Assessment of non-destructive nuclear and non-nuclear asphalt density testing devices for Australian road construction

Noriko Wood [ab]*, Azizur Rahman [a], Ryan H. L. Ip [ac] and Jackie Graham [d]

[a] School of Computing, Mathematics and Engineering, Charles Sturt University, NSW 2650, Australia; [b] Fulton Hogan, Dandenong, VIC 3175, Australia; [c] Department of Mathematical Sciences, Auckland University of Technology, Auckland 1010, New Zealand; [d] Statistical Edge Pty Ltd, Frankston South, VIC 3199, Australia

* Corresponding author. Email: noriko.wood@gmail.com

Abstract

Non-nuclear density gauges, also known as Pavement Quality Indicators (PQIs), offer several advantages over widely used destructive test methods and nuclear density gauges (NDGs) for asphalt density measurement. Replacing NDGs with PQIs would simplify regulatory requirements and reduce the cost and complexity of in-situ asphalt density measurement. This is due to the absence of nuclear materials and the ability of PQIs to provide measurements rapidly. However, PQIs have not gained widespread use in Australia due to concerns regarding accuracy. This work assesses the suitability of PQIs for both acceptance and quality control testing by analysing a dataset collected over an extended period and across a wide range of construction sites in Australia. Accuracy was evaluated based on the strength of the relationship between gauge measurements and core density results. Although some evidence supporting the use of PQIs for quality control purposes has been found, their use for acceptance testing remains unwarranted. Recommendations for establishing standard testing methods are provided to encourage wider adoption of PQIs.

Keywords: asphalt density; statistical analysis; measurement accuracy; pavement quality indicators, nuclear gauge

1. Introduction

The compaction of asphalt during the construction process increases density and reduces air void content, which improves the pavement’s fatigue resistance, durability, and longevity [1]. The optimal air void content for a particular mix design is developed using a laboratory mix design method, which provides a target range for field air voids. In situ air voids above the target range can reduce fatigue resistance and durability, while void content below the target range can increase the risk of asphalt rutting and instability [2]. For this reason, state road authorities and pavement asset owners typically set minimum criteria for asphalt compaction, which are used to assess the acceptance of pavement lots constructed by contractors [3]. In assessing the quality of pavements, together with stiffness and modulus, density is considered one of the most important measurements [4]. It is therefore desirable to both construction contractors and asset owners that an accurate method of assessing pavement compaction is available during construction so that under-compacted areas can be identified and rectified before the asphalt cools and can no longer be compacted. After construction works are completed, it is also desirable that appropriate methods are available for determining if the entire pavement lot is of sufficient quality for acceptance by the asset owners [5].

Assessing the degree of compaction of asphalt during construction is usually part of the quality control procedures performed by the construction contractor during asphalt construction works. In many Australian road authority jurisdictions, the contractor is also responsible for carrying out the acceptance tests that are used by the asset owners to accept or reject the completed work [6,7,8].

The objectives of quality control testing and acceptance testing for asphalt pavement construction are different. Quality control (or process control) checks are intended to prevent the occurrence of defects and to ensure the pavement, upon completion, meets design intent [9]. In contrast, acceptance testing is intended to enable asset owners to identify defects in the completed pavement lot and calculate incentives or penalties that may be applied based on the quality of work

[10]. In most Australian states, these two related objectives are devolved into a single test performed by the contractor. Such an approach is thought to discourage contractors from implementing process improvements that would eliminate the underlying causes of defects but may instead motivate a focus on meeting only the minimum standards [5].

Nuclear density gauges (NDGs) are widely adopted in the asphalt industry as acceptance testing tools but have several limitations on their application to quality control. These limitations relate to the need for specialised training, licensing, handling, and storage due to the presence of nuclear materials [11]. Core extraction, another widely used method of density measurement, takes several hours to return a result, which precludes its use for quality control purposes [12]. For this reason, the use of simple, user-friendly devices that enable workers to directly monitor asphalt compaction in real-time, such as non-nuclear density gauges, also known as pavement quality indicators (PQIs), would be expected to reduce defects by enabling under- or over-compaction to be detected and rectified during pavement construction. In fact, other non-destructive testing devices such as falling weight deflectometer (FWD), distributed fiber optic sensors (DFOS), and traffic speed deflections devices have been shown to be suitable for the assessment of pavement cracks, thickness and/ or critical strains [13,14,15,16]. Compared to NDGs, the non-nuclear nature of PQIs is a major benefit as it ensures the safety of operators. A more detailed comparison between various testing methods will be provided in Section 2.

Despite the advantages of PQIs, the performances of these tools are sometimes unsatisfactory [17], which may explain the lower adoption in the industry. The aim of this paper is to assess the suitability of PQIs for both quality control and acceptance testing purposes via a comparison with NDGs. To achieve this aim, field record data supplied by an Australian construction contractor, covering a wide range of locations and time periods, were analysed. Since PQI is a relatively new method, analyses of field data, especially collected in Australia, are rarely found in the literature. The results of the analyses will lead to several recommendations for improving the adoption of PQIs in Australia. Therefore, the results of this study will contribute to the literature and provide insights for both the scientific community and the industry.

The rest of the paper is organised as follows. Section 2 provides a brief review of common asphalt density testing methods and recent findings on the performances of NDGs and PQIs. Section 3 describes the data sets used in this work and the analytical techniques employed. Results are provided in Section 4. Discussions and recommendations are provided in Section 5. A brief conclusion is provided in Section 6.

2. Common Testing Methods and Tools

Measurement of asphalt compaction is generally achieved using either destructive testing via the extraction of core samples that are then tested in a laboratory, or non-destructive testing using field-applied density measurement devices [18].

2.1. Destructive density testing

Destructive testing via core extraction, although widely accepted as the most accurate method for assessing asphalt density, has several key disadvantages compared to non-destructive testing. These include damaging the integrity of the pavement, introducing points of weakness, being time-consuming and costly and negatively affecting construction productivity due to delays in receiving laboratory results, which may take several hours. This also eliminates the opportunity to correct inadequately compacted areas during construction [11,12].

Several well-established standard test methods for determining the density of extracted asphalt core samples include (1) mensuration: the sample’s physical dimensions and mass are used to compute density; (2) saturation (underwater): the sample is weighed in water and in air in a ‘saturated surface dry’ state and the Archimedes principle is used to determine density; and (3) vacuum sealing and coating: the sample is either placed in a vacuum-sealed plastic bag or coated with a water impermeable material and weighed in water and air, using a similar density determination principle as the saturation method, but eliminating the influence of surface texture and voids [18].

Each method has known biases [18], and precision values vary: the repeatability range for ASTM D2726-17 (saturation method) is 0.006 t/m^3 [19]; while for ASTM D6752-18 (vacuum sealing method) it is 0.035 t/m^3 [20]. Despite the method chosen, core density measurement is considered more accurate than indirect or non-destructive methods and is widely used for acceptance testing [21].

2.2.Non-destructive density testing

Non-destructive testing mitigates the limitations of destructive testing by avoiding damage to the pavement surface and by providing immediate results during paving [14]. This paper focuses on two non-destructive testing methods for asphalt density measurements: NDGs, which are currently used in the state of Victoria for acceptance testing, and PQIs, which are currently used on a more limited basis by construction contractors for quality control testing. For more details regarding other non-destructive testing methods including FWD, DFOS and ground penetrating radars (GPR), readers may consult [22,23].

2.2.1. Nuclear density gauges (NDGs)

NDGs are widely used as a non-destructive testing technique for asphalt density measurement. The method of using NDGs to test asphalt density in Australia and New Zealand is specified by the Australian Standard test method AS/NZS 2891.14.2:2013. Laboratories conducting asphalt density testing using this method in Australia are generally required by state road authorities to be certified by the Australian National Association of Testing Authorities.

NDGs infer density by emitting and detecting gamma rays, typically in two modes: direct transmission and backscattered [24]. The former requires placing the source and detector on two sides of the asphalt. The gamma radiation is emitted from the source, passes through the material, and is detected. This approach is partially destructive as a hole needs to be formed [24]. Since this work focuses on non-destructive methods, the direct transmission method will not be discussed further. The backscattered method is non-destructive in nature. The density is determined by applying gamma rays to the pavement surface and measuring the scattering of gamma photons that result from collision with electrons in the material being measured [21]. To take a measurement, the gauge is placed on the paved surface parallel to the direction of movement of compaction rollers and the backscattered radiation is measured by the gauge for up to four minutes [25]. It is standard industry practice to adjust raw gauge results using an offset value determined from the mean difference between a set of core density and gauge results for each asphalt mix type [12].

Despite reducing some of the drawbacks of destructive testing, particularly with the ability to provide results quicker, the use of radioactive material in NDGs presents various safety and logistical challenges, including the risk of radiation exposure to operators, and the need for appropriate licensing, transport, and storage arrangements [11]. Some other drawbacks of the backscattered method include restricted measurement depth, measurements being biased towards the top layer of the pavement, sensitivity to surface roughness and lower accuracy compared to the direct transmission method [24]. For these reasons, alternative non-destructive methods that do not contain radioactive elements, including intelligent compaction rollers, GPR and PQIs, have been assessed for their suitability for use on road and airport pavement projects in recent years [11,14,17,26,27,28,29,30,31].

2.2.2. Pavement quality indicators (PQIs)

PQIs provide a non-radioactive and non-destructive alternative to NDGs and core extraction for density measurement. Several models of PQI that utilise electromagnetic radiation for density measurements, including the TransTech Systems Pavement Quality Indicator PQI-380, are currently available commercially. The operating principle of the TransTech PQI-380 device is based on the relationship between the dielectric constant of asphalt and its density. An electric field is applied to the pavement surface via an electrode in the centre of the PQI-380's base plate, which is placed onto the asphalt surface. Measurement of the capacitance of the material to which the field is applied enables the dielectric constant to be calculated, and hence the asphalt density can be inferred [32].

In using the TransTech PQI-380, the measurements can be taken in either a single discrete location in single reading mode, or in five locations in a 'cloverleaf' pattern for average reading mode. Unlike NDGs, the manufacturer does not recommend that measurements be taken in a particular direction relative to the placement of the asphalt [33]. By default, the depth of measurement is set at 5in. (9.5mm) but the user could specify a depth within the range of 0.75in. to 6.0in. (19.05mm – 152.4mm) [33]. Density measurements in single reading mode are available well within a minute, compared to up to four minutes for NDGs [25,33]. Use of a mix-specific offset value based on the mean difference between gauge and core density results for a control lot is also recommended, similar to that for NDGs [33].

2.3.Commonly used methods to compare density measurements

Several analytical techniques are available for assessing the suitability of PQIs for measuring asphalt density. The primary method of assessment in the majority of recent studies is the calculation of correlation coefficient (r) and/ or coefficient of determination (R^2) to determine the strength of a linear relationship between core density and PQI gauge results [4,11,13,34,35]. This method assesses how closely the PQI gauge results track core density over a range of density values. Similarly, the calculation of slope and intercept values from the linear regression of core against PQI results provides an indication of whether an offset is required to align raw gauge results with core densities.

Paired t-test is an alternative method for assessing PQI suitability. This method indicates if the mean of a set of PQI results is equivalent to the mean of a matched set of core density results. Although this method provides a general indication of equivalency, it is not considered suitable as a sole measure of suitability due to the potential to provide misleading results when the PQI accurately tracks core density results but requires an offset, or when the study is performed over a narrow range of densities so that the lack of a linear relationship has not been detected [36].

The analyses could be done based on a single site and mix type [11] or multiple sites where correlational relationships are developed for each individual project [36]. Results from multiple projects and mix types could also be combined into a single dataset [4,27,37]. Below we summarise findings from some recently published studies.

2.3.1. Recent findings

A summary of some recently published R^2 values between various devices and cores with various density measurement methods is shown in Table 1. Of particular note, a wide range of R^2 values (between 0 and greater than 0.99) has been reported.

Recent studies that compare TransTech PQI-380 density results to core densities have mostly found moderately strong correlations, with R^2 values of at least 0.7 [11,33,35,38]. Lower R^2 values appear to be associated with studies containing very few replicates (less than 20), studies where results from many different projects and mix types have been pooled [37,39], or where results were obtained from routine construction contractor quality assurance testing [17]. Studies with high R^2 values (greater than 0.99) were conducted on airport pavements where a single asphalt mix type was constructed under uniform conditions [34].

Several studies that have undertaken correlation calculations have also attempted to determine the influence of test conditions on the strength of the correlation. In particular, the influence of ambient temperature [35], pavement temperature [11], moisture [35] and single or average reading mode method [11] have been studied in relation to PQI-380 gauges. Earlier models, including the TransTech PQI-300 and PQI-301, that use the same operating principle as the PQI-380, have also been investigated for the influence of moisture [36,37,40], direction of measurement [40], and asphalt mix characteristics including bitumen type [37], stone type and nominal maximum aggregate size (NMAS) [36]. The bulk of studies determined that the key confounding factors for PQI-380 density measurements are temperature, particularly pavement temperature, and the presence of moisture [11,35,40]. The use of mix-specific offsets has been found to mitigate the effects of asphalt mix characteristics [36].

2.4. Acceptance of PQIs

Criteria based on correlation have been developed for the acceptance or otherwise of PQIs as a replacement for NDGs. These include definitions of acceptable correlations or R^2 values between core and PQIs densities [36], and guidance from state road authority's surveys on R^2 values for core and non-nuclear density gauge densities that are acceptable to asset owners [27].

An extensive field investigation that compared core and PQI-300 density results was used to define an acceptable correlation strength as $r \geq 0.85$, and unacceptable correlation as $r \leq 0.60$ [36]. These values were selected by considering the wide confidence intervals for r that generally resulted from the small sample sizes used for core and gauge density correlation studies. Typically, sample sizes of six are used for lot conformance testing in Australian State Road Authority jurisdictions [6], while in fact, a sample size as small as five could be used, according to Asphalt Institute [2]. These values align with surveys taken of forty US state and provincial Departments of Transport in 2015 in which road authorities indicated a preferred R^2 value between 0.7 and 0.99 for the strength of correlation between core density measurement and any new method proposed to replace it [27]. In the same study, the road authorities also pointed to accuracy as the most important criterion for assessing the suitability of an asphalt density measurement device and reported a maximum acceptable deviation from core density values of between 0.5 and 2.0 lb/ft^3 (equivalent to 0.008-0.032 t/m^3).

Investigation of PQI-380 repeatability has indicated ranges of 0.010 t/m^3 for five replicates [11]. This is within the acceptable repeatability range for electromagnetic density measurements for asphalt of 0.058 t/m^3 published by ASTM [41].

3. Methods

3.1.Data sources

The suitability of PQIs for use in asphalt quality control and acceptance testing was assessed via the analysis of data provided by a construction contractor operating across three states of Australia: Victoria (Vic), South Australia (SA), and Queensland (Qld). The approximate range of locations of the construction sites for which data were provided is shown in Figure 1. The construction contractor provided three data sets:

1) Paired core and NDG results from acceptance testing undertaken in Victoria in the period 2018 to 2022 using Troxler Nuclear Density Gauges as per standard test method AS/NZS 2891.14.2:2013 for which the facility undertaking the testing was accredited by the Australian National Association of Testing Authorities,
2) Paired core and PQI results from quality control testing undertaken in Queensland in 2017 using TransTech PQI-380 gauges as per a local work instruction developed by the contractor, and
3) Paired core and PQI results for quality control testing undertaken in South Australia in 2021 using TransTech PQI-380 gauges as per a local work instruction developed by the contractor.

The data sets included data for various asphalt mix compositions that were measured at various local and highway road projects between 2017 and 2021. Each data set included density results obtained from either PQIs or NDGs and corresponding core samples extracted from the area of the pavement that was tested using the gauge. The densities of the core samples were measured as per the Australian Standard test method, ‘Determination of bulk density of compacted asphalt - Presaturation method’, AS/NZS 2891.9.2:2014.

The data sets included offsets that were applied to the gauge results as well as raw data obtained from the gauge. For NDG results, the supplied offset for each NDG lot had been calculated directly from the lot’s data points. This meant that the lot’s average core density was equal to the average NDG density with offset applied because by definition the offset is the average difference between core and gauge densities. The PQI data sets also had an offset applied, but these were obtained from prior offset calculations for which data were unavailable. Specifically, these offset values were calculated by taking six measurements of asphalt density using a PQI gauge and then extracting 100mm diameter core samples from these locations. The cores are tested for density in a laboratory using a water displacement method based on the Archimedes principle following the standards outlined in AS/NZS 2891.9.2. The results of each density measurement type were averaged, and the offset was the difference between the averages. Due to this difference in offset methodology, most analysis to compare NDGs and PQIs was undertaken without offsets applied. The contractor’s offsets for PQI measurements were applied to the raw data for the purpose of assessing if the offset improved equivalence between core density and gauge measurements. The features of the supplied data sets are summarised in Table 2.

