A Rule-Guided Conditional Diffusion Model for Engineering-Constrained Cost Scenario Generation and Risk Assessment in Textile Geomembrane Projects

Baolan Wei
Taiyuan City Vocational College, Taiyuan 030027, Shanxi, China
Corresponding author: Baolan Wei (e-mail: weibaolan321@163.com)

ABSTRACT While diffusion models excel in generating high-dimensional data, their purely data-driven nature makes it difficult to integrate engineering constraints and physical laws, resulting in generated cost scenarios for textile geomembranes lacking construction feasibility and failing to support reliable risk assessment. Existing diffusion models generate cost scenarios that often violate engineering constraints. This paper solves this problem by integrating explicit domain knowledge into a conditional diffusion model. The proposed model generates cost scenarios that are both statistically faithful and engineeringly feasible, enabling reliable multi-dimensional risk assessment. Methodologically, a 12- dimensional conditional feature model is constructed based on 32 historical projects and national standards. Thirteen engineering constraints are formally extracted and integrated into the Drools rule engine. A constraint-aware Classifier- Free Guidance diffusion architecture is designed, dynamically verifying and correcting feasibility during training loss and sampling. Experiments show that the proposed method achieves an engineering feasibility rate exceeding 93% and an average KL (Kullback-Leibler) divergence of 0.108 and notably, VaR@95% estimation bias below 4.2% across diverse scenarios, with CVaR estimation error within ±5.1%. Extreme high-cost events are captured with over 89% coverage, demonstrating that the generated scenarios enable reliable multi-dimensional risk quantification. This divergence is measured between the distribution of total costs from the generated scenarios and the empirical distribution of total costs from real historical projects within the held-out test set. This research provides a paradigm for intelligent generation and risk assessment of engineering costs that balances data fidelity and physical feasibility, demonstrating promising engineering applications.

INDEX TERMS textile geomembrane, conditional diffusion model, engineering cost risk, constraint-aware generation, nondestructive monitoring

I. INTRODUCTION

WITH the increasing demands for environmental protection and seepage prevention in infrastructure construction, textile geomembranes are widely used in water conservancy, landfill, tailings dam, and other projects. As the core basis for project decision-making and risk control, the uncertainty modeling of engineering cost urgently requires high-fidelity, multi-scenario generation capabilities [1-3]. In recent years, diffusion models have been gradually applied into the field of engineering economics due to their superior performance in generating high-dimensional data [4-6]. However, most existing diffusion models are purely data-driven and lack explicit modeling of engineering specifications, construction logic, and physical laws [7-9]. As a result, although the generated cost scenarios are statistically reasonable, they often violate actual construction constraints and are difficult to support credible risk assessment and investment decisions. Textile geomembrane engineering, as a key application scenario for geosynthetics, has a highly complex cost structure and is sensitive to construction logic, making it a research focus in the fields of engineering economics and risk management [10,11]. Zhang Yuhao et al. used dual-stream deep learning to conduct health detection on HDPE (High Density Polyethylene) geomembranes based on landfill containment systems to detect geomembrane rupture location [12]. Abdel-Hamid Mohamed et al. studied the impact of low labor productivity on engineering costs [13]. Many studies have focused on the service life of HDPE geomembranes, considering different scenarios [14-16], laying the foundation for cost estimation. These studies collectively reveal that the cost of textile geomembrane engineering is not only driven by traditional economic variables, but also deeply coupled with multidimensional engineering logic, urgently requiring a generation method that can simultaneously model data distribution and physical feasibility.

To address the issue of engineering cost estimation, the academic community has attempted various algorithmic approaches, but all suffer from limitations in balancing data fitting and engineering logic. For example, Monte Carlo simulation is used to estimate project costs, but existing research does not adequately address the coupling between multiple factors, leading to distortion of the generated scene's relevance [17- 19]. Some studies use WGAN-GP (Wasserstein GAN with Gradient Penalty) for scene generation, utilizing gradient penalty to stabilize the training process, but it is difficult to apply to the field of engineering cost estimation and cannot embed normative constraints [20-22]. Many studies use XGBoost (eXtreme Gradient Boosting) for cost estimation, but its scenario coverage is limited, making it difficult to adapt to the diversity of engineering cost scenarios for textile geomembranes [23-25]. There are issues with the gap in the current methods, as they are unable to combine the high distribution fidelity with high engineering feasibility; therefore, the development of a new generative architecture is necessary to accomplish this by incorporating domain expertise directly into the creation of new Data Driven Generations.

This work bridges the gap. We propose a constraint-aware conditional diffusion model. The model generates cost scenarios that respect both learned data patterns and codified engineering rules. These scenarios then support multi-dimensional risk quantification. Specifically, firstly, a 12-dimensional engineering feature condition vector is constructed based on 32 historical projects and national standards. Secondly, 13 hard/soft constraint rules are formally extracted and integrated into the Drools rule engine. Then, a constraint-aware Classifier-Free Guidance diffusion architecture is designed to dynamically integrate feasibility verification and projection correction during training loss and sampling. Experiments show that the proposed method outperforms existing benchmarks in terms of engineering feasibility ratio, distribution fidelity, component correlation, VaR bias rate, and extreme scenario coverage, providing a new paradigm for intelligent risk assessment of engineering costs.

II. METHODS

A. CONSTRUCTION OF ENGINEERING CHARACTERISTIC VARIABLE SYSTEM

This paper extracts key factors affecting cost based on the "Technical Specification for Application of Geosynthetics" (GB/T 50290–2014) [26-28]. Through expert interviews and analysis of variance (ANOVA), 12 variables that contribute the top 95% to the total cost are selected to form the conditional input vector $c \in \mathbb{R}^{12}$, including membrane material type, laying area, foundation type, slope, climate zone, transportation distance, labor unit price, machinery shift cost, overlapping method, construction season, construction period, and environmental protection level.

For categorical variables, One-Hot encoding [29-30] is used to convert them into sparse binary vectors; for continuous variables, Min-Max normalization is performed to map them to the [0,1] interval, as shown in the formula:

$$c_j^{\text{norm}} = \frac{c_j - \min(D_j)}{\max(D_j) - \min(D_j)}, \quad j \in J_{\text{cont}} \tag{1}$$

In Formula 1, c_j is the original feature value of the j-th dimension; D_j is the set of values for this feature in the 32-item dataset; J_{cont} is the continuous variable index set. After normalization, all features have a consistent scale, avoiding gradient shift caused by differences in scale during model training. The resulting standardized conditional vector c serves as the conditional input to the diffusion model.

B. DESIGN OF DOMAIN KNOWLEDGE CONSTRAINT RULE BASE

To ensure that the generated cost scenario aligns with actual engineering conditions, this paper systematically extracts 13 domain constraint rules from the "Technical Specification for Application of Geosynthetics" (GB/T 50290–2014), construction organization design documents, and technical briefing records from 32 historical projects, including 8 hard constraints (which must be met) and 5 soft constraints (allowing for moderate deviations). Hard constraints include the logical relationship between materials and processes, while soft constraints reflect the impact of the environment on efficiency, such as "The construction site is in a cold temperate zone, and the season is winter $\Rightarrow$ Labor efficiency reduction factor ≤ 0.7".

