

Optimum crumb rubber content for fly ash and stone dust stabilised subgrade: a field study

M Harishkumar

Holy Mary Institute of Technology and Sciences

R Samundeswari

Karunya Institute of Technology and Sciences

Arunraj E

Karunya Institute of Technology and Sciences

Godson caleb J

Karunya Institute of Technology and Sciences

Vincent Sam Jebadurai S

svsjsamy@gmail.com

Karunya Institute of Technology and Sciences

G Hemalatha

Karunya Institute of Technology and Sciences

C Mightraj

Holy Mary Institute of Technology and Sciences

Research Article

Keywords: Subgrade stabilisation, Crumb rubber, Fly ash, Stone dust, Plate load test, California bearing ratio, Surrogate model validation, Waste valorisation, Sustainable development goals

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11192827/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Optimum crumb rubber content for fly ash and stone dust stabilised subgrade: a field study

Harishkumar M¹, Samundeswari R², Arunraj E², Godson caleb J², Vincent Sam Jebadurai S², Hemalatha G², Mightraj C¹

¹ Holy Mary Institute of Technology and Sciences, Hyderabad, India 501301

² Karunya Institute of Technology and Sciences, Tamilnadu, India 64114

Abstract

A two-stage experimental programme quantifies the dosage response of a crumb-rubber-amended, fly ash and stone dust stabilised lateritic subgrade, and then uses that response as an external criterion against which a surrogate model built on a synthetic corpus for the same material system is audited. Laboratory screening on triplicate specimens fixed the base mix at 20% stone dust and 20% fly ash, raising unconfined compressive strength from 0.47 to 1.12 MPa at a coefficient of variation of 2.97%, the lowest of the four candidate combinations screened at equal total replacement. Six field sections were then constructed on that base mix with crumb rubber substituted at 0, 1, 2, 3, 4 and 5% by dry weight, and each was characterised in place by California bearing ratio and by 300 mm plate load testing. Both measurements resolve an interior optimum at 2% rubber, at which bearing ratio increases by 10.9% and the modulus of subgrade reaction by 11.9% relative to the rubber-free base mix. Beyond the optimum both decline monotonically, and at 5% rubber the bearing ratio falls 5.1% below the base mix, so that excess rubber is detrimental. The two tests share no apparatus, loading geometry, depth of influence or governing standard, and every reported quantity is recomputed from the primary load-penetration and load-settlement records. A stacked ensemble with Gaussian process residual correction, trained on an 800-record synthetic corpus and reaching nested cross-validated R² of 0.934, 0.939 and 0.888 with calibrated prediction intervals, nevertheless predicts monotonic decline across the same dosage range, resolves no interior optimum, and overpredicts the field bearing ratio 8.1-fold. Internal validation statistics computed within a synthetic corpus are therefore not diagnostic of physical validity.

Keywords: Subgrade stabilisation; Crumb rubber; Fly ash; Stone dust; Plate load test; California bearing ratio; Surrogate model validation; Waste valorisation; Sustainable development goals

1. Introduction

Pavement foundations consume large volumes of quarried granular material, and subgrade improvement using industrial and post-consumer waste offers a route to reducing that demand while diverting fly ash, quarry fines and end-of-life tyre rubber from disposal. The three materials considered here act through distinct mechanisms. Fly ash contributes pozzolanic reaction products that cement the soil matrix progressively with curing. Stone dust contributes angular fines that improve the gradation and raise the frictional component of shear strength. Crumb rubber contributes neither: it is an elastomeric inclusion that reduces unit weight, increases damping and increases compressibility.

That distinction is the reason rubber content should be expected to exhibit an interior optimum rather than a monotonic effect. At low dosage, rubber granules occupy inter-particle voids and introduce a compliant phase that redistributes contact stress without interrupting the granular skeleton, so measured bearing capacity may rise. Above some dosage the elastomeric phase becomes continuous enough to carry load paths itself, soil grain contact is progressively lost, and both strength and stiffness decline. The dosage at which this transition occurs depends on host soil gradation, rubber particle size and aspect ratio, and compaction energy, and it cannot be predicted from constituent properties alone. Tasalloti et al. [41], reviewing the soil-rubber literature, identify rubber content and the ratio of rubber to soil median particle size as the two governing variables and note that published results have not been compiled in a form that permits direct cross-study comparison. The transition dosage is therefore a quantity that must be measured for a given soil, rubber grading and compaction procedure.

In parallel, mix design for waste stabilisation has moved rapidly toward data-driven surrogate modelling. The motivation is practical: a mix space spanned by three replacement fractions, a rubber grading, a compaction state and a curing duration contains far more candidate designs than any laboratory programme can test, whereas a surrogate trained on a modest dataset can be searched exhaustively. A recurring obstacle is data scarcity, and a common response is to generate a synthetic design-of-experiments corpus, whether by simulation, by interpolation of published values or by generative modelling, and to train on that corpus.

This practice creates a validation gap that is seldom closed. A surrogate trained on a synthetic corpus and validated by cross-validation within that corpus is being tested for self-consistency rather than for physical fidelity. Held-out synthetic records are drawn from the same generator as the training records, so any systematic departure of the generator from the true response surface is invisible to every internal metric, including the coefficient of determination, the root-mean-square

error and the prediction interval coverage. The question relevant to engineering use is whether the surrogate reproduces the features on which a designer would act, and answering it requires measurements that the generator never saw.

1.1 Objectives and contributions

The objectives of this study were, first, to determine experimentally whether crumb rubber addition to a fly ash and stone dust stabilised lateritic subgrade produces an interior optimum, and to locate it; and second, to test whether a surrogate model trained on a synthetic corpus for the same material system reproduces that feature.

The contributions are as follows.

1. An interior optimum in crumb rubber dosage is established for a fly ash and stone dust stabilised lateritic subgrade under field placement and compaction, resolved at 1% increments across 0 to 5% by dry weight and located concordantly by two mechanically independent in-situ tests that share no apparatus, loading geometry, depth of influence or governing standard. Dosage-resolved field verification at this granularity is not otherwise available for rubber-amended stabilised subgrade, and the concordance of two independent measurements excludes an apparatus-specific artefact as the origin of the optimum.
2. The divergence between that measured response and a surrogate trained on a synthetic corpus for the same material system is quantified rather than asserted. The surrogate attains nested cross-validated R^2 of 0.888 to 0.939 with empirically calibrated prediction intervals, yet returns a monotonic dosage response containing no stationary point and overpredicts the field bearing ratio 8.1-fold. Internal validation statistics are thereby shown to be non-diagnostic of divergence between a corpus and the material it represents.
3. A regularised linear baseline fitted to the same engineered features reproduces the accuracy of the full stacked ensemble, establishing that the reported accuracy is a property of the corpus rather than of model capacity. This control is frequently omitted in applied geotechnical machine learning, and without it an accuracy figure is uninterpretable.
4. A reproducibility protocol for reported geotechnical quantities is specified and applied to the primary records underlying this study, recomputing every reported value from its source curve. It identified a misaligned load column, a dimensionally incorrect unit, one section whose reported stiffness is not derivable from its own load-settlement curve, and a resilient modulus that is a correlation of bearing ratio rather than an independent measurement. Faults of this class propagate silently into any corpus generated from reported values, and the protocol is transferable to any study that does so.

5. The claims of this paper are bounded as follows. The measured optimum is established for one soil, one base mix, one rubber gradation and one compaction regime; its position is a property of the compacted skeleton and is not asserted to transfer to other systems. The surrogate audit establishes that high internal validation scores were compatible with physical invalidity in this instance; it does not establish that synthetically generated corpora are invalid in general, which a single case cannot demonstrate. The synthetic corpus is the object of examination in this study rather than a source of evidence, and no statement made here about the behaviour of the stabilised subgrade rests upon it.

2. Materials and experimental programme

2.1 Materials

The host soil was a lateritic subgrade obtained from the project site, classified as a fine-grained soil of low plasticity with a liquid limit of approximately 32% and a plasticity index of approximately 8.5% in its natural state. Fly ash was sourced from a thermal power station, stone dust from a local granite crushing unit as the fine fraction of crushing, and crumb rubber from a tyre recycling unit as a shredded granulate. Figure 1 documents the materials and the field programme.



Constituent materials: subgrade soil, stone dust, fly ash, crumb rubber



Field test plots: conventional (E) and Mix A to Mix D



Spreading of stabilisers



Compaction of the layer



300 mm plate and gauges



Loading via reaction vehicle

Figure 1. Experimental programme.

Constructing all six sections adjacent to one another on the same subgrade is the design feature that makes the comparison between rubber contents interpretable. Testing sections at separate locations or on separate dates would confound rubber content with natural variability in the host soil, which for a lateritic deposit can exceed the effect under investigation.

2.2 Laboratory screening and selection of the base mix

The fly ash and stone dust levels were fixed before rubber was introduced, by a replicated laboratory screening programme. Unconfined compressive strength was determined in accordance with IS 2720 (Part 10): 1991 on cylindrical specimens 38 mm in diameter and 76 mm in height, a height to diameter ratio of 2:1, compacted at optimum moisture content, with three specimens tested at each replacement level. Stone dust was varied from 0 to 30% and fly ash from 0 to 40% of the dry soil weight, each in isolation, after which four combinations were tested. Figure 2 reports the results.

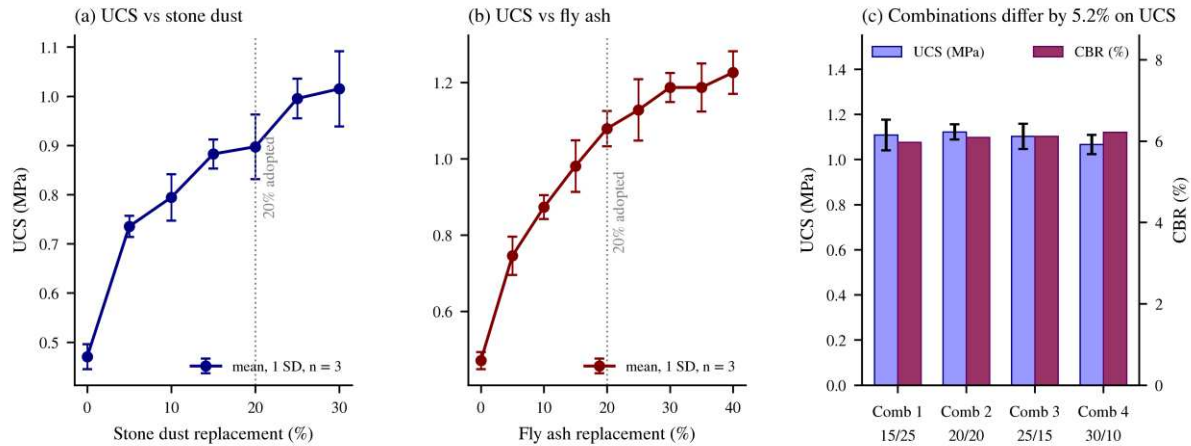


Figure 2. Laboratory screening of the base mix.