3.2.Statistical analysis

To assess the suitability of PQIs as a replacement for NDGs or core extraction, R^2 and r values were calculated for each pavement lot and for pooled data of each gauge type in the supplied data set. Given that the data were collected over a large geographical area with different mix types and plants involved, the calculation of correlation coefficients was done for individual lots. The intention was to help mitigate the influence of factors that are known to influence PQI results, including mix characteristics, presence of moisture, and temperature [11,35,36,40], as the data within each lot were supposed to be relatively more homogeneous. A minimum of six data points was used for the calculation, and the actual amount of data points varied across the lots. As a comparison, the same was also performed on the pooled data to determine an overall strength of correlation between core and PQI densities.

Confidence intervals for linear regression coefficients were calculated to determine if an intercept of zero and a slope of one could be assumed. This will indicate if the gauge provides a directly equivalent result to core density results without the need for an offset. In the absence of this relationship, linear regression coefficients would enable suitable offsets to be calculated. Paired t-tests were also undertaken to determine the number of lots for which gauge results could be considered to be equivalent to core density results. As the data came from different plants, the effects of the plants on differences between core density and PQI gauge densities were assessed via analysis of variance followed by Tukey's honest significance test. The level of significance was set at 5%.

Outliers were identified via the calculation of standardised residuals and removed from the data set if the standardised residual value exceeded |3|. This methodology resulted in 15 data points from the Victorian data set being excluded from the analysis. These outliers came from various gauges, locations and dates. In other words, no obvious common characteristics were found among these outliers. No outliers were detected in the combined Qld and SA PQI dataset.

4. Results

4.1.Relationship between gauge and core densities

4.1.1. NDG and core densities

Scatterplots of core density versus NDG density for data obtained from Victoria are shown in Figure 2. Different colours are used to indicate the three main categories of asphalt mix type included in the analysis: Dense-Graded (DG), Fine Gap-Graded (FGG) and Stone Mastic Asphalt (SMA) in Figure 2(a) and the asphalt plant source in Figure 2(b). The scatterplot indicates the presence of a moderately strong, positive relationship between core density and NDG results. This is also indicated by a correlation coefficient of 0.86 (Table 3).

At the 5% level of significance, the strength of the linear relationship for the three Victorian asphalt plants was found to be statistically similar, as indicated by overlapping 95% confidence intervals for correlation coefficients (Table 3). Information provided by the contractor supplying the data indicated that different stone types with differing densities were used to manufacture asphalt at each of the three plants. The lack of statistically significant difference in correlation coefficients therefore indicates that the strength of the relationship between NDG and core density measurements was not influenced by different stone sources and plant mixing methods. In contrast, a weaker relationship between gauge and core results was observed for SMA compared to DG and FGG asphalt types (Table 3). Since SMA mixes have a higher texture depth than DG and FGG in Australian mix designs [42], this may imply that the strength of the relationship was influenced by mix composition most likely due to very different surface texture and offset values of these mix types. In fact, this finding is consistent with [43] which suggested that surface texture depth is a contributing factor to greater variability for SMA mixes.

Plants A and C were found to have statistically similar linear relationships between NDG and core measurements, as indicated by overlapping 95% confidence intervals for slope and intercept coefficients (Table 3). Similar linear relationships were also observed for all mix types. This suggests that the linear relationship between NDG and core density measurements remains consistent across different asphalt types.

In terms of accuracy, paired t-test results without offsets, as shown in Table 4, indicate that NDGs' mean results were equivalent to the mean core density results, with a 95% confidence interval of (0.084, 0.088). On average, the core densities were 0.086 t/m^3 higher than the NDG's measurement. When the contractor's offsets were applied, the NDG and core densities were not significantly different (p=0.811; Table 4). This result is somewhat expected given that the offset calculation explicitly reduces the difference between gauge and core mean results by lot.

4.1.2. PQI and core densities

Scatterplots for core density versus PQI density are shown in Figures 2(c) and (d). Unlike the NDG data, the PQI results were found to be strongly dependent on the asphalt plant source. For example, despite having the same mix type (DG), the asphalt produced by Qld Plant B tended to be denser than those produced by SA plant (Figures 2(c) and (d)). This justifies the motivation of looking at the statistical patterns on a lot-by-lot basis.

The scatterplots indicate a generally weak, positive correlation between core and PQI density measurements. The correlation coefficient for the pooled Queensland and South Australia data was 0.66 (Table 3). Correlation coefficients for individual states' data were 0.31 and 0.63 for SA and Qld, respectively. When the contractor's offset was applied to the results, the correlation coefficient decreased to 0.57 for the pooled data set, 0.41 for SA, and 0.33 for Qld (Table 3). This result indicates that the strength of the relationship between PQI measurements and core density was weaker than between NDG measurements and core density. It is however unknown how accurately the contractor established the offset, which could be a critical influencing factor.

The application of the contractor's offset to the PQI measurements was found to decrease the strength of the correlation for Qld Plants B and C, and to increase the strength of the correlation for SA and Qld Plant A. Applying the contractor's offset to the Qld Plant C results produced no linear relationship between gauge and core measurements, as indicated by zero being within the 95% confidence interval for the correlation coefficient (Table 3). This implies that the methodology used to calculate the offset varied between locations and did not uniformly achieve the intended aim of bringing gauge measurements in line with core density.

Different linear relationships between PQI and core density measurements were also observed for different mix types. For example, the slope and intercept coefficients for DG asphalt were significantly different to EME asphalt, at the 5% level of significance (Table 3).

Regarding accuracy, results based on paired t-test without offsets applied indicate that the

PQI and core density measurements were not equivalent, with a 95% confidence interval of (-0.014, -0.003). On average, the PQI measurements were 0.0085 t/m^3 different from the core measurements. Such a difference was closer to zero and around 90% smaller than the one for NDGs (Table 4). However, it is found that the 95% confidence interval with the offset applied, (0.025, 0.036), was further from zero than the raw PQI results. Again, this casts doubt on whether the offsets were established correctly.

4.2. Effects of plants on the differences between core and gauge densities

Results from one-way analysis of variance (ANOVA) followed by Tukey's honest significance test indicate significant effects of plants on the differences between the core and gauge densities for both NDGs and PQIs (Tables 5 and 6, Figure 3).

For NDG results, although the differences between core and gauge densities were close (Figure 3(a)), a statistically significant difference between Vic Plants A and B, and Plants B and C was observed at the 5% level of significance. No statistically significant differences were observed between Vic Plants A and C (Table 5).

For PQI results, statistically significant differences between the plants are observed except between Qld Plants A and C (Table 6). In particular, Qld Plant B had the most positive median difference while SA plant had the most negative median difference and Qld Plant A had a median difference closet to zero (Figure 3(b)).

It is possible that observable differences between plants are the result of multiple associated variables that influence density measurement. These include stone density, mix composition and frequency of manufacture of particular mix types, ambient conditions, construction methodologies, and measurement techniques. Further investigation would be required to determine the influence and interaction of these factors with regard to density measurement.

4.3. Acceptability of NDGs and PQIs

Based on the analysis results undertaken for each pavement lot, defined as material of the same composition, tested on the same day with the same gauge [6], Table 7 and Figure 4 show a greater proportion of lots tested with NDGs than PQIs that had a correlation coefficient greater than the minimum limit of acceptability (0.8), as defined in [27]. It is notable that a majority (64%) of lots tested with NDGs, which are currently widely used in asphalt acceptance testing in Australia, did not meet the minimum criteria for acceptance of new density testing methods recommended by [27].

5. Discussion

5.1. Suitability of PQIs for use in acceptance testing

Acceptance testing is commonly used to determine pay factors where contractors are paid on the basis of the degree of compaction achieved [10]. This requires that the methods used to measure the degree of compaction are appropriately accurate and demonstrate a strong correlation to core density results, which are used as a proxy for the true density of the pavement. The acceptable accuracy for non-nuclear gauges suggested by US state road authorities is a minimum R^2 value of 0.7, equivalent to a correlation coefficient of around 0.8 [27]. In the absence of a similar figure provided by Australian state road authorities, this has been adopted as a preliminary criterion for determining the acceptability of PQI gauges as alternatives to NDGs.

Correlation coefficients for pooled core and PQI measurements with no offset applied were found to be 0.66, as reported in Table 3, which did not meet the minimum limit of 0.8. This result was in line with a similar study of Australian construction lot data where the reported range of correlation coefficients for PQI-core results grouped by mix type was between 0.2 and 0.8 [17].

