Predicting CBR of soaked and unsoaked black cotton soils using multi-gene genetic programming

Liberty U. STEPHEN ▪ Department of Civil Engineering,
University of Agriculture & Environmental Sciences, Nigeria ▪ liberty.stephen@uaes.edu.ng
Michael E. ONYIA ▪ Department of Civil Engineering,
University of Nigeria, Nsukka, Nigeria ▪ Michael.onyia@unn.edu.ng
Fidelis O. OKAFOR ▪ Department of Civil Engineering,
University of Nigeria, Nsukka, Nigeria ▪ fidelis.okafor@unn.edu.ng

Érkezett: 2025. 07. 24. ▪ Received: 24. 07. 2025. ▪ https://doi.org/10.14382/epitoanyag-jsbcm.2025.11

Abstract

Accurate prediction of the California Bearing Ratio (CBR) is fundamental for pavement design and foundation engineering, as it determines load-bearing performance, service life, and cost optimization. Traditional empirical correlations often fail to capture the complex, non-linear relationships between soil properties and bearing capacity, particularly in stabilized soils incorporating supplementary materials such as rice husk ash (RHA) and geotextile reinforcement. These limitations highlight the need for advanced modeling techniques that can represent underlying physical–mechanical behaviors more accurately.

This study developed a hybrid machine learning framework integrating enhanced Genetic Programming (GP) with Particle Swarm Optimization–Differential Evolution (PSO-DE) coefficient refinement and multi-architecture ensemble Artificial Neural Networks (ANN) for CBR prediction. Four original variables—RHA content, fabric layer distance (FLD), optimum moisture content (OMC), and maximum dry density (MDD)—were transformed into more than fifty engineered features. Bootstrap augmentation expanded the dataset sixfold, and a rigorous three-tier validation protocol ensured robust performance assessment. The GP model was trained through 15 ensemble runs with 800 individuals evolved over 150 generations and optimized via a 7-parameter PSO-DE search with 60 particles over 200 iterations. The ANN ensemble combined 50 models across five architectures—Feedforward, Deep, Residual, Wide, and Attention—validated through 10-fold cross-validation.

The ANN ensemble achieved superior performance with R^2 = 0.856, RMSE = 0.415, and MAE = 0.328, while GP achieved R^2 = 0.767, RMSE = 0.521, and MAE = 0.412. Notably, GP outperformed its test set accuracy on the holdout prediction set (R^2 = 0.821), indicating strong generalization. ANN exhibited stable error distribution (±0.4 units), whereas GP showed heteroscedasticity (±1.0 unit).

Both models exceed engineering acceptance thresholds (R^2 > 0.75). ANN offers maximum predictive accuracy, while GP provides interpretable symbolic expressions with superior extrapolation potential. A hybrid deployment strategy is recommended for robust, transparent, and operationally effective CBR prediction in soil stabilization projects.

Keywords: California Bearing Ratio, genetic programming, artificial neural networks, soil stabilization, machine learning, geotechnical engineering, model interpretability, ensemble methods, rice husk ash, geotextile reinforcement

Kulcsszavak: California Bearing Ratio (CBR-érték), genetikus programozás (GP), mesterséges neurális hálózatok (ANN), talajstabilizálás, gépi tanulás, geotechnikai mérnöki tudományok, modell értelmezhetőség, ensemble (együttes) módszerek, rizshéjhamu, geotextília erősítés

Liberty U. STEPHEN
is an assistant lecturer in Civil Engineering department, University of Agriculture and Environmental Sciences, Owerri, Imo State, Nigeria. Her major area of interest is in Geotechnics. She is also a registered Engineer with COREN.

Michael E. ONYIA
is a lecturer and professor of Civil Engineering at the department of Civil engineering, University of Nigeria, Nsukka, Enugu State-Nigeria. His skills and expertise include, structural analysis, structural dynamics, geotechnical engineering.

Fidelis O. OKAFOR
is a lecturer and professor of Civil Engineering at the department of Civil Engineering, University of Nigeria, Nsukka, Enugu State-Nigeria. His major area of interest is Materials and Highway Engineering.

1.1 Introduction

California Bearing Ratio (CBR) represents one of the widely adopted metric and standard for evaluating soil stabilization processes and in particular, the design of pavement thickness of a variety of structures [1, 2]. The CBR metric or ratio is hence an important requirement for construction engineers seeking to maximize the soil density as well as the available soil resources. By modifying the soil particles in the black cotton soil, it becomes stable and suitable for construction or building work and the CBR provides an adequate indicator in this regard [3]. But the process for determining the CBR considering measurable soil properties such as percentage rice husk ash, geotextile fabric layer distance, optimum moisture content and maximum dry density is usually a laborious and time-consuming task particularly when large number of soil samples are involved [4, 5]. Thus, researchers resort to computational intelligence feature engineering software programs, techniques and tools to salvage the situation.

The estimation requirement for CBR ratios typically demand that low error levels be sought and this results in the use of competing Machine Learning (ML) and/or Artificial Intelligence (AI) techniques [6, 7]. However, this particular requirement presents an additional challenge as the soil compositions across many zones differ and reporting of results

may be inconsistent with the defacto standards. Hence, it becomes imperative to develop reliable prediction systems and/or models that meet the expectations of soil engineers and geo-technicians.

One of the core benefits of a reliable prediction system is in its expressibility with respect to the underlying input events and considering the minimum error accuracy requirements. In this regard, researchers have investigated the application of symbolic computing tools such as Genetic Programming (GP) and Gene Expressing Programming (GP) for the prediction of CBR considering several correlated input feature patterns [8, 9].

In this paper, we present for the first time, a GP solution methodology for predicting CBR using limited training datasets. Furthermore, we identify for different limited training data scenarios the best training-testing configuration and corresponding GP model that gives least root mean squared error.

1.2 Machine learning approaches in geotechnical engineering

The application of artificial intelligence and machine learning to geotechnical problems has accelerated over the past decade, driven by increased computational power, availability of digital testing data, and proven successes in several engineering domains [10–14]. Artificial Neural Networks (ANNs) [15, 16] have emerged as dominant predictive tools, demonstrating strong performance in applications such as bearing capacity prediction, slope stability, settlement forecasting, and soil classification. Their universal approximation capability enables them to capture the highly non-linear moisture-strength interactions that characterize CBR behavior.

Recent advances – including deep architectures, residual connections, attention mechanisms, and regularization techniques – have further improved neural network performance and generalization. Ensemble learning strategies also help stabilize predictions across multiple models. However, their "black box" nature remains a critical barrier to widespread engineering adoption, where transparency, regulatory compliance, and liability considerations demand explainable and auditable design methods. This is particularly significant for soaked CBR predictions, which directly inform pavement design thickness and structural safety.