Three features of these results governed the design of the field programme. Strength rises sharply at low replacement and then saturates: stone dust raised compressive strength from 0.471 MPa untreated to 0.883 MPa at 15% and only to 1.015 MPa at 30%, while fly ash raised it from 0.471 MPa to 0.981 MPa at 15% and 1.226 MPa at 40%. Replacement beyond approximately 20% therefore yields little additional strength for a large mass of added material. Second, the four combinations are not distinguishable from one another at this level of replication, with a 5.2% spread in strength against coefficients of variation between 2.97 and 6.11%, so the selection among them rests on economy and workability rather than on strength. Third, the replicate scatter is small and consistent, with coefficients of variation between 2.93 and 7.54% across all levels of both single-additive series. This last figure establishes that the testing procedure is repeatable to within a few percent, and it provides the scale against which the rubber effect reported in Section 3.3 must be judged.

The combination of 20% stone dust and 20% fly ash was adopted as the base mix for all field sections. It gave the highest mean compressive strength of the four combinations at 1.122 MPa, the lowest coefficient of variation at 2.97%, and it weights the two additives equally, which avoids reliance on either the pozzolanic or the gradational mechanism alone.

2.3 Field test sections

Six field sections were constructed on the adopted base mix, differing only in crumb rubber content. Table 1 gives the designations used throughout.

Designation	Crumb rubber (%)	Stone dust (%)	Fly ash (%)
Conventional	0	20	20

Mix A	1	20	20
Mix B	2	20	20
Mix C	3	20	20
Mix D	4	20	20
Mix E	5	20	20

Table 1. Field mix designations.

Each plot was excavated, the stabilisers were spread and blended to the full layer depth, and the layer was compacted at optimum moisture content. Rubber content is expressed as a percentage of the dry weight of the soil, consistent with the convention used for the other two additives.

2.4 California bearing ratio testing

Bearing ratio was determined from the load-penetration response following IS 2720 (Part 16), with load recorded at 0.5 mm penetration increments to 5.0 mm. The bearing ratio is the applied load at a specified penetration expressed as a percentage of a standard load, taken as 1370 kg at 2.5 mm penetration and 2055 kg at 5.0 mm, with the governing value taken as the larger of the two.

2.5 Plate load testing

Plate load tests followed IS 1888 using a 300 mm diameter rigid bearing plate loaded in increments against a reaction vehicle, with settlement read on dial gauges. The modulus of subgrade reaction, k , is the applied bearing pressure divided by the settlement at the 1.25 mm datum:

$$k = p / s, \text{ where } s = 1.25 \text{ mm, } p = P / A, \text{ and } A \text{ is the plate area (1)}$$

with P the applied load and A the area of the 300 mm plate. The dimensions of k are pressure divided by length, so the appropriate unit is kN/cm^3 . The unit kN/cm^2 that appears in the authors' earlier records is dimensionally incorrect; Section 4.4 sets out how this was identified.

2.6 Data reduction and verification protocol

Because the central claim of this paper is a comparison between measured and modelled behaviour, the measured quantities were not accepted as recorded. Four checks were applied to every reported value before it entered the analysis.

1. Recomputation from primary data. Each reported bearing ratio and modulus of subgrade reaction was recalculated from its own tabulated load-penetration or load-settlement curve using the standard formulae of Sections 2.4 and 2.5.
2. Dimensional consistency. Each reported unit was checked against the dimensions of the formula used to obtain the quantity.
3. Circularity. Each reported quantity was tested against empirical correlations of the other reported quantities, to establish whether it constitutes an independent measurement.
4. Internal consistency of primary data. Load and settlement columns were checked for duplication between sections and for alignment against their independent variable.

All four checks returned findings, which are reported in Section 4.4 rather than being applied silently. Where a reported and a recomputed value differ, both are given.

2.7 The synthetic corpus

Independently of the field programme, a surrogate was developed on an 800-record synthetic corpus describing the same material system. Each record specifies a base soil through its liquid limit, plasticity index and specific gravity; a replacement mix through fly ash, crumb rubber and stone dust percentages together with a categorical rubber grading; compaction and moisture state through water to cement ratio and maximum dry density; and a curing duration. Four responses are recorded: compressive strength, bearing ratio, resilient modulus and fatigue life.

This corpus is synthetic. It was not obtained by laboratory testing, and it is used here for a single purpose, which is to construct the best surrogate such a corpus supports so that the confrontation with the field data in Section 3.6 constitutes a fair test of the synthetic-training approach rather than a test of a deliberately weak model. Section 4.3 establishes the provenance of the corpus.

2.8 Screening of the synthetic corpus

Every synthetic record was tested against four admissibility criteria before any model was fitted. Each criterion derives from a definitional constraint rather than from a distributional assumption, so a violation constitutes a data fault and not an outlier. Table 2 states the criteria and the outcome.

Check	Criterion	Basis	Records	Action
Negative plasticity index	$PI < 0$	$PI = LL - PL$ with $PL \geq 0$	2 of 800	Removed
Plasticity index not below liquid limit	$PI \geq LL$	$PI = LL - PL$, so $PI < LL$ always	1 of 800	Removed

Missing categorical field	soil class or grading absent	Encoding would create a merge-artefact indicator	0 of 800	None required
Zero-variance column	one distinct value in corpus	Carries no information	1 of 800	Column dropped (compactionindex)

Table 2. Physical consistency screening applied to the synthetic corpus of 800 records.

Screening removed three records and dropped one invariant column, leaving 797 records and 15 usable input fields, a loss of 0.4% of the corpus. No records had missing categorical fields, so no missingness indicator was created. This is worth recording because a missingness indicator arising from a table merge is a known failure mode in which the strongest apparent predictor in a model becomes the identity of the source file from which a row originated.

Figure 3 characterises the screened corpus as a modelling basis and relates its response scales to measured values.

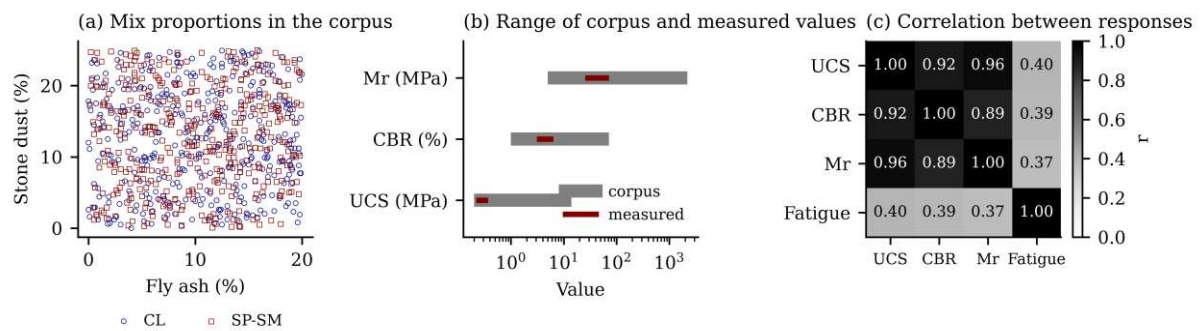


Figure 3. Composition, value ranges and response correlations of the synthetic corpus.

3. Surrogate modelling and results

3.1 Feature construction

Nine measured inputs were retained and eight derived features were added. Each derived feature encodes a specific mechanical hypothesis rather than an arbitrary algebraic combination, and Table 3 states that hypothesis alongside the definition so that the feature set can be audited or transferred to another dataset.

Feature	Definition	Mechanical reasoning
totalreplacementpet	fly ash + rubber + stone dust	Total mass fraction displaced from the host soil; sets how far the mix departs from the parent material.

$\text{rubber}_{\text{tostone}}$	$\text{rubber} / (\text{stone dust} + 1)$	Balance between the elastomeric and angular-granular components competing for the same void space.
$\text{ash}_{\text{total}}$	$\text{fly ash} / (\text{total replacement} + 1)$	Share of the replacement that is pozzolanically reactive rather than inert.
$\text{granular}_{\text{fraction}}$	$(\text{stone dust} + \text{fly ash}) / (\text{total replacement} + 1)$	Share of the replacement contributing load-bearing skeleton rather than compliance.
$\text{plasticity}_{\text{ratio}}$	PI / LL	Normalised plasticity; separates clay mineralogy from bulk water content.
$\text{log}_{\text{curing}}$	$\ln(\text{curing days})$	Pozzolan strength gain is approximately logarithmic in time, not linear.
$\text{compaction}_{\text{energyproxy}}$	$\text{MDD} / (\text{w/c ratio})$	Higher dry density at lower water content indicates more effective compaction.
$\text{phyS}_{\text{modulusprior}}$	Eq. (2)	Empirical pavement-style stiffness expression supplied as a soft physics prior.

Table 3. Derived features, their definitions and the mechanical reasoning each encodes.

The modulus prior in Table 3 is a soft physics feature, an empirical pavement-style expression

$$M_{R,\text{prior}} = 300 (\text{MDD})^{2.5} (t_c)^{0.12} (1 - 0.030 R) (1 + 0.012 F) (1 - 0.006 \text{PI}) \quad (2)$$

in which MDD is the maximum dry density in Mg/m^3 , t_c the curing duration in days, R the rubber percentage, F the fly ash percentage and PI the plasticity index. It is supplied to the learners as one input among many and is not imposed as a constraint.

One implementation detail requires emphasis because it produced a catastrophic and initially obscure failure. Ratio features must use a denominator floor of order unity rather than a small constant. With a floor of 10^{-6} , mixes containing no stone dust generate a rubber to stone dust ratio of order 10^7 . Tree ensembles are invariant to monotonic rescaling and absorb this without visible harm, whereas any standardised linear or neural learner is destroyed by it, in this case driving the coefficient of determination below -10^3 and, through the stacking layer, degrading the ensemble as well. The symptom is a hybrid model that performs far worse than its own components. Raising the floor to unity resolved the failure completely. The relevance to the wider literature is that a hybrid model reported as underperforming its components should be examined for unbounded engineered features before its architecture is revised.

3.2 Base learners and combination

Five base learners spanning three model families were used, so that the ensemble is not a collection of near-identical tree variants. Table 4 gives the configuration of each.