Results based on various groupings of PQI results by asphalt plant source and mix type also did not achieve the minimum limit of 0.8, except for the SMA mix type which was not considered to be significant due to the small lot size and wide 95% confidence interval (n=7, (0.503, 0.987)). In contrast, the correlation coefficients for the pooled core and NDG measurements and various groupings by asphalt plant and mix type all exceeded the minimum limit.

Analysis results for individual pavement lots, as summarised in Table 7, also indicate that the majority of correlation coefficients for lots tested using PQI gauges did not meet the minimum criteria. This contrasts with NDG results, where a much greater proportion of *r* values by lot can be seen to exceed 0.8 (Figure 4).

It is expected that improvement can be achieved with offsets correctly established and regularly validated. Solely based on the analysis results of this work, PQI measurements do not seem to sufficiently correlate to core density results to justify their direct substitution for NDGs for acceptance testing. Further investigation is required to measure improvements in correlation that may be brought about by the introduction of standardised procedures for undertaking testing and applying offsets.

5.2. Suitability of PQIs for use in quality control testing

In contrast to acceptance testing, where the actual measured density value is important for the purpose of determining pay factors, quality control testing is more concerned with detecting inconsistencies in density across the paved area. A key purpose of quality control testing is to detect areas of under- or over-compaction so that they can be corrected during construction. For this reason, the accuracy requirements of quality control testing applications are less stringent than those for acceptance testing.

The suitability of PQI gauges for use in quality control testing was assessed based on paired t-test results. A statistically significant difference between mean PQI and mean core values would provide an indication that the PQI gauge would not adequately be able to detect areas of under- or over-compaction.

The analysis presented in Section 4.1 indicated that there was a statistically significant difference between PQI and core measurements at the 5% level of significance. This implies that the gauge measurements, generally speaking, do not reflect the core measurements. Furthermore, applying the contractor’s offset to the results appeared to further reduce the equivalence between gauge and core measurements. This is counterintuitive as one would have expected the opposite. It is possible that the methodology used by the contractor to develop the offset during the period in which data were collected was flawed.

Since the application of lot-by-lot offsets helps improve the equivalence between NDG measurements and core density, it is expected that the correct application of offsets would result in a statistical equivalence between PQI and core measurements as well. Since offsets can be influenced by stone type, surface texture, mix blend composition, and other material characteristics, the use of site- or lot-specific offsets is expected to be beneficial. Combined with the result that the 95% confidence interval for the paired differences between PQI and core densities results without offsets applied was closer to zero (i.e., closer to the true mean core density) than that for NDGs, PQIs may have a potential to perform better than NDGs, if guidance on offset calculations is available.

5.3.Recommendations

Although the results of this study show that PQIs cannot substitute for NDGs immediately, especially for acceptance testing purposes, the benefits of PQIs and their potential for quality control testing should not be overlooked. In fact, the better performance of NDGs can at least be partially attributed to the existence of standard test methods (AS/NZS 2891.14.2:2013 in Australia, for example). Without rigorous training and standard protocols, NDGs may not be able to remain an industry standard today.

Therefore, authorities are recommended to establish standard testing methods for PQIs, especially regarding the calculation of offset values aiming at improving the equivalence between gauge and core results. While developing these methods, the effects of various factors on the offset values should be investigated. For example, a recent study shows that PQI measurements are acceptably accurate only when the asphalt is dry and has cooled down [44]. Thus, more research is required to determine the effectiveness of using PQI during the compaction phase. The results of this study may serve as a foundation for future research. Once standard methods are established, construction contractors should provide rigorous training to workers operating the devices to minimise other potential sources of bias.

5.4.Limitations

Some known factors affecting the performance of non-destructive testing tools include the temperature and moisture of the asphalt [45,46,47], gauge orientation [48] and the depth of measurement [49]. In particular, the gauge reading would represent a combined density value of the asphalt mix and the underlying base material if the depth of measurement is set to be larger than the thickness of the pavement [46]. In fact, even different nuclear gauges may measure the pavement at a different depth, causing variability in the density readings [50]. Thus, the feasibility of controlling the depth of measurement precisely [33] is considered an advantage of PQIs [50]. Yet, it is important for the operators to set the asphalt depth to be at least 5 mm less than the nominal thickness of the layer to improve the consistency of the results [47]. Nonetheless, one of the limitations in the current study lies in the lack of information regarding the above factors. Since the data were collected by the contractor over a range of locations and time, it is possible that some of the above factors may have introduced biases. Therefore, the results presented in this work shall be used with caution.

6. Conclusions

In summary, this work assesses the suitability of PQI as a tool for both acceptance testing and quality control purposes. By comparing paired NDG-core and PQI-core asphalt density measurements based on datasets of Australian construction records, this work indicates weak correlations between PQI and core density measurements, in contrast to the strong correlations between NDG and core density measurements. As a result, there is a lack of support for the direct substitution of NDGs with PQI gauges for the purpose of acceptance testing. Yet, this study shows that raw PQI density measurements can provide a reasonably accurate estimate of the core densities, indicating a potential for PQIs to be used in quality control testing during asphalt laying for the purpose of improving compaction consistency in an efficient manner. Considering the advantages of PQIs, especially the safety of the operators, the establishment of standard methods for calculating the offset values of PQIs is recommended. The creation of a test method and standard operating procedures for PQIs, similar to that currently in use for NDGs, would be expected to further improve the equivalence between PQI and core density measurements. Further investigation in the form of a field study with paired NDG, PQI, and core density measurements is required to confirm this conclusion.

Acknowledgements

The authors wish to acknowledge the financial support provided by the Data Science Research Unit (DSRU) Summer Scholarship at Charles Sturt University, Australia. We are thankful for all constructive comments provided by the reviewers and editors, which have improved this manuscript.

Disclosure statement

The first author is employed by Fulton Hogan, an infrastructure company based in Australia. The company had no further involvement in this work. The views and opinions expressed in this article are those of the authors and do not necessarily reflect the position of Fulton Hogan. The project was designed and conducted under the DSRU Summer Scholarship Program.

Data availability statement

The secondary data that were used in this study are available from the first author (N.W.). Restrictions apply to the availability of these data, which were used under license for this study. Data are available from N.W., subject to the permission of the associated contractor(s).

References

[1] Linden, R.N., Mahoney, J.P., and Jackson, N.C., 1989. Effect of compaction on asphalt concrete performance. *Transportation Research Record: Journal of the Transportation Research Board*, **1217**, 20-28. https://doi.org/10.1177/0361198196154300105

[2] Asphalt Institute, n.d. Asphalt Pavement Construction. https://www.asphaltinstitute.org/engineering/frequently-asked-questions/asphalt-pavement-construction/

[3] Schmitt, R.L., Russell, J.S., Hanna, A.S., Bahia, H.U., Jung, G.A., 1998. Summary of current quality control/quality assurance practices for hot-mix asphalt construction. *Transportation Research Record*, **1632**(1), 22-31. https://doi.org/10.3141%2F1632-03

[4] Lin, S., Ashlock, J., Williams, R.C., Lee, H.D., Wang, Y., 2018. Evaluation of three non-destructive testing techniques for quality assessment of asphalt pavements. *Nondestructive Testing and Evaluation*, **33**(4), 361-375. https://doi.org/10.1080/10589759.2018.1484921

[5] Wani, S.S. and Gharaibeh, N.G., 2013. Evaluating the use of contractor-performed test results in highway construction and material acceptance decisions. *Journal of Infrastructure Systems*, **19**(1), 67-73. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000113

[6] VicRoads, 2017. *Acceptance of Field Compaction Code of Practice RC 500.05*. Melbourne: Department of Transport. https://www.vicroads.vic.gov.au/-/media/files/technical-documents-new/codes-of-practice-rc500/code-of-practice-rc-50005--acceptance-of-field-compaction.ashx

[7] Department of Transport and Main Roads, 2022. *Technical Specification Transport and Main Roads Specifications MRTS30 Asphalt Pavements*. Queensland Government. https://www.tmr.qld.gov.au/-/media/busind/techstdpubs/Specifications-and-drawings/Specifications/5-Pavements-Subgrade-and-Surfacing/MRTS30.pdf?la=en

[8] Department of Planning, Transport and Infrastructure, 2018. *Specification: Part R28 Construction of Asphalt Pavements*. https://www.dit.sa.gov.au/__data/assets/pdf_file/0005/488381/Part_R28_Construction_of_Asphalt_Pavements_Effective_July_2018.pdf

[9] Mohammad, L.N., Elseifi, M., Cooper, S.B., Raghavendra, A., 2013. Levels of variability in volumetric and mechanical properties of asphalt mixtures. *Journal of Materials in Civil Engineering*, **25**(10), 1424-1431.

