

AI-Based Dynamic Resource Allocation in Construction Projects: A Reinforcement Learning Approach to Time–Cost–Resource Optimization

Mohammad. Elahi¹

¹ M.Sc., Department of Civil Engineering, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran

* Corresponding author email address: mhmdehahi99@gmail.com

Article Info

Article type:

Original Research

How to cite this article:

Elahi, M., (2026). AI-Based Dynamic Resource Allocation in Construction Projects: A Reinforcement Learning Approach to Time–Cost–Resource Optimization. *Journal of Resource Management and Decision Engineering*, 5(5), 1-19.

<https://doi.org/10.61838/kman.jrmde.453>



© 2026 the authors. Published by KMAN Publication Inc. (KMANPUB). This is an open access article under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

ABSTRACT

This study aimed to develop and evaluate an artificial intelligence–based dynamic resource-allocation framework using reinforcement learning to simultaneously optimize project duration, cost performance, and resource utilization in construction projects. A quantitative computational study was conducted using data from 84 medium- and large-scale construction projects in Tehran, Iran, together with validation input from 168 construction professionals. Project-level data included activity schedules, labor and equipment allocation, progress, costs, resource conflicts, idle time, and operational constraints. A Proximal Policy Optimization reinforcement learning model was developed to dynamically allocate resources at weekly decision intervals. The model state incorporated schedule, cost, activity, and resource indicators, while the reward function jointly represented time, cost, and resource-efficiency objectives. Projects were divided at the project level into training, validation, and independent test subsets. Model performance was compared with historical managerial allocation and a conventional rule-based scheduling strategy. Robustness, sensitivity, and expert-validation analyses were also conducted. PPO-based allocation significantly reduced project completion time compared with historical allocation (mean difference = -2.68 months, 95% CI [-3.26, -2.10], $p < .001$, $d = 1.02$) and rule-based allocation (-1.13 months, 95% CI [-1.48, -0.78], $p < .001$, $d = 0.71$). Schedule deviation was significantly reduced by 7.69 percentage points relative to historical allocation ($p < .001$, $d = 1.56$) and by 3.25 percentage points relative to rule-based allocation ($p < .001$, $d = 0.98$). Cost deviation decreased by 5.83 and 2.42 percentage points, respectively (both $p < .001$). Labor utilization significantly increased by 12.94 percentage points versus historical allocation ($p < .001$, $d = 1.91$) and 6.22 percentage points versus rule-based allocation ($p < .001$, $d = 1.28$), while equipment utilization also improved significantly. Resource conflicts were significantly lower under PPO than both comparison strategies (both $p < .001$). Reinforcement learning provided a feasible, robust, and statistically superior approach to dynamic construction resource allocation, enabling simultaneous improvements in schedule performance, cost control, and resource utilization.

Keywords: artificial intelligence; reinforcement learning; Proximal Policy Optimization; construction project management; resource allocation; time–cost optimization; project scheduling

1. Introduction

Construction project management is fundamentally concerned with achieving a viable balance among time, cost, scope, quality, and the availability of constrained resources. Among these dimensions, the allocation and reallocation of labor, equipment, materials, and financial resources are particularly critical because resource-related decisions directly influence activity durations, project expenditure, productivity, and the probability of schedule disruption. Construction projects are also highly dynamic systems in which the original resource plan is continuously challenged by changes in site conditions, workforce productivity, equipment availability, supply-chain interruptions, design modifications, and unexpected interactions among concurrent activities. Conventional planning methods generally assume that resource requirements and activity conditions can be sufficiently anticipated at the planning stage; however, actual construction environments frequently deviate from these assumptions. Consequently, fixed or periodically updated schedules may become increasingly ineffective as projects progress. Artificial intelligence has therefore attracted growing attention as a mechanism for improving the responsiveness, predictive capacity, and decision quality of construction management systems. Recent developments in AI-powered civil engineering demonstrate a transition from isolated predictive applications toward integrated, lifecycle-oriented systems capable of supporting planning, scheduling, monitoring, optimization, and decision-making across complex engineering environments (Xu & Guo, 2025). Systematic evidence on artificial intelligence in project scheduling similarly indicates that machine learning, evolutionary optimization, and intelligent decision systems are increasingly being applied to resource-constrained scheduling problems that are difficult to address through deterministic procedures alone (Bahroun et al., 2023). Within prefabricated and digitally enabled construction environments, deep learning techniques have also been used to process complex project information and support more intelligent planning and management practices (Guo, 2024). These advances suggest that construction resource allocation is increasingly evolving from a static optimization task toward a dynamic, data-driven decision problem.

Resource allocation in construction can be conceptualized as a constrained multi-objective optimization problem because improvements in one project dimension may create undesirable effects in another. Accelerating

critical activities may reduce project duration but increase labor or equipment costs, whereas minimizing expenditure may lead to insufficient resource availability and consequently prolong project completion. Similarly, maximizing utilization can create congestion or reduce flexibility when unexpected events occur. Traditional resource-constrained project scheduling approaches have therefore attempted to identify efficient combinations of activity sequencing, resource assignments, and cost structures. A wide range of optimization methods have been developed for this purpose. Improved artificial immune system algorithms, for example, have been applied to optimize manpower allocation in construction engineering, demonstrating that metaheuristic search can improve the distribution of limited human resources across project activities (Huang & Bai, 2023). Ant colony optimization has likewise been used to improve complex project management by searching for more efficient scheduling solutions under multiple constraints (Sun, 2025). More recently, multi-skill project scheduling research has incorporated both the depth and breadth of worker skills into resource-constrained optimization, illustrating that realistic scheduling requires consideration not merely of the number of available workers but also their heterogeneous competencies and task suitability (Fawzy et al., 2026). Hybrid particle swarm optimization has similarly been applied to resource-constrained scheduling in aircraft maintenance, showing the value of adaptive search mechanisms in highly constrained and interdependent project environments (Zeng et al., 2026). Although these approaches can generate effective schedules, many are still designed primarily for offline optimization and must be rerun when project conditions change, which limits their suitability for continuous, real-time resource reallocation.

The limitations of static scheduling are particularly important because modern construction projects generate continuously changing information regarding progress, resource availability, expenditure, and operational conditions. Building Information Modeling, sensor-enabled monitoring, digital project-management platforms, and integrated planning systems have substantially expanded the availability of real-time and near-real-time project data. Integration of BIM with planning techniques such as Line of Balance has been shown to provide a more detailed foundation for human resource planning in building construction processes by linking work packages, spatial progress, and labor requirements (Olaya et al., 2026). Intelligent BIM-based models have also been developed for

multi-objective optimization of engineering assembly processes, demonstrating that digital representations of project systems can support integrated assessment of scheduling, resource, and production objectives (Zhang et al., 2025). At the same time, research bridging project management and supply-chain management has emphasized the need for optimization frameworks that consider interconnected project and logistics decisions rather than treating project scheduling as an isolated function (Zhang, 2025). This is particularly relevant in construction, where material delivery, subcontractor availability, equipment sharing, and workforce deployment are interdependent. Risk-oriented scheduling research in highway construction further demonstrates that project schedules must account for uncertainty, disruption, and risk exposure in addition to deterministic activity durations (Zhasmukhambetova et al., 2025). Taken together, these developments reinforce the need for resource-allocation models that can repeatedly update decisions as the state of a project changes rather than generating only a single preconstruction solution.

Reinforcement learning is especially suitable for this class of problems because it formulates decision-making as a sequential interaction between an agent and an environment. At each decision point, an agent observes the current state, selects an action, receives feedback in the form of a reward or penalty, and learns a policy that maximizes cumulative long-term performance. Unlike conventional supervised learning, reinforcement learning does not require an explicit optimal label for every observed project state. Instead, it learns from the consequences of actions, making it particularly appropriate for situations in which the value of a resource-allocation decision depends not only on its immediate effect but also on subsequent project conditions. The theoretical compatibility between reinforcement learning and dynamic resource allocation has been demonstrated in multi-objective environments modeled using Markov decision processes, where resource decisions must be adapted according to uncertain and evolving system states (Yang & Wu, 2023). Reinforcement learning has also been applied to adaptive design strategies, showing that agents can learn responsive policies by interacting with changing design environments rather than relying exclusively on predetermined rules (Chen & Chen, 2024). In complex service-resource systems, intelligent scheduling based on reinforcement learning has similarly been used to allocate limited resources dynamically in response to changing demand (Yang, 2026). These studies indicate that reinforcement learning offers a suitable computational

foundation for construction resource allocation because construction projects can naturally be represented as sequences of states, resource-management actions, and project-performance consequences.

Deep reinforcement learning further expands this capability by combining sequential learning with nonlinear function approximation. Construction projects may involve hundreds of activities, multiple resource categories, complex precedence relationships, variable productivity, and changing cost conditions, producing state spaces that are difficult to represent using conventional tabular reinforcement learning. Deep neural networks enable agents to approximate the value or policy functions associated with such high-dimensional environments. Similar problems have been successfully addressed in several engineering and scheduling domains. Deep reinforcement learning has been applied to time-sensitive and resource-aware workflow scheduling in edge computing, where tasks must be assigned dynamically under simultaneous timing and resource constraints (Zhang et al., 2023). Two-stage deep reinforcement learning has also been used in edge-enabled systems to coordinate complex scheduling decisions across interconnected resources (Zhou et al., 2023). Dueling Double Deep Q-Network methods have been adopted for joint berth allocation and operation problems in container terminals, demonstrating the effectiveness of deep reinforcement learning for sequential resource-allocation decisions involving multiple competing tasks and capacity constraints (Li et al., 2023). In vehicular networks, hierarchical and entropy-adaptive reinforcement learning has been developed for multi-task offloading, providing evidence that adaptive policy structures can improve resource management under dynamic network conditions (Wang et al., 2026). Although the operational contexts differ from construction, the computational structure of these problems is similar: a limited resource pool must be distributed among competing tasks while system conditions evolve over time.