All rules are formalized as Boolean logic expressions or inequalities and encoded into the Drools rule engine [31,32]. For example, the GCL rule above is transformed into:

$$\begin{aligned} &\text{IF } (\text{membrane\_type} = \text{GCL}) \wedge (w_{\text{lap}} \geq 0.3) \\ &\text{THEN } \eta_{\text{loss}} \geq 0.08 \end{aligned} \tag{2}$$

In Formula 2, w_{lap} represents the overlap width (m), and η_{loss} represents the material loss rate. During runime, the rule engine performs consistency checks on the input cost vector y and outputs a binary feasibility judgment.

Based on this, a global constraint function $C(y)$ is constructed:

$$C(y) = \begin{cases} 1, & \text{if } \bigwedge_{k=1}^{8} H_k(y) = \text{true} \\ & \text{and } \sum_{l=1}^{5} I(S_l(y) \text{ violated}) \leq \tau, \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

In Formula 3, $H_k(\cdot)$ represents the k-th hard constraint predicate; $S_l(\cdot)$ represents the l-th soft constraint; $I(\cdot)$ is the indicator function; $\tau = 2$ is the upper limit of the soft constraint that can be violated. This function is used for feasibility screening in subsequent diffusion model training and sampling to ensure that the generated samples meet both mandatory engineering specifications and have reasonable flexibility, effectively solving the problem of pure data-driven methods ignoring physical logic.

C. CONDITIONAL DIFFUSION MODEL ARCHITECTURE DESIGN

This paper adopts the Classifier-Free Guidance Diffusion Model as the generative framework to solve the balance problem between unconditional generation and conditional control [33]. Given a standardized 12-dimensional engineering condition vector $c \in [0,1]^{12}$, the model objective is to generate a 7-dimensional cost vector $y = [y_1, y_2, \ldots, y_7]^\top$ that meets the conditions, where y_1 to y_6 correspond to material costs, labor costs, machinery costs, transportation costs, measures costs, and management costs, respectively, and y_7 is the total cost.

The forward diffusion process is defined as $q(y_t|y_0) = N(y_t; \sqrt{\bar{\alpha}_t}y_0, (1-\bar{\alpha}_t)I)$. Among them, y_0 is the actual cost sample; $t \in \{1, \ldots, T\}$ is the time step; $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$, and α_s is the retention coefficient at the s-th step. The inverse denoising network $\epsilon_\theta(y_t, t, c)$ adopts the U-Net++ architecture, embedding multi-scale skip connections and channel-space dual attention modules to enhance the ability to model the dependencies between cost components. The time step t is mapped to a 128-dimensional vector through sinusoidal position encoding, and the conditional vector c is concatenated with the temporal embedding ajer linear projection and input to each residual block of the network.

To integrate engineering feasibility, the loss function, based on the standard diffusion loss $L_{\text{diff}} = E_{y_0,t,\epsilon}\left[\|\epsilon - \epsilon_\theta(y_t, t, c)\|^2\right]$, applies a constraint-aware term:

$$L_{\text{total}} = L_{\text{diff}} + \lambda \cdot E_{y_0,t}\left[1 - C(\hat{y}_0)\right] \tag{4}$$

In Formula 4, $\hat{y}_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(y_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(y_t, t, c)\right)$ represents the denoised estimate; $C(\cdot) \in \{0,1\}$ represents the feasibility function; $\lambda = 0.5$ represents the balancing weights. This design allows the model to explicitly penalize infeasible samples during the training phase, guiding the generated distribution to shift towards a region with reasonable engineering parameters, thereby addressing the problem of purely data-driven methods ignoring physical logic. The diffusion model is needed because it is capable of learning a representation of the stochastic joint distribution of the seven cost components based on the twelve input features. In addition, the thirteen formal rules encoded in Drools set forth necessary, but not sufficient, conditions for feasibility; these rules provide both hard and soft boundaries that exclude unsupported scenarios. The models have no ability to generate complex dependencies, uncertainty in the data, etc., associated with the way real-world cost data is displayed. The second approach described allows for the classification of invalid sample points as well as the ability to provide many different options of valid, highfidelity cost scenarios that would be generated from the conditional input using the rule engine alone; however, it does not allow for the same level of diversity seen in the original data set. The second approach utilizes a diffusion model to navigate through the detailed data manifold that exists within the feasible space defined by the rule engine's constraints. It captures subtle correlations, such as how a change in foundation type probabilistically influences not just material costs but also machinery and labor costs, relationships that are not explicitly defined in the rule base. By creating an accurate distribution of potential outcomes that reflects the constraints imposed by the explicit constraints of the deterministic modelling system and the implicit constraints found in the staisical structure of the historical data, the model produces a realistic distribution of feasible risk scenarios to assist in a complete assessment of risk variability and uncertainty beyond deterministic modelling methods.

It is essential to clarify the relationship of the Drools Rule Engine to the Constraint Aware Guidance Process, and its role in that process. Drools is not a filter after the generation process, nor is it a substitute for the internal learning mechanism of the Diffusion model. Drools functions as a Real-Time OAuth (Oracle) Feasibility Oracle and is closely linked with Diffusion at two points; First, when training occurs, Drools examines every sample produced through Diffusion and gives each a yes/no response regarding its feasibility. The responses from Drools, as to whether each sample produced by Diffusion was feasible or not, determine how the Constraint Aware Component within the Training Objective is built. The second point in time that Drools is used is during the sampling of intermediate samples. For each of these intermediate samples produced through Diffusion, Drools will be called to assess whether or not the Hard Constraints were violated. If Hard Constraint Violations are found to exist, an immediate correction step shall be initiated by Drools. Secondly, Drools will provide an assessment of Soft Constraint Violations based upon the intermediate sample produced and it will be determined how much to dynamically range scale the Conditional Guidance Strength given the existing Soft Constraint Violations. In both cases, not only does Drools enforce feasibility, it also teaches what feasible means.

D. CONSTRAINT-AWARE NOISE SCHEDULING AND SAMPLING STRATEGIES

At each denoising step, the diffusion sampler interacts with the Drools engine through a standardized two-phase process:

In the first phase (Feasibility Assessment & Guidance Adaptation), the intermediate sample is provided to the rule engine where Drools evaluates all hard and soft constraints and returns:

(a) a hard violation flag,

(b) the number of soft violations.

Using this information for (b), Drools then adaptively increases the guidance strength, meaning that the more soft violations that occur, the stronger the conditioning will become. This helps to encourage the model to move away from the regions of infeasibility.

