Proceedings of the 14th Australia and New Zealand Conference on Geomechanics, Cairns 2023 (ANZ2023)

Earthworks testing and the density illusion

Burt Look
Formerly FSG – Geotechnics and Foundations, Australia, lookburt@hotmail.com

ABSTRACT: Density has historically satisfied a 1 parameter need in quality control assessment. Other quality parameters are assessed independently. But modulus and performance-based subgrade design aggregates influences such as density, moisture content, material quality, thickness influence, underlying material and equipment used. Comparing density ratio (DR) to other equipment measurements often leads to a poor correlation, since a multivariate relationship is required - Dendrogram analysis is used in this paper to illustrate this inter-relationship. Yet field supervisors often ask for a correlation to provide a linkage with DR as the de facto standard. A higher DR does not necessarily produce a higher strength or modulus – although this is an implicit assumption. Data from case studies with project trials over a period of 5 years are used to show the density illusion, which impedes the implementation of other modern testing. Issues associated with alternative tests are also discussed. Parallel testing results over several sites collectively indicate both Type I and Type II errors. Using 1-to-1 parallel testing can lead to DR results demonstrating a "pass" whilst LWD (or other tools) assessment could report a "fail" (or vice versa).

KEYWORDS: Dry density ratio, dendrogram analysis, subgrade modulus, compaction, test equipment.

1 INTRODUCTION

Quality assessment of earthworks remains dominated by the 1930s Proctor laboratory density compaction model coupled with the California Bearing Ratio (CBR) test. Over the past few decades, field compaction and testing equipment have moved ahead of these commonly employed quality techniques routinely used. Intelligent compaction (IC) now combines field compaction as a "test" in terms of another unit of measurement in the Intelligent Compaction Measurement Value (ICMV).

In compaction mechanics, relative compaction is measured as the density ratio (DR = ratio of field dry density to the laboratory maximum dry density). This is the most common quality test for compaction and is also one of the more precise tests as evidenced by its repeatability and reproducibility. Understanding of the energy – dry density and moisture content relationship for a particular soil is required during the compaction process.

The emphasis on density has led to the (incorrect) belief that it is the key parameter, yet it is an index only, i.e., we assume an increased density ratio means an increased strength or modulus or reduced permeability. Technology has now advanced to measure those parameters directly, yet road and approving authorities use density testing as the main quality evaluation parameter because of our longstanding experience.

Various in-situ devices have been available to industry for the past 2 decades and research has shown these have significant benefits. However, studies then try to correlate those measured parameters with the density ratio, with often poor correlations. Correlating to density is flawed as DR is not a fundamental parameter. An alternative testing approach is required which encapsulates the quality assurance (QA) required and avoids the paired correlation approach. DR is a Quality control (QC) index.

Look (2019, 2021) describe a methodology based on matching Probability density functions (PDFs). This first requires a best fit distribution functions as density are normally distributed, but other more accurate tests are non-normally distributed. Comparative strength and modulus values for varying compaction levels are presented using this methodology.

This shows a 95% density ratio does not have a singular strength or modulus value as measured with one-to-one testing. The material origin affects the strength or modulus.

The illusion of density as a reliable index of strength or modulus is shown. Type 1 and Type 11 errors occur if paired correlations are used to compare more accurate tests with the precise (but inaccurate) DR typically applied in quality control. Type 1 error is a good result being rejected and Type II error is when a poor result is not rejected. Ideally good results are accepted, and poor results are rejected. The several reasons for these poor correlations are shown.

2 WHAT DO ENGINEERS WANT FROM A QUALITY MEASUREMENT TEST

A survey of engineers ranking what attributes are desirable in a test equipment showed that accuracy is the most preferred attribute of any test (Look, 2018). The preference ranking order for the 8 attributes of a test equipment from that survey was:

1. Accuracy
2. Precision
3. Time to conduct test / Ease of use
4. Time to process results / Ease to process and report
5. Amount of data obtained / Capital cost of equipment

In any measurement, accuracy refers to closeness of the measurements to a "true" value, while precision refers to the closeness of the measurements to each other (repeatability). DR is precise and assumed to be accurate. How does one assess whether the DR is accurate? But more importantly, can DR be the key parameter to accurately predict the overall performance of the earth structure?

Equipment tested (Figure 1) can be broadly classified as penetration and surface-based tests. The Plate Load test (PLT) and Density Ratio (DR) or more correctly the Dry Density Ratio (DDR) are the reference axis (the spine) as these are the most accurate and precise tests, respectively. These 2 tests are also the historical anchor points in implementation of alternative testing methods. Research has shown a poor correlation of many test equipment (including ICMV) with DR.