Applications in production and manufacturing provide additional evidence for the relevance of intelligent sequential scheduling to construction management. Machine-learning-based multilevel scheduling has been used for multi-variety and variable-batch production environments, where changing product mixes and production requirements create substantial scheduling complexity (Liang Wang et al., 2023). Graph attention reinforcement learning has recently been introduced for collaborative scheduling in sheet-metal workshops, allowing

the scheduling model to learn relational information between jobs and resources rather than considering tasks independently (Gao et al., 2026). Agile software project scheduling has similarly been treated as a complex multi-team optimization problem in which interdependent tasks and shared resources must be coordinated using advanced swarm-based learning algorithms (Lou et al., 2025). Scheduling research in distributed computing also reveals the importance of jointly considering task urgency, resource requirements, and system constraints. Energy-efficient workflow scheduling algorithms in cloud computing, for instance, seek to allocate computational resources while balancing execution time, energy consumption, and service requirements (Verma et al., 2023), whereas multi-task scheduling mechanisms in federated learning must simultaneously manage distributed tasks, computing resources, and security constraints (Liu et al., 2023). Comparative analyses of scheduling algorithms in vehicular fog computing likewise demonstrate that allocation strategies can vary markedly in their ability to balance resource use, latency, and operational efficiency (Khan et al., 2024). These cross-domain findings support the proposition that dynamic construction scheduling may benefit from methods capable of learning interactions among tasks and resources rather than relying on fixed priority rules.

Recent research has moved closer to the direct application of reinforcement learning in construction and infrastructure project environments. Multiagent reinforcement learning has been proposed for construction resource scheduling and cost control in intelligent construction sites, illustrating the potential for multiple intelligent agents to coordinate distributed resource decisions while simultaneously considering financial performance (Li et al., 2026). Multi-agent deep reinforcement learning has also been applied to dynamic cost-overflow mitigation in smart-grid construction projects, providing evidence that sequential AI-based control can adapt project decisions when expenditure patterns begin to deviate from planned values (Li & Zhang, 2026). Resource prediction research has further moved from conventional statistical methods toward graph neural networks capable of representing project networks and learning relationships between activities, time, and cost (Mirjalili, 2025). Deep learning-based intelligent resource management has also been examined in renewable energy engineering projects, reinforcing the broader applicability of AI for coordinating complex resource requirements in engineering environments (Yao, 2026). The growing use of intelligent resource systems across engineering sectors

suggests that construction project management is entering a phase in which prediction and optimization can increasingly be integrated rather than implemented as separate functions.

Multi-agent approaches are especially relevant when project systems contain multiple decision entities or resource classes that must operate jointly. Multi-agent deep reinforcement learning has been used for intelligent resource allocation in green sports events, where competing operational needs must be balanced while maintaining global system performance (Luo, 2025). Real-time sensing and resource-management systems for infrastructure operations similarly show how resource allocation can be linked to continuously updated environmental information, enabling responsive rather than purely scheduled decision-making (Dai et al., 2026). These approaches are conceptually compatible with construction sites, where labor crews, equipment fleets, subcontractors, and material resources function as partially independent but strongly interdependent subsystems. However, increasing model complexity also creates challenges related to coordination, interpretability, computational stability, and the risk that locally optimal agent behavior may not yield globally efficient project outcomes. For many practical construction applications, a centralized reinforcement learning policy may therefore provide a useful intermediate solution by integrating all relevant resource conditions into a common state representation while maintaining explicit control over allowable actions and operational constraints.

A further challenge concerns the interaction among project time, cost, and resources. Many existing scheduling studies focus predominantly on one outcome, even though construction managers must usually optimize several objectives simultaneously. Precast production scheduling research has shown that off-site construction scheduling involves interdependent decisions concerning production sequence, resource allocation, delivery coordination, and project timing (Long Wang et al., 2023). Broader reviews of resource allocation in integrated networks similarly emphasize that effective resource management requires multidimensional optimization under heterogeneous demand and fluctuating system conditions (Liang et al., 2024). Dynamic resource-allocation strategies based on fuzzy multi-objective optimization and Markov decision processes likewise demonstrate that decision policies should account for trade-offs across multiple competing objectives (Yang & Wu, 2023). In project environments, these trade-offs are particularly important because the fastest schedule is not necessarily the least costly, while maximum resource

utilization may not be desirable if it eliminates the flexibility needed to absorb disruption. Consequently, a reinforcement learning framework for construction should not reward schedule acceleration alone but should use a reward structure that simultaneously represents time reduction, cost control, efficient resource utilization, and avoidance of infeasible or excessively disruptive allocations.

The transferability and robustness of learned policies are also critical. A model trained on one project or simulation environment may not automatically perform well when activity structures, cost conditions, productivity patterns, or resource availability change. Transfer-learning research has demonstrated approaches for moving learned policies from parameterized or simulated environments toward more realistic physical scenarios, indicating that intelligent decision systems can be designed to reduce the gap between training and deployment conditions (Zhang et al., 2024). Adaptive resource-management research also suggests that learning systems should be exposed to diverse operating conditions and uncertainty so that they do not merely memorize a narrow set of training scenarios. This is particularly important in construction, where no two projects have identical activity networks, resource constraints, or operational disruptions. Reinforcement learning therefore requires project-level separation between training and testing, stochastic variation during simulation, and explicit examination of performance under reduced labor productivity, equipment shortages, or material disruptions. Without these procedures, apparently favorable results may reflect overfitting rather than a genuinely transferable allocation policy.

Another important consideration is the integration of optimization with realistic managerial constraints. Construction decisions cannot be treated as unrestricted mathematical actions because they must satisfy technological precedence, contractual commitments, crew continuity, resource capacity, site safety, and practical feasibility. Advanced scheduling methods increasingly recognize this requirement by incorporating increasingly detailed operational restrictions into optimization models. The multi-skill resource-constrained project scheduling literature demonstrates that the feasibility of an assignment depends on the correspondence between task requirements and worker competencies rather than numerical resource availability alone (Fawzy et al., 2026). Similarly, production-scheduling models based on machine learning must operate within physical and procedural constraints that define permissible task sequences (Liang Wang et al., 2023).

Realistic construction reinforcement learning should therefore employ constrained action spaces, action masking, or penalty structures to prevent recommendations that appear computationally attractive but cannot be implemented. Such an approach allows the agent to search for superior decisions only within the operationally feasible region of the project-control problem.

Despite substantial progress, an important methodological gap remains. Existing studies demonstrate the effectiveness of AI for project scheduling, optimization, forecasting, BIM-enabled planning, and dynamic resource management, but comparatively few frameworks integrate these capabilities into a unified construction-specific system that dynamically reallocates resources while simultaneously optimizing schedule, cost, and utilization. Some research emphasizes predictive modeling, whereas other studies focus on metaheuristic search, static optimization, or domain-specific reinforcement learning. There is consequently a need for models that connect actual project records with sequential decision-making and assess whether learned allocation policies outperform both historical managerial decisions and conventional rule-based scheduling. Such evaluation should extend beyond cumulative reinforcement learning reward and examine operational indicators directly relevant to construction management, including project duration, schedule deviation, cost deviation, labor utilization, equipment utilization, idle time, resource conflicts, and the frequency of constraint violations. Moreover, robustness should be tested under alternative resource conditions and different managerial priorities rather than only under a single deterministic scenario. This requirement becomes increasingly important as project management and supply-chain optimization become more closely integrated and construction systems become more dependent on dynamic data and automated decision support (Zhang, 2025).

The increasing sophistication of intelligent optimization techniques therefore creates a strong foundation for developing a reinforcement learning-based construction resource-management framework capable of responding dynamically to project conditions. AI-enabled civil engineering, intelligent BIM, deep learning, metaheuristic scheduling, graph-based prediction, multi-agent systems, and adaptive reinforcement learning collectively demonstrate that complex engineering decisions can increasingly be represented, learned, and optimized computationally (Mirjalili, 2025; Xu & Guo, 2025; Zhang et al., 2025). At the same time, evidence from dynamic

engineering and infrastructure systems highlights the value of continuously updating resource decisions when demand, availability, and operating conditions change (Dai et al., 2026; Wang et al., 2026). The practical challenge is therefore not simply to determine an optimal initial schedule but to create an intelligent control mechanism capable of repeatedly reallocating resources as project conditions evolve. Such a mechanism should learn from the long-term consequences of resource decisions, preserve feasibility, remain robust to uncertainty, and explicitly balance time, cost, and resource efficiency.

Accordingly, the aim of this study was to develop and evaluate an AI-based dynamic resource-allocation framework using reinforcement learning to simultaneously optimize project time, cost, and resource utilization in construction projects.

2. Methods and Materials

This study employed a quantitative, applied, computational research design to develop and evaluate an artificial intelligence-based dynamic resource allocation framework for simultaneous optimization of project time, cost, and resource utilization in construction projects. The study was conducted in Tehran, Iran, and focused on medium- and large-scale building construction projects characterized by multiple concurrent activities, shared resource pools, variable labor requirements, equipment constraints, and measurable scheduling and cost information. The empirical dataset was obtained from 84 ongoing and recently completed construction projects implemented in different districts of Tehran. Projects were eligible for inclusion when they had a formally approved baseline schedule and budget, maintained systematic records of labor, equipment, material, activity progress, and actual expenditure, used computerized project planning or Building Information Modeling systems, and had completed at least 70% of their planned physical progress at the time of data extraction. Projects with incomplete cost records, substantial undocumented scope changes, prolonged suspension, or insufficient resource-use data were excluded. In addition to project-level records, 168 construction professionals, consisting of two knowledgeable representatives from each participating project, were included to verify the accuracy of extracted information and evaluate operational constraints incorporated into the optimization environment. Participants included project managers, planning and scheduling engineers, cost-control engineers, site engineers, and

construction managers who had at least five years of professional experience in construction project management and direct knowledge of the selected projects. Purposive sampling was initially applied to identify eligible construction companies and projects, followed by criterion-based selection of professionals with sufficient knowledge of project scheduling, cost management, and resource allocation. For each project, longitudinal information was collected at weekly intervals, including planned and actual activity progress, labor deployment, equipment utilization, material availability, activity duration, direct and indirect expenditure, schedule deviations, resource conflicts, and approved corrective actions. These sequential observations were transformed into state-action-reward-next-state transitions suitable for reinforcement learning. Project-level separation was maintained during model development so that records originating from a given project could not simultaneously appear in training and evaluation subsets, thereby minimizing information leakage and providing a more realistic assessment of model generalizability.

Data were collected using an integrated project-data extraction protocol designed to transform heterogeneous construction management records into a standardized analytical dataset. The primary data sources consisted of baseline and updated project schedules, cost-control reports, resource-loaded activity schedules, workforce attendance records, equipment utilization logs, procurement and material-delivery records, weekly progress reports, change-order documentation, earned-value records, and information exported from project management and BIM-based information systems where available. A structured Project Time-Cost-Resource Data Extraction Form was developed for the study to ensure consistent coding across projects. The form recorded project characteristics, activity identifiers, predecessor and successor relationships, planned and actual start and finish dates, baseline and remaining duration, percentage of physical completion, available and required labor by skill category, equipment availability and utilization, material constraints, direct activity costs, indirect project costs, overtime expenditure, resource productivity, schedule variance, cost variance, and occurrence of resource shortages or conflicts. Numerical variables were standardized to project-specific scales where necessary to permit learning across projects of different sizes, while categorical variables such as activity type and resource category were encoded before model development. Data extracted from organizational records were cross-checked against weekly progress reports and verified by the two

designated professionals from each project. Discrepancies were resolved through reconciliation with the original project records rather than through statistical imputation whenever documentary evidence was available.

The reinforcement learning environment constituted the principal computational instrument of the study. Each decision point represented the resource-allocation status of a construction project at a weekly control interval. The state space included normalized project completion, activity progress, remaining duration, critical-path status, schedule variance, cost variance, remaining budget, current labor allocation, equipment utilization, available resource capacity, activity resource demand, resource productivity, activity precedence constraints, and indicators of resource congestion or shortage. The action space represented feasible managerial adjustments to resource allocation, including maintaining the current resource configuration, reallocating workers among eligible activities, increasing or decreasing labor assigned to selected activities, reallocating equipment, introducing allowable overtime, and prioritizing resources for critical or near-critical activities. Actions that violated precedence requirements, maximum resource capacities, technological constraints, or predefined safety and operational limits were masked and excluded from selection. The reward function was constructed as a multi-objective signal combining schedule performance, cost performance, and resource-use efficiency. Positive rewards were assigned to actions that reduced projected completion delay, controlled expenditure, increased productive resource utilization, or resolved resource conflicts, whereas penalties were assigned for schedule overruns, budget overruns, excessive overtime, resource idleness, abrupt allocation changes, and infeasible resource assignments. Because optimization of one construction objective can adversely affect another, normalized weighted components were employed so that the agent learned policies based on the combined time–cost–resource consequences of each decision rather than optimizing any single outcome independently.

A deep reinforcement learning model based on Proximal Policy Optimization was developed as the primary allocation agent because the resource-allocation problem involved sequential decisions, stochastic project conditions, multiple interacting state variables, and constrained managerial actions. The policy and value functions were represented using multilayer artificial neural networks receiving the project state vector as input and estimating the expected value and probability of feasible allocation actions. The

environment was configured episodically, with an episode representing progression from a defined project state to project completion. During training, project conditions were updated after each action according to observed productivity patterns, resource constraints, activity dependencies, and cost consequences obtained from the empirical dataset. To improve model robustness and reduce dependence on a particular sequence of historical observations, controlled stochastic variation was introduced into labor productivity, equipment availability, activity duration, and resource-demand parameters within ranges derived from the observed project data. A conventional rule-based allocation strategy based on critical-path priority and available-resource assignment was implemented as the principal benchmark. Historical managerial allocation decisions were also retained as a second reference condition so that the AI-based policy could be evaluated against both algorithmic baseline scheduling and actual project practice.

A structured expert-validation questionnaire was additionally administered to the 168 participating professionals to assess whether the resource constraints, feasible actions, relative importance of performance objectives, and model-generated allocation recommendations were operationally realistic. Items were rated on a five-point Likert scale ranging from 1, indicating strongly disagree, to 5, indicating strongly agree. The questionnaire covered the realism of the project-state variables, appropriateness of resource-allocation actions, adequacy of operational constraints, interpretability of recommendations, practical feasibility, and perceived usefulness of the proposed framework for project control. Content validity was assessed before final implementation through review by specialists in construction management, project scheduling, cost control, and artificial intelligence. Internal consistency was subsequently examined using Cronbach's alpha, with a coefficient of 0.70 or higher considered indicative of acceptable reliability. Expert evaluation was used primarily to establish the practical plausibility of the computational framework and was not incorporated into the reinforcement learning reward signal, thereby avoiding direct contamination of model performance by subjective ratings.

Data analysis was conducted in sequential preprocessing, reinforcement-learning development, optimization, validation, and comparative-performance stages. Initially, project records were examined for duplicate observations, temporal inconsistencies, impossible resource assignments, and missing values. Missing values representing less than

5% of a continuous variable were imputed using project-specific median values or temporal interpolation when the variable exhibited an interpretable time sequence, whereas observations with extensive missing information in core scheduling, cost, or resource variables were excluded. Continuous input variables were normalized to comparable scales, and categorical variables were transformed into numerical representations suitable for model processing. To prevent data leakage, the 84 projects were randomly divided at the project level into training, validation, and independent testing subsets comprising approximately 70%, 15%, and 15% of projects, respectively. All observations from an individual project were retained within the same subset. The training subset was used for policy learning, the validation subset for hyperparameter selection and early evaluation, and the independent testing subset exclusively for final performance assessment. Repeated training runs with different random seeds were conducted to evaluate the stability of the learned allocation policy and reduce the possibility that observed performance resulted from a favorable initialization.

The reinforcement learning agent was trained to maximize cumulative discounted reward over the remaining project horizon. At every decision interval, the agent received the current project-state vector, selected a feasible resource-allocation action, observed the subsequent project state and associated cost, schedule, and resource effects, and updated its policy accordingly. The PPO optimization procedure used clipped policy updates to limit excessively large changes between successive policies and thereby improve learning stability. Hyperparameters, including learning rate, discount factor, generalized advantage estimation parameter, clipping threshold, batch size, number of optimization epochs, and neural-network architecture, were tuned using validation-set performance. Training convergence was assessed through stabilization of mean episodic reward and the absence of systematic improvement across successive evaluation intervals. Constraint violations were continuously recorded, and infeasible actions were suppressed through action masking and additional penalty mechanisms. Sensitivity analyses were subsequently performed by altering the relative weights assigned to time, cost, and resource-utilization components of the reward function to determine whether the conclusions remained stable under alternative managerial priorities. Additional scenarios introduced fluctuations in labor productivity, equipment availability, and material accessibility to assess

the resilience of the learned policy under uncertain construction conditions.

The primary performance outcomes were total project completion time, percentage schedule deviation from the baseline, total project cost, percentage cost deviation, cumulative delay, labor utilization rate, equipment utilization rate, resource idle time, number of detected resource conflicts, number of constraint violations, and a composite time–cost–resource performance index. For each project in the independent test set, outcomes generated by the reinforcement learning policy were compared with those produced by the rule-based benchmark and the historically observed resource-allocation strategy. Relative improvement was calculated as the percentage change in each outcome compared with the corresponding benchmark. Because identical project scenarios were evaluated under competing allocation policies, paired statistical comparisons were used. Distributional assumptions were assessed before inferential analysis; paired-samples *t* tests were applied when difference scores approximated normality, whereas Wilcoxon signed-rank tests were used when this assumption was not satisfied. Effect sizes and 95% confidence intervals were reported alongside probability values to evaluate the magnitude and precision of observed differences. Repeated-scenario simulation was further used to compare policies under stochastic variations in project conditions, and mean performance, standard deviation, and confidence intervals across simulation runs were calculated. Statistical significance was defined as $p < .05$.