[10] Kim, Y.R., Lee, S., Seo, Y., El-Haggan, O., 2008. Determination of price reduction factors for density-deficient asphalt pavements. *Journal of Testing and Evaluation*, **36**(4), 335-344. https://doi.org/10.1520/JTE101460

[11] White, G., 2019. Evaluation of a non-nuclear density gauge as an alternative to destructive coring for airport asphalt acceptance testing. *SN Applied Sciences*, **1**:921. https://doi.org/10.1007/s42452-019-0958-6

[12] Australian Asphalt Pavement Association, 2004. *Nuclear density testing of asphalt pavements – Pavement Work Tips No. 42*. https://www.afpa.asn.au/wp-content/uploads/2017/03/Worktips-42-Nuclear-density-testing-of-asphalt-pavements.pdf

[13] Plati, C., Liozos, A. and Gkyrtis, K., 2020. Integration of non-destructive testing methods to assess asphalt pavement thickness. *NDT and E International*, **115**:102292. https://doi.org/10.1016/j.ndteint.2020.102292

[14] Plati, C., Gkyrtis, K. and Liozos, A., 2021. Integration of non-destructive testing data to produce pavement critical strains. *Nondestructive Testing and Evaluation*, **36**(5), 546-570. https://doi.org/10.1080/10589759.2020.1834555

[15] Tan, X., Abu-Obeidah, A., Bao, Y., Nassif, H. and Nasreddine, W., 2021. Measurement and visualization of strains and cracks in CFRP post-tensioned fiber reinforced concrete beams using distributed fiber optic sensors. *Automation in Construction*, **124**, 103604. https://doi.org/10.1016/j.autcon.2021.103604

[16] Tan, X., Du, J., Zhang, Q., Meng, W. and Bao, Y., 2024. Monitoring restrained shrinkage and cracks of ultra-high-performance concrete (UHPC) using distributed fiber optic sensors. *Construction and Building Materials*, **422**, 135789. https://doi.org/10.1016/j.conbuildmat.2024.135789

[17] Tarr, K. and Grobler, J., 2021. *Annual Summary Report. P59 - Best Practice in Non-destructive Testing of Assurance of Asphalt – Year 3 (2019-20)*. Queensland Department of Transport and Main Roads. https://www.nacoe.com.au/wp-content/uploads/2021/08/P59-PQI-Report-Y3-FINAL.pdf

[18] Alderson, A., 2011. *The influence of compaction on the performance of dense-graded asphalt* (Report AP-T194-11). Austroads. https://austroads.com.au/publications/pavement/ap-t194-11/media/AP-T194-11_The_Influence_of_Compaction_on_the_Performance_of_Dense_Graded_Asphalt.pdf

[19] ASTM International, 2017. ASTM D2726/D2726M-17. *Standard Test Method for Bulk Specific Gravity and Density of Non-Absorptive Compacted Asphalt Mixtures*. https://www.astm.org/d2726_d2726m-17.html

[20] ASTM International, 2018. ASTM D6752/6752M-18. *Standard Test Method for Bulk Specific Gravity and Density of Compacted Asphalt Mixtures Using Automatic Vacuum Sealing Method*. https://www.astm.org/d6752_d6752m-18.html

[21] Li, M. Huang. L. and Al-Jumaily, A., 2021. Methods for asphalt road density measurement: A review. *2021 27th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, 269-274. https://doi.org/10.1109/M2VIP49856.2021.9664991

[22] Goel, A. and Das, A., 2008. Nondestructive testing of asphalt pavements for structural condition evaluation: a state of the art. *Nondestructive Testing and Evaluation*, **23**(2), 121 - 140. https://doi.org/10.1080/10589750701848697

[23] Yan, M., Tan, X., Mahjoubi, S. and Bao, Y., 2022. Strain transfer effect on measurements with distributed fiber optic sensors. *Automation in Construction*, **139**, 104262. https://doi.org/10.1016/j.autcon.2022.104262

[24] Department of Transport and Main Roads, 2023. *Nuclear Gauge Testing Manual* (ed. 4, amendment 1). Queensland Government. https://www.tmr.qld.gov.au/business-industry/Technical-standards-publications/Nuclear-gauge-testing-manual

[25] Troxler Electronic Laboratories, Inc., 2009. *Manual of Operation and Instruction Model 4640-B Thin Layer Density Gauge* (PN 105005, ed. 6.7).

[26] Foroutan, M., Bijay, K.C., and Ghazanfari, E., 2020. Evaluation of correlations between intelligent compaction measurement values and in situ spot measurements. *Geo-Congress 2020 GSP 318*. 602-611. https://doi.org/10.1061/9780784482810.062

[27] Wen, H., Rose, M., Timm, A., 2015. *Review of non-nuclear density gauges as possible replacement for ITD's nuclear density gauges* (Report RP 210). Idaho Transportation Department. https://apps.itd.idaho.gov/apps/research/Completed/RP210.pdf

[28] Sivagnanasuntharam, S., Sounthararajah, A., Kodikara, J., 2023. A new approach to maximising the benefits of current intelligent compaction technology for asphalt materials. *Construction and Building Materials* **393**, 132031. http://doi.org/10.1016/j.conbuildmat.2023.132031

[29] Cheng, L., Lu, J., and Zhou, C. Non-destructive compaction quality evaluation of runway construction based on GPR data. *Nondestructive Testing and Evaluation*, in press. https://doi.org/10.1080/10589759.2023.2255363

[30] Hu, J., Vennapusa, P., White, D. and Beresnev, I., 2016. Pavement thickness and stabilised foundation layer assessment using ground-coupled GPR. *Nondestructive Testing and Evaluation*, **31**(3), 267-287. https://doi.org/10.1080/10589759.2015.1111890

[31] Plati, C. and Loizos, A., 2012. Using ground-penetrating radar for assessing the structural needs of asphalt pavement. *Nondestructive Testing and Evaluation*, **27**(3), 273-284. https://doi.org/10.1080/10589759.2012.695784

[32] TransTech Systems, 2016. *TransTech Systems' Pavement Quality Indicator*^{TM} *Technical Application Brief.* http://transtechsys.com/pdf/techBrief380.pdf

[33] TransTech Systems, 2021. *Pavement Quality Indicator PQI-380 Operator's Handbook* (Vers. 1.6). http://www.transtechsys.com/pdf/PQI380Manual.pdf

[34] White, G., and Alrashidi, F., 2021. Field evaluations of nuclear and non-nuclear gauges as alternates to destructive coring for airport asphalt density testing. *Green and Intelligent Technologies for Sustainable and Smart Asphalt Proceedings of the 5th International Symposium on Frontiers of Road and Airport Engineering, 12-14 July, 2021*, 241-246. https://doi.org/10.1201/9781003251125

[35] Zaman, S., Hassaan, M., Hussain, J., Hayat, U., 2020. Efficacy of non-nuclear methods used for hot mix asphalt density determination. *Civil Engineering Journal*, **6** (6), 1166-1179. http://doi.org/10.28991/cej-2020-03091538

[36] Romero, P., 2002. Evaluation of non-nuclear gauges to measure density of hot-mix asphalt pavements. The University of Utah, Salt Lake City, Utah, USA. https://portal.ct.gov/-/media/DOT/documents/dresearch/SPRMarylandFINALREPORTV4pdf.pdf

[37] Smith, B. and Diefenderfer, B., 2008. Comparison of nuclear and non-nuclear pavement density testing devices. *Transportation Research Record: Journal of the Transportation Research Board*. **2081**, 121-129. https://doi.org/10.3141/2081-13

[38] Zaman, S., Hussain, J., Zaidi, S.B.A., Ejaz, N., 2019. Field assessment of non-nuclear methods used for hot mix asphalt density measurement. *Civil Engineering Journal*, **5** (8), 1820-1831. http://doi.org/10.28991/cej-2019-03091374

[39] Ziari, H., Behbahani, H., Izadi, A., Divandary, H., 2010. Field evaluation of non-destructive tests in measuring the pavement layers density. *Journal of Applied Sciences*, **10** (23), 3091-3097. https://doi.org/10.3923/jas.2010.3091.3097

[40] Leng, Z., Zhang, Z., Zhang, Y., Wang, Y., Yu, H., Ling, T., 2018. Laboratory evaluation of electromagnetic density gauges for hot-mix asphalt mixture density measurement. *Construction and Building Materials*, **158**, 1055-1064. https://doi.org/10.1016/j.conbuildmat.2017.09.186

[41] ASTM International, 2010. ASTM D7113/D7113M-10. *Standard Test Method for Density of Bituminous Paving Mixtures in Place by the Electromagnetic Surface Contact Methods*. https://www.astm.org/d7113_d7113m-10.html

[42] Austroads, 2009. *Guide to Pavement Technology Part 3: Pavement Surfacings*. Austroads Ltd., Sydney. https://austroads.com.au/publications/pavement/agpt03/media/AGPT03-09_Guide_to_Pavement_Technology_Part_3_Pavement_Surfacings_Ed1.2.pdf

[43] Li, C.M., De, T., Chen, Z.Y., 2013. The improvement of PQI density testing method of considering texture depth of asphalt pavement. *Advanced Materials Research*, **718-720**, 1914–1918. https://doi.org/10.4028/www.scientific.net/amr.718-720.1914.