Poor correlation between modulus and density measurements was reported in Meehan et al. (2012). McLain and Gransberg (2016) searched for correlations between potential alternatives to the nuclear density gauge (NDG). The tests included LFWD modulus and Clegg Impact Values (CIV). The results showed that no definite relationship among NDG and Modulus or CIV results could be found. Mazari et al. (2013) quantifies the equipment and operator related variabilities from a database of stiffness measurements made with four devices on eighteen separate specimens. Most devices are repeatable and reproducible if the moisture content and density are rigidly controlled. This conclusion leads to parallel testing of modulus and density rather than replacement testing.

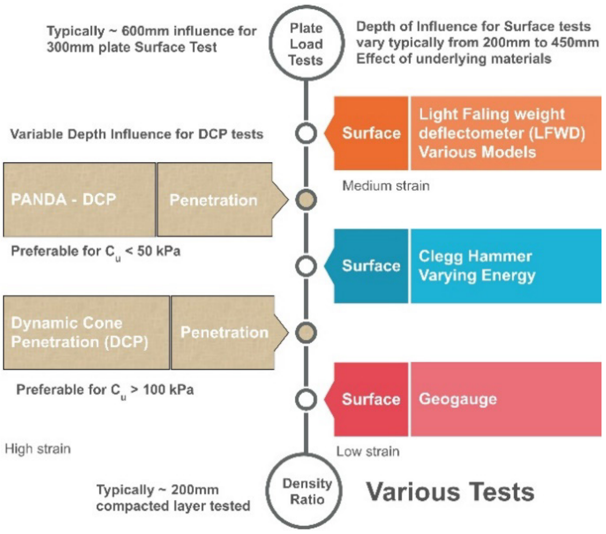

Figure 1. Alternative tests considered and the 2 historical anchor points of DR and PLT. Static and dynamic modulus vary.

Nazarian et al. (2014) reported on a Modulus-Based Construction Specification for Compaction of Earthwork and Unbound Aggregate. They concluded the adaptation of the modulus-based specification needs to be approached in the context of the levels of uncertainty associated with the current well-established density criteria. It was shown that achieving quality compaction (defined as achieving adequate layer modulus) is only weakly associated with achieving density.

The emphasis on in-situ quality testing does not reduce the requirement for design parameters which may require soaked lab values, while field values are at the time of compaction.

Field result trials were used to benchmark the various alternative equipment and compared with traditional DR measurements for various sites. The lot coefficient of variation (COV = standard deviation / mean) at each site was used to judge the precision, and the DR test is the standout leader (COV < 3.0%), with the Plate Load test (PLT), the most variable (COV of 77%) as compared to other tests which had a COV between 20% to 80%. The commonly used soaked CBR test had a typical COV of 40% but varying from 17% to 58% for the various lot tests. The results are summarized in Table 1.

Table 1. COV for various alternative tests considered over 5 sites.

Test	*Coefficient of Variation (%)*		
	Median	*Low*	*High*
Density Ratio	2.0	1.8	2.9
Geogauge	26.5	19.1	34.5
Prima LFWD	33.5	15.0	35.7
CBR	40	17.0	58
Zorn (LFWD)	34.1	21.6	51
Clegg	36.0	26.0	54
PANDA (50 – 100mm)	53	34.0	74
(150 – 200mm)	50	48	92
DCP (50 – 100mm)	38.0	28.0	97
(150 – 200mm)	53	34.0	74
Plate Load Test	77	14.0	142

The high precision of density was recognized in the early days and a governing factor in its implementation in quality control. For example, Selig and Truesdale (1967) examined the independent and joint effects for

- M – Moisture level; T – Lift thickness
- S – Soil type ; C – Compactive effort
- E – Compaction equipment

Table 2 summarizes this variation for the properties measured by Selig and Truesdale (1967). This ranking shows density has a low variability compared to other measurements. Moisture was the most significant factor influencing the strength and stiffness of the soil, but with poor repeatability. Thus, precision took preference over the more useful but wider variability of other soil measurements which were also more difficult to measure with the technology of the day.

Table 2. Range / average ratio of properties for all effects.

Measurement	*Range / Average (%)*
Dry Density	19
Wet Density	23
Seismic Velocity	75
Moisture Density	87
Plate Load	105
Moisture Content	112
Penetration Resistance	145
Field CBR	177

However, precision is the 2nd preference attribute as compared to accuracy. Accuracy was assessed in terms of how well the results compared ranked with each other for similar order of the high to the low values for the various test sites. The PLT as an industry accepted reference test was the most accurate test but is not as precise as the DR test (Look, 2018).