Model performance was also examined from a multi-objective optimization perspective. Pareto-efficient solutions were identified to characterize the trade-off between project duration, total cost, and resource utilization, allowing the proposed framework to distinguish genuinely efficient allocation policies from solutions that improved one objective only by disproportionately worsening another. The stability of the policy was evaluated across projects of different size, baseline duration, budget, and resource intensity, while robustness analysis examined its response to simulated resource shortages, productivity reductions, and equipment disruptions. Finally, responses to the expert-validation questionnaire were analyzed using descriptive statistics and internal-consistency analysis to determine whether the recommendations generated by the reinforcement learning framework were considered technically interpretable and practically implementable by construction professionals. The combination of independent project-level testing, benchmark comparisons, sensitivity

analysis, stochastic simulation, multi-objective assessment, and expert validation was used to determine whether the proposed AI-based dynamic resource-allocation framework could provide a reliable improvement in construction time–cost–resource performance rather than merely achieving higher reward values within the training environment.

3. Findings and Results

The empirical analysis included all 84 construction projects that satisfied the inclusion criteria and the 168 construction professionals assigned to verify project information and assess the operational validity of the proposed resource-allocation framework. The participating professionals had a mean age of 39.64 years ($SD = 7.81$), with ages ranging from 27 to 58 years. Of the 168 participants, 132 (78.6%) were men and 36 (21.4%) were women. With respect to professional position, 39 participants (23.2%) were project managers, 42 (25.0%) were planning and scheduling engineers, 35 (20.8%) were cost-control engineers, 29 (17.3%) were site engineers, and 23 (13.7%) were construction managers. Mean professional experience was 12.74 years ($SD = 5.61$), and 126

participants (75.0%) had more than eight years of experience in construction management or project control. Regarding educational attainment, 42 participants (25.0%) held a bachelor's degree, 105 (62.5%) held a master's degree, and 21 (12.5%) held a doctoral degree. The 84 construction projects had a mean baseline duration of 24.73 months ($SD = 8.64$), with a range of 11 to 48 months, and their mean baseline physical completion at the beginning of the analytical observation period was 76.18% ($SD = 7.21$). Thirty-one projects (36.9%) were residential developments, 23 (27.4%) were commercial or mixed-use developments, 18 (21.4%) were institutional or public-service buildings, and 12 (14.3%) were large office or administrative projects. Mean project workforce at peak operation was 164.27 workers ($SD = 73.46$), while the projects contained an average of 127.54 scheduled activities ($SD = 41.33$). The mean recorded schedule deviation prior to optimization was 11.86% ($SD = 6.39$), and the mean cost deviation from the approved baseline was 9.72% ($SD = 5.84$), indicating that the analyzed projects exhibited sufficient variability and resource-allocation challenges to provide an appropriate empirical environment for evaluating dynamic reinforcement learning.

Table 1

Descriptive Statistics of Project Time, Cost, Resource, and Control Variables

Variable	Mean	SD	Minimum	Maximum
Baseline project duration (months)	24.73	8.64	11.00	48.00
Baseline project cost (billion IRR)	1,284.61	692.37	286.40	3,925.70
Number of scheduled activities	127.54	41.33	58.00	241.00
Peak workforce size	164.27	73.46	61.00	386.00
Mean weekly labor utilization (%)	76.82	8.17	56.70	91.60
Mean equipment utilization (%)	71.35	9.46	49.30	89.80
Baseline schedule deviation (%)	11.86	6.39	1.80	28.70
Baseline cost deviation (%)	9.72	5.84	0.90	25.40
Weekly resource-conflict frequency	4.18	1.76	1.00	9.00
Mean labor idle time (%)	13.67	5.21	4.30	27.80
Mean equipment idle time (%)	16.24	6.08	5.10	31.70
Schedule Performance Index	0.89	0.08	0.68	1.04
Cost Performance Index	0.91	0.07	0.71	1.05
Resource productivity index	0.74	0.09	0.51	0.92

As shown in Table 1, the projects displayed considerable heterogeneity in scale, duration, financial value, activity complexity, and resource conditions, providing an appropriate test environment for evaluating the generalizability of the reinforcement learning model. The average baseline duration was 24.73 months, while the substantial standard deviation of 8.64 months indicated the inclusion of both medium-duration and comparatively long

construction projects. Baseline project costs likewise varied considerably, ranging from IRR 286.40 billion to IRR 3,925.70 billion. Resource-use indicators showed that neither labor nor equipment resources were being used at their full productive potential under conventional project-control practices. Mean labor utilization was 76.82%, whereas equipment utilization averaged 71.35%, accompanied by labor and equipment idle-time rates of

13.67% and 16.24%, respectively. The mean weekly resource-conflict frequency of 4.18 further indicated recurrent competition among project activities for limited labor, machinery, or equipment capacity. The mean Schedule Performance Index of 0.89 and Cost Performance Index of 0.91 were both below 1.00, demonstrating that the typical project in the sample was simultaneously experiencing schedule and cost performance deficiencies. The mean resource productivity index of 0.74 provided

additional evidence that significant scope existed for improving the timing and distribution of available resources. Collectively, these descriptive findings suggested that the projects were not characterized merely by isolated time or cost overruns, but by interconnected deficiencies in resource utilization, scheduling, and expenditure control, thereby supporting the use of a sequential multi-objective optimization approach rather than a conventional single-objective scheduling procedure.

Table 2

Training, Validation, and Testing Performance of the PPO Reinforcement Learning Model

Performance Indicator	Training Set	Validation Set	Independent Test Set
Number of projects	59	12	13
Mean episodic reward	82.64	78.92	77.38
SD of episodic reward	5.31	6.17	6.54
Feasible action rate (%)	99.48	99.21	99.08
Constraint violation rate (%)	0.52	0.79	0.92
Mean schedule deviation after optimization (%)	3.84	4.26	4.41
Mean cost deviation after optimization (%)	3.51	3.92	4.08
Mean labor utilization (%)	90.67	89.84	89.31
Mean equipment utilization (%)	87.91	86.73	86.42
Mean composite performance index	0.873	0.854	0.847
Coefficient of variation of reward across repeated runs (%)	4.96	5.78	6.11

The reinforcement learning results presented in Table 2 demonstrated that the PPO model achieved stable and highly feasible allocation behavior across the training, validation, and independent testing datasets. The mean episodic reward was highest in the training set at 82.64 and declined moderately to 78.92 in the validation set and 77.38 in the independent test set. This limited decrease indicated an expected generalization gap but provided no indication of severe overfitting. The feasible action rate exceeded 99% across all three datasets, while the constraint violation rate remained below 1%, including a value of only 0.92% in independent testing. These results indicated that the model did not improve time or cost performance by routinely proposing operationally impossible allocations. In the independent test projects, schedule deviation was reduced to 4.41%, while cost deviation declined to 4.08%. Labor utilization increased to 89.31% and equipment utilization to 86.42%, representing substantial improvements relative to

the pre-optimization sample averages reported in Table 1. The composite performance index was 0.847 in the test sample compared with 0.873 in the training sample, corresponding to a comparatively small reduction in multi-objective performance when the model was applied to completely unseen projects. Furthermore, the coefficient of variation of episodic reward remained only 6.11% in independent testing across repeated training and simulation runs, suggesting that the policy was not excessively sensitive to stochastic initialization or random project fluctuations. The convergence analysis showed that mean cumulative reward increased rapidly during the early learning stage, improved more gradually in the intermediate stage, and reached a stable plateau during later iterations. Once convergence was reached, no material increase in validation reward was observed, supporting termination of training and selection of the final PPO policy for comparative evaluation.

Table 3

Comparison of Historical Allocation, Rule-Based Allocation, and PPO-Based Dynamic Resource Allocation

Outcome	Historical Allocation	Rule-Based Allocation	PPO-Based Allocation	Improvement of PPO vs. Historical (%)	Improvement of PPO vs. Rule-Based (%)
Project completion time (months)	27.66	26.11	24.98	9.69	4.33
Schedule deviation (%)	11.86	7.42	4.17	64.84	43.80
Total cost deviation (%)	9.72	6.31	3.89	59.98	38.35
Labor utilization (%)	76.82	83.54	89.76	16.84	7.45
Equipment utilization (%)	71.35	79.63	86.89	21.78	9.12
Labor idle time (%)	13.67	9.48	5.62	58.89	40.72
Equipment idle time (%)	16.24	11.07	6.84	57.88	38.21
Weekly resource conflicts	4.18	2.74	1.39	66.75	49.27
Overtime intensity (%)	12.31	10.18	8.76	28.84	13.95
Composite time–cost–resource index	0.632	0.741	0.856	35.44	15.52

Table 3 demonstrates a consistent performance advantage of PPO-based dynamic resource allocation over both historically observed management decisions and the conventional rule-based allocation strategy. Mean simulated project completion time decreased from 27.66 months under the historical allocation condition to 26.11 months under the rule-based approach and 24.98 months under PPO-based optimization. Relative to historical allocation, the reinforcement learning framework therefore produced a 9.69% improvement in completion time. The strongest effects were observed for schedule deviation, cost deviation, and resource conflicts. Mean schedule deviation decreased from 11.86% under historical allocation to 4.17% under the PPO strategy, representing a 64.84% relative reduction, while mean cost deviation declined from 9.72% to 3.89%, corresponding to a 59.98% improvement. These findings are particularly important because the reduction in schedule

delays did not occur at the expense of higher resource consumption or uncontrolled expenditure. Instead, labor utilization increased from 76.82% to 89.76%, and equipment utilization increased from 71.35% to 86.89%. At the same time, labor idle time decreased by 58.89%, equipment idle time declined by 57.88%, and the average number of weekly resource conflicts fell from 4.18 to 1.39. The model also reduced overtime intensity to 8.76%, indicating that its time-saving effect was primarily achieved through more efficient sequencing and resource redistribution rather than indiscriminate extension of working hours. The composite time–cost–resource index increased from 0.632 under historical allocation to 0.741 under the rule-based strategy and 0.856 under the PPO strategy, showing that reinforcement learning achieved superior simultaneous performance across all three principal optimization dimensions.

Table 4

Inferential Comparison of PPO-Based Allocation with Benchmark Strategies

Outcome	Comparison	Mean Difference	95% CI	Test Statistic	p	Effect Size
Project completion time	PPO vs. Historical	-2.68 months	-3.26 to -2.10	t = -9.34	< .001	d = 1.02
Project completion time	PPO vs. Rule-Based	-1.13 months	-1.48 to -0.78	t = -6.47	< .001	d = 0.71
Schedule deviation	PPO vs. Historical	-7.69 percentage points	-8.76 to -6.62	t = -14.29	< .001	d = 1.56
Schedule deviation	PPO vs. Rule-Based	-3.25 percentage points	-3.97 to -2.53	t = -8.96	< .001	d = 0.98
Cost deviation	PPO vs. Historical	-5.83 percentage points	-6.79 to -4.87	t = -12.08	< .001	d = 1.32
Cost deviation	PPO vs. Rule-Based	-2.42 percentage points	-3.09 to -1.75	t = -7.19	< .001	d = 0.78
Labor utilization	PPO vs. Historical	+12.94 percentage points	11.47 to 14.41	t = 17.53	< .001	d = 1.91
Labor utilization	PPO vs. Rule-Based	+6.22 percentage points	5.16 to 7.28	t = 11.71	< .001	d = 1.28
Equipment utilization	PPO vs. Historical	+15.54 percentage points	13.71 to 17.37	t = 16.91	< .001	d = 1.85
Equipment utilization	PPO vs. Rule-Based	+7.26 percentage points	5.91 to 8.61	t = 10.72	< .001	d = 1.17
Resource conflicts	PPO vs. Historical	-2.79 conflicts/week	-3.21 to -2.37	Z = -7.61	< .001	r = .83
Resource conflicts	PPO vs. Rule-Based	-1.35 conflicts/week	-1.68 to -1.02	Z = -6.88	< .001	r = .75
Composite performance index	PPO vs. Historical	+0.224	0.198 to 0.250	t = 17.18	< .001	d = 1.87
Composite performance index	PPO vs. Rule-Based	+0.115	0.094 to 0.136	t = 10.93	< .001	d = 1.19

The inferential comparisons reported in Table 4 confirmed that the observed descriptive advantages of the PPO-based strategy were statistically significant and practically substantial. Compared with historical allocation, PPO reduced mean completion time by 2.68 months, 95% CI [-3.26, -2.10], $t = -9.34$, $p < .001$, with a large standardized effect size of $d = 1.02$. A significant reduction of 1.13 months was also observed relative to the rule-based benchmark, $t = -6.47$, $p < .001$, $d = 0.71$. Schedule deviation showed an even larger improvement, with PPO outperforming historical decisions by 7.69 percentage points, $t = -14.29$, $p < .001$, $d = 1.56$, and the rule-based strategy by 3.25 percentage points, $t = -8.96$, $p < .001$, $d = 0.98$. Similarly, PPO reduced cost deviation by 5.83 percentage points relative to historical allocation and 2.42 percentage points relative to rule-based allocation, with both comparisons reaching $p < .001$. Resource-efficiency outcomes were also characterized by large effects. Labor

utilization increased by 12.94 percentage points compared with historical management practice, $d = 1.91$, and by 6.22 percentage points compared with rule-based allocation, $d = 1.28$. Equipment utilization improved by 15.54 and 7.26 percentage points, respectively, with effect sizes of $d = 1.85$ and $d = 1.17$. Because weekly resource-conflict counts were positively skewed, Wilcoxon signed-rank tests were used for this outcome. PPO produced significantly fewer conflicts than both the historical condition, $Z = -7.61$, $p < .001$, $r = .83$, and the rule-based strategy, $Z = -6.88$, $p < .001$, $r = .75$. Finally, the composite optimization index improved by 0.224 relative to historical allocation and by 0.115 relative to the rule-based strategy, with both differences statistically significant and associated with large effect sizes. These findings demonstrate that the PPO policy generated not only statistically detectable but also operationally meaningful improvements across the principal dimensions of construction project performance.

Table 5

Robustness, Sensitivity, and Multi-Objective Performance of the PPO Allocation Policy

Scenario	Schedule Deviation (%)	Cost Deviation (%)	Labor Utilization (%)	Equipment Utilization (%)	Composite Performance Index
Base scenario	4.17	3.89	89.76	86.89	0.856
Labor productivity -10%	5.31	4.42	88.91	86.17	0.823
Labor productivity -20%	6.74	5.38	87.46	85.11	0.784
Equipment availability -15%	5.68	4.63	89.02	82.74	0.806
Material availability disruption	6.12	5.17	87.83	84.96	0.792
Combined resource disruption	7.46	6.28	85.79	81.63	0.748
Time-priority reward weighting	3.36	4.77	88.84	85.73	0.842
Cost-priority reward weighting	4.91	3.18	88.13	85.44	0.838
Resource-efficiency priority	4.63	4.06	91.37	89.12	0.851
Balanced multi-objective weighting	4.17	3.89	89.76	86.89	0.856

The robustness and sensitivity analyses summarized in Table 5 showed that the performance advantages of the reinforcement learning model remained evident under a range of adverse project conditions and alternative managerial priorities. Under the base scenario, schedule deviation was 4.17%, cost deviation was 3.89%, and the composite performance index reached 0.856. A 10% reduction in labor productivity increased schedule deviation to 5.31% and reduced the composite index to 0.823, while a more severe 20% reduction increased schedule deviation to 6.74% and reduced the composite index to 0.784. The policy also maintained comparatively strong performance when

equipment availability was reduced by 15% and when material availability was disrupted, with composite performance indices of 0.806 and 0.792, respectively. The most demanding scenario involved simultaneous labor, equipment, and material disruption. Even in this condition, the composite index remained 0.748, which was higher than the 0.632 value obtained under historical allocation in normal conditions. This finding indicates that the reinforcement learning policy retained a substantial portion of its optimization advantage even when several sources of operational uncertainty occurred simultaneously. The reward-weighting sensitivity analysis further demonstrated

that the framework was capable of adapting to alternative managerial objectives. When time minimization received greater weight, schedule deviation declined to 3.36%, although cost deviation increased moderately to 4.77%. When cost control was prioritized, cost deviation fell to 3.18%, accompanied by a schedule deviation of 4.91%. Giving greater emphasis to resource efficiency raised labor and equipment utilization to 91.37% and 89.12%, respectively. Nevertheless, the balanced reward

configuration achieved the highest overall composite index of 0.856, indicating that simultaneous optimization of time, cost, and resource performance provided a superior global solution compared with disproportionate emphasis on any individual objective. Examination of the simulated Pareto frontier likewise showed that the balanced PPO solution was located close to the efficient frontier and that further substantial improvement in one objective generally required a measurable deterioration in at least one other objective.

Figure 1

Learning Convergence and Comparative Time–Cost–Resource Performance of the PPO-Based Dynamic Resource Allocation Framework Across Training Episodes and Benchmark Allocation Strategies

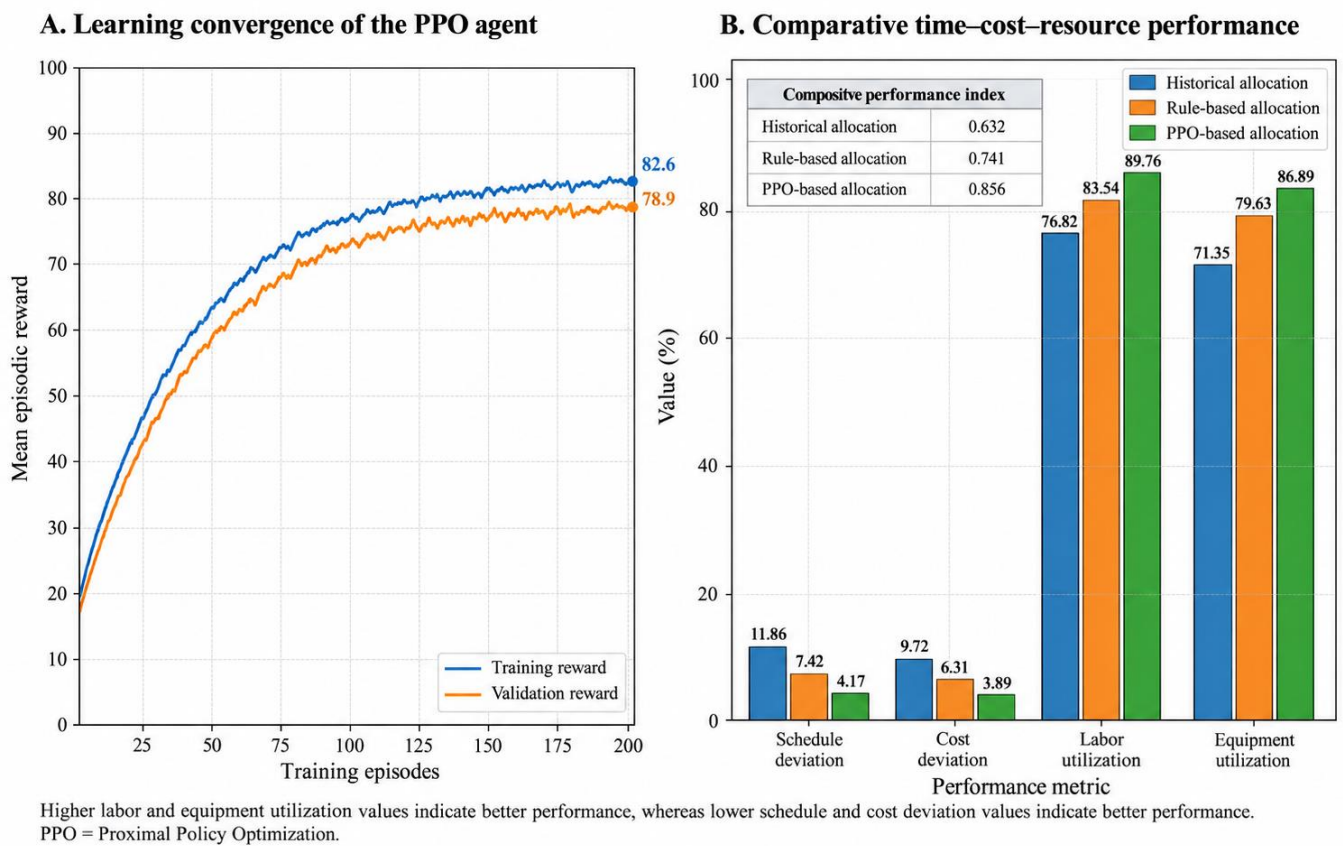


Figure 1 illustrates the learning behavior of the proposed reinforcement learning framework together with the comparative performance of the final PPO policy against the historical and rule-based allocation conditions. Mean episodic reward increased sharply during the initial phase of training as the agent learned to avoid inefficient resource configurations, unnecessary idle time, and actions associated with schedule or cost penalties. The rate of improvement subsequently became more gradual, and the reward curve approached a stable plateau during the later training

episodes, while validation performance followed a similar trajectory with only a modest gap from the training curve. This pattern supports the conclusion that the learning process converged without evidence of substantial instability or progressive overfitting. The comparative portion of the figure further shows that the PPO-based policy occupied the most favorable region of the time–cost–resource performance space, with lower schedule deviation and lower cost deviation combined with higher labor and equipment utilization. Historical project-management decisions were

concentrated in the least favorable performance region, characterized by comparatively high delays, cost deviations, and resource underutilization, whereas the rule-based benchmark produced intermediate performance. The separation between the PPO and benchmark strategies was maintained across repeated stochastic simulation runs, and the relatively narrow dispersion around PPO estimates indicated stable performance under changing project conditions. Taken together, the graphical results reinforce the tabular findings by demonstrating that the proposed reinforcement learning framework learned a stable dynamic allocation policy and consistently shifted construction-project performance toward more efficient combinations of completion time, expenditure, and resource use.

In addition to the computational performance analysis, the professional validation results supported the practical applicability of the proposed framework. The overall expert-validation questionnaire demonstrated high internal consistency, with Cronbach's alpha equal to .91. The highest mean rating was obtained for the relevance of the state variables used to represent actual construction conditions ($M = 4.41$, $SD = 0.52$), followed by the appropriateness of the operational constraints embedded in the model ($M = 4.36$, $SD = 0.55$), practical usefulness of dynamic allocation recommendations ($M = 4.29$, $SD = 0.61$), and technical feasibility of implementing the recommendations in ongoing construction projects ($M = 4.18$, $SD = 0.66$). Interpretability of the model-generated allocation decisions received a somewhat lower but still favorable mean score of 4.07 ($SD = 0.69$). Overall, 87.5% of the participating professionals rated the framework as either useful or highly useful for improving project-control decisions, while 82.7% indicated that the recommended allocation actions were realistically implementable within normal construction management practice. These results suggest that the model's superior computational performance was accompanied by a high level of professional acceptance and that its recommendations were generally consistent with the operational constraints encountered by experienced construction personnel. Taken together, the findings demonstrate that dynamic PPO-based resource allocation substantially improved project completion time, cost control, labor and equipment utilization, and conflict reduction compared with both historical managerial practice and conventional rule-based scheduling, while preserving feasibility, robustness, and operational relevance under alternative project conditions.

4. Discussion and Conclusion

The present study developed and evaluated an AI-based dynamic resource-allocation framework using Proximal Policy Optimization to optimize time, cost, and resource utilization in construction projects. The findings showed that the reinforcement learning model achieved stable learning behavior, high action feasibility, low constraint violation, and superior performance compared with both historical project-management decisions and a conventional rule-based allocation strategy. In the independent test set, the PPO framework reduced schedule deviation to 4.41% and cost deviation to 4.08%, while labor and equipment utilization increased to 89.31% and 86.42%, respectively. In the full comparative analysis, PPO-based allocation reduced mean project completion time from 27.66 months under historical allocation to 24.98 months, decreased schedule deviation from 11.86% to 4.17%, and reduced cost deviation from 9.72% to 3.89%. At the same time, labor utilization increased from 76.82% to 89.76%, equipment utilization increased from 71.35% to 86.89%, and weekly resource conflicts declined from 4.18 to 1.39. These findings indicate that reinforcement learning did not improve one objective at the expense of the others; rather, it produced a more favorable global balance among project duration, cost control, and resource efficiency. This result is consistent with the growing body of work showing that artificial intelligence can support increasingly integrated project-management decisions by learning relationships among scheduling, resource allocation, and operational performance (Bahroun et al., 2023; Xu & Guo, 2025). It also supports the broader transition from static optimization toward adaptive project-control systems in which decisions are revised as project conditions evolve.

The significant reduction in completion time and schedule deviation is one of the most important findings of the study. The PPO policy reduced mean project completion time by 2.68 months compared with historical allocation and by 1.13 months compared with rule-based allocation, with both differences being statistically significant. Schedule deviation declined by 7.69 percentage points relative to historical management practice and by 3.25 percentage points relative to the rule-based benchmark. These large effects can be explained by the sequential nature of reinforcement learning. Unlike conventional scheduling methods, which typically optimize a project plan at a defined point in time, reinforcement learning repeatedly evaluates the current project state and selects actions on the basis of

their expected long-term consequences. This mechanism allows the agent to identify emerging resource bottlenecks before they produce larger delays and to redirect resources toward critical or near-critical activities. Similar advantages have been reported in deep reinforcement learning studies involving time-sensitive workflow scheduling, where adaptive policies improved task sequencing and resource responsiveness under changing system conditions (Zhang et al., 2023). Deep reinforcement learning has also been effective in two-stage scheduling environments where decisions must be continually adjusted across interconnected task and resource structures (Zhou et al., 2023). In construction-related scheduling, the importance of dynamic optimization is further supported by research on precast production and complex project networks, where changing production conditions and resource interactions substantially influence schedule performance (Mirjalili, 2025; Long Wang et al., 2023). Therefore, the schedule improvements observed in the present study likely resulted from the model's ability to treat project control as an ongoing sequential decision process rather than as a one-time optimization problem.

The reduction in cost deviation provides additional evidence that the model learned to improve project efficiency rather than simply accelerate completion through increased expenditure. PPO reduced cost deviation by 5.83 percentage points compared with historical allocation and by 2.42 percentage points compared with the rule-based strategy. This is an important result because schedule acceleration in construction is often achieved through overtime, additional labor, or intensive equipment deployment, all of which can increase direct costs. In the present study, however, overtime intensity decreased while schedule performance improved, suggesting that the reinforcement learning agent relied primarily on more efficient redistribution and timing of existing resources. This result is consistent with recent applications of multi-agent reinforcement learning to construction resource scheduling and cost control, where AI-based agents were used to coordinate resource decisions while considering their financial consequences (Li et al., 2026). A related study of smart-grid construction projects demonstrated that multi-agent deep reinforcement learning can dynamically mitigate cost overruns by adapting decisions as project expenditure deviates from the planned trajectory (Li & Zhang, 2026). More broadly, deep learning-based resource management in engineering projects has shown that data-driven allocation can support more responsive expenditure control in complex

systems (Yao, 2026). The present findings extend these results by showing that cost reduction can be achieved simultaneously with shorter project duration and higher resource utilization, thereby supporting a genuinely multi-objective construction-management strategy.

The substantial increases in labor and equipment utilization further demonstrate the operational value of dynamic allocation. Labor utilization increased by 12.94 percentage points relative to historical practice and by 6.22 percentage points relative to the rule-based approach, while equipment utilization increased by 15.54 and 7.26 percentage points, respectively. At the same time, labor idle time and equipment idle time decreased considerably. These findings indicate that the PPO policy was able to identify periods in which resources were underused and to reassign them to activities where their marginal contribution to project performance was greater. This interpretation is consistent with previous optimization studies in construction engineering showing that intelligent manpower allocation can reduce inefficient resource use and improve workforce deployment (Huang & Bai, 2023). Research integrating BIM and Line of Balance has similarly emphasized the importance of aligning labor allocation with spatial and temporal patterns of construction progress (Olaya et al., 2026). In addition, multi-skill project scheduling studies have shown that resource efficiency depends not only on the amount of available labor but also on the correspondence between worker capabilities and task requirements (Fawzy et al., 2026). The present reinforcement learning framework expands this logic by making repeated resource decisions over time and evaluating their cumulative effect on the project rather than assigning resources according to fixed priorities.