[44] Sivagnanasuntharam, S., Sounthararajah, A., Kodikara, J., 2023. In-situ spot test measurements and ICMVs for asphalt pavement: lack of correlations and the effect of underlying support. *International Journal of Pavement Engineering*, **24** (1), 2198770. https://doi.org/10.1080/10298436.2023.2198770

[45] Kvasnak, A., Williams, R., Ceylan, H. and Gopalakrishnan, K., 2007. *Investigation of Electromagnetic Gauges for Determining In-Place HMA Density*. *Final Report*. Report No. IHRB Project TR-547. Iowa State University. https://rosap.ntl.bts.gov/view/dot/23882

[46] Sargand, S., Kim, S., and Farrington, S., 2005. *A Working Review of Available Non-Nuclear Equipment for Determining In-Place Density of Asphalt.* Report No. FHWA/OH-2005/18. The Ohio Department of Transportation. https://trid.trb.org/view/781627

[47] Tarr, K., 2023. *Final Report. P59 – Best Practice Non-destructive Testing in Assurance of Asphalt: Evaluation of the Pavement Quality Indicator (2021-22)*. Queensland Department of Transport and Main Roads. https://www.nacoe.com.au/wp-content/uploads/2023/05/P59_21_22_Rep_Best-practise-non-descructive-testing-in-assurance-of-aspahlt_Y4_For-Publication.pdf

[48] Williams, S., 2008. *Non-nuclear Methods for HMA Density Measurements. Final Report*. Report No. MBTC 2075. University of Arkansas. https://rosap.ntl.bts.gov/view/dot/17014

[49] Darra, S., 2009. *Development of an Improved System for Oregon to Accurately Quantify Dense-Graded Hot Mix Asphalt Pavement Density*. Masters Thesis, Oregon State University. https://ir.library.oregonstate.edu/concern/graduate_thesis_or_dissertations/1z40kz10b

[50] Padlo, P., Mahoney, J., Aultman-Hall, L. and Zinke, S., 2005. *Correlation of Nuclear Density Readings with Cores Cut from Compacted Roadways. Final Report*. Report No. CT-2242-F-05-5. https://trid.trb.org/View/788441s

Table 1. Some recently published coefficients of determination (R^2) between core density (except where specified) and non-destructive methods.

Author and Year	**Device**	**Method Type**	**Project Type**	**Number of Samples**	***R^2***
White and Alrashidi (2021)	PQI-380	PQI in single mode	Airport	210 mat (2 projects, pooled)	>0.99
	PQI-380	PQI in single mode	Airport	45 joint (2 projects, pooled)	>0.99
	Troxler	NDG	Airport	210 mat (2 projects, pooled)	>0.99
	Troxler	NDG	Airport	45 joint (2 projects, pooled)	>0.99
Tarr and Grobler (2021)	PQI-380	PQI (‘high’ i.e. R^2>0.72 results found in 0% of mixes)	Road, various	6-173 per location (7 material types). No offset applied, collated data from contractor QA procedures.	0.06-0.71
Foroutan *et al.* (2020)	PQI	PQI vs Intelligent Compaction Measurement Value	Road HMA	118	0.01
Zaman *et al.* (2020)	PQI-380	PQI	Road: 25-45ºC ambient	30	0.82
	PQI-380	PQI	Road: 5-25ºC ambient	15	0.84
Zaman *et al.* (2019)	PQI-380	PQI	Road: Dense graded, ‘hotter’ conditions	150	0.75
	PQI-380	PQI	Dense graded road, ‘cooler’ conditions	45	0.81
White (2019)	PQI-380	PQI in single mode	Airport	14 (10 mat, 4 joint)	0.71
	PQI-380	PQI in average mode	Airport	14 (10 mat, 4 joint)	0.62
	PQI-380	PQI in single mode	Airport	111	0.74
Lin *et al.* (2018)	PaveTracker/ GeoGauge	PQI	Various road	2-20 per project (10 projects), averaged by lot and pooled	0.24
Wen *et al.* (2015)	PQI-301	PQI in average mode	Road: NMAS 12.5mm and 19.0mm	5-7 per project (14 projects, results pooled)	0.88

Author and Year	**Device**	**Method Type**	**Project Type**	**Number of Samples**	***R^2***
	PaveTracker	NNDG	Road: NMAS 12.5mm and 19.0mm	5-7 per project (12 projects, results pooled)	0.83
	Various	NDG	Road: NMAS 12.5mm and 19.0mm	5-7 per project (11 projects, results pooled)	0.89
Ziari *et al.* (2010)	PQI-301	PQI	Road	10	0.20
	Troxler HS-5001EZ	NDG	Road	10	0.01
Smith and Diefenderfer (2008)	Troxler 4640-B	NDG	Road: NMAS 9.5mm	73 (8 projects, pooled)	0.28
	PQI-300	PQI	Road: NMAS 9.5mm	73 (8 projects, pooled)	0.04
	PQI-300	PQI	Road: NMAS 9.5mm	<10 per project (8 projects)	0.02-0.92
	PQI-300	PQI	Road: 9.5mm NMAS	<20 per project (8 projects)	0.003-0.95
Romero (2002)	PQI-300	PQI (‘high’ i.e. R^2>0.72 results found in 17% of projects)	Various road	3-15 per project (74 projects)	0-0.98
	Various	NDG (‘high’ i.e. R^2>0.72 results found in 43% of projects)	Various road	3-15 per project (74 projects)	0-0.98
	PQI-300	Laboratory-prepared slab density vs PQI	Laboratory slabs: granite, gravel, and limestone; NMAS 12.5mm, 19mm, 25mm	8-18 per material type	0.93-0.99

Table 2. Data set features

Gauge type	**NDG**	**PQI**
State	Victoria	Queensland and South Australia
No. data points	2698	609 (330 Qld, 279 SA)
No. lots[1]	447	88 (49 Qld, 39 SA)
No. mix types	182	23 (17 Qld, 6 SA)

1. A lot is defined as a section of pavement of the same material composition, constructed during a single shift under uniform conditions and tested using the same gauge [6].

Table 3. Regression coefficients (slope and intercept), sample sizes (n), correlation coefficients (r), and residual standard error (RSE) values for core and gauge densities.

Gauge	**State**	**Data set**	**Offset applied**	**Slope (95%CI)**	**Intercept (95%CI)**	***n***	***r* (95%CI)**	**RSE**
NDG	Vic	All, pooled	No	1.005 (0.982, 1.028)	-0.098 (-0.150, -0.045)	2683	0.859 (0.848, 0868)	0.0467
		Plant A	No	0.937 (0.898, 0.975)	0.057 (-0.031, 0.144)	1033	0.830 (0.810, 0.845)	0.0443
		Plant B	No	1.043 (1.009, 1.077)	-0.183 (-0.262, -0.103)	1382	0.851 (0.835, 0.865)	0.0491
		Plant C	No	0.927 (0.862, 0.992)	0.074 (-0.076, 0.224)	268	0.863 (0.829, 0.891)	0.0397
		DG	No	0.982 (0.958, 1.006)	-0.042 (-0.097, 0.014)	2279	0.861 (0.850, 0.871)	0.0450
		FGG	No	1.065 (0.877, 1.254)	-0.198 (-0.617, 0.221)	18	0.941 (0.845, 0.978)	0.0232
		SMA	No	1.062 (0.982, 1.142)	-0.242 (-0.424, -0.059)	386	0.798 (0.758, 0.831)	0.0534
PQI	SA & Qld	All, pooled	No	0.519 (0.471, 0.566)	1.162 (1.049, 1.276)	609	0.657 (0.609, 0.670)	0.0526
			Yes	0.300 (0.266, 0.333)	1.645 (1.567, 1.727)	609	0.573 (0.517, 0.624)	0.0373
		DG	No	0.491 (0.440, 0.542)	1.226 (1.105, 1.347)	491	0.651 (0.597, 0.699)	0.0514
			Yes	0.241 (0.207, 0.275)	1.779 (1.697, 1.860)	491	0.530 (0.463, 0.591)	0.0347
		EME	No	0.710 (0.566, 0.854)	0.711 (0.359, 1.062)	111	0.679 (0.564, 0.768)	0.0514
			Yes	0.423 (0.347, 0.498)	1.380 (1.197, 1.563)	111	0.726 (0.624, 0.804)	0.0268
		SMA	No	0.535 (0.323, 0.747)	1.036 (0.507, 1.564)	7	0.911 (0.503, 0.987)	0.0132