The DR and PLT remain the benchmarks for precision and accuracy, respectively, but overall did not reflect the other attributes desirable from testing equipment such as time or ease of use. These attributes are summarized in Figure 2, which compares the No.1 attribute of accuracy with the other considerations for traditional and alternative quality tests.

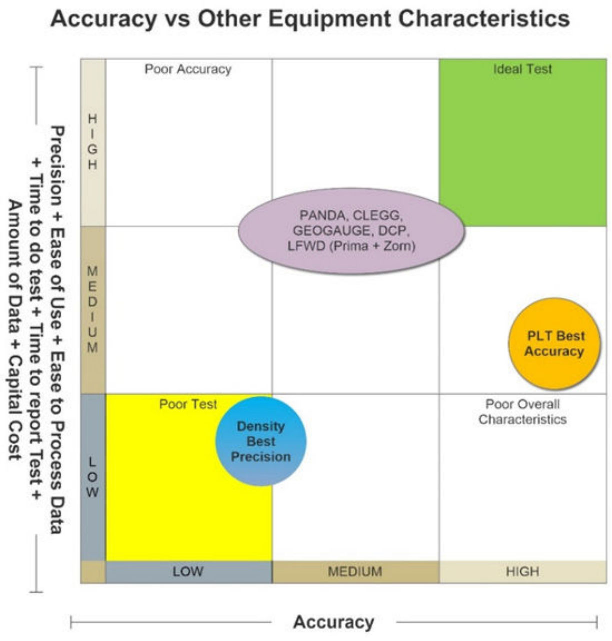
Accuracy vs Other Equipment Characteristics

Figure 2. Traditional (DR & PLT) and alternative tests compared with desirable attributes.

None of the test equipment showed both high accuracy and high precision. However, the results collectively indicate that a higher compacted density at a given site is not necessarily producing a higher value measured by the other equipment. In comparing the results at 5 sites, the order of the highest to lowest DR did not match the rank order as per the other tests carried out. Either DR is inaccurate, or all these alternative tests are collectively inaccurate.

Given these (unexpected) inconsistencies, one is now obligated to question

1. Why is DR not well corelated with these other tests?
2. Why is DR still the de facto standard of quality testing?
3. Does DR satisfy the current needs of the industry given these nascent technologies?

3 CORRELATING WITH DENSITY RATIO

The first question relates to other tests measuring multiple properties simultaneously. The depths of measurement are all different (Figure 3). The alternative tests are also measuring:

1. The density ratio, as well as
2. The moisture condition, and
3. The quality of the compacted material, and
4. The material underlying the layer being tested

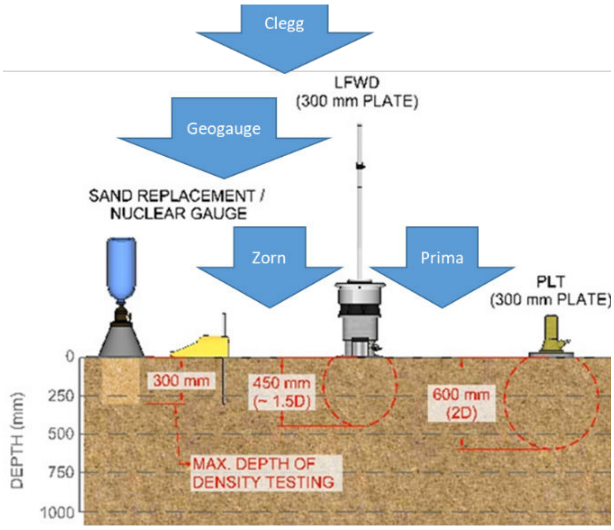

Figure 3. Zone of influence varies with each surface-based test. Dynamic cone penetrometer (DCP) and variable energy PANDA DCP would have extended depths

All the above factors are combined into one measurement, while the current traditional approach is to measure each of the above independently (Figure 4). Depending on the equipment there can be different proportions of these factors being measured. DR is the main target; however, this should not be interpreted as the only or most useful target. The former is part of the rationale in why this test has been the de facto standard. The traditional approach is for each quality to be measured separately, while alternative tests measure combined qualities.

