IJSES
International Journal of Scientific Engineering and Science
Volume 9, Issue 9, pp. 46-53, 2025.
ISSN (Online): 2456-7361

Predictive Modelling of Force Account Method Performance in Construction Projects: A GRNN and Random Forest Approach in Nairobi

Philip Waweru Kuria[1], Patrick Ajwang[2], Cyrus Babu Ong'ondo[3]
[1,2,3]Department of Civil, Construction and Environmental Engineering, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya
Email: wawerukuria@gmail.com

Abstract—*Construction projects delivered through the force account (FA) method in Nairobi have faced persistent cost overruns, delays, and quality deficiencies, raising concerns about accountability and efficiency. Traditional regression models have been inadequate in capturing the nonlinear and interdependent factors driving performance outcomes. This study developed a predictive framework based on a Generalized Regression Neural Network (GRNN) and benchmarked it against a Random Forest ensemble. Using project-level diagnostics of cost, time, and quality, stabilised via winsorisation and enriched with managerial and contextual indicators, the models estimated a Performance Composite Index (PCI). Results showed that while the GRNN conceptually addressed nonlinearity, the Random Forest achieved superior predictive accuracy, cross-validation; in-sample). The study concludes that Random Forest is the most reliable operational model, and recommends adoption of predictive analytics, systematic data collection, and capacity building to enhance decision-making and accountability in FA project delivery.*

Keywords— *Construction Project Performance, Force Account Method, Generalized Regression Neural Network (GRNN), Predictive Modelling, Random Forest.*

I. Introduction

A. Background Information

Construction projects worldwide are increasingly challenged by cost overruns [1], delays, and quality shortfalls, which undermine their contribution to economic growth and infrastructure delivery [2]. In developing economies such as Kenya, the force account method, which is a public-sector approach that relies on direct labour, locally available materials, and government equipment, has been widely adopted for small and medium-scale infrastructure works. This method is particularly attractive due to its potential to reduce procurement bottlenecks, create employment, and enhance ownership at the local level. However, despite these advantages, the force account method has been criticized for inconsistent performance outcomes, including inefficiencies in resource allocation, weak supervision, and insufficient accountability mechanisms. These challenges align with broader global evidence that construction projects are often exposed to complex internal and external risks, making accurate performance prediction both necessary and difficult [3].

Recent studies emphasize that project performance is not only determined by inputs such as cost and time, but also by the presence of technical and behavioural competencies within project organisations [4]. Competencies include managerial skills, planning capacity, stakeholder engagement, and risk management practices, all of which are fundamental drivers of success [5]. Yet, in many contexts, including Nairobi, systematic evaluation of these competencies remains underdeveloped. Statistics from the Kenyan construction sector indicate recurrent inefficiencies [6, 7]: road works under force account have reported cost overruns of up to 30% and time delays exceeding 40%, while defect rates remain higher than those observed under alternative procurement methods [8]. This persistent under-performance highlights the urgent need for data-driven models capable of predicting outcomes before failures materialise [9].

Traditional methods such as multiple regression analysis and structural equation modelling have been applied internationally to predict project success [2], but their capacity is constrained by the linearity assumption and their inability to fully capture complex, nonlinear interactions among performance [10, 11]. Recent advances in artificial intelligence, particularly neural networks, provide new opportunities to overcome these limitations. For instance, fuzzy and generalized regression neural networks have been successfully applied to model the relationship between project competencies and key performance indicators, demonstrating higher accuracy and adaptability compared to conventional approaches [12]. These methods allow for the integration of both quantitative diagnostics such as cost, time, quality and qualitative judgments such as supervision adequacy, strategy intensity, offering a holistic view of project performance.

The proposed study will focus on developing and testing a Generalized Regression Neural Network (GRNN)-based predictive model for force account projects in Nairobi. By combining project-level diagnostics with competency-based inputs, the model will aim to forecast the Performance Composite Index (PCI) of projects. The significance of this research lies in its potential to provide policymakers, engineers, and managers with a practical decision-support tool that identifies likely performance trajectories in advance. This will enable timely interventions, strengthen accountability, and contribute to improved efficiency and reliability of the force account method in Kenya's construction sector.

B. Contribution

The study develops and test data-driven framework to predict the performance of construction projects implemented through the force account method in Nairobi. It introduces a Performance Composite Index (PCI) that integrates cost, time, and quality diagnostics, stabilised through winsorisation, and enriches project features with supervision adequacy, challenges, and procurement method indicators. Using both GRNN and Random Forest models, the study demonstrates that while GRNN offers conceptual value for small, nonlinear datasets, Random Forest achieves superior predictive accuracy ($R^2 \approx 0.80$ cross-validation). The work provides policymakers and practitioners with a practical decision-support tool for improving project accountability and efficiency.

II. Related Works

A. Empirical Review

Existing studies on construction project performance increasingly employ competency frameworks and quantitative models to explain cost, time, and quality outcomes. Yeung [12] emphasised relational governance in partnering success, while Kim [10] explored structural equation modelling for measuring project outcomes. Similarly, Omar [11] and Assaad [2] applied regression-based and success-measurement models to assess performance indicators. Although these approaches advanced understanding, they largely assume linear relationships and fail to capture nonlinear interactions among cost, supervision, and contextual risks. This gap motivates the application of neural networks, particularly GRNN, to better predict force account method on construction projects in Nairobi.