		Yes	0.535 (0.323, 0.747)	1.036 (0.507, 1.564)	7	0.911 (0.503, 0.987)	0.0132
SA	All, pooled	No	0.455 (0.289, 0.622)	1.314 (0.927, 1.702)	279	0.307 (0.197, 0.410)	0.0458
		Yes	0.453 (0.333, 0.573)	1.282 (1.003, 1.562)	279	0.407 (0.304, 0.501)	0.0332
Qld	All, pooled	No	0.597 (0.518, 0.676)	0.966 (0.771, 1.160)	330	0.633 (0.564, 0.694)	0.0573
		Yes	0.170 (0.118, 0.223)	1.969 (1.839, 2.098)	330	0.330 (0.230, 0.423)	0.0383
Qld	Plant A	No	0.314 (-0.005, 0.634)	1.639 (0.876, 2.402)	126	0.171 (-0.004, 0.336)	0.0468
		Yes	0.498 (0.288, 0.707)	1.186 (0.686, 1.687)	126	0.386 (0.226, 0.525)	0.0307
	Plant B	No	0.645 (0.445, 0.849)	0.843 (0.335, 1.351)	176	0.430 (0.301, 0.543)	0.0663
		Yes	0.292 (0.162, 0.423)	1.659 (1.330, 1.988)	176	0.316 (0.176, 0.443)	0.0429
	Plant C	No	0.121 (0.047, 0.196)	2.090 (1.912, 2.267)	28	0.529 (0.194, 0.753)	0.0127
		Yes	0.129 (-0.014, 0.271)	2.087 (1.749, 2.425)	28	0.329 (-0.050, 0.625)	0.0241

Table 4. Paired t-test results for the difference between core and NDG density (i.e. core – NDG density), and core and PQI density (i.e. core – PQI density), with and without offsets applied.

Gauge Type	**Offset**	**t-statistic**	**95%CI**	**p-value**
NDG	No offset	95.211	0.084, 0.088	$< 2.2 \times 10^{-16}$
	Offset applied by lot	-0.239	-0.001, 0.001	0.811
PQI	No offset	-3.098	-0.014, -0.003	0.002
	Offset applied as per industry dataset	10.819	0.025, 0.036	$< 2.2 \times 10^{-16}$

Table 5. Differences between core density and NDG densities among different plants based on Tukey's honest significance test at the 5% family-wise error rate.

Plant	vs. Plant	Mean Difference	Std. Error	p-value	95% C.I.
VIC Plant A	Vic Plant B	0.005	0.002	0.026	(0.000, 0.009)
	Vic Plant C	-0.006	0.003	0.132	(-0.014, 0.001)
VIC Plant B	Vic Plant C	-0.011	0.003	0.001	(-0.018, -0.004)

Table 6. Differences between core density and PQI gauge densities among different plants based on Tukey's honest significance test at the 5% family-wise error rate.

Plant	vs. Plant	Mean Difference	Std. Error	p-value	95% C.I.
Qld Plant A	Qld Plant B	-0.046	0.007	<0.001	(-0.063, -0.030)
	Qld Plant C	0.006	0.012	0.958	(-0.024, 0.035)
	SA	0.044	0.006	<0.001	(0.029, 0.059)
Qld Plant B	Qld Plant C	0.052	0.011	<0.001	(0.023, 0.081)
	SA	0.091	0.005	<0.001	(0.077, 0.104)
Qld Plant C	SA	0.038	0.011	0.003	(0.010, 0.066)

Table 7. Acceptable lots based on correlation results for core and NDG, and core and PQI results.

	NDG	**PQI**
Lots with acceptable correlation coefficients ($r \geq 0.8$) [1]	162 (36.2%)	10 (11.5%)
Lots with unacceptable correlation coefficients ($r < 0.8$) [1]	285 (63.8%)	77 (88.5%)
Total lots	447	87

[1] As defined by a survey of US state road authorities [27]

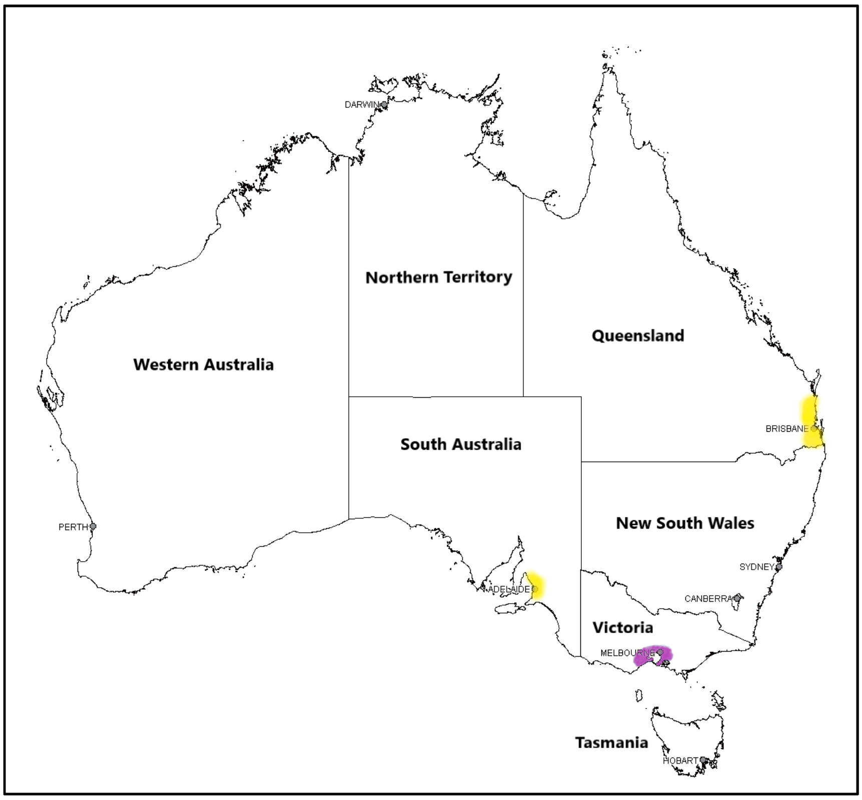
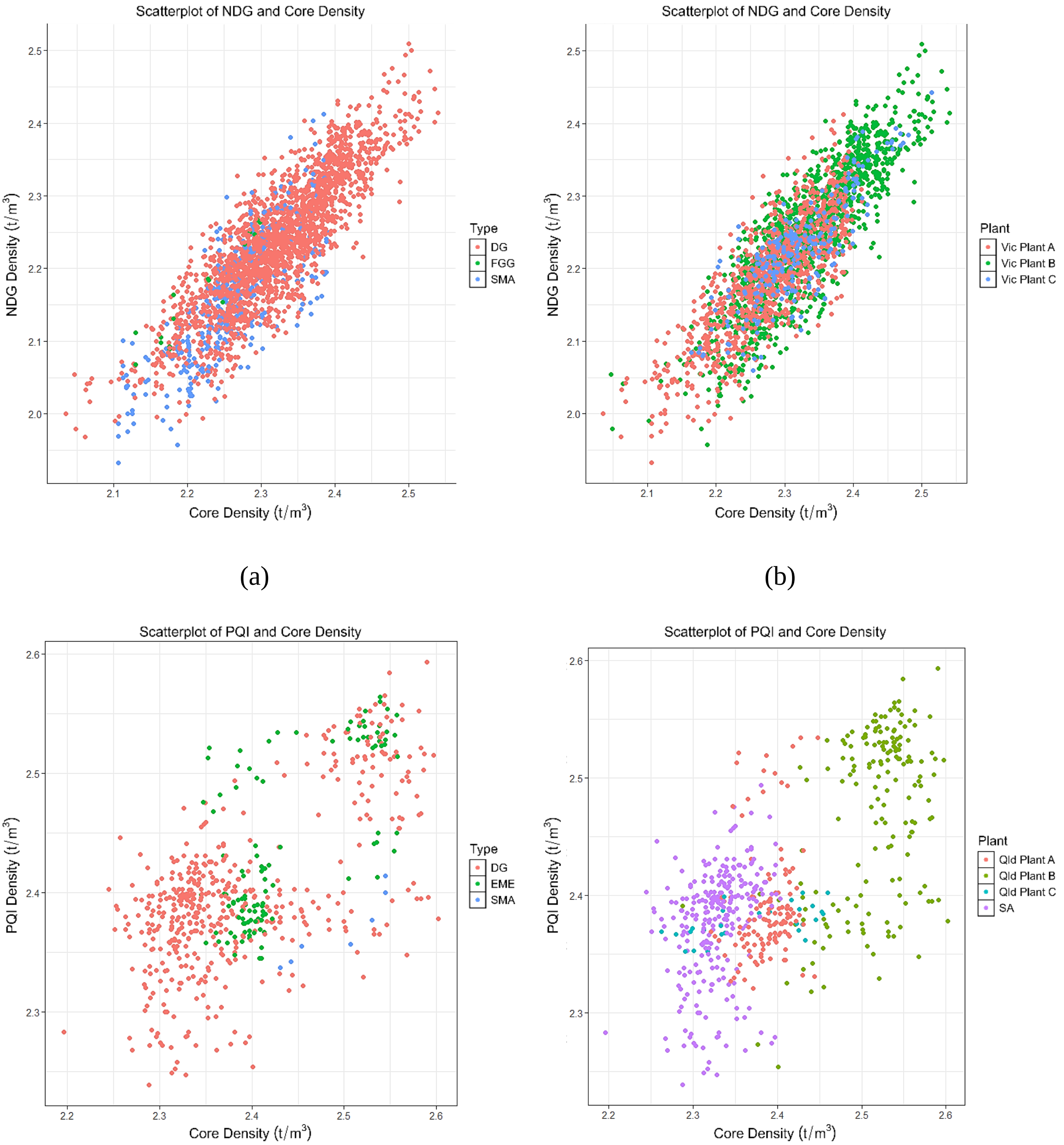