Component	Family	Configuration
-----------	--------	---------------

Ridge	Penalised linear	Generalised cross-validated penalty over 25 values from 10^{-3} to 10^3
Random forest	Bagged trees	300 trees, minimum leaf size 2
XGBoost	Gradient-boosted trees	450 trees, depth 4, learning rate 0.06, subsample and column sample 0.85, L2 penalty 1.5
LightGBM	Leaf-wise boosted trees	450 trees, 24 leaves, learning rate 0.06, matched sampling, L2 penalty 1.5
MLP	Feed-forward network	192 and 96 hidden units, ReLU, AdamW, penalty 10^{-3} , early stopping after 20 stale epochs
Ridge meta-learner	Stacking combiner	Fitted only on inner-fold out-of-fold base predictions (Section 3.2)
GP residual corrector	Gaussian process	Matern kernel, smoothness 3/2, with ARD and white noise, 350 to 400 subsampled residuals

Table 4. Base learner configurations.

Predictions were combined by a ridge meta-learner in the stacked regression framework. Fatigue life was modelled on a logarithmic scale because of its pronounced right skew, with the back-transformation clipped in log space to prevent the unconstrained linear meta-learner from emitting values that overflow on exponentiation.

A Gaussian process residual corrector with a Matern kernel of smoothness 3/2 and automatic relevance determination was fitted to the residuals of the stack, supplying a predictive standard deviation alongside each point prediction. All performance figures derive from a nested five-fold cross-validation in which the outer folds report while the inner folds generate the out-of-fold base predictions on which the meta-learner is trained, so that the stacker never observes its own training targets. This protocol matters because a stacker trained on in-fold base predictions reports optimistic accuracy that does not survive deployment.

3.3 Surrogate accuracy on the synthetic corpus

All seven predictors were evaluated under the single nested protocol of Section 3.2, so their scores are directly comparable. Table 5 reports the outcome.

Model	UCS R^2	UCS RMSE	CBR (%) R^2	CBR (%) RMSE	M_R (MPa) R^2	M_R (MPa) RMSE	Fatigue (kcyc) R^2	Fatigue (kcyc) RMSE
Ridge	0.932	0.649	0.941	3.479	0.892	118.198	0.513	29.565
RandomForest	0.890	0.827	0.921	4.020	0.845	141.565	0.451	31.419

XGBoost	0.913	0.733	0.922	4.004	0.863	133.159	0.447	31.533
LightGBM	0.915	0.728	0.924	3.945	0.861	133.959	0.436	31.833
MLP	0.924	0.685	0.923	3.989	0.871	128.995	0.484	30.457
Stacked	0.934	0.639	0.940	3.513	0.892	118.470	0.521	29.348
Stacked + GP	0.934	0.641	0.939	3.541	0.888	120.603	0.526	29.195

Table 5. Nested five-fold cross-validated performance of every learner on the screened synthetic corpus of 797 records. All learners share an identical engineered feature matrix and resampling partition.

The stacked ensemble with Gaussian process correction reached R^2 of 0.934 for compressive strength, 0.939 for bearing ratio, 0.888 for resilient modulus and 0.526 for fatigue life. The first three are strong, while fatigue life is predicted only moderately, which is consistent with its weak correlation to the strength family in Figure 3c and with the greater intrinsic scatter characteristic of cyclic testing.

Ridge regression on the same engineered features attained R^2 of 0.932, 0.941, 0.892 and 0.513. It falls below the full hybrid on two responses, by 0.0018 on compressive strength and by 0.012 on fatigue life, and rises above it on the other two, by 0.0021 on bearing ratio and 0.0044 on resilient modulus. Only the fatigue gap exceeds 0.005; on the remaining three responses the two models are indistinguishable at the third decimal place. Figure 4 presents that comparison across all learners and, in its second panel, the ablation discussed below.

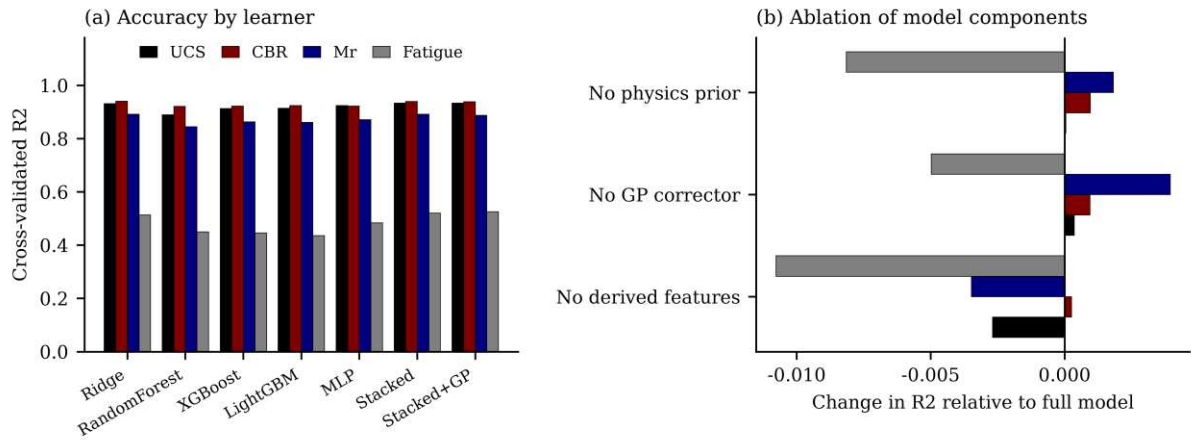


Figure 4. Cross-validated accuracy by learner, and change in accuracy when each model component is removed.

This result is reported rather than suppressed. The engineered features, and not the ensemble architecture, carry the predictive performance on this corpus. A hybrid architecture is

demonstrated to be necessary only when it is compared against a linear model supplied with the same features.

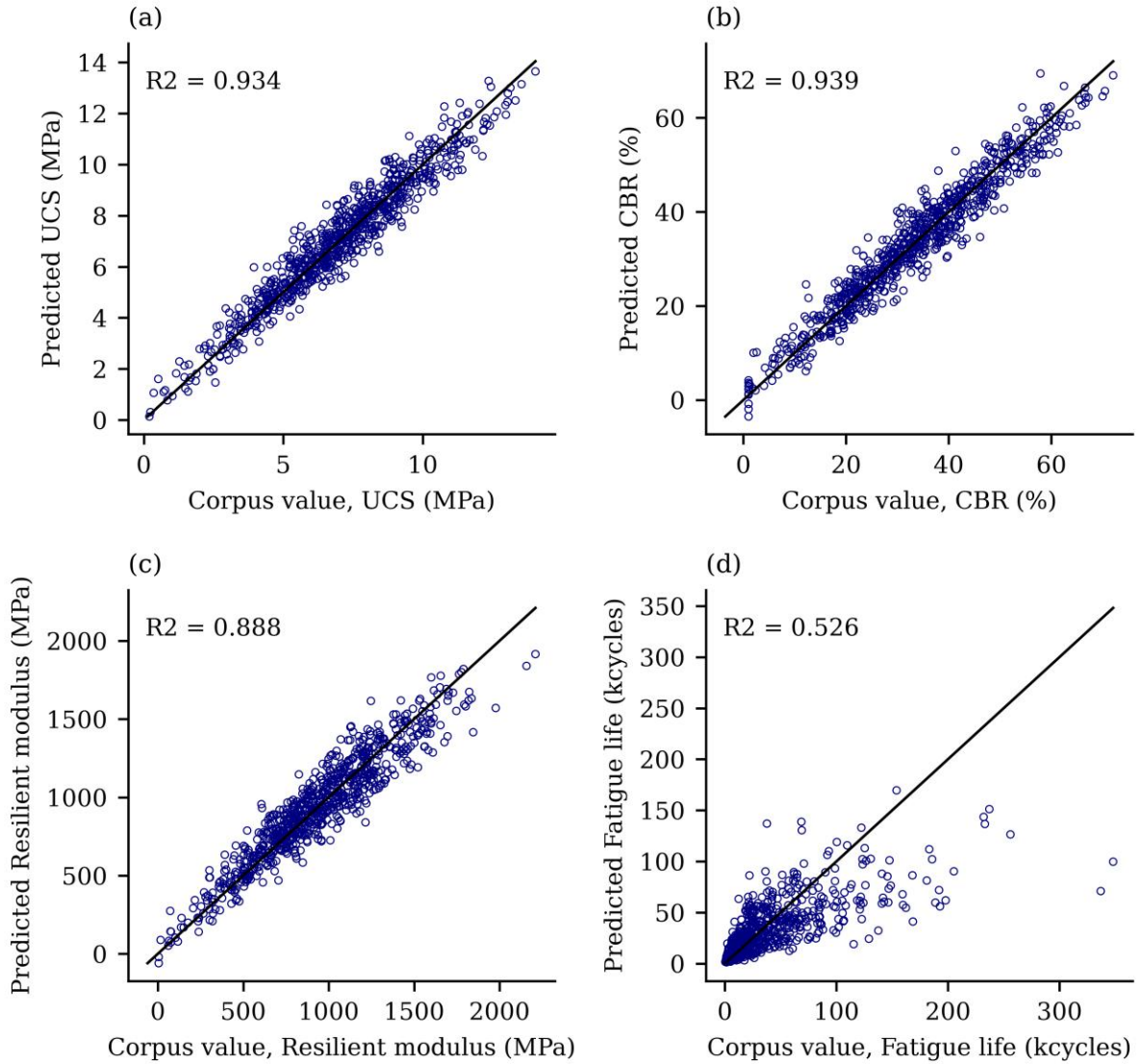


Figure 5. Out-of-fold predictions against corpus values for the stacked ensemble with Gaussian process correction.

Table 6 gives the ablation values underlying Figure 4b.

	CBR (%)	M_R (MPa)	UCS	Fatigue (kcyc)
A full model	0.939	0.888	0.934	0.526
B no derived features	0.939	0.884	0.931	0.515
C no GP corrector	0.940	0.892	0.934	0.521

D no physics prior	0.940	0.889	0.934	0.517
--------------------	-------	-------	-------	-------

Table 6. Ablation of the stacked ensemble with Gaussian process correction. Values are nested cross-validated R^2 and the full model is configuration A.

Removing the derived features is the most costly single change, while removing the Gaussian process corrector or the physics prior costs almost nothing in point accuracy. The corrector is retained not for accuracy but for the prediction intervals it supplies, and those intervals must be shown to be trustworthy before they are used.