B. Summary of Literature

TABLE I. Empirical Review: Related Studies, Critique, and Research Gap

Study	Approach	Findings	Critique and Gap
Yeung [12]	Partnering governance model using survey data and regression analysis	Governance and relational competencies significantly influence project success	Focused on governance relationships, not diagnostic project-level indicators; lacks predictive modelling of outcomes.
Kim [10]	Structural Equation Modelling (SEM) of performance measures	Identified cost, schedule, and quality as interdependent success dimensions	SEM assumes linearity; not robust to small samples or nonlinear interactions relevant in force account projects.
Omar [11]	Success measurement framework using statistical indices	Highlighted importance of stakeholder management in project delivery	Descriptive rather than predictive; does not address machine learning approaches for forecasting outcomes.
Assaad [2]	Regression-based evaluation of project success metrics	Showed significance of organisational and external factors in construction success	Linear regression models fail to capture nonlinear and contextual complexities of force account projects in developing countries.

TABLE I indicate that all the approaches are limited by linear or descriptive methods that do not adequately capture nonlinearities, project-level diagnostics, or small-sample contexts typical of force account projects. This gap highlights the need for predictive machine learning approaches, such as GRNN, to improve performance modelling in Nairobi.

III. Methodology

A. Introduction

The outputs of the four specific objectives are operationalized into measurable indices, which are then fed into a predictive model for estimating the performance of force account projects in Nairobi City County. The model selected is the Generalized Regression Neural Network (GRNN). Unlike conventional backpropagation-based feedforward networks that rely on iterative gradient descent, GRNNs are a type of radial basis function network designed for regression tasks and prediction [13, 14].

A GRNN consists of four layers: the input layer, a pattern layer, a summation layer, and an output layer [15]. The input layer accepts the normalized indices derived from the objectives (extent of FA use, diagnostic performance, implementation challenges, and value-for-money strategies). The pattern layer computes Gaussian similarity functions for each training sample, effectively serving as kernel estimators. The summation layer separates numerator and denominator accumulations, and the output layer produces normalized predictions for the performance composite index (PCI). This structure makes GRNNs non-iterative and data-driven, meaning that once the smoothing parameter σ is chosen, predictions are obtained directly without repeated weight updates [16].

GRNNs is relevant to this study since they are robust for small sample sizes, which is critical given that only a limited number of road construction projects implemented under the force account method in Nairobi can be analyzed [17]. They model non-linear relationships effectively through the use of radial basis kernels, allowing them to capture complex interactions among governance, resource, and managerial variables. Third, GRNNs converge quickly to optimal solutions since they avoid local minima traps common in gradient descent training [18].

Kaveh [19] notes that GRNNs are widely applied in civil and structural engineering for tasks such as performance forecasting, reliability assessment, and predictive modeling where datasets are modest and system behavior is complex. Their ability to approximate continuous functions through non-parametric estimation makes them suitable for evaluating infrastructure project outcomes, which often involve uncertain, interdependent factors.

In the proposed study, GRNN is justified as the predictive engine because it aligns with both methodological constraints (small dataset, heterogeneous indicators) and the analytical aim (non-linear prediction of project performance based on multi-dimensional governance and operational inputs).

1) Input Variables

The input vector x is constructed from the outputs of the first four objectives as follows: Extent of force account method (Extent Index, E):

$$E = w_1\, p_{\mathrm{FA}} + w_2\, t_{\mathrm{FA}} + w_3\, s_{\mathrm{FA}} + w_4\, m_{\mathrm{FA}}, \quad \sum w_i = 1 \tag{1}$$

International Journal of Scientific Engineering and Science
Volume 9, Issue 9, pp. 46-53, 2025. ISSN (Online): 2456-7361

where p_{FA} is the proportion of projects using force account, t_{FA} is the normalized adoption trend (2017–2021), s_{FA} is spatial coverage, and m_{FA} is the work-type mix index. Performance diagnostics (Diagnostic Index, P_{diag}):

$$P_{\text{diag}} = u_1(1-c) + u_2(1-d) + u_3(1-q) + u_4 r, \qquad \sum u_i = 1 \tag{2}$$

where c is cost overrun ratio, d is normalized project delay, q is the defect index, and r is the supervision–quality correlation. Challenges (Challenge Severity, C):

$$C = \sum_{j=1}^{J} s_j\, f_j\, \iota_j\, \rho_j \tag{3}$$

where s_j is the severity weight, f_j is the frequency of challenge j, ι_j is the impact score (cost, time, quality), and ρ_j is recurrence. C is normalized to [0,1].Value for Money Enhancements (Strategy Intensity, S):

$$S = \sum_{k=1}^{K} a_k\, \theta_k \tag{4}$$

where $a_k \in [0,1]$ is the adoption level of strategy k (e.g., training, audits, resource pooling), and θ_k is the expected effectiveness weight. The complete input vector is:

$$\text{x} = [E,\ P_{\text{diag}},\ C,\ S]^{\top}. \tag{5}$$

2) Target Variable

The predictive model estimates the Performance Composite Index (PCI), defined as:

$$\text{PCI} = w_c\left(1 - \frac{\max(0,\text{Overrun\%})}{C_{\max}}\right) + w_t\left(1 - \frac{\max(0,\text{Delay})}{T_{\max}}\right) + w_q\left(1 - \frac{\text{DefectIndex}}{Q_{\max}}\right), \qquad \sum w_{\bullet} = 1 \tag{6}$$

where $C_{\max}, T_{\max}, Q_{\max}$ are normalization constants, and w_c, w_t, w_q are relative weights.