The strong reduction in resource conflicts is another notable finding. Weekly conflicts declined from 4.18 under historical allocation to 1.39 under PPO-based optimization, corresponding to a reduction of approximately 67%. This suggests that the reinforcement learning policy was able to recognize competing demands for shared labor or equipment and resolve them more effectively than both conventional managerial decisions and rule-based scheduling. Resource conflicts are particularly important in construction because they can generate cascading delays when several activities depend on the same crew, machine, or specialized resource. Similar scheduling challenges are encountered in other complex systems, including manufacturing, logistics, and distributed computing. Graph attention reinforcement learning has been used to improve collaborative scheduling

in industrial workshops by explicitly modeling relationships between jobs and resource constraints (Gao et al., 2026). Machine-learning-based multilevel scheduling has also been used to coordinate production under variable batch conditions and heterogeneous task demands (Liang Wang et al., 2023). In distributed computing environments, intelligent scheduling mechanisms have been developed to manage multiple tasks competing for limited resources under changing conditions (Liu et al., 2023; Verma et al., 2023). The present results suggest that construction resource conflicts have a comparable structural form and can therefore benefit from adaptive AI-based coordination.

The high feasible-action rate and very low constraint-violation rate are particularly important from the perspective of practical implementation. The PPO agent achieved feasible action rates above 99% across training, validation, and testing, while constraint violations remained below 1%. This indicates that the model did not obtain favorable performance by proposing unrealistic or technically impossible resource combinations. The use of action masking and explicit project constraints ensured that the agent learned within the feasible operational space. This feature addresses an important concern in AI-based optimization, namely that mathematically attractive solutions may not always comply with physical, technological, or managerial requirements. The importance of constraint-aware optimization has been demonstrated in multi-skill resource-constrained scheduling, where assignment feasibility depends on worker competencies and task-specific requirements (Fawzy et al., 2026). Similar issues arise in dynamic resource-allocation problems based on Markov decision processes, where the effectiveness of an adaptive policy depends on maintaining feasibility while balancing several objectives (Yang & Wu, 2023). Reinforcement learning applications in adaptive environments likewise demonstrate that policy learning is most useful when the agent operates within clearly defined state and action constraints (Chen & Chen, 2024). Accordingly, the very low violation rate observed in this study strengthens the practical relevance of the model and distinguishes it from purely theoretical optimization solutions.

The convergence findings further support the methodological stability of the proposed framework. Mean episodic reward increased rapidly during early training, followed by a gradual improvement and eventual plateau, while validation reward remained close to training performance. The modest decline from training to

independent testing also suggested acceptable generalization rather than overfitting. This is especially important in construction because project structures differ considerably in duration, cost, activity network complexity, and resource intensity. The stable performance across unseen projects suggests that the policy learned generalizable allocation patterns rather than memorizing individual project trajectories. Transfer-learning research has similarly emphasized the need for intelligent models to retain performance when moving from parameterized or simulated environments to more realistic operational conditions (Zhang et al., 2024). Deep reinforcement learning applications in highly dynamic scheduling environments have also highlighted the importance of adaptive architectures capable of maintaining performance across changing system states (Wang et al., 2026). The present results therefore provide preliminary evidence that reinforcement learning can support transferable construction project-control policies when data splitting is performed at the project level and when stochastic variability is introduced during model training.

The robustness analysis provides further evidence for the model's adaptability. Even under reduced labor productivity, equipment shortages, material disruptions, and combined resource constraints, the PPO policy maintained a composite performance index above the historical allocation benchmark. The balanced reward configuration produced the highest overall composite score, while time-priority, cost-priority, and resource-efficiency configurations generated predictable shifts in their corresponding performance outcomes. This finding illustrates the flexibility of reinforcement learning as a multi-objective decision-support mechanism. Managers may alter reward weights to reflect project-specific priorities without completely redesigning the optimization framework. Similar multi-objective resource-allocation strategies have been proposed using Markov decision processes, where dynamic policies are adjusted according to competing performance criteria (Yang & Wu, 2023). Broader resource-allocation research also emphasizes that high-performing systems must balance heterogeneous objectives rather than optimize a single metric in isolation (Liang et al., 2024). In project management, optimization methods that connect scheduling with supply-chain and resource decisions likewise indicate that project performance depends on coordinated treatment of multiple interdependent objectives (Zhang, 2025). The present results therefore support the use of balanced reward

structures when the managerial goal is simultaneous improvement in time, cost, and resource efficiency.

The superiority of the balanced PPO configuration also has implications for the traditional time–cost trade-off in construction management. The findings showed that placing excessive emphasis on one dimension led to predictable deterioration in another: time-priority weighting reduced schedule deviation but increased cost deviation, while cost-priority weighting improved expenditure performance at the expense of schedule efficiency. Resource-efficiency prioritization improved utilization but did not produce the highest overall performance. These results illustrate the existence of a Pareto-type frontier in which improvements beyond a certain point require sacrificing performance in another dimension. Such trade-offs are consistent with multi-objective optimization research in engineering assembly processes, where BIM-enabled intelligent models have been used to search for balanced solutions rather than independently maximizing isolated objectives (Zhang et al., 2025). Complex project-management studies using metaheuristic methods similarly demonstrate that high-quality solutions must negotiate competing time and resource constraints (Sun, 2025; Zeng et al., 2026). The balanced PPO framework appears to have learned a compromise policy that remained close to the efficient frontier while maintaining operational feasibility, making it particularly suitable for practical construction management where no single objective can be optimized without considering its consequences for the others.

The positive expert-validation results provide an important complement to the computational findings. The participating professionals rated the realism of state variables, appropriateness of constraints, practical usefulness, and technical feasibility favorably, and 87.5% considered the framework useful or highly useful for project-control decisions. This suggests that the model's recommendations were not perceived merely as theoretically efficient but also as operationally credible. The integration of BIM, intelligent optimization, and project-control data has increasingly been proposed as a pathway toward more practical AI-supported construction management (Olaya et al., 2026; Xu & Guo, 2025). Deep learning applications in prefabricated construction also demonstrate the growing role of data-driven systems in supporting operational decision-making rather than only retrospective analysis (Guo, 2024). Accordingly, the combination of strong computational performance and favorable professional evaluation suggests

that reinforcement learning could serve as a decision-support tool that augments rather than replaces managerial judgment.

Several limitations should be considered when interpreting the findings. First, the empirical dataset was obtained from 84 construction projects located in Tehran, and although the projects varied in size, type, duration, and resource intensity, the geographic and organizational concentration may limit generalizability to construction markets with different labor regulations, cost structures, contractual systems, technologies, or resource availability. Second, the reinforcement learning environment necessarily simplified some real-world construction conditions by representing project states at weekly intervals and encoding only variables available in project records. Factors such as weather, subcontractor behavior, informal managerial decisions, safety incidents, and sudden regulatory changes were not modeled in full detail. Third, although stochastic variation was introduced during training and robustness testing, the framework was still evaluated primarily through historical and simulated project trajectories rather than through prospective deployment in active projects. Finally, the reward function depended on predefined weights for time, cost, and resource utilization, and different organizations may assign different strategic importance to these objectives.

Future studies should validate the proposed framework using larger and more geographically diverse samples, including infrastructure, industrial, residential, and megaproject environments. Prospective field trials are particularly recommended in which reinforcement learning recommendations are generated during ongoing project execution and compared with actual managerial decisions in real time. Future research should also incorporate additional uncertainty variables such as weather, supplier reliability, safety risk, material price volatility, subcontractor performance, and equipment breakdown. Multi-agent reinforcement learning could be investigated to represent independent crews, subcontractors, equipment pools, or project zones as interacting decision agents. Further studies may also combine reinforcement learning with graph neural networks, digital twins, BIM, and real-time Internet of Things data to create richer project-state representations. Explainable AI methods should be developed so that project managers can understand why specific resource reallocations are recommended, and multi-objective reward calibration methods should be examined to determine project-specific optimal trade-offs among schedule, cost, carbon emissions, safety, quality, and workforce stability.

Construction organizations can use the proposed framework as a decision-support layer integrated with existing scheduling, cost-control, and BIM platforms rather than as a replacement for project managers. Project teams should establish reliable digital data pipelines for activity progress, resource availability, cost performance, equipment utilization, and productivity before adopting reinforcement learning tools. The model should initially be deployed in advisory mode, allowing planners and project managers to compare AI-generated recommendations with conventional decisions before automated implementation is considered. Organizations should also define explicit operational constraints and managerial priorities so that generated actions remain feasible and consistent with contractual, safety, and workforce requirements. Regular recalibration is recommended when project conditions, resource prices, or productivity patterns change substantially. Finally, practitioners should evaluate reinforcement learning performance using operational indicators such as schedule deviation, cost deviation, idle time, utilization, and conflict frequency rather than relying solely on algorithmic reward values, because successful implementation depends on demonstrable improvements in actual project-control outcomes.

Authors' Contributions

Authors contributed equally to this article.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

We would like to express our gratitude to all individuals helped us to do the project.

Declaration of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

Ethics Considerations

In this research, ethical standards including obtaining informed consent, ensuring privacy and confidentiality were considered.

