Intelligent Material Testing and Predictive Quality Monitoring for Sustainable Road and Highway Infrastructure in Saudi Arabia

Abstract

Road infrastructure quality is determined by the interaction of constituent materials, construction processes, environmental exposure and traffic loading. In Saudi Arabia, the combination of demanding climatic conditions, industrial freight movements and expanding transport networks creates a strong rationale for integrating material testing with predictive quality monitoring. This integrative review examines laboratory performance testing, intelligent compaction, thermal profiling, dielectric measurements, pavement condition surveying and machine learning as complementary components of a quality assurance system. The review draws on selected published reviews, Saudi pavement research and technical guidance, rather than original field measurements or an exhaustive systematic search. A proposed framework connects material qualification, georeferenced construction measurements, independent verification and operational condition monitoring through a common quality information architecture. Particular attention is given to uncertainty, sensor calibration, environmental transferability, contractual acceptance, and the distinction between forecasting and demonstrating compliance. Sustainability is treated as a lifecycle outcome involving material consumption, construction rework, maintenance requirements and service continuity. An illustrative monitoring allocation and a phased validation protocol are presented to support implementation without implying measured benefits. The analysis indicates that intelligent monitoring is most defensible when used to identify spatial variability, prioritize verification and accelerate corrective action while retaining independent engineering judgment. The proposed framework provides a research agenda for climate-adapted, traceable and performance-oriented road quality management in support of Saudi Vision 2030.

Keywords: asphalt pavement; intelligent compaction; material testing; predictive quality; highway infrastructure; Saudi Arabia; sustainability.

UNDER PEER REVIEW IN IJAR

1. Introduction

A highway is a manufactured engineering system assembled under conditions that vary across time and space. Unlike a factory product, pavement is placed continuously over a foundation whose moisture, stiffness and geometry cannot be assumed uniform. Material production, transportation, paving, rolling and subsequent traffic exposure introduce additional variability. Conventional quality assurance addresses these uncertainties through specifications, sampling, inspection and laboratory testing. These mechanisms remain essential, but isolated test results may fail to describe the condition of an entire lane or identify the origin of a localized defect. A satisfactory average density, for example, does not establish that longitudinal joints or truck exchange locations were compacted adequately. The central challenge is therefore not simply increasing the number of tests, but connecting measurements to the physical processes and decisions that determine long-term performance.

Saudi Arabia provides an important setting for this challenge. High pavement temperatures can reduce the resistance of asphalt mixtures to permanent deformation, while repeated loading on industrial and logistics corridors can intensify accumulated damage. Regional variations in foundation conditions, drainage and construction materials require careful engineering characterization rather than a single national prediction model. The Saudi mechanistic-empirical pavement design study [1] illustrates the importance of combining traffic, climate, structural thickness and material properties when evaluating performance. Road sector policy also emphasizes network quality, safety and efficiency [2,3]. Translating these objectives into practical

construction decisions requires measurable quality evidence at the level of individual projects and material lots.

Digital instruments offer opportunities to extend the spatial and temporal coverage of traditional testing. Roller-mounted sensors can record compaction-related responses and coverage; thermal imaging can reveal temperature differences during placement; dielectric profiling can support estimates of in-place density; and mobile survey systems can document surface deterioration after opening. Reviews of intelligent paving and compaction [4,5] demonstrate the breadth of these technologies. Nevertheless, these instruments measure different physical phenomena, and their outputs cannot automatically substitute for standardized acceptance tests. A stiffness-related roller signal may reflect the underlying layer as well as the material being compacted, whereas a dielectric reading depends on mixture composition and calibration. A meaningful quality architecture must preserve those distinctions.

Machine learning introduces a further opportunity: construction observations can be used to estimate the probability of nonconformance before an acceptance test becomes available. This changes the timing of quality intervention. Instead of waiting until a completed layer fails a test, a team may investigate an anomalous pattern while additional compaction or another corrective measure is still practicable. Predictive models can also associate construction records with later observations of rutting, cracking or roughness. Published work on asphalt property prediction [6] and intelligent pavement surveying [7] provides a foundation, but the operational value of such models depends on the quality of the underlying labels, the independence of evaluation data and the ability to respond to alerts.

The contribution of this paper is an integrated engineering framework that treats measurement, prediction, verification and corrective action as linked stages. It does not presume that artificial intelligence itself makes a road sustainable. Rather, sustainability emerges when reliable decisions reduce unnecessary material replacement, improve durability and allow maintenance to be undertaken at appropriate times. The analysis is directed toward flexible pavements on highways and industrial roads, including their granular foundations, while recognizing that different principles may be necessary for rigid pavement structures.