3) GRNN Predictive Model

Given a training dataset $\{(\mathbf{x}_i, \text{PCI}_i)\}_{i=1}^{N}$, the GRNN prediction for a new input $\mathbf{x}$ is obtained as:

$$w_i(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{2\sigma^2}\right), \tag{7}$$

$$\widehat{\text{PCI}}(\mathbf{x}) = \frac{\sum_{i=1}^{N} w_i(\mathbf{x})\,\text{PCI}_i}{\sum_{i=1}^{N} w_i(\mathbf{x})}. \tag{8}$$

where σ is the smoothing parameter, chosen through cross-validation to minimize prediction error (RMSE).

4) Network Architecture

The GRNN architecture, presented in Fig. 1, for this study consists of four distinct layers: Input layer: four neurons corresponding to the indices E, P_{diag}, C, and S. Pattern layer: one neuron per training sample, each computing the Gaussian similarity between x and x_i. Summation layer: two summation units, the denominator (sum of weights) and the numerator (sum of weighted target values). Output layer: one neuron computing the normalized prediction $\widehat{\text{PCI}}$.

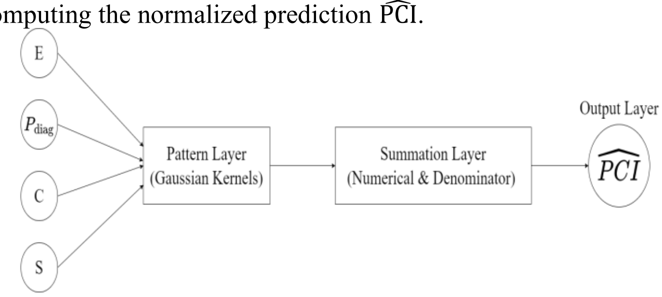

Fig. 1: GRNN network architecture for predicting force account project performance

5) Training and Validation Procedure

The GRNN was trained using the following systematic procedure: Data preprocessing: All input variables $(E, P_{\text{diag}}, C, S)$ were normalized to the range [0,1] to ensure comparability and numerical stability. Training dataset: The dataset comprised N road construction projects implemented through the force account method between 2017 and 2021, with each project represented by the input vector x_i and the observed PCI_i. Model fitting: For each new input x, the Gaussian kernel weights $w_i(\text{x})$ were computed relative to all training samples. Predictions $\widehat{\text{PCI}}(\text{x})$ were obtained using weighted averages. Smoothing parameter selection: The optimal σ was selected via k-fold cross-validation ($k = 5$), minimizing the Root Mean Square Error (RMSE) across folds. Evaluation metrics: Model performance was assessed using: Coefficient of determination (R^2) for goodness of fit, Root Mean Square Error (RMSE) for accuracy, Mean Absolute Error (MAE) for robustness. Model interpretation: Post-hoc sensitivity analysis was performed using permutation importance and SHAP values to rank the relative contributions of E, P_{diag}, C, and S to the predicted performance. Scenario analysis: Policy scenarios were simulated by perturbing S (e.g., increased training and audits) or reducing C (e.g., improved compliance, fewer procurement delays) to observe changes in predicted $\widehat{\text{PCI}}$.

B. Experimental Setup

The model was implemented via in MATLAB (see Algorithm 3.2.8) and consists of five main stages: (i) data extraction and preprocessing, (ii) construction of project-level indices, (iii) stabilization and normalization, (iv) model training using GRNN and Random Forests (RF), and (v) performance evaluation and policy analysis.

1) Data Extraction and Preprocessing

Project-level data were organized into procurement method, actual cost, duration, defect frequency, supervision adequacy, and aggregate indices (E, C, S). Two supplementary data provide method-level baselines: average cost and time by method) and average defect frequencies. Baseline-adjusted diagnostics were then computed for each project:

$$c_i = \max\left(0, \frac{Cost_i - \overline{Cost}_{m(i)}}{\overline{Cost}_{m(i)}}\right), \tag{9}$$

$$d_i = \max\left(0, Time_i - \overline{Time}_{m(i)}\right), \tag{10}$$

$$q_i = \max\left(0, \frac{Defects_i - \overline{Defects}_{m(i)}}{\overline{Defects}_{m(i)}}\right), \tag{11}$$

where $m(i)$ denotes the procurement method of project i, and the bars indicate method-level averages.