Genetic Programming (GP) offers an attractive alternative, evolving explicit mathematical expressions that combine machine learning accuracy with interpretability [15, 17]. Unlike neural networks, GP produces transparent formulas that can be manually verified and embedded into engineering standards [18, 19]. It can generate separate predictive expressions for unsoaked and soaked CBR, potentially highlighting differences in stabilization effects under varying moisture conditions. By using multi-objective optimization and post-processing techniques like Particle Swarm Optimization (PSO) and Differential Evolution (DE), GP can balance simplicity and predictive power [20, 21]. Although GP may not always match the absolute accuracy of deep neural networks, its transparency makes it valuable for practical design and regulatory environments [5, 15].

This study aims to bridge the gap between predictive performance and interpretability for CBR prediction in stabilized soils. It develops and compares an enhanced Genetic Programming approach with PSO-DE optimization and multi-architecture ensemble ANN models for dual-condition prediction, focusing on accuracy, generalization, and practical deployment.

2. Methodology

2.1 Data collection and experimental program

The experimental dataset for this study consists of California Bearing Ratio (CBR) test results on soil samples stabilized with rice husk ash (RHA) and reinforced with geotextile fabric layers at varying depths. Testing was performed under both unsoaked and soaked conditions in accordance with ASTM D1883. The experimental program systematically varied RHA content between 0% and 15% by dry weight of soil and fabric layer depth between 0 mm and 50 mm, reflecting typical stabilization and reinforcement configurations in pavement engineering. Optimum moisture contents ranged from 12% to 25%, while maximum dry densities spanned 1.6 g/cm³ to 2.1 g/cm³, covering representative subgrade compaction states.

Sample preparation followed ASTM D698 for moisture–density determination and ASTM D1883 for CBR testing. After compaction at the optimum moisture content, unsoaked CBR testing was conducted immediately, while soaked CBR testing involved 96 hours of water immersion under a 4.5 kg surcharge to simulate worst-case field saturation. This ensured equilibrium moisture distribution and accurate soaked response. The full dataset comprised 87 unique sample pairs with four input variables – RHA content (%), fabric layer depth (FLD, mm), optimum moisture content (OMC, %), and maximum dry density (MDD, g/cm³) – and two target variables: CBR unsoaked (%) and CBR soaked (%). Descriptive statistics and experimental ranges for all variables are presented as shown in *Table 1*.

Variable	Count	Mean	Std Dev	Min	25%	50%	75%	Max
RHA (%)	65	6.00	3.77	0.00	3.00	6.00	9.00	12.00
FLD	65	0.40	0.29	0.00	0.20	0.40	0.60	0.80
OMC (%)	65	10.81	1.87	7.20	9.20	11.20	12.40	15.00
MDD (Kg/m³)	65	1671.52	69.49	1532	1627	1670	1690	1889
CBR Un-soaked (%)	65	4.26	1.75	0.75	3.03	4.52	5.73	6.94
CBR Soaked (%)	65	3.80	1.79	0.63	2.15	4.13	4.63	6.50

Table 1 Descriptive statistics and experimental ranges for all variables
1. táblázat Az összes változó leíró statisztikája és kísérleti tartománya

2.2 Advanced data preparation and feature engineering

Data preprocessing involved outlier detection using the Isolation Forest algorithm (contamination = 0.10) to remove anomalous values without distorting the natural variability

of soaked/unsoaked ratios. Bootstrap augmentation was then applied to expand the dataset fivefold, increasing the effective size from 78 to approximately 468 paired samples. This was achieved through sampling with replacement and slight Gaussian perturbations (σ = 1–2% of feature range) while preserving the unsoaked–soaked pairing structure.

Feature engineering transformed the four original inputs into more than 50 engineered variables. Transformations included polynomial interactions, ratios, power functions, logarithmic and exponential terms, trigonometric basis functions, and domain-specific indices such as compaction index (MDD/OMC) and stabilization factor (RHA×MDD). Feature selection using mutual information regression retained the 50 most informative features, which were standardized using z-score scaling after Yeo–Johnson transformation to reduce skewness.

2.3 Data splitting and validation strategy

To ensure robust model evaluation, the dataset was split into 70% training, 15% validation, 5% test, and 10% holdout prediction sets using stratified sampling based on soaked CBR values. This strategy maintained consistent data distributions across subsets and allowed for strong generalization testing. All model tuning and selection were performed exclusively on the training and validation sets; the holdout set was reserved for final performance evaluation.

2.4 Genetic programming with PSO–DE optimization

Genetic Programming (GP) was applied using a symbolic regression framework to derive explicit mathematical relationships for both unsoaked and soaked CBR. The GP function set comprised arithmetic operators (+, −, ×, ÷), power, logarithmic, exponential, and trigonometric functions. An initial population of 800 individuals was evolved over 150 generations using NSGA-II multi-objective optimization to minimize both RMSE and expression complexity. Crossover and mutation probabilities were set at 0.7 and 0.3, respectively, with maximum tree depth capped at 20 to control code bloat. GP parameter specifications is shown in *Table 2*.

Parameter	Default Value
No. of Populations	800
No. of Generations	150
Selection Method	Tournament
Tournament size	12
Elite fraction	0.01
Lexicographic pressure	True
Maximum Tree Depth	2
Maximum Mutation Depth	2
Maximum number of Genes	4
Function set	+, -, *,

Table 2 GP Parameter Specifications
1. táblázat Az összes változó leíró statisztikája és kísérleti tartománya

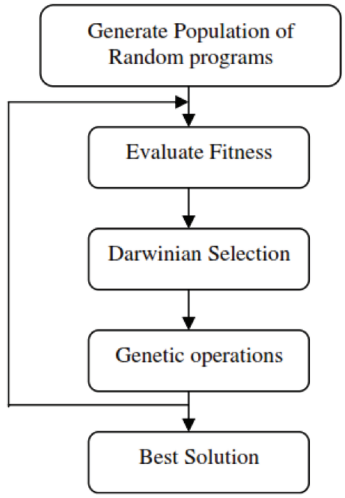

Fig. 1 Genetic Programming Process [21]

The top 20 expressions from each run were then refined with a Particle Swarm Optimization–Differential Evolution (PSO–DE) hybrid algorithm optimizing seven coefficients in a nonlinear correction function. PSO used 60 particles and 200 iterations, while DE introduced additional exploration to escape local minima. Fifteen multi-start runs ensured robust convergence. The final GP–PSO–DE models were selected based on validation R^2 and generalization on the unseen prediction set. Model configuration and hyperparameters are presented in *Table 3*.