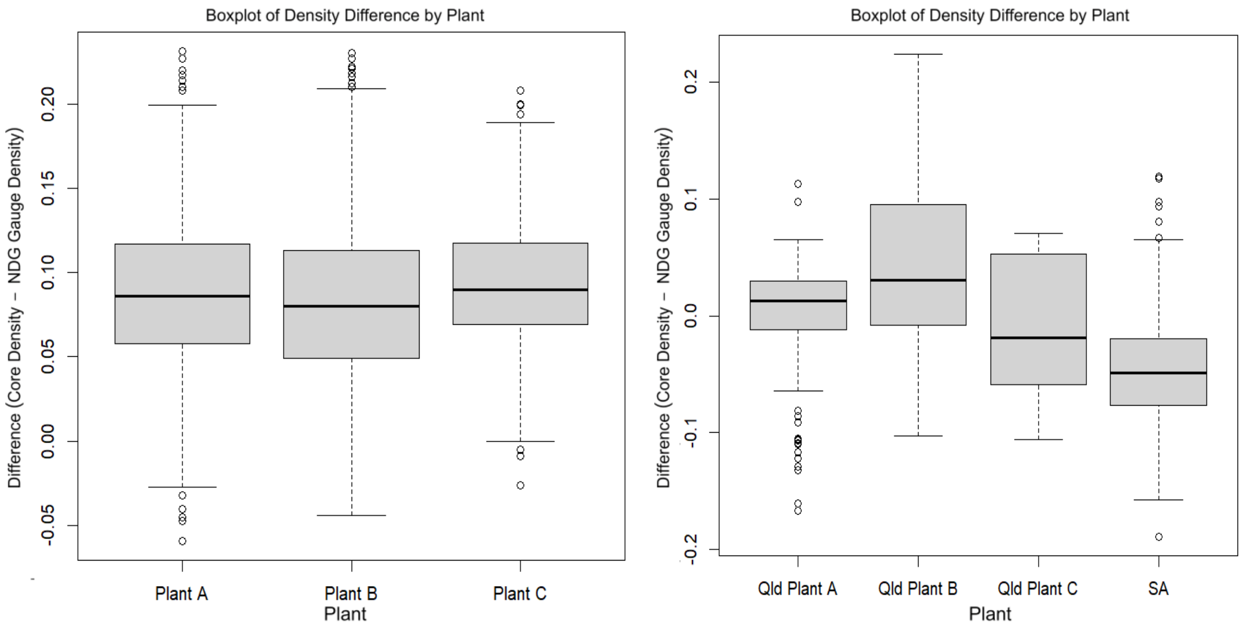
Figure 1. Map of Australia. Shaded yellow regions in Queensland and South Australia indicate approximate range of locations of PQI test sites; shaded purple region in Victoria indicates approximate range of locations of NDG test sites.



(a) (b)

(c) (d)

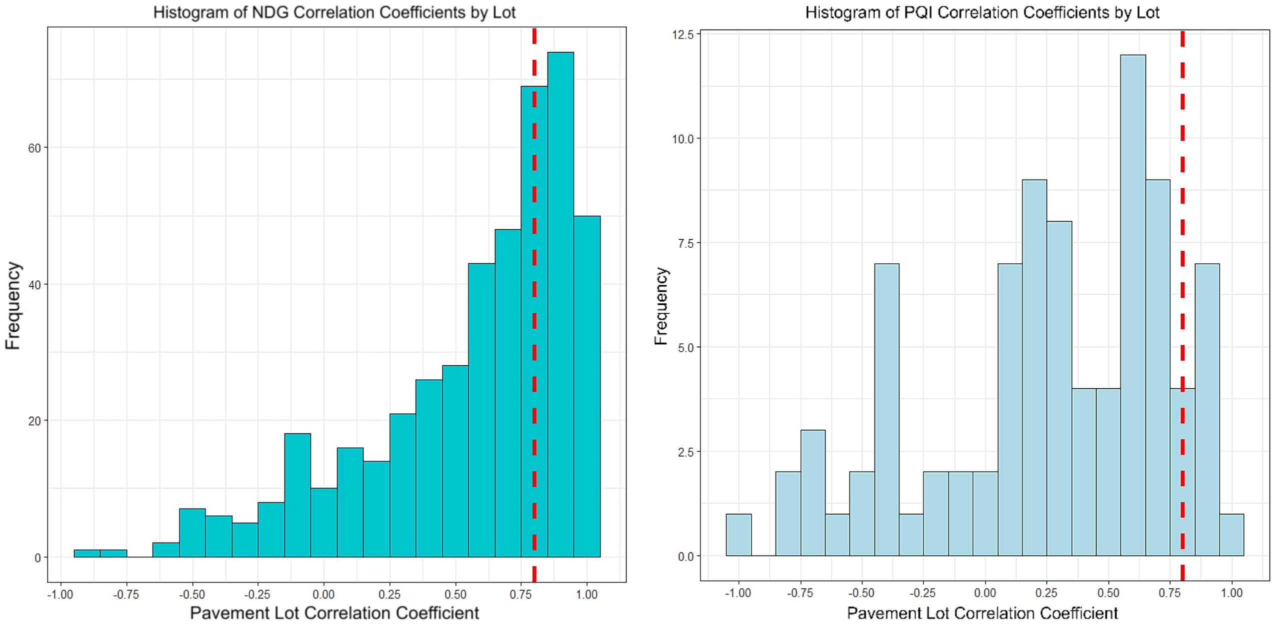
Figure 2. Scatterplot of laboratory-determined core density versus (a) Nuclear Density Gauge density with results highlighted by asphalt mix type; (b) Nuclear Density Gauge density with results highlighted by asphalt plant source; (c) Pavement Quality Indicator density with results highlighted by asphalt mix type; (d) Pavement Quality Indicator density with results highlighted by asphalt plant source.



(a)

(b)

Figure 3. Boxplot of (a) difference in density result (Core Density - Gauge Density) by plant for NDG Vic data, and (b) difference in density result by plant for PQI Qld and SA data. Outliers are indicated based on difference to the median exceeding 1.5 x Inter-Quartile Range.



(a) (b)

Figure 4. Histograms of r values by lot for (a) lots tested with NDGs, and (b) lots tested with PQIs. The red dotted line indicates the minimum acceptable correlation coefficient value (0.8) for alternative methods of density testing, as defined by [27].