A model that emits a standard deviation has not thereby established that the standard deviation is meaningful. The relevant test is whether an interval stated to contain 90% of outcomes does so. Table 7 reports coverage at four nominal levels.

index	50%	80%	90%	95%
CBR (%)	0.484	0.804	0.896	0.939
M_R (MPa)	0.493	0.783	0.880	0.937
UCS	0.506	0.797	0.877	0.941
Fatigue (kcyc)	0.484	0.795	0.905	0.959

Table 7. Prediction interval coverage probability at four nominal levels. Well-calibrated intervals give coverage equal to the nominal level.

Coverage tracks the nominal level closely, at 0.877 to 0.905 for the 90% interval and 0.937 to 0.959 for the 95% interval. Figure 6 presents the reliability diagram. This result justifies retaining the corrector despite its neutral effect on point accuracy, and it is what makes the reliability-adjusted mix search of Section 3.7 meaningful rather than decorative.

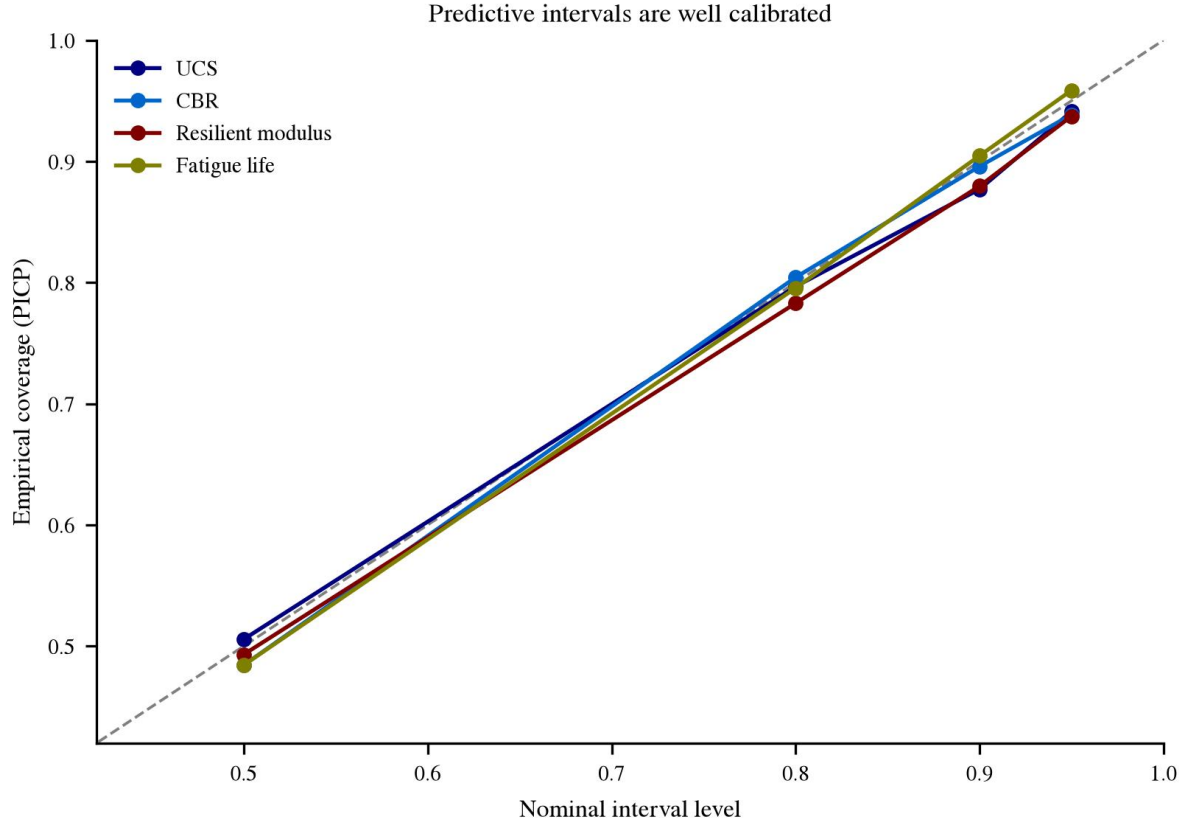


Figure 6. Empirical against nominal interval coverage. Points on the dashed diagonal indicate calibrated intervals.

Shapley attribution on the tree component orders the contributors in a manner consistent with the underlying mechanics. Table 8 reports the mean absolute Shapley contribution of each influential feature to each of the four responses.

Feature	UCS, mean abs. SHAP (MPa)	CBR, mean abs. SHAP (%)	M _R , mean abs. SHAP (MPa)	Fatigue, mean abs. SHAP (kcycles)
soil _{typeCL} (high-plasticity clay)	0.646	7.699	76.7	0.192
granular _{fraction}	0.756	5.488	154.6	not in top 8
phy _{Smodulusprior}	0.828	1.143	61.8	0.051
liquid _{limit}	0.244	not in top 8	35.7	0.437
total _{replacementpct}	0.304	0.755	58.6	0.279
fly _{ashpct}	0.248	not in top 8	25.9	0.118
stone _{dustpct}	0.106	1.422	24.1	not in top 8

rubber _{pct}	not in top 8	0.455	not in top 8	0.152
mdd _{Mgm3}	0.174	not in top 8	16.8	not in top 8
rubber _{lostone}	not in top 8	1.519	not in top 8	not in top 8
plasticity _{index}	not in top 8	not in top 8	not in top 8	0.044
plasticity _{ratio}	not in top 8	not in top 8	not in top 8	0.046
soil _{typeSP-SM} (silty sand)	not in top 8	0.732	not in top 8	not in top 8

Table 8. Mean absolute Shapley contribution of each feature to each response, expressed in the units of that response. A value of 155 in the resilient modulus column means that varying that feature moves the predicted modulus by 155 MPa on average. Values are therefore comparable down a column but not across columns, because the four responses have different units and different magnitudes. Rows list the union of the eight most influential features for each response, ordered by importance averaged over the four responses after scaling each column to its own maximum, so that features influencing several responses appear first. An entry of "not in top 8" indicates that the feature falls outside the eight most influential for that response.

Three patterns are visible. The granular fraction ranks first or second for all three strength-family responses, first for resilient modulus and second for compressive strength and bearing ratio, which is the expected result if strength is governed by the proportion of the replacement that contributes a load-bearing skeleton. The base soil class appears among the eight most influential features for all four responses and ranks first for bearing ratio, confirming that the corpus does not treat the two soil types as interchangeable. Fatigue life is governed instead by the index properties of the host soil, with the liquid limit ranking first and mix composition contributing less; this is consistent with the weak correlation between fatigue life and the strength family in Figure 3c, and it explains why fatigue life is the one response on which the ensemble outperforms the linear baseline. The physics prior of Eq. (2) ranks first for compressive strength and third for resilient modulus but only fifth for bearing ratio, so its usefulness is confined to the two responses it was constructed to approximate.

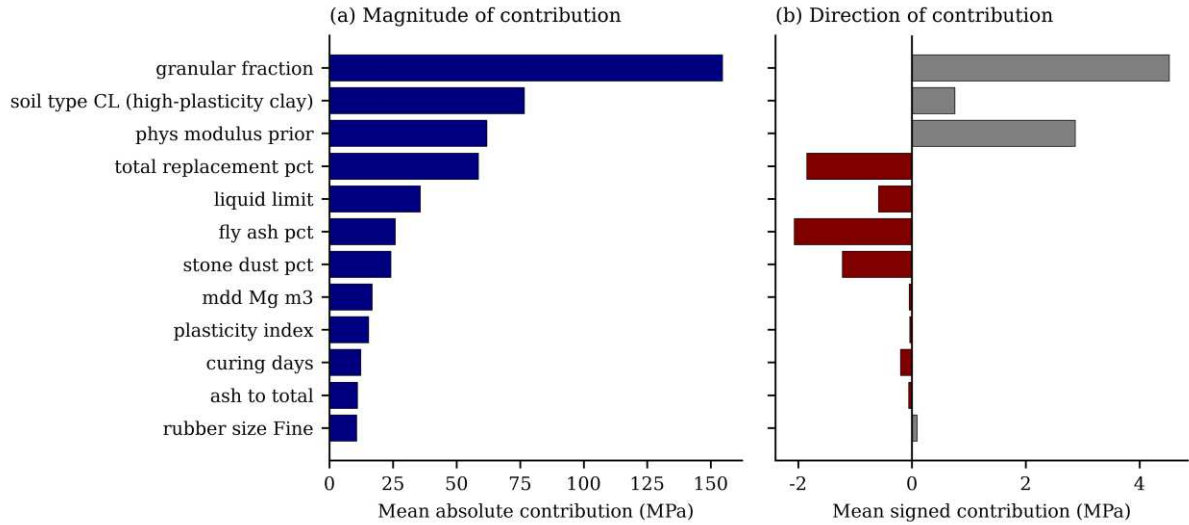


Figure 7. Magnitude and mean direction of feature contributions to predicted resilient modulus.

3.4 Field results

Table 9 reports both measured quantities for all six field sections, as recorded and as recomputed under the protocol of Section 2.6.

Mix	Rubber (%)	CBR reported (%)	CBR recomputed (%)	k reported (kN/cm ³)	k recomputed (kN/cm ³)	M _R via IRC:37 (MPa)
Conventional	0	5.76	5.68	6.09	5.99	53.49
Mix A	1	6.09	6.01	6.34	5.92	55.45
Mix B	2	6.30	6.30	6.71	6.71	57.16
Mix C	3	6.23	6.15	6.45	6.45	56.26
Mix D	4	5.80	5.80	6.18	6.07	54.19
Mix E	5	5.60	5.39	5.83	5.90	51.75

Table 9. Field results for the six sections. The bearing ratio is governed by the 2.5 mm penetration in every case. The modulus of subgrade reaction is computed from Eq. (1). Recomputed values are independent back-calculations from the tabulated load-penetration and load-settlement curves. Resilient modulus is obtained from the recomputed bearing ratio through the IRC:37 correlation and is therefore not an independent measurement; see Section 4.4.

Figure 8 presents the underlying response curves and the resulting trend with rubber content.