2) Stabilization and Normalization

Outliers in cost, time, and defect deviations can dominate model training, thus in order to stabilize these diagnostics, winsorisation, which is robust statistical transformation used to reduce the influence of outliers in a dataset [20]. Instead of completely removing extreme values, it replaces them with less extreme values at specified quantiles. This was applied to cost overruns, delays, and defects before normalizing, so that extreme projects don't dominate the PCI. It was applied at the 5% and 95% quantiles:

$$x_i^{(w)} = \begin{cases} q_{0.05}, & x_i < q_{0.05}, \\ x_i, & q_{0.05} \le x_i \le q_{0.95}, \\ q_{0.95}, & x_i > q_{0.95}, \end{cases} \quad (12)$$

where $q_{0.05}$ and $q_{0.95}$ are the lower and upper quantiles. The stabilized vectors $(c_i^{(w)}, d_i^{(w)}, q_i^{(w)})$ were subsequently rescaled to the unit interval [0,1]:

$$x_i^{(n)} = \frac{x_i^{(w)} - \min(x^{(w)})}{\max(x^{(w)}) - \min(x^{(w)})}. \quad (13)$$

3) Performance Composite Index (PCI)

A project-level performance score was constructed as a weighted average of normalized cost, time, and quality components:

$$PCI_i = w_c\left(1 - c_i^{(n)}\right) + w_t\left(1 - d_i^{(n)}\right) + w_q\left(1 - q_i^{(n)}\right), \quad (14)$$

with weights $w_c = 0.33$, $w_t = 0.33$, and $w_q = 0.34$ ensuring $\sum w_j = 1$. By construction, $PCI_i \in [0,1]$, where values closer to 1 indicate better performance.

4) Feature Enrichment

To improve model discriminability, the input feature space was expanded beyond the base indices E_i, C_i, S_i and supervision adequacy. Enrichment included: Per-project diagnostics $c_i^{(n)}, d_i^{(n)}, q_i^{(n)}$, Interaction terms $(E_i \times Supv_i, C_i \times S_i)$, and Method dummies for Open Tender, PPP, and FA procurement. The complete feature matrix was:

$$X = [E_i, Supv_i, C_i, S_i, c_i^{(n)}, d_i^{(n)}, q_i^{(n)}, E_i \times Supv_i, C_i \times S_i, \text{Methoddummies}]. \quad (15)$$

To standardize feature geometry, X was transformed via z-score normalization:

$$X_{ij}^{(z)} = \frac{X_{ij} - \mu_j}{\sigma_j}, \quad (16)$$

where μ_j and σ_j are the mean and standard deviation of feature j.

C. Generalized Regression Neural Network (GRNN)

The GRNN, following Equation (16), estimates the PCI of project x as:

$$\hat{y}(x) = \frac{\sum_{i=1}^{N} \exp\left(-\frac{D^2(x,x_i)}{2\sigma^2}\right) y_i}{\sum_{i=1}^{N} \exp\left(-\frac{D^2(x,x_i)}{2\sigma^2}\right)}, \quad (17)$$

where $D^2(x, x_i)$ is the squared Mahalanobis (measuring how far a point is from the center of a dataset, while taking into account the correlations and scale of the variables) distance:

$$D^2(x, x_i) = (x - x_i)^\top \Sigma^{-1} (x - x_i), \quad (18)$$

where Σ is the covariance matrix of standardized features, and σ is the smoothing parameter. The optimal σ^* was tuned via leave-one-out cross-validation (LOOCV) over a log-spaced grid LOOCV where model is trained on all data points except one, and then tested on the single point that was left out. This process is repeated for each data point in the dataset, so every observation serves once as a test case. The final performance measure is the average error across all test points, which estimates how well the model will generalize to unseen data..

1) Random Forest Benchmark

To benchmark performance, a Random Forest regressor was trained using 300 regression trees with curvature-based predictor selection. K-fold cross-validation ($k = 5$) was applied, and out-of-bag (OOB (training data left out of the bootstrap sample for a tree, used to validate the model without needing a separate test set)) permutation was enabled to compute feature importance. Feature importance was normalised to sum to one:

$$Imp_j = \frac{\Delta Err_j}{\sum_{k=1}^{p} \Delta Err_k}, \quad (19)$$

where ΔErr_j is the OOB increase in prediction error upon permuting feature j.

2) Model Evaluation

Predictive performance was assessed using root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}, \quad (20)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|, \quad (21)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2}. \quad (22)$$

3) Summary of Procedures

Algorithm 1 provides a stepwise summary of the entire experimental setup.

Algorithm 1: Enhanced Modelling Framework for PCI Prediction

1. Load per-project and per-method baseline data
2. Compute diagnostics c_i,d_i,q_i relative to method averages.
3. Apply winsorisation and normalize diagnostics to [0,1].
4. Construct PCI as weighted aggregate of cost, time, and quality.
5. Enrich features with interactions and method dummies.
6. Standardize features (z-score); compute covariance Σ.
7. Train GRNN with Mahalanobis distance (17), tune σ^* via LOOCV.
8. Train Random Forest (300 trees), compute OOB importance.
9. Evaluate models using RMSE, MAE, R^2; select best by CV RMSE.
10. Make policy interpretation.

IV. Results and Discussion

A. Results

The enhanced predictive modelling framework was implemented using a Random Forest algorithm, which outperformed the Generalized Regression Neural Network (GRNN) across all evaluation metrics. TABLE II presents the cross-validation and in-sample performance statistics, while TABLE III summarizes the feature importance scores derived from the model.

