Calibration of AASHTOWare Pavement ME Design Performance Prediction Models for Flexible Pavements in Quebec

Syrine Chabchoub
Postdoctoral researcher
Department of Civil Engineering, Laval University
Québec, Québec/Canada
syrine.chabchoub.1@ulaval.ca

Jean Pascal Bilodeau
Assistant professor
Department of Civil Engineering, Laval University
Québec, Québec/Canada
jean-pascal.bilodeau@gci.ulaval.ca

Erdrick Leandro Pérez-González
Assistant professor
Department of Civil Engineering, Laval University
Québec, Québec/Canada
erdrick.perez-gonzalez@gci.ulaval.ca

Olivier Sylvestre
Engineer
Ministère des Transports et de la Mobilité Durable
Québec, Québec/Canada
olivier.Sylvestre@transports.gouv.qc.ca

Prepared paper for the session
Design, construction and management of a sustainable road network
2025 TAC Conference & Exhibition
Quebec City (Quebec)

Abstract

The common practice for designing asphalt pavement structures in Quebec has been based on the AASHTO 93 method for over twenty years. This approach was developed to determine the components of a pavement structure to ensure, from a structural standpoint, adequate performance throughout the pavement's expected life. Validating and calibrating a mechanistic-empirical (ME) method would notably allow better consideration of Quebec's specific climatic conditions. In the current practice in Quebec, the structural design method is limited to dividing the province into two climatic zones (South and North). In contrast, the ME method proposed by the AASHTO simulates the effects of hourly temperature distribution, precipitation, sunlight, wind speed, and even the depth of the water table . The ME approach

aims to establish links between the characteristics of pavement materials, climatic conditions, and traffic loading to predict the occurrence and progression of various degradation mechanisms. Before implementing this pavement design method, it is essential to calibrate the prediction models to Quebec's conditions to reduce biased measurements that often reflect inadequate correlations between field results and predictions. This study will undertake a local calibration for flexible pavements in Quebec using AASHTOWare Pavement Mechanistic-Empirical Design (PMED) to ensure accurate predictions under local conditions. The Calibration Assistance Tool (CAT) of the AASHTOware Pavement ME Design has been used to assess the local calibration factors for the performance models, including Bottom-Up Fatigue Cracking, Total Rutting, Transverse Cracking, and International Roughness Index (IRI). A total of 353 flexible pavements under the authority of the Ministère des Transports et de la Mobilité Durable (MTMD) have been retained concerning several criteria. Better accuracy of pavement performance has been yielded based on the goodness of fit between local and national calibration findings. Results show enhanced predictions in the case of Fatigue cracking, Rutting and International Roughness Index compared to the global calibration.

1. Introduction

Lack of initiatives to locally calibrate performance models in Quebec was the driving force behind this project, initiated by the Ministry of Transportation and Sustainable Mobility of Quebec (MTMD) and Laval University. There is a need to accurately reflect the local conditions of Quebec, as the PMED national calibration did not reflect a proper correlation between the field-measured and predicted data. At the provincial level, research in Canada to calibrate pavement performance models locally is limited: Ahmed et al.[1] conducted local calibration of cracking models for flexible pavements in the Province of Ontario. The aim was to optimize the calibration coefficients and minimize bias and standard error between field measurements and predictions using MEPDG. The calibration process involved bottom-up, top-down fatigue cracking and thermal longitudinal cracking. The locally calibrated model did not achieve satisfactory performance. The authors recommended further optimization of database for the load related cracking calibration as well as upgrading input level from level 3 to level 1 to achieve more accuracy. Dong et al.[2] put forward an advanced local calibration method for Mechanistic-Empirical Pavement Design (MEPD) to overcome the predictions' misrepresentativity when using global calibration coefficients. A Local Calibration (LC) method that integrates a jackknife sampling procedure and an iteratively weighted least squares technique into one coherent LC process was proposed. The study considered both bottom-up fatigue cracking and joint faulting. To capture the local conditions of Ontario's road network, 44 flexible pavement sections and 27 rigid pavement sections were selected. Better predictions with lower bias and Residual Sum of Squares (RSS) were observed compared to the traditional data-split method findings. Parallel to studies conducted in Canada, many attempts have been made by the American Highway Agencies to develop local calibration factors that better represent local conditions. Darter et al[3] attempted to overcome the under-prediction of Transverse Cracking of pavements in Arizona yielded by national calibration factors. Bias was accordingly reduced, and better fit was observed.

[1] Ahmed, S.E., Yuan, X.X., Lee, W., & Li, N. (2016). Local Calibration of Cracking Models in MEPDG for Ontario's Flexible Pavement, *Proceedings of the Transportation Association of Canada Conference,* Toronto, Ontario, Canada.

[2] Dong, Shi., Yuan, X. X., & Hao, PW. (2020). An advanced local calibration method for mechanistic-empirical pavement design. *Computer-Aided Civil and Infrastructure Engineering*. 35. 10.1111/mice.12574.

[3] Darter, M.I., Titus-Glover, L., Von Quintus, H., Bhattacharya, BB., & Jagannath. M. (2014). Calibration and Implementation of the AASHTO Mechanistic-Empirical Pavement Design Guide in Arizona. Arizona Department Transportation Research Center,https://apps.azdot.gov/ADOTLibrary/publications/project_reports/PDF/AZ606.pdf.

Surveys such as that conducted by Mallela et al.[4] have shown a significant correlation between the field-measured and predicted values for Jointed Plain Concrete Pavement (JPCP) in Colorado at a national scale. Therefore, global calibration factors have been retained. After being nationally calibrated to Florida's conditions, Oh and Fernando's[5] comprehensive study concluded that IRI and Joint Faulting were underpredicted, while the Transverse Cracking yielded an excellent correlation.

Although the literature on the local calibration under AASHTOWare Pavement ME Design is extensive, no study has examined the Quebec context. In this research work, the first attempt to improve the prediction accuracy of flexible pavement in Quebec is undertaken using the version 3 of PMED . The transfer function coefficients of performance models, including Bottom-Up Fatigue Cracking, Total Rutting, Transverse Cracking and IRI, are adjusted based on statistical bases involving bias, Standard Error of the Estimate (SEE) and the coefficient of determination.

Climate change poses a challenge to pavement performance. Bituminous materials, which are highly sensitive to temperature fluctuations and moisture, are particularly vulnerable. Addressing this problem requires proactive and coordinated strategies at both the national and local levels. Some studies focused on the change in binder grade due to temperature rise in southern Canada[6]. The authors revealed that climate change does not affect low-temperature cracking in future years. In Ontario, Basit et al.[7] reclaimed a change in performance-grade asphalt cement with the increase in temperature. Other studies examined the impact of climate change on surface Hot Mix Asphalt layer modulus. A declining trend was observed with the increase in temperature, and pavements are more prone to deformation[8,9] . In Qiao et al.[10], the authors investigated the effect of increased precipitations on the bonding quality of Hot Mix Asphalt Layers. Indeed, water ingress through initiated cracks in the pavement surface will likely weaken the structure and result in potholes. In line with previous works, this paper emphasizes the impact of climate change on the mechanical performance of flexible pavements in Quebec over 80 years. Section one gives a brief overview of the site selection process. The primary and second objectives of this study are then outlined: First, the performance models are locally calibrated considering local conditions of Quebec, particularly traffic and climate specific to the region. Once calibrated, this paper assesses the long-term evolution of climate parameters as well as performance models over an 80-year horizon. The subsequent sections will draw results, discussion, and main conclusions.

[4] Mallela, J., Titus-Glover, L., Sadasivam, S., Bhattacharya, B.B., Darter, M., & Von Quintus, H. (2013). Implementation of the AASHTO Mechanistic-Empirical Pavement Design Guide for Colorado. Technical Report, Colorado Department of Transportation and Research Branch.

[5]Oh, J., & Fernando, E.G. (2008). Development of Thickness Design Tables Based on the M-E PDG, Florida Department of Transportation.

[6] Mills, B.N., Tighe, S.L., Andrey, J., Smith, J.T., & Huen, K. (2009). Climate Change Implications for Flexible Pavement Design and Performance in Southern Canada. *Journal of Transportation Engineering*, vol. 135, p. 773-782.

[7] Abdul Basit, M. S., Rashid, B., & Matthew., A. (2022). Climate change and asphalt binder selection across Ontario: A quantitative analysis towards the end of the century. *Construction and Building Materials*, Volume 361.

[8] Kumlai, S., Jitsangiam, P., & Pichayapan, P. (2017). The implications of increasing temperature due to climate change for asphalt concrete performance and pavement design. *KSCE Journal of Civil Engineering 21*, 1222–1234. https://doi.org/10.1007/s12205-016-1080-6

[9] Mallick, R. B., Radzicki, M. J., Daniel, J. S., & Jacobs, J. M. (2014). Use of system dynamics to understand long-term impact of climate change on pavement performance and maintenance cost. *Transportation Research Record*, 2455, 1–9. https://doi.org/10.3141/2455-01

[10] Qiao, Y., Casey, D.B., Kuna, K.K., Kelly, K., & Macgregor, I.D. (2016). Climate resilience of flexible pavement highways: assessment of current practice. *Proceedings of the LJMU Annual International Conference on Asphalt*, Pavement Engineering and Infrastructure, Liverpool, UK, Liverpool John Moores University (LJMU). pp. 1–14.

2. Site selection

A total of 353 flexible pavements were selected across the province of Quebec for the distress prediction models. A complete database for the local calibration under PMED, covering different traffic levels (low, intermediate and high) and different climatic zones (1, 2 and 3), was thus established. The criteria used to decide whether the pavement section is to be retained or eliminated include:

- Pavement surfacing must be between 6 and 25 years old to be retained ;
- The trend of the damage response versus time ;
- The relevance of the data and the information available on the pavements ;
- The presence or absence of a bridge at the beginning of the pavement section ;
- The presence of geotextiles and/or an insulating layer ;
- The intersection with railways.

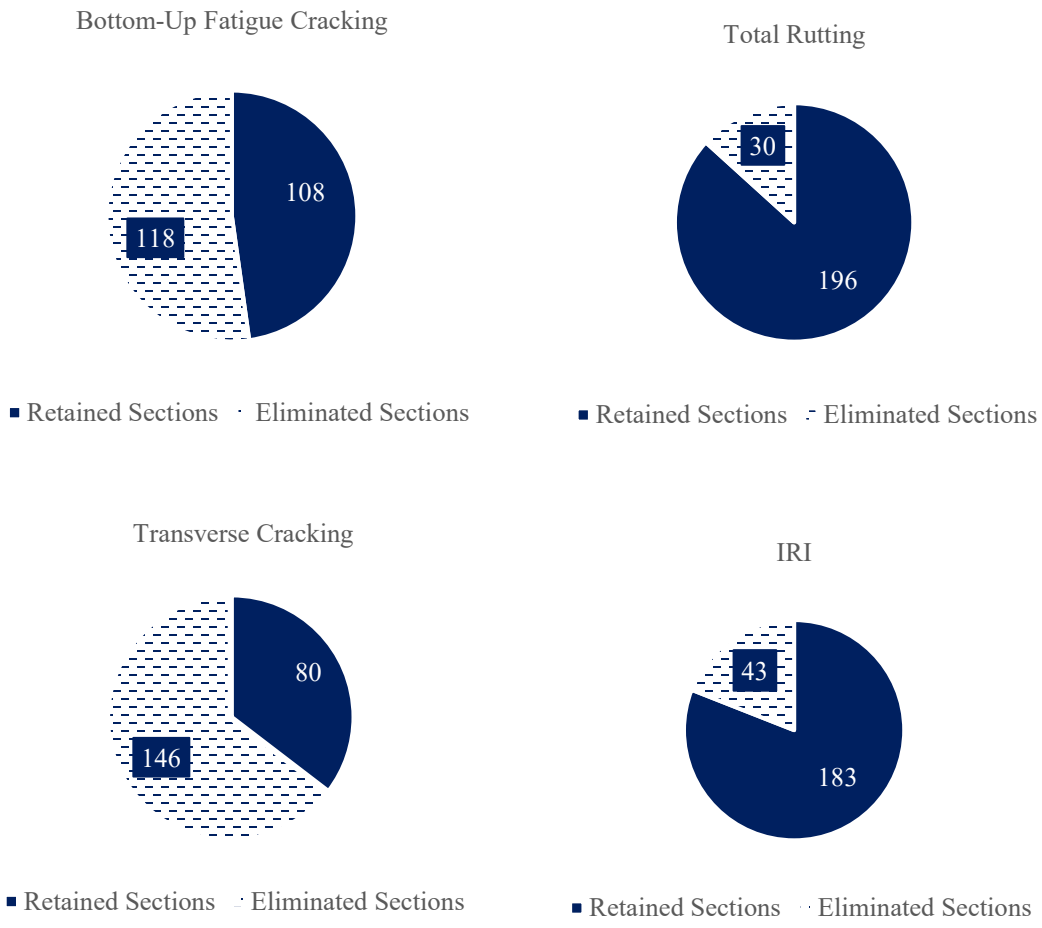
Statistics on retained sections are reported in Figure 1. The number of retained sections varies across different distress types as one or more performance models may be excluded from a given section (e.g., Bottom-Up Fatigue Cracking), while it may still be included for other distress types, such as Total Rutting, Transverse Cracking and IRI.

It should be noted that the pavement layer materials, subgrade characteristics, pavement distresses data as well as traffic and axle load data were sourced from the documentation provided by the Ministère des Transports et de la Mobilité Durable (MTMD). Data were extracted from the roadway management system database “GCH-6011” maintained by the MTMD. Geometric and physical characteristics of road segments were evaluated using the image diffusion DIR system “DIR”. Besides the GCH-6011 database, the traffic data were also sourced from the MTMD’s CIR-6002 system.

Based on the Guide for the Local Calibration of the Mechanistic-Empirical Pavement Design Guide[11], the retained sections were split automatically for the calibration-validation procedure considering the jackknifing approach. For the various prediction models: Bottom-Up Fatigue Cracking (FC), Total Rutting (R), Transverse Cracking (TC) and International Roughness Index (IRI), 80% of the retained sections were used int the calibration step where the remaining 20% were used in the validation process.

[11] 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.

Figure 1. Statistics on selected sites for the various prediction models



3. Optimization of local calibration coefficients for flexible pavements in Quebec

3.1. Optimization approach

The calibration exercise was based mainly on the approach described in the AASHTO 2010 guide[11]. It examines the correlation between the predicted values and those measured in the field for different optimization scenarios. Initial verifications for the various prediction models regarding the global calibration coefficients were first conducted. The Calibration Assistance Tool (CAT) of the AASHTOware Pavement ME Design was used. The accuracy of the predictions is assessed using several statistical measures, including bias, the Standard Error of Estimate (SEE), the coefficient of determination R^2 and the null hypothesis. The bias and SEE reflect the errors in the calculation and the quality of the dispersion of the data around their mean value. A low bias and SEE, ideally close to zero, increase the probability that the predictions converge toward the field measurements. In addition to the bias and the SEE, the quality of the prediction was examined using the coefficient of determination R^2 and null hypothesis or the probability of occurrence of a given event p-value. It defines the probability of reaching a predicted value equal to the measured one. The first measure reflects the linear relationship between the predicted and measured data in the field. It is expressed as the ratio between the sum of the squares of the standard deviations predicted by the CAT and their mean. The second parameter checks whether the measured

and predicted values come from different populations. It should be greater than or equal to 0.05 for the simulation to be accepted[11].

At this stage, the decision to undertake the local calibration is conditioned by the statistical analysis results yielded by the initial verification and the scatter plot around the line of equality. A local calibration seems necessary in the case of a rejected null hypothesis with an aberrant bias and SEE that indicates an overestimation or underestimation of the real measurements in the field. Minimum and maximum values are set for the different calibration coefficients. They define an interval that covers the predefined global value of the calibration coefficient. Each time the proposed values (Minimum and maximum) are adjusted until the statistical measures and data dispersion are improved. In fact, during the calibration process, different iterations involving one or a combination of calibration coefficients are run under the CAT to reduce the statistical measures and optimize the dispersion of data around the line of equality. This survey allows for the adjustment of the tested calibration coefficients each time. The set of calibration parameters considered for each prediction model was mainly:

- β_{f1}, β_{f2}, β_{f3}, C_1 and C_2 for Bottom-Up Fatigue Cracking ;
- β_{r1}, β_{r2}, β_{r3} and β_{s1} for Total Rutting ;
- k_t for Transverse Cracking ;
- C_1, C_2, C_3 and C_4 for the International Roughness Index.

The local calibration process is then completed once the suggested values sufficiently reduce the bias and the Standard Error of Estimation, the null hypothesis is accepted, and the quality of measurement dispersion is satisfactory.

3.2. Bottom-Up Fatigue Cracking

Although Bottom-Up and Top-Down Fatigue Cracking mechanisms are fundamentally distinct, it is important to note that both categories were merged during the calibration process. This could be explained by the fact the identification of crack types, whether originate from the bottom or the top, requires an effective field identification method, that remains an evolving process. In practice, cracking data from sections with a high likelihood of Bottom-Up Fatigue are used to calibrate the performance model associated with this type of cracking. Similarly, data from sections with a high probability of Top-Down cracking are used to calibrate the corresponding model. No individual crack-level classification is performed; instead, all cracks within a reference section are assumed to be of the most probable type. Since this information was not obtained within the scope of this project, both categories were combined during the calibration process. In several optimization attempts using the CAT [12,13], identical results were reported for Bottom-Up Cracking despite the separate consideration of Top-Down and Bottom-Up cracking.

The global calibration of the Bottom-Up Fatigue Cracking model showed poor data dispersion, with most points straying from the 1:1 line. Figure 2-a illustrates the scatter plot clusters along the horizontal and vertical axes. Statistical measures yielded -1.58 and 5.73 for the bias and SEE, respectively. The null

[12] Gong, H., Huang, B., Shu, X., & Udeh. S. (2017). Local calibration of the fatigue cracking models in the Mechanistic-Empirical Pavement Design Guide for Tennessee. *Road Materials and Pavement Design*, 18(sup3), 130–138. https://doi.org/10.1080/14680629.2017.1329868

[13] Idaho Transportation Department. (2019). Calibration of the AASHTOWare Pavement ME Design Software for PCC Pavements in Idaho, 122.p.

hypothesis was rejected as the p-value value of 0.00229 was lower than 0.05. As per the Local Calibration Guide [11], the reduction of bias needs the optimization of C_2 and β_{f1} and reduction of the Standard Error of Estimate involves β_{f2}, β_{f3} and C_1. The process is iterative: several simulations were run under the CAT tool in terms of a single calibration coefficient β_{f1} or β_{f2} or β_{f3} or C_1 or C_2 or a combination of them. The iteration involving C_1 = 0.5 and C_2 = 0.5 for layer thicknesses between 5 and 12 inches significantly improved the dispersion of data around the line of equality. To further refine this result, the impact of the coefficients β_{f1}, β_{f2}, and β_{f3} was tested. The most promising scenario was thus obtained in the case of C_1= 0.5, C_2 (for 5–12 inches) = 0.5, and β_{f3} = 0.85. The other coefficients were kept at their global values and were not modified. The optimum scenario was retained based mainly on a significant improvement in the distribution of data points around the line of equality (Figure 2-b and Table 2). Bias and SEE were improved compared to global calibration: 0.66 and 4.15 against -1.58 and 5.73, respectively. The parameter R^2 remained equal to 0.01 in both initial verification and local calibration process. It should be noted that the null hypothesis was a no-decisive criterion in the case of local calibration of the Bottom-Up Fatigue Cracking model. The locally calibrated Bottom-Up Fatigue Cracking model was validated based on a set of 21 out of 107 sections. This process examines the adequacy of the local calibration results and ensures that requirements in terms of dispersion of data as well as the statistical measures are met. The validation sample size was arbitrarily selected under the CAT tool for 80/20 split rule. The bias and SEE were reduced to 0.08 and 2.97. A satisfactory scatter plot was also observed. Comparing the model's predictions with observed measurements yielded promising results (Figure 2-c and Table 1), confirming the reliability of the selected calibration factors.

Figure 2. Measured vs. predicted values for the Bottom-Up Fatigue Cracking model: a) Initial verification, b) local calibration and c) validation results

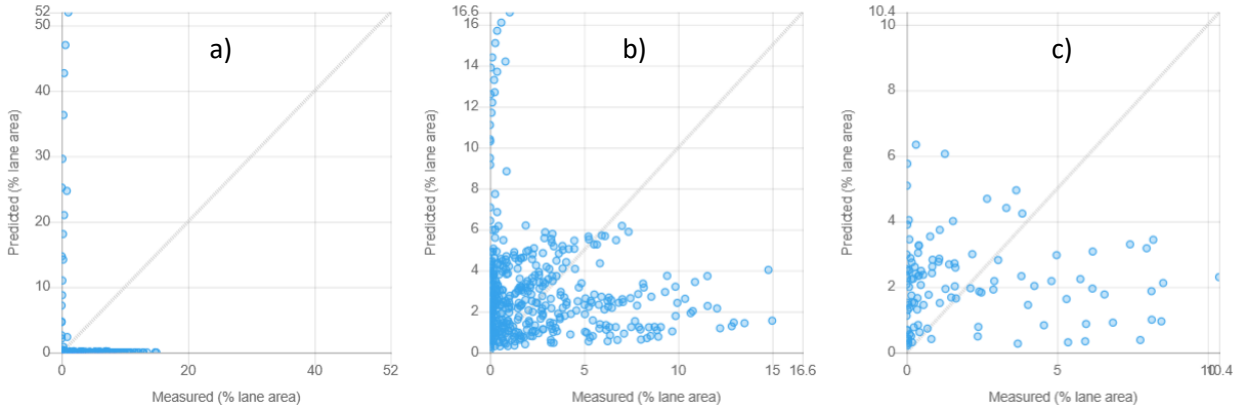


Table 1. Summary of statistical measures for the Bottom-Up Fatigue Cracking model

	Initial verification	Local calibration	validation
Number of sections	107	86	21
Bias	-1.58	0.66	0.08
SEE	5.73	4.15	2.97
R^2	0.01	0.01	0.00
p-value	0.00000	0.00058	0.78123

Table 2. Initial verification and local calibration results for the Bottom-Up Fatigue Cracking model

Coefficients	β_{f1}			β_{f2}	β_{f3}	C_1	C_2		
	$Th. < 5\ inch$	$5inch \leq TH. \leq 12inch$	$Th. > 12\ inch$				$Th. < 5\ inch$	$5inch \leq TH. \leq 12inch$	$Th. > 12\ inch$
Global value	0.02054	0	0.001032	1.38	0.88	1.31	2.1585	0	3.9666
Local Value	0.02054	0	0.001032	1.38	**0.85**	**0.5**	2.1585	**0.5**	3.9666

3.3. Total Rutting

Agreement between predictions and observed measurements was yielded by the initial verification of the Total Rutting model (Figure 3-a). 192 out of 226 pavement sections were used to this end. Bias and SEE were 5.02 and 7.14, respectively. The primary focus was to enhance the scatter plot around the line of equality and reduce the statistical measures. Based on the Local Calibration Guide recommendations, the best way to reduce the Bias should involve the parameter β_{r1}, while the reduction of the SEE depends on the β_{r2}, β_{r3}, β_{s1} and β_{sg1}coefficients[11]. The approach followed in the calibration process of the Total Rutting model was iterative: adjustments were made several times until the predicted measures matched most effectively experimental data. Initially, various combinations of calibration coefficients were tested. Only combinations involving both β_{s1} for the Granular Base and $\beta_{sg1}$the Subgrade allowed better correlation with experimental data. Multiple variants were explored. Once setting the coefficient β_{s1} for the Granular Base and β_{sg1} for the Subgrade to 0.3 instead of 1, the most optimum scenario was depicted (Figure *3*-b). Interestingly, Bias and SEE were optimized: -0,03 and 4.59, respectively (global values: 5.02 and 7.14) (Table 3). However, the null hypothesis was accepted: 0.83065 compared to 0.05. A set of 38 sections were used in the validation process of the Total Rutting model. Statistical measures of calibrated and validated Total Rutting model against initial verification were reported in Figure *3* and Table 3.

Figure 3. Measured vs. predicted values for the Total Rutting model: a) Initial verification, b) local calibration and c) validation results

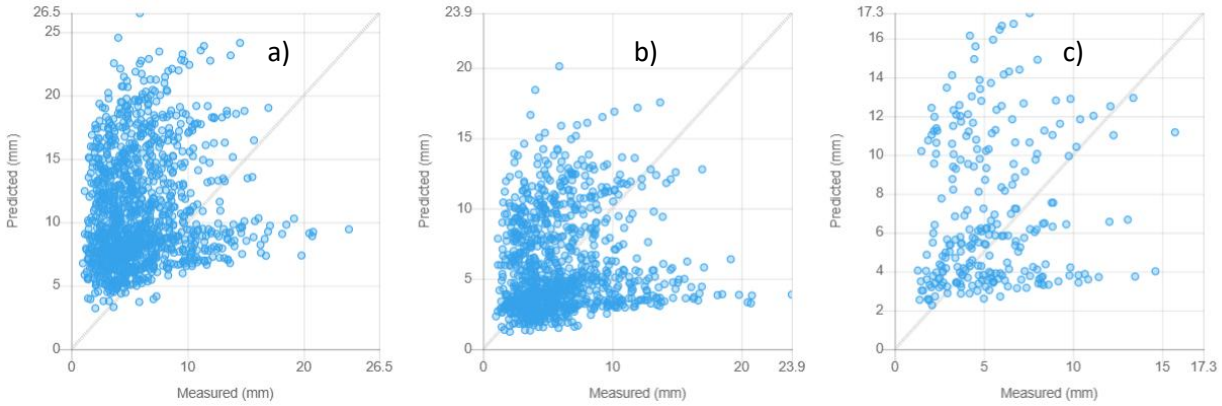


Table 3. Summary of statistical measures for the Total Rutting model

	Initial verification	Local calibration	Validation
Number of sections	192	154	38
Bias	5.02	-0.03	1.63
SEE	7.14	4.59	4.67
R^2	0.01	0.01	0.01
p-value	0.00000	0.83065	0.00000

Table 4. Initial verification and local calibration results for the Total Rutting model

Coefficients	β_{r1}	β_{r2}	β_{r3}	β_{s1} Granular Base	β_{sg1} Subgrade
Global value	0.4	0.52	1.36	1	1
Local Value	0.4	0.52	1.36	**0.3**	**0.3**

3.4. Transverse cracking

k_t is the key parameter in the calibration procedure of the Transverse Cracking model. It is highly dependent on the average annual air temperature $MAAT$. Three levels (1, 2 and 3) and two temperature ranges ($> 57°F$ and $\leq 57°F$) are used to define the k_t parameter under the CAT. Referring to Table *5*, the Transverse Cracking of the flexible pavement in Quebec was underpredicted. Initial verification conducted by mean of the global coefficient $k_t = 0$ (Table 5) yielded a very high statistical measurement (Bias =-51.63 and SEE= 80.15). Nevertheless, measured thermal cracks were observed without leading to predictions (Figure *4*-a). Notably, 80 sections were retained during the initial verification procedure. Different ranges for k_t have been tested in order to optimize the preliminary results. No significant improvement has been observed. Data points were always clustered along the x-axis. In such circumstances, wider ranges of k_t parameter have been proposed. Multiple simulations were run, for instance a minimum value of 1, a maximum value of 50 and an intermediate value within the range [1,50] was tested. Despite those efforts, no thermal cracking was predicted, while significant thermal cracking was observed in the field (Figure *4*-b). In addition, no change in the statistical measures was depicted in the locally calibrated Transverse Cracking model. This apparent lack of correlation can be justified because the Transverse Cracking model under the Pavement ME Design is inappropriate for Quebec. No validation was undertaken in this case. These results correlate with previous research conducted across the United States that further supports the idea that local calibration of Transverse Cracking cannot provide satisfactory predictions[14,15].

[14] Schwartz, C.W., Li, R., Kim, S., Ceylan, H., & Gopalakrishnan, K. (2011). Sensitivity evaluation of MEPDG performance prediction. Contractor's Final Report of National Cooperative Highway Research Program 1-47, Transportation Research Board, National Research Council, Washington, D.C.

[15] Hall, K. D., Xiao, D. X., & Wang, K. C. P. (2011). Calibration of the MEPDG for flexible pavement design in Arkansas. Transportation Research Record, 2226, 135-141, Transportation Research Board, National Research Council, Washington, D.C.

Figure 4. Measured vs. predicted values for the Transverse Cracking model: a) Initial verification and b) local calibration

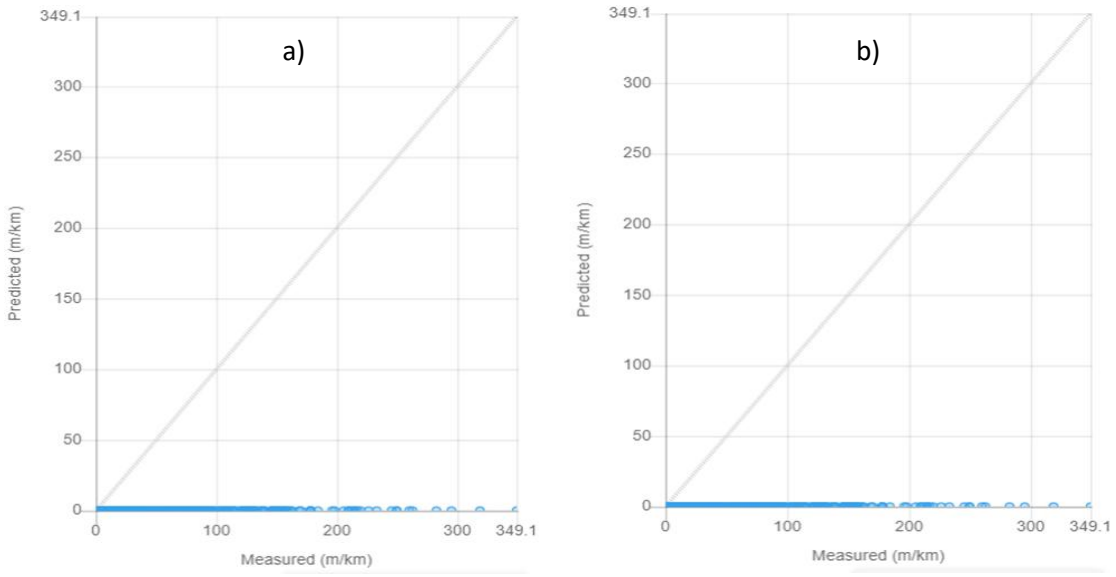


Table 5. Summary of statistical measures for the Transverse Cracking model

Coefficients	k_t
Global value	0
Local Value	Not defined

Table 6. Initial verification and local calibration results for the Transverse Cracking model

	Initial verification	Local calibration
Number of sections	80	Not defined
Bias	-51.63	Not defined
SEE	80.15	Not defined
R^2	0.00	Not defined
p-value	0.00	Not defined

3.5. International Roughness Index

Particular attention was paid to the initial roughness value IRI_0 to be defined during the calibration process. This parameter is not considered in the CAT. Road infrastructure agencies have made several attempts to provide answers regarding initial IRI to be considered during optimization. One of the approaches adopted was based on introducing initial IRIs measured in the field as a level 1 input for calibration. The Ontario Ministry of Transportation[16] provides a range between 0.65 and 1 m/km for the initial IRI of new pavements. This is consistent with the default value defined under the CAT. A wider

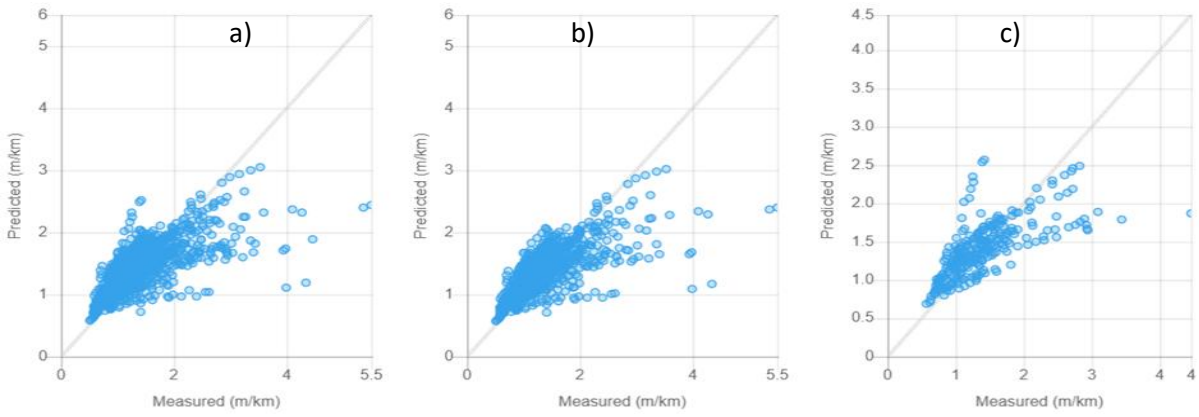
[16] Ontario Ministry of Transportation. (2014). Ontario's Default Parameters for AASHTOWare Pavement ME Design, Interim Report. Downsview, ON: Ontario Ministry of Transportation, Pavement and Foundations Section, Materials Engineering and Research Office.

range, between 0.72 and 1.97 m/km, is defined for the IRI of the first year of the pavement in the Superpave calibration database. For this study, an effort to specify the initial roughness value IRI_0 for all the selected projects for the IRI model calibration was made. This allows a rigorous reflection of field-measured data instead of adopting the default value of 1 m/km under CAT, which is consistent with the manual's specifications[17].

It should be noted that, the optimization process of the IRI model was conducted following the local calibration of Bottom-Up Fatigue Cracking, Total Rutting and Transverse Cracking. Local calibration in the case of the IRI model was performed to optimize the initial verification results, mainly data variability in the predicted vs. field measurement scatter plot. Global calibration was undertaken in terms of $C_1 = 40$, $C_2 = 0.4$, $C_3 = 0.008$ and $C_4 = 0.015$ revealed a Bias of 0.03 and a Standard Error of Estimate of 0.41. The null hypothesis was rejected (Figure 5-a and Table *7*). First, attempts were made to optimize the C_1 coefficient; the bias and the standard error of estimate decreased from 0.03 and 0.41 to 0 and 0.40, respectively. The p-value was 0.88162. This result was for the scenario involving a value of C_1=35. The correlation between actual and predicted measurements for iterations based on the C_4 coefficient generated an optimal scenario for a value of 0.01. For this scenario, the p-value of 0.21233 greatly exceeded 0.05. In turn, the bias was -0.02. It should be noted that iterations involving the C_2 and C_3 coefficients separately or combined with the other C_i coefficients were not satisfactory. After several simulations, the statistical measures were found to be satisfactory for the scenario involving the coefficients $C_1 = 35$ and $C_2 = 0.5$ (Figure 5-b, Table *7* and Table *8*). A zero bias, a standard deviation of 0.4 and a p-value of 0.99 were observed.

A set of 36 sections were used to validate the International Roughness Index model. A p-value of 0.63305 generated during validation allowed the acceptance of the null hypothesis condition. A lower bias was observed compared to the initial verification (Figure *5*-c and Table *8*).

Figure 5. Measured vs. predicted values for the IRI model: a) Initial verification, b) local calibration and c) validation results



[17] American Association of State Highway and Transportation Officials. (2014). AASHTOware Pavement M-E Design v2.0 Help Manual. American Association of State Highway and Transportation Officials.

Table 7. Summary of statistical measures for the IRI model

	Initial verification	Local calibration	Validation
Number of sections	183	147	36
Bias	0.03	0.00	0.01
SEE	0.41	0.40	0.41
R^2	0.52	0.52	0.48
p-value	0.00242	0.99094	0.63305

Table 8. Initial verification and local calibration results for the IRI model

Coefficients	C_1	C_2	C_3	C_4
Global value	40	0.4	0.008	0.015
Local Value	35	0.5	0.008	0.015

4. Climate projections

The local calibration coefficients identified in the previous sections for the various performance models were explored to assess the impact of climate change on the response of flexible pavements in Quebec. Besides the default climate baseline of the Modern Era Retrospective Reanalysis 2 (MERRA 2) model integrated into PMED, two future climate databases RCP4.5 and RCP8.5, developed by McGill University specifically for the period 1989-2069, were used. Both define different greenhouse gas emission scenarios (GHG). The trend either gradual or rapid correspond to a stabilized scenario (RCP4.5) or a high emission or more pessimistic scenario (RCP8.5), respectively. RCP4.5 features low to moderate GHG emissions, while RCP8.5 is characterized by high GHG emissions. Five physical entities were considered: air temperature, cloud cover, wind speed, relative humidity, and precipitation. For that, several .hcd files were generated using MATLAB software. They cover the geographical area of Quebec where the roadways monitored by the MTMD are located; this geographical area corresponds to southern Quebec. Each *.hcd file corresponds to a weather station. In total, six road sections were defined. They were considered typical and representative of the conditions and practices of Quebec in terms of functional classes, traffic, soils, climatic regions, etc.

First, the impact of climate change on future temperature, precipitation, the number of freeze-thaw cycles, the freezing index, the cloud cover and percent sunshine under baseline scenario (MERRA 2) and future scenarios (RCP 4.5 and RCP 8.5) was examined. Then, the focus was to depict the effect of climate change on the performance of flexible pavements in Quebec.

4.1. Evolution of climatic parameters over an 80-year horizon

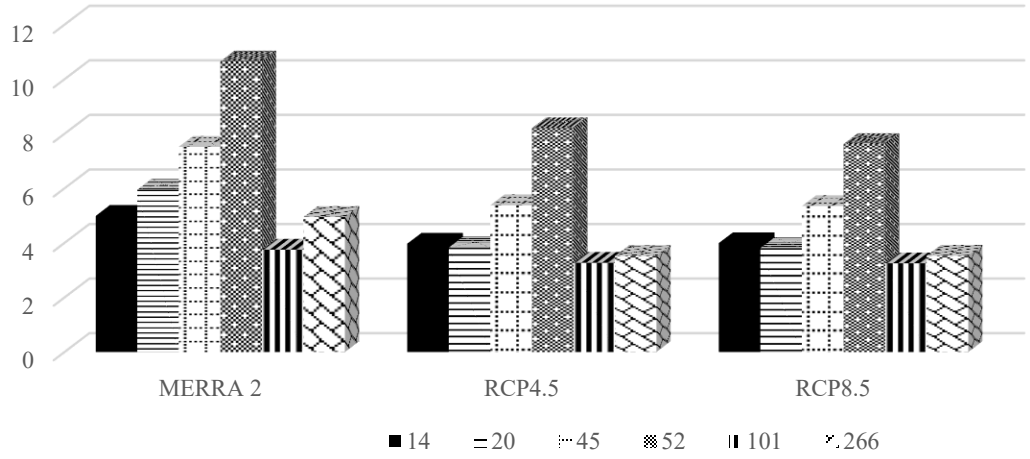
Over an 80-year horizon, climate projections have revealed an increase in temperature. In the long term, this would impact winter conditions in Quebec because milder winter temperatures will mark winters. A decrease in frost depth would also be expected. A decrease in precipitation has also been observed. As a result, water content in road surfaces and groundwater levels are reduced. A decrease in the number of freeze-thaw cycles and the frost index in the future will likely generate positive effects associated with

climate change in winter. According to several studies [18,19,20,21], future precipitation trends will vary by region and season, possibly increasing or decreasing depending on geographic location.

4.2. Analysis of Bottom-Up Fatigue Cracking over an 80-year horizon

Compared to baseline climate (MERRA 2), a reduction ranging between 0.48 and 3.00 % in Bottom-Up Fatigue Cracking over an 80-year horizon was observed for the flexible pavements in Quebec (Figure *6*). Even though this decrease is relatively small, it still reflects a potential improvement in long-term pavement durability. Therefore, climate change could have a slight positive effect on fatigue related distress of flexible pavements in Quebec. This could be explained not only by the decrease in precipitation but also by the reduction in the freezing index and the number of freeze-thaw cycles observed for the province of Quebec [6,22,23]. In the future, pavements are expected to be less prone to water infiltration which makes the structure more resistant and could mitigate crack initiation. Yet, reduced freeze episodes are likely to enhance the pavement's thermal resistance.

Figure 6. Evolution of Bottom-Up Fatigue Cracking according to different climate scenarios



[18] CARBONBRIEF. (2018). Climate Modeling, explainer: What climate models tell us about future rainfall [Online]. [https://www.carbonbrief.org/explainer-what-climate-models-tell-us-about-future-rainfall/]

[19] Easterling, D. R., Kunkel, K. E., Arnold, J. R., Knutson, T. R., LeGrande, A. N., Leung, L. R., Vose, R. S., Waliser, D. E., & Wehner, M. (2017). Precipitation change in the United States. Climate Science Special Report: Fourth National Climate Assessment, I, 207–230

[20] Government of Canada. (2019). Change in precipitations [Online] [https://www.canada.ca/fr/environnement-changement-climatique/services/changements-climatiques/centre-canadien-services-climatiques/essentiels/tendances-projections/changements-precipitations.html]

[21] Qiao, Y., Guo. Y., Stoner, A. M.K., & Santos, J. (2022). Impacts of future climate change on flexible road pavement economics: A life cycle costs analysis of 24 case studies across the United States. *Sustainable Cities and Society*, vol. 80, https://doi.org/10.1016/j.scs.2022.103773

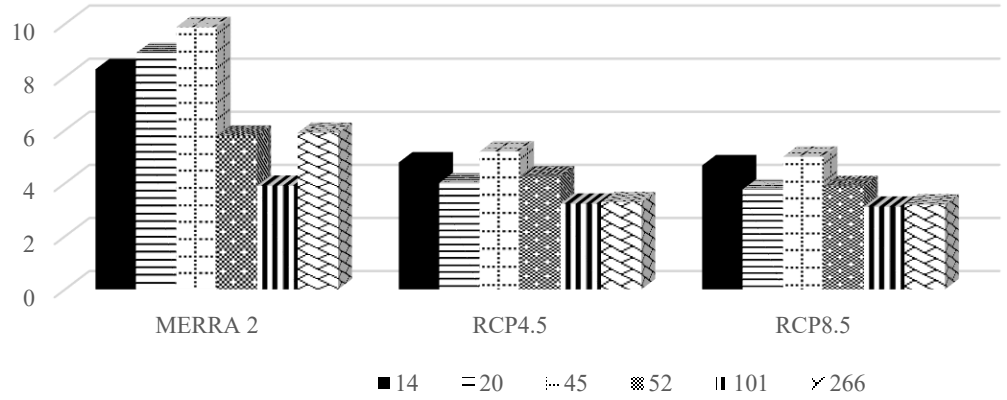
[22] Swarna, S. T. (2021). Influence of Climate Change on Pavement Design and Materials in Canada. Ph.D. thesis, Memorial University of Newfoundland.

[23] Swarna, S, T., Hossain, K., Mehta, Y.A., & Bernier, A. (2022). Climate Change Adaptation Strategies for Canadian Asphalt Pavements; Part 1: Adaptation strategies. *Journal of Cleaner Production*, vol. 363.

4.3. Analysis of Total Rutting over an 80-year horizon

Over an 80-year horizon, the Total Rutting of flexible pavements in Quebec will likely decrease due to climate change (Figure 7). Even though an increase in future air temperature is revealed in this study, the decreasing trend of Total Rutting may be explained by other key factors. Increased wind speed and cloud cover can lower the pavement surface temperature[24,25]. This cooling effect could change the stiffness of the asphalt surface layer as well as the mechanical behavior of the pavement structure. This could be attributable also to a reduction in the freezing index, the number of freeze-thaw cycles, and decreased precipitation[6,22,23]. Taken together, these factors lead to less water infiltration and improve overall pavement resistance to rutting.

Figure 7. Evolution of Total Rutting according to different climate scenarios



4.4. IRI analysis over an 80-year horizon

Compared to the reference climate, the International Roughness Index has decreased over 80 years (Figure 8). Considering the anticipated reduction in Total Rutting and Bottom-Up Fatigue Cracking, this result was expected.

In addition, the increase in temperatures as well as the reduction in the freezing index and the number of freeze-thaw cycles can indirectly lead to a reduction in IRI over time. A similar decreasing trend in IRI was observed in the provinces of Quebec, Newfoundland and Labrador [6,26,27,28].

[24] Qin Y., & Hiller, J.E. (2013). Ways of formulating wind speed in heat convection significantly influencing pavement temperature prediction. *Heat Mass Transf*, vol 9(5), pp 745–752.

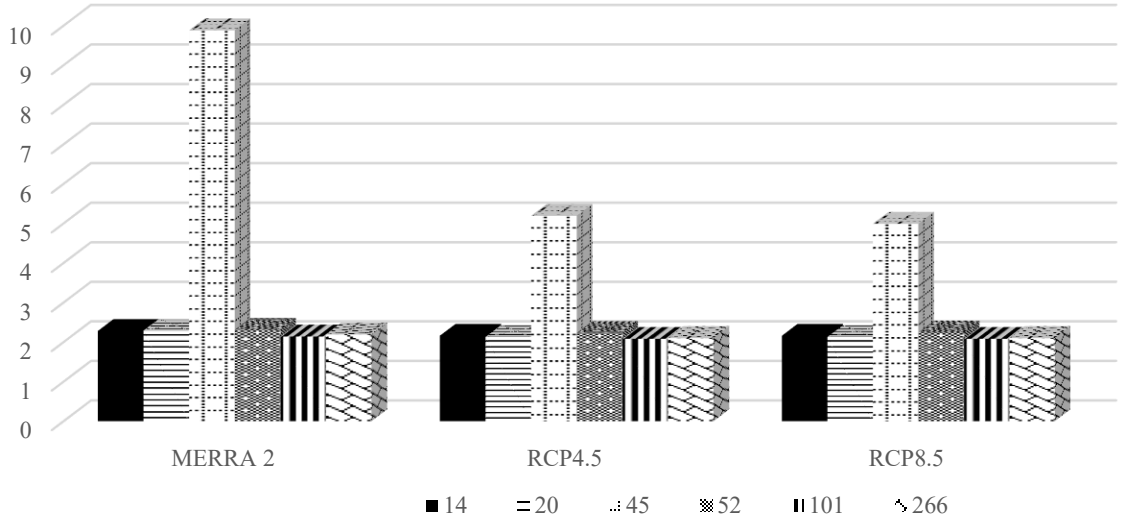
[25] Ling, J., Xing, X., Zhang, J., & Liu, S. (2024). An Experimental Validation-Based Study of Airport Pavement Icing Mechanisms in Saline Environments and the Development of a Simplified Prediction Model. *Journal of Applied Science*, vol 14, 8867. https://doi.org/10.3390/app14198867

[26] Bizjak, K. F., Dawson, A., Hoff, I., Makkonen, L., Yihaisi, J. S., & Carrera, A. (2014). The impact of climate change on the European road network. *Proceedings of the Institution of Civil Engineers – Transport*. https://doi.org/10.1680/tran.11.00072

[27] Doré, G., Bilodeau, J. P., Thiam, P. M., & Drolet, F. P. (2014). Impact des changements climatiques sur les chaussées des réseaux routiers québécois. Université Laval.

[28] Drolet, F. P. (2015). Effets des changements climatiques sur la performance à long terme des chaussées souples au Québec – Volet 2 : Effet de l'augmentation de la température en hiver et d'une hausse du nombre d'épisodes de redoux hivernaux. Mémoire de maîtrise en génie civil, Université Laval.

Figure 8. Evolution of the IRI according to different climate scenarios



4.5. Analysis of Transverse Cracking over an 80-year horizon

This project's measurements remained relatively constant across all sites under all three climate scenarios[6]. A value of 40.97 m/km was recorded.

5. Conclusion

This study leverages AASHTOWare PMED, locally calibrated, to enhance the accuracy of flexible-pavement performance predictions in Quebec. Four performance models were evaluated: Bottom-Up Fatigue Cracking (FC), Total Rutting (R), Transverse Cracking (TC) and the International Roughness Index (IRI). A dataset of 353 flexible-pavement sections managed by the Ministère des Transports et de la Mobilité Durable (MTMD) was assembled. They were selected according to established quality and representativeness criteria. Transverse and longitudinal cracking models were held fixed to maximize consistency between global and regional calibrations (i.e., not recalibrated). The main findings are as follows:

- Data points clustered in the global calibration of the Bottom-Up Fatigue Cracking (FC) model, indicating poor dispersion around the 1:1 line. Local calibration substantially improved the alignment of measurements with the equality line.
- Using global coefficients, the CAT tool overestimated the Total Rutting of flexible pavements in the Quebec road network. During local calibration, the most optimal scenarios involved both the β_{s1} (Granular Base) and β_{sg1} (Subgrade) coefficients. The one that offered the best dispersion of measurements around the equality line, thus a p-value greater than 0.05 of the order of 0.83, was ultimately retained.
- Due to the Transverse Cracking prediction model present in AASHTOware PMED, null predictions were generated. The inadequacy of the K_t relation to Quebec conditions was put forward to explain such a result. In addition, the initial bitumen class of the PG 58-34 project may be inappropriate for local Quebec climatic conditions. Future efforts to develop new expressions specific to local Quebec conditions are strongly recommended.

- In the International Roughness Index (IRI) case, the global calibration yielded a strong correlation between predicted and observed values, with low bias and standard error of estimate. Local refinement of the calibration coefficients focused on maximizing the p-value, ultimately achieving p = 0.99 and satisfying the null hypothesis test. Crucially, each project's initial IRI was specified during local calibration—replacing the CAT tool's default value of 1.0. Building on Ontario's findings[29], we strongly recommend adopting a standardized initial IRI in future designs—tailored to road-agency specifications—to enhance consistency. However, this approach may understate residual variability and thus overstate the apparent reliability of the calibrated model.

Long-term projections for Quebec's flexible pavements indicate declines in Bottom-Up Fatigue Cracking, Total Rutting and the International Roughness Index (IRI).

- Bottom-Up Fatigue cracking is expected to decrease mainly because of reduced precipitation, a lower freezing index and fewer freeze-thaw cycles.
- Total Rutting (and the resulting IRI improvement) should diminish, driven by the same freeze-thaw factors, reduced rainfall, rising wind speeds, and increased cloud cover.

[29] Yuan, X.X., Lee, W., & Li, N. (2017). Ontario's local calibration of the MEPDG distress and performance models for flexible roads: A summary

References

1. Ahmed, S.E., Yuan, X.X., Lee, W., & Li, N. (2016). Local Calibration of Cracking Models in MEPDG for Ontario's Flexible Pavement, *Proceedings of the Transportation Association of Canada Conference*, Toronto, Ontario, Canada.
2. Dong, Shi., Yuan, X. X., & Hao, PW. (2020). An advanced local calibration method for mechanistic-empirical pavement design. *Computer-Aided Civil and Infrastructure Engineering*. 35. 10.1111/mice.12574.
3. Darter, M.I., Titus-Glover, L., Von Quintus, H., Bhattacharya, BB., & Jagannath. M. (2014). Calibration and Implementation of the AASHTO Mechanistic-Empirical Pavement Design Guide in Arizona. Arizona Departement Transportation Research Center, https://apps.azdot.gov/ADOTLibrary/publications/project_reports/PDF/AZ606.pdf.
4. Mallela, J., Titus-Glover, L., Sadasivam, S., Bhattacharya, B.B., Darter, M., & Von Quintus, H. (2013). Implementation of the AASHTO Mechanistic-Empirical Pavement Desgin Guide for Colorado. Technical Report, Colorado Department of Transportation and Research Branch.
5. Oh, J., & Fernando, E.G. (2008). Development of Thickness Design Tables Based on the M-E PDG, Florida Department of Transportation.
6. Mills, B.N., Tighe, S.L., Andrey, J., Smith, J.T., & Huen, K. (2009). Climate Change Implications for Flexible Pavement Design and Performance in Southern Canada. *Journal of Transportation Engineering*, vol. 135, p. 773-782.
7. Abdul Basit, M. S., Rashid, B., & Matthew., A. (2022). Climate change and asphalt binder selection across Ontario: A quantitative analysis towards the end of the century. Construction and Building Materials, Volume 361.
8. Kumlai, S., Jitsangiam, P., & Pichayapan, P. (2017). The implications of increasing temperature due to climate change for asphalt concrete performance and pavement design. KSCE Journal of Civil Engineering 21, 1222–1234. https://doi.org/10.1007/s12205-016-1080-6
9. Mallick, R. B., Radzicki, M. J., Daniel, J. S., & Jacobs, J. M. (2014). Use of system dynamics to understand long-term impact of climate change on pavement performance and maintenance cost. *Transportation Research Record*, 2455, 1–9. https://doi.org/10.3141/2455-01
10. Qiao, Y., Casey, D.B., Kuna, K.K., Kelly, K., & Macgregor, I.D. (2016). Climate resilience of flexible pavement highways: assessment of current practice. *Proceedings of the LJMU Annual International Conference on Asphalt, Pavement Engineering and Infrastructure*, Liverpool, UK, Liverpool John Moores University (LJMU). pp. 1–14.
11. 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.
12. Gong, H., Huang, B., Shu, X., & Udeh. S. (2017). Local calibration of the fatigue cracking models in the Mechanistic-Empirical Pavement Design Guide for Tennessee. *Road Materials and Pavement Design,* 18(sup3), 130–138. https://doi.org/10.1080/14680629.2017.1329868
13. Idaho Transportation Department. (2019). Calibration of the AASHTOWare Pavement ME Design Software for PCC Pavements in Idaho, 122.p.
14. Schwartz, C. W., Li, R., Kim, S., Ceylan, H., & Gopalakrishnan, K. (2011). Sensitivity evaluation of MEPDG performance prediction. Contractor's Final Report of National Cooperative Highway Research Program 1-47, Transportation Research Board, National Research Council, Washington, D.C.

15. Hall, K. D., Xiao, D. X., & Wang, K. C. P. (2011). Calibration of the MEPDG for flexible pavement design in Arkansas.Transportation Research Record, 2226, 135-141, Transportation Research Board, National Research Council, Washington, D.C.
16. Ontario Ministry of Transportation. (2014). Ontario's Default Parameters for AASHTOWare Pavement ME Design, Interim Report. Downsview, ON: Ontario Ministry of Transportation, Pavement and Foundations Section, Materials Engineering and Research Office.
17. American Association of State Highway and Transportation Officials. (2014). AASHTOware Pavement M-E Design v2.0 Help Manual. American Association of State Highway and Transportation Officials.
18. CARBONBRIEF. (2018). Climate Modeling, explainer: What climate models tell us about future rainfall [Online]. [https://www.carbonbrief.org/explainer-what-climate-models-tell-us-about-future-rainfall/]
19. Easterling, D. R., Kunkel, K. E., Arnold, J. R., Knutson, T. R., LeGrande, A. N., Leung, L. R., Vose, R. S., Waliser, D. E., & Wehner, M. (2017). Precipitation change in the United States. Climate Science Special Report: Fourth National Climate Assessment, I, 207–230
20. Government of Canada. (2019). Change in precipitations [Online] [https://www.canada.ca/fr/environnement-changement-climatique/services/changements-climatiques/centre-canadien-services-climatiques/essentiels/tendances-projections/changements-precipitations.html]
21. Qiao,Y., Guo. Y., Stoner, A. M.K., & Santos, J. (2022). Impacts of future climate change on flexible road pavement economics: A life cycle costs analysis of 24 case studies across the United States. *Sustainable Cities and Society*, vol. 80, https://doi.org/10.1016/j.scs.2022.103773
22. Swarna, S. T. (2021). Influence of Climate Change on Pavement Design and Materials in Canada. Ph.D thesis, Memorial University of Newfoundland.
23. Swarna, S, T., Hossain, K., Mehta, Y.A., & Bernier, A. (2022). Climate Change Adaptation Strategies for Canadian Asphalt Pavements; Part 1: Adaptation strategies. *Journal of Cleaner Production*, vol. 363.
24. Qin Y., &Hiller, J.E. (2013). Ways of formulating wind speed in heat convection significantly influencing pavement temperature prediction. *Heat Mass Transf*, vol 9(5), pp 745–752.
25. Ling, J., Xing, X., Zhang, J., & Liu, S. (2024). An Experimental Validation-Based Study of Airport Pavement Icing Mechanisms in Saline Environments and the Development of a Simplified Prediction Model. *Journal of Applied Science*, vol 14, 8867. https://doi.org/10.3390/app14198867
26. Bizjak, K. F., Dawson, A., Hoff, I., Makkonen, L., Yihaisi, J. S., & Carrera, A. (2014). The impact of climate change on the European road network. *Proceedings of the Institution of Civil Engineers – Transport*. https://doi.org/10.1680/tran.11.00072
27. Doré, G., Bilodeau, J. P., Thiam, P. M., & Drolet, F. P. (2014). Impact des changements climatiques sur les chaussées des réseaux routiers québécois. Université Laval.
28. Drolet, F. P. (2015). Effets des changements climatiques sur la performance à long terme des chaussées souples au Québec – Volet 2 : Effet de l'augmentation de la température en hiver et d'une hausse du nombre d'épisodes de redoux hivernaux. Mémoire de maîtrise en génie civil, Université Laval.
29. Yuan, X.X., Lee, W., & Li, N. (2017). Ontario's local calibration of the MEPDG distress and performance models for flexible roads: A summary