3.1 *Dendrograms relationships*

In statistics, hierarchical clustering builds cluster trees (Dendrograms) to represent clustered data. The Cluster analysis searches for patterns in a data set to classify observations or variables into groups of related items. The analysis supports a variety of agglomerative hierarchical methods and distance measures. The clade is a branch in the tree. Clades that are close to the same height are like each other and clades with different heights are dissimilar.

Laboratory CBR testing is statically analyzed below with 55 data points from a "uniform" CH Cooroy clay with 5-point-soaked test values. This data is found in Look (1996). Look (2021) shows the dendrogram analysis for the Cooroy (CH) clay Soaked CBR. This clustering provides visual evidence that the CBR is more closely clustered to the compaction moisture and the OMC rather than the density (Figure 5).

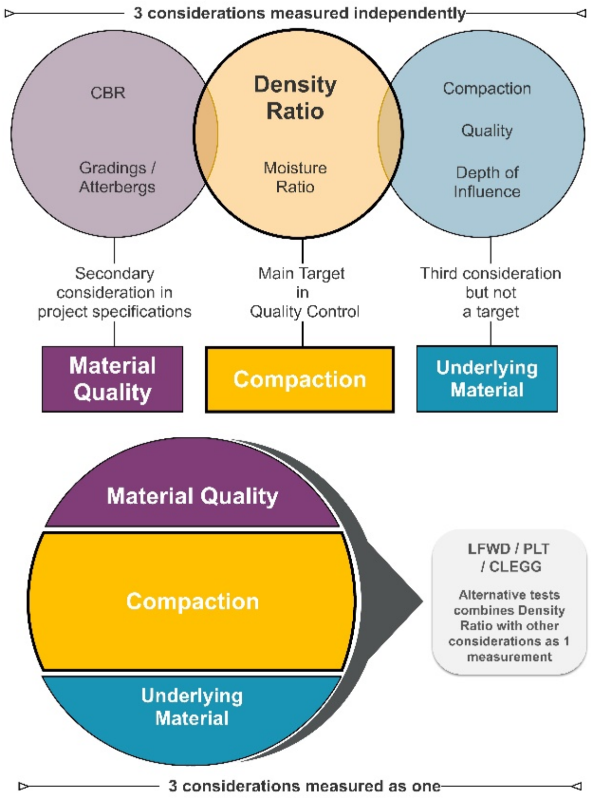

Figure 4. Current tests measure various attributes separately with DR as key target. Alternative tests measure multiple qualities

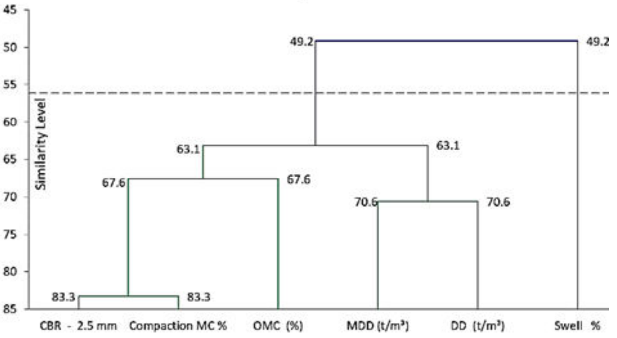

Figure 5. Dendrogram of key measurements in a soaked CBR test

The similarity level provides a visual representation of the closeness of clusters. The closer together suggests a high correlation. This could also be shown numerically in a correlation matrix (Table 3) and highlights this Cooroy clay is most strongly correlated with the compaction moisture content (0.69) and least with the dry density (0.04). However, the CBR is also strongly (negatively) correlated with swell (-0.83). This shows that the CBR for this expansive clay is most correlated to the swell value after soaking for the 4 days. Yet CBR is the key design input for pavement design and shown to be poorly correlated to the density measurements – the key QC parameter.

Figure 6 shows the same data dendrogram with the clustering for 15 relationships. As expected, the soaked CBR at 2.5mm and 5.0mm are clustered together and are the most strongly related. The CBR is most strongly related to the Moisture ratio (MR) at compaction, the compaction moisture content (MC) the Degree of saturation (DOS) before soaking, the density ratio (DR) when soaked and the DOS after soaking. The CBR is least related to the DR at compaction, the dry density and the MDD.

Table 3. Correlation Matrix for 6 No. test outputs.