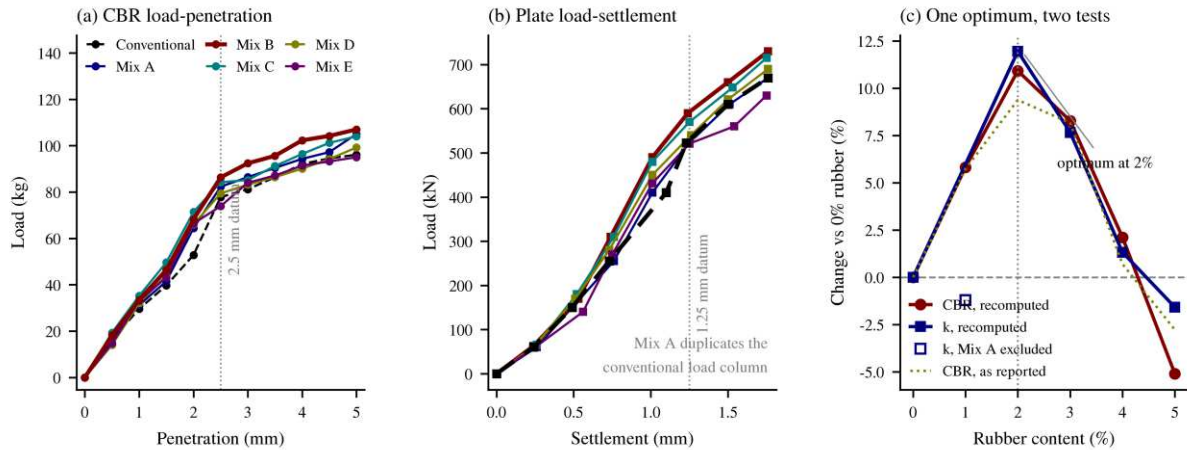


Figure 8. Field measurements. (a) Bearing ratio load-penetration curves for the six sections, with the 2.5 mm datum marked. Mix B is drawn heavier. (b) Plate load-settlement curves with the 1.25 mm datum marked. The conventional section is drawn on top in heavy dashes to show that the Mix A load column duplicates it at every increment, which is why the Mix A plate load result is excluded (Section 4.4). (c) Both measurements expressed as a change relative to the unamended base mix, on a single axis so that relative magnitudes are directly comparable. Bearing ratio and modulus of subgrade reaction peak together at 2% rubber. The originally reported bearing ratio series is overlaid as a dotted line and gives the same optimum at the same position with magnitudes approximately one sixth smaller.

Bearing ratio rises from 5.68% at 0% rubber to a maximum of 6.30% at 2%, an increase of 10.9%, and then falls monotonically to 5.39% at 5%, which is 5.1% below the unamended base mix. The modulus of subgrade reaction follows the same path, rising by 11.9% to its maximum at 2% and declining thereafter. The optimum is interior, and its position is identical in two measurements that share no apparatus, no specimen preparation and no governing standard.

These figures are recomputed from the tabulated curves under the protocol of Section 2.6 rather than taken from the originally reported values. Recomputation reproduces the reported bearing ratio exactly for only two of the six sections, Mix B at 6.30% and Mix D at 5.80%. For the other four it does not, giving 5.68% against 5.76% reported for the conventional section, 6.01% against 6.09% for Mix A, 6.15% against 6.23% for Mix C, and 5.39% against 5.60% for Mix E. Because the conventional section is the reference against which every change is normalised, the recomputed series is used throughout so that all six sections rest on a single consistent derivation. Substituting the reported values would give an increase of 9.4% and a decline of 2.8%, that is, the same optimum at the same position with magnitudes approximately one sixth smaller. Figure 8c plots both series.

3.5 Magnitude of the effect relative to test repeatability

The screening programme of Section 2.2 established coefficients of variation between 2.93 and 7.54%, with the adopted base mix at 2.97%. The observed increases of 10.9% in bearing ratio and 11.9% in stiffness therefore exceed the demonstrated repeatability of the testing procedures, though not by a large factor. With a single test of each type per section, the position of the optimum is more strongly supported than its magnitude. No formal confidence interval can be attached to either without replication, and Section 4.5 states the replication that would be required. Rubber content was varied in integer steps, so the optimum is located to within plus or minus 1%.

3.6 The surrogate confronted with the field data

The surrogate of Sections 3.1 to 3.3 was evaluated at the field mix design, namely 20% stone dust, 20% fly ash, the measured index properties of the host soil, 28 days curing and fine rubber grading, across 0 to 5% rubber in steps of 0.25%. Figure 9 places the two side by side.

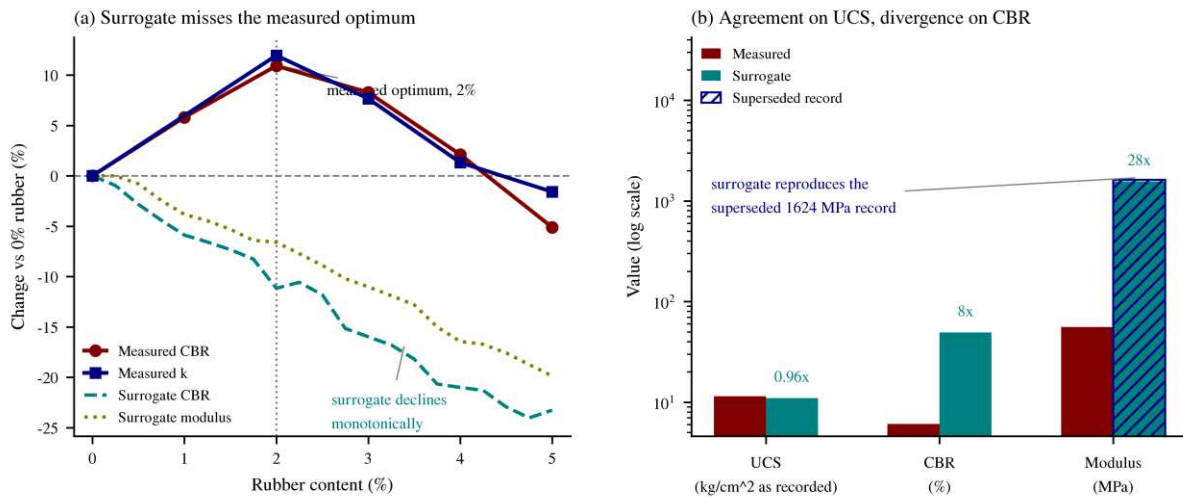


Figure 9. Surrogate against experiment. (a) Measured bearing ratio and modulus of subgrade reaction against the surrogate predictions of bearing ratio and resilient modulus, all normalised to their own value at 0% rubber. The measurements peak at 2% while the surrogate declines monotonically and contains no interior optimum. (b) Absolute values at the field mix design. The surrogate reproduces the screening compressive strength and the previously reported resilient modulus almost exactly while overpredicting bearing ratio by a factor of eight.

The surrogate does not reproduce the measured behaviour. Its predicted optimum lies at 0% rubber for bearing ratio, compressive strength and resilient modulus in the high plasticity clay class, and at 0.25% in the silty sand class. In every case the optimum falls at or adjacent to the boundary of the search interval, which is the signature of a monotonic response rather than an interior optimum. Across 0 to 5% rubber the predicted bearing ratio falls by 24% and the resilient

modulus by 20%, monotonically. The single feature of this response surface on which a designer would act is absent from the model.

This outcome does not reflect a failure of model fitting. The same model attains R^2 of 0.934, 0.939 and 0.888 on the three strength-family responses under nested cross-validation, carries calibrated prediction intervals, and survives ablation with no single component proving load-bearing. By every metric computable from the corpus it is a satisfactory model. The deficiency lies in the corpus.

Panel 9b indicates the reason. At the field mix design the surrogate predicts a compressive strength of 10.97 in the units of the corpus against the screening value of 11.44 in the same units, and a resilient modulus of 1561 MPa against the 1624 MPa reported in the authors' earlier records, an agreement within 4% in both cases. It nonetheless predicts a bearing ratio of 49.2% against a measured 6.09%, an overprediction by a factor of 8.1. A corpus that independently modelled this material system would not reproduce two reported quantities to within 4% while departing from the third by a factor of eight. The pattern indicates that the corpus was generated from those reported values, and therefore reproduces both their scale and their errors, while the bearing ratio was generated on a different basis.

3.7 Constrained mix search

For completeness, the mix optimisation that the surrogate supports is reported, subject to the finding of Section 3.6 that its rubber dimension is unreliable. A grid of 55,566 candidate mixes was scored and non-dominated sorting was applied to resilient modulus against fatigue life.

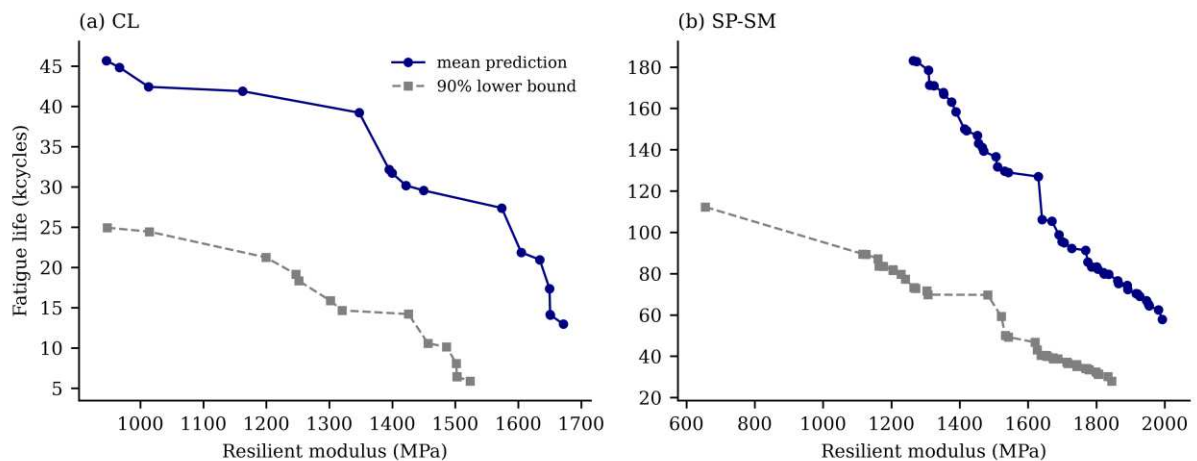


Figure 10. Non-dominated sets of stiffness against fatigue life for both soil classes.