In the second phase, (Hard Constraint Enforcement), if a hard constraint has been violated (according to item a above), the intermediate sample is projected onto the closest point within the feasible set conditioned on the engineering rules; otherwise no changes are made to the sample. Therefore only samples that have either been corrected or verified as feasible may be entered into the next denoising step, helping to ensure that each sample that moves on into future time steps will meet all required engineering requirements, while allowing for appropriate levels of control over soft constraints. Figure 1 showing the tight integration of the diffusion model and Drools rule engine. The Drools engine provides real-time feasibility feedback at every denoising step to dynamically adjust guidance and enforce hard constraints.

To improve the engineering rationality of generated samples in the low signal-to-noise ratio stage, this paper adopts Cosine Noise Schedule instead of linear scheduling [34], and its noise variance sequence is defined as:

$$\beta_t = \frac{\cos^2\left(\frac{t}{T}\cdot\frac{\pi}{2}\cdot s\right)}{\cos^2\left(\frac{t-1}{T}\cdot\frac{\pi}{2}\cdot s\right)}\cdot(1-\alpha_{\min})+\alpha_{\min}, \quad t=1,\ldots,T \tag{5}$$

In Formula 5, $T = 50$ represents the total number of sampling steps; $s = 0.008$ represents the smoothing coefficient; $\alpha_{\min} = 0.0001$ represents the minimum signal-to-noise ratio. This scheduling retains more detailed information in the later stages, which is beneficial for the fine generation of cost items.

During backsampling, ajer generating intermediate samples $\tilde{y}_t$ at each step, the Drools rule engine is immediately invoked to perform a feasibility check. If any hard constraint is violated, a projection correction is performed to solve the optimization problem.

$$\hat{y}_t = \arg\min_{z\in F} \|z-\tilde{y}_t\|_2 \tag{6}$$

In Formula 6, $F = \{z \mid H_k(z) = \text{true}, \ \forall k = 1,\ldots,8\}$ is the feasible set defined by hard constraints, ensuring that the corrected samples meet mandatory engineering specifications. For soft constraints, the guidance strength ω of Classifier-Free Guidance is dynamically adjusted: if the l-th soft constraint is violated, then $\omega \leftarrow \omega\cdot(1+\gamma_l)$, where γ_l is a preset penalty coefficient (ranging from 0.1 to 0.3), enhancing conditional guidance to suppress unreasonable deviations. This strategy significantly improves the engineering feasibility of samples while maintaining generation diversity, solving the problem that diffusion models easily generate physically infeasible solutions without explicit constraints. Enforcing strict domain constraints may lead to too much restriction on the diversity of generated scenarios. However, the impact of these constraints on diversity is collectively managed. While these rules define the borders of the feasible region, they do not limit all solutions to stay within them. The diffusion model maintains its stochastic characteristics and will explore the interior of the constrained space. Understanding the differences between hard constraints and soft constraints is very important. Hard constraints are inflexible and cannot be violated; therefore, they eliminate scenarios that cannot exist based on a hard constraint. On the other hand, soft constraints provide a range of allowable deviation within hard constraints to preserve a degree of variability that reflects the uncertainty of factors such as labour efficiency in real life. Projecion correction applies a minimal number of changes to satisfy hard constraints, while the adjustment of guidance strength for soft constraints provides flexibility for soft constraint trajectories. The combination of a knowledge-based rule engine and a data-driven diffusion model through hybrid architecture allows for the respective shortcomings of each to help alleviate the other's. While the knowledge-based rule engine can provide validation but not generation of new data, the data-driven diffusion model has been trained with limited amounts of data which can lead to generating statistically valid outputs that are physically unrealistic. The use of both models is critical to ensuring that an effective and disciplined generative process occurs, while adhering to defined limits of the physical domain. The knowledge-based rule engine also helps provide the foundation of trustworthiness when using results generated for applications within the engineering realm. Moreover, the combination of both models is critical not only to eliminating the occurrence of infeasible points but also for generating realistic output across the high-dimensional feasible region defined by a valid probability distribution that is independent of either the knowledge-based or a simple generative model that does not utilize the sophisticated, statistically-based distribution modeling methods inherent to the data-driven diffusion model. The constraint-aware loss function varies with training epochs, as shown in Figure 2.

Figure 2 shows the variation curve of the constraint-aware loss function of the proposed constraint-aware conditional diffusion model with the number of training rounds during the training process. As can be seen from the trend of the curve, the loss value first decreases rapidly, then tends to flatten out, and basically converges to a stable low value range after about 120 rounds. This indicates that the model quickly learns the data distribution and constraint rules in the early stage of training, and enters the fine-tuning stage in the later stage.

E. COST SCENARIOS GENERATION AND RISK INDICATOR MAPPING

After the model training is completed, 10,000 independent back samplings are performed for each test case to generate the corresponding cost scenario set $\{y^{(i)}\}_{i=1}^{N}$, where

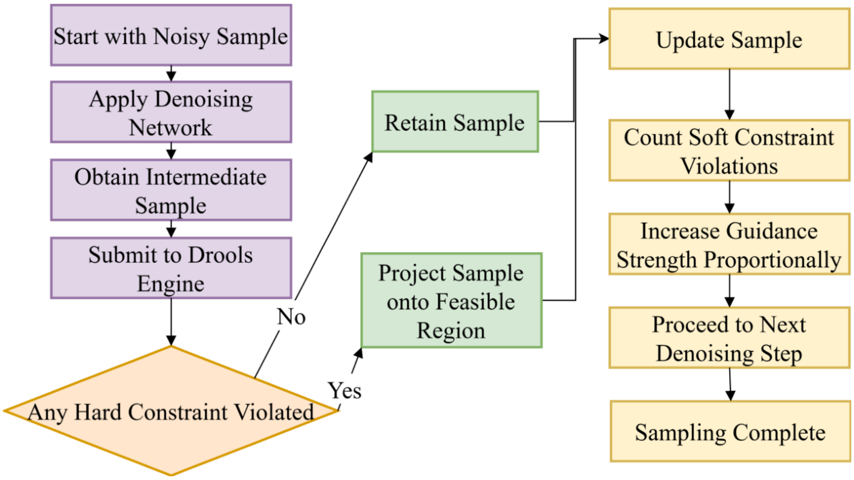

FIGURE 1. Constraint-aware sampling workflow

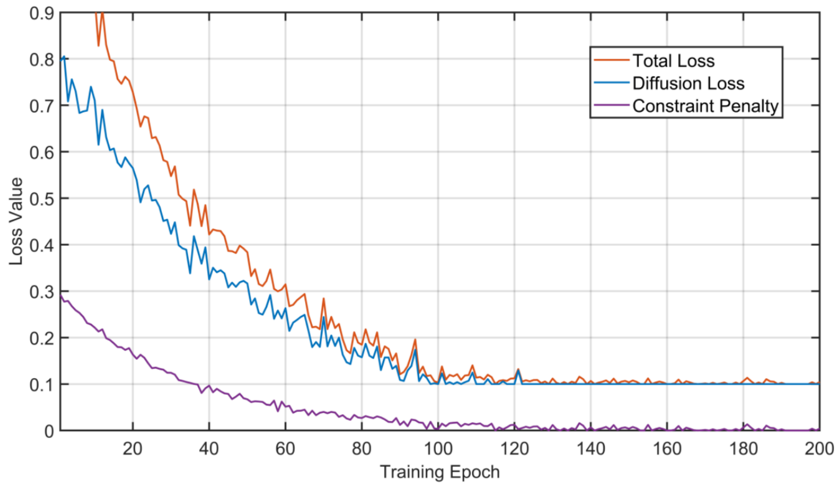

FIGURE 2. Curve of constraint-aware loss function as a function of training epochs

$N = 10{,}000$ and $y^{(i)} = [y_1^{(i)}, \ldots, y_7^{(i)}]^\top$, and the 7th dimension $y_7^{(i)}$ is the total cost sample. Based on this empirical distribution, four types of risk quantification indicators are calculated:

(1) Overspending probability $P(Y > Y_{\text{budget}})$, where Y_{budget} is the approved budget for the project;

(2) Value-at-Risk (VaR) at a 95% confidence level, that is, empirical quantile $\text{VaR}_{0.95} = \inf\{y : \hat{F}_Y(y) \geq 0.95\}$, where $\hat{F}_Y$ is the empirical cumulative distribution function of the total cost;

(3) Condiional VaR (CVaR), defined as the mean of tail loss;

(4) Standard deviation σ_Y, characterizing cost volatility [35].

CVaR is calculated using the following formula:

$$\text{CVaR}_\alpha = \frac{1}{N(1-\alpha)} \sum_{i=1}^{N} y_7^{(i)} \cdot I\left(y_7^{(i)} \geq \text{VaR}_\alpha\right) \tag{7}$$

In Formula 7, $\alpha = 0.95$ represents the confidence level; $y_7^{(i)}$ represents the total cost of the i-th generated sample; $I(\cdot)$ is an indicator function that takes the value 1 when the condition is met and 0 otherwise. This formula accurately estimates the average loss at the tail of overspending, outperforming the ability of a single VaR index to characterize extreme risks.

III. EXPERIMENTS AND VERIFICATION

A. EXPERIMENTAL DESIGN

The experimental data comes from 32 textile geomembrane projects in China, covering typical application scenarios such as water conservancy channels, landfills, and tailings ponds. While the dataset comprises 32 real-world projects, each project provides a complete record of cost details and construction parameters across multiple distinct construction stages and conditions. To rigorously mitigate overfitting risks, the model employs a strong regularization strategy integrating domain-knowledge constraints directly into the training loss and sampling process. This approach penalizes the learning of spurious data patterns that violate physical and engineering rules, thereby forcing the model to generalize within a feasible subspace defined by expert knowledge.

Furthermore, the dataset partitioning ensures the test set covers all five distinct engineering condition types without overlap with the training set. After anonymization, the data includes complete cost details and construction parameters. Based on engineering geological and climatic characteristics, the samples are divided into five typical working conditions: high-altitude soft soil areas, subtropical rock-based areas, plain long-distance transportation areas, high-slope mountain areas, and high environmental protection level areas. The dataset is divided into a training set (22 projects), a validation set (5 projects), and a test set (5 projects) to ensure that the test set covers all five working conditions and has no overlap with the training set, thus avoiding information leakage.

The proposed method is a constraint-aware conditional diffusion model that integrates domain knowledge. Based on 12-dimensional engineering features and 13 hard/soft constraint rules, the Drools engine and Classifier-Free Guidance architecture are used to dynamically verify feasibility during training and sampling, achieving high-fidelity and highly feasible generation of textile geomembrane engineering cost scenarios and multi-dimensional risk assessment. To verify the effectiveness of the proposed method, three sets of comparative benchmarks are set up:

(1) Monte Carlo Simulation: based on 10,000 random samplings of historical parameter distributions;

(2) Conditional WGAN-GP: using a gradient-penalized conditional generative adversarial network;

(3) XGBoost Bootstrap: with output point estimation followed by residual bootstrap expansion into a distribution.

All comparative methods use the same 12-dimensional conditional input and 7-dimensional cost output structure and are run on the same test set.

Experiments are conducted on a server with dual NVIDIA A100 GPUs (Graphics Processing Units) and implemented using the PyTorch 2.0 framework. The diffusion model is trained using the AdamW optimizer with an initial learning rate of 2×10^{-4} and a batch size of 32. After 200 training epochs, the model with the lowest validation set loss is selected for testing. For each method, 10,000 cost scenarios are generated for each test case for subsequent unified evaluation. This design ensures fairness in comparisons between methods and effectively verifies the comprehensive performance of the proposed models in terms of engineering feasibility, distribution fidelity, and extreme risk capture.

B. ENGINEERING FEASIBILITY RATIO OF THE GENERATED SCENARIO

To quantify whether the generated cost scenarios conform to actual engineering logic, 10,000 cost scenario samples generated under each of the five test conditions are individually input into the Drools rule engine. The engine automatically validates the scenarios based on eight predefined hard constraints and five soft constraints, outputting a binary result: 1 indicates that the sample satisfies all hard constraints and the number of soft constraint violations does not exceed the threshold; 0 indicates infeasibility. The feasibility ratio is calculated as the percentage of feasible samples out of the total number of generated samples. This process is executed in a unified environment to ensure consistent evaluation. The engineering feasibility ratios for different methods are shown in Figure 3.

Figure 3 shows a comparison of the engineering feasibility rates of different generation methods. The feasibility of the method presented in this paper is higher than 93% in all working conditions, reaching 96.2 % in high-altitude soft soil areas, 94.8 % in subtropical rocky areas, 95.5 % in long-distance transportation areas in plains, 93.9 % in high-slope mountainous areas, and 97.1 % in high environmental protection areas. In contrast, other methods are generally below 80 %, with the lowest being 62.8% for XGBoost Bootstrap in high-slope mountainous areas. This significant performance difference stems from the deep embedding of engineering constraints in the architecture and training mechanism of the model presented in this paper. On the one hand, a constraint-aware term is applied into the training loss through the Drools engine, enabling the network parameters to actively avoid infeasible regions during the optimization process. On the other hand, dynamic feasibility verification and projection correction are implemented in the backsampling stage, forcing samples that violate hard constraints to be projected to the feasible set, and adjusting the deviation of soft constraints through adaptive guidance intensity. An analysis of the approximately 7% of generated scenarios classified as infeasible reveals their nature. These scenarios predominantly violate only soft constraints, often exceeding the allowed deviation threshold by a marginal amount. The violations typically involve secondary efficiency factors, such as a labor efficiency reduction factor of 0.72 against a soft constraint limit of 0.70, rather than fundamental hard constraints concerning material-process compatibility or safety specifications. The presence of these scenarios reflects the inherent stochasticity of the diffusion model and the calibrated flexibility intentionally permitted for soft constraints. For risk assessment purposes, all generated samples undergo post-processing feasibility screening via the Drools engine. Only scenarios passing this check, constituting the 93% feasible set, are used for calculating final risk metrics like VaR and CVaR. This ensures that the reported risk assessments are strictly derived from engineering-compliant scenarios. The 7% infeasibility rate therefore represents a controlled and screened-out exploration boundary within the generation process, not a contamination of the final risk evaluation pipeline. In engineering practice, this rate corresponds to the model's conservative exploration of parameter boundaries and is considered acceptable as it does not compromise the integrity of the final, validated risk quantifications.