Correlation Matrix	*Comp. MC %*	*DD (t/m³)*	*OMC (%)*	*MDD (t/m³)*	*CBR @2.5mm*	*Swell %*
Comp. MC %	1.00					
DD (t/m³)	-0.30	1.00				
OMC (%)	0.23	-0.38	1.00			
MDD (t/m³)	-0.04	0.46	-0.34	1.00		
CBR@2.5mm	0.69	0.04	0.40	0.32	1.00	
Swell %	-0.85	0.06	-0.14	-0.38	-0.83	1.00

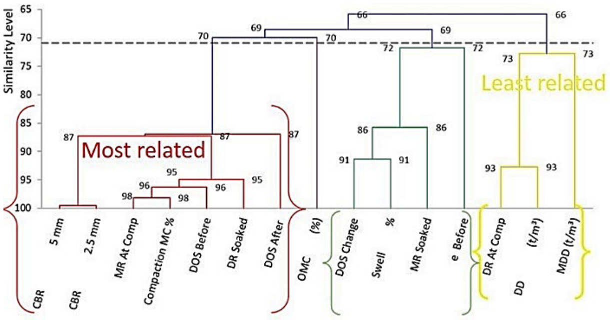

Figure 6. Dendrogram of 15 parameters in a soaked CBR test.

Similarly, dendrogram analysis comparing various insitu tests at a Ballina project site showed test equipment were comparable to each other, but least with the density ratio (Figure 7).

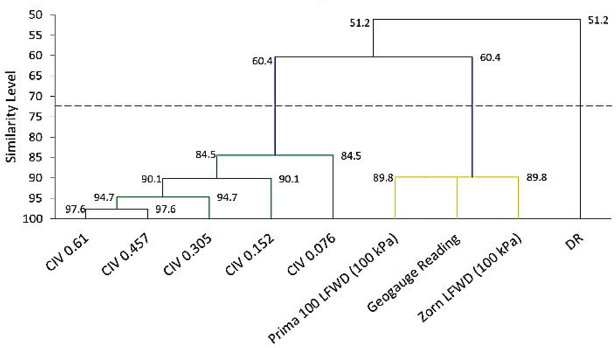

Figure 7. Dendrogram similarity for test equipment used at Ballina site.

The Clegg impact value (CIV) represents one grouping for various drop heights and as expected are similar to each other. The Prima LFWD, Zorn LFWD and Geogauge which measure modulus but with different strain and applied pressure are clustered together. The DR represented the furthest cluster, thus suggesting the least similarity with these other tests.

4 DR - THE CURRENT DE FACTO STANDARD

The second question can also be answered by the fact that DR is the most precise test as shown in Tables 1 and 2. DR has served the industry well in the past, and we are accustomed to this test with well-established standards and procedures. Standards are less developed for most of the other tests aside from the PLTs and DCPs. DR provides comfort to the past, simplicity, and high precision. Tests on the underlying material or the material quality are separate considerations. There is also no universal clear pass / fail value for the other tests, such as is currently used as in a 95% density ratio.

Comparing any 1 parameter (say compacted DR) to the other equipment measurement leads to a poor correlation, since a multivariate relationship is required. This approach would be impractical to implement in a field-based quality control system. Yet field supervisors often ask for a paired bivariate correlation to provide a linkage with DR as the de facto standard.

Additionally, a peak modulus or strength does not necessarily coincide with the maximum dry density (MDD), or compaction (and energy) used in the field. The MDD test requires curing periods and removal of oversize to have test repeatability and reproducibility. Neither of these occur during field compaction.

"New" technologies are more sensitive to change than traditional density tests. Figure 8 shows that 2.5% change in moisture content results in:

- 6% change in density, but.
- 300% change in modulus

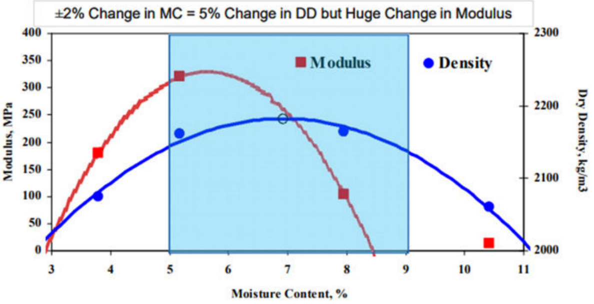

Figure 8. Impact of moisture on density and modulus (Nazarian and Correia, 2009).

The modulus is highly dependent on the moisture content. Drying or wetting after compaction can significantly influence the modulus result. A passing DR could have a low-test modulus if rainfall occurred shortly after the test. More layers should not be placed even with passing DR tests. Conversely, the test modulus can increase if considerable time (say ½ day) occurred after compaction was completed and with sun and wind drying.