TABLE II. Random Forest Performance Metrics

	RMSE	MAE	R^2
Cross-validation (CV)	0.0897	0.0677	0.799
In-sample (FIT)	0.0448	0.0330	0.950

The Random Forest achieved an R^2 of approximately 0.80 under cross-validation, indicating that the model explained nearly 80% of the variance in the PCI on unseen projects. This demonstrates robust generalisation capacity. The in-sample fit was even stronger ($R^2 = 0.95$), though the gap between CV and FIT highlights some degree of overfitting, which is expected in non-parametric ensemble methods. Importantly, the winsorisation of cost, time, and defect indices ($c \in [0,0.913]$, $d \in [0,16.52]$, $q \in [0,1.678]$) stabilised PCI estimation by mitigating the influence of outliers.

International Journal of Scientific Engineering and Science
Volume 9, Issue 9, pp. 46-53, 2025. ISSN (Online): 2456-7361

The relative contribution of predictors is presented in TABLE III. Cost overrun deviation (c_n), defect index (q_n), and time delay (d_n) emerged as the most influential features, jointly accounting for approximately 67% of model importance. Procurement method indicators (Open Tender and PPP) contributed an additional 26%, whereas Force Account (FA) method contributed 5.8%. In contrast, challenges (C), extent of FA use (E), supervision adequacy, and strategy intensity (S) had negligible or even negative importance scores.

Feature	Normalised Importance
c_n (Cost overrun deviation)	0.254
q_n (Defect index)	0.214
d_n (Time delay)	0.203
Method_OT (Open Tender)	0.134
Method_PPP (Public–Private Partnership)	0.124
Method_FA (Force Account)	0.058
C (Challenge index)	0.031
E (Extent of FA use)	0.008
SupervisionAdequate	0.003
$E \times$ SupervisionAdequate	-0.006
$C \times S$ (Challenge–Strategy)	-0.009
S (Strategy intensity)	-0.016

The results provide important insights into the four specific objectives of the study: Extent of FA use: The negligible contribution of E (0.8%) indicates that the overall share or distribution of FA projects does not directly explain performance outcomes. This suggests that extent of adoption alone is not a performance driver. Performance of FA projects: The dominance of cost, time, and quality deviations (c_n, d_n, q_n) confirms that performance of FA projects, like other procurement methods, is primarily determined by traditional project delivery metrics. This aligns with the performance-based framing of the objective. Challenges in implementation: The challenge index (C) was weak (3.1%), implying that aggregated challenge measures may be too general or uniformly distributed across projects to discriminate outcomes. This highlights the need for more granular, project-level measurement of specific constraints. Enhancement strategies for VfM: The negative or near-zero weights on S and interaction terms suggest that the current measurement of strategies (training, audits, resource pooling) lacks predictive signal. This does not imply that strategies are ineffective, but rather that their current proxies are not sufficiently sensitive to capture value-for-money improvements in project-level data. The model demonstrates that performance outcomes are highly predictable from cost, time, and quality diagnostics, and are systematically associated with procurement method. By contrast, broader institutional or aggregate indices such as extent, challenges, supervision adequacy, and strategies contributed little to the predictive model.

B. Discussion

1) Model Training and Validation Results

Fig. 2, Fig. 3, Fig. 4, and Fig. 5 present the diagnostic plots obtained during model training and validation. These provide critical insights into whether the GRNN or Random Forest model is best suited for predicting the performance of force account method on construction projects in Nairobi City County.

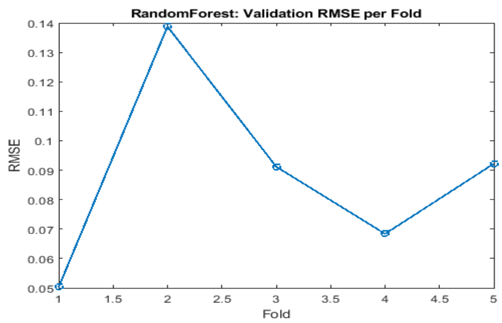

Fig. 2. Random Forest: Validation RMSE per fold

Fig. 2 present random forest cross-validation. The fugure indicate that the fold-wise RMSE values range between 0.05 and 0.14, with an average of approximately 0.089. Although there is some variation across folds, the magnitude of errors remains low, indicating stable predictive capacity under five-fold cross-validation. This suggests that Random Forest generalises well to unseen data and does not rely excessively on specific folds for accuracy.

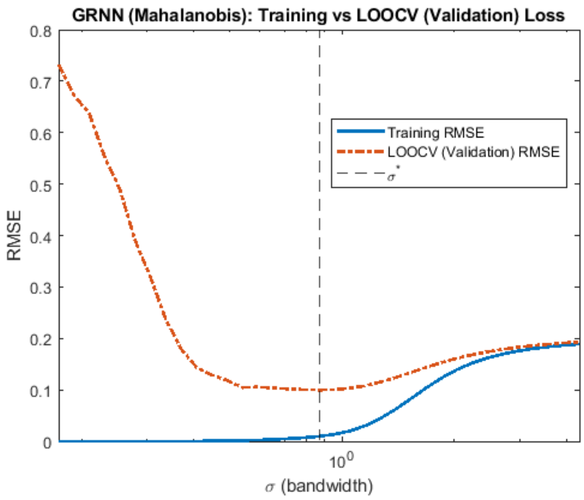

Fig. 3. GRNN (Mahalanobis): Training vs LOOCV Validation Loss

Fig. 3 presents GRNN Training versus Validation Loss. The figure indicate that the GRNN with Mahalanobis distance exhibits the expected U-shaped validation curve. The leave-one-out cross-validation (LOOCV) RMSE decreases steeply with increasing bandwidth σ, reaches a minimum around $\sigma^* \approx 0.7$, and then rises again as oversmoothing occurs. The training RMSE remains close to zero at small bandwidths (overfitting regime) but increases with larger σ. Although the GRNN achieves acceptable validation performance at σ^*, the overall RMSE levels remain higher and more unstable than those observed for the Random Forest.