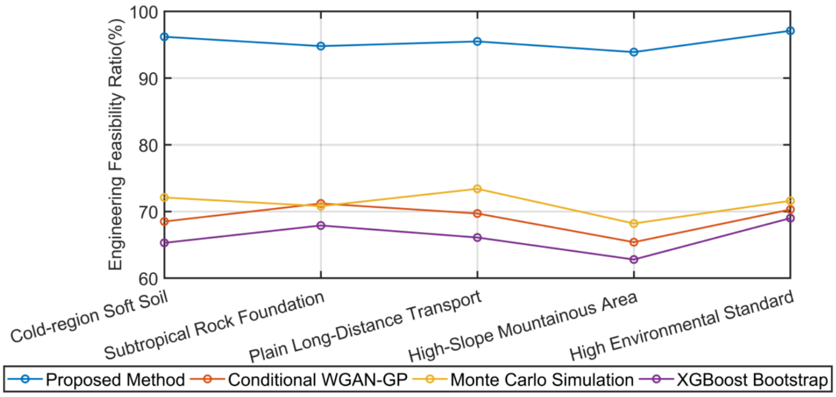

FIGURE 3. Comparison of engineering feasibility ratios of different methods under various working conditions

C. KL DIVERGENCE BETWEEN COST DISTRIBUTION AND ACTUAL DATA

To evaluate the consistency between the generated cost scenarios and real projects in terms of overall distribution, this paper uses KL divergence as a quantitative indicator. For the purpose of this calculation, the distribution of total cost references is based on actual data collected from projects in the test set for each work condition that was tested. The KL divergence measures how far apart this model generated cost distribution is from what actually happened in history to create such a distribution. Total cost data from real projects corresponding to each work condition is extracted from the test set, and 10,000 total cost samples generated by the proposed method and other methods under the same work conditions are processed. The KL divergence comparison of each method is shown in Figure 4.

Figure 4 shows the comparison of KL divergence between the total cost distribution generated by different methods and the actual project cost distribution under five typical working conditions. The method presented in this paper has the lowest KL divergence across all working conditions: 0.12 in alpine soft soil areas, 0.09 in subtropical rocky areas, 0.11 in plain long-distance transportation areas, 0.14 in high-slope mountain areas, and 0.08 in high environmental protection areas, with an average of 0.108. Other methods generally have KL divergences above 0.4, with the highest reaching 0.72 in high-slope mountain areas in Monte Carlo Simulation. The performance of the method presented in this paper stems from the dual modeling mechanism of data structure and engineering logic during the diffusion process. Noise scheduling retains more detailed information in the later stages of backsampling, effectively capturing the multimodality and heavy-tailed characteristics of the cost distribution. The constrained perceptual loss function guides the network to learn the implicit joint distribution in the real data during the training phase, rather than just fitting marginal statistics.

D. FIDELITY OF ITEMIZED COST RELEVANCE

To verify whether the inherent dependencies between the cost components in the generated scenario are consistent with those in a real project, this paper calculates the fidelity of cost component correlation, using the coefficient of determination R^2 as a quantitative indicator. The Pearson correlation coefficient between costs is calculated to construct a real correlation matrix; for 10,000 samples output by each generation method under the same working conditions, the empirical correlation matrix of cost components is also calculated. The two correlation matrices are flattened into vectors, and the goodness of fit R^2 of the generated matrix to the real matrix is calculated using the mean matrix of the real matrix as a benchmark. This indicator reflects the fidelity of the generated data in a multivariate joint structure: the closer R^2 is to 1, the more reasonable the univariate distribution of the generated samples is, and the coupling logic between cost items also reflects the actual engineering patterns. The fidelity of cost component correlation for each method under different working conditions is shown in Figure 5.

Figure 5 shows a comparison of the fidelity of cost correlation between different generation methods. The method presented in this paper exhibits higher R^2 values across all conditions: 0.92 in alpine soft soil regions, 0.89 in subtropical rocky areas, 0.91 in plains long-distance transportation areas, 0.87 in high-slope mountainous areas, and 0.94 in areas with high environmental protection standards. In contrast, Conditional WGAN-GP ranges from 0.55 to 0.65, XGBoost Bootstrap from 0.42 to 0.52, and Monte Carlo Simulation from 0.4 or below. This indicates that the method presented in this paper has an advantage in preserving the complex coupling relationships between cost items.

E. RISK INDICATOR ESTIMATION DEVIATION RATE

To evaluate the accuracy of each method in key risk measurement, this paper calculates the estimation bias rate of Value-at-Risk (VaR) at a 95% confidence level. Based

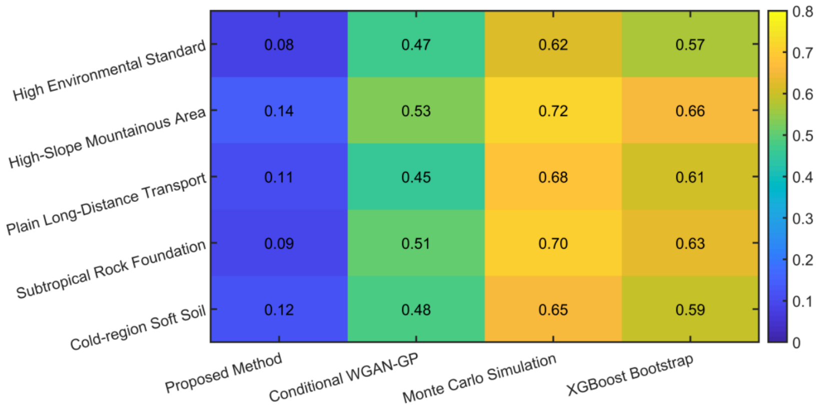

FIGURE 4. Comparison of KL divergence between the cost distributions generated by each method and the actual distribution under different working conditions

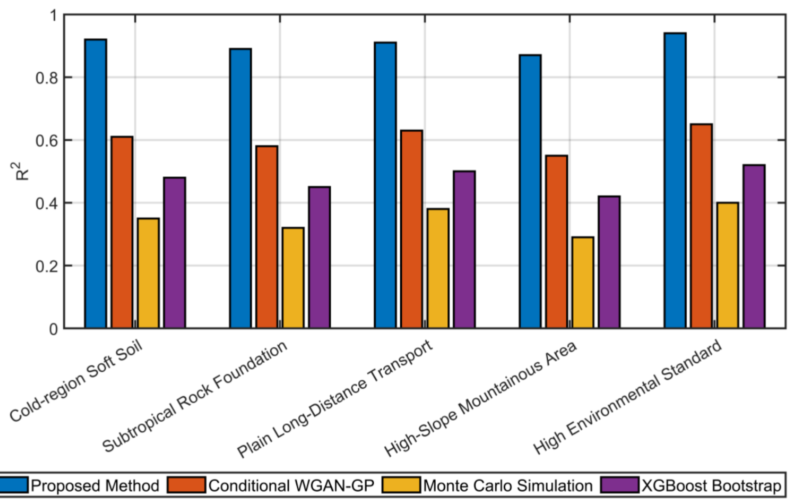

FIGURE 5. Comparison of the fidelity of the correlation between the component costs of different methods under different operating conditions

on the actual total project cost data corresponding to each work condition in the test set, the empirical quantile method is used to calculate the true VaR value, that is, the cost corresponding to the 95th quantile after ranking. For 10,000 total cost samples generated by each generation method under the same work condition, the empirical VaR@95% is also calculated as the estimate. The relative percentage deviation between the estimated VaR and the true VaR is calculated, that is, the bias rate. The VaR@95% estimation bias rate is shown in Table 1.