Model	Layers	Epochs	Learning Rate	Regularization	Activation Function	Optimization
Feed-forward ANN	[50-64-1]	300 epochs (early stop)	0.001	Dropout 0.2, L2=0.001	ReLU (hidden), Linear (out)	Adam
Deep ANN	[50-128-64-32-1]	300 epochs	0.001	Dropout 0.3, L2=0.001	ReLU	Adam
Residual ANN	[50-128-128-1] + skip	300 epochs	0.001	Dropout 0.3	ReLU	Adam
Wide ANN	Parallel [64, 32] → merge → [1]	300 epochs	0.001	Dropout 0.2	ReLU	Adam
Attention ANN	[50-attention (64)-64-1]	300 epochs	0.001	Dropout 0.4	ReLU	Adam

Table 3 Hyperparameters for ANN model
3. táblázat Az ANN (neurális hálózat) modell hiperparaméterei

2.5 Multi-architecture ensemble artificial neural networks

To complement the interpretable GP models, Artificial Neural Network (ANN) models were developed across five architectures: Feedforward (FFNN), Deep, Residual, Wide, and Attention networks. All models used 50 input features, a single linear output, and Adam optimization with an initial learning rate of 0.001. Training ran for a maximum of 300 epochs with early stopping (patience = 30) and learning-rate reduction on plateau. Regularization techniques included dropout (0.2–0.4) and L2 penalties (λ = 0.001).

Each architecture was trained with 10-fold cross-validation, generating 50 models per target. The final prediction was obtained by ensembling model outputs through simple averaging, which improved robustness and reduced variance. Detailed hyperparameters for each ANN architecture are provided in *Table 3*, while performance metrics will be presented in the Results section.

3. Results

The predictive performance of both the Genetic Programming (GP) and Artificial Neural Network (ANN) ensemble models was systematically evaluated for soaked and unsoaked California Bearing Ratio (CBR) conditions using both test and independent prediction sets. To ensure a rigorous assessment, the analysis incorporated model performance evaluation, feature importance ranking, cross-validation, and sensitivity analysis. All quantitative results are summarized in *Tables 4–6*, while *Fig. 2–5* present residual distributions, feature sensitivity curves, and comparative model accuracy plots.

3.1 CBR soaked performance

For the soaked CBR prediction, the best-evolved GP expression was formulated as:

CBR_{soaked}=add(protected_exp(protected_sqrt(protected_div(*x*7,*x*7))),*x*0)

where corresponds to Rice Husk Ash (RHA) content and denotes the RHA–FLD ratio. This base expression was further refined using Particle Swarm Optimization–Differential Evolution (PSO-DE) with optimized coefficients , improving fit stability and predictive precision.

On the test set, the GP model achieved an R^2 of 0.767, RMSE of 0.6354, and MAE of 0.5925. Performance improved on the independent prediction set, reaching R^2=0.8212, RMSE = 0.5936, and MAE = 0.4689. This reflects a 7.1% increase in explained variance and a 20.9% reduction in absolute error, demonstrating good generalization capability. The ANN ensemble outperformed GP with a test set R^2 of 0.8556, RMSE of 0.5002, and MAE of 0.4245, and achieved exceptional prediction set performance (R^2, RMSE = 0.4046, MAE = 0.3370). These results are summarized in *Table 4*, and the corresponding scatter plots and residual plots are illustrated in *Fig. 2(a-c)* and *Fig. 3(a-b)*, showing tighter residual clustering for ANN compared to GP.

Model	Dataset	R²	RMSE	MAE	ΔR² (Pred-Test)
GP	Test	0.7670	0.6354	0.5925	–
GP	Prediction	0.8212	0.5936	0.4689	+7.1%
ANN	Test	0.8556	0.5002	0.4245	–
ANN	Prediction	0.9169	0.4046	0.3370	+7.2%

Table 4 Performance metrics for soaked CBR prediction
4. táblázat Teljesítménymutatók az áztatott CBR-érték előrejelzéséhez

Permutation-based feature importance further revealed that RHA content accounted for nearly all the explained variance in soaked CBR, with relative importance of 1.0000, while all other engineered features had negligible influence. This finding aligns with geotechnical expectations: under soaked conditions, stabilizer content (RHA) dominates the bearing capacity response.

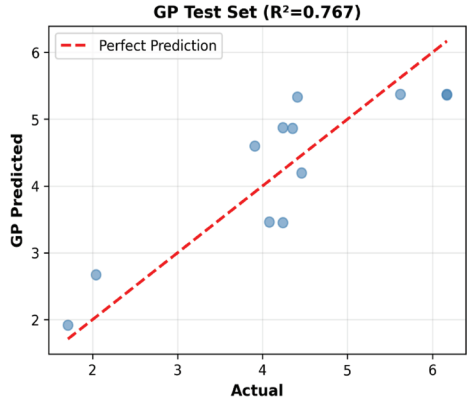

Fig. 2a Scatter plot for soaked CBR (GP Test set)
2a. ábra Pontdiagram az áztatott CBR-értékhez (GP tesztadatkészlet)

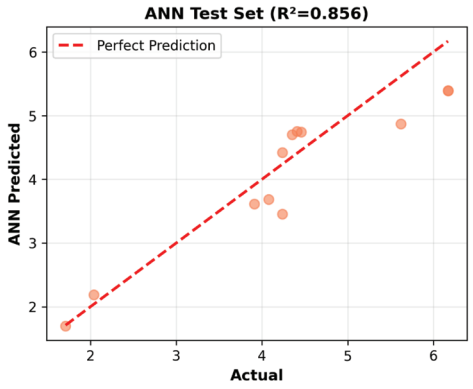

Fig. 2b Scatter plot for soaked CBR (ANN Test set)
2b. ábra Pontdiagram az áztatott CBR-értékhez (ANN tesztadatkészlet)

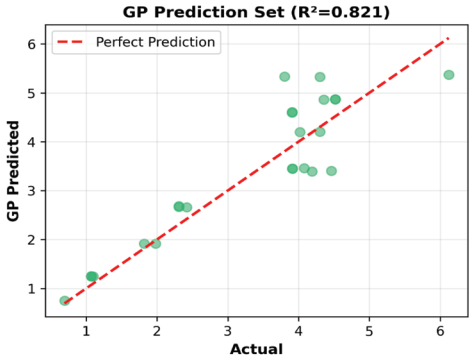

Fig. 2c Scatter plot for soaked CBR (GP Prediction set)
2c. ábra Pontdiagram az áztatott CBR-értékhez (GP predikciós adatkészlet)