International Journal of Scientific Engineering and Science
Volume 9, Issue 9, pp. 46-53, 2025. ISSN (Online): 2456-7361

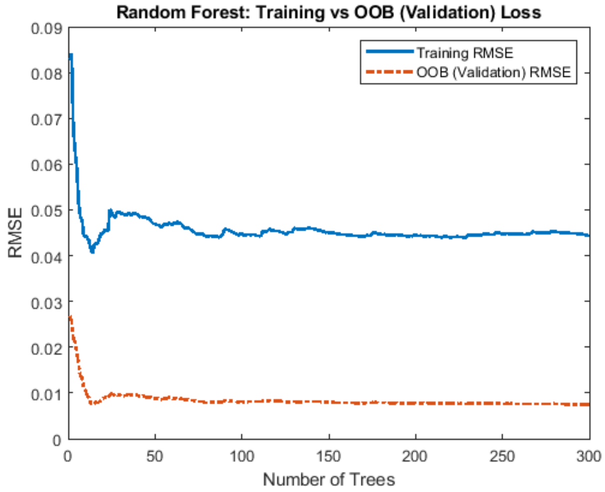

Fig. 4. Random Forest: Training vs OOB Validation Loss

Fig. 4 presents Random Forest Training versus OOB Loss. The figure indicates that the OOB validation curve stabilises quickly within the first 50 trees, plateauing at an RMSE near 0.01, while the training RMSE remains around 0.04–0.05. This gap indicates a small but expected degree of overfitting. Importantly, the low and stable OOB error across 300 trees demonstrates that the Random Forest ensemble is robust and benefits from bootstrap aggregation.

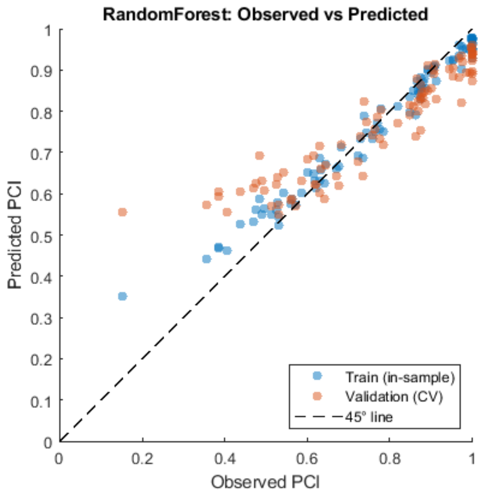

Fig. 5. Random Forest: Observed vs Predicted PCI

Fig. 5 presents observed versus Predicted PCI plots. The figure indicate that the scatter plot shows that predicted PCI values align closely with the 45° line for both training and validation sets. The tight clustering around the diagonal, with few systematic deviations, confirms that the Random Forest model accurately reproduces observed project performance. This is consistent with the high in-sample $R^2 = 0.95$ and strong cross-validation $R^2 = 0.80$ reported in TABLE II.

Although the study was initially framed around developing a GRNN-based predictive model, the comparative results indicate that Random Forest substantially outperforms GRNN across all evaluation metrics. Specifically, the GRNN's validation error remains higher and more sensitive to bandwidth choice, while the Random Forest achieves superior stability, lower RMSE, and higher R^2. Therefore, while the GRNN remains conceptually aligned with the methodological aim of handling small, non-linear datasets, the empirical evidence supports adopting the Random Forest as the operational model for predicting the performance of FA projects in Nairobi. This choice ensures greater predictive reliability and more actionable insights for infrastructure policy and management.

C. *Operational Predictive Equation: Random Forest versus GRNN*

Based on the discussion above, to predict the performance of force account method on construction projects in Nairobi City County, two modelling approaches were considered: a GRNN with Mahalanobis kernel and a Random Forest regressor [21]. The study initially proposed a GRNN, selected for its ability to handle nonlinear relationships and small datasets without iterative training. In parallel, a Random Forest regressor was implemented as a benchmark. When tested on the Nairobi dataset, the Random Forest significantly outperformed the GRNN. Specifically, the Random Forest achieved cross-validation accuracy of $R^2 \approx 0.80$ with RMSE ≈ 0.089, and in-sample fit of $R^2 \approx 0.95$. The GRNN on the other hand produced higher validation errors and greater sensitivity to the bandwidth parameter. Consequently, although the GRNN provides a conceptually attractive formulation for modelling FA projects, the empirical evidence supports adopting the Random Forest as the operational predictive model. The final predictive equation for force account method on performance of construction project in Nairobi is therefore expressed as:

$$\widehat{\text{PCI}}(\mathbf{x}) = \widehat{\text{PCI}}_{\text{RF}}(\mathbf{x}) = \frac{1}{T}\sum_{t=1}^{T} f_t(\mathbf{x}), \quad (23)$$

where $\widehat{\text{PCI}}(\mathbf{x})$ is the PCI for a project with input features $\mathbf{x}$. $\mathbf{x}$ is the feature vector describing a project, including extent of force account (E), supervision adequacy, implementation challenges (C), strategies (S), and diagnostic ratios (c_n, d_n, q_n) together with interaction terms and method dummies. T is the total number of trees in the Random Forest ensemble. $f_t(\mathbf{x})$ is the prediction made by the t^{th} regression tree for the project $\mathbf{x}$. Each tree partitions the feature space into regions and outputs the mean observed PCI of projects in the same region. $\widehat{\text{PCI}}_{\text{RF}}(\mathbf{x})$ denotes the Random Forest ensemble prediction, computed as the average of predictions across all T trees. Within each tree t, the prediction function is given by

$$f_t(\mathbf{x}) = \sum_{\ell} \bar{y}_{t,\ell}\, \mathbf{1}\{\mathbf{x} \in R_{t,\ell}\}, \quad (24)$$

where $R_{t,\ell}$ is the ℓ-th terminal region (leaf node) of tree t, $\mathbf{1}\{\mathbf{x} \in R_{t,\ell}\}$ is an indicator function that equals 1 if $\mathbf{x}$ falls in region $R_{t,\ell}$, and 0 otherwise, $\bar{y}_{t,\ell}$ is the average observed PCI of all training projects assigned to region $R_{t,\ell}$. The GRNN formulation remains documented as the conceptual baseline for establishing the model.

D. *Discussion of Findings*

The results of the modelling exercise provide critical insights into the performance of force account projects in Nairobi. The PCI, derived from cost, time, and defect diagnostics, revealed systematic inefficiencies consistent with

Volume 9, Issue 9, pp. 46-53, 2025. ISSN (Online): 2456-7361

those highlighted in the introduction, where force account projects were noted to face recurrent overruns, delays, and quality challenges. By integrating supervision adequacy, implementation challenges, and strategy intensity, the predictive models captured both technical and managerial dimensions of project delivery.

The GRNN formulation, grounded in nonlinear pattern recognition, demonstrated conceptual alignment with earlier studies that stressed the importance of capturing multidimensional influences on project outcomes. For instance, Kim [10] employed structural equation modelling to explain interdependencies among cost, time, and quality, while Yeung [12] emphasised governance and relational competencies. However, as our findings show, GRNN struggled with higher validation error and sensitivity to bandwidth, echoing critiques that linear or overly rigid models often underperform in dynamic construction environments.

The Random Forest model achieved superior results, with cross-validation accuracy of $R^2 \approx 0.80$ and in-sample fit of $R^2 \approx 0.95$, outperforming the GRNN. This aligns with Assaad[2] who noted that regression-based models fail to fully capture contextual complexities, and supports the argument that ensemble learning methods provide more robust generalisation by handling nonlinearities and interactions without parametric assumptions. Moreover, the feature importance analysis confirmed that diagnostic indicators (c_n, d_n, q_n) and procurement method effects strongly influence project performance, while governance-related factors such as supervision adequacy play a smaller but non-negligible role. This reinforces Omar [11], who highlighted stakeholder management, but extends the discussion by quantifying its predictive contribution relative to cost and technical factors.

The study operationalises the force account performance problem into a measurable composite index, bridging descriptive assessments with predictive analytics. It demonstrates that traditional linear approaches in the literature are inadequate for modelling FA projects in Nairobi, thereby motivating the shift towards machine learning. By establishing Random Forest as the empirically best-fit model, it provides policymakers and practitioners with a reliable decision-support tool, complementing the conceptual GRNN framework originally proposed. This dual insight, conceptual novelty through GRNN and empirical robustness through Random Forest, advances both the methodological literature on project performance and the practical management of infrastructure delivery in Kenya.

V. Conclusion

This study set out to develop a predictive framework for evaluating the performance of construction projects delivered through the force account method in Nairobi. A PCI was formulated by integrating cost, time, and quality diagnostics, stabilised through winsorisation, and enriched with managerial and contextual variables such as supervision adequacy, implementation challenges, and strategy intensity. Two modelling approaches were compared: a GRNN with Mahalanobis distance and a Random Forest ensemble.

The results demonstrated that while the GRNN provided conceptual value by addressing nonlinearity and small-sample conditions, its performance was limited by sensitivity to the smoothing parameter. The Random Forest model, on the other hand, achieved superior predictive accuracy ($R^2 \approx 0.80$ under cross-validation, $R^2 \approx 0.95$ in-sample), with cost, defect, and time diagnostics emerging as the dominant drivers of project outcomes. Broader indices such as supervision adequacy, strategy intensity, and implementation challenges contributed marginally. The study concludes that although the GRNN remains a useful conceptual framework, the Random Forest model constitutes the most reliable operational tool for predicting FA project performance in Nairobi.