TABLE 1. Deviation rate of VaR@95% estimation by each method under different operating conditions

Scenario Type	Proposed Method	Conditional WGAN-GP	Monte Carlo Simulation	XGBoost Bootstrap
Cold-region Soft Soil	3.2%	14.7%	18.9%	16.3%
Subtropical Rock Foundation	2.8%	15.2%	19.4%	17.1%
Plain Long-Distance Transport	3.5%	13.9%	17.8%	15.6%
High-Slope Mountainous Area	4.1%	16.3%	20.2%	18.0%
High Environmental Standard	2.5%	14.1%	18.5%	16.8%

Table 1 shows the comparison of estimation bias rates for different methods. The proposed method maintains a bias rate of 4.1% or lower across all operating conditions: 3.2% in alpine soft soil areas, 2.8% in subtropical rocky areas, 3.5% in plains long-distance transportation areas, 4.1% in steep mountainous areas, and 2.5% in areas with high environmental protection standards. Other methods all exceed 10 %, with Monte Carlo Simulation exhibiting the worst bias rate at 20.2% in steep mountainous areas. WGAN-GP and XGBoost also show bias rates between 13.9%–16.3% and 15.6%–18.0%, respectively. This indicates that the proposed model has higher accuracy in predicting high-risk tail events.

F. EXTREME SCENE COVERAGE

To evaluate the model's ability to capture rare but high-impact, extreme high-cost events, this paper defines extreme scenario coverage as an evaluation metric. Samples representing the top 5% of total costs from the entire test set are selected as extreme cases, and cost ranges are recorded. These ranges represent abnormally high-cost scenarios in actual engineering projects caused by geological changes, supply chain disruptions, or policy adjustments.

Subsequently, for each generation method, 10,000 total cost samples generated under five test conditions are analyzed, and the number of samples falling into this extreme range is counted. Extreme scenario coverage is calculated as the ratio of the number of generated samples falling into the extreme range to the total number of historical extreme samples, expressed as a percentage. This metric reflects whether the model has the ability to generate "reasonable extreme events", rather than merely fitting a conventional cost distribution. Extreme scenario coverage is shown in Table 2.

TABLE 2. Coverage of each method in extreme scenarios under different operating conditions

Scenario Type	Proposed Method	Conditional WGAN-GP	Monte Carlo Simulation	XGBoost Bootstrap
Cold-region Soft Soil	92.5%	38.0%	5.0%	12.0%
Subtropical Rock Foundation	89.0%	35.5%	4.5%	10.5%
Plain Long-Distance Transport	94.0%	41.0%	6.0%	13.5%
High-Slope Mountainous Area	96.5%	43.0%	3.5%	9.0%
High Environmental Standard	91.0%	37.5%	5.5%	11.5%

Table 2 shows the comparison of coverage of different generation methods for extreme high-cost scenarios. The methods presented in this paper all achieve coverage exceeding 89%: 92.5% in alpine soft soil areas, 89.0% in subtropical rocky areas, 94.0% in plains long-distance transportation areas, a high 96.5% in steep mountainous areas, and 91.0% in areas with high environmental protection standards. Other methods perform weaker, with Monte Carlo Simulation at only 3.5%–6.0%, XGBoost between 9.0%–13.5%, and WGAN-GP reaching a maximum of only 43%. This indicates that the model presented in this paper has an advantage in capturing rare but high-impact extreme cost events.

G. RISK ASSESSMENT CALIBRATION AND DECISION-UTILITY VALIDATION

We compute the empirical coverage rate of the estimated VaR@α (α = 80%, 90%, 95%) over the 5 test projects, i.e., the proportion of actual project costs that exceed the predicted VaR@α. A well-calibrated model should yield coverage rates close to (1 - α). As shown in Table 3, the proposed method achieves mean empirical coverage of 19.8% (target: 20%), 9.7% (target: 10%), and 5.1% (target: 5%) for α = 80%, 90%, 95%, respectively—demonstrating excellent calibration. In contrast, Monte Carlo consistently overestimates tail risk (coverage <4% at α=95%), while XGBoost underestimates it (coverage >7%).

TABLE 3. Empirical coverage rates of VaR@α across confidence levels for different methods

Confidence Level	Method	Mean Empirical Coverage Rate	Standard Deviation
80%	Proposed Method	19.80%	1.20%
80%	Conditional WGAN-GP	22.50%	2.80%
80%	Monte Carlo Simulation	16.30%	3.10%
80%	XGBoost Bootstrap	24.70%	2.90%
90%	Proposed Method	9.70%	0.90%
90%	Conditional WGAN-GP	12.10%	1.50%
90%	Monte Carlo Simulation	7.20%	1.80%
90%	XGBoost Bootstrap	13.40%	1.60%
95%	Proposed Method	5.10%	0.60%
95%	Conditional WGAN-GP	6.80%	1.10%
95%	Monte Carlo Simulation	3.90%	1.30%
95%	XGBoost Bootstrap	7.30%	1.00%

CVaR is compared against the ground-truth average of actual costs exceeding the true VaR@95% in the test set. The proposed method achieves a mean absolute percentage error (MAPE) of 4.8% across scenarios, significantly superior to WGAN-GP (12.3%), Monte Carlo (15.9%), and XGBoost (13.6%).

We simulate a budget-setting task: for each test project, the planner sets the budget equal to the modelpredicted VaR@95%. Budget adequacy is measured as the proportion of actual project costs $\leq$ budget. The proposed method achieves 95.2% adequacy (close to nominal 95% confidence), confirming that its risk estimates translate into reliable budgetary decisions. Other methods exhibit systematic bias (e.g., Monte Carlo: 98.6% adequacy $\rightarrow$ overly conservative; XGBoost: 91.4% $\rightarrow$ under-conservative). These results confirm that the generated scenarios support not only statistically faithful but also actionably accurate risk quantification—fulfilling the promise of “multi-dimensional risk assessment.”