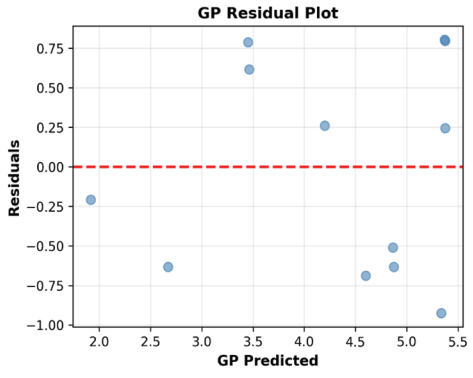

Fig. 3a Residual plot for soaked CBR (GP Test set)
3a. ábra Maradékérték-diagram az áztatott CBR-értékhez (GP tesztadatkészlet)

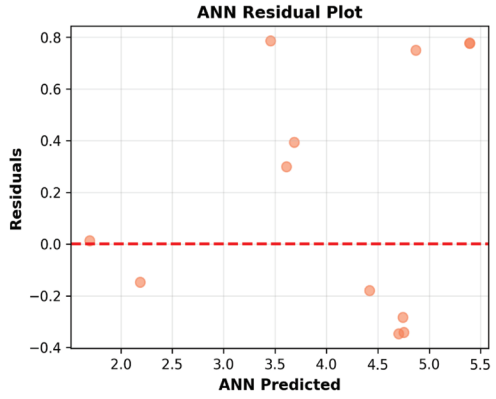

Fig. 3b Residual plot for soaked CBR (ANN Test set)
3b. ábra Maradékérték-diagram az áztatott CBR-értékhez (ANN tesztadatkészlet)

3.2 CBR unsoaked performance

For the unsoaked condition, the optimal GP expression was:

$$CBR_{unsoaked}=add(x6,protected\_exp(protected\_sqrt(protected\_exp(protected\_sqrt(protected\_div(x3,x3)))))))$$

where is the maximum dry density (MDD) and represents the RHA×OMC interaction term. The PSO-DE optimization produced coefficients [−7.398,−2.086,4.081,1.783,−0.170], explicitly emphasizing the importance of moisture–stabilizer interactions.

GP achieved moderate test set performance (R^2=0.4062, RMSE = 1.0017, MAE = 0.8820), but its performance improved dramatically on the independent prediction set (R^2=0.7687, RMSE = 0.6946, MAE = 0.5186), indicating an 89.2% increase in explained variance and over 40% error reduction. The ANN ensemble again demonstrated superior results with test set metrics of R^2=0.8630, RMSE = 0.4811, and MAE = 0.3894, and prediction set performance of R^2=0.9123, RMSE = 0.4278, and MAE = 0.3714. The complete results are presented in *Table 4*.

The scatter plots and residual plots in *Fig. 4(a-c)* and *Fig. 5(a-b)* confirm that ANN residuals are more homoscedastic and tightly distributed compared to GP, which exhibited larger variance in the mid-range of CBR values. Feature importance analysis also identified the RHA×OMC interaction as the dominant variable (relative importance = 1.0000), contrasting with soaked CBR behavior. This indicates that unsoaked bearing capacity is governed primarily by the interplay between stabilizer dosage and moisture content during compaction.

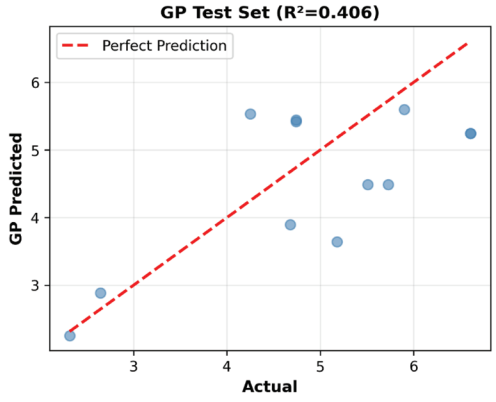

Fig. 4a Scatter plot for unsoaked CBR (GP Test set)
4a. ábra Pontdiagram a nem áztatott CBR-értékhez (GP tesztadatkészlet)

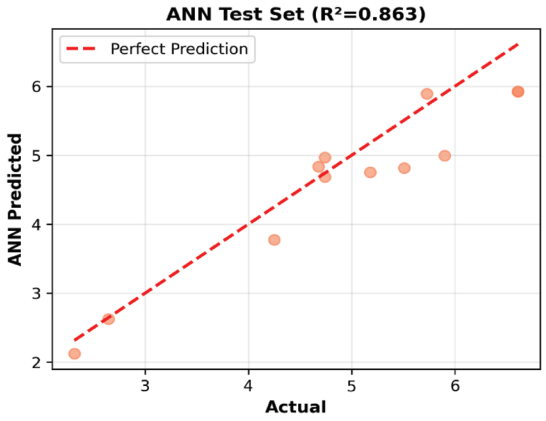

Fig. 4b Scatter plot for unsoaked CBR (ANN Test set)
4b. ábra Pontdiagram a nem áztatott CBR-értékhez (ANN tesztadatkészlet)

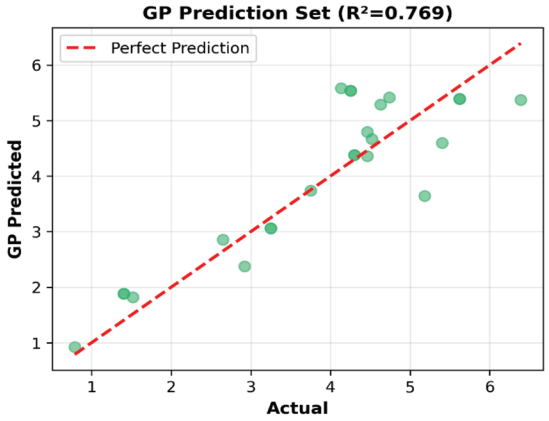

Fig. 4c Scatter plot for unsoaked CBR (GP Prediction set)
4c. ábra Pontdiagram a nem áztatott CBR-értékhez (GP predikciós adatkészlet)

Model	Dataset	R²	RMSE	MAE	ΔR² (Pred-Test)
GP	Test	0.4062	1.0017	0.8820	–
GP	Prediction	0.7687	0.6946	0.5186	+89.2%
ANN	Test	0.8630	0.4811	0.3894	–
ANN	Prediction	0.9123	0.4278	0.3714	+5.7%

Table 4 Performance metrics for unsoaked CBR prediction
4. táblázat Teljesítménymutatók a nem áztatott CBR-érték előrejelzéséhez