This is also evident when 5-point CBR tests are caried out, the peak CBR is often not coincident with the OMC / MDD as shown in Figure 9 for this CH Cooroy clay. A high plasticity clay would have its high strength wet of optimum and below MDD (Figure 5), while a granular material would have its peak strength dry of optimum (below MDD).

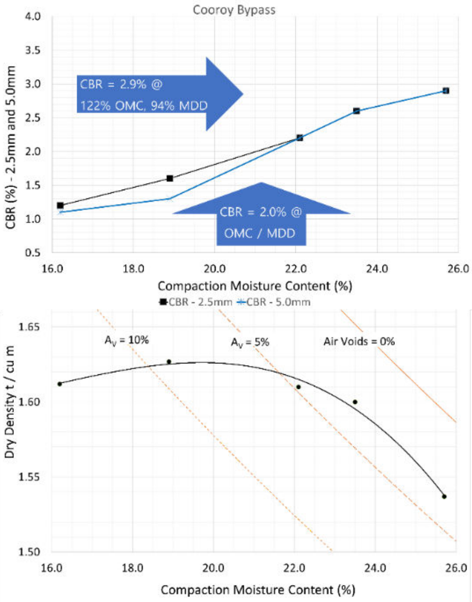

Figure 9. Maximum CBR is not coincident with MDD/OMC for CH clay.

5 DOES DR REPRESENT BEST PRACTICE GIVEN NASCENT TECHNOLOGIES

The third question is the ability of nascent technology to supplant DR tests. As the density ratio increases the strength or modulus is expected to increase. Typically, a 95% DR is specified at subgrade level. However, this DR does not correlate directly with modulus values as shown in Figure 10 for 3 materials in trials tested with 3 different source materials of residual soils.

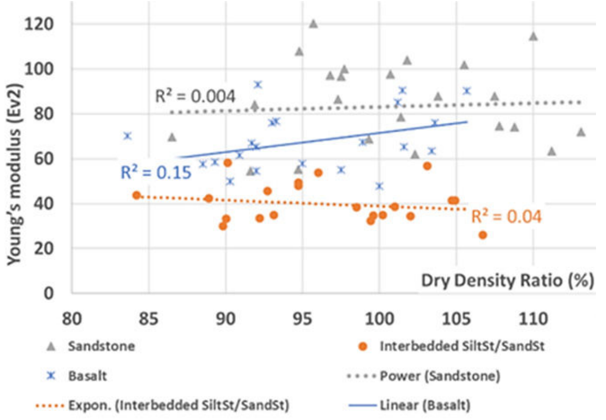

Figure 10. Poor correlations between modulus with DR.

Figure 10 shows the poor correlations with PLT E_{v2} results for different material origin. This often creates conflicts with both Type 1 and Type II errors. The current high esteem given to DR should be in question given its lack of accuracy (poor correlation) and other advanced attributes available from modern quality control testing. The outstanding precision of DR has overshadowed its poor accuracy even as an index parameter.

Look (2019) highlight this issue for granular materials, when the density ratio may be increasing due to reduction in the maximum dry density (MDD), and not from improved compaction. Crushing of the large particles occurs rather than compaction. Figure 11 shows this effect with the MDD decreasing at 8 No. passes for the residual soil / weathered sandstone material in Figure 10.

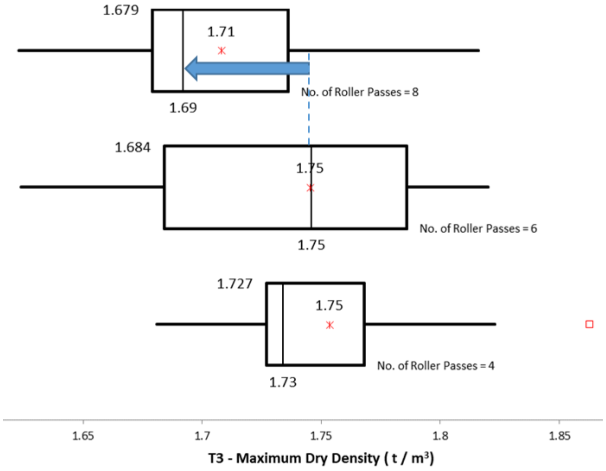

Figure 11. MDD may decrease although DR is increasing

Table 4 shows different source materials have dissimilar strength and modulus values irrespective of the same 95% density ratio achieved. This Table also highlights that different equipment may have different modulus values and absolute values may still need to be assessed with the PLT.