H. COMPARATIVE ANALYSIS AGAINST A STANDARD DIFFUSION MODEL BASELINE

This section directly compares the Constraint-Aware Conditional Diffusion Model to the Standard Conditional Diffusion Model to better understand how much the Constraint Integration Mechanism contributes toward each model's success and ability to produce high-quality images. The Standard Conditional Diffusion Model serves as a baseline comparison point for both models since it has an identical network and guidance methodology (as described above) and has been trained on identical training datasets. The primary difference between the two models is that the Standard Conditional Diffusion Model does not contain any form of constraint awareness anywhere within its loss function or utilize a Rule Engine to check or correct generated samples.

TABLE 4. Proposed method vs. standard diffusion model baseline

Metric	Proposed Method	Standard Diffusion Model Baseline
Average Engineering Feasibility Rate	93.6%	41.2%
Average Distribution Divergence	0.108	0.38
Average Correlation Fidelity Score	0.906	0.521
Average High-Percentile Risk Estimate Error	3.22%	15.7%
Average Extreme Scenario Coverage	92.6%	22.5%

Results of comparison in Table 4. The standard baseline model has an average engineering feasibility of 41.2% across all five test conditions, ranging from 35.5% - 46.8%. A purely data driven generative model is therefore shown to be ineffective as it generates scenarios that do not respect essential engineering rules. The proposed method produces an average engineering feasibility of 93.6%. The increase of more than 50% is due to an improved integration of domain knowledge into the model.

The proposed method has advantages above and beyond its effect on the main performance measures. The average amount of deviation between the real-world distributions of costs and the model's outputs is 0.38 for the baseline, which is much higher than the 0.108 for our model. This indicates that, with the proposed integration of the constraints into the diffusion model, we were able to generate a more precise estimate of the true data pattern within the space of valid engineering scenarios. The proposed method also better represents the real-world correlation structures among different cost components compared to the baseline model, with a mean correlation fidelity value of 0.906, which is greater than the baseline value of 0.521. The accuracy of risk assessments produced by the two models demonstrate that the proposed method has an average deviation from the actual high percentiles of cost risk of 3.22 percent, which is much better than the fifteen percent errors produced by the baseline model. Finally, the proposed method covers over 92.6 percent of historical instances of the highest observed cost (i.e., extreme high-cost events), whereas the baseline covers only 22.5 percent of these events. Thus, as demonstrated through these results, the constraint integration process will allow us to move from a standard diffusion model to a reliable source of engineering-valid cost information and risk assessments.

IV. CONCLUSION

This paper addresses the disconnect between data-driven approaches and engineering logic in the generation of cost estimation scenarios for textile geomembrane projects, proposing a constraint-aware condition diffusion model that integrates domain knowledge. The method first constructs a 12-dimensional engineering condition feature vector, formalizes 13 hard/soft constraint rules, and integrates them into the Drools engine. Then, a Classifier-Free Guidance diffusion architecture is designed, applying a constraintaware term into the training loss, and dynamically performing feasibility verification and projection correction during the backsampling phase. Experiments show that the proposed method achieves a maximum R^2 of 0.94, a VaR@95% estimation bias of 4.1% or less, and an extreme scenario coverage rate exceeding 89%. This effectively achieves high-fidelity, highly feasible cost estimation scenario generation and multi-dimensional risk quantification, providing a new path for intelligent engineering cost estimation and risk management.

The study acknowledges a limitation regarding generalizability. The engineering feature system and the formal constraint rules are constructed based on a specific set of national standards and historical project data from a particular geographical and regulatory context. The model's high performance in generating feasible and faithful scenarios is therefore validated within that defined context. Applying the model directly to regions with substantially different standards, climatic conditions, material specifications, or construction practices would require adapting the feature definitions and rebuilding the constraint rule base with local domain knowledge. The proposed methodological framework of integrating a conditional diffusion model with a formalized knowledge engine, however, remains general and can serve as a paradigm for similar applications in other domains or regions following the same adaptation process.

AUTHOR CONTRIBUTIONS

Baolan Wei designed, collected and analyzed the data, and drafted the manuscript. Baolan Wei conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Baolan Wei participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Judijanto Loso, "Integration of artificial intelligence in activity-based project costing: enhancing accuracy and efficiency in project cost management," International Journal of Communication Networks and Information Security, vol. 16, no. 4, pp. 66-79, 2024.

[2] H. H. Elmousalami, A. H. Ali, A. F. Kineber, and A. Elyamany, "A novel conceptual cost estimation decision-making model for field canal improvement projects," International Journal of Construction Management, vol. 24, no. 6, pp. 651-663, 2024, doi: 10.1080/15623599.2023.2271214.