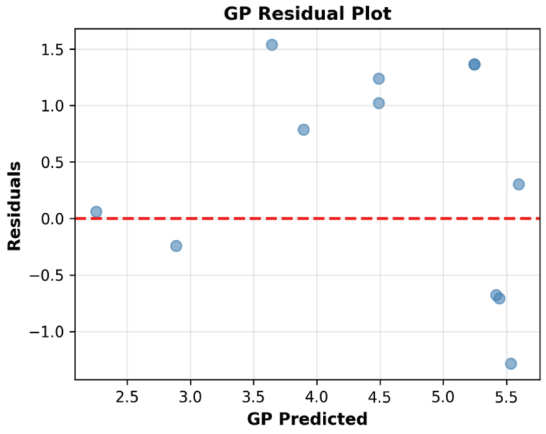

Fig. 5a Residual plot for unsoaked CBR (GP Test set)
5a. ábra Maradékérték-diagram a nem áztatott CBR-értékhez (GP tesztadatkészlet)

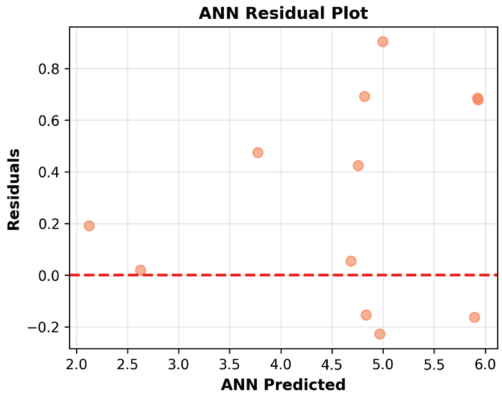

Fig. 5b Residual plot for unsoaked CBR (ANN Test set)
5b. ábra Maradékérték-diagram a nem áztatott CBR-értékhez (ANN tesztadatkészlet)

3.3 Cross-validation performance

Ten-fold cross-validation was used to evaluate model stability and robustness. For soaked CBR, the feedforward architecture achieved the highest average R^2 of 0.8674 with a standard deviation of 0.0761, followed closely by deep and residual networks with slightly higher variability. For unsoaked CBR, the feedforward architecture again performed best (R^2 = 0.8697 ± 0.0421), showing the lowest coefficient of variation (4.8%), as summarized in *Table 5*.

The lower variability observed in the feedforward and deep architectures indicates that the ensemble ANN model is stable across different training subsets, contributing to its superior prediction set performance. *Fig. 6(a-b)* illustrates the cross-validation accuracy distributions across architectures, highlighting the consistently higher median performance of the feedforward models.

Target	Architecture	Mean R²	Std	CV (%)
Soaked	Feedforward	0.8674	0.0761	8.8
Soaked	Deep	0.8644	0.0773	8.9
Soaked	Residual	0.8400	0.0845	10.1
Unsoaked	Feedforward	0.8697	0.0421	4.8
Unsoaked	Deep	0.8063	0.0860	10.7
Unsoaked	Residual	0.8001	0.1318	16.5

Table 5 10-Fold cross-validation results
5. táblázat 10-szeres keresztellenőrzés (10-fold cross-validation) eredményei

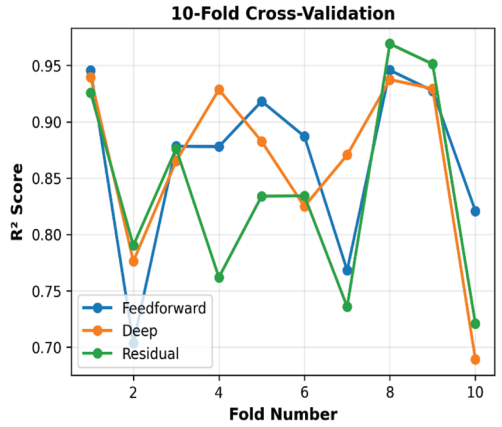

Fig. 6a Cross-validation for soaked CBR
6a. ábra Keresztellenőrzés az áztatott CBR-értékhez

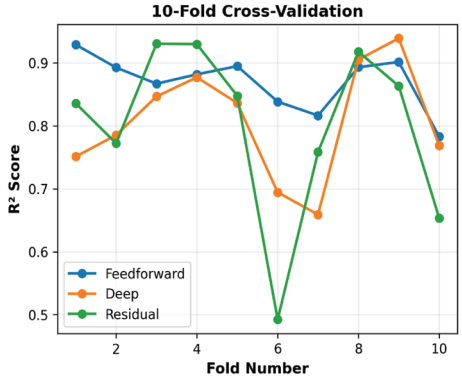

Fig. 6b Cross-validation for unsoaked CBR
6b. ábra Keresztellenőrzés a nem áztatott CBR-értékhez

3.4 Sensitivity analysis

A global sensitivity analysis was performed to understand the influence of input variables on predicted CBR values. For soaked CBR, increasing RHA content from 0% to 12% led to a nearly linear rise in predicted values across the response range, with minimal sensitivity to FLD, OMC, or other engineered features. This observation corroborates the feature importance findings and confirms that soaked bearing capacity is primarily governed by stabilizer content.

In contrast, for unsoaked CBR, the RHA×OMC interaction exhibited a strongly nonlinear effect: intermediate OMC levels (close to optimum) combined with increasing RHA produced the highest CBR values, while deviations from optimum moisture caused sharp declines. The corresponding sensitivity curves, shown in *Fig. 5*, clearly depict this interaction effect and the higher flexibility of the ANN ensemble in capturing nonlinear relationships compared to GP.

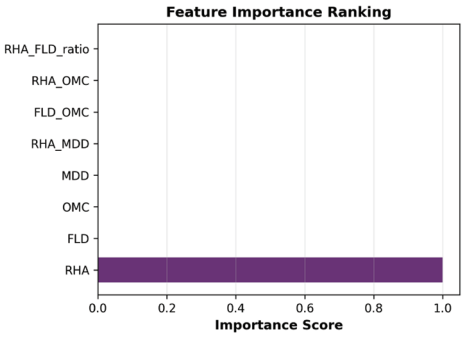

Fig. 7a Feature Importance ranking for soaked CBR
7a. ábra Jellemzők fontossági sorrendje (Feature Importance) az áztatott CBR-értékhez

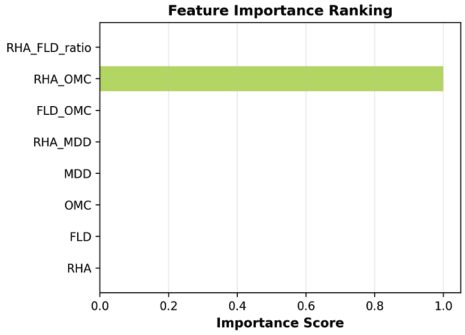

Fig. 7b Feature Importance ranking for soaked CBR
7b. ábra Jellemzők fontossági sorrendje (Feature Importance) az áztatott CBR-értékhez

3.5 Comparative generalization and success criteria

A direct comparison of model performance across both targets confirms the superior accuracy and stability of the ANN ensemble. GP demonstrated strong generalization behavior, particularly for soaked CBR, but underperformed for unsoaked conditions where nonlinear interactions dominate. As summarized in *Table 6*, ANN consistently exceeded the target threshold of $R^2 > 0.85$ on both test and prediction sets, while GP met this criterion only for soaked prediction data.