Non-dominated sorting returns 47 mixes for silty sand and 16 for high plasticity clay on the mean front, reducing to 39 and 14 respectively when the fronts are computed on lower confidence bounds. Within the corpus, higher rubber content and coarser grading favour fatigue life while lower rubber content and finer grading favour stiffness. This trade-off is directionally consistent with the role of rubber as a compliant phase, although the corpus places the stiffness optimum at zero rubber where measurement places it at 2%. Table 10 extracts representative points.

soil	objective	fly _{ashp} ct	rubber _p ct	rubber _{si} ze	stone _{dustp} ct	MR _{MPa}	MR _{lcb}	fatigue _{kc} yc	fatigue _l cb	UCS _M Pa	CBR _p ct
CL (high-plasticity clay)	max stiffness	20.00	0.00	Fine	25.00	1671.70	1523.60	13.00	5.90	11.50	51.90
CL (high-plasticity clay)	balance (knee)	19.00	4.50	Fine	23.75	1347.50	1199.40	39.20	21.20	9.42	39.20
CL (high-plasticity clay)	max fatigue	19.00	10.00	Coarse	25.00	945.80	798.00	45.70	21.90	6.17	25.20
SP-SM (silty sand)	max stiffness	20.00	0.00	Fine	25.00	1993.20	1845.00	57.90	28.00	14.17	72.10
SP-SM (silty sand)	balance (knee)	16.00	3.00	Fine	21.25	1630.20	1482.10	127.00	69.70	11.76	59.20
SP-SM (silty sand)	max fatigue	19.00	10.00	Coarse	25.00	1263.90	1115.80	183.10	89.50	8.98	43.80

Table 10. Representative mixes from the Pareto fronts: the two extremities and the knee, defined as the candidate maximising the sum of the two min-max normalised objectives, together with their reliability-adjusted counterparts. These are surrogate predictions on a synthetic corpus and are not design recommendations.

4. Discussion

4.1 Mechanism of the interior optimum

The measured response has the form predicted by the constituent behaviour set out in Section 1. At 1 to 2% by dry weight, rubber granulate occupies inter-particle voids and introduces a compliant phase that redistributes contact stress without interrupting the granular skeleton, and both bearing ratio and stiffness improve. By 4 to 5% the elastomeric phase is continuous enough to carry load paths itself, soil grain contact is progressively lost, and both quantities fall below those of the unamended mix. For this soil, this rubber grading and this compaction procedure, the transition lies between 2 and 3%.

Two observations support this interpretation rather than an alternative in which the improvement arises from improved compaction. First, the improvement and the subsequent decline are both present in the plate load test, which loads a large volume of the compacted layer in situ, as well as in the bearing ratio test. A compaction artefact would not be expected to reverse sign at higher rubber content. Second, the decline below the unamended value at 5% is inconsistent with a purely densification-driven mechanism, under which additional fines would not reduce capacity below the baseline.

4.2 Comparison with published work

Direct quantitative comparison of the optimum rubber content across studies is not attempted here, and the reason is itself a finding of the review literature. Tasalloti et al. [41] observe that published soil-rubber results have not been compiled in a form permitting direct comparison, and identify rubber content together with the ratio of rubber to soil median particle size as the governing variables. Reported optima are therefore conditional on host soil gradation, rubber particle size distribution, mixing method and test type, none of which are standardised across the literature. Studies on granulated rubber in fine-grained soils [15, 17] and on tyre-derived aggregate in expansive clays [27, 30] report improvements at low rubber content followed by degradation at higher content, which is qualitatively consistent with the present result, but the dosages are not transferable between soils.

What the present study contributes to that literature is not a universal optimum but a demonstration that the optimum is interior and locatable with a small field programme, together with the observation that its position is reproduced by two independent tests. The practical implication is that the optimum must be determined for the specific soil and rubber grading in

use, which is an argument for short targeted field programmes rather than for extrapolation from published dosages.

4.3 Provenance of a synthetic corpus determines what a surrogate can learn

Section 3.6 established that the corpus reproduces the authors' reported compressive strength and resilient modulus to within 4% while overpredicting bearing ratio by a factor of eight. Two consequences follow, and both generalise beyond this study.

The first is that a corpus generated from reported values inherits the errors in those values. The authors' earlier records report compressive strengths of 4.80 to 12.50 in units labelled MPa. Unconfined compressive strength of a natural fine-grained subgrade is normally 0.05 to 0.5 MPa, and stabilised subgrade rarely exceeds 2 MPa, so those figures are not credible in the units stated. Interpreted as kg/cm^2 they become 0.471 to 1.226 MPa, which is the expected range and is consistent with the bearing ratio loads having been recorded in kg. An independently measured crumb rubber series gives 0.22 to 0.37 MPa, a conflict of a factor of 30 with the labelled values and a close match to the corrected ones. The corrected values are used throughout this paper. The corpus, however, was generated on the uncorrected scale, which is why its compressive strength appears to agree with the authors' records while disagreeing with physical expectation, and why in Figure 3b the corpus and the measured series reconcile for compressive strength only after the unit correction is applied.

The authors' records also contain two irreconcilable resilient moduli: 1577 to 1624 MPa for the stone dust and fly ash combinations, and 53 to 57 MPa for the rubber sections. Obtaining 1624 MPa from the IRC:37 correlation would require a bearing ratio of 1176%. For a stabilised subgrade, 53 to 57 MPa is the plausible range while 1624 MPa is characteristic of a cement-treated base. The corpus follows the implausible value, predicting 1561 MPa at the field mix design.

The second consequence is that no internal validation metric detects any of this. Cross-validated R^2 , root-mean-square error and interval coverage all measure the agreement between the surrogate and the generator. They are silent on the agreement between the generator and the material. A surrogate can be simultaneously well-fitted, well-calibrated, robust to ablation and physically wrong, and the present study is an instance.

4.4 Findings of the verification protocol

The protocol of Section 2.6 returned four findings.

The plate load method was confirmed and its unit corrected. Recomputing k from Eq. (1) against the tabulated load-settlement curves reproduces Mix B as 6.708 against 6.71 reported and Mix C as 6.451 against 6.45, an agreement to three decimal places that confirms both the 1.25 mm datum and the 300 mm plate diameter. The dimensions of the quantity are pressure divided by length, so the unit is kN/cm^3 ; the kN/cm^2 appearing in the records is a labelling error carried through every table.

One section's plate load record is not supported by its own data. The Mix A load column is identical to that of the conventional section at all eight increments, while its settlements differ slightly and are marginally larger at the datum. The tabulated data therefore give k of $5.92 \text{ kN}/\text{cm}^3$, below the conventional section, whereas 6.34 is reported, above it. The reported value cannot be derived from the tabulated curve. The Mix A plate load result is accordingly excluded from the quantitative claims of Section 3.4, which rest on the bearing ratio series for all six sections and on the plate load series for the remaining five. The exclusion does not alter the conclusion, since the optimum remains at 2% in both measurements.

The reported resilient modulus is a correlation rather than a measurement. Substituting the authors' reported bearing ratios into the IRC:37 expression $M_R = 17.6 (\text{CBR})^{0.64}$ reproduces the tabulated resilient moduli to within 0.004 MPa at every section: 53.974 against 53.97, 55.933 against 55.93, 57.160 against 57.16, 56.753 against 56.75, 54.213 against 54.21, and 53.009 against 53.01. The agreement is exact to the printed precision, so these values were computed and not measured. Resilient modulus is consequently a strictly increasing function of bearing ratio, with a correlation of 0.99999, and carries no information independent of it. Any statement that bearing ratio, resilient modulus and plate load stiffness independently confirm a conclusion therefore reduces to two independent confirmations rather than three, and two is what is claimed throughout this paper.

A bearing ratio column was misaligned, and three further values do not reproduce. In the Mix C load-penetration table the load column was displaced by one penetration step, recording a non-zero load at zero penetration and leaving the deepest reading blank. Correcting the alignment and recomputing gives 6.15% against the 6.23% originally reported. Applying the same recomputation to the remaining five sections reproduces the reported value exactly for Mix B at 6.30% and Mix D at 5.80%, but not for the conventional section at 5.68% against 5.76%, Mix A at 6.01% against 6.09%, or Mix E at 5.39% against 5.60%. Three of the four discrepancies, namely the conventional section, Mix A and Mix C, are identical at 0.08 percentage points, which suggests a systematic rounding or interpolation convention rather than independent transcription

errors. The discrepancy of 0.21 for Mix E is larger and cannot be accounted for. Because the conventional section is the reference against which every reported change is normalised, an error of 0.08 there propagates into all five comparisons.

These findings are reported rather than silently corrected, because the checks that exposed them are inexpensive, general, and on this evidence necessary. Recomputing a reported derived quantity from its own tabulated inputs, and confirming that a reported unit is dimensionally consistent with its defining formula, would have identified all four.

4.5 Limitations

The field programme provides a single test of each type per section, so the position of the optimum is more strongly supported than its magnitude. The screening coefficients of variation of 2.93 to 7.54% indicate the repeatability of the procedures, and the observed increases of 10.9 and 11.9% exceed that band, but a replicated programme of three sections per rubber content, or three tests per section, would be required to attach a confidence interval to the optimum and to resolve whether the true peak lies at 2% or between 2 and 3%. Rubber contents between the integer values were not tested, so the optimum is located to within plus or minus 1%.

Resilient modulus was not measured directly for the rubber sections, and the tabulated values are a correlation of bearing ratio as established in Section 4.4. Repeated-load triaxial testing would make stiffness a third independent measurement rather than a restatement of the first. Compressive strength was measured for the screening programme but not for the rubber sections, so the interior optimum is established for bearing capacity and subgrade stiffness and not for compressive strength. Durability was not assessed. Rubber-amended soils raise questions concerning long-term creep under sustained load and concerning leaching, neither of which this programme addresses, and both of which would need to be resolved before field adoption.

On the modelling side, the surrogate is trained on a synthetic corpus and its numerical predictions should not be used for design. Its role in this study is as the object of a validation test, and the result of that test is negative. Retraining it on the field data was not attempted, because six sections cannot support a model of a six-dimensional mix space. The constructive path is the reverse one, of using measured optima as acceptance criteria for a candidate corpus.

4.6 Implications for practice

For subgrade stabilisation with this combination of wastes, the practical finding is that crumb rubber should be dosed low. At 2% by dry weight of soil, on a base of 20% stone dust and 20% fly ash, both bearing capacity and subgrade stiffness improve by approximately 10% relative to

the unamended base mix. At 5% both fall below it, so rubber is not a material to be added in quantity for its own sake, even where waste diversion is an objective.

For surrogate-based mix design, the implication is methodological. Validation of a surrogate trained on generated data should be reported against measured features of the response surface, such as the position of an optimum, the sign of a gradient or a known limiting behaviour, and not solely against held-out records from the same generator. A short targeted physical programme, of the order of the six sections reported here, is sufficient to test whether a surrogate contains the design-relevant structure. Where a corpus is generated from published or previously reported values, those values should first be checked for dimensional consistency and for derived-quantity circularity, since both propagate into the corpus without trace.