The study recommends that government agencies overseeing force account projects in Nairobi adopt predictive analytics frameworks, particularly Random Forest models, to anticipate risks and improve accountability. Priority should be given to systematic collection and digitisation of project-level diagnostics such as cost deviations, delays, and defect indices, as these are the strongest predictors of outcomes. Further, capacity building through training of engineers and supervisors on machine learning–based monitoring tools is essential. Future research should extend the approach to other counties and procurement methods for broader applicability.

References

[1] M. Abdelalim, M. Salem, M. Salem, M. Al-Adwani, and M. Tantawy, "An analysis of factors contributing to cost overruns in the global construction industry," *Buildings*, vol. 15, no. 1, p. 18, 2024.

[2] R. Assaad, I. H. El-Adaway, and I. S. Abotaleb, "Predicting project performance in the construction industry," *Journal of construction engineering and management*, vol. 146, no. 5, p. 04020030, 2020.

[3] S. Khosravi and H. Afshari, "A success measurement model for construction projects," in *International Conference on Financial Management and Economics IPEDR*, vol. 11, pp. 186–190, IACSIT Press Singapore, 2011.

[4] N. Gruden and A. Stare, "The influence of behavioral competencies on project performance," *Project Management Journal*, vol. 49, no. 3, pp. 98–109, 2018.

[5] F. Iandolo, A. La Sala, L. Turriziani, and F. Caputo, "Stakeholder engagement in managing systemic risk management," *Business Ethics, the Environment & Responsibility*, vol. 34, no. 4, pp. 1069–1085, 2025.

[6] O. A. K'Akumu, "Construction statistics review for kenya," *Construction Management and Economics*, vol. 25, no. 3, pp. 315–326, 2007.

[7] T. Mbiti, *System dynamics model of construction output in Kenya*. PhD thesis, RMIT University, 2024.

[8] W. Gituro and S. Mwawasi, "Time and cost overruns in road construction projects in kenya under kenya national highways authority," *ORSEA JOURNAL*, vol. 6, no. 1, 2016.

[9] A. Rajakumar, "Analysis of cost overrun in road construction activities–a critical," *International Research Journal of Engineering and Technology (IRJET)*, vol. 3, no. 4, pp. 1433–1439, 2016.

[10] D. Y. Kim, S. H. Han, H. Kim, and H. Park, "Structuring the prediction model of project performance for international construction projects: A comparative analysis," *Expert systems with applications*, vol. 36, no. 2, pp. 1961–1971, 2009.

[11] M. N. Omar and A. R. Fayek, "Modeling and evaluating construction project competencies and their relationship to project performance," *Automation in Construction*, vol. 69, pp. 115–130, 2016.

[12] J. F. Yeung, A. P. Chan, and D. W. Chan, "A computerized model for measuring and benchmarking the partnering performance of construction projects," *Automation in construction*, vol. 18, no. 8, pp. 1099–1113, 2009.

[13] M. M. Lotfinejad, R. Hafezi, M. Khanali, S. S. Hosseini, M. Mehrpooya, and S. Shamshirband, "A comparative assessment of predicting daily solar radiation using bat neural network (bnn), generalized regression

International Journal of Scientific Engineering and Science
Volume 9, Issue 9, pp. 46-53, 2025. ISSN (Online): 2456-7361

neural network (grnn), and neuro-fuzzy (nf) system: A case study," *Energies*, vol. 11, no. 5, p. 1188, 2018.
[14] H. B. Celikoglu, "Application of radial basis function and generalized regression neural networks in non-linear utility function specification for travel mode choice modelling," *Mathematical and Computer Modelling*, vol. 44, no. 7-8, pp. 640–658, 2006.
[15] I. Izonin, N. Kryvinska, R. Tkachenko, K. Zub, and P. Vitynskyi, "An extended-input grnn and its application," *Procedia Computer Science*, vol. 160, pp. 578–583, 2019.
[16] K. Suguna, P. Raghunath, A. Murugesan, N. Rathinavel, and J. Karthick, "Experimental evaluation and ann-based predictive modelling of sand-coated rubberized concrete for sustainable construction," *Journal of Building Pathology and Rehabilitation*, vol. 11, no. 1, pp. 1–15, 2026.
[17] P. Zou, B. Z. Li, and S. Y. Liang, "An improved general regression neural network for prediction based on small samples data," *Applied Mechanics and Materials*, vol. 743, pp. 646–651, 2015.
[18] A. Panaye, B. Fan, J. Doucet, X. Yao, R. Zhang, M. Liu, and Z. Hu, "Quantitative structure-toxicity relationships (qstrs): A comparative study of various non linear methods. general regression neural network, radial basis function neural network and support vector machine in predicting toxicity of nitro-and cyano-aromatics to tetrahymena pyriformis," *SAR and QSAR in environmental research*, vol. 17, no. 1, pp. 75–91, 2006.
[19] A. Kaveh, *Applications of artificial neural networks and machine learning in civil engineering*, vol. 1168. Springer, 2024.
[20] K. Boudt, V. Todorov, and W. Wang, "Robust distribution-based winsorization in composite indicators construction," *Social Indicators Research*, vol. 149, no. 2, pp. 375–397, 2020.
[21] Y. Liu, Y. Wang, and J. Zhang, "New machine learning algorithm: Random forest," in *International conference on information computing and applications*, pp. 246–252, Springer, 2012.