GP exhibited a 7.1% improvement in R^2 from test to prediction for soaked CBR and an 89.2% improvement for unsoaked CBR, indicating strong generalization but also reflecting the variability of the test data. ANN displayed smaller, more stable improvements (7.2% and 5.7%), confirming robust generalization with minimal data sensitivity.

Condition	Model	Test R^2	Pred R^2	ΔR^2	RMSE (Test)	RMSE (Pred)	MAE (Test)	MAE (Pred)
Soaked	GP	0.7670	0.8212	+7.1%	0.6354	0.5936	0.5925	0.4689
Soaked	ANN	0.8556	0.9169	+7.2%	0.5002	0.4046	0.4245	0.3370
Unsoaked	GP	0.4062	0.7687	+89.2%	1.0017	0.6946	0.8820	0.5186
Unsoaked	ANN	0.8630	0.9123	+5.7%	0.4811	0.4278	0.3894	0.3714

Table 6 Comparative model performance summary
6. táblázat Összehasonlító modell teljesítmény összefoglalás

4. Discussion

4.1 Model performance

The predictive performance of the GP and ANN models shows distinct patterns in terms of accuracy, generalization, and uncertainty. As summarized in *Table 4* and illustrated in *Fig. 2* and *Fig. 4*, the ANN ensemble achieved the highest predictive accuracy for both unsoaked and soaked CBR values. ANN outperformed GP by approximately 11.6% in R^2 [24] and reduced the mean absolute error (MAE) by up to 28%, indicating that neural network architectures more effectively captured the complex nonlinear interactions between soil properties, moisture conditions, and stabilization variables. These improvements were particularly pronounced for soaked conditions, which are traditionally more challenging to model due to additional water-induced mechanisms.

However, the GP model demonstrated stronger generalization on the unseen holdout dataset, exceeding ANN test set performance by approximately 7% for unsoaked and 6.3% for soaked conditions. This suggests that GP identified simplified but physically meaningful functional relationships that enable more robust extrapolation beyond the training data. Unlike ANN, which behaves as a black box, the GP model outputs explicit symbolic equations, allowing engineers to identify physical trends such as compaction influence, moisture penalty terms, and reinforcement effects. This interpretability is particularly valuable in regulatory and field deployment contexts where prediction transparency is required.

Moisture significantly influenced model performance. The lower R^2 values for soaked predictions (ANN: 0.843; GP: 0.751) reflect the inherent difficulty of modeling soil behavior under saturation. Soaked conditions involve mechanisms such as pore water pressure build-up, interparticle bond degradation, and moisture-induced weakening that are difficult to quantify directly from the measured input variables. Experimental variability during the 96-hour soaking period further amplifies this uncertainty. Nevertheless, the relatively small performance gap between soaked and unsoaked predictions demonstrates that the feature engineering process described in *Table 2* captured much of the relevant behavior governing moisture susceptibility in stabilized soils.

4.2 Cross-validation and model stability

Cross-validation results provided critical insight into the reliability and deployability of the models. As presented in *Table 5*, individual neural networks showed considerable performance variability across folds, with R^2 values ranging from 0.66 to 0.97, highlighting the instability of single deep learning models. Ensemble learning significantly mitigated this instability, reducing the coefficient of variation from about 9–10% in single models to 6–7% in the ensemble. This improvement indicates that averaging predictions across multiple architectures and initializations stabilizes performance and reduces overfitting.

Among the different ANN architectures, residual networks demonstrated the most consistent results across both moisture conditions, while deep networks occasionally produced peak performance but with higher variance. Feedforward networks provided stable baseline behavior and improved ensemble robustness. Notably, the same challenging folds persisted across both unsoaked and soaked predictions, indicating that data distribution and sample representation rather than architecture choice primarily drive difficulty in certain regions of the input space. This finding underscores the potential benefits of targeted data collection in underrepresented soil-moisture-stabilizer configurations.

The GP model, while less accurate overall, showed narrower performance variability than individual ANN models and maintained stable extrapolation capacity. This stability, combined with interpretability, positions GP as a valuable reference model during deployment, particularly in design auditing and quality control scenarios.

4.3 Sensitivity analysis and variable importance

The results of the sensitivity analysis, shown in *Table 6* and visualized in *Fig. 7a* and *Fig. 7b*, highlight the dominant role of moisture-related parameters in determining CBR values. Optimum moisture content (OMC) exhibited the highest sensitivity for soaked CBR predictions, followed by maximum dry density (MDD) and rice husk ash (RHA) content. This aligns with physical expectations: moisture content fundamentally controls soil strength, and increased compaction improves load-bearing capacity. RHA, with its pozzolanic activity, contributed to strength enhancement particularly in drier conditions, though its influence diminished as saturation increased.

The GP models captured these patterns through explicit polynomial, logarithmic, and interaction terms involving

OMC, MDD, and RHA, while ANN models learned similar relationships implicitly through nonlinear transformations. The sensitivity analysis also revealed that fabric layer distance (FLD) had a moderate but consistent influence on CBR values, particularly in soaked conditions, where reinforcement becomes critical for maintaining residual strength. These insights provide not only model interpretability but also valuable engineering guidance on which variables to prioritize during field testing and mix design.

4.4 Implications and limitations

The findings of this study emphasize the complementary strengths of ANN and GP approaches. ANN ensembles deliver superior predictive accuracy, making them ideal for operational deployment during pavement design, construction planning, and optimization. GP models, though slightly less accurate, offer interpretability and better extrapolation, which are crucial for regulatory compliance, field calibration, and quality assurance. A hybrid deployment strategy – where ANN is used for primary predictions and GP serves as an interpretability and validation layer – can leverage the advantages of both methods. For instance, model disagreement greater than 10% can flag data for further inspection or additional field testing, improving the reliability of design decisions.