Note this is <u>not a rock fill</u> – and reference to rock type is only to show the origin of the fill. All source materials of residuals soils and weathered rock were observed to breakdown and conform to “soil” specifications during placement and compaction. and would all classify as a GW /GP/ GC based on its grading. The influence of the type of roller is also shown.

Table 4. In-situ modulus and strength E at 95% DR (Look, 2021).

Fill Material Origin	*Plate Load Test (PLT) E_{V2} (MPa)*	*In situ angle of friction φ (˚)*	
		Smooth	Padfoot
Sandstone	70	45	45
Interbedded Siltstone/Sandstone	40	41	39
Basalt	65	39	43

Figures 12 show the changing modulus with the corresponding DR in a box and whisker plot. This applies the method of matching PDFs (Look, 2019). For example, the lower quartile value of the sandstone is 70MPa and the corresponding DR of 95% is shown. Similarly, the median value and quartile values of 87 MPa and 100 MPa correspond to a DR value of 98% and 106%. respectively.

The median E_{v2} modulus was 62 MPa from 66 valid tests of all 3 geological source materials and for a typical DR of 97%.

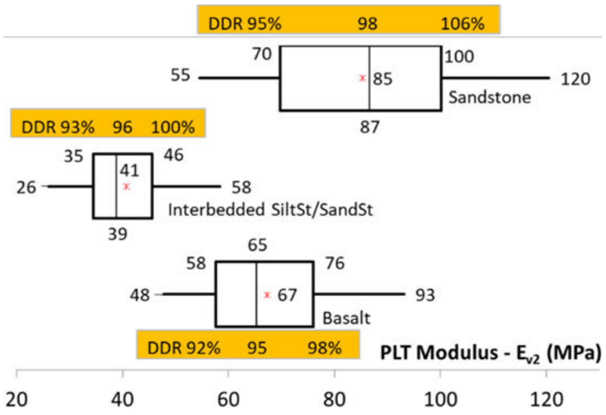

Figure 12. Material type affects Modulus at the DR.

Figure 12 shows that increasing DR is not matched with a singular increasing modulus for the varying test conditions. The effect of material origin is evident. The basalt had the lowest DR, but it was the interbedded material which had the lowest modulus at any given compaction DR. A compaction at 95% DR represents an E_{v2} of 70, 38 and 65 MPa for the sandstone, interbedded siltstone/sandstone and basalt, respectively. At 100% DDR there is a factor of 2 (90 to 46 MPa) between the Sandstone and the interbedded materials.

A similar analysis was carried out for the other materials and the results are summarized in Table 4. The sandstone achieves a significantly higher friction angle. The roller type has an effect for the interbedded siltstone /sandstone and basalt, but no effect for the sandstone material. The padfoot was better for the basalt, but less so for the interbedded material at 90% to 95% DDR.

DR tests are also a lag indicator, while other tests can provide more timely results. A lab MDD is required as part of the DR test. In granular materials little change in test results occurs with curing time, while for highly plastic clays, 4 to 7 days of curing results in large difference in OMC, MDD and CBR test results as compared to no curing time. Using a nuclear density gauge with an assumed MDD may be applicable for an in-situ assessment of processed granular materials but has significant errors with natural materials where greater variability occurs.

To answer the current needs of industry, one should consider a hierarchy of needs.

6 HIERARCHY OF NEEDS

In the field of behavioral psychology, Maslow introduced his concept of a hierarchy of needs. This hierarchy suggests that people are motivated to fulfill basic needs before moving on to other, more advanced needs. There are five distinct levels of Maslow's hierarchy of needs. The hierarchy is often depicted as a pyramid to represent the need to fulfill the lower levels before an individual can move up to the next level. Without fulfillment at the lower level below, one is unlikely to progress because of the lack of motivation to do so. Each need builds on the last, allowing greater fulfillment. The lowest to highest needs are

- Basic needs which involve
  - Physiological needs (Food, water, warmth, and rest) and Safety needs (security and safety)
- Psychological needs which involve
  - Relationships, love and belonging
- Esteem. Confidence, respect of and by others
- Self-actualization

This hierarchy of need concept can also be used to understand our current testing quality procedures and the need to advance to higher levels. Figure 13 uses this analogy to find where DR sits in the needs of industry. Density satisfies the basic needs in terms of its reliability as the test is both easy and precise. However, density alone cannot account for the performance of the earth structure which is the highest-level need.

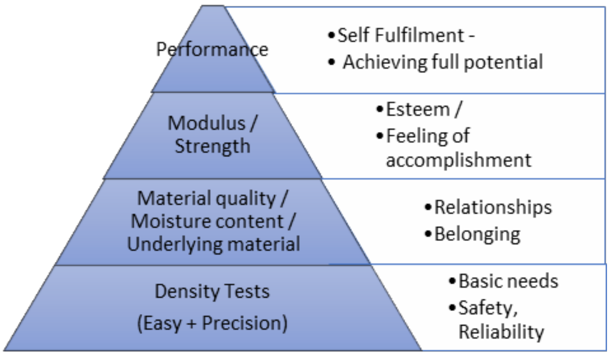

Figure 13. Testing hierarchy requirements with Maslow analogy.

Relationship needs are of a higher order than the basic need. The integration of density with the material quality, moisture content and the underlying material is a higher order need which is required to assess the design criteria of modulus / strength. The latter influences the goal of performance of the earth structure.

Use of alternative quality tests therefore requires letting go of the "low level" need of DR as the de facto standard for quality assessment of earthworks (Figure 14). Overall, DR testing is associated with quality control while other testing is more associated with quality assurance (Look, 2021).

Figure 14. Shackled to the past when paired correlations used with DR.

7 SUMMARY

Dry Density ratios are applied widely in quality control for earthworks testing. Yet because of its widespread usage, this now acts as an impediment to the application of alternative methods of testing. Density testing is a lag indicator and unable to provide a reliable indication on the ground strength or modulus when geotechnical and pavement designs are based on these parameters. DR has low accuracy, but high precision and simplicity. The measurement technology of the time was limited foe other key parameters and DR was chosen as the key QC measurement.

Several devices have been available for the past decades and while used in design should also be used in quality control testing. Density has historically satisfied our basic needs in quality assessment. Performance-based subgrade design aggregates influences in

- Density
- Moisture Content
- Material quality
- Underlying material

Correlating to the basic needs testing of density only, is therefore blind to the significant influences of the other factors influencing strength and modulus – the higher needs design basis. Most alternative measurements combine all the above in one measurement, while the current approach is to measure each of the above independently. Depending on the equipment there can be different proportions of these factors being measured. A 95% density ratio does not have a singular strength or modulus value as measured with one-to-one testing.

The decision is between the past comforts and success of using density measurements (confined to our basic needs) vs the higher order needs and benefits of alternative modern testing equipment.

8 REFERENCES

Look, B. G. (2018). Compaction QA Limitations: Benefits of alternative testing methods, *AGS Victoria Symposium on Geotechnics and Transport Infrastructure.* 7.

Look, B.G. (2019). Overcoming the current density testing impediment to alternative quality testing in earthworks. *Australian Geomechanics Journal,* Vol. 55, No. 1, pp. 55-74.

Look, B.G. (2021). An earthworks quality assurance methodology which avoids unreliable correlations. *4th International Conference on Transportation Geotechnics,* Chicago, USA. In: Tutumluer, E., Nazarian, S., Al-Qadi, I., Qamhia, I.I. (eds) Advances in Transportation Geotechnics IV. Lecture Notes in Civil Engineering, vol 166. pp 179 – 192, Springer

Look, Burt G (2021). Strength and modulus changes associated with compaction of residual soils and weathered rock. *Australian Geomechanics Journal*, Vol 56, No.1, pp 79 – 104

Mazari M, Garcia G, Garibay G, Abdallah I and Nazarian S, (2015). Impact of modulus-based device variability on quality control of compacted geomaterials using measurement system analysis, *TRB 2013 annual meeting*.

McLain K., and Gransberg, D.D. (2016). Contractor-Furnished Compaction Testing: Searching for Correlations Between Potential Alternatives to the Nuclear Density Gauge in Missouri Highway Projects, *International Journal of Quality and Innovation, Vol 3 (1).*

Meehan C, Faraz T and Farshid V (2012). A comparison of density – base and modulus base in situ tests for earthworks quality control. *Geotechnical Testing Journal*, Vol 35, No. 3, pp 387 to 399.

Nazarian S, Mazari M, Abdallah I, Puppala I, Mohammad L, and Abu-Farsakh M, (2014). Modulus-Based Construction Specification for Compaction of Earthwork and Unbound Aggregate. Draft Final report, *National Cooperative Highway Research Program NCHRP Project 10-84*, Transportation Research Board.

Selig E T, and Truesdale W B (1967). Properties of Field Compacted Soils. Highway Research Record, Issue 177, *National Research Council (U.S.). Highway Research Board*, pp 77 – 97.