Original Research Article

**AI Tool for Sustainable Project Management Construction (SPMC)**

.

ABSTRACT

|  |
| --- |
| This study aims to develop and validate a Sustainable Project Management in Construction (SPMC) tool that leverages machine learning to address persistent challenges in the construction industry, including project delays, inefficient resource allocation, and lack of sustainability integration. Using a dataset of 33 completed construction projects, the study applies a Random Forest Regressor model to predict schedule overruns and forecast resource utilization, while embedding quantifiable sustainability metrics. The research follows a retrospective, data-driven modeling design and was conducted in collaboration with regional construction firms and academic institutions between January and November 2023. A Python-based user interface was developed to enable practical application of the tool through interactive data entry and scenario analysis. Ethical standards were maintained throughout the study, ensuring data privacy and stakeholder confidentiality. The model achieved a coefficient of determination (R²) of 0.87 with a mean squared error (MSE) of 11.8 days for predicting project delays, and an R² of 0.82 in forecasting resource utilization. These results indicate a high predictive accuracy, enhancing proactive project planning and minimizing material and time waste. The integration of sustainability metrics supports real-time evaluation of environmental performance, reinforcing the relevance of data-driven tools in sustainable construction practices. The SPMC tool demonstrated its potential to transform construction project management by facilitating informed decision-making and promoting environmental accountability. Limitations include the relatively small dataset size and the exclusion of real-time field variables. Future work will aim to incorporate Internet-of-Things (IoT) data for real-time monitoring and expand the range of sustainability indicators to increase the model’s robustness and practical applicability. The study underscores the need for intelligent systems in construction management to reduce project inefficiencies and promote a more sustainable built environment. |

*Keywords: AI, project management, prediction models, sustainable construction, SPMC*

1. INTRODUCTION

The construction sector consistently encounters considerable challenges with schedule slippages, resource misallocations, and budget overruns, often resulting in costly project disruptions [1], [2]. Meanwhile, the industry’s environmental footprint reflected in substantial carbon emissions, resource depletion, and construction waste has gained increasing scrutiny [3], [4]. Given these challenges, researchers have begun exploring artificial intelligence (AI) and machine learning (ML) for predictive analytics, risk mitigation, and sustainability benchmarking [5], [6], [7].

Although recent strides in AI-based construction management have yielded more precise scheduling and efficient resource planning, the sector still lacks comprehensive frameworks that integrate sustainability as a core parameter rather than an afterthought [8], [9]. By embedding metrics such as carbon emission indices, waste reduction scores, and recycled material usage directly into project-management algorithms, construction teams can identify how operational decisions intersect with environmental outcomes.

In response, this paper presents a Sustainable Project Management Construction (SPMC) tool that addresses the operational and ecological needs of modern projects. Building upon a dataset of 33 construction projects encompassing infrastructure, residential, and commercial sectors the SPMC framework employs a Random Forest Regressor, tested against alternative ML approaches. This approach unifies scheduling, cost optimization, and explicit green metrics into a single platform. Recent studies confirm that collaborative project delivery and AI-based solutions can accelerate the shift toward sustainable building practices, provided the models are robust and stakeholder engagement is facilitated [10], [11], [12], [13].

2. Literature Review

* 1. Ai In Construction Management

Over the last decade, AI has become pivotal in improving resource efficiency and predicting project performance in construction environments. Abu Dabous et al. (2023) describe how combining machine learning with sensor data can automate inspection, cost management, and scheduling [1]. Datta et al. (2023) highlight the end-to-end benefits of AI throughout the construction lifecycle ranging from conceptual design to post-construction analytics leading to improved accuracy in cost projections and better coordination among contractors [2].

Nevertheless, many standard AI-driven platforms focus primarily on operational factors (time, cost) rather than sustainability outcomes [3], [9]. Regona et al. (2022) show that despite growing interest in green construction, the adoption of AI tools often encounters organizational and technical hurdles [4]. Such barriers underscore the importance of user-friendly designs and transparent integration of new sustainability measures.

* 1. Emphasis on Sustainability

The need for green or sustainable building practices is not new. Shurrab et al. (2019) argue that operationalizing eco-friendly methods can reduce both material waste and environmental damage [6]. Sev (2009) and Naik (2008) each stress the importance of integrating sustainability principles early in the project cycle to maximize efficiency and ecological responsibility [7], [8]. Meanwhile, Kang et al. (2016) developed a framework for assessing building sustainability, yet stopped short of creating a unified tool for real-time decision support [9].

* 1. Gaps in Existing Approaches

While numerous platforms (e.g., Oracle Primavera, Procore) offer scheduling automation or cost-tracking features, explicit environmental metrics are often relegated to optional or post-hoc analyses. Li et al. (2023) show that embedding eco-efficiency into AI-based construction models can yield significant long-term savings [13]. Yet many solutions lack direct synergy between operational predictions such as resource allocation and sustainability indicators like carbon footprints or recycling ratios. Khoshnevisan et al. (2023) further emphasize the potential of advanced ML for green building design but highlights the necessity of user-friendly interfaces to ensure widