterplot. Two-dimensional representation of the final nondominated solution set for four-objective calibration: SSE on Florida LTPP SPS-1 versus SSE on FDOT APT data.



Source: FHWA.

Figure 33. Scatterplot. Two-dimensional representation of the final nondominated solution set for four-objective calibration: STE on Florida LTPP SPS-1 versus STE on FDOT APT data.

Table 34 shows some viable candidate solutions from the final nondominated front. The solution that has the minimum difference in kurtosis between the predicted and measured SPS-1 data is exhibiting a high difference in skewness. Therefore, other candidates with a better balance were sought. Some solutions show a very low value for the calibration factors for either base or subgrade, and those solutions were avoided. Therefore, it seems that the solution with the minimum difference in kurtosis is the best candidate, as it exhibits reasonable calibration factors with relatively good quality distribution of predicted values (in addition to being on the final nondominated front of the multi-objective optimization). Note that this solution offers more reasonable calibration factors for base and subgrade layers compared to the solutions on the single-objective and two-objective calibration (where all of the solutions had insignificant values for the calibration coefficient of the base layer):

β_{r1} = 3.84, β_{r2} = 0.96, β_{r3} = 0.56, β_{GB} = 0.14, and β_{SG} = 0.79.

Table 34. Candidate solutions from the four-objective nondominated front with difference in skewness and kurtosis between the predicted and measured data distributions.

Candidate Solutions	β_{r1}	β_{r2}	β_{r3}	β_{GB}	β_{SG}	On LTPP Data SSE (mm^2)	On LTPP Data STE (mm)	Skewness Difference (%)	Kurtosis Difference (%)
Minimum difference in kurtosis	3.8397	0.9549	0.5625	0.1377	0.7861	539.53	2.05	105.03	29.52
Minimum difference in skewness	3.8476	0.6471	0.8332	1.4298	0.4571	1,288.64	3.03	11.41	53.79
Other viable solution	1.2356	0.6223	0.7848	0.7718	0.7178	893.73	2.64	30.82	44.60
Other viable solution	6.2883	0.5961	0.9930	0.0104	0.5231	736.54	2.27	139.10	35.43
Other viable solution	1.0952	1.4996	0.5051	1.3954	0.0140	2,354.96	3.99	58.50	52.62
Other viable solution	1.0952	1.4559	0.5016	0.0560	0.0189	823.28	2.28	151.96	41.42
Other viable solution	0.9257	0.8089	1.0421	0.0206	0.1975	1,090.185	1.87	139.77	38.17

The final results in table 35 demonstrate a high *p* value for a paired *t*-test of measured versus predicted total rut depth on both calibration and validation datasets (0.18 and 0.51, respectively). Therefore, there is not enough evidence to reject the null hypothesis of the bias being insignificant. The negative bias indicates that the calibrated model underpredicts rut depth values on average by 0.28 mm, which is insignificant compared to the accuracy of manual rut depth measurements that is 1 mm in the LTPP program.

Based on the final selected solution, figure 34 and figure 35 show the measured versus predicted total rut depth on Florida SPS-1 calibration and validation datasets, respectively. Similar to the single-objective and two-objective calibration results, there is a significant amount of scatter in these plots, and the goodness-of-fit indicator (R^2) is poor. As explained before, this scatter indicates the lack of precision of the rutting model. However, the overall STE is lower than the national global calibration results (0.078 inch < 0.129 inch).

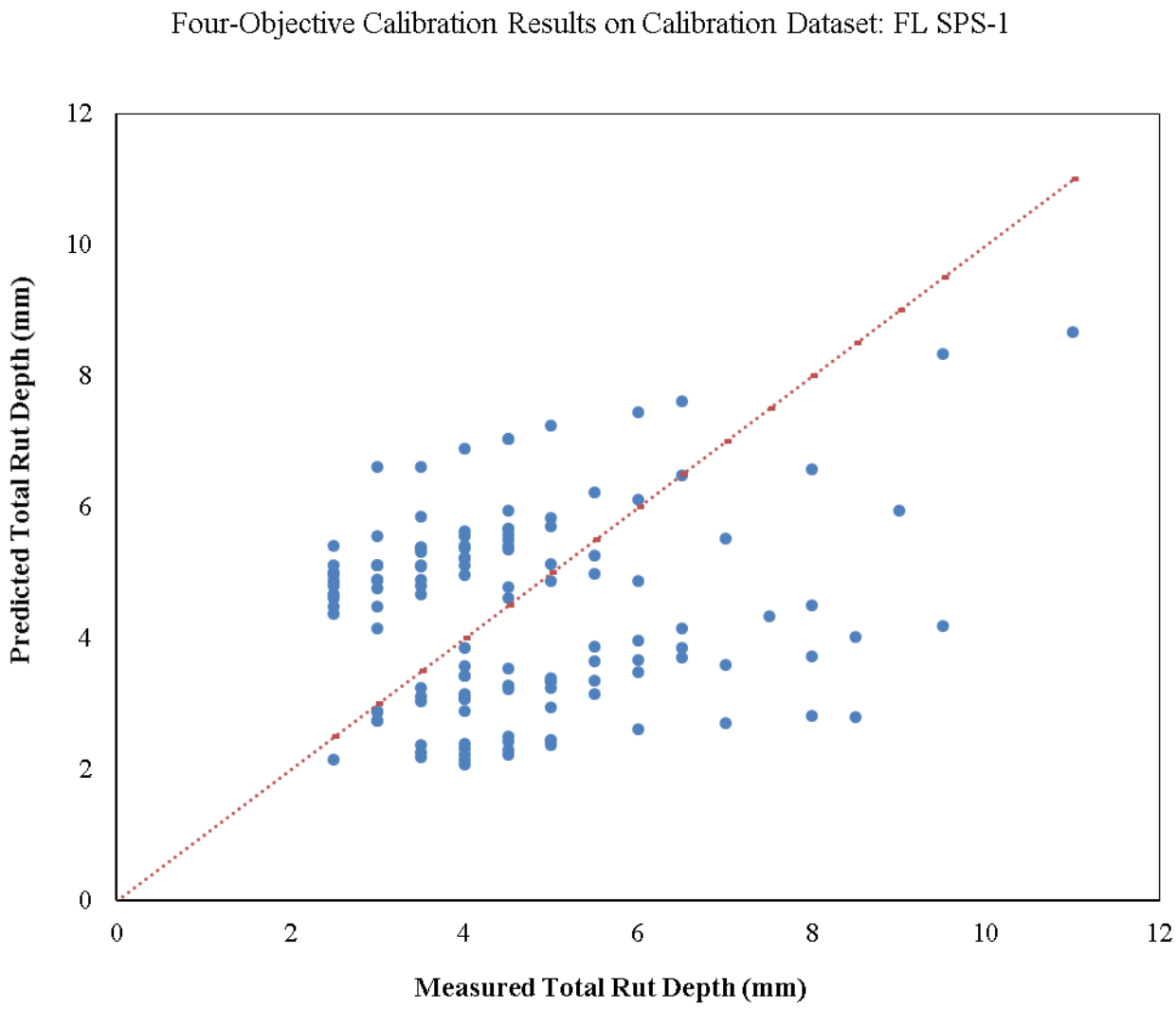
Table 35. Four-objective calibration results of rutting models for new pavements on Florida SPS-1 data.

Statistic	Calibration Data	Validation Data	Combined Data
Data records	128	26	154
SSE	539.53 mm^2 (0.84 $inch^2$)	71.13 mm^2 (0.11 $inch^2$)	610.66 mm^2 (0.95 $inch^2$)
RMSE	2.05 mm (0.081 inch)	1.65 mm (0.065 inch)	1.99 mm (0.078 inch)
Bias	–0.24 mm (–0.009 inch)	–0.24 mm (–0.009 inch)	–0.24 mm (–0.009 inch)
p value (paired t-test) for bias	0.18 > 0.05; insignificant bias	0.51 > 0.05; insignificant bias	0.125 > 0.05; insignificant bias
STE	2.05 mm (0.081 inch)	1.67 mm (0.066 inch)	1.98 mm (0.078 inch)
Generalization capability	N/A	N/A	100%
R^2 (goodness of fit)	0.0263	0.1478	0.0382

N/A = not applicable.

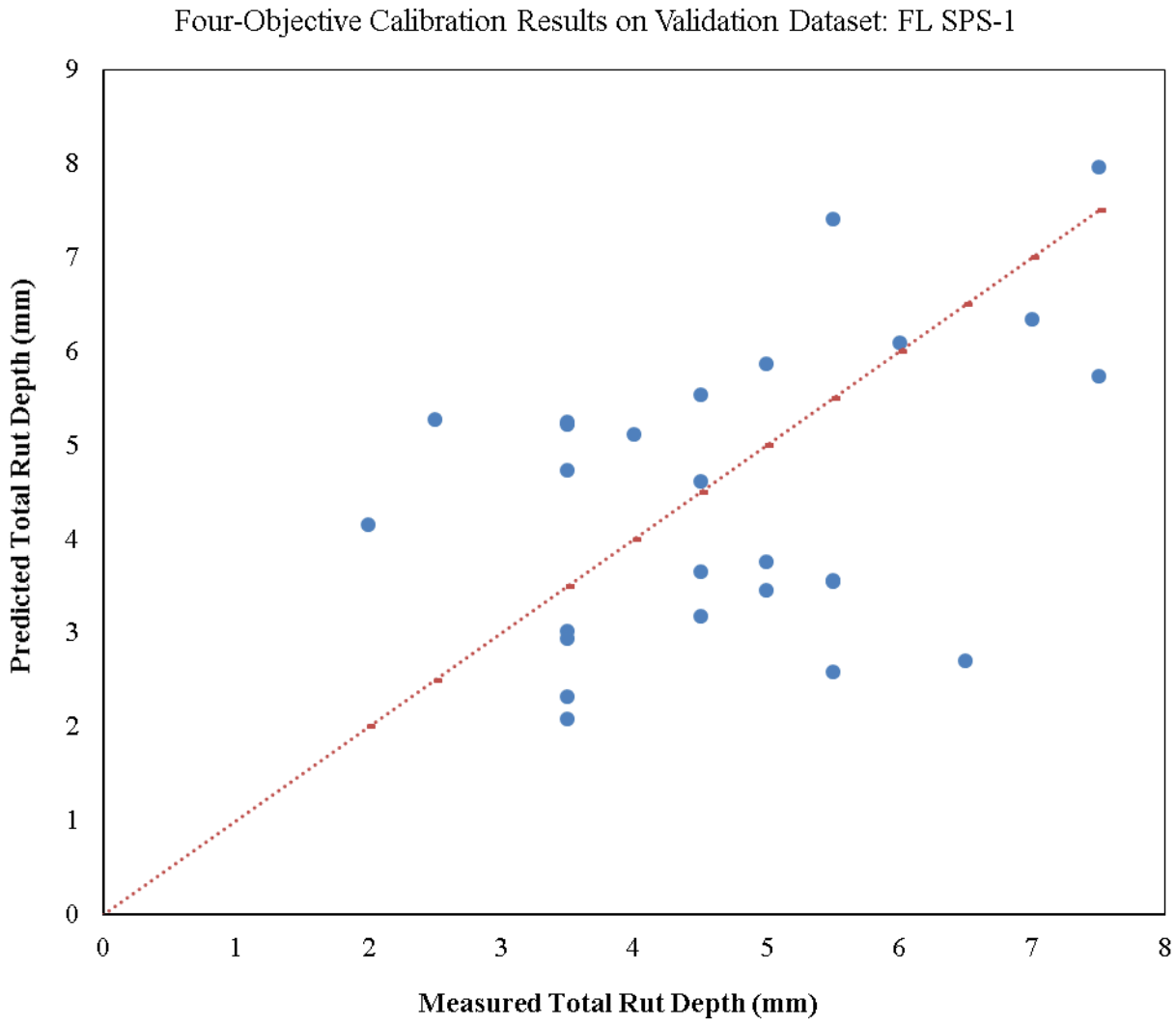
In comparison to the single-objective calibration results, the four-objective calibration has resulted in increased accuracy (lower bias and even higher *p* value), but lower precision (a higher STE), of the final model on both the calibration and validation datasets. This is because the four-objective scenario has included APT data in addition to the LTPP data, which are different in nature. The main improvement in the results of four-objective calibration compared to single-objective calibration is the increased generalization capability of the calibrated model. The model that was calibrated using four-objective optimization had more similar bias values on calibration and validation data, compared to the single-objective results.

While the four-objective calibration results are not superior to the single-objective results in terms of precision, the simultaneous minimization of error on both LTPP and APT data has actually resulted in higher accuracy and generalization capability of the calibrated model.



Source: FHWA.

Figure 34. Scatterplot. Measured versus predicted four-objective calibration results of rutting models for new pavements on calibration dataset for Florida SPS-1.



Source: FHWA.

Figure 35. Scatterplot. Measured versus predicted four-objective calibration results of rutting models for new pavements on validation dataset for Florida SPS-1.

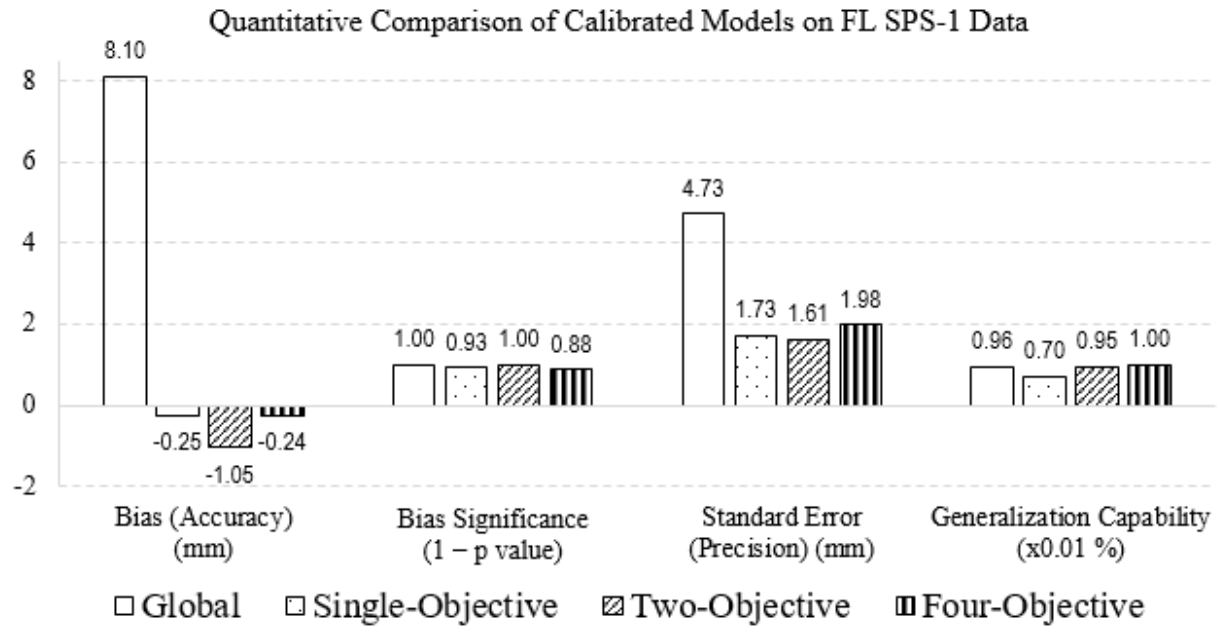
COMPARISON OF RESULTS

To investigate the advantages of the novel calibration approach recommended in this study, the final single-objective and multi-objective calibrated models were compared through a comprehensive framework. This framework (figure 17) includes quantitative measures of accuracy, precision, and generalization capability. A model that has high accuracy and precision might not necessarily have an adequate generalization capability, meaning that it would not be able to reproduce the same accuracy when predicting the performance on other similar test sections. The employed qualitative measures include the goodness of fit and similarity of the distribution shape between measured and predicted performance. The quality of the solutions needs to be also examined in terms of the engineering reasonableness of the selected calibration factors. Finally, a combination of quantitative and qualitative criteria is used by comparing the

performance trends prediction. The capability of the final calibrated model in predicting future performance trends is an essential feature for pavement design and therefore a significant indicator of the quality of the developed model.

Quantitative Comparison

Figure 36 shows a comparison of the quantitative criteria among the single-objective, two-objective, and four-objective calibrated models for rutting in new pavements on combined (calibration and validation) SPS-1 data.

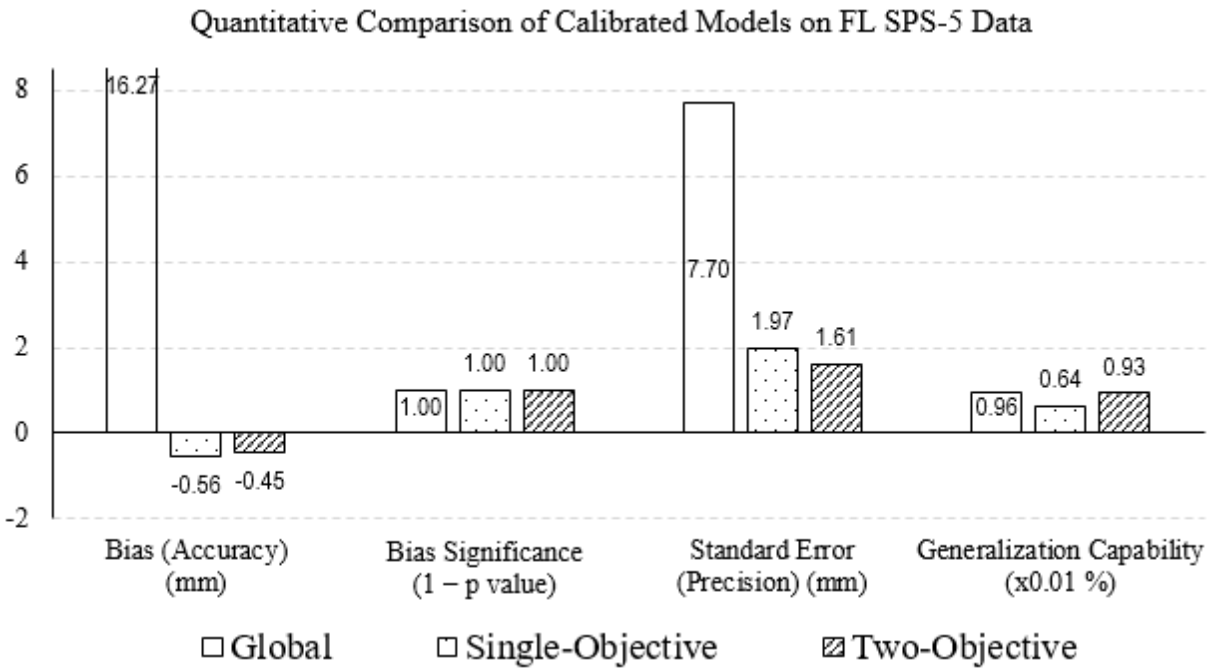


Source: FHWA.

Figure 36. Bar chart. Comparison of the quantitative criteria for the calibrated rutting models on SPS-1.

It is evident in this figure that the incorporation of the APT data in calibration of rutting models on LTPP SPS-1 data has led to a model with the lowest bias (and least significant) and the highest generalization capability. On the other hand, the two-objective minimization of SSE and STE has resulted in the lowest STE (highest precision), but a significant bias (low accuracy).

Figure 37 shows a comparison of the quantitative criteria between the single-objective and two-objective calibrated models for rutting in overlaid pavements on combined (calibration and validation) SPS-5 data. There were no APT data available for this calibration of rutting models for overlaid pavements. Unlike the case in calibration on SPS-1 data, the two-objective calibration on SPS-5 data has resulted in both an increase in accuracy (lower bias but still significant), an increase in precision (lower STE), and an increase in generalization capability of the rutting model compared to single-objective calibration.

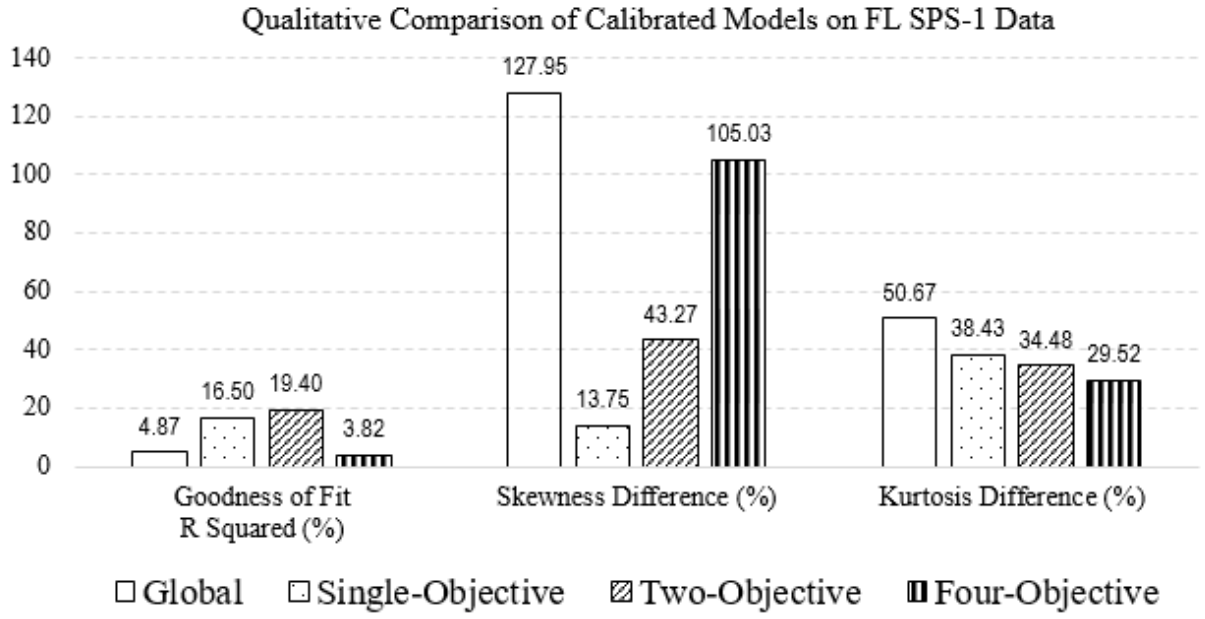


Source: FHWA.

Figure 37. Bar chart. Comparison of the quantitative criteria for the calibrated rutting models on SPS-5.

Qualitative Comparison

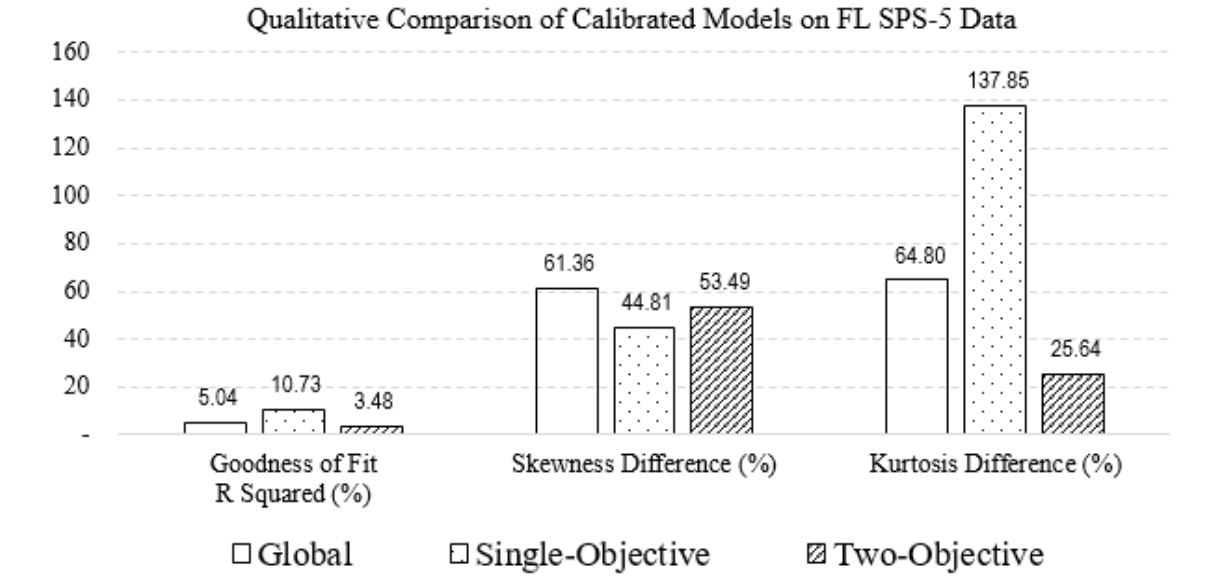
A model that exhibits the best quantitative results might not necessarily show the most desirable qualitative performance. Figure 38 shows a comparison of the qualitative criteria among the single-objective, two-objective, and four-objective calibrated models for rutting in new pavements on combined (calibration and validation) SPS-1 data. The two-objective results show the best goodness of fit, but as explained before, there is a lack of precision of the mechanistic model, and all of the calibration approaches resulted in significant scatter. The single-objective results have the least difference in skewness between the distributions of measured and predicted rutting. The four-objective results show the least difference in kurtosis (flatness) between the distributions of measured and predicted rutting.



Source: FHWA.

Figure 38. Bar chart. Comparison of the qualitative criteria for the calibrated rutting models on SPS-1.

Figure 39 shows a comparison of the qualitative criteria between the single-objective and two-objective calibrated models for rutting in overlaid pavements on combined (calibration and validation) SPS-5 data. The single-objective results show the best goodness of fit and the lowest skewness difference between measured and predicted data. The two-objective results show the least difference in kurtosis (flatness) between the distributions of measured and predicted rutting.



Source: FHWA.

Figure 39. Bar chart. Comparison of the qualitative criteria for the calibrated rutting models on SPS-5.

In all cases, using the multi-objective approach has resulted in predicted rutting distributions that are more similar in flatness to the measured rutting distributions.

Table 36 shows the final selected calibration factors. It is evident that the calibration factors selected for the rutting in base and subgrade layers were insignificant in single-objective calibration. However, the multi-objective calibration has resulted in more reasonable calibration factors for rutting in unbound layers.

Table 36. Final selected calibration factors.

Model	β_{r1}	β_{r2}	β_{r3}	β_{GB}	β_{SG}
New pavement (SPS-1) global	1.0	1.0	1.0	1.0	1.0
New pavement (SPS-1) single-objective	0.522	1.0	1.0	0.011	0.171
New pavement (SPS-1) two-objective	0.536	0.795	1.165	0.010	0.087
New pavement (SPS-1) four-objective	3.840	0.955	0.563	0.138	0.786
Overlaid pavement (SPS-5) global	1.0	1.0	1.0	1.0	1.0
Overlaid pavement (SPS-5) single-objective	0.500	1.0	1.0	0.074	0.155
Overlaid pavement (SPS-5) two-objective	2.984	0.759	0.507	0.588	0.089

Quantitative and Qualitative Comparison

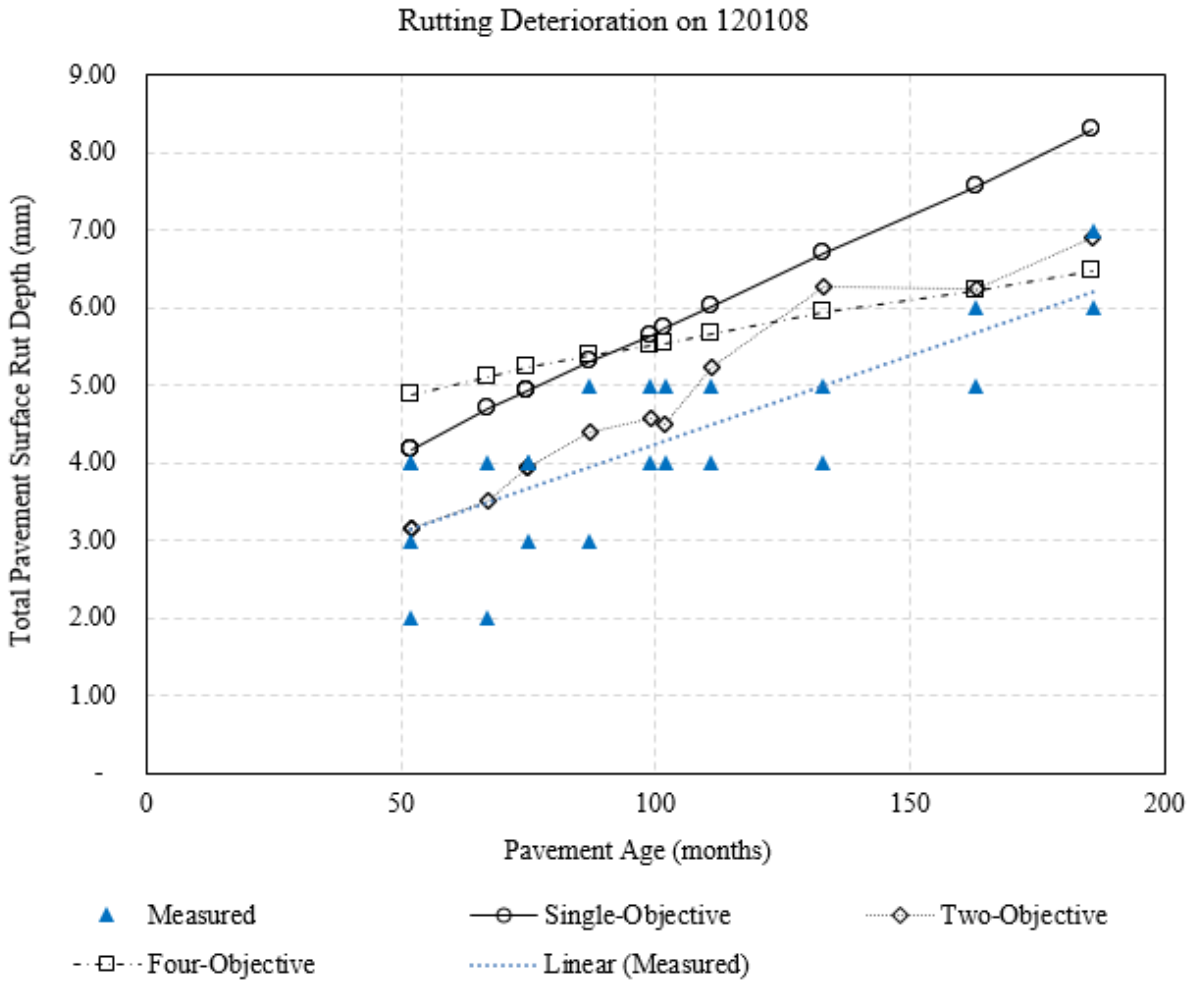
To combine the quantitative and qualitative success metrics of a performance prediction model, the average absolute error (AAE) of the calibrated models in predicting the rate of change in pavement rutting was calculated. The rate of change in rut depth was estimated using measured data and by dividing the change in rutting by the amount of time (months) passed. Only the positive rates were considered in this investigation. Table 37 shows the AAE of the calibrated models in predicting the rate of change in pavement rutting. While the two-objective calibration on SPS-5 data has significantly improved the prediction of rutting deterioration rates compared to single-objective calibration, the multi-objective calibration results on SPS-1 do not exhibit the same quality. This investigation reveals that simply evaluating the bias and STE is not adequate for a comprehensive evaluation of performance prediction models.

Table 37. AAE of calibrated models in predicting the rutting deterioration rates.

Model	AAE in Predicting the Rate of Change in Rutting (%)
New pavement (SPS-1) single-objective	51.93
New pavement (SPS-1) two-objective	60.53
New pavement (SPS-1) four-objective	79.47
Overlaid pavement (SPS-5) single-objective	100.81
Overlaid pavement (SPS-5) Two-objective	66.26

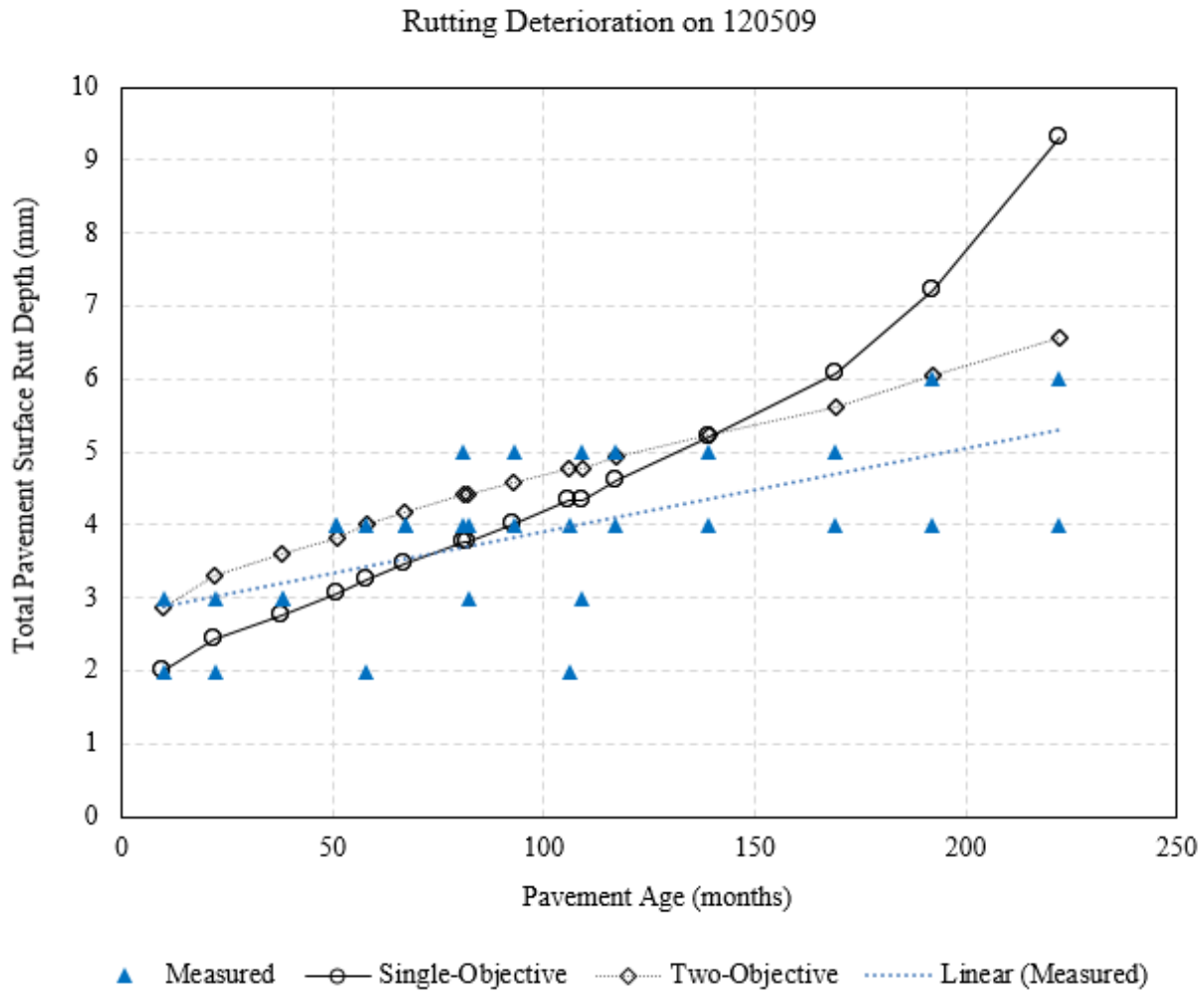
To visualize these deterioration trends, figure 40 and figure 41 demonstrate the measured and predicted rut depth trends on sample test sections of the Florida SPS-1 and SPS-5 sites, respectively. It is evident that on the SPS-1 test section 120108, the two-objective and four-objective calibrations have predicted deterioration trends that are closer to the monitored performance, compared to single-objective calibration results. It should be noted that the predicted rutting data are from the simulated calculations (see chapter 4 for the description of the simulated calculations), and that is why a rutting decrement is observed for the model predictions

from the two-objective calibration around the 100th month. The trends on SPS-5 reveal that the two-objective calibration has predicted deterioration trends that better match the measured values, compared to single-objective calibration results.



Source: FHWA.

Figure 40. Chart. Predicted and measured rutting deterioration on FL SPS-1 section 120108.



Source: FHWA.

Figure 41. Chart. Predicted and measured rutting deterioration on FL SPS-5 section 120509.

CHAPTER 6. SUMMARY OF FINDINGS AND RECOMMENDATIONS

SUMMARY OF FINDINGS

This report presented the results of an FHWA LTPP data analysis project with the objective of using multi-objective optimization to enhance calibration of the MEPDG performance models. The AASHTO recommended single-objective calibration approach was conducted on the MEPDG rutting models for new pavements using the Florida SPS-1 data and overlaid pavements using the Florida SPS-5 data. In the first alternative scenario for multi-objective calibration, SSE and standard deviation of prediction error were simultaneously minimized to reduce the bias and STE (increase model accuracy and precision) at the same time. In the second scenario, the SSE and STE in predicting permanent deformation of SPS-1 and FDOT APT pavement structures were used as separate objective functions to be minimized simultaneously.

Although there was no fundamental way to prove whether there was a theoretical conflict between the selected objective functions, the shape of the final nondominated front indicated that the selected objective functions conflicted with one another, and therefore, the application of a multi-objective optimization approach was justified. In the first multi-objective calibration scenario, the simultaneous minimization of bias and STE resulted in calibrated models that had higher precision (lower STE) and higher generalization capability (lower difference in bias between calibration and validation data), compared to the single-objective calibration. While this scenario was more successful in calibration of rutting models for overlaid pavements on Florida SPS-5 data, it did not result in desirable accuracy levels for rutting models on new pavements using Florida SPS-1 data. The results of the second multi-objective scenario demonstrated that incorporation of the disparate source of performance data (FDOT APT data) as a separate objective function has significantly improved the prediction accuracy, precision, and generalization capability of the calibrated rutting model on SPS-1 data.

The qualitative comparison of the calibrated models showed that using the multi-objective approach has resulted in predicted rutting distributions that are more similar in flatness (kurtosis) to the measured rutting distributions. However, the same was not true about skewness. The low goodness-of-fit indicator for scatterplots of predicted versus measured rutting in the case of all calibration approaches reveals that the MEPDG rutting models have an inherent lack of precision that might not be addressed with the calibration process. This is perhaps because the variability in pavement materials has not been captured in these models. The final selected calibration factors for rutting in unbound pavement layers (base and subgrade) were more reasonable in the multi-objective approach, compared to insignificant values achieved through single-objective calibration. Once again, this possibility of applying engineering judgment demonstrates the value of the multi-objective calibration in providing a final pool of solutions to choose from.

To combine the quantitative and qualitative success metrics of a performance prediction model, the measured and predicted rutting deterioration trends were examined. While the two-objective calibration on SPS-5 data had significantly improved the prediction of rutting deterioration rates compared to single-objective calibration, the multi-objective calibration results on SPS-1 did not exhibit the same quality. This investigation revealed that simply evaluating the bias and STE is not adequate for a comprehensive evaluation of performance prediction models. Therefore, it is

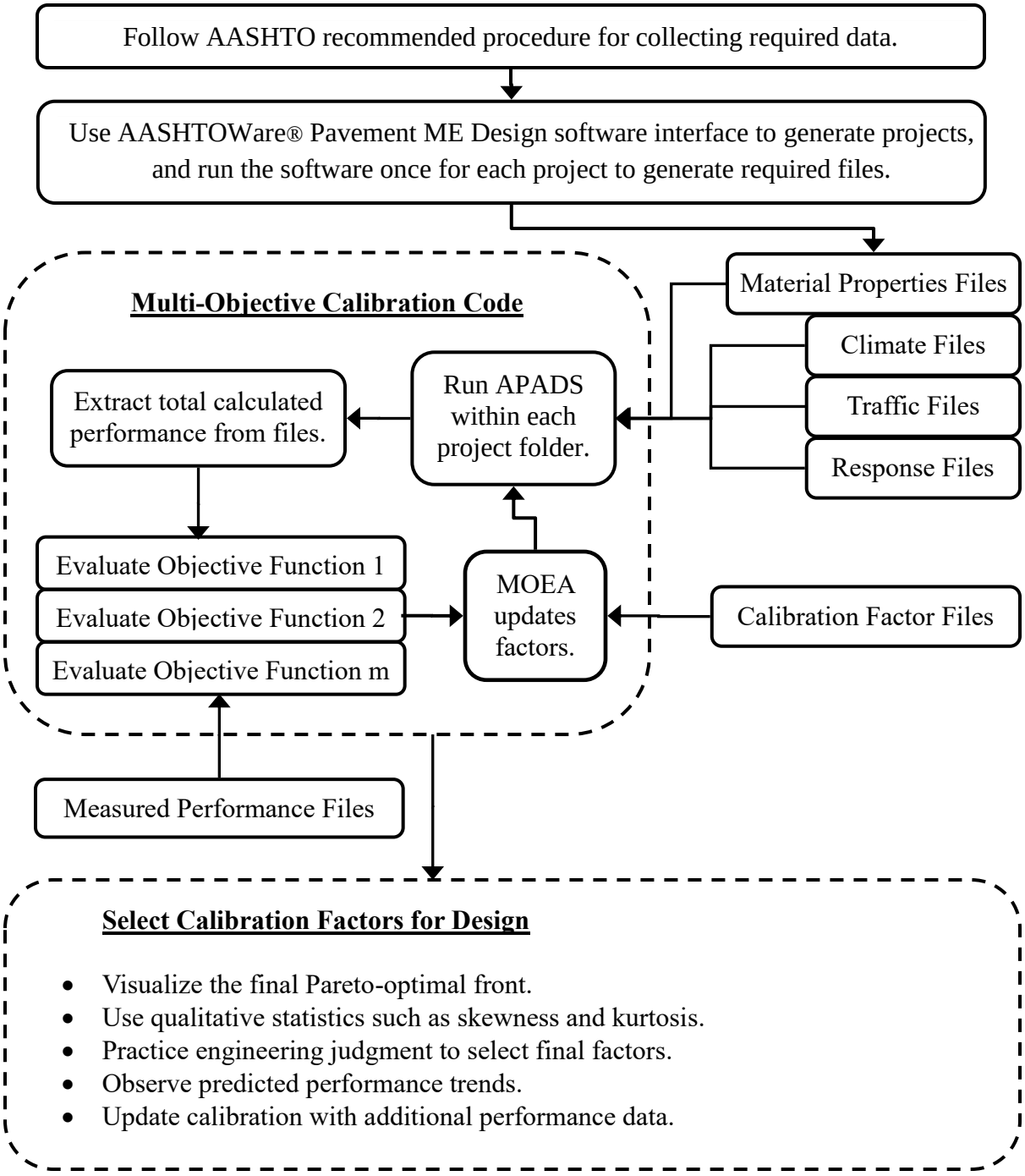
recommended that the comprehensive comparison framework presented in this study be used when selecting suitable performance prediction models.

RECOMMENDATIONS FOR STATE HIGHWAY AGENCIES

Based on the findings of this study, there are several advantages in adopting a multi-objective calibration approach for pavement design within each State or local highway agency:

- The existing trial-and-error approach will be replaced with a systematic framework for calibration of AASHTOWare® Pavement ME Design software.
- Accuracy and precision of the models can be optimized at the same time.
- Multiple sources of performance data with disparate data collection protocols (i.e., LTPP, State PMS, APT, etc.) can be incorporated simultaneously, without the need for reconciliation among different data sources.
- Following the quantitative multi-objective optimization process, engineering judgment and qualitative criteria can be used to select the most reasonable calibration coefficients among the final pool of solutions.

The flowchart in figure 42 demonstrates a potential multi-objective calibration framework that highway agencies can adopt for enhanced performance models.



Source: FHWA.

Figure 42. Flowchart. Framework for implementation of multi-objective calibration.

Using this multi-objective calibration approach results in a final pool of tradeoff solutions. This way, none of the possible sets of calibration factors are eliminated prematurely, and all of the nondominated solutions are included in the final tradeoff front. Exploring the final front might reveal unknown aspects of this calibration problem and result in more reasonable calibration coefficients that could not be identified using single-objective approaches. This study

demonstrates the application of engineering judgment and qualitative criteria to select reasonable calibration coefficients from the final pool of solutions that result from the multi-objective optimization. More reasonable calibration factors result in a more justifiable pavement design considering multiple aspects of pavement performance.

RECOMMENDATIONS FOR FURTHER RESEARCH AND VALIDATION

The following are recommendations for future research into enhancing the multi-objective calibration process:

- Other EAs and parameter settings could be tested to identify the most suitable one for this calibration application. Specifically, the stopping criteria need to be evaluated to make sure a good solution has been achieved. In addition, the precision (epsilon value) corresponding to each objective function needs to be examined to avoid additional computational cost that is beyond the required precision.

- Other multi-objective calibration scenarios could be considered in future research. For example, statistical distribution shape descriptors (skewness and kurtosis, which were used in this study as qualitative comparison criteria) can be used in conjunction with SSE and STE to better match the distribution of measured and calculated rutting data on pavement test sections.

- Other sources of data such as State PMS data could be incorporated into this multi-objective calibration framework to investigate its feasibility in that regard.

- The same multi-objective calibration framework can also be applied to other performance prediction models.

APPENDIX A. DETAILS OF CALIBRATION INPUT DATA

ASPHALT CONCRETE DYNAMIC MODULUS

Computed AC dynamic modulus is available in the LTPP database for the surface layers. However, for asphalt stabilized layers, dynamic modulus values are not included in the LTPP. In those cases, the ANNACAP software was run to obtain the dynamic modulus at five temperatures (–10, 4.4, 21.1, 37.8, and 54.4 ºC) and for four frequencies (0.1, 1, 10, and 25).[64] This software calculates the dynamic modulus based on an ANN algorithm. The input data are the resilient modulus at 5, 25, and 40 ºC and the shift factors. The details of the software and the ANN model are reported by Kim et al.[12] Table 38 shows the resulting estimated dynamic modulus values for SPS-1 and SPS-5 test sections.

Table 38. Dynamic modulus for the Florida SPS-1 and SPS-5 experiment test sections.

Temperature (ºC)	Frequency (Hz)	Dynamic Modulus (MPa) SPS-1 ATB Material Project Layer Code H	Dynamic Modulus (MPa) SPS-5 ID 120503 Material Project Layer Code E
–10.0	25.0	27,598	30,128
–10.0	10.0	26,894	29,534
–10.0	5.0	26,286	29,011
–10.0	1.0	24,589	27,513
–10.0	0.5	23,726	26,730
–10.0	0.1	21,394	24,550
4.4	25.0	21,263	24,424
4.4	10.0	19,721	22,929
4.4	5.0	18,467	21,685
4.4	1.0	15,322	18,458
4.4	0.5	13,905	16,952
4.4	0.1	10,616	13,331
21.1	25.0	10,399	13,086
21.1	10.0	8,628	11,052
21.1	5.0	7,375	9,579
21.1	1.0	4,863	6,529
21.1	0.5	3,977	5,419
21.1	0.1	2,386	3,370
37.8	25.0	3,084	4,279
37.8	10.0	2,284	3,237
37.8	5.0	1,803	2,598
37.8	1.0	1,020	1,535
37.8	0.5	797	1,224
37.8	0.1	454	736
54.4	25.0	817	1,253
54.4	10.0	591	933
54.4	5.0	465	752
54.4	1.0	275	473
54.4	0.5	223	394
54.4	0.1	144	272

ASPHALT CONCRETE CREEP COMPLIANCE

MEPDG at input level 1 requires the AC creep compliance modulus at three temperatures (–20, –10, and 10 ºC) and seven loading frequencies (1, 2, 5, 10, 20, 50, and 100 s). The LTPP creep compliance test was done under a different set of temperatures (–10, 5, and 25 ºC) but similar loading frequencies to the MEPD input values. Therefore, a conversion process that is suggested in the literature is applied to generate the master curve and then fit a generalized viscoelastic model.(65) The log time–temperature shift factor is considered linearly correlated with the temperature. The sections included in the analysis are located in areas where the yearly lowest temperature is above 0 ºC. Therefore, the creep compliance conversion was not performed, and input level 3 was considered for the MEPDG analysis. However, for some other locations, where the temperature falls below 0 ºC, the conversion process needs to be applied.

The conversion process begins with the calculation of the reduced time that can be calculated using equation 48:

$$t_T = t_{T_0} e^{[2.3026\beta(T-T_0)]} \tag{48}$$

Where:

t_T = time to obtain the creep compliance at temperature T.
t_{T_0} = time to obtain the creep compliance at temperature T_0 (1, 2, 5, 10, 20, 50, and 100 s).
β = slope of the straight line of the relationship between the time–temperature shift factor and the temperature, which is calculated using equation 49:

$$\beta = \frac{\log\left(\frac{t_T}{t_{T_0}}\right)}{T - T_0} \tag{49}$$

Where:

T = temperature.
T_0 = reference temperature.

The generalized model can be composed by a number of connected Kelvin models. The creep compliance model for a constant stress condition is expressed as in equation 50:

$$D(t) = G_0 + \sum_{i=1}^{n} G_i \left(1 - e^{\left(-\frac{t_T}{T_i}\right)}\right) \tag{50}$$

Where:

$D(t)$ = creep compliance at time t.
G_i = creep compliance coefficients.
n = number of Kelvin models.
T_i = time durations to cover the range of interest.

Creep compliance coefficients are calculated solving the simultaneous equations as in equation 51:

$$\{G_i\} = \left[[E]^T[E]\right]^{-1}[E]^T\{D_i\} \quad (51)$$

Where:

$\{G_i\}$ = vector of creep compliance coefficients.

$\{D_i\}$ = vector of creep compliance.

$[E]$ = creep compliance matrix as in equation 52:

$$[E] = \begin{bmatrix} 1 - e^{\left(-\frac{t_{T1}}{T_1}\right)} & \dots & 1 - e^{\left(-\frac{t_{T1}}{T_n}\right)} \\ \vdots & \ddots & \vdots \\ 1 - e^{\left(-\frac{t_{Tm}}{T_1}\right)} & \dots & 1 - e^{\left(-\frac{t_{Tm}}{T_n}\right)} \end{bmatrix} \quad (52)$$

RESILIENT MODULUS FOR UNBOUND AND SUBGRADE LAYERS

Table 39 shows the available resilient modulus values for unbound materials in the LTPP database.

Table 39. Calculated resilient modulus of unbound materials.

Sample Code	k_1	k_2	k_3	M_r (psi) Subgrade	M_r (psi) Granular
BS05	685.27	0.52	0.02	7,902	12,928
BG03	1,178.20	0.73	N/A	13,050	25,584
BS04	752.32	0.48	N/A	9,209	14,234
BS02	737.88	0.44	N/A	9,164	13,685
BS01	700.22	0.60	N/A	8,165	14,163
BGX01	968.66	0.38	0.37	5,747	11,399
BGX02	1,215.13	0.60	N/A	14,183	24,549
BS03	735.40	0.53	N/A	8,817	14,318
BS06	714.43	0.58	N/A	8,391	14,308
BGX01	1,099.94	0.61	0.00	12,766	22,382
BGX06	1,019.99	0.74	N/A	11,284	22,185
BGX05	1,274.99	0.64	N/A	14,640	26,345
BGX04	1,011.04	0.48	0.12	9,667	16,751
BS57	766.96	0.49	N/A	9,340	14,614
TS01	791.19	0.70	N/A	8,895	16,830
BS55	1,229.38	0.52	N/A	14,789	23,824
BG**	1,318.69	0.53	N/A	15,777	25,747
BG56	1,188.30	0.56	N/A	14,105	23,456
BG**	889.01	0.72	N/A	9,919	19,111
BG55	1,117.46	0.68	N/A	12,655	23,535
BS**	848.57	0.89	N/A	8,855	20,005
BG**	909.19	0.66	N/A	10,365	18,973
BG56	825.94	0.62	N/A	9,549	16,907
BG**	543.51	0.77	0.92	901	4,221
BG55	1,131.11	0.68	N/A	12,772	23,919
TS01	854.43	0.50	N/A	10,361	16,375
BS**	1,208.62	0.38	N/A	15,321	21,790
TS03	982.82	0.23	N/A	13,234	16,304
BG**	804.25	0.60	N/A	9,368	16,294

Sample Code	k_1	k_2	k_3	M_r (psi) Subgrade	M_r (psi) Granular
TS01	751.50	0.58	N/A	8,850	14,997
BS55	965.17	0.46	N/A	11,872	18,139
BG56	1,000.85	0.53	N/A	11,981	19,528
BG**	706.06	0.78	0.10	6,232	13,966
BG55	1,253.38	0.68	N/A	14,153	26,502
BS**	1,291.75	0.52	N/A	15,515	25,087
TS03	838.60	0.55	N/A	9,980	16,493
BG**	1,416.59	0.49	N/A	17,245	27,006
BG56	1,347.88	0.55	N/A	16,004	26,596
BG55	1,381.21	0.51	N/A	16,698	26,583
BS**	1,028.25	0.46	0.32	6,546	13,401
TS02	513.33	0.62	0.78	1,206	4,338
BG**	747.88	0.70	N/A	8,394	15,947
BG56	773.47	0.69	N/A	8,726	16,376
BG**	1,315.23	0.55	N/A	15,636	25,906
BG55	2,052.86	0.56	N/A	24,285	40,712
TS01	683.52	0.17	0.81	1,776	4,357
TS04	723.25	0.56	0.64	2,319	6,939
TS01	1,124.28	0.57	N/A	13,290	22,319
TS03	1,229.06	0.65	N/A	14,088	25,456
BG**	1,296.16	0.55	N/A	15,401	25,548
TS01	824.28	0.61	0.23	5,946	12,863
BS**	834.87	0.35	N/A	10,717	14,794
BG**	1,743.84	0.44	N/A	21,642	32,371
TS01	662.98	0.62	0.48	2,871	7,869
TS03	753.00	0.75	0.14	6,197	14,000
BG**	979.84	0.53	N/A	11,723	19,132
BG55	697.57	0.68	N/A	7,905	14,679
TS01	816.71	0.38	0.25	6,172	11,011
TS03	1,043.36	0.35	N/A	13,411	18,454
BS**	1,109.58	0.57	N/A	13,109	22,044
BS55	1,190.33	0.51	0.42	6,034	14,212
BG**	1,165.05	0.59	N/A	13,634	23,453
BG56	1,257.62	0.56	N/A	14,893	24,903
BG**	1,231.76	0.55	N/A	14,645	24,258
TS01	556.69	0.44	0.67	1,744	4,829
TS03	606.55	0.52	0.58	2,233	6,120
BG**	953.39	0.55	N/A	11,329	18,791
BG55	968.58	0.56	N/A	11,474	19,171
TS01	489.95	0.43	0.69	1,478	4,124
TS03	508.58	0.49	0.77	1,273	4,050
BG**	1,175.50	0.57	N/A	13,876	23,380
BG56	928.50	0.62	N/A	10,750	18,967
BG**	1,599.55	0.42	N/A	20,037	29,314
BG55	1,265.45	0.53	N/A	15,176	24,628
BS**	703.77	0.92	1.27	533	3,960
BS55	922.19	0.55	0.70	2,609	8,210
BG**	1,546.27	0.65	N/A	17,701	32,084
BG56	1,594.46	0.49	N/A	19,387	30,447
BG**	405.93	0.97	N/A	4,100	10,009
BG55	547.74	1.03	N/A	5,414	13,916
BS**	1,014.08	0.54	N/A	12,120	19,829

Sample Code	k_1	k_2	k_3	M_r (psi) Subgrade	M_r (psi) Granular
BG**	1,210.00	0.57	N/A	14,267	24,105
TS01	794.16	0.69	0.53	2,994	9,159
TS03	878.04	0.60	0.46	3,948	10,447
BG**	1,062.14	0.76	0.15	8,535	19,694
BG56	1,055.31	0.60	N/A	12,308	21,341
BG**	1,372.03	0.58	N/A	16,131	27,441
BG55	1,403.67	0.66	N/A	15,981	29,346
TS01	751.00	0.53	0.62	2,507	7,189
BG**	1,157.71	0.66	N/A	13,190	24,180
BG55	993.02	0.57	N/A	11,741	19,707
BS**	N/A	0.90	0.89	N/A	N/A
TS03	N/A	0.90	0.89	N/A	N/A
BG**	N/A	0.33	0.58	N/A	N/A
BS**	N/A	0.90	0.89	N/A	N/A
BS55	N/A	0.90	0.89	N/A	N/A
BG**	N/A	0.46	0.71	N/A	N/A
BG56	N/A	0.46	0.71	N/A	N/A
BG**	N/A	0.46	0.71	N/A	N/A
BG55	N/A	0.33	0.58	N/A	N/A
TS01	0.00	0.93	0.87	0.00	0.00
BS55	920.88	0.50	N/A	11,153	17,680
BG**	N/A	0.33	0.58	N/A	N/A
BG56	N/A	0.33	0.58	N/A	N/A
BG**	N/A	0.33	0.58	N/A	N/A
BG55	N/A	0.33	0.58	N/A	N/A
TS01	N/A	0.93	0.87	N/A	N/A
TS03	N/A	0.92	0.87	N/A	N/A
BG**	N/A	0.33	0.58	N/A	N/A
BG56	N/A	0.38	0.63	N/A	N/A
BG**	N/A	0.38	0.63	N/A	N/A
BG55	N/A	0.38	0.63	N/A	N/A
TS01	N/A	0.90	0.89	N/A	N/A
BS**	N/A	0.86	0.89	N/A	N/A
TS01	N/A	0.90	0.89	N/A	N/A
TS03	N/A	0.90	0.89	N/A	N/A
BG**	1,324.28	0.50	N/A	16,052	25,396
BG55	1,272.95	0.47	N/A	15,593	24,059
TS01	N/A	0.90	0.89	N/A	N/A
TS02	N/A	0.92	0.88	N/A	N/A
BG**	3.77	3.58	0.82	3	145
BG56	4.28	3.46	0.84	3	153
BG**	0.05	6.44	0.84	0	8
BG55	746.38	0.99	N/A	7,498	18,543

N/A = no adequate data.

RUTTING DATA

SPS-1

Table 40, table 41, and table 42 list the average rut depth values measured through the monitoring period for every test section on the Florida SPS-1 site.

Table 40. Average measured rut depth for Florida SPS-1 test sections 120107 to 120111.

120107 Date	120107 Rut (mm)	120108 Date	120108 Rut (mm)	120106 Date	120106 Rut (mm)	120110 Date	120110 Rut (mm)	120111 Date	120111 Rut (mm)
2/9/2000	4	2/9/2000	4	2/9/2000	5	2/9/2000	4	2/9/2000	4
2/17/2000	3	2/17/2000	4	2/17/2000	3	2/17/2000	3	2/16/2000	4
5/10/2001	4	5/10/2001	4	5/10/2001	4	5/10/2001	4	5/9/2001	4
1/18/2002	5	1/18/2002	4	1/17/2002	5	1/17/2002	5	1/17/2002	5
1/21/2002	5	1/21/2002	4	1/21/2002	4	1/21/2002	4	1/21/2002	6
1/23/2003	5	1/23/2003	5	1/23/2003	5	1/23/2003	6	1/22/2003	6
1/22/2004	6	1/22/2004	5	1/22/2004	5	1/22/2004	5	1/22/2004	4
4/22/2004	6	4/22/2004	5	4/22/2004	6	4/22/2004	6	4/21/2004	6
1/19/2005	7	1/19/2005	5	1/19/2005	5	1/18/2005	6	1/18/2005	7
11/9/2006	8	11/9/2006	5	11/9/2006	6	11/9/2006	7	11/9/2006	7
5/8/2009	10	5/8/2009	6	5/8/2009	7	5/7/2009	8	5/7/2009	8
4/4/2011	12	4/4/2011	7	4/4/2011	9	4/4/2011	9	3/30/2011	9

Table 41. Average measured rut depth for Florida SPS-1 test sections 120112 to 120105.

120112 Date	120112 Rut (mm)	120109 Date	120109 Rut (mm)	120104 Date	120104 Rut (mm)	120103 Date	120103 Rut (mm)	120105 Date	120105 Rut (mm)
2/9/2000	3	2/9/2000	4	2/9/2000	4	2/9/2000	6	2/9/2000	5
2/16/2000	3	2/16/2000	3	2/16/2000	4	2/16/2000	5	2/15/2000	5
5/9/2001	3	5/9/2001	3	5/9/2001	4	5/9/2001	6	5/9/2001	5
1/16/2002	4	1/16/2002	3	1/16/2002	5	1/16/2002	7	1/15/2002	6
1/21/2002	6	1/21/2002	4	1/21/2002	5	1/21/2002	6	1/21/2002	6
1/22/2003	4	1/22/2003	3	1/22/2003	5	1/21/2003	7	1/21/2003	6
1/22/2004	4	1/22/2004	5	1/22/2004	4	1/22/2004	6	1/22/2004	4
4/21/2004	4	4/21/2004	4	4/21/2004	6	4/20/2004	7	4/20/2004	6
1/18/2005	5	1/18/2005	4	1/17/2005	6	1/17/2005	8	1/17/2005	6
11/8/2006	5	11/8/2006	4	11/8/2006	6	11/8/2006	8	11/8/2006	7
5/7/2009	6	5/7/2009	5	5/6/2009	8	5/6/2009	10	5/6/2009	8
3/30/2011	7	3/30/2011	6	3/30/2011	9	3/29/2011	11	3/29/2011	9

Table 42. Average measured rut depth for Florida SPS-1 test sections 120101 to 120161.

120101 Date	120101 Rut (mm)	120102 Date	120102 Rut (mm)	120161 Date	120161 Rut (mm)
2/9/2000	4	2/9/2000	3	2/9/2000	4
2/15/2000	4	2/15/2000	3	2/15/2000	3
5/8/2001	4	5/8/2001	3	5/8/2001	3
1/15/2002	5	1/15/2002	3	1/15/2002	3
1/21/2002	5	1/21/2002	4	1/21/2002	4
1/21/2003	5	1/21/2003	3	1/21/2003	3
1/22/2004	5	1/22/2004	4	1/22/2004	4
4/20/2004	6	4/20/2004	4	4/20/2004	3
1/17/2005	6	1/17/2005	4	1/14/2005	4
11/3/2006	7	11/3/2006	4	11/3/2006	4
5/6/2009	8	5/6/2009	5	N/A	N/A
3/29/2011	9	3/29/2011	5	N/A	N/A

N/A = no adequate data.

SPS-5

Table 43, table 44, and table 45 list the average rut depth values measured through the monitoring period for every test section on the Florida SPS-5 site.

Table 43. Average measured rut depth for Florida SPS-5 test sections 120502 to 120565.

120502 Date	120502 Rut (mm)	120561 Date	120561 Rut (mm)	120503 Date	120503 Rut (mm)	120508 Date	120508 Rut (mm)	120565 Date	120565 Rut (mm)
1/21/1996	4	1/21/1996	3	1/21/1996	4	1/21/1996	3	1/21/1996	4
1/21/1997	6	1/21/1997	5	1/21/1997	4	1/21/1997	4	1/21/1997	4
5/18/1998	4	5/18/1998	5	5/18/1998	5	5/18/1998	4	5/19/1998	4
7/14/1999	4	7/14/1999	6	7/14/1999	5	7/15/1999	4	7/15/1999	5
2/9/2000	5	2/9/2000	4	2/9/2000	4	2/9/2000	4	2/9/2000	4
11/1/2000	6	11/1/2000	7	11/1/2000	7	11/1/2000	5	11/1/2000	5
1/10/2002	6	1/9/2002	7	1/9/2002	6	1/9/2002	5	1/9/2002	6
1/21/2002	5	1/21/2002	6	1/21/2002	6	1/21/2002	4	1/21/2002	5
1/15/2003	6	1/15/2003	7	1/15/2003	7	1/15/2003	5	1/15/2003	6
1/23/2004	3	1/23/2004	5	1/23/2004	5	1/23/2004	4	1/23/2004	5
4/26/2004	6	4/26/2004	8	4/26/2004	6	4/26/2004	5	4/26/2004	6
1/11/2005	6	1/11/2005	8	1/11/2005	7	1/11/2005	5	1/11/2005	6
11/6/2006	6	11/6/2006	8	11/6/2006	7	11/6/2006	5	11/6/2006	6
5/4/2009	6	10/1/2013	9	5/4/2009	7	5/4/2009	5	11/7/2006	6
3/31/2011	7	N/A	N/A	3/31/2011	8	3/31/2011	6	10/1/2013	7
10/1/2013	7	N/A	N/A	10/1/2013	8	10/1/2013	6	10/1/2013	6

N/A = no adequate data.

Table 44. Average measured rut depth for Florida SPS-5 test sections 120509 to 120504.

120509 Date	120509 Rut (mm)	120506 Date	120506 Rut (mm)	120566 Date	120566 Rut (mm)	120507 Date	120507 Rut (mm)	120504 Date	120504 Rut (mm)
1/21/1996	3	1/21/1996	3	1/21/1996	3	1/21/1996	4	1/21/1996	4
1/21/1997	3	1/21/1997	3	1/22/1997	4	1/22/1997	5	1/22/1997	4
5/18/1998	3	5/18/1998	4	5/18/1998	4	5/19/1998	5	5/19/1998	4
7/15/1999	4	7/15/1999	4	7/15/1999	3	7/15/1999	5	7/15/1999	4
2/9/2000	4	2/9/2000	3	2/9/2000	3	2/9/2000	5	2/9/2000	4
11/2/2000	4	11/2/2000	4	11/2/2000	4	11/2/2000	5	11/2/2000	4
1/9/2002	5	1/10/2002	4	1/10/2002	4	1/10/2002	5	1/10/2002	4
1/21/2002	4	1/21/2002	3	1/21/2002	3	1/21/2002	5	1/21/2002	4
1/15/2003	5	1/15/2003	4	1/16/2003	4	1/16/2003	5	1/16/2003	5
1/23/2004	4	1/23/2004	4	1/23/2004	4	1/23/2004	4	1/23/2004	3
4/26/2004	5	4/26/2004	4	4/27/2004	4	4/27/2004	5	4/27/2004	5
1/11/2005	5	1/11/2005	4	1/13/2005	4	1/13/2005	6	1/12/2005	5
11/6/2006	5	11/7/2006	4	11/7/2006	5	11/7/2006	6	11/7/2006	5
5/4/2009	5	5/4/2009	4	10/2/2013	5	5/4/2009	6	5/5/2009	6
4/1/2011	6	3/31/2011	4	1/29/2014	5	3/31/2011	6	4/1/2011	6
10/1/2013	6	10/2/2013	5	3/31/2011	4	10/2/2013	6	N/A	N/A
N/A	N/A	N/A	N/A	10/2/2013	5	10/2/2013	5	N/A	N/A

N/A = no adequate data.

Table 45. Average measured rut depth for Florida SPS-5 test sections 120562 to 120564.

120562 Date	120562 Rut (mm)	120505 Date	120505 Rut (mm)	120563 Date	120563 Rut (mm)	120564 Date	120564 Rut (Mm)
1/21/1996	4	1/21/1996	3	1/21/1996	3	1/21/1996	3
1/22/1997	4	1/22/1997	4	1/22/1997	4	1/23/1997	5
5/19/1998	5	5/19/1998	4	5/19/1998	4	5/19/1998	5
7/19/1999	3	7/19/1999	3	7/19/1999	3	7/19/1999	4
2/9/2000	3	2/9/2000	3	2/9/2000	3	2/9/2000	4
11/6/2000	4	11/6/2000	4	11/6/2000	4	11/6/2000	5
1/10/2002	4	1/14/2002	4	1/14/2002	4	1/14/2002	5
1/21/2002	4	1/21/2002	4	1/21/2002	4	1/21/2002	4
1/16/2003	4	1/16/2003	4	1/16/2003	4	1/16/2003	5
1/23/2004	4	1/23/2004	6	1/23/2004	6	1/23/2004	4
4/27/2004	4	4/27/2004	4	4/27/2004	4	4/28/2004	6
1/12/2005	4	1/12/2005	5	1/12/2005	4	1/12/2005	5
11/7/2006	5	11/7/2006	5	11/7/2006	5	11/7/2006	6
10/2/2013	7	5/5/2009	5	10/2/2013	6	10/3/2013	6
N/A	N/A	4/1/2011	6	N/A	N/A	N/A	N/A
N/A	N/A	10/3/2013	6	N/A	N/A	N/A	N/A

N/A = no adequate data.

APT ARB

Table 46 shows the measured rut depth on the APT ARB experiment sections with HVS passes.

Table 46. Average measured rut depth (mm) for FDOT ARB experiment sections.

HVS Passes	1 PG 76-22 PMA	2 ARB-5	3 30% RAP	4 Hybrid B	5 Hybrid A-L	6 GTR-C	7 Hybrid A-H
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100	0.37	1.37	0.45	0.95	0.92	0.54	1.18
300	0.76	1.85	0.71	1.22	1.32	0.77	1.66
500	1.15	2.13	0.93	1.37	1.55	0.90	1.91
700	1.23	2.32	1.12	1.51	1.67	1.04	2.09
1,000	1.48	2.53	1.30	1.63	1.83	1.19	2.25
1,300	1.71	2.78	1.45	1.74	1.96	1.26	2.41
1,600	1.84	2.92	1.58	1.84	2.07	1.32	2.49
2,000	1.99	3.13	1.71	1.91	2.18	1.42	2.60
2,500	2.16	3.25	1.84	1.99	2.32	1.54	2.68
3,000	2.30	3.34	1.94	2.06	2.42	1.56	2.69
3,500	2.43	3.53	2.03	2.12	2.51	1.66	2.74
4,000	2.54	3.57	2.10	2.19	2.59	1.71	2.91
4,500	2.68	3.69	2.18	2.25	2.68	1.78	3.03
5,000	2.76	3.78	2.26	2.28	2.78	1.83	3.10
6,000	2.95	3.93	2.37	2.36	2.91	1.88	3.21
7,000	3.12	4.06	2.46	2.42	3.02	1.95	3.27
8,000	3.25	4.19	2.55	2.49	3.09	2.00	3.31
9,000	3.39	4.27	2.59	2.53	3.22	2.06	3.38
10,000	3.50	4.35	2.66	2.59	3.27	2.11	3.44
11,000	3.62	4.45	2.71	2.62	3.33	2.14	3.51
12,000	3.72	4.49	2.77	2.70	3.40	2.22	3.54
13,000	3.77	4.57	2.83	2.74	3.45	2.24	3.61

HVS Passes	1 PG 76-22 PMA	2 ARB-5	3 30% RAP	4 Hybrid B	5 Hybrid A-L	6 GTR-C	7 Hybrid A-H
14,000	3.87	4.62	2.89	2.76	3.51	2.28	3.65
15,000	3.92	4.67	2.94	2.77	3.56	2.30	3.68
16,000	3.99	4.73	2.98	2.81	3.59	2.34	3.75
18,000	4.11	4.83	3.05	2.85	3.67	2.38	3.72
20,000	4.20	4.92	3.10	2.89	3.75	2.42	3.83
22,000	4.33	5.01	3.22	2.93	3.83	2.46	3.89
24,000	4.42	5.09	3.25	2.97	3.88	2.56	3.92
26,000	4.48	5.21	3.34	3.01	3.93	2.64	3.97
28,000	4.58	5.28	3.41	3.05	4.00	2.64	4.01
32,000	4.70	5.41	3.57	3.11	4.06	2.69	4.03
36,000	4.84	5.52	3.67	3.18	4.19	2.76	4.10
38,000	4.88	5.64	3.69	3.21	4.22	2.80	4.15
40,000	4.94	5.73	3.71	3.24	4.23	2.83	4.20
45,000	5.06	5.73	3.84	3.32	4.34	2.92	4.26
50,000	5.18	5.84	3.99	3.41	4.43	3.00	4.32
55,000	5.24	5.99	4.07	3.49	4.49	3.08	4.40
60,000	5.35	6.02	4.17	3.54	4.57	3.19	4.38
65,000	5.47	6.07	4.27	3.59	4.65	3.23	4.48
70,000	5.54	6.16	4.36	3.65	4.68	3.30	4.58
80,000	5.67	6.32	4.50	3.74	4.86	3.39	4.69
90,000	5.82	6.44	4.64	3.84	4.97	3.47	4.80
100,000	5.93	6.58	4.76	3.92	5.09	3.57	4.97

APT DASR

Table 47 shows the measured rut depth on the APT DASR experiment sections with HVS passes.

Table 47. Average measured rut depth (mm) for FDOT DASR experiment sections.

HVS Passes	DASR Porosity 44%	HVS Passes	DASR Porosity 42%	HVS Passes	DASR Porosity 49%	HVS Passes	DASR Porosity 56%
0	0.00	0	0.00	0	0.00	0	0.00
100	1.43	100	1.45	100	2.12	100	2.87
300	2.11	300	2.88	300	3.23	300	4.70
500	2.57	500	3.54	500	3.91	500	5.95
700	2.91	700	4.01	700	4.43	700	7.06
1,000	3.35	1,000	4.71	1,000	4.95	1,000	8.32
1,300	3.68	1,600	5.14	1,300	5.44	1,300	9.48
1,423	3.81	2,000	5.64	1,600	5.82	1,600	10.50
1,600	3.98	3,000	6.10	2,000	6.26	2,000	11.80
2,000	4.30	3,500	6.47	2,500	6.73	2,500	13.22
2,500	4.73	2,500	6.49	3,000	7.16	3,000	14.57
3,000	5.03	4,000	6.99	3,500	7.48	3,500	15.75
3,500	5.29	4,500	7.26	4,000	7.79	4,000	16.82
4,000	5.61	5,000	7.51	4,500	8.06	4,500	17.83
4,500	5.72	7,000	8.02	5,000	8.33	5,000	18.79
5,000	6.08	8,000	8.39	6,000	8.75	6,000	20.59
6,000	6.36	9,000	8.59	7,000	9.07	7,000	22.24
7,000	6.60	10,000	8.88	8,000	9.37	8,000	23.74
8,000	6.77	11,000	8.95	9,000	9.65	9,000	25.20

HVS Passes	DASR Porosity 44%	HVS Passes	DASR Porosity 42%	HVS Passes	DASR Porosity 49%	HVS Passes	DASR Porosity 56%
9,000	6.95	12,000	9.27	10,000	9.91	10,000	26.33
10,000	7.14	13,000	9.43	11,000	10.14	12,000	28.46
11,000	7.31	14,000	9.61	12,000	10.33	13,000	29.59
12,000	7.39	15,000	9.83	13,000	10.61	22,000	34.52
13,000	7.61	16,000	9.96	14,000	10.72	24,000	36.21
14,000	7.71	18,000	10.11	15,000	11.03	N/A	N/A
15,000	7.83	20,000	10.37	16,000	11.07	N/A	N/A
16,000	7.97	22,000	10.59	18,000	11.34	N/A	N/A
18,000	8.11	24,000	10.71	20,000	11.56	N/A	N/A
20,000	8.23	26,000	10.89	22,000	11.82	N/A	N/A
22,000	8.31	28,000	11.09	24,000	12.09	N/A	N/A
24,000	8.46	32,000	11.29	26,000	12.33	N/A	N/A
26,000	8.59	36,000	11.57	27,546	12.42	N/A	N/A
28,000	8.62	38,000	11.69	28,000	12.45	N/A	N/A
32,000	8.95	40,000	11.88	32,000	12.80	N/A	N/A
36,000	9.10	45,000	12.08	34,044	13.02	N/A	N/A
38,000	9.18	50,000	12.34	36,000	13.28	N/A	N/A
40,000	9.25	55,000	12.51	38,000	13.51	N/A	N/A
45,000	9.43	60,000	12.78	40,000	13.61	N/A	N/A
50,000	9.57	65,000	12.89	N/A	N/A	N/A	N/A
55,000	9.69	N/A	N/A	N/A	N/A	N/A	N/A
60,000	9.84	N/A	N/A	N/A	N/A	N/A	N/A
65,000	9.90	N/A	N/A	N/A	N/A	N/A	N/A
70,000	9.98	N/A	N/A	N/A	N/A	N/A	N/A
75,000	10.10	N/A	N/A	N/A	N/A	N/A	N/A
79,788	10.20	N/A	N/A	N/A	N/A	N/A	N/A
80,000	10.21	N/A	N/A	N/A	N/A	N/A	N/A
90,000	10.31	N/A	N/A	N/A	N/A	N/A	N/A
95,000	10.35	N/A	N/A	N/A	N/A	N/A	N/A

N/A = no adequate data.

APPENDIX B. PROGRAMMING CODES

In this research project, MOEAs have been implemented to optimize the multiple objective functions involved. MOEA Framework (moeaframework.org) is a free and open-source Java framework for multi-objective optimization using a variety of EAs, including GAs and ESs. In this object-oriented framework, an instance of the "abstract problem" class needs to be created, in which the calculation process for the multiple objective functions is implemented. Then an instance of the "problem execution" class is created, where the abstract multi-objective optimization problem is solved using a selected "algorithm." The user should employ an integrated development environment (IDE) to add the "MOEAFramework-2.10.jar" to the build path before running the codes discussed in this appendix.

The following sections in this appendix describe the source codes developed for this study. In addition, the "ReadFromFile.java" and "WriteToFile.java" source codes were developed to read the calibration data and factors and to write the rutting calculation results and final calibration factors into corresponding files.

To use these codes, the user needs to store the calibration and validation datasets in two separate folders and provide the folder path in the "problem execution" code. Depending on the number of objective functions, different source codes should be used as explained below. The calibration dataset comprises one folder per test section, which contains both the measured rutting through time and the intermediate pavement response files calculated using the AASHTOWare® Pavement ME Design software. The validation dataset only includes measured rutting in each folder for each test section.

All the noted source codes and the calibration and validation data files used for this study are available online through https://www.fhwa.dot.gov/.

PROBLEM EXECUTION: SOLVING THE MULTI-OBJECTIVE OPTIMIZATION PROBLEM

This code is the main code that needs to be executed. An executable file was not created for this project because this was a research experiment, and the developed code does not include a front-end user interface. Therefore the "problem execution" code needs to be run through an IDE. The user needs to make sure that all the other corresponding source codes are located in the same folder and that the folder path for the calibration and validation datasets is specified within this source code. Table 48 shows the names of the "problem execution" codes and the corresponding "problem" codes for the calculation of various numbers of objective functions.

Table 48. Developed source codes for multi-objective calibration of MEPDG rutting models.

Problem To Be Solved	Problem Execution Code	Rutting Problem Objective Functions Code
AASHTO calibration method for new pavements on SPS-1	ProblemExecution_SPS1_ 1Obj.java	RuttingProblem_SPS1_ 1Obj.java
Two-objective calibration for new pavements on SPS-1	ProblemExecution_SPS1_ 2Obj.java	RuttingProblem_SPS1_ 2Obj.java
Four-objective calibration for new pavements on SPS-1 and FDOT APT	ProblemExecution_SPS1_ 4Obj.java	RuttingProblem_SPS1_ 4Obj.java
AASHTO calibration method for overlaid pavements on SPS-5	ProblemExecution_SPS5_ 1Obj.java	RuttingProblem_SPS5_ 1Obj.java
Two-objective calibration for overlaid pavements on SPS-5	ProblemExecution_SPS5_ 2Obj.java	RuttingProblem_SPS5_ 2Obj.java

RUTTING CALIBRATION PROBLEM: SETTING UP PROBLEM SOLUTIONS AND OBJECTIVE FUNCTIONS

As noted above, an instance of the "abstract problem" class was created, in which the calculation process for the multiple objective functions is implemented. Table 48 includes the file names for the Java codes created for rutting problems according to the pavement type (new AC models calibrated on SPS-1 and overlaid AC models calibrated on SPS-5 data) and the number of objective functions used (single-objective calibration according to the AASHTO method, two-objective minimization of bias and STE, or four-objective optimization using LTPP and APT data simultaneously). These source codes set up the problem solutions (calibration factors) and the way in which the various objective functions are calculated. These source codes are called from the corresponding problem execution codes, and they in turn call for the source code that provides total pavement deformation estimates.

CLASS FOR DEFORMATION CALCULATIONS

The source code called TotalDeformation.java adds the cumulative rutting developed in all pavement layers (including asphalt-bound and unbounded materials) up to each month through pavement age. This source code calls the corresponding codes for calculating the cumulative rutting in AC and unbounded layers. This source code is called by the problem codes to provide the specific calculations of objective functions.

CALCULATION OF CUMULATIVE RUTTING IN ASPHALT CONCRETE LAYERS

The source code called CumulativeAC.java provides a simulated process for calculating the amount of permanent deformation accumulated in asphalt-bound layers through time and under specific traffic and climatic conditions using the intermediate pavement response files provided by the AASHTOWare® Pavement ME Design software. The detailed process is described in chapter 4, under the heading "Simulating Permanent Deformation in Asphalt Concrete Layers." This source code is called by the source code for calculating the total pavement deformation.

CALCULATION OF CUMULATIVE RUTTING IN UNBOUND LAYERS

The source code called CumulativeUnbounded.java provides a simulated process for calculating the amount of permanent deformation accumulated in unbounded layers through time and under specific traffic and climatic conditions using the intermediate pavement response files provided by the AASHTOWare® Pavement ME Design software. The detailed process is described in chapter 4, under the heading “Simulating Permanent Deformation in Unbound Materials.” This source code is called by the source code for calculating the total pavement deformation.

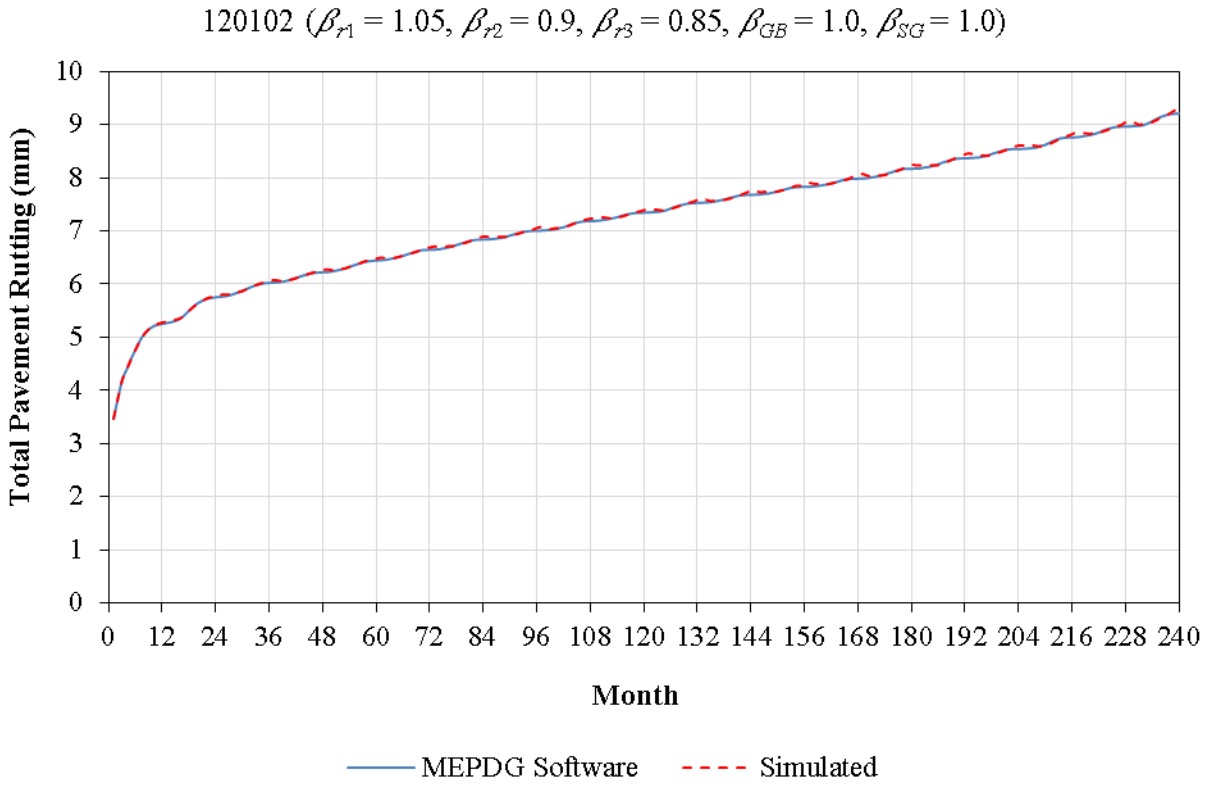
APPENDIX C. COMPARISON OF SIMULATED RUTTING CALCULATIONS TO ME SOFTWARE RESULTS

This appendix presents example comparisons between the simulated monthly rutting calculations and the AASHTOWare® Pavement ME Design software calculations. The intent is to demonstrate that the simulated calculation approach produces results that are almost identical to the software results on various pavement structures (LTPP SPS-1 and SPS-5 sections in Florida) and with different local calibration factors. The literature review indicated that the previous calibration efforts had found the calibration factors to be within the ranges in table 49. These ranges were used to change the calibration factors one by one and observe the simulated rutting calculations compared to the Pavement ME software results.

Table 49. Range of the calibration factors reported in the literature.

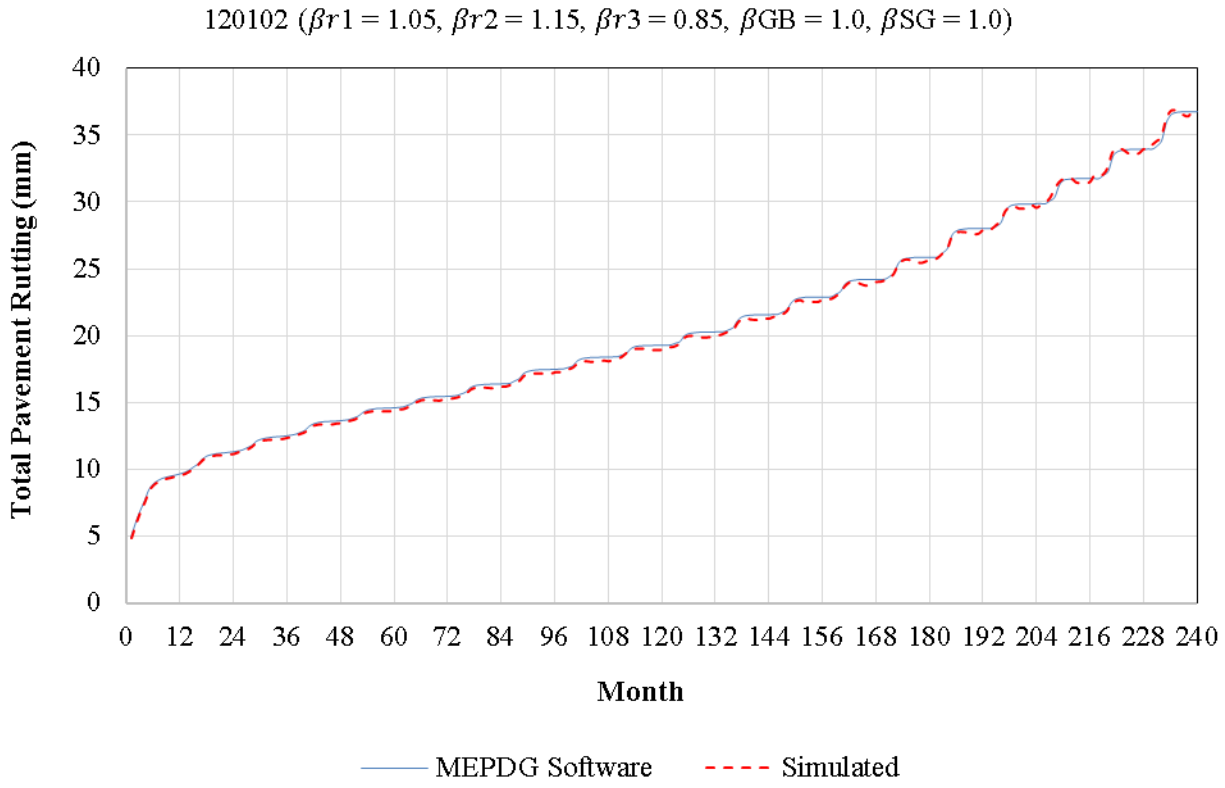
Statistic	HMA Rutting β_{r1}	HMA Rutting β_{r2}	HMA Rutting β_{r3}	Base Rutting β_{GB}	Subgrade Rutting β_{SG}
Average	1.7757	1.0445	0.9273	0.4039	0.4569
Range	0.51 to 7	1 to 1.15	0.7 to 1.1	0.0 to 1.5803	0.0 to 1.38

A handful of these comparisons have been demonstrated in figure 43 through figure 53. In these figures, the solid line represents the rut depth values calculated using the MEPDG software as a function of pavement age, and dashed lines represent the simulated values for the same pavement sections with the same calibration factors. These figures show that overall, the simulated process is successfully estimating the rutting progression trend very similar to the Pavement ME software. While there are some intermediate decrements in rutting values simulated within each month, the total accumulated rutting at the end of each month is very close to the software output. The reason for the intermediate decrements (which do not comply with the theory of rutting accumulation) is the assumptions made for the simulation, which may not be inherently correct. One of those assumptions is that traffic and rutting values are increasing linearly within each month, which might not be true. As explained before, the actual subseason pavement response data for each layer are not provided by the AASHTOWare® software, and therefore, an exact calculation (according to the MEPDG equations) could not be conducted for this project.



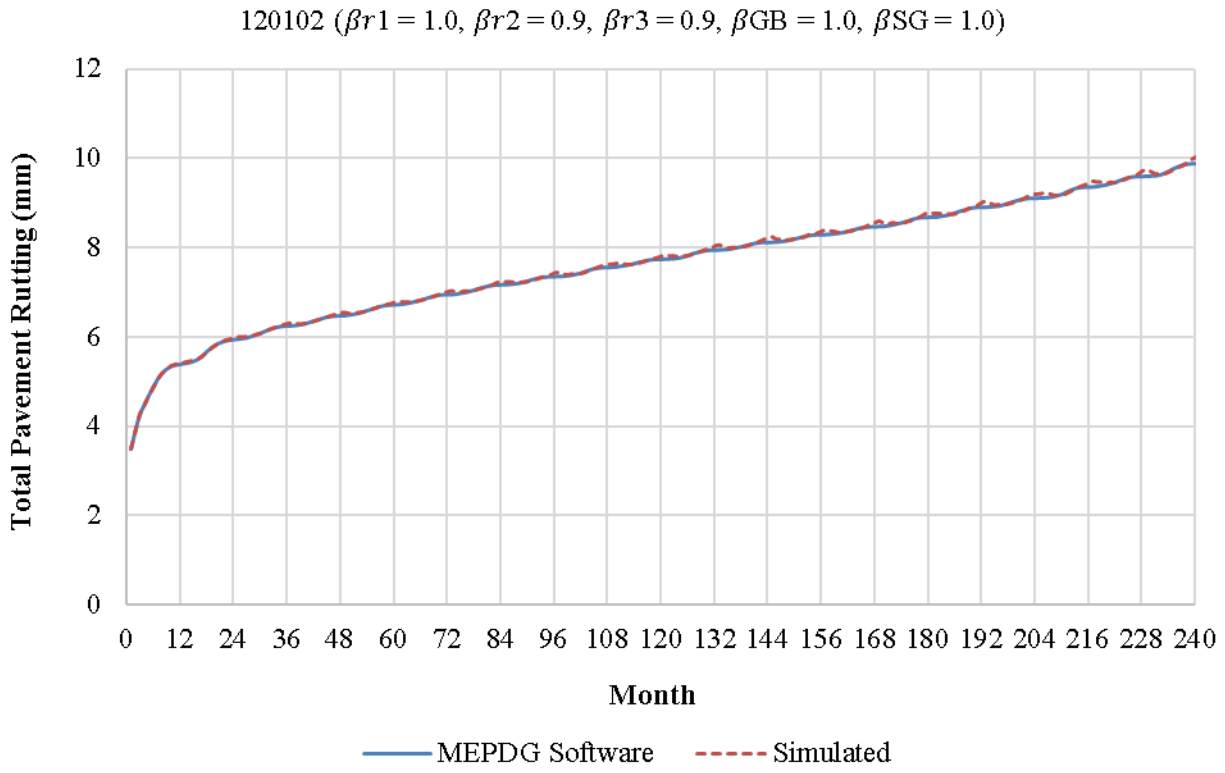
Source: FHWA.

Figure 43. Chart. Comparison of simulated rutting calculations to ME software results for test section 120102 with β_{r1} = 1.05, β_{r2} = 0.9, β_{r3} = 0.85, β_{GB} = 1.0, β_{SG} = 1.0.



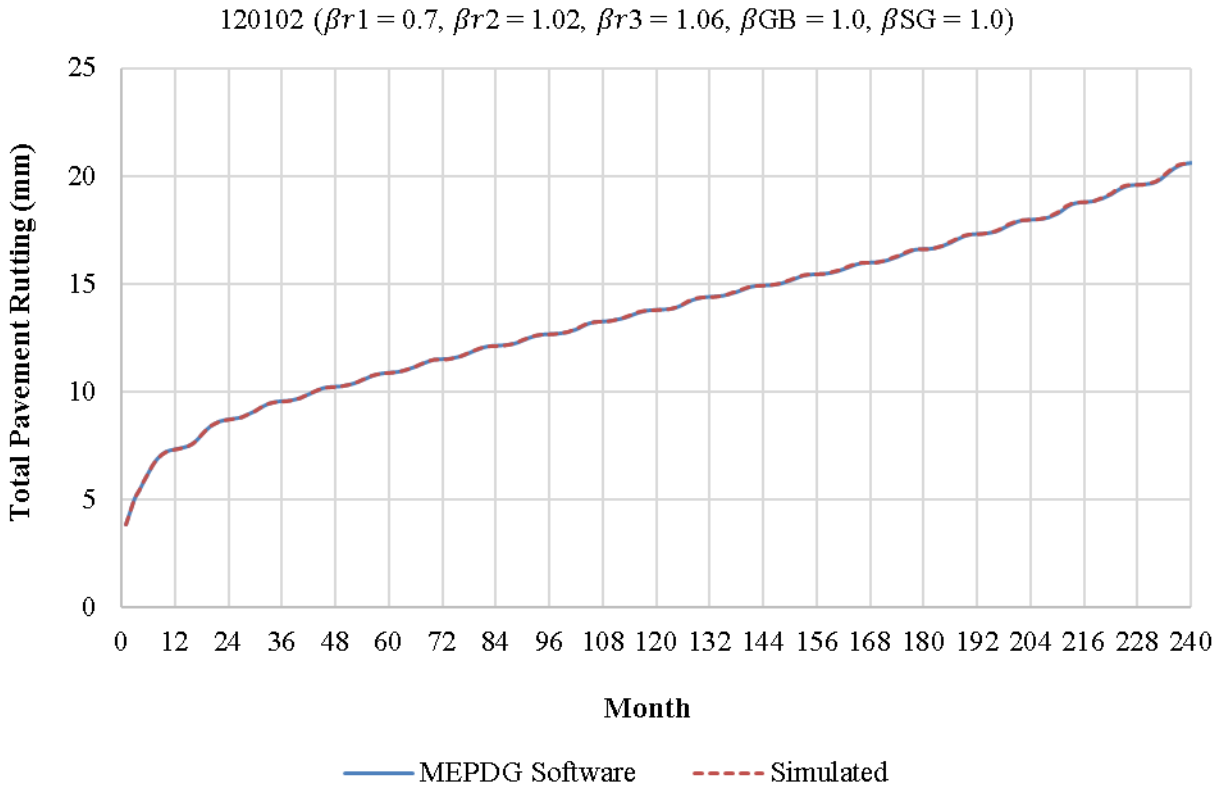
Source: FHWA.

Figure 44. Chart. Comparison of simulated rutting calculations to ME software results for test section 120102 with β_{r1} = 1.05, β_{r2} = 1.15, β_{r3} = 0.85, β_{GB} = 1.0, β_{SG} = 1.0.



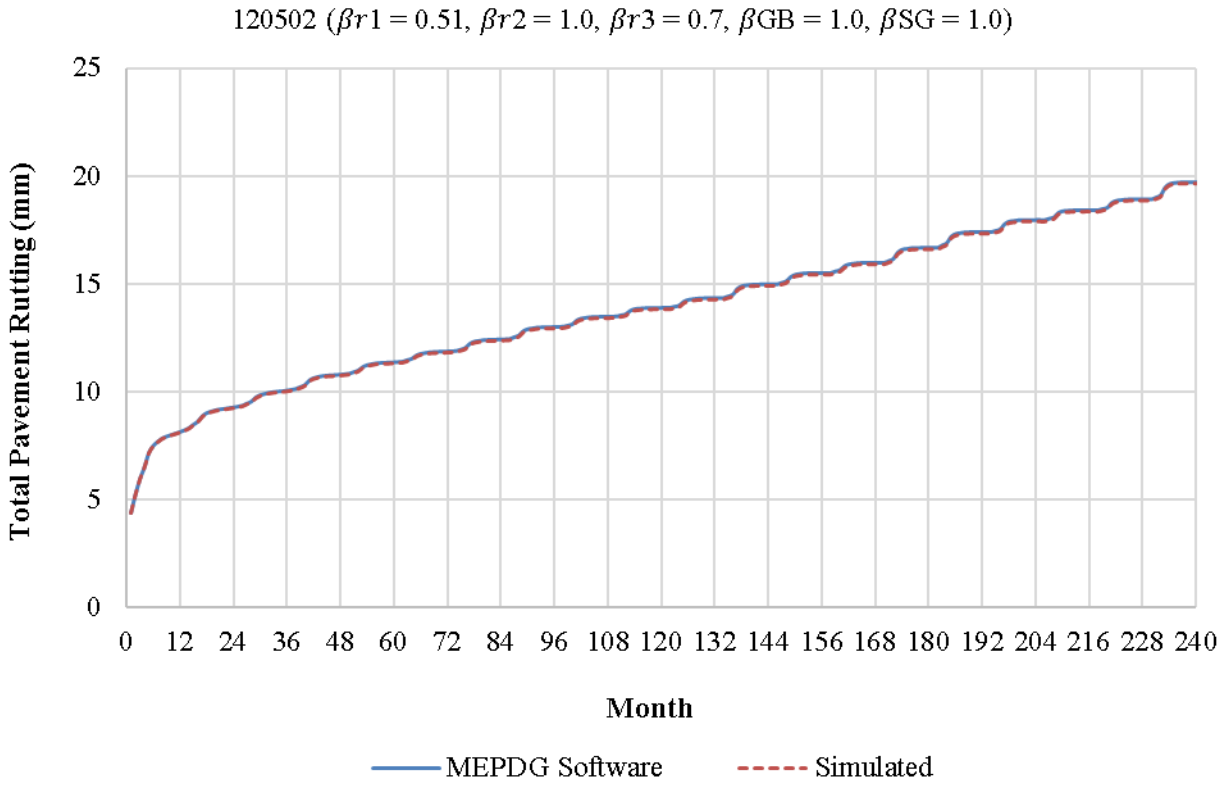
Source: FHWA.

Figure 45. Chart. Comparison of simulated rutting calculations to ME software results for test section 120102 with β_{r1} = 1.0, β_{r2} = 0.9, β_{r3} = 0.9, β_{GB} = 1.0, β_{SG} = 1.0.



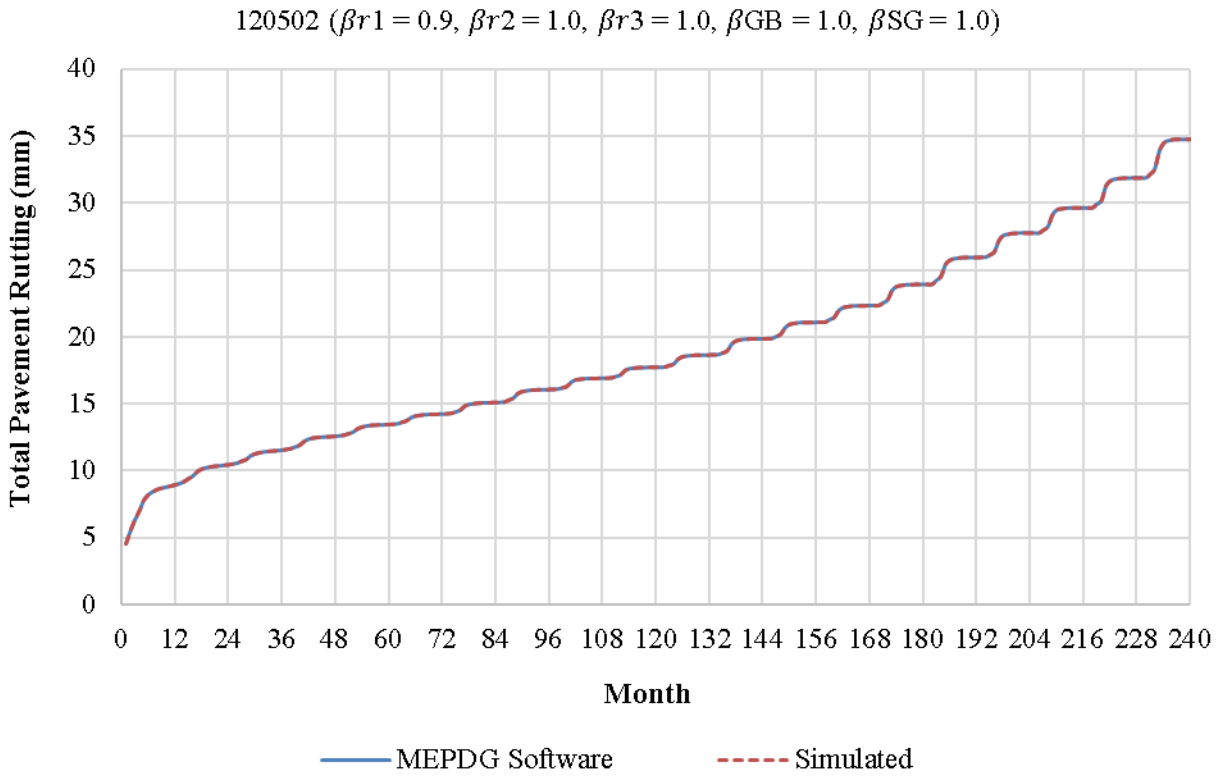
Source: FHWA.

Figure 46. Chart. Comparison of simulated rutting calculations to ME software results for test section 120102 with β_{r1} = 0.7, β_{r2} = 1.02, β_{r3} = 1.06, β_{GB} = 1.0, β_{SG} = 1.0.



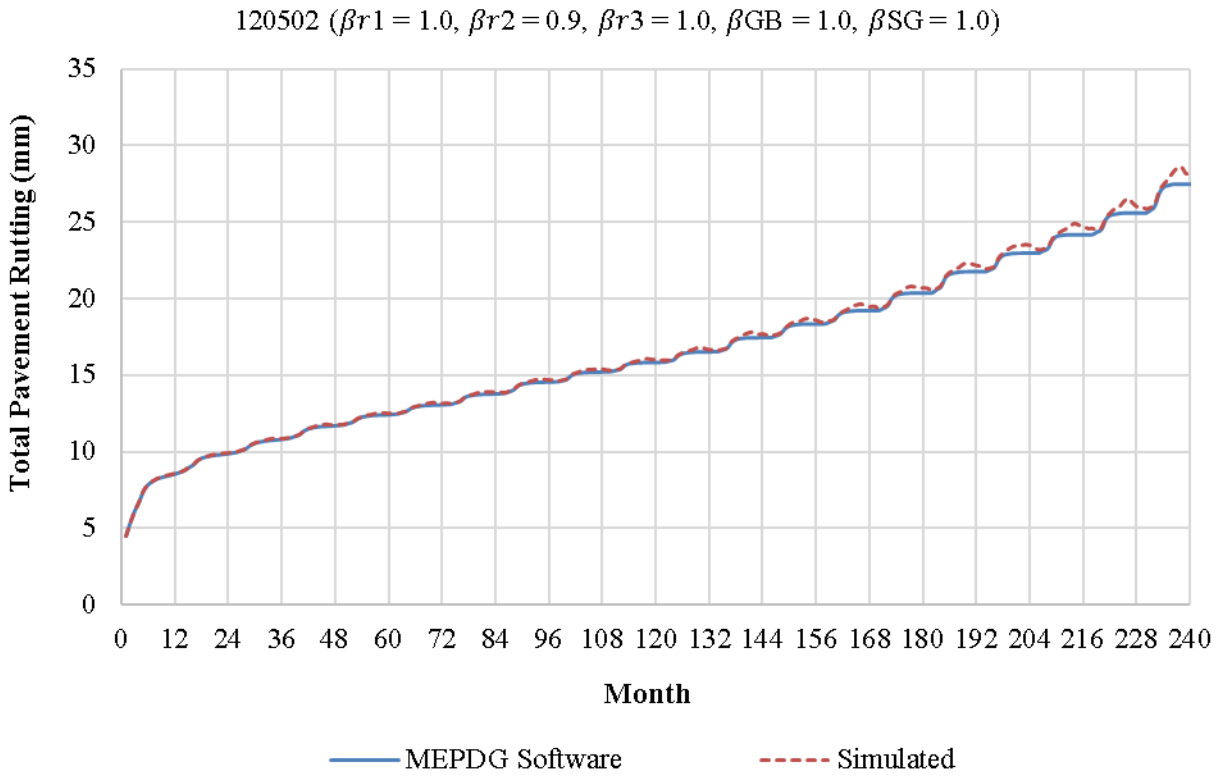
Source: FHWA.

Figure 47. Chart. Comparison of simulated rutting calculations to ME software results for test section 120502 with β_{r1} = 0.51, β_{r2} = 1.0, β_{r3} = 0.7, β_{GB} = 1.0, β_{SG} = 1.0.



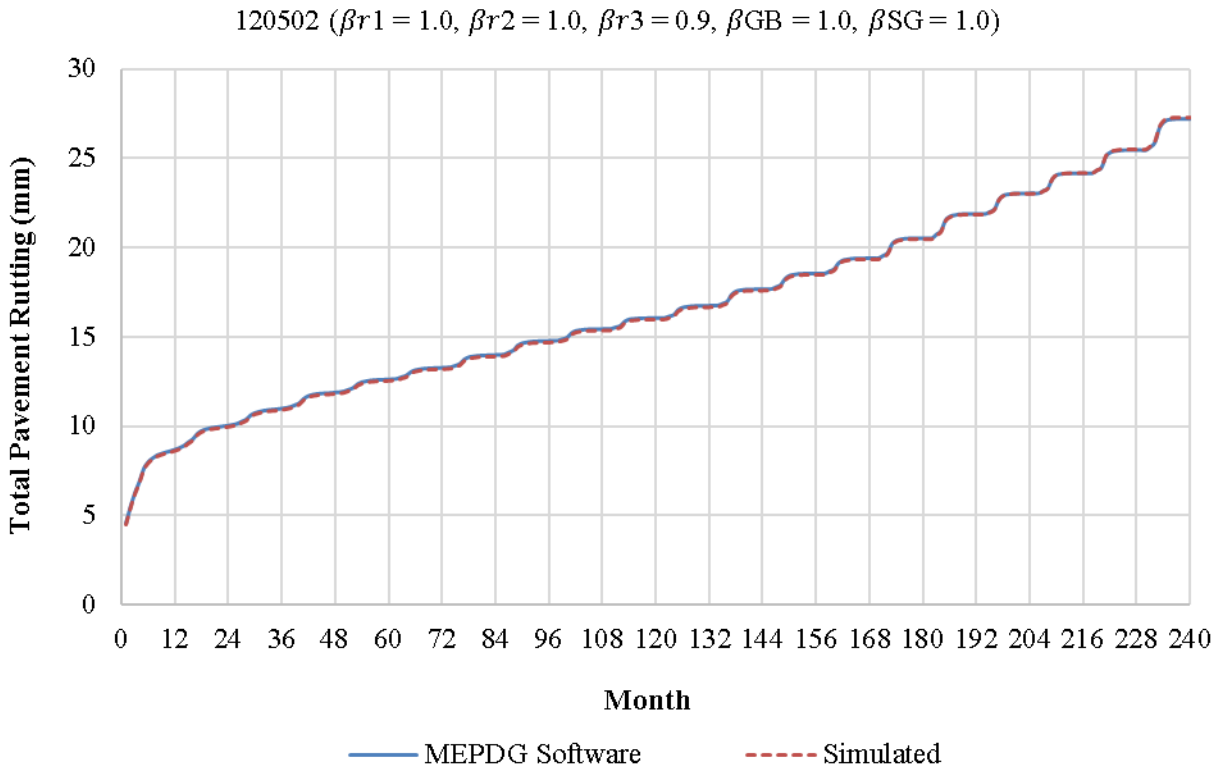
Source: FHWA.

Figure 48. Chart. Comparison of simulated rutting calculations to ME software results for test section 120502 with β_{r1} = 0.9, β_{r2} = 1.0, β_{r3} = 1.0, β_{GB} = 1.0, β_{SG} = 1.0.



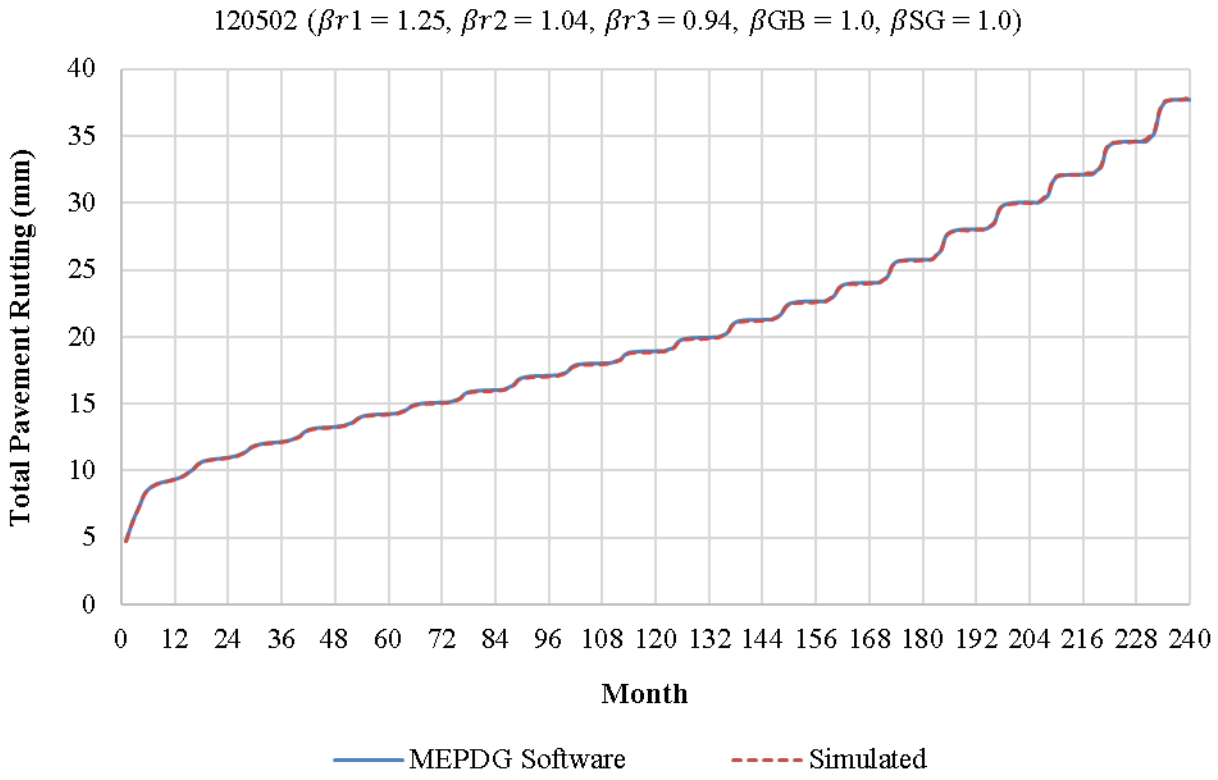
Source: FHWA.

Figure 49. Chart. Comparison of simulated rutting calculations to ME software results for test section 120502 with β_{r1} = 1.0, β_{r2} = 0.9, β_{r3} = 1.0, β_{GB} = 1.0, β_{SG} = 1.0.



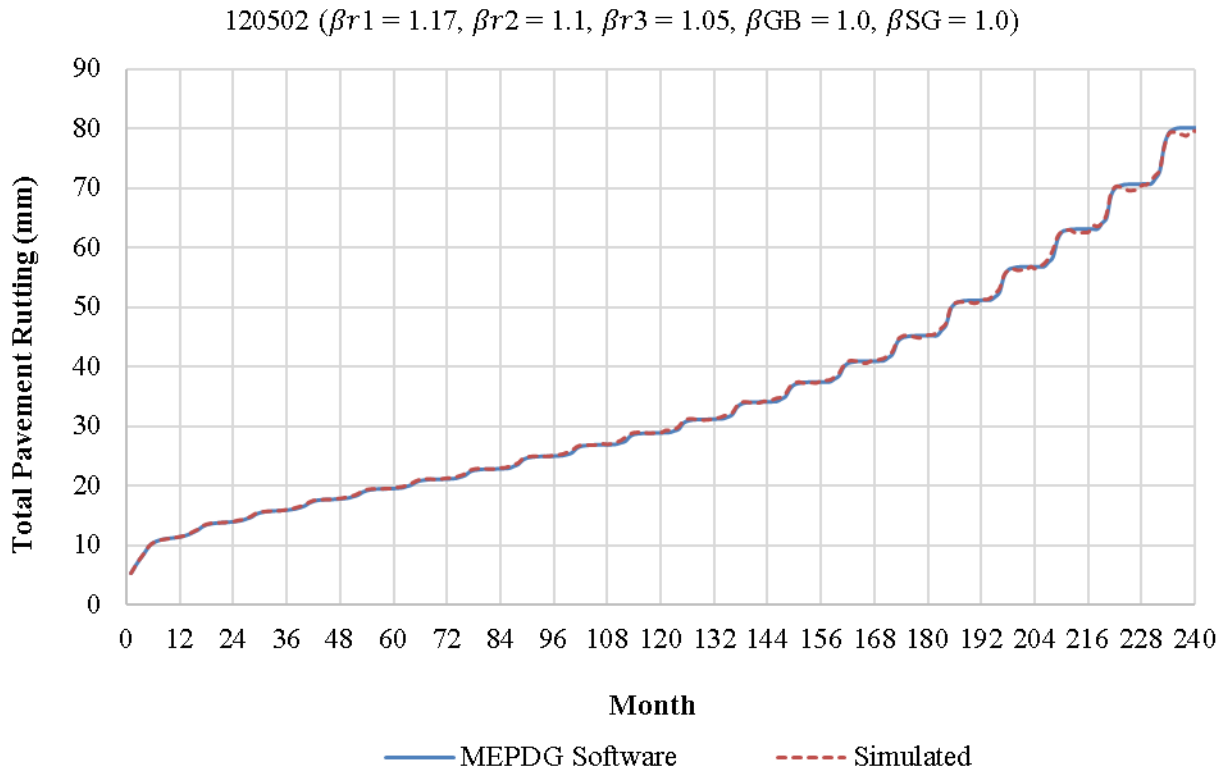
Source: FHWA.

Figure 50. Chart. Comparison of simulated rutting calculations to ME software results for test section 120502 with β_{r1} = 1.0, β_{r2} = 1.0, β_{r3} = 0.9, β_{GB} = 1.0, β_{SG} = 1.0.



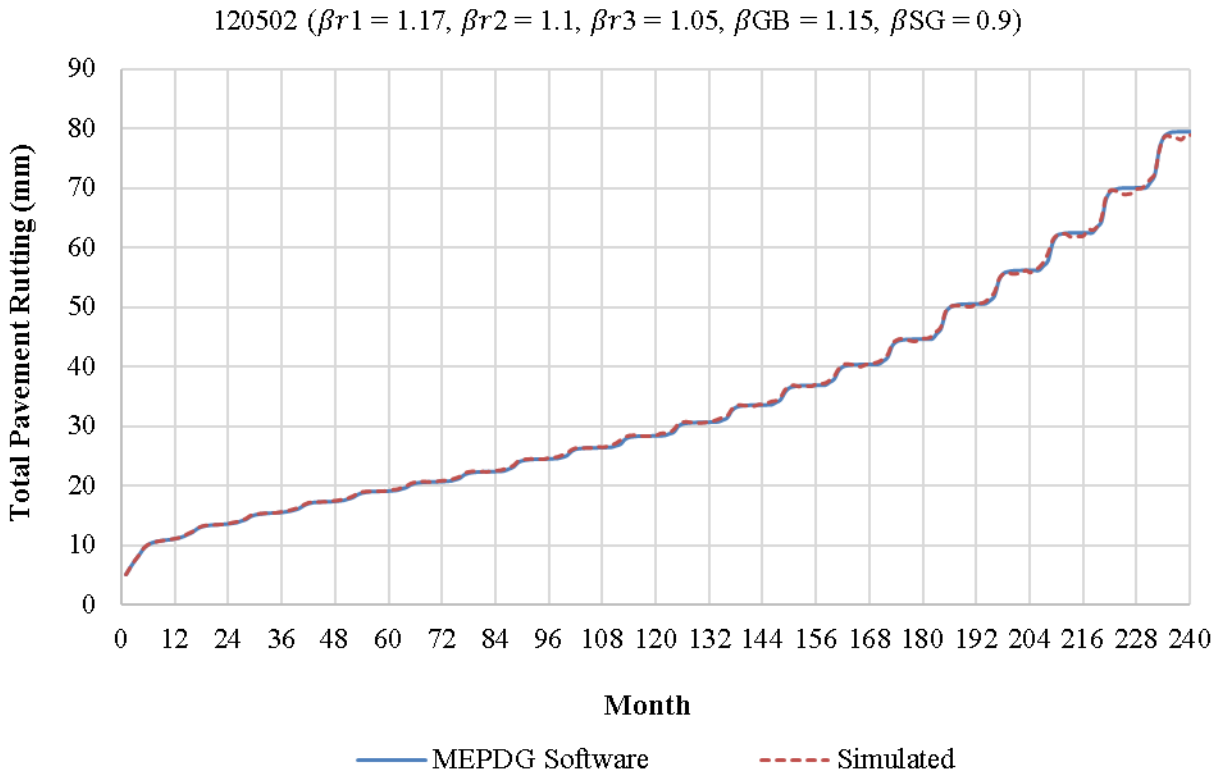
Source: FHWA.

Figure 51. Chart. Comparison of simulated rutting calculations to ME software results for test section 120502 with β_{r1} = 1.25, β_{r2} = 1.04, β_{r3} = 0.94, β_{GB} = 1.0, β_{SG} = 1.0.



Source: FHWA.

Figure 52. Chart. Comparison of simulated rutting calculations to ME software results for test section 120502 with β_{r1} = 1.17, β_{r2} = 1.1, β_{r3} = 1.05, β_{GB} = 1.0, β_{SG} = 1.0.



Source: FHWA.

Figure 53. Chart. Comparison of simulated rutting calculations to ME software results for test section 120502 with β_{r1} = 1.17, β_{r2} = 1.1, β_{r3} = 1.05, β_{GB} = 1.15, β_{SG} = 0.9.

REFERENCES

1. American Association of State Highway and Transportation Officials. (2015). AASHTOWare® Pavement ME Design, 2.2 Build 2.2.4, American Association of State Highway and Transportation Officials, Washington, DC.

2. National Cooperative Highway Research Program. (2004). *Guide for Mechanistic–Empirical Design of New and Rehabilitated Pavement Structures,* Final Report, NCHRP Project 1-37A, National Cooperative Highway Research Program, National Research Council, Washington, DC.