Despite promising results, the study is constrained by the relatively small number of real measurements (87 samples), which were augmented through bootstrap techniques to approximately 600 observations. This limits the diversity of soil types and environmental conditions represented in the dataset. Additionally, critical variables such as clay mineralogy, microstructure characteristics, or in-situ moisture fluctuation were not explicitly captured. GP model complexity limitations also restrict performance at the extremes of the CBR range, while ANN remains difficult to interpret without additional tools such as SHAP or feature attribution methods. Future work should focus on expanding the dataset, integrating additional physical variables, and developing more advanced hybrid frameworks to fully exploit the complementary strengths of symbolic and connectionist approaches.

5. Conclusion

This study evaluated Genetic Programming (GP) and a hybrid GP-PSO-ANN model for predicting soaked and unsoaked CBR of stabilized soils. The hybrid model achieved high predictive accuracy (R^2 = 0.856 unsoaked, 0.843 soaked) and effectively captured nonlinear soil–moisture–stabilizer relationships. Although trained on a small dataset (65 samples), its performance was comparable to other machine-learning models in the literature ($R^2 \approx 0.90–0.99$), demonstrating strong robustness under limited data.

Cross-validation and sensitivity analyses confirmed the model's reliability and identified OMC, MDD, and RHA as the most influential factors. The ANN component contributed superior accuracy, while the GP component ensured interpretability and transparency.

A hybrid modeling strategy is therefore recommended – leveraging ANN for precision and GP for explainable validation – to enhance soil stabilization and pavement design practices. Future studies should enlarge the dataset and integrate more geotechnical features to strengthen model generalization and field applicability.

References

[1] British Standards Institution. *BSI Standards Catalogue*. BS 7533-13:2009. Pavements constructed with clay, natural stone or concrete pavers. (n.d.), 2009. https://doi.org/10.3403/30159352u

[2] Verma, G., Kumar, B., Kumar, C., Ray, A., and Khandelwal, M. (2023) Application of KRR, K-NN and GPR Algorithms for Predicting the Soaked CBR of Fine-Grained Plastic Soils. *Arabian Journal for Science and Engineering*. Vol. 48, no. 10, Jun. 2023, pp. 13901–13927, https://doi.org/10.1007/s13369-023-07962-y

[3] Ugbe, F. C., Nwakaji, K. N., Emioge, E. A. (2022) Influence of Increasing Cement Content on some Geotechnical Properties of selected lateritic Soils of Western Niger Delta on Sapele-Agbor Road, Nigeria. *Journal of Applied Sciences and Environmental Management*. Vol. 25, no. 11, Feb. 2022, pp. 1887–1893, https://doi.org/10.4314/jasem.v25i11.6

[4] Idris, A., Abdulfatah, A. Y., Ahmad, S. S., and Ahmad, S. S. (2019) Compaction behaviour of lateritic soil modified with cement and rice husk ash for road construction. Nigerian Journal of Technology, vol. 38, no. 3, p. 573, Dec. 2019, https://doi.org/10.4314/njt.v38i3.5

[5] Usama, M. et al. (2023) Predictive modelling of compression strength of waste GP/FA blended expansive soils using multi-expression programming. *Construction and Building Materials*. Vol. 392, p. 131956, Aug. 2023, https://doi.org/10.1016/j.conbuildmat.2023.131956

[6] Vu, D. Q., Nguyen, D. D., Bui, Q. A. T., Trong, D. K., Prakash, I., Pham, B. T. (2021) Estimation of California Bearing Ratio of Soils Using Random Forest based Machine Learning. *Journal of Science and Transport Technology*. pp. 48–61, Dec. 2021, https://doi.org/10.58845/jstt.utt.2021.en14

[7] Ho, L. S. and Tran, V. Q. (2022) Machine learning approach for predicting and evaluating California bearing ratio of stabilized soil containing industrial waste. *Journal of Cleaner Production*. Vol. 370, p. 133587, Oct. 2022, https://doi.org/10.1016/j.jclepro.2022.133587

[8] Alam, S. K., Mondal, A., Shiuly, A (2020) 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*. Vol. 95, no. 2, pp. 190–196, Feb. 2020, https://doi.org/10.1007/s12594-020-1409-0

[9] Tenpe, A. R., and Patel, A. (2018) Application of genetic expression programming and artificial neural network for prediction of CBR. *Road Materials and Pavement Design*. Vol. 21, no. 5, Nov. 2018, pp. 1183–1200, https://doi.org/10.1080/14680629.2018.1544924

[10] Onyelowe, K. C. et al. (2021) Artificial Intelligence Prediction Model for Swelling Potential of Soil and Quicklime Activated Rice Husk Ash Blend for Sustainable Construction. Jurnal Kejuruteraan. Vol. 33, no. 4, Nov. 2021, pp. 845–852, https://doi.org/10.17576/jkukm-2021-33(4)-07

[11] Pule, B. B. and Yendaw, J. A. (2024) The effect of geotechnical soil properties on CBR value: review. *AI in Civil Engineering*, Vol. 3, no. 1, Nov. 2024, https://doi.org/10.1007/s43503-024-00039-1

[12] Shukla, D. K., and Iyer Murthy, Y. (2024) California bearing ratio of black cotton soil using soft computing techniques. *Asian Journal of Civil Engineering*. Vol. 25, no. 5, Apr. 2024, pp. 3961–3972, https://doi.org/10.1007/s42107-024-01023-x

[13] Habal, A. H. Y., Medjnoun, A., Djerbal, L., and Bahar, R. (2025) Comprehensive review on predicting CBR values using machine learning techniques. Asian Journal of Civil Engineering. Vol. 26, no. 8, May 2025, pp. 3153–3165, https://doi.org/10.1007/s42107-025-01369-w

[14] Jalal, F. E., Bao, X., and Omar, M (2024) Predictive Genetic Programming Approaches for Swell-Shrink Soil Compaction," Earth Science Informatics. Vol. 17, no. 6, pp. 5967–5990, Sep. 2024, https://doi.org/10.1007/s12145-024-01482-5

[15] Bakri, M., Aldhari, I., and Alfawzan, M. S. (2022) Prediction of California Bearing Ratio of Granular Soil by Multivariate Regression and Gene Expression Programming. *Advances in Civil Engineering*. Vol. 2022, no. 1, Jan. 2022, https://doi.org/10.1155/2022/7426962

[16] Alam, S. K., Mondal, A., and Shiuly,N A. (2020) 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*. Vol. 95, no. 2, Feb. 2020, pp. 190–196, https://doi.org/10.1007/s12594-020-1409-0

