ISSN: 2524-1354 (Online), ISSN: 2519-7851 (Print)
Africa Habitat Review Journal
Volume 20 Issue 3 (December 2025)
http://uonjournals.uonbi.ac.ke/ojs/index.php/ahr

A Machine Learning–Based Computational Framework for Road Performance Assessment in Developing Countries: *Evidence from Kenya*

***Amollo Otieno Kenneth, Simpson Nyambane Osano and Tony Ndungu**

Received on 8th August, 2025; Received in revised form 29th September, 2025; Accepted on 27th October, 2025.

Abstract

Road infrastructure assessment in developing countries remains fragmented. Existing methods focus on isolated dimensions such as pavement condition or geometric design, limiting their utility for comprehensive planning. This study develops the first holistic road quality assessment framework, applying knowledge distillation techniques to Kenya's national road dataset (260,773 segments). Six infrastructure dimensions—surface condition, material type, geometric design, facilities, usage patterns, and functional classification—were integrated into a universal scoring system. The Random Forest teacher model (R^2 = 0.93; MAE = 0.0051) was successfully distilled into an interpretable polynomial regression formula. Results show that surface condition has the highest influence (37.3%), followed by infrastructure facilities (20.0%), usage patterns (14.6%), and functional classification (11.6%). The framework captures the heterogeneity of Kenya's road network, where 85.1% of roads are unpaved. This work represents the first synthesis of multiple infrastructure criteria into a unified, quantitative formula. The distillation of complex machine learning into an interpretable polynomial equation enables consistent, reproducible road quality scoring. The methodology is transferable—other countries can adapt the six-dimension framework by recalibrating weights to reflect local infrastructure priorities. This provides an evidence-based tool for infrastructure planning, maintenance prioritization, and policy decision-making. Countries with existing road inventory systems can implement this framework by retraining the model on their data to derive context-appropriate dimensional weights.

Keywords: Road assessment, machine learning, knowledge distillation, infrastructure management, developing countries, road quality, pavement management and pavement condition

INTRODUCTION

Road infrastructure assessment in developing countries has traditionally relied on fragmented approaches that evaluate individual aspects of road quality in isolation. Existing methods often emphasize pavement distress, geometric design, or structural capacity without integrating these dimensions into a unified assessment (Shahin, 2005; Haas et al., 1994). Such siloed approaches create gaps in infrastructure evaluation, leading to suboptimal resource allocation, inconsistent maintenance prioritization, and limited comparability across networks.

Conventional tools such as the Pavement Condition Index (PCI) and the International Roughness Index (IRI) illustrate these limitations. PCI focuses exclusively on surface distress (ASTM, 2018), while IRI requires specialized equipment, restricting its coverage to a fraction of road networks in low- and middle-income countries (Sayers & Karamihas, 1998). Visual inspection, though widely applied, remains subjective and resource-intensive (McGhee, 2004).

More critically, developing countries face networks with extreme heterogeneity in design standards, materials, and usage patterns. In Kenya, for example, 85% of roads are unpaved, spanning international corridors, rural access roads, and urban arterials (Kenya Roads Board [KRB], 2023). Frameworks designed for homogeneous, paved networks in developed countries struggle to capture such variability, undermining the effectiveness of infrastructure management

**Corresponding author:*
***Amollo Otieno Kenneth** M.Sc. Student, PhD Civil Engineering (Transportation Engineering), Department of Civil and Construction Engineering, University of Nairobi Kenya*
Email: kennethamollo27@gmail.com

(African Development Bank [AfDB], 2018).

This study responds to these challenges by developing a holistic, data-driven framework for road performance assessment that integrates six dimensions—surface condition, material type, geometric design, facilities, usage patterns, and functional classification. Using machine learning and knowledge distillation, the framework not only enhances analytical rigor but also ensures interpretability for policy and planning in resource-constrained environments.

THEORY

Established Frameworks and their Limitations

Road assessment practices have historically relied on manual and visual inspection methods. Walking and windshield surveys provide detailed, field-level observations of pavement distress, drainage conditions, and safety features, making them foundational to condition assessment. However, they remain time-consuming, resource-intensive, and inherently subjective, with assessor variability limiting consistency across large networks (Shahin, 1994; ASTM, 2018).

The Pavement Condition Index (PCI) standardizes surface distress evaluation through weighted deduction values for specific defect types, offering reproducibility and benchmarking capability (ASTM, 2018). Yet it often struggles with contextual adaptation in tropical and low-resource settings where moisture-induced deterioration patterns differ from temperate climates. More critically, PCI focuses narrowly on surface distress without accounting for drainage adequacy, safety facilities, geometric design, or functional roles (Loizos & Plati, 2007).

The International Roughness Index (IRI) provides an objective, vehicle-response-based measure of ride quality that correlates well with user comfort and vehicle operating costs (Paterson, 1987). Its mechanical measurement reduces subjectivity, making IRI widely adopted in pavement management systems globally. However, IRI offers a reliable measure of surface roughness but provides little insight into other performance dimensions such as geometric design, traffic patterns, structural capacity, or infrastructure facilities. A road with acceptable IRI may still have inadequate shoulders, poor drainage, or unsafe horizontal curves.

Broader frameworks like HDM-4 and iRAP attempt more comprehensive evaluation but remain limited in scope. HDM-4 integrates deterioration modeling with economic analysis (Kerali, 2006), yet requires extensive calibration and high-quality input data rarely available in developing contexts. iRAP emphasizes safety through star ratings (Hughes et al., 2019), but its single-dimension focus overlooks broader quality dimensions such as surface condition and facilities. National pavement management systems, while systematic, often reduce performance evaluation to a few isolated indices, again failing to capture road quality as a multidimensional construct (Zimmerman & Peshkin, 2003).

The main limitation of these approaches is not merely their cost or technical demands, but their tunnel-vision focus on isolated indicators, which prevents holistic and generalizable assessment across diverse networks. No existing method integrates surface condition, geometric design, facilities, material type, usage patterns, and functional classification into a unified, empirically weighted quality metric - a critical gap for evidence-based infrastructure planning in developing countries.

Emerging Computational Approaches

Recent advances in computational methods are beginning to address the fragmentation inherent in traditional road assessment frameworks. Python-based machine learning ecosystems—including scikit-learn, TensorFlow, and PyTorch—have enabled data-driven models for pavement deterioration prediction and automated distress detection from image data (Pedregosa et al., 2011; Abadi et al., 2016). These approaches leverage large datasets to identify complex patterns across multiple road performance dimensions simultaneously, moving beyond the single-indicator limitations of PCI and IRI.

Geospatial platforms such as QGIS, ArcGIS, and Google Earth Engine facilitate the integration of road condition data with terrain models, climate variables, and remote sensing layers (Longley et al., 2005; Gorelick et al., 2017). This spatial integration allows planners to visualize network-wide degradation patterns and prioritize maintenance activities based on multiple criteria rather than isolated condition scores.

Edge computing and IoT-based monitoring represent another promising direction. Daud et al. (2023) demonstrated a cost-effective, real-time pavement condition monitoring system using MEMS-based sensors mounted on vehicle axles. The system integrated an ESP32 microcontroller, GPS module, and MPU6050 accelerometer to capture vertical acceleration and geospatial data at high sampling rates. On-device preprocessing using Fast Fourier Transform (FFT) extracted spectral features, which were then fed into ensemble models (Random Forest, LightGBM, XGBoost) to predict IRI at 0.1-mile intervals. The XGBoost model achieved RMSE of 16.89 in/mi and MAPE of 20.3%, with field validation on Missouri road segments showing ride-quality classification accuracies up to 97.1% and repeatability (CV = 8.3%). While this approach demonstrates the feasibility of continuous, automated data collection, it remains focused on ride quality metrics and does not synthesize broader dimensions such as geometric design, facilities, and functional classification into holistic assessment scores.

Despite these computational advances, a fundamental gap remains: no existing framework synthesizes multiple infrastructure dimensions into a unified, empirically weighted quality metric. Current tools either focus on specific indicators (roughness, safety, economics) or serve as integration platforms without providing standardized, comparable quality scores across diverse road types and contexts.

Regional Context

In Africa, assessment challenges are compounded by underfunding, heavy reliance on unpaved networks, and fragmented data collection. The AfDB (2018) notes that more than half of Africa's classified roads are in poor condition, yet evaluation methods remain fragmented and inconsistent across countries. In Kenya, for example, 85% of the network is unpaved, but assessment tools still privilege paved-road indices that fail to capture the functional and social importance of gravel and earth roads (KRB, 2023). Recent studies highlight not just data scarcity but the lack of a comprehensive evaluation framework that accommodates the heterogeneity of African networks (Mbugua et al., 2020).

The Road Infrastructure Condition Survey (RICS), widely adopted by the Kenya Roads Board (KRB) and counties to support Road Asset Management Systems (RAMS), combines visual protocols with digital tools such as GPS/GNSS mapping and video logging. While it provides systematic national coverage, its scoring framework remains highly subjective and dependent on surveyor judgment, as illustrated in the official RICS form on **Figure 1**. This underscores the absence of a consistent, holistic framework for road condition assessment in heterogeneous networks. A sample of the KRB RICS forms is shown in **Figure 1**.

It can be observed that the nature of the assessment is heavily subjective, relying mostly on the surveyor's temperament and discernment.

RESEARCH METHODS

Data Collection and Processing

The study utilized the Road Inventory and Condition Survey (RICS 2023) dataset maintained by the Kenya Roads Board (KRB), comprising 260,773 records and 44 variables. The RICS data was collected through a comprehensive national survey conducted by KRB in 2023, building upon previous surveys to ensure methodological consistency and temporal comparability. For model validation purposes, additional field data was collected from a stratified sample of 150 road segments across five counties: Nairobi (35 segments), Kiambu (30 segments), Nakuru (25 segments), Mombasa (30 segments), and Kisumu (30 segments). These validation roads represented diverse conditions including urban arterials (A104 Thika Road, B3 Mombasa Road), rural trunk routes (A1 Nakuru-Eldoret Highway, C27 Kiambu- Limuru Road), secondary roads (D412 Karen-Ngong Road, E463 Kisumu-Kericho Road), and tertiary access roads. These counties were selected to represent a range of contexts: major urban centers (Nairobi, Mombasa), peri-urban networks (Kiambu), agricultural corridors (Nakuru), and lake-region transport hubs (Kisumu). This diversity provided a balanced validation sample covering trunk highways, secondary roads, and rural access routes.

Data processing involved extracting attribute information from GIS shapefiles as shown in

A1. KRB ARICS Form – Paved Roads

ROAD CONDITION SURVEY - PAVED *ARICS P (F)*

REGION: SOUTH-RIFT COUNTY: NAKURU

C707 ROAD SECTION NAME: KARAGITA SECTION LENGTH (km): 1.0

SECTION START; CHAINAGE: 0+000 LOCATION: Karagita

SECTION END; CHAINAGE: 1+000 LOCATION: Karagita

SHEET: 1 OF 20 CARRIAGEWAY WIDTH: 7m F✓ R H

Chainage per 200 meters	SHOULDER G	SHOULDER R	ON/OFF-CARRIAGEWAY (Rate of Deterioration) 1	2	3	4	5	REMARKS (SPOT IMPROVEMENT)	STRUCTURES – CULVERTS N	RR	HR	NH	G	B	REMARKS/ OTHER STRUCTURES
0+200		✓		✓									3	1	
0+400		✓			✓								1		Concrete-lined outfall
0+600		✓			✓								1		
0+800		✓		✓											
1+000		✓			✓								1		
1+200															
1+400															
1+600															
1+800															
2+000															
2+200															
2+400															
2+600															
2+800															
3+000															
3+200															
3+400															
3+600															
3+800															
4+000															
4+200															
4+400															
4+600															
4+800															
5+000															
S km	0	1	-	0.4	0.6	-	-	S No	0	0	0	0	6	1	

SECTION KM: 1.0

	1	2	3	4	5
S %	-	0.4	0.6	-	-

Average Rate of Deterioration: 2.60

PRIORITY FOR SPOT IMPROVEMENT IN THIS SECTION: Ch: Ch: Ch:

PRIORITY FOR STRUCTURES IN THIS SECTION: Ch: Ch: Ch:

COMPILED BY: SIGN: DATE:

FIGURE 1
Sample KRB ARICS Form – Paved Roads
Source: Kenya Roads Board (KRB, 2023)

Figure 2 and converting to CSV format for machine learning applications. Variables as presented on **Table 1** encompassed route identifiers, network classification, geometric measurements, surface characteristics, infrastructure facilities, and condition assessments.

Dataset Characteristics

The RICS 2023 dataset provides comprehensive coverage of Kenya's road network across 47 counties. Analysis revealed network heterogeneity: 85.1% unpaved roads (46.1% earth, 39.1% gravel), with only 14.9% paved infrastructure. Surface condition assessment showed 73.6% of roads rated as Fair or Poor condition, indicating substantial infrastructure challenges.

Key infrastructure gaps included limited shoulder coverage (6.0% of roads), minimal footpath provision (2.2%), and extremely rare cycling lanes (0.1%) as presented in **Figure 3**. Drainage systems existed on 47.0% of roads, though 40.7% of drainage systems were rated as Poor condition. Usage patterns showed 70.9% of roads experience regular traffic, 19.9% rarely used, and 9.2% busy traffic conditions.

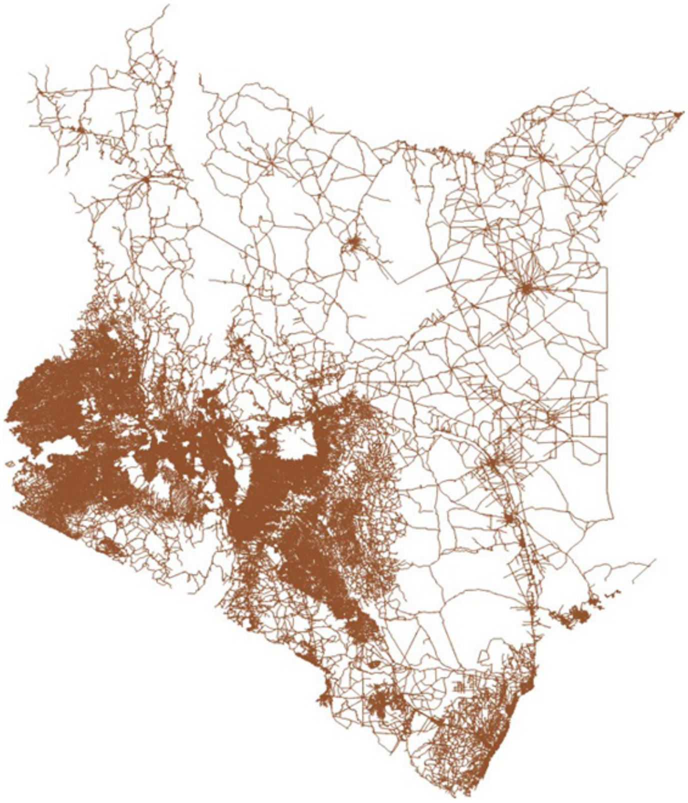

FIGURE 2
Kenya's National Road Network (RICS 2023)
Source: Kenya Roads Board (KRB, 2023)

TABLE 1
Variable distribution by category

Category	Variables	Numeric	Categorical
Administrative	7	0	7
Network Classification	6	0	6
Physical Dimensions	6	6	0
Surface Characteristics	6	0	5
Shoulder Infrastructure	4	1	3
Drainage Systems	5	0	5
Traffic Control & Safety	6	1	5
Additional Infrastructure	4	0	4
Total	**44**	**9**	**35**

Source: Author Analysis, 2025

Universal Scoring Framework Development

A comprehensive scoring framework was developed to synthesize diverse infrastructure, usage, and condition attributes into a single continuous metric ranging from 0 (poor quality) to 1 (excellent quality). The universal score is computed as a weighted average of six normalized sub-scores:

$$universalscore = \frac{1}{6}\sum_{i=1}^{6} w_i s_i \quad (1)$$

where w_i represents component weights (initially set to 1 for equal contribution) and S_i denotes individual component scores normalized to the [0,1] interval.

AFRICA HABITAT REVIEW 20(3) 2025

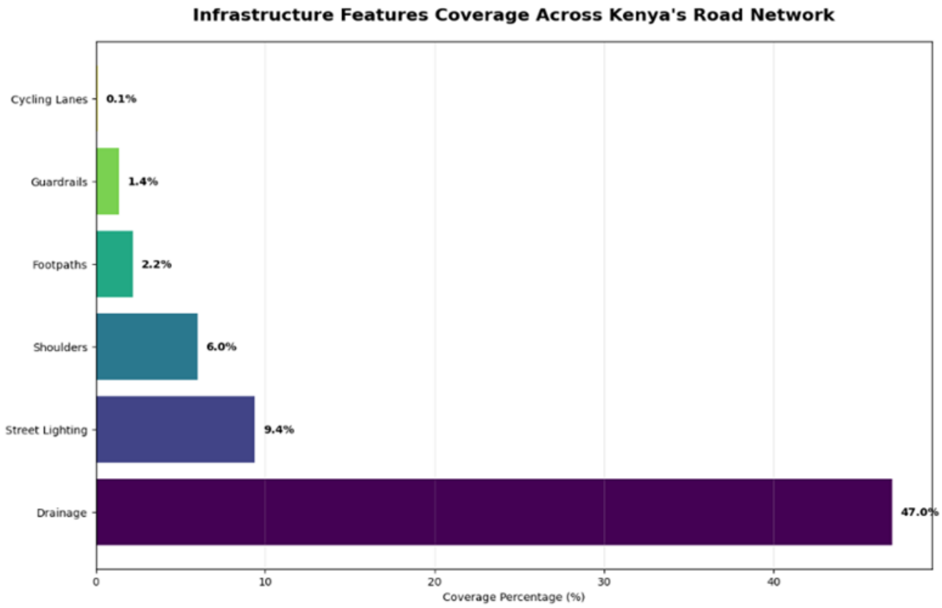

FIGURE 3
Infrastructure Coverage
Source: Author Analysis, 2025

Component Score Methodologies
Condition Score (S1): Surface Quality Assessment
The condition score implements a hierarchical evaluation system prioritizing quantitative measurements over categorical assessments. For roads with IRI measurements (18.5% of network):

$$S_i^{IRI} = max\left(0, min\left(1, \frac{(8-IRI)}{8.0}\right)\right) \qquad (2)$$

This inverse linear relationship maps IRI values where:

- IRI = 0 m/km→S1 =1.0(perfect condition.
- IRI = 2 m/km→S1 =0.75(good condition threshold).
- IRI = 5.5 m/km →S1 =0.35 (poor condition threshold).
- IRI 8m/km→S1 =0.0(failed condition)

For roads without IRI data (81.5% of network), categorical conditions are prioritized with mapping: Excellent (1.0), Good (0.8), Fair (0.6), Poor (0.4), Very Poor (0.2), Under Construction (0.5).

Surface Type Score (S_2): Material Performance Rating
Surface material scoring reflects expected durability, maintenance requirements, and performance characteristics: Asphalt (1.0), Concrete (0.95), Surface Dressing (0.85), Gravel (0.6), Earth (0.3).

Design Score (S_3): Geometric Standards Assessment
Design score evaluates road geometry against functional requirements using normalized dimensional parameters:

$$S_{3=0.3S_3^{width}+0.3S_3^{carriageway}+0.3S_3^{shoulders}+0.3S_3^{lanes}} \qquad (3)$$

Where:
Road Width Normalization (S_3^{width}):

$$0.3S_3^{width} = min\left(1.0, \frac{RdWidth}{12.0}\right)$$

Based on 12m as optimal width for two-way traffic with shoulders.

Carriageway Width Normalization ($S_3^{carriageway}$)

$$S_3^{carriageway} = min\left(1.0, \frac{CWWidth}{7.0}\right)$$

Using 7m as the standard for comfortable two-way traffic flow.

Lane Configuration (S_3^{lanes})

- 1 lane: 0.5 (single track, limited capacity).
- 2 lanes: 1.0 (standard two-way operation).
- 3+lanes: 1.0 (enhanced capacity maintained at maximum)

Shoulder Width Assessment ($S_3^{shoulder}$)

$$S_3^{shoulder} = min\left(0, \frac{Shouldwidth}{2.0}\right)$$

Recognizing 2m shoulders as optimal for safety and maintenance access.

Infrastructure Score (S_4): Facility Enhancement Rating

Infrastructure score quantifies additional facilities through binary presence indicators:

$$S_{4} = 0.25(I_{draingage} + I_{footpath} + I_{cycling} + I_{shoulder}) \quad (4)$$

where I_x represents binary indicators for facility presence.

Usage Score (S_5): Operational Importance Weighting

Usage scoring reflects traffic intensity and operational criticality: Busy (1.0), Used (0.7), Rare (0.3), Unconfirmed (0.5).

Class Score (S_6): Functional Hierarchy Rating

Road classification scoring follows Kenya's hierarchical system: Class A (1.0), Class B (0.9), Class C (0.8), class D (0.7) and class E (0.6), with urban variants ("u" suffix) under KURA jurisdiction receiving identical scores.

Knowledge Distillation Process

To refine the initial heuristic score, a two-stage machine learning approach was employed using knowledge distillation principles as presented on **Figure 4**.

Stage 1: Random Forest Teacher Model

A Random Forest Regressor was trained to predict the Universal Score using the six component scores as input variables. The dataset was split into training (80%) and testing (20%) sets, including only records with complete values (n = 208,618). Permutation importance and Shapley Additive Explanation (SHAP) analysis were used to diagnose the empirical contribution of each component.

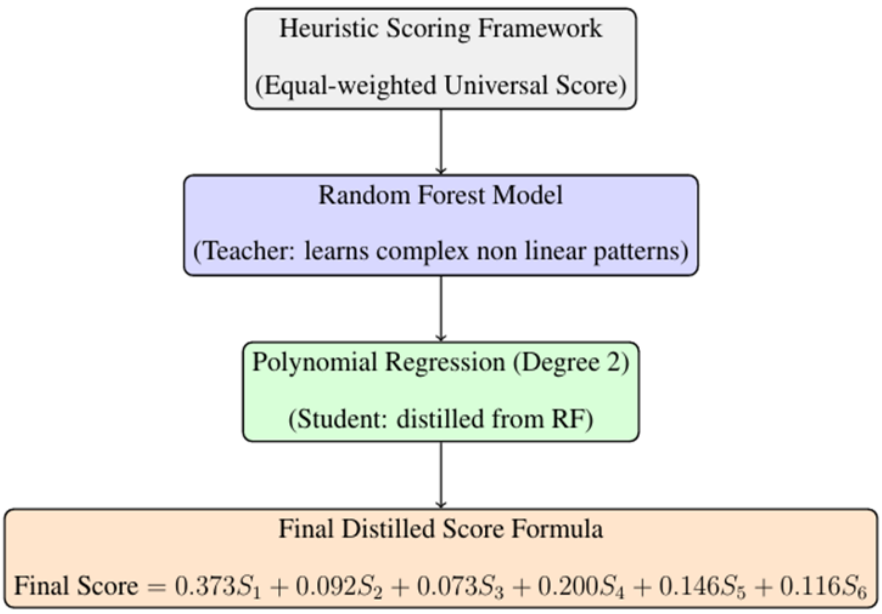

FIGURE 4
Knowledge Distillation Pipeline
Source: Author Analysis, 2025

AFRICA HABITAT REVIEW 20(3) 2025

Stage 2: Polynomial Regression Student Model

To ensure interpretability, the Random Forest served as a teacher model for a polynomial regression student model. The student model was fit to the teacher's predictions to produce a transparent, weighted formula:

$$y = \beta_0 + \sum_{j=1}^{1} \quad \beta_j S_j \quad (5)$$

The resulting coefficients, normalized to sum to one, provide a data-driven weighting scheme for the final score.

RESULTS

Model Performance

The Random Forest teacher model as shown in **Figure 5** achieved exceptional performance in learning the heuristic scoring framework:

- $R^2 = 0.92986$
- Mean Absolute Error (MAE) = 0.0051

Cross-validation stability across 5 folds

These metrics indicate successful pattern learning from the heuristic framework while identifying non-linear relationships between components.

Feature Importance Analysis

Permutation importance analysis revealed the empirical contribution hierarchy, confirmed through SHAP analysis on 1,000 test records. The analysis confirmed the dominance of Condition Score, which comes second to usage score in the Random Forest model's feature importance, while revealing interaction effects between Design and Class components **Figure 6**. The mean absolute SHAP values provided further validation of component contributions.

Distilled Assessment Formula

The polynomial regression student model successfully captured the Random Forest's knowledge in interpretable form. The final, empirically weighted Universal Road Score formula is:

$$\text{Road Score} = 0.373 \cdot S_1 + 0.200 \cdot S_4 + 0.146 \cdot S_5 + 0.116 \cdot S_6 + 0.092 \cdot S_2 + 0.073 \cdot S_3 \quad (6)$$

Results demonstrate surface condition dominance (37.3%) followed by infrastructure facilities (20.0%), validating expert intuition while revealing empirical weight relationships. The significant infrastructure component weight highlights the importance of drainage, shoulders, and pedestrian facilities in comprehensive quality assessment.

DISCUSSION

The empirical weights reveal surface condition

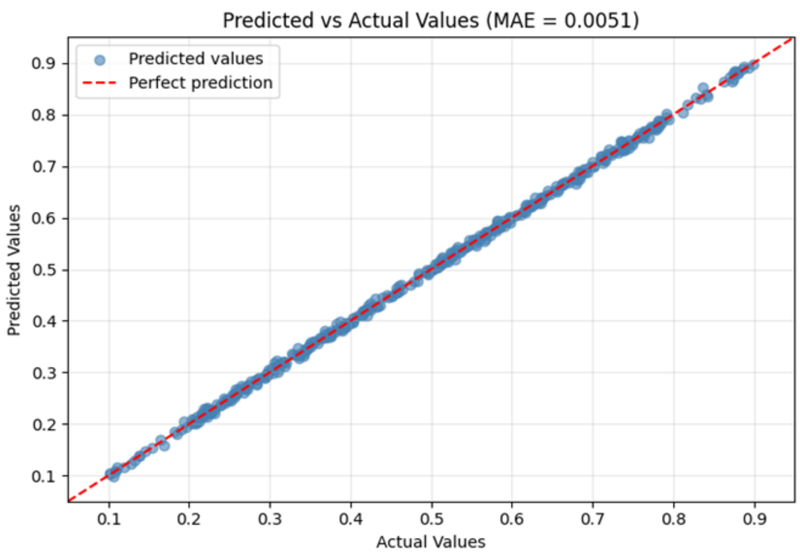

FIGURE 5
Model Performance
Source: Author Analysis, 2025

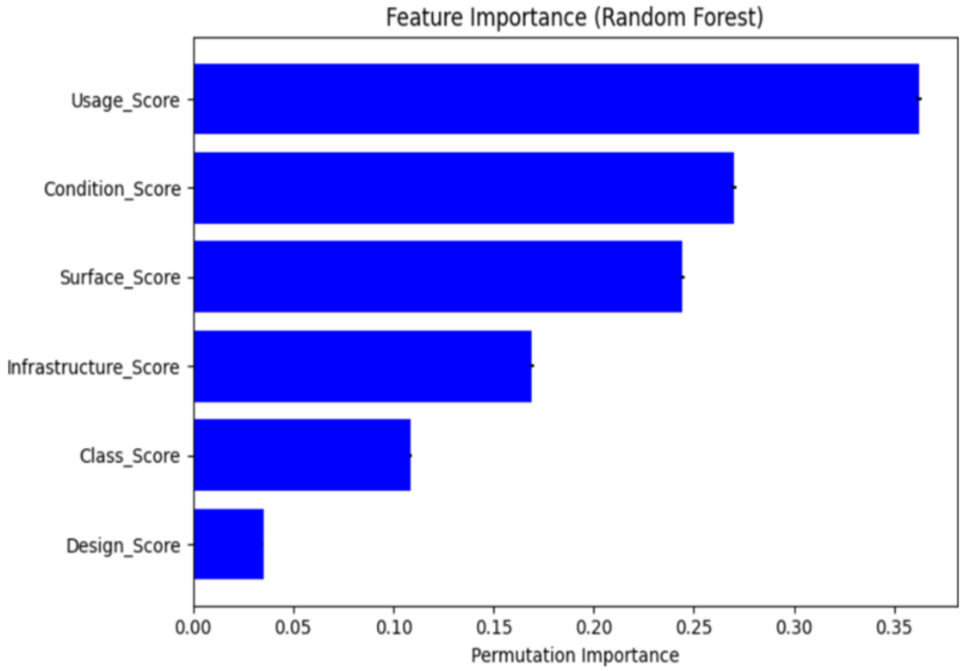

FIGURE 6
Feature Importance Analysis
Source: Author Analysis, 2025

as the dominant factor (37.3%) in road quality assessment, validating expert intuition while providing quantitative confirmation of its primacy. However, the substantial weight assigned to infrastructure facilities (20.0%) highlights a dimension often overlooked in traditional frameworks focused solely on pavement condition. This finding underscores the importance of drainage systems, shoulders, and pedestrian facilities in comprehensive quality evaluation. Such elements critical to road functionality but absent from PCI or IRI metrics.

Usage patterns (14.6%) and functional classification (11.6%) together account for over a quarter of the quality score, reflecting the reality that road performance cannot be divorced from its intended purpose and traffic demands. A rural access road serving light traffic requires different standards than a national highway, and the framework captures this contextual dependency. The relatively lower weights for material type (9.2%) and geometric design (7.3%) suggest these factors, while important, are secondary to surface condition and functional adequacy in the Kenyan context.

Practical Implications

The framework addresses critical developing country challenges through several key advantages:

i. Heterogeneity Accommodation: Successfully integrates quantitative IRI measurements with categorical assessments, enabling evaluation across diverse road types from paved highways to earth tracks.
ii. Resource Optimization: Provides systematic methodology for maintenance prioritization using available data sources, reducing reliance on expensive specialized equipment.
iii. Policy Support: Transparent, interpretable formula enables evidence-based resource allocation and performance bench-marking across different road classifications and regions.

The 92.99% model accuracy demonstrates that complex multi-dimensional quality assessments can be systematically automated while maintaining full interpretability for policy applications. The methodology's applicability extends beyond Kenya to other developing countries facing similar infrastructure assessment challenges.

While this study presents a methodological advancement in holistic road assessment, several limitations warrant acknowledgment. The heuristic scoring framework, though informed by engineering standards and domain expertise, relies on expert-defined component weights for the initial teacher model training. This introduces an element of subjectivity in the normalization and aggregation of individual dimension scores. Although the Random Forest model learns patterns from this framework and SHAP analysis validates the resulting weights empirically, the foundation remains partially heuristic rather than purely data-driven from observed road performance outcomes.

The component weights derived from Kenya's road network reflect local infrastructure characteristics, climate conditions, and usage patterns. While the six-dimension framework is generalizable, direct application of the formula's specific weights (Equation 6) to other contexts without recalibration may not capture regional priorities accurately. Countries with different unpaved-to-paved ratios, climatic deterioration mechanisms, or traffic compositions may require retraining to establish context-appropriate weights.

The framework is constrained by available data dimensions. Other potentially relevant factors such as structural capacity, drainage system functionality beyond presence/absence indicators, or climate resilience metrics were not included due to data limitations. Future research could expand the dimensional scope as more comprehensive datasets become available, particularly incorporating longitudinal performance data to validate predictive capability.

A tabular comparison of the PCI, IRI, RICS survey and the machine learning based framework was as shown on **Table 2**:

CONCLUSION

This study developed the first holistic framework for road performance assessment tailored to the heterogeneous networks of developing countries. By integrating six critical dimensions—surface condition, surface type, geometric design, facilities, usage patterns, and functional class—into a universal scoring system, it addressed the limitations of fragmented, indicator-specific approaches such as PCI and IRI.

Through a knowledge distillation process, a Random Forest teacher model ($R^2 = 0.93$; MAE = 0.0051) was successfully distilled into a transparent polynomial regression formula, balancing predictive accuracy with interpretability. Results confirmed that surface condition is the most influential factor (37.3%), while infrastructure facilities (20.0%) and usage patterns (14.6%) also play substantial roles.

The framework's adaptability to Kenya's predominantly unpaved network demonstrates its potential for broader application across resource-constrained contexts. By enabling systematic, data-driven assessment, the study advances the state of knowledge on infrastructure evaluation and provides an actionable tool for evidence-based planning.

RECOMMENDATIONS

Based on the study findings the following recommendations are drawn:

i. Standardize assessment across agencies: Adopt the universal scoring framework as a national benchmark to harmonize road quality evaluation across counties and agencies, reducing inconsistencies in maintenance prioritization.
ii. Strengthen Road Asset Management Systems (RAMS): Integrate the scoring methodology into digital platforms, allowing real-time visualization of condition data and evidence-based allocation of maintenance funds.
iii. Invest in objective and scalable data collection: Expand the use of digital RICS surveys, video logging, and remote sensing to minimize subjectivity in condition ratings and improve data comparability across regions.
iv. Tailor framework to regional contexts: Adjust component weights to reflect local conditions such as climate, terrain, and usage intensity, ensuring that the framework remains flexible for adoption across Africa and other developing regions.

CITED REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., & Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. In *12th USENIX*

TABLE 2
Comparison between PCI, IRI and RICS

Criteria	PCI	IRI	Kenya RICS	Proposed Method
Assessment Scope	Surface Distress	Ride Quality	Multi-element (deterioration rate, shoulders, structures, surface type, facilities)	Holistic integration of six RICS-derived dimensions into unified framework
Output Metric	Single index (0-100) based on distress deductions	Roughness value (mm/km or in/mi)	Separate scores per element; Average Rate of Deterioration; no unified quality metric	Universal Road Score (0-1) with empirically weighted integration of all dimensions
Relationship to Infrastructure	Independent assessment protocol requiring separate surveys	Independent measurement requiring specialized equipment (profilometer)	Primary national data source—260,773 segments surveyed using standardized forms.	Built on RICS data—no additional field surveys required; provides analytical layer for existing national inventory
Key Innovation and Limitation	Standardized distress taxonomy globally recognized	Objective, reproducible ride quality measurement	Comprehensive multi-dimensional field data collection	First systematic synthesis of multi-dimensional RICS data with empirical weights (surface 37.3%, facilities 20.0%, usage 14.6%, classification 11.6%, material 9.2%, geometry 7.3%)
Limitation	No account of geometric design , facilities or functional context of the road	Measures a single dimension, a road may perform well on IRI but may fail on safety , drainage, etc.	No integration of the data collected. Elements are recorded separately and are interpreted as such	Requires RICS data infrastructure

Source: Author Analysis, 2025

Symposium on Operating Systems Design and Implementation (OSDI 16) (pp. 265-283). Retrieved from https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

African Development Bank (AfDB). (2018). *Road infrastructure in Africa: Challenges and opportunities.* Abidjan, *Côte d'Ivoire: Author*. Retrieved from https://www.afdb.org/en/documents/road-infrastructure-africa-challenges-and-opportunities

American Society for Testing and Materials (ASTM). (2018). *Standard practice for roads and parking lots pavement condition index surveys* (ASTM D6433-18). *West Conshohocken, PA: ASTM International.* Retrieved from https://www.astm.org/d6433-18.html

Archondo-Callao, R. (2019). *Applying the HDM-4 model to strategic planning of road works.* Washington, DC: World Bank. Retrieved from https://documents.worldbank.org

Daud, A., Amo-Boateng, M., Owor, N. J., Aboah, A., & Adu-Gyamfi, Y. (2023). Edge computing-enabled road condition monitoring: System development and evaluation. *arXiv preprint* arXiv:2310.05321. Retrieved from https://arxiv.org/abs/2310.05321

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment,* 202, 18-27. https://doi.org/10.1016/j.rse.2017.06.031

Haas, R., Hudson, W. R., & Zaniewski, J. (1994). *Modern pavement management.* Malabar, FL: Krieger Publishing.

Hahn, B. D., & Valentine, D. T. (2002). *Essential MATLAB for engineers and scientists* (2nd ed.). Burlington, MA: Butterworth-Heinemann.

Hughes, B. P., Anund, A., Falkmer, T., & Wundersitz, L. (2019). *iRAP and the evolution of road safety assessment models. Accident Analysis & Prevention*, 122, 124–132. https://doi.org/10.1016/j.aap.2018.09.002

Kenya Roads Board (KRB). (2023). *Road inventory and condition survey (RICS 2023).* Nairobi, Kenya: Author. Retrieved from https://www.krb.go.ke

Kerali, H. G. R. (2006). *Overview of HDM-4: Highway Development and Management Model.* Paris, France: World Road Association (PIARC). Retrieved from https://www.piarc.org

Krajzewicz, D., Erdmann, J., Behrisch, M., & Bieker, L. (2012). Recent development and applications of SUMO-Simulation of Urban MObility. *International Journal on Advances in Systems and Measurements*, 5(3&4), 128-138. Retrieved from https://www.iaria.org/journals/home.html

Loizos, A., & Plati, C. (2007). Pavement evaluation using advanced technologies. In *Proceedings of the 4th International Conference on Managing Pavement Assets*, Athens, Greece.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2005). *Geographic information systems and science* (2nd ed.). Chichester, UK: John Wiley & Sons.

Mbugua, W., Mwaura, G., & Mugo, S. (2020). Assessment of road condition data availability for Kenya's classified road network. *African Journal of Civil and Construction Engineering,* 7(2), 45–57. Retrieved from https://journals.uonbi.ac.ke/index.php/ajcce/article/view/1547

McGhee, K. H. (2004). *Automated pavement distress collection techniques* (NCHRP Synthesis 334). Washington, DC: Transportation Research Board. Retrieved from https://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_334.pdf

Paterson, W. D. O. (1987). *Road deterioration and maintenance effects: Models for planning and management.* Baltimore, MD: Johns Hopkins University Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830. Retrieved from https://jmlr.org/papers/v12/pedregosa11a.html

Sayers, M. W., & Karamihas, S. M. (1998). *The little book of profiling: Basic information about measuring and interpreting road profiles.* Ann Arbor, MI: University of Michigan Transportation Research Institute. Retrieved from https://deepblue.lib.umich.edu/handle/2027.42/61126

Shahin, M. Y. (2005). *Pavement management for airports, roads, and parking lots* (2nd ed.). New York, NY: Springer.

Systèmes, D. (2023). *Abaqus Unified FEA.* Providence, RI: Dassault Systèmes Simulia Corp. Retrieved from https://www.3ds.com/products-services/simulia/products/abaqus/

Zimmerman, K. A., & Peshkin, D. G. (2003). *Systems for pavement condition data collection* (NCHRP Synthesis 326). Washington, DC: Transportation Research Board. Retrieved from https://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_326.pdf