2. Aim, objectives and research questions

The aim is to develop a technically defensible review framework for integrating intelligent material testing and predictive quality monitoring into Saudi road and highway construction and asset management. The framework is intended to support both engineering research and the practical responsibilities of consultants, contractors, laboratories and road authorities.

The objectives are to identify material and process variables that govern pavement durability under Saudi operating conditions; evaluate the capabilities and limitations of laboratory, nondestructive and continuous sensing methods; formulate a prediction and verification architecture that preserves independent quality assurance; establish lifecycle indicators for evaluating sustainability and economic value; and propose a field validation program capable of testing the framework across different projects and climatic settings.

Three research questions guide the analysis. First, which combinations of conventional tests and intelligent observations provide complementary information on pavement quality? Second, how should predictive models be calibrated, validated and governed so that their outputs support timely and defensible decisions? Third, how can construction quality information be connected to lifecycle performance and sustainability without attributing unmeasured benefits to digital technology? These questions intentionally separate measurement validity, prediction validity and decision effectiveness. A sensor may be accurate without producing a useful prediction, and a statistically accurate prediction may have little operational value if corrective action is unavailable.

UNDER PEER REVIEW IN IJAR

3. Review methodology

This manuscript adopts a targeted integrative review design. It combines selected research on asphalt material behavior, intelligent compaction, machine learning and automated pavement inspection with technical guidance and Saudi road policy. Sources were located through targeted web searches using combinations of asphalt pavement, intelligent compaction, physical property prediction, pavement distress detection, dielectric profiling, Saudi Arabia, rutting and sustainable materials. Publisher records, DOI-linked article records and official technical publications were prioritized. The main emphasis is on research and guidance published during 2020–2025, with earlier or subsequent contextual material considered only where directly relevant.

The intended inclusion criteria are a clear connection to pavement material characterization, construction process measurement, predictive analysis or lifecycle condition monitoring; identifiable publication provenance; and relevance to the proposed quality management framework. General marketing claims, unsupported performance statistics and sources lacking an identifiable technical contribution are excluded. Review articles are used to map the state of practice, while technical guidance is used to distinguish implementation and acceptance considerations. Saudi-specific studies are used for contextual interpretation rather than as evidence that the proposed integrated system has already been deployed nationally.

The synthesis is organized around four functional stages: preconstruction qualification, construction monitoring, acceptance verification and in-service feedback. Within each stage, technologies are compared by the physical quantity measured, the likely sources of error, the applicable engineering decision and the requirements for independent confirmation. Cross-cutting themes include spatial data integration, model transferability, sustainability assessment and quality governance. No pooled effect estimate is calculated because the selected sources address different outcomes and measurement systems, and their reported performance cannot be treated as directly comparable.

This is not a completed systematic review. No comprehensive Scopus or Web of Science database export, duplicate screening log, full-text appraisal matrix or formal risk-of-bias assessment is claimed. Several sources were consulted at abstract or record level, which limits the specificity of conclusions that can be drawn from them. Accordingly, the paper makes methodological recommendations and presents a conceptual synthesis rather than reporting exhaustive evidence coverage. Before journal submission, an author-led database search should document the search dates, complete Boolean strings, eligibility decisions, full-text appraisal and any changes to the review scope. A second reviewer should independently assess a sample of inclusion decisions and resolve disagreements through discussion.

4. Saudi pavement conditions and material quality

4.1 Environmental and loading mechanisms

Asphalt is a temperature-sensitive composite in which mineral aggregates provide a load-bearing skeleton and bituminous binder contributes cohesion and viscoelastic response. Under elevated temperatures and repeated wheel loads, inadequate shear resistance may allow progressive permanent deformation. The resulting rut can originate in the surface mixture, lower asphalt layers or unbound foundation. A surface measurement alone is insufficient to distinguish these mechanisms. Material qualification should therefore be linked to structural analysis and evidence from the relevant layer. Saudi pavement design research [1] supports the consideration of local traffic and climate inputs rather than transferring assumptions from unrelated networks.

Industrial corridors introduce distinct loading concerns. Slow heavy vehicles, repeated turning and braking, loading areas and intersections may impose demanding shear and repeated loading conditions. Axle configuration and loading distribution are more informative than vehicle counts alone. A monitoring plan

UNDER PEER REVIEW IN IJAR

should identify these locations as distinct engineering strata, rather than combining them with ordinary through lanes. Foundation moisture, drainage outlets and utility crossings also merit attention because their behavior may differ from the adjacent roadway. The purpose of risk stratification is to allocate verification resources without abandoning representative sampling.

4.2 Laboratory qualification and performance testing

Material qualification begins with representative sampling and traceability. Aggregate gradation, particle shape, cleanliness, durability and resistance to abrasion affect mixture structure. Binder characterization provides evidence of temperature-dependent behavior and aging susceptibility. Asphalt mixture testing addresses volumetric properties, moisture sensitivity, stiffness and resistance to deformation or cracking. For unbound layers, gradation, moisture-density relationships, bearing or stiffness indicators and field density contribute to the assessment of foundation suitability. Each test must be interpreted in relation to the relevant specification and the actual layer function.

Performance-oriented testing can provide additional information beyond conventional compositional checks. Wheel tracking evaluates the tendency of a compacted mixture to accumulate deformation under prescribed conditions; moisture conditioning examines susceptibility to deterioration in the presence of water; and cracking-related tests characterize resistance to selected fracture mechanisms. No single procedure captures all field damage modes. Laboratory specimens differ from field pavement in compaction history, boundary conditions, aging and loading. Therefore, the most useful strategy combines complementary tests and recognizes the uncertainty involved in extrapolating laboratory performance to service life.

Material approval should establish a baseline for subsequent predictive monitoring. Each approved mix design needs a stable identifier, documented constituent sources, production tolerances and the calibration relationships required by field instruments. Changes in aggregate source, binder grade, recycled content or additives may invalidate an existing sensor relationship. The quality information system should identify such changes automatically and require engineering review before reusing the previous model. Table 1 summarizes the relationship between conventional testing and complementary intelligent observations. The listed methods are examples, not a substitute for the applicable Saudi Road Code or project acceptance schedule.

Table 1. Complementary testing and monitoring methods.

Layer / property	Reference test	Intelligent observation	Interpretive caution
Subgrade	Moisture-density; field density	Roller response; coverage	Affected by moisture and deeper layers
Granular base	Gradation; density; stiffness	Intelligent compaction	Calibrate by material and equipment
Asphalt binder	Rheology; aging tests	Plant temperature and batch records	Process records do not prove rheology
Asphalt mixture	Volumetrics; wheel tracking; moisture tests	Plant and thermal profiles	Performance requires reference tests
Placed asphalt	Core density and thickness	Dielectric profiling; roller records	Mix-specific calibration required
In-service pavement	Reference distress and profile surveys	Mobile imaging; automated profiling	Surface damage does not identify cause

UNDER PEER REVIEW IJAR

4.3 Sustainable materials and performance balance

Reclaimed asphalt pavement, recycled additives and lower-temperature production methods may reduce demand for virgin resources or process energy. However, a material is not sustainable merely because it incorporates waste. The performance consequences of stiffness, aging, moisture susceptibility and cracking resistance must be evaluated alongside potential environmental benefits. Research from Saudi Arabia on recycled plastic-modified binders [8] provides an example of investigating rheological performance under demanding thermal conditions. Binder-level improvement should not automatically be interpreted as demonstrated mixture durability or field service life.

A responsible materials strategy establishes performance equivalence or a justified improvement for the intended application. The design team should examine interactions between reclaimed binder, fresh binder and aggregate structure, and assess whether variability can be controlled through plant procedures. The monitoring system should distinguish changes caused by recycled content from changes due to temperature, production settings or source materials. This requires a mixture-specific calibration and sufficient test results to support comparisons. Experimental mixtures should be introduced through controlled trial sections rather than widespread implementation based only on laboratory predictions.

5. Intelligent construction and quality measurement

5.1 Intelligent compaction

Intelligent compaction combines instrumented rollers, positioning and data recording to characterize the compaction process. Depending on the equipment, observations may include roller passes, speed, vibration settings, temperature and a measurement value related to the dynamic response of the roller and supporting material. Such records can reveal areas that were missed or received different rolling patterns. The state-of-the-art review of intelligent compaction measurement values [5] highlights the continuing development of methods and the need to understand the physical interpretation of their outputs.

A roller measurement is not identical to density. Its response may be influenced by underlying stiffness, layer thickness, vibration mode, moisture, speed and equipment characteristics. A high value may reflect a stiff foundation rather than adequate compaction of the current layer. Equally, a low value could indicate a weak underlying zone even if the surface material has reached the required density. These differences make the measurement useful for investigation, but they also caution against a universal numerical acceptance threshold.

A trial section should be established before production use. The team should document roller settings and collect co-located reference measurements at different stages of compaction. The relationship between sensor response and the target property should be evaluated separately for distinct materials and layers. Where the relationship is weak, the data may still be valuable for monitoring coverage or identifying relative variability, but should not be presented as a validated density estimate. An equipment change or substantial alteration in layer conditions should trigger a review of the calibration.

5.2 Thermal profiling and placement uniformity

Paver-mounted thermal profiling records the spatial pattern of surface temperature during asphalt placement. It can identify cooler zones associated with interruptions, truck changes or nonuniform delivery. These patterns help investigate whether paving operations are producing conditions likely to impair compaction. Federal Highway Administration guidance discusses thermal profiling together with other nondestructive construction technologies [9,10]. The practical advantage is that a temperature anomaly can be observed before all subsequent construction operations are completed.

UNDER PEER REVIEW IN IJAR

Temperature maps require careful interpretation. The measured surface temperature may be influenced by emissivity assumptions, environmental conditions, sensor positioning and the interval between placement and measurement. A cooler region does not necessarily fail density requirements, and an apparently uniform surface does not prove adequate mixture composition. Therefore, thermal alarms should be connected to a defined inspection or testing response. The responsible engineer should determine whether additional rolling remains feasible within the approved compaction window or whether a different remedy is required.

5.3 Dielectric profiling and independent verification

Dielectric profiling systems use electromagnetic measurements to characterize asphalt pavement and support estimates of in-place density. The physical relationship depends on the mixture and measurement configuration, making calibration against reference samples important. Technical guidance [9,10] discusses the role of dielectric profiling in construction quality programs. Continuous or closely spaced observations can offer an improved description of spatial uniformity compared with a small number of cores.

Independent acceptance sampling should remain separate from model development and contractor process control. Randomly selected verification points protect against biased sampling, while additional risk-based points help investigate anomalies. The two categories should be identified in the database because targeted samples do not provide an unbiased estimate of lot-level conformity. A contract may authorize an alternative acceptance method only after its validity, reproducibility and decision rules have been established. The present framework assumes conventional acceptance remains in force unless expressly changed by the relevant authority.

5.4 Structural and surface condition monitoring

After opening to traffic, condition monitoring provides information on the consequences of construction and exposure. Mobile imaging can document surface cracking and other visible defects, while longitudinal profile measurements characterize roughness. Deflection testing and subsurface sensing may be required to investigate structural weakness. Reviews of intelligent condition surveying and automated pavement inspection [7,11,12] describe advances in data acquisition and automated interpretation. Their relevance to construction quality arises when observations are linked back to the original material lots and layer records.

Computer vision is particularly useful for repeatable screening of large image collections, but its performance depends on camera geometry, illumination, dust, shadows and annotation practices. A detection model trained on one road environment may not perform similarly on another. A crack classification does not establish the underlying cause, and an image-based estimate of severity requires geometric calibration. Human verification and periodic reference surveys remain necessary. The monitoring program should define how ambiguous observations are recorded and when additional structural investigation is warranted.

6. Predictive quality monitoring framework

6.1 Data architecture and traceability

The proposed framework is organized around a georeferenced pavement segment and its construction history. Every test result and sensor observation should retain a project identifier, lane, chainage or coordinates, layer, date, material lot, instrument identifier and quality status. Additional metadata should include measurement units, calibration version, source system and the person or automated process responsible for recording the information. A shared identifier structure allows data from the plant, field laboratory, contractor and consultant to be connected without relying on manual file names.

UNDER PEER REVIEW IN IJAR

The database must preserve original observations separately from derived features and corrected values. A revised calibration should generate a new derived dataset rather than silently overwrite historical results. Missing readings should be distinguished from readings outside the operating range, and both should be distinguished from actual nonconformance. This distinction is crucial for auditability because a missing sensor record is not evidence that the pavement is defective, although it may create a need for supplementary inspection.

Figure 1 presents the proposed information and decision flow. The architecture includes a feedback connection from independently confirmed results and in-service performance to future calibration and material qualification. It is not a description of an existing Saudi road authority platform. Its purpose is to identify the interfaces that a pilot implementation would need to establish.

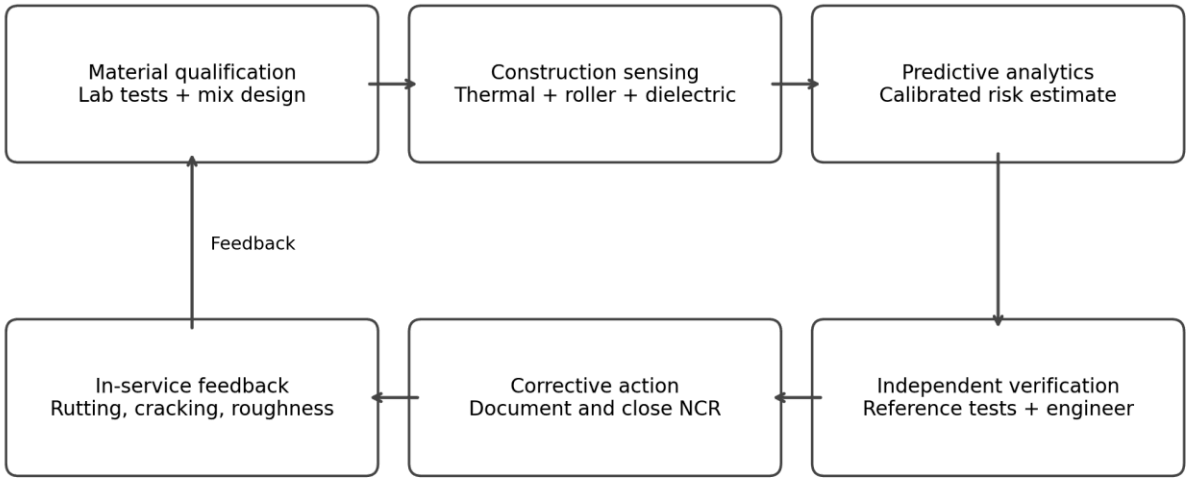


Figure 1. Proposed closed-loop quality monitoring framework (conceptual synthesis).

6.2 Model development and target definition

Predictive monitoring should begin with a precisely defined target. For construction control, a suitable example is the probability that an independently measured density will fall below the applicable acceptance limit. Other targets might include a continuous estimate of density, the probability of an out-of-tolerance thickness, or the likelihood that a laboratory performance indicator will fail a specified criterion. These targets are not interchangeable. A model that predicts density accurately cannot be presumed to predict rutting or service life without additional evidence.

Potential predictors include temperature history, roller coverage, speed, vibration settings, foundation measurements, mixture composition, paving interruptions and relevant environmental variables. Features must be available at the time the prediction is issued. Including a later laboratory result as an input to a construction-time prediction would create information leakage. Similarly, a feature representing the final disposition of a nonconformance report cannot be used to forecast that same nonconformance before it occurs.

A simple statistical baseline should precede complex machine learning. Linear or logistic regression can reveal whether basic physical and process variables have useful predictive relationships. Tree-based ensembles can represent nonlinear interactions, while more complex spatial or temporal models may become appropriate if sufficient independent observations exist. Research reviewing machine learning for asphalt properties [6]

supports investigating a variety of model families, but algorithm choice should depend on the target, sample size, measurement quality and operational constraints rather than reported accuracy from unrelated datasets.

Model training should address class imbalance because confirmed defects may be less frequent than satisfactory observations. Overall accuracy can be misleading when the majority of segments comply. Precision, recall, false-alarm frequency and calibration should be examined alongside conventional discrimination metrics. For continuous predictions, mean absolute error, bias and interval coverage should be reported. Performance should also be evaluated near contractual limits, where prediction errors have the greatest practical consequences.

6.3 Spatial validation and uncertainty

Construction observations are spatially dependent. Nearby records may share the same material load, equipment settings and environmental conditions. Randomly dividing individual sensor points into training and testing sets can therefore produce overly optimistic estimates of predictive performance. A more defensible evaluation holds out complete lots, construction days, road sections or projects. The choice of grouping should reflect the intended use: predicting an adjacent portion of a current project is a different task from transferring a model to a new contractor or region.

Probability calibration matters because a risk score is used to allocate attention. Among comparable observations assigned a particular probability, the observed frequency of the target outcome should be consistent with that probability within sampling uncertainty. Calibration can deteriorate when production practices or defect prevalence change. Periodic reassessment should therefore include both discrimination and calibration. Explanatory feature contributions can help identify suspicious relationships, but such explanations describe model behavior and do not prove physical causation.

6.4 Engineering decisions and corrective action

Prediction is useful only when a responsible party can act on it. The framework distinguishes routine monitoring, targeted investigation and formal acceptance decisions. A low-risk result permits normal process monitoring but does not eliminate mandatory tests. An elevated-risk result triggers review of sensor quality, construction history and the need for supplementary verification. A confirmed failure is handled through the contractual nonconformance procedure. Where continuing construction may conceal a potential defect, an authorized engineer may impose a hold point according to the approved inspection plan.

Alert thresholds should reflect the cost and consequences of missed defects and unnecessary investigations. They should not be copied from another project without validation. The operational team should record the alert, investigation, reference result, corrective action and closure evidence. This creates a dataset for evaluating whether monitoring actually changes outcomes. An alert that is technically accurate but arrives after the opportunity for correction has passed may offer little construction-stage benefit.

Table 2 proposes performance indicators for the integrated system. The indicators are intentionally defined without numerical targets because suitable thresholds require baseline measurements, project specifications and agreement between the relevant parties. Reporting should distinguish sensor coverage from measurement quality and model performance from engineering compliance.

Table 2. Proposed validation and operational indicators.

Indicator	Definition	Purpose
Coverage	Valid mapped area / intended area	Identify monitoring gaps
Density MAE	Mean absolute error against independent cores	Quantify prediction accuracy

UNDER PEER REVIEW IN IJAR

Recall	Confirmed failures alerted / all confirmed failures	Measure missed-defect risk
Precision	Confirmed failures / investigated positive alerts (with defined verification sampling)	Assess investigation burden
Warning lead time	Time from alert to loss of feasible correction opportunity	Assess actionability
Rework intensity	Confirmed replacement quantity per lane-km	Track material consequences
Lifecycle condition	Repeated rut depth, cracking and roughness	Assess durability over time

7. Sustainability and lifecycle evaluation

7.1 Environmental accounting

Environmental evaluation should compare road alternatives that deliver an equivalent engineering function. A suitable functional unit may be a lane-kilometre providing a defined level of service over a stated analysis period. The assessment boundary should include material extraction, mixture production, transportation, construction, maintenance and end-of-life treatment where relevant. Without a common functional unit and analysis period, a comparison may favor a low-impact initial mixture that requires more frequent replacement.

The proposed monitoring framework can improve the inventory used for such an assessment by linking material quantities, production records and confirmed rework to specific construction segments. However, a reduction in predicted defects is not equivalent to a measured reduction in greenhouse gas emissions. Environmental impacts require actual quantities and defensible emission factors. Any estimate of avoided future rehabilitation must be derived from a documented performance scenario and tested for sensitivity to assumptions about service life.

7.2 Economic and social implications

Lifecycle cost analysis should include initial construction, monitoring, testing, corrective action, maintenance and rehabilitation. Future expenditures need a consistent discounting convention. User delay and disruption may be included when reliable traffic and work-zone information is available. A contractor's immediate savings and a road authority's long-term costs are different perspectives; the distribution of costs and benefits should be reported rather than hidden in an aggregate figure.

Early detection of nonuniformity may reduce the scale of corrective work when an effective remedy remains available. Yet additional investigations may also increase costs and delay operations if alerts are poorly calibrated. The economic case depends on defect prevalence, the consequences of missed defects, investigation costs and the effectiveness of available remedies. A pilot study should collect these quantities rather than assume a fixed percentage reduction in rework.

7.3 Illustrative allocation of monitoring effort

Figure 2 presents a hypothetical allocation of 100 monitoring effort units across five activities. The purpose is to demonstrate how a project might discuss resource allocation among conventional verification, continuous sensing, data management and validation. The allocation is not based on survey results, measured expenditure

UNDER PEER REVIEW IN IJAR

or Saudi national statistics. It is a planning example and should not be interpreted as the relative contribution of any technology to pavement quality.

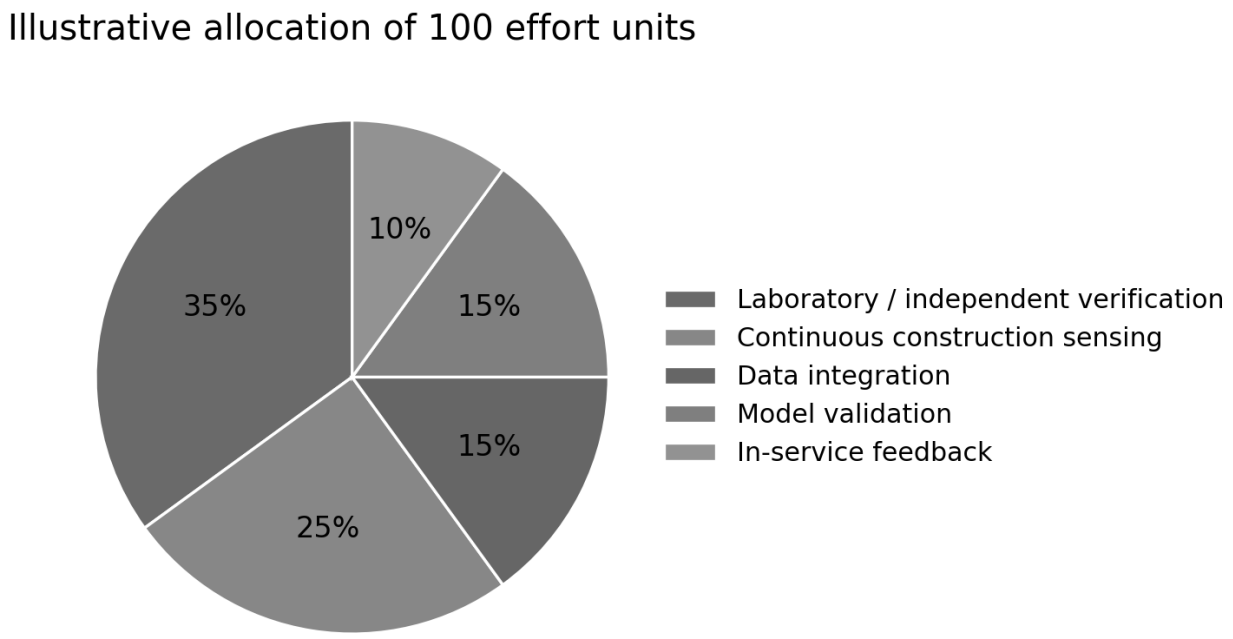


Figure 2. Hypothetical planning allocation; not observed data or recommended universal proportions.

8. Field validation and implementation in Saudi Arabia

A phased implementation is preferable to immediate contractual adoption. During the first phase, the project team should establish a data dictionary, instrument inventory, reference testing plan and baseline description of existing quality practices. A trial section should test data capture, positioning accuracy, calibration stability and the ability to match records from different organizations. The team should document practical constraints, including equipment availability, laboratory turnaround, network connectivity and the timing of construction decisions.

During the second phase, sensors and predictive models should operate in shadow mode. Predictions are recorded but do not independently determine acceptance. Conventional quality assurance proceeds normally, enabling an evaluation of the predictions against independent results. The study protocol should specify the outcomes and analysis methods before examining the final evaluation dataset. Sample size should be justified using the expected variability and frequency of the target outcome, rather than a convenient count of sensor readings.

The third phase should evaluate decision effectiveness under controlled operational conditions. Comparable sections or staged implementation periods may be used to examine whether alerts result in earlier investigations, fewer repeated defects or more efficient allocation of testing resources. Differences in mixture, contractor, traffic and season need to be considered when interpreting outcomes. Where random allocation is not practical, matched comparisons and transparent adjustment for confounding variables can strengthen the analysis, although causal uncertainty remains.

The final phase should connect verified construction records with periodic condition surveys. Observations of roughness, rutting and cracking should be collected with consistent methods and linked to the corresponding

UNDER PEER REV

road segments. Changes in traffic loading, maintenance history and drainage condition should be recorded to avoid attributing all deterioration to construction quality. Longer follow-up is required for conclusions about durability than for evaluating construction-stage detection performance.

9. Discussion and research gaps

The literature supports the view that continuous construction observations and automated condition surveys can expand the evidence available to pavement engineers [4–7,11,12]. The main unresolved challenge is the integration of these measurements into a validated quality decision process. Existing technologies often address particular stages, such as compaction or surface inspection, whereas a lifecycle framework must preserve the relationships between material source, construction history, acceptance and later performance. The proposed architecture addresses this integration conceptually, but its effectiveness has not been demonstrated through original field experiments.

A significant research gap concerns local transferability. Saudi projects encompass different climates, material supply chains, traffic patterns and construction practices. Models trained on a narrow dataset may not remain reliable across these conditions. Collaborative datasets with documented material properties and independent reference measurements could support more robust evaluation. Data sharing should use appropriate permissions and remove confidential commercial information where necessary. A shared metadata standard may be more feasible initially than pooling all raw project data.

A second gap concerns the relationship between construction variability and long-term deterioration. Construction sensors provide dense observations, but later damage may be influenced by structural design, drainage, loading and maintenance. Longitudinal studies should distinguish these mechanisms and examine whether the additional construction data improves performance forecasts beyond established engineering variables. Linking records over time requires stable spatial identifiers and documentation of repairs or changes in lane geometry.

A third gap concerns the evaluation of interventions rather than algorithms alone. Many predictive studies emphasize statistical performance, whereas road authorities need evidence that a model improves decisions. Future work should measure investigation turnaround, actionable warning time, confirmed defect detection, unnecessary interventions and the environmental and economic consequences of corrective action. These outcomes should be compared against conventional practice using a clearly defined study design.

A fourth gap is the development of defensible uncertainty thresholds for operational deployment. Models should communicate where evidence is insufficient and provide a route for engineering review. Investigating how uncertainty changes across material types, sensors and project conditions could help define practical limits of applicability. An additional research priority is the reproducibility of sensor calibration procedures and the portability of data between equipment vendors.

10. Limitations and conclusion

This review is limited by its targeted source selection, reliance in part on publication abstracts and absence of original laboratory or field data. It does not establish the prevalence of defects in Saudi highways, quantify achievable reductions in rework or emissions, or validate a specific predictive algorithm. The illustrative resource allocation is a conceptual planning device. The proposed framework therefore requires independent testing, full-text literature appraisal and review against the specifications of any intended implementation project.

UNDER PEER REVIEW IN IJAR

Intelligent material testing and predictive quality monitoring offer a credible direction for strengthening road construction quality when deployed as complementary engineering tools. Laboratory tests establish material suitability, continuous sensors describe construction variability, independent verification protects acceptance integrity, and condition monitoring provides feedback on operational performance. Connecting these activities through traceable spatial records can support more timely investigations and better-informed lifecycle decisions. The key requirement is not the sophistication of the technology but the validity of the evidence and the accountability of the decisions derived from it.

For Saudi Arabia, a climate-aware and performance-oriented implementation is consistent with the broader objectives of durable infrastructure, efficient logistics and sustainable development under Vision 2030. The next research step is a controlled multi-project validation program that evaluates measurement accuracy, predictive reliability, decision effectiveness and lifecycle outcomes. Only after such validation should specific performance gains or alternative acceptance procedures be proposed. The framework provides a starting point for this work while retaining the central role of professional engineering judgment and independent quality assurance.

References

[1] Application of Mechanistic Empirical Pavement Design Guide Software in Saudi Arabia. (2022). Applied Sciences, 12(16), 8165. https://doi.org/10.3390/app12168165

[2] Saudi Press Agency. (2023, November 14). Road Sector Strategy Bound to Enhance Road Quality in Saudi Arabia, Improve its Global Ranking. https://www.spa.gov.sa/en/N1997515

[3] Saudi National Platform. (n.d.). Transportation and Logistics. https://my.gov.sa/en/content/transport

[4] Intelligent paving and compaction technologies for asphalt pavement. Automation in Construction. https://www.sciencedirect.com/science/article/pii/S0926580523003412

[5] A state-of-the-art review of intelligent compaction measurement values (ICMVs) for subgrade and pavement: Advances and challenges. (2024). Construction and Building Materials. https://doi.org/10.1016/j.conbuildmat.2024.137174

[6] Machine learning models for predicting physical properties in asphalt road construction: A systematic review. (2024). Construction and Building Materials. https://www.sciencedirect.com/science/article/pii/S095006182402539X

[7] Intelligent pavement condition survey: Overview of current researches and practices. (2024). Journal of Road Engineering. https://doi.org/10.1016/j.jreng.2024.04.003

[8] Asphalt Binder Rheological Performance Properties Using Recycled Plastic Wastes and Commercial Polymers. (2025). Construction Materials. https://doi.org/10.3390/constrmater5040075

[9] Federal Highway Administration. (2023). A Practice for including Intelligent Construction Equipment in Quality Assurance Programs. FHWA-HIF-23-007. https://rosap.ntl.bts.gov/view/dot/75217

[10] Federal Highway Administration. (2024). NDT and ICT for Asphalt Pavement Construction. FHWA-HIF-24-031. https://rosap.ntl.bts.gov/view/dot/78767

[11] Advances in smart technologies and materials for automated asphalt pavement inspection: Toward transport infrastructure digitalisation. (2025). Automation in Construction. https://doi.org/10.1016/j.autcon.2025.106523

UNDER PEER REVIEW IN IJAR

[12] Automation in road distress detection, diagnosis and treatment. (2024). Journal of Road Engineering. https://doi.org/10.1016/j.jreng.2024.01.005

UNDER PEER REVIEW IN IJAR