4.7 Waste diversion and the sustainable development goals

The three amendments used here are all residues: fly ash from coal combustion, stone dust as the fine fraction of aggregate crushing, and crumb rubber from end-of-life tyres. At the adopted proportions the stabilised subgrade contains 200 kg of fly ash, 200 kg of stone dust and 20 kg of crumb rubber per tonne, so 420 kg of every tonne placed is material that would otherwise require disposal, and a corresponding mass of in-situ soil is not excavated or imported. This is directly relevant to United Nations Sustainable Development Goal 12, and specifically to target 12.5 on substantially reducing waste generation through reuse and recycling, and it contributes to Goal 9 on resilient infrastructure and Goal 11 on reducing the adverse environmental impact of cities, under which the accumulation of coal combustion residues and quarry fines is an established burden in economies that are urbanising rapidly.

The dosage ceiling reported in this paper qualifies that contribution in a way that is not usually made explicit. Waste diversion and engineering performance are not co-monotonic. Raising the rubber fraction from 2% to 5% diverts a further 30 kg of tyre rubber per tonne, but converts a 10.9% improvement in bearing ratio into a 5.1% deficit relative to the rubber-free base mix, so the additional diversion is obtained at the cost of the engineering benefit that justified the intervention. A specification that treats waste content as a quantity to be maximised will therefore produce a subgrade inferior to one containing no rubber at all, while appearing more sustainable on a mass-diverted basis. Sustainability claims for waste-amended subgrade require a dosage bound established by measurement.

These statements are claims about material diversion by mass and about measured mechanical performance. No life-cycle assessment was undertaken in this study, and the transport distances,

processing energy and any leaching behaviour of the three amendments are outside its scope; net environmental benefit cannot be asserted on the present evidence and would require a separate assessment.

5. Conclusions

1. For a lateritic subgrade stabilised with 20% stone dust and 20% fly ash, crumb rubber produces an interior optimum at 2% by dry weight. Bearing ratio rises by 10.9% and the modulus of subgrade reaction by 11.9% relative to the unamended base mix, and both decline at higher dosage, with bearing ratio falling 5.1% below the base mix at 5%. The optimum is located identically by two mechanically independent tests, and its position is unchanged whether the field values are recomputed from the tabulated curves or taken as originally reported.
2. The laboratory screening programme fixes the base mix. Compressive strength rises from 0.471 MPa untreated to 1.122 MPa at 20% stone dust with 20% fly ash, at a coefficient of variation of 2.97%, and the four candidate combinations differ by only 5.2%, so the selection among them is not constrained by strength.
3. A stacked ensemble with a Gaussian process residual corrector, trained on an 800-record synthetic corpus for the same material system, reaches nested cross-validated R^2 of 0.934, 0.939, 0.888 and 0.526 with calibrated prediction intervals at 0.877 to 0.905 coverage for the 90% interval, yet predicts monotonic decline in every response across 0 to 5% rubber and contains no interior optimum. At the field mix design it overpredicts bearing ratio by a factor of 8.1.
4. The failure is attributable to the provenance of the corpus and not to the model architecture. The corpus reproduces previously reported compressive strength and resilient modulus to within 4% while departing from bearing ratio by a factor of eight, which identifies it as generated from those reported values, including a unit mislabelling and a resilient modulus that a dimensional check shows to be in error by a factor of 27.
5. No internal validation metric detects this. Cross-validated accuracy, interval calibration and ablation robustness all measure agreement between the surrogate and the generator, and none measures agreement between the generator and the material.
6. Regularised linear regression on the same engineered features differs from the full ensemble by less than 0.005 in R^2 on three of four responses, exceeding it on bearing ratio at 0.941 against 0.939 and on resilient modulus at 0.892 against 0.888, and falling 0.0018 short on compressive strength. Fatigue life is the only response on which the ensemble holds a

material advantage, and there by 0.012. Hybrid architectures should be reported against a matched-feature linear baseline before their complexity is claimed to be necessary.

7. Independent recomputation of reported quantities from their own tabulated inputs identified a misaligned load column, a dimensionally incorrect unit, one section whose reported stiffness cannot be derived from its own curve, and a resilient modulus that is a correlation of bearing ratio rather than a measurement. These checks are inexpensive and should be routine, particularly where reported values will be used to seed a generated corpus.

Declarations

Data availability. The field and laboratory data, the screened synthetic corpus, all result tables and the complete analysis code are provided as supplementary files.

Declaration on data provenance. Every value reported as a measurement in Sections 2.2, 3.4 and 4.4 was obtained by physical testing and is recomputed in this paper from its primary load-penetration, load-settlement or compression record; those primary records are supplied as supplementary data so that each reported quantity can be reproduced independently. The 800-record corpus described in Section 2.7 is synthetically generated and is not presented as measurement. It is the object of the audit reported in Sections 3.1 to 3.3 and 3.6, and no conclusion drawn in this paper about the material behaviour of the stabilised subgrade depends upon it.

Declaration of competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. The work received no specific funding.

CRedit authorship contribution statement. Harishkumar M: Conceptualization, Methodology, Investigation, Formal analysis, Software, Writing - original draft, Visualization. Samundeswari R: Investigation, Validation, Data curation. Godson Caleb J: Investigation, Resources, Data curation. Vincent Sam Jebadurai S: Conceptualization, Supervision, Project administration, Writing - review and editing. Hemalatha G: Supervision, Resources, Writing - review and editing. Mightraj C: Software, Formal analysis, Validation, Visualization.

Declaration of generative AI and AI-assisted technologies in the writing process. During the preparation of this work the authors used Claude and ChatGpt in order to draft and edit the manuscript text and to assist with data analysis and figure generation, as described in Section 3.

After using this tool, the authors reviewed and edited the content as needed, verified all reported values against the source data, and take full responsibility for the content of the publication.

Note on figure preparation. All plotted figures were produced by scripted rendering of the measured and computed data reported in this paper, using standard scientific plotting software.. The photographs in Figure 1 are unmodified field photographs, cropped only to a common aspect ratio for layout. No image content was generated, retouched or otherwise altered.

References

1. Breiman L. Stacked regressions. Machine Learning 1996;24:49-64. <https://doi.org/10.1007/bf00117832>
2. Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 2002;6:182-197. <https://doi.org/10.1109/4235.996017>
3. Rasmussen CE, Williams CKI. Gaussian Processes for Machine Learning. The MIT Press eBooks 2005. <https://doi.org/10.7551/mitpress/3206.001.0001>
4. Horpibulsuk S, Phetchuay C, Chinkulkijniwat A, Cholphatsorn A. Strength development in silty clay stabilized with calcium carbide residue and fly ash. Soils and Foundations 2013;53:477-486. <https://doi.org/10.1016/j.sandf.2013.06.001>
5. Yao Z, Ji X, Sarker PK, Tang J, Ge L, Xia M, et al. A comprehensive review on the applications of coal fly ash. Earth-Science Reviews 2014;141:105-121. <https://doi.org/10.1016/j.earscirev.2014.11.016>
6. Mudgal A. Effect of Lime and Stone Dust in the Geotechnical Properties of Black Cotton Soil. International Journal of Geomate 2014. <https://doi.org/10.21660/2014.14.140402>
7. Correia AAS, Oliveira PJV, Custódio DG. Effect of polypropylene fibres on the compressive and tensile strength of a soft soil, artificially stabilised with binders. Geotextiles and Geomembranes 2014;43:97-106. <https://doi.org/10.1016/j.geotextmem.2014.11.008>
8. Reddy SB, Krishna AM. Recycled Tire Chips Mixed with Sand as Lightweight Backfill Material in Retaining Wall Applications: An Experimental Investigation. International Journal of Geosynthetics and Ground Engineering 2015;1. <https://doi.org/10.1007/s40891-015-0036-0>
9. Mashiri MS, Vinod JS, Sheikh MN, Tsang H. Shear strength and dilatancy behaviour of sand-tyre chip mixtures. Soils and Foundations 2015;55:517-528. <https://doi.org/10.1016/j.sandf.2015.04.004>

10. Shahriari B, Swersky K, Wang Z, Adams RP, Freitas ND. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE* 2015;104:148-175. <https://doi.org/10.1109/jproc.2015.2494218>
11. al-Swaidani AM, Hammoud I, Meziab A. Effect of adding natural pozzolana on geotechnical properties of lime-stabilized clayey soil. *Journal of Rock Mechanics and Geotechnical Engineering* 2016;8:714-725. <https://doi.org/10.1016/j.jrmge.2016.04.002>
12. Sharma AK, Sivapullaiah PV. Ground granulated blast furnace slag amended fly ash as an expansive soil stabilizer. *Soils and Foundations* 2016;56:205-212. <https://doi.org/10.1016/j.sandf.2016.02.004>
13. Chen T, Guestrin C. XGBoost. *nan* 2016. <https://doi.org/10.1145/2939672.2939785>
14. Lundberg S, Lee S. A Unified Approach to Interpreting Model Predictions. *arXiv (Cornell University)* 2017. <https://doi.org/10.48550/arxiv.1705.07874>
15. Anvari SM, Shooshpasha I, Kutanaei SS. Effect of granulated rubber on shear strength of fine-grained sand. *Journal of Rock Mechanics and Geotechnical Engineering* 2017;9:936-944. <https://doi.org/10.1016/j.jrmge.2017.03.008>
16. Firoozi AA, Olgun CG, Firoozi AA, Baghini MS. Fundamentals of soil stabilization. *International Journal of Geo-Engineering* 2017;8. <https://doi.org/10.1186/s40703-017-0064-9>
17. Yadav JS, Tiwari SK. Influence of crumb rubber on the geotechnical properties of clayey soil. *Environment Development and Sustainability* 2017;20:2565-2586. <https://doi.org/10.1007/s10668-017-0005-y>
18. Choobbasti AJ, Kutanaei SS. Microstructure characteristics of cement-stabilized sandy soil using nanosilica. *Journal of Rock Mechanics and Geotechnical Engineering* 2017;9:981-988. <https://doi.org/10.1016/j.jrmge.2017.03.015>
19. Tenpe AR, Patel A. Application of genetic expression programming and artificial neural network for prediction of CBR. *Road Materials and Pavement Design* 2018;21:1183-1200. <https://doi.org/10.1080/14680629.2018.1544924>
20. Ghadir P, Ranjbar N. Clayey soil stabilization using geopolymer and Portland cement. *Construction and Building Materials* 2018;188:361-371. <https://doi.org/10.1016/j.conbuildmat.2018.07.207>
21. Arulrajah A, Yaghoubi M, Disfani MM, Horpibulsuk S, Bo MW, Leong M. Evaluation of fly ash- and slag-based geopolymers for the improvement of a soft marine clay by deep soil mixing. *Soils and Foundations* 2018;58:1358-1370. <https://doi.org/10.1016/j.sandf.2018.07.005>