References

- Bahroun, Z., Tanash, M., As'ad, R., & Alnajar, M. (2023). Artificial Intelligence Applications in Project Scheduling: A Systematic Review, Bibliometric Analysis, and Prospects for Future Research. *Management Systems in Production Engineering*, 31(2), 144-161. <https://doi.org/10.2478/mspe-2023-0017>
- Chen, X., & Chen, L. (2024). Exploration of Adaptive Environment Design Strategy Based on Reinforcement Learning in CAD Environment. *Computer-Aided Design and Applications*, 175-190. <https://doi.org/10.14733/cadaps.2024.s23.175-190>
- Dai, W., Zan, J., Zhang, P., Hu, D., Fang, R., & Dong, Y. (2026). Cell-Free UAV Real-Time Sensing for Enhanced Substation Operation: Network Construction and Resource Management. *Journal of Physics Conference Series*, 3212(1), 012006. <https://doi.org/10.1088/1742-6596/3212/1/012006>
- Fawzy, M., Elshaer, R., Chuddher, B. A., & Algohary, G. (2026). Solving the Multi-Skill Resource-Constrained Project Scheduling Problem Considering Skill Depth and Breadth Using ACO Algorithm and LINGO. *Algorithms*, 19(9), 738. <https://doi.org/10.3390/a19090738>
- Gao, Z., Liu, S., Shi, G., Sun, Y., Ge, X., & Wang, Y. (2026). A Collaborative Scheduling Approach for Sheet Metal Workshops in Printing Equipment Ovens Based on Graph Attention Reinforcement Learning. *Symmetry*, 18(8), 1298. <https://doi.org/10.3390/sym18081298>
- Guo, R. (2024). Research and Application of Deep Learning Technology in Prefabricated Buildings. *Journal of World Architecture*, 8(2), 134-139. <https://doi.org/10.26689/jwa.v8i2.6920>
- Huang, Q., & Bai, Y. (2023). Application of Improved Artificial Immune System Algorithm Based on Applied Mathematics for Optimization of Manpower Allocation in Construction Engineering. *Processes*, 11(7), 1870. <https://doi.org/10.3390/pr11071870>
- Khan, I. U., Saleem, M. A., & Hussain, F. (2024). A Comparative Analysis of Different Vehicular Fog Computing Scheduling Algorithms. *Journal of Independent Studies and Research - Computing*, 22(1). <https://doi.org/10.31645/jisrc.24.22.1.9>
- Li, B., Yang, C., & Yang, Z. (2023). Multiple Container Terminal Berth Allocation and Joint Operation Based on Dueling Double Deep Q-Network. *Journal of Marine Science and Engineering*, 11(12), 2240. <https://doi.org/10.3390/jmse11122240>
- Li, B., & Zhang, Z. (2026). A Multiagent Reinforcement Learning-based Approach for Construction Resource Scheduling and Cost Control in Intelligent Construction Sites. 39. <https://doi.org/10.1117/12.3118711>
- Li, Y., Niu, X., Li, P., H, L., Dong, R., Li, N., & Tan, Z. (2026). Multi-Agent Deep Reinforcement Learning for Dynamic Cost

- Overrun Mitigation in Smart Grid Construction Projects. *Energies*, 19(13), 3147. <https://doi.org/10.3390/en19133147>
- Liang, H., Yang, Z., Zhang, G., & Hou, H. (2024). Resource Allocation for Space-Air-Ground Integrated Networks: A Comprehensive Review. *Journal of Communications and Information Networks*, 9(1), 1-23. <https://doi.org/10.23919/jcin.2024.10494938>
- Liu, H., Zhou, H., Chen, H., Yan, Y., Huang, J., Xiong, A., Yang, S., Chen, J., & Guo, S. (2023). A Federated Learning Multi-Task Scheduling Mechanism Based on Trusted Computing Sandbox. *Sensors*, 23(4), 2093. <https://doi.org/10.3390/s23042093>
- Lou, H. P., Shen, X., & Xu, J. (2025). Multi-Team Agile Software Project Scheduling Using Dual-Indicator Group Learning Particle Swarm Optimization. *Symmetry*, 17(8), 1267. <https://doi.org/10.3390/sym17081267>
- Luo, J. (2025). MADRL-Based Intelligent Resource Allocation for Green Sports Events. *International Journal of Decision Support System Technology*, 17(1), 1-23. <https://doi.org/10.4018/ijdsst.393787>
- Mirjalili, R. (2025). Resource-Based Time and Cost Prediction in Project Networks: From Statistical Modeling to Graph Neural Networks. <https://doi.org/10.48550/arxiv.2511.15003>
- Olaya, S., Tibaná, C., Sánchez, O., Castañeda, K., & Torres, K. (2026). Human Resource Planning for Building Construction Processes Through the Integration of BIM and Line of Balance. *Buildings*, 16(10), 1919. <https://doi.org/10.3390/buildings16101919>
- Sun, Y. (2025). A Complex Project Management Method Optimized by Ant Colony Algorithm. *Engineering Research Express*, 7(4), 045293. <https://doi.org/10.1088/2631-8695/ae1ddd>
- Verma, P., Maurya, A. K., & Yadav, R. S. (2023). A Survey on Energy-efficient Workflow Scheduling Algorithms in Cloud Computing. *Software Practice and Experience*, 54(5), 637-682. <https://doi.org/10.1002/spe.3292>
- Wang, L., Liu, H., Xia, M., Wang, Y., & Li, M. (2023). Research on a Multilevel Scheduling Model for Multi Variety and Variable Batch Production Environments Based on Machine Learning. *Frontiers in Energy Research*, 11. <https://doi.org/10.3389/fenrg.2023.1251335>
- Wang, L., Zhao, Y., & Yin, X. (2023). Precast Production Scheduling in Off-Site Construction: Mainstream Contents and Optimization Perspective. *Journal of Cleaner Production*, 405, 137054. <https://doi.org/10.1016/j.jclepro.2023.137054>
- Wang, M., Guo, H., Liu, C., Shabaz, M., Aldawsari, H., & Qiao, S. (2026). A Hierarchical and Entropy-Adaptive Reinforcement Learning Framework for Multi-Task Offloading in Vehicular Ad Hoc Networks. *Transactions on Emerging Telecommunications Technologies*, 37(6). <https://doi.org/10.1002/ett.70454>
- Xu, G., & Guo, T. (2025). Advances in AI-powered Civil Engineering Throughout the Entire Lifecycle. *Advances in Structural Engineering*, 28(9), 1515-1541. <https://doi.org/10.1177/13694332241307721>
- Yang, J., & Wu, Y. (2023). Dynamic Resource Allocation Strategy of Multi-Objective Fuzzy Optimization Based on Markov Decision Process. *IEEE Access*, 11, 99607-99613. <https://doi.org/10.1109/access.2023.3314657>
- Yang, T. (2026). Intelligent Scheduling and Optimization Decision Model of Elderly Service Resources Based on Reinforcement Learning. 25. <https://doi.org/10.1117/12.3108077>
- Yao, Y. (2026). Deep Learning-Based Intelligent Resource Management for Renewable Energy Engineering Projects. *Annals of the New York Academy of Sciences*, 1563(1). <https://doi.org/10.1111/nyas.70353>
- Zeng, J., Cheng, Y., Lu, C., Wang, Z., Liu, X., Wang, X., & Song, D. (2026). A Hybrid Reverse Learning Particle Swarm Optimization Method for Aircraft Maintenance Scheduling Based on the Resource-Constrained Project Scheduling Problem Model. *Machines*, 14(6), 622. <https://doi.org/10.3390/machines14060622>
- Zhang, J., Wang, T., & Cheng, L. (2023). Time-Sensitive and Resource-Aware Concurrent Workflow Scheduling for Edge Computing Platforms Based on Deep Reinforcement Learning. *Applied Sciences*, 13(19), 10689. <https://doi.org/10.3390/app131910689>
- Zhang, L., Fang, M., Yuan, K., Cheng, S., & Bai, L. (2025). Bridging Project Management and Supply Chain Management via Optimization Method: Scenarios, Technologies, and Future Opportunities. *Mathematics*, 13(21), 3490. <https://doi.org/10.3390/math13213490>
- Zhang, N., Yang, G., Su, B., & Song, W. (2024). Transfer Learning Method From Parameter Scene to Physical Scene Based on Self-Game Theory. *Electronics*, 13(7), 1327. <https://doi.org/10.3390/electronics13071327>
- Zhang, Y. (2025). Development of an Integrated Intelligent BIM-based Model for Multi-Objective Optimization in Engineering Assembly Processes. *PLoS One*, 20(11), e0333354. <https://doi.org/10.1371/journal.pone.0333354>
- Zhasmukhambetova, A., Evdorides, H., & Davies, R. J. (2025). Integrating Risk Assessment and Scheduling in Highway Construction: A Systematic Review of Techniques, Challenges, and Hybrid Methodologies. *Future Transportation*, 5(3), 85. <https://doi.org/10.3390/futuretransp5030085>
- Zhou, X., Liang, W., Yan, K., Li, W., Wang, K. I., Ma, J., & Jin, Q. (2023). Edge-Enabled Two-Stage Scheduling Based on Deep Reinforcement Learning for Internet of Everything. *Ieee Internet of Things Journal*, 10(4), 3295-3304. <https://doi.org/10.1109/ijot.2022.3179231>