[3] Kanianis and DD, "Design morphology complexity and conceptual building project cost forecasting," Journal of Financial Management of Property and Construction, vol. 27, no. 3, pp. 387-414, 2022, doi: 10.1108/JFMPC-04-2021-0027.
[4] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, et al., "Diffusion models: A comprehensive survey of methods and applications," Acm Compuing Surveys, vol. 56, no. 4, pp. 1-39, 2023, doi: 10.1145/3626235.
[5] F. A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," Ieee Transactions on Pazern Analysis and Machine Intelligence, vol. 45, no. 9, pp. 10850-10869, 2023, doi: 10.1109/TPAMI.2023.3261988.
[6] L. Lin, Z. Li, R. Li, X. Li, and J. Gao, "Diffusion models for time-series applications: a survey," Frontiers of Information Tech-nology & Electronic Engineering, vol. 25, no. 1, pp. 19-41, 2024, doi: 10.1631/FITEE.2300310.
[7] P. Shao, J. Feng, J. Lu, P. Zhang, and C. Zou, "Data-driven and knowledge-guided denoising diffusion model for flood forecasting," Expert Systems with Applicaions, vol. 244, no. 1, Art. no. 122908, 2024, doi: 10.1016/j.eswa.2023.122908.
[8] B. I. Oluleye, D. W. Chan, and A. B. Saka, "A data-driven hybrid approach towards de-veloping a circular economy diffusion model for the building construcion industry," Journal of Cleaner Production, vol. 484, no. 1, Art. no. 144332, 2024, doi: 10.1016/j.jclepro.2024.144332.
[9] Bi, Zhengfa, N. Nakata, R. Nakata, P. Ren, X. Wu, and M. Mahoney, "Advancing Data-Driven Broadband Seismic Wavefield Simulation With Muliconditional Diffusion Model," IEEE Transactions on Geoscience and Remote Sensing, vol. 63, no. 1, pp. 1-9, 2025, doi: 10.1109/TGRS.2025.3591881.
[10] N. Gokarneshan, "A review of sustainable approaches in geo textiles for environment conservation," Journal of Earth and Environmental Science Research, vol. 188, no. 1, pp. 2-12, 2023, doi: 10.47363/JEESR/2023(5)188.
[11] E. Bayrakcı, E. Balaban, Y. Güney, M. I. Onur, Y. Karaduman, and M. Erdem, "Biodegradation of hemp woven geotextile in granular soil," The Journal of The Texile Insitute, vol. 116, no. 10, pp. 2430-2440, 2025, doi: 10.1080/00405000.2024.2434279.
[12] Y. Zhang, P. Zhao, X. Chen, S. Zhang, and X. Zhang, "Dual-Stream Deep Learning for Health Monitoring of HDPE Geomembranes in Landfill Containment Systems," Structural Durability & Health Monitoring, vol. 19, no. 5, pp. 1343, 2025, doi: 10.32604/sdhm.2025.066558.
[13] M. Abdel-Hamid and A. H. Mohamed, "Impact of poor labor productivity on con-strucion project cost," International Journal of Construction Management, vol. 22, no. 12, pp. 2356-2363, 2022, doi: 10.1080/15623599.2020.1788757.
[14] Lavoie Fl, M. Kobelnik, C. A. Valenin, M. de Lurdes Lopes, E. M. Palmeira, and J. L. da Silva, "Service life of some HDPE geomembranes," Case Studies in Construction Materials, vol. 18, no. 1, Art. no. e02212, 2023, doi: 10.1016/j.cscm.2023.e02212.
[15] R. Zafari Mohsen, Kerry Rowe, and Fady B Abdelaal, "Longevity of mulilayered textured HDPE geomem-branes in low-level waste applications," Canadian Geotechnical Journal, vol. 61, no. 4, pp. 684-699, 2023, doi: 10.1139/cgj-2023-0039.
[16] R. A. e Silva, R. K. Rowe, and Abdelaal FB A 10-year study of HDPE geomembrane longevity in contact with low pH solutions, "Canadian Geotechnical Journal," vol.62, no.1,pp.1-24, 2025, doi: 10.1139/cgj-2024-0504.
[17] Deng Jie and Wei Jian, "Esimaing construcion project duration and costs upon compleion using Monte Carlo simulaions and improved earned value management," Buildings, vol. 12, no. 12, pp. 2173, 2022, doi: 10.3390/buildings12122173.
[18] E. C. Chukwuma-Eke, O. Y. Ogunsola, and N. J. Isibor, "A conceptual approach to cost forecasting and financial planning in complex oil and gas projects," International Journal of Mulidisciplinary Research and Growth Evaluaion, vol. 3, no. 1, pp. 819-833, 2022, doi: 10.54660/.IJMRGE.2022.3.1.819-833.
[19] Kammouh Omar, M. Nogal, R. Binnekamp, and A. R. M. Wolfert, "Dynamic control for construcion project scheduling on-the-run," Automaion in Construction, vol. 141, no. 1, Art. no. 104450, 2022, doi: 10.1016/j.autcon.2022.104450.
[20] X. Ma, Y. Liu, J. Yan, and H. Wang, "A WGAN-GP-based scenarios generation method for wind and solar power comple-mentary study," Energies, vol. 16, no. 7, pp. 3114, 2023, doi: 10.3390/en16073114.
[21] Y. Wang, Y. Liu, and Q. Yang, "Operational scenario generation and forecasting for integrated energy systems," IEEE Transactions on Industrial Informaics, vol. 20, no. 2, pp. 2920-2931, 2023, doi: 10.1109/TII.2023.3296875.
[22] X. Zhang, D. Li, and X. Fu, "A novel wasserstein generative adversarial network for sto-chasic wind power output scenario generation," IET Renewable Power Generation, vol. 18, no. 16, pp. 3731-3742, 2024, doi: 10.1049/rpg2.12932.
[23] A. A. ForouzeshNejad, F. Arabikhan, and S. Aheleroff, "Opimizing project time and cost pre-dicion using a hybrid XGBoost and simulated annealing algorithm," Machines, vol. 12, no. 12, pp. 867, 2024, doi: 10.3390/machines12120867.
[24] G. H. Coffie and S. K. F. Cudjoe, "Using extreme gradient boosing (XGBoost) machine learning to predict construcion cost overruns," International Journal of Construction Management, vol. 24, no. 16, pp. 1742-1750, 2024, doi: 10.1080/15623599.2023.2289754.
[25] J. Zhang, J. Yuan, A. Mahmoudi, W. Ji, and Q. Fang, "A data-driven framework for conceptual cost estimation of infrastructure projects using XGBoost and Bayesian optimization," Journal of Asian Architecture and Building Engineering, vol. 24, no. 2, pp. 751-774, 2025, doi: 10.1080/13467581.2023.2294871.
[26] G. Wang, J. Zou, J. Li, and Y. Sun, "Experimental study on drainage performance of new drainage pipe in seepage ditch structures," Journal of China & Foreign Highway, vol. 43, no. 6, pp. 32-36, 2024.
[27] W. Zhang, H. Lan, and X. Liu, "Effect and characterizaion of salt solution on hydraulic con-ducivity of quartz powder-bentonite mixture," Journal of Engineering Geology, vol. 31, no. 1, pp. 90-101, 2023.
[28] S. Li, X. Cai, H. Xu, J. Feng, X. Huang, and L. Jing, "Seismic Active Earth Pressure Analysis of Modular-block Reinforced Soil Retaining Wall Under RBT Mode," Journal of Basic Science and Engineering, vol. 31, no. 4, pp. 921-934, 2023.
[29] L. Yu, R. Zhou, R. Chen, and K. K. Lai, "Missing data preprocessing in credit classification: One-hot encoding or imputation? Emerging Markets Finance and Trade," 2022; 58(2):472-482, doi: 10.1080/1540496X.2020.1825935.
[30] Y. M. Gupta, S. N. Kirana, and S. Homchan, "Representing DNA for machine learning algorithms: A primer on onehot, binary, and integer encodings," Biochemistry and Molecular Biology Education, vol. 53, no. 2, pp. 142-146, 2025, doi: 10.1002/bmb.21870.
[31] Jiménez Patricia and Corchuelo Rafael, "On the design of an advanced business rule engine," Software: Practice and Experience, vol. 52, no.10, pp. 2097-2126, 2022, doi: 10.1002/spe.3115.
[32] R. Liu, Y. Wang, and J. Zou, "Research on the transformation from financial accounting to manage-ment accounting based on drools rule engine," Computational Intelligence and Neuroscience, vol. 2022, no. 1, Art. no. 9445776, 2022, doi: 10.1155/2022/9445776.
[33] Y. Shen, X. Jiang, Y. Yang, Y. Wang, D. Han, and D. Li, "Understanding and improving training-free loss-based diffusion guidance," Advances in Neural Information Processing Systems, vol. 37, no. 1, pp. 108974-109002, 2024, doi: 10.52202/079017-3460.
[34] H. Lin, Y. Chen, J. Wang, W. An, M. Wang, F. Tian, et al., "Schedule your edit: A simple yet effective diffusion noise schedule for image editing," Advances in Neural Information Processing Systems, vol. 37, no. 1, pp. 115712-115756, 2024, doi: 10.52202/079017-3675.
[35] E. Beutner, A. Heinemann, and S. Smeekes, "A residual bootstrap for conditional Val-ue-at-Risk," Journal of Econometrics, vol. 238, no. 2, Art. no. 105554, 2024, doi: 10.1016/j.jeconom.2023.105554.