22. Ghorbani A, Hasanzadehshooiili H. Prediction of UCS and CBR of microsilica-lime stabilized sulfate silty sand using ANN and EPR models; application to the deep soil mixing. *Soils and Foundations* 2018;58:34-49. <https://doi.org/10.1016/j.sandf.2017.11.002>
23. Arshad M. Development of a Correlation between the Resilient Modulus and CBR Value for Granular Blends Containing Natural Aggregates and RAP/RCA Materials. *Advances in Materials Science and Engineering* 2019;2019:1-16. <https://doi.org/10.1155/2019/8238904>
24. Guo Y, Markine VL, Qiang W, Zhang H, Jing G. Effects of crumb rubber size and percentage on degradation reduction of railway ballast. *Construction and Building Materials* 2019;212:210-224. <https://doi.org/10.1016/j.conbuildmat.2019.03.315>
25. Ikeagwuani CC, Nwonu DC. Emerging trends in expansive soil stabilisation: A review. *Journal of Rock Mechanics and Geotechnical Engineering* 2019;11:423-440. <https://doi.org/10.1016/j.jrmge.2018.08.013>
26. Arrieta AB, Díaz-Rodríguez N, Ser JD, Bennetot A, Tabik S, Barbado A, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 2019;58:82-115. <https://doi.org/10.1016/j.inffus.2019.12.012>
27. Soltani A, Deng A, Taheri A, Mirzababaei M, Vanapalli SK. Swell-Shrink Behavior of Rubberized Expansive Clays during Alternate Wetting and Drying. *Minerals* 2019;9:224. <https://doi.org/10.3390/min9040224>
28. Zhou JP, Qiu Y, Zhu S, Armaghani DJ, Khandelwal M, Mohamad ET. Estimation of the TBM advance rate under hard rock conditions using XGBoost and Bayesian optimization. *Underground Space* 2020;6:506-515. <https://doi.org/10.1016/j.undsp.2020.05.008>
29. Lundberg S, Erion G, Chen H, DeGrave AJ, Prutkin JM, Nair BG, et al. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* 2020;2:56-67. <https://doi.org/10.1038/s42256-019-0138-9>
30. Soltani A, Taheri A, Deng AZ, O'Kelly BC. Improved Geotechnical Behavior of an Expansive Soil Amended with Tire-Derived Aggregates Having Different Gradations. *Minerals* 2020;10:923. <https://doi.org/10.3390/min10100923>
31. Jalal FE, Xu Y, Jamhiri B, Memon SA. On the Recent Trends in Expansive Soil Stabilization Using Calcium-Based Stabilizer Materials (CSMs): A Comprehensive Review. *Advances in Materials Science and Engineering* 2020;2020. <https://doi.org/10.1155/2020/1510969>
32. Alam SK, Mondal A, Shiuly A. Prediction of CBR Value of Fine Grained Soils of Bengal Basin by Genetic Expression Programming, Artificial Neural Network and Krigging Method.

Journal of the Geological Society of India 2020;95:190-196. <https://doi.org/10.1007/s12594-020-1409-0>

33. Zhang W, Wu C, Zhong H, Li Y, Wang L. Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. *Geoscience Frontiers* 2020;12:469-477. <https://doi.org/10.1016/j.gsf.2020.03.007>
34. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* 2020;17:261-272. <https://doi.org/10.1038/s41592-019-0686-2>
35. Basack S, Goswami G, Khabbaz H, Karakouzian M, Baruah P, Kalita N. A Comparative Study on Soil Stabilization Relevant to Transport Infrastructure using Bagasse Ash and Stone Dust and Cost Effectiveness. *Civil Engineering Journal* 2021;7:1947-1963. <https://doi.org/10.28991/cej-2021-03091771>
36. Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning* 2021;110:457-506. <https://doi.org/10.1007/s10994-021-05946-3>
37. Miraki H, Shariatmadari N, Ghadir P, Jahandari S, Tao Z, Siddique R. Clayey soil stabilization using alkali-activated volcanic ash and slag. *Journal of Rock Mechanics and Geotechnical Engineering* 2021;14:576-591. <https://doi.org/10.1016/j.jrmge.2021.08.012>
38. Bardhan A, Samui P, Ghosh K, Gandomi AH, Bhattacharyya S. ELM-based adaptive neuro swarm intelligence techniques for predicting the California bearing ratio of soils in soaked conditions. *Applied Soft Computing* 2021;110:107595. <https://doi.org/10.1016/j.asoc.2021.107595>
39. Liang M, Chang Z, Wan Z, Gan Y, Schlangen E, Šavija B. Interpretable Ensemble-Machine-Learning models for predicting creep behavior of concrete. *Cement and Concrete Composites* 2021;125:104295. <https://doi.org/10.1016/j.cemconcomp.2021.104295>
40. Zhang J, Huang Y, Ma G, Nener B. Mixture optimization for environmental, economical and mechanical objectives in silica fume concrete: A novel frame-work based on machine learning and a new meta-heuristic algorithm. *Resources Conservation and Recycling* 2021;167:105395. <https://doi.org/10.1016/j.resconrec.2021.105395>
41. Tasalloti A, Chiaro G, Murali A, Banasiak LJ. Physical and Mechanical Properties of Granulated Rubber Mixed with Granular Soils, A Literature Review. *Sustainability* 2021;13:4309. <https://doi.org/10.3390/su13084309>
42. Yang X, Guan J, Ding L, You Z, Lee VCS, Hasan MRM, et al. Research and applications of artificial neural network in pavement engineering: A state-of-the-art review. *Journal of*

Traffic and Transportation Engineering (English Edition) 2021;8:1000-1021.
<https://doi.org/10.1016/j.jtte.2021.03.005>

43. Formela K. Sustainable development of waste tires recycling technologies - recent advances, challenges and future trends. *Advanced Industrial and Engineering Polymer Research* 2021;4:209-222. <https://doi.org/10.1016/j.aiepr.2021.06.004>
44. Sudarsanan N, Kim YR. A critical review of the fatigue life prediction of asphalt mixtures and pavements. *Journal of Traffic and Transportation Engineering (English Edition)* 2022;9:808-835. <https://doi.org/10.1016/j.jtte.2022.05.003>
45. Ekanayake IU, Meddage DPP, Rathnayake U. A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). *Case Studies in Construction Materials* 2022;16:e01059. <https://doi.org/10.1016/j.cscm.2022.e01059>
46. Moasas AM, Amin MN, Khan K, Ahmad W, Al-Hashem MN, Deifalla AF, et al. A worldwide development in the accumulation of waste tires and its utilization in concrete as a sustainable construction material: A review. *Case Studies in Construction Materials* 2022;17:e01677. <https://doi.org/10.1016/j.cscm.2022.e01677>
47. Zhong X, Gallagher BJ, Liu S, Kailkhura B, Hiszpanski AM, Han TY. Explainable machine learning in materials science. *npj Computational Materials* 2022;8. <https://doi.org/10.1038/s41524-022-00884-7>
48. Wakjira TG, Ibrahim M, Ebead UA, Alam MS. Explainable machine learning model and reliability analysis for flexural capacity prediction of RC beams strengthened in flexure with FRCM. *Engineering Structures* 2022;255:113903. <https://doi.org/10.1016/j.engstruct.2022.113903>
49. Li G, Zhou C, Ahmad W, Usanova KI, Karelina MY, Mohamed AM, et al. Fly Ash Application as Supplementary Cementitious Material: A Review. *Materials* 2022;15:2664. <https://doi.org/10.3390/ma15072664>
50. Phoon K, Cao Z, Ji J, Leung YF, Najjar SS, Shuku T, et al. Geotechnical uncertainty, modeling, and decision making. *Soils and Foundations* 2022;62:101189. <https://doi.org/10.1016/j.sandf.2022.101189>
51. Rudin C, Chen C, Chen Z, Huang H, Semenova L, Zhong C. Interpretable machine learning: Fundamental principles and 10 grand challenges. *Statistics Surveys* 2022;16. <https://doi.org/10.1214/21-ss133>

52. Ho LS, Tran VQ. Machine learning approach for predicting and evaluating California bearing ratio of stabilized soil containing industrial waste. *Journal of Cleaner Production* 2022;370:133587. <https://doi.org/10.1016/j.jclepro.2022.133587>
53. Khasawneh MA, Al-Akhrass HI, Rabab'ah SR, Al-sugaier AO. Prediction of California Bearing Ratio Using Soil Index Properties by Regression and Machine-Learning Techniques. *International Journal of Pavement Research and Technology* 2022;17:306-324. <https://doi.org/10.1007/s42947-022-00237-z>
54. Lyngdoh GA, Zaki MFM, Krishnan NMA, Das S. Prediction of concrete strengths enabled by missing data imputation and interpretable machine learning. *Cement and Concrete Composites* 2022;128:104414. <https://doi.org/10.1016/j.cemconcomp.2022.104414>
55. Barman D, Dash SK. Stabilization of expansive soils using chemical additives: A review. *Journal of Rock Mechanics and Geotechnical Engineering* 2022;14:1319-1342. <https://doi.org/10.1016/j.jrmge.2022.02.011>
56. Yadav VK, Gacem A, Choudhary N, Rai A, Kumar P, Yadav KK, et al. Status of Coal-Based Thermal Power Plants, Coal Fly Ash Production, Utilization in India and Their Emerging Applications. *Minerals* 2022;12:1503. <https://doi.org/10.3390/min12121503>
57. Cao Y, Su FY, Antwi-Afari MF, Lei J, Wu X, Liu Y. Enhancing mix proportion design of low carbon concrete for shield segment using a combination of Bayesian optimization-NGBoost and NSGA-III algorithm. *Journal of Cleaner Production* 2024;465:142746. <https://doi.org/10.1016/j.jclepro.2024.142746>
58. Utkarsh, Jain P. Enhancing the properties of swelling soils with lime, fly ash, and expanded polystyrene -A review. *Heliyon* 2024;10:e32908. <https://doi.org/10.1016/j.heliyon.2024.e32908>
59. Shi XW, Yang P, Li L, Geng X, Liu X, Zhao J. Strength and microscopic pore structure characterization of cement-fly ash stabilized organic soil under freeze-thaw cycles. *Construction and Building Materials* 2024;420:135635. <https://doi.org/10.1016/j.conbuildmat.2024.135635>