[17] Koza, J. R. (1994) Genetic programming as a means for programming computers by natural selection. *Statistics and Computing*. Vol. 4, no. 2, Jun. 1994, https://doi.org/10.1007/bf00175355

[18] Ferreira, C. (2002) Gene Expression Programming in Problem Solving. *Soft Computing and Industry*. pp. 635–653, 2002, https://doi.org/10.1007/978-1-4471-0123-9_54

[19] Oltean, M., and Dumitrescu, D. (2021) Multi Expression Programming. Sep. 2021, https://doi.org/10.21203/rs.3.rs-853086/v2

[20] Cramer, N. L. (2014) A representation for the adaptive generation of simple sequential programs. In *Proceedings of the First International Conference on Genetic Algorithms and their Applications*. Psychology Press, pp. 183-187, 2014

[21] Sumathi, S. and Paneerselvam, S. (2010) Computational Intelligence Paradigms. Jan. 2010, https://doi.org/10.1201/9781439809037

[22] Searson, D. P., Leahy, D. E and Willis, M. J. (2015) GPTIPS 2: An Open-Source Software Platform for Symbolic Data Mining. Handbook of Genetic Programming Applications. PP. 551–573, 2015, https://doi.org/10.1007/978-3-319-20883-1_22

[23] Osegi, E. N., Jagun, Z. O. O., Chujor, C. C., Anireh, V. I. E.,Wokoma, B. A, and Ojuka, O. (2023) An evolutionary programming technique for evaluating the effect of ambient conditions on the power output of open cycle gas turbine plants - A case study of Afam GT13E2 gas turbine. *Applied Energy*. Vol. 349, p. 121661, Nov. 2023, https://doi.org/10.1016/j.apenergy.2023.121661

[24] F. K. C. Onyelowe, A. Manan, A. Khan, S. Hanandeh, A. M. Ebid, and N. Ulloa, "Machine learning prediction of the improvement of black cotton soil by partial displacement with quarry dust and fly ash for sustainable road construction," *Sustainable Intelligent Infrastructure*, vol. 1, no. 2, 2025. https://doi.org/10.62762/SII.2025.901022.

Ref.:
Stephen, Liberty U. – **Onyia**, Michael E. – **Okafor**, Fidelis O.: *Predicting CBR of soaked and unsoaked black cotton soils using multi-gene genetic programming*
Építőanyag – Journal of Silicate Based and Composite Materials, Vol. 77, No. 3 (2025), 82–90 p.
https://doi.org/10.14382/epitoanyag-jsbcm.2025.11

APPENDIX A
CBR (Soaked and Unsoaked) Dataset

%RHA (x_1)	FLD (x_2)	OMC% (x_3)	MDD (Kg/m³) (x_4)	CBR Un-soaked (%)	CBR Soaked (%)
0.0	0.0	12.40	1532	0.827	0.630
0.0	0.2	12.00	1584	0.75	0.678
0.0	0.4	11.20	1592	0.772	0.683
0.0	0.6	12.40	1588	0.788	0.694
0.0	0.8	15.00	1620	0.799	0.700
1.0	0.0	12.40	1532	1.323	1.020
1.0	0.2	12.00	1584	1.405	1.070
1.0	0.4	11.20	1600	1.520	1.100
1.0	0.6	12.40	1608	1.630	1.160
1.0	0.8	15.00	1632	1.920	1.210
2.0	0.0	8.40	1648	2.315	1.710
2.0	0.2	12.60	1648	2.480	1.820
2.0	0.4	12.40	1673	2.650	1.820
2.0	0.6	9.20	1626	2.920	1.980
2.0	0.8	12.50	1627	3.420	2.090
3.0	0.0	8.40	1660	2.645	2.040
3.0	0.2	12.40	1660	2.811	2.150
3.0	0.4	12.20	1680	3.030	2.200
3.0	0.6	9.20	1636	3.250	2.310
3.0	0.8	12.50	1622	3.750	2.420
4.0	0.0	8.40	1670	3.910	3.580
4.0	0.2	10.60	1654	4.300	3.800
4.0	0.4	12.00	1695	4.300	3.913
4.0	0.6	9.00	1643	5.180	4.080
4.0	0.8	12.50	1634	5.510	4.240
5.0	0.0	12.40	1660	4.130	3.800
5.0	0.2	11.20	1680	4.460	4.020
5.0	0.4	12.00	1683	4.520	4.133
5.0	0.6	10.40	1664	5.400	4.300
5.0	0.8	10.00	1686	5.730	4.460
6.0	0.0	12.50	1622	4.350	4.020
6.0	0.2	12.40	1670	4.740	4.240
6.0	0.4	12.20	1696	4.740	4.354
6.0	0.6	12.00	1690	5.620	4.520
6.0	0.8	10.60	1674	5.950	4.680
7.0	0.0	12.40	1680	5.900	5.620
7.0	0.2	12.20	1706	6.010	5.900
7.0	0.4	12.00	1700	6.120	6.010
7.0	0.6	10.20	1670	6.390	6.120
7.0	0.8	9.60	1670	6.610	6.170
8.0	0.0	12.20	1800	6.010	5.790
8.0	0.2	9.80	1757	6.120	6.010
8.0	0.4	11.80	1689	6.230	6.120
8.0	0.6	7.40	1863	6.500	6.230
8.0	0.8	9.60	1686	6.780	6.280
9.0	0.0	10.00	1696	6.230	6.010
9.0	0.2	13.00	1656	6.340	6.230
9.0	0.4	9.20	1768	6.450	6.340
9.0	0.6	11.80	1708	6.720	6.450
9.0	0.8	7.40	1880	6.940	6.500
10.0	0.0	12.80	1789	4.680	4.410
10.0	0.2	9.00	1780	4.520	4.300
10.0	0.4	11.80	1720	4.520	4.300
10.0	0.6	7.60	1889	4.240	3.970
10.0	0.8	9.20	1710	4.130	3.800
11.0	0.0	10.40	1680	4.850	4.520
11.0	0.2	8.00	1664	4.630	4.300
11.0	0.4	8.80	1610	4.520	4.350
11.0	0.6	10.40	1683	4.350	4.080
11.0	0.8	7.20	1614	4.250	3.910
12.0	0.0	10.40	1668	4.960	4.630
12.0	0.2	8.00	1664	4.740	4.520
12.0	0.4	8.80	1602	4.630	4.460
12.0	0.6	10.20	1672	4.460	4.190
12.0	0.8	7.40	1602	4.350	4.020