3. American Association of State Highway and Transportation Officials. (2010). *Guide for the Local Calibration of the Mechanistic–Empirical Pavement Design Guide.* American Association of State Highway and Transportation Officials, Washington, DC.

4. National Cooperative Highway Research Program. (2009). *User Manual and Local Calibration Guide for the Mechanistic–Empirical Pavement Design Guide and Software,* NCHRP Project 1-40B, Final Report, National Cooperative Highway Research Program, Transportation Research Board, Washington, DC.

5. Von Quintus, H.L. and Moulthrop, J.S. (2007). *Mechanistic–Empirical Pavement Design Guide Flexible Pavement Performance Prediction Models for Montana: Volume II Reference Manual,* Report No. FHWA/MT-07-008/8158-2, Fugro Consultants, Inc.-Montana Department of Transportation, Austin, TX.

6. Darter, M.I., Mallela, J., Titus-Glover, L., Rao, C., Larson, G., Gotlif, A., Von Quintus, H.L., et al. (2006). *NCHRP Research Results Digest 308: Changes to the Mechanistic–Empirical Pavement Design Guide Software Through Version 0.900, July 2006.* Transportation Research Board of the National Academies, Washington, DC.

7. Kang, M., Adams, T.M., and Bahia, H.U. (2007). *Development of a Regional Pavement Performance Database for the AASHTO Mechanistic–Empirical Pavement Design Guide: Part 2: Validations and Local Calibration,* Report No. MRUTC 07-01, Department of Civil & Environmental Engineering, University of Wisconsin-Madison-Midwest Regional University Transportation Center, Madison, WI.

8. Banerjee, A., Aguiar-Moya, J.P., Smit, A.d.F., and Prozzi, J.A. (2010). *Development of the Texas Flexible Pavements Database,* Report No. FHWA/TX-10/0-5513-2, Center for Transportation Research, Texas Department of Transportation, Austin, TX.

9. Federal Highway Administration. (2010). *Local Calibration of the MEPDG Using Pavement Management Systems,* Report No. HIF-11-026, Office of Asset Management, Federal Highway Administration, Washington, DC.

10. Pierce, L.M. and McGovern, G. (2014). *NCHRP Synthesis 457: Implementation of the AASHTO Mechanistic–Empirical Pavement Design Guide and Software.* Transportation Research Board of the National Academies, Washington, DC.

11. Miller, J.S. and Bellinger W.Y. (2014). *Distress Identification Manual for the Long-Term Pavement Performance Program* (Fifth Revised Edition). Report No. FHWA-HRT-13-092, Office of Infrastructure Research and Development, Federal Highway Administration, Washington, D.C.

12. Kim, Y.R., Underwood, B.S., Far, M.S., Jackson, N., and Puccinelli, J. (2011). *LTPP Computed Parameter: Dynamic Modulus,* Report No. FHWA-HRT-10-035, Nichols Consulting Engineers, Chtd., Federal Highway Administration, Reno, NV.

13. Von Quintus, H.L., Mallela, J., Sudavisan, S., and Darter, M.I. (2013). *Verification and Local Calibration/Validation of the MEPDG Performance Models for Use in Georgia, Literature Search and Synthesis, Task—Interim Report,* Report No. GADOT-TO-01-Task 1, Applied Research Associates, Inc.-Georgia Department of Transportation, Champaign, IL.

14. Schwartz, C.W., Li, R., Kim, S., Ceylan, H., and Gopalakrishnan, K. (2011). *Sensitivity Evaluation of MEPDG Performance Prediction,* NCHRP Project 1-47, Final Report, National Cooperative Highway Research Program, Washington, DC.

15. Li, J., Pierce, L.M., and Uhlmeyer, J.S. (2009). "Calibration of Flexible Pavement in Mechanistic–Empirical Pavement Design Guide for Washington State." *Transportation Research Record: Journal of the Transportation Research Board, 2095*, pp. 73–83, Transportation Research Board, Washington, DC.

16. Ceylan, H., Kim, S., Gopalakrishnan, K., and Ma, D. (2013). *Iowa Calibration of MEPDG Performance Prediction Models,* Report No. InTrans Project 11-401, Institute for Transportation, Iowa State University, Iowa Department of Transportation, Ames, IA.

17. Von Quintus, H.L., Schwartz, C., McCuen, R.H., and Andrei, D. (2004). *Experimental Plan for Calibration and Validation of Hot-Mix Asphalt Performance Models for Mix and Structural Design*, NCHRP Project 9-30, Report No. Final, National Cooperative Highway Research Program Project, Washington, DC.

18. Von Quintus, H.L. (2008). "Local Calibration of MEPDG An Overview of Selected Studies (With Discussion)." *Journal of the Association of Asphalt Paving Technologists, 77*, pp. 935–974, Association of Asphalt Paving Technologists, Lino Lakes, MN.

19. Corley-Lay, J., Jadoun, F.M., Mastin, J.N., and Kim, Y.R. (2010). "Comparison of Flexible Pavement Distresses Monitored by North Carolina Department of Transportation and Long-Term Pavement Performance Program." *Transportation Research Record: Journal of the Transportation Research Board, 2153*, pp. 91–96, Transportation Research Board, Washington, DC.

20. Mamlouk, M.S. and Zapata, C.E. (2010). *"The Need to Carefully Assess Use of State PMS Data for MEPDG Calibration Process."* Presented at the 89th Annual Meeting of the Transportation Research Board, Transportation Research Board, Washington, DC.

21. Thompson, M.R., Carpenter, S.H., Dempsey, B.J., and Elliot, R.B. (2006). *Independent Review of the Mechanistic–Empirical Design Guide and Software,* NCHRP Project 1-40A, *Research Results Digest 307,* National Cooperative Highway Research Program, Transportation Research Board, Washington, DC.

22. Muthadi, N. and Kim, Y. (2008). "Local Calibration of Mechanistic–Empirical Pavement Design Guide for Flexible Pavement Design." *Transportation Research Record: Journal of the Transportation Research Board, 2087,* pp. 131–141, Transportation Research Board, Washington, DC.

23. Banerjee, A., Aguiar-Moya, J.P., and Prozzi, J.A. (2009). "Texas Experience Using LTPP for Calibration of the MEPDG Permanent Deformation Models." *Transportation Research Record: Journal of the Transportation Research Board, 2094,* pp. 12–20, Transportation Research Board, Washington, DC.

24. Titus-Glover, L. and Mallela, J. (2009). *Guidelines for Implementing NCHRP 1-37A ME Design Procedures in Ohio: Volume 4-MEPDG Models Validation & Recalibration,* Report No. FHWA/OH-2009/9D, Applied Research Associates, Inc., Champaign, IL.

25. Souliman, M.I., Mamlouk, M.S., El-Basyouny, M.M., and Zapata, C.E. (2010). *"Calibration of the AASHTO MEPDG for Flexible Pavement for Arizona Conditions."* Presented at the 89th Annual Meeting of the Transportation Research Board, Washington, DC.

26. Hoegh, K., Khazanovich, L., and Jense, M. (2010). "Local Calibration of Mechanistic–Empirical Pavement Design Guide Rutting Model: Minnesota Road Research Project Test Sections." *Transportation Research Record: Journal of the Transportation Research Board, 2180,* pp. 130–141, Transportation Research Board, Washington, DC.

27. Hall, K.D., Xiao, D.X., and Wang, K.C.P. (2011). "Calibration of the Mechanistic–Empirical Pavement Design Guide for Flexible Pavement Design in Arkansas." *Transportation Research Record: Journal of the Transportation Research Board, 2226,* pp. 135–141, Transportation Research Board, Washington, DC.

28. Williams, R.C. and Shaidur, R. (2013). *Mechanistic–Empirical Pavement Design Guide Calibration for Pavement Rehabilitation,* Report No. FHWA-OR-RD-13-10, Institute for Transportation Iowa State University, Oregon Department of Transportation, Ames, IA.

29. Mallela, J., Titus-Glover, L., Sadasivam, S., Bhattacharya, B.B., Darter, M.I., and Von Quintus, H.L. (2013). *Implementation of the AASHTO Mechanistic–Empirical Pavement Design Guide for Colorado,* Report No. CDOT-2013-4, Applied Research Associates, Inc.-Colorado Department of Transportation, Champaign, IL.

30. Li, J., Luhr, D.R., and Uhlmeyer, J.S. (2010). "Pavement Performance Modeling Using Piecewise Approximation." *Transportation Research Record: Journal of the Transportation Research Board, 2153,* pp. 24–29, Transportation Research Board, Washington, DC.

31. Von Quintus, H.L., Mallela, J., Bonaquist, R.F., Schwartz, C.W., and Carvalho, R.L. (2012). *NCHRP Report 719: Calibration of Rutting Models for Structural and Mix Design.* Transportation Research Board of the National Academies, Washington, DC.

32. Jadoun, F.M. and Kim, Y.R. (2012). "Calibrating Mechanistic–Empirical Pavement Design Guide for North Carolina—Genetic Algorithm and Generalized Reduced Gradient Optimization Methods." *Transportation Research Record: Journal of the Transportation Research Board, 2305,* pp. 131–140, Transportation Research Board, Washington, DC.

33. Coello Coello, C.A. (2001). "A Short Tutorial on Evolutionary Multiobjective Optimization, International Conference on Evolutionary Multi-Criterion Optimization." *EMO,* Eds. Zitzler, Thiele, Deb, Coello Coello, and Corne. Springer-Verlag, Berlin-Heidelberg, Germany.

34. Fwa, T.F., Chan, W.T., and Hoque, K.Z. (2000). "Multiobjective Optimization for Pavement Maintenance Programming." *Journal of Transportation Engineering, 126,* pp. 367–374, American Society of Civil Engineers, Reston, VA.

35. Deshpande, V., Damnjanovic, I.D., and Gardoni, P. (2010). "Reliability-Based Optimization Models for Scheduling Pavement Rehabilitation." *Computer-Aided Civil and Infrastructure Engineering, 25,* pp. 227–237, Blackwell Publishers, Malden, MA.

36. Gao, L., Xie, C., Zhang, Z., and Waller, S.T. (2012). "Network-Level Road Pavement Maintenance and Rehabilitation Scheduling for Optimal Performance Improvement and Budget Utilization." *Computer-Aided Civil and Infrastructure Engineering, 27,* pp. 278–287, Wiley Online Library, doi:10.1111/j.1467-8667.2011.00733.x.

37. Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., and Kollat, J.B. (2013). "Evolutionary Multiobjective Optimization in Water Resources: The Past, Present, and Future." *Advances in Water Resources, 51,* pp. 438–456, Elsevier, doi:10.1016/j.advwatres.2012.01.005.

38. Tang, Y., Reed, P., and Wagener, T. (2006). "How Effective and Efficient Are Multiobjective Evolutionary Algorithms at Hydrologic Model Calibration?" *Hydrology and Earth System Sciences Discussions, 10,* pp. 289–307, Hydrology and Earth System Sciences, doi:10.5194/hess-10-289-2006.

39. Efstratiadis, A. and Koutsoyiannis, D. (2010). "One Decade of Multi-Objective Calibration Approaches in Hydrological Modelling: A Review." *Hydrological Sciences Journal, 55,* pp. 58–78, Taylor & Francis, doi:10.1080/02626660003526292.

40. Federal Highway Administration. (2016). "LTPP Infopave™." (website) Washington, DC. Available online: https://infopave.fhwa.dot.gov/, last accessed December 15, 2016.

41. Elkins, G.E., Thompson, T., Ostrom, B., Simpson, A., and Visintine, B.A. (2015). *Long-Term Pavement Performance Information Management System User Guide,* Report No. FHWA-RD-03-088 (revision), Federal Highway Administration, McLean, VA.

42. Selezneva, O.I. and Hallenbeck, M. (2013). *Long-Term Pavement Performance Pavement Loading User Guide (LTPP PLUG) Software Manual,* Report No. FHWA-HRT-13-089, Federal Highway Administration, McLean, VA.

43. Brown, E.R., Kandhal, P.S., Roberts, F.L., Kim, Y.R., Lee, D., and Kennedy, T.W. (2009). *Hot Mix Asphalt Materials, Mixture Design, and Construction.* NAPA Research and Education Foundation, Lanham, MD.

44. Asphalt Institute. (1989). *The Asphalt Handbook, Manual Series No. 4 (MS-4).* Asphalt Institute, Lexington, KY.

45. Mirza, M.W. and Witczak, M.W. (1995). "Development of a Global Aging System for Short and Long Term Aging of Asphalt Cements (With Discussion)." *Journal of the Association of Asphalt Paving Technologists, 64,* pp. 393–430, Association of Asphalt Paving Technologists, Lino Lakes, MN.

46. Federal Highway Administration. (2015). *Determination of In-Place Elastic Layer Modulus: LTPP Backcalculation Methodology and Procedures,* Publication No. FHWA-HRT-15-037, Federal Highway Administration, Washington, DC.

47. Loulizi, A., Flintsch, G.W., and McGhee, K.K. (2006). *Determination of the In-Place Hot-Mix Asphalt Layer Modulus for Rehabilitation Projects Using a Mechanistic–Empirical Procedure,* Report No. FHWA/VTRC 07-CR1, Virginia Tech Transportation Institute, Virginia Department of Transportation, Blacksburg, VA.

48. Loulizi, A., Al-Qadi, I.L., Lahouar, S., and Freeman, T. (2002). "Measurement of Vertical Compressive Stress Pulse in Flexible Pavements: Representation for Dynamic Loading Tests." *Transportation Research Record: Journal of the Transportation Research Board, 1816,* pp. 125–136, Transportation Research Board, Washington, DC.

49. Katicha, S., Flintsch, G.W., Loulizi, A., and Wang, L. (2008). "Conversion of Testing Frequency to Loading Time Applied to the Mechanistic–Empirical Pavement Design Guide." *Transportation Research Record: Journal of the Transportation Research Board, 2087,* pp. 99–108, Transportation Research Board, Washington, DC.

50. Ji, Y. (2005). "Frequency and Time Domain Backcalculation of Flexible Pavement Layer Parameters." Ph.D. dissertation. Michigan State University, East Lansing, MI.

51. Witczak, M.W. (2004). *NCHRP Research Results Digest 285: Laboratory Determination of Resilient Modulus for Flexible Pavement Design,* Transportation Research Board of the National Academies, Washington, DC.

52. American Association of State Highway and Transportation Officials. (2015). *Mechanistic–Empirical Pavement Design Guide: A Manual of Practice,* 2nd Edition. American Association of State Highway and Transportation Officials, Washington, DC.

53. Greene, J., Chun, S., and Choubane, B. (2014). "Enhanced Gradation Guidelines to Improve Asphalt Mixture Performance." *Transportation Research Record: Journal of the*

Transportation Research Board, 1, pp. 3–10, Transportation Research Board, Washington, DC.

54. Greene, J., Chun, S., Nash, T., and Choubane, B. (2014). *"Evaluation and Implementation of PG 76-22 Asphalt Rubber Binder in Florida."* Presented at the 94th Annual Meeting of the Transportation Research Board, Washington, DC.

55. Pareto, V. (1896). *Cours D'Economie Politique.* Rouge, Lausanne.

56. Bäck, T., Fogel, D.B., and Michalewicz, Z. (1997). *Handbook of Evolutionary Computation.* Oxford University Press/IOP Publishing, New York, NY.

57. Kollat, J.B. and Reed, P.M. (2006). "Comparing State-of-the-Art Evolutionary Multi-Objective Algorithms for Long-Term Groundwater Monitoring Design." *Advances in Water Resources, 29,* pp. 792–807, Elsevier, doi:10.1016/j.advwatres.2005.07.010.

58. Kollat, J.B. and Reed, P.M. (2005). *The Value of Online Adaptive Search: A Performance Comparison of NSGAII, ε-NSGAII and εMOEA, International Conference on Evolutionary Multi-Criterion Optimization,* Eds. Coello Coello, Hernandez and Zitzler. Springer-Verlag, Berlin-Heidelberg, Germany.

59. Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II." *IEEE Transactions on Evolutionary Computation, 6,* pp. 182–197, IEEE, doi:10.1109/4235.996017.

60. Srinivas, N. and Deb, K. (1994). "Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms." *Evolutionary Computation, 2,* pp. 221–248, MIT Press, Cambridge, MA.

61. Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, Boston, MA.

62. Reed, P., Minsker, B.S., and Goldberg, D.E. (2003). "Simplifying Multiobjective Optimization: An Automated Design Methodology for the Nondominated Sorted Genetic Algorithm-II." *Water Resources Research, 39,* Wiley Online Library, doi:10.1029/2002WR001483.

63. Kargah-Ostadi, N. and Stoffels, S.M. (2015). "A Framework for Development and Comprehensive Comparison of Alternative Pavement Performance Models," *Journal of Transportation Engineering, 141*(8) 04015012, American Society of Civil Engineers, Reston, VA.

64. Nichols Consulting Engineers. (2010). *LTPP: Artificial Neural Networks for Asphalt Concrete Dynamic Modulus Prediction-ANNA CAP,* V.1.2, Federal Highway Administration, Washington, DC.

65. Huang, Y.H. (2004). *Pavement Analysis and Design.* Pearson/Prentice Hall, Upper Saddle River, NJ.

HRDI-30/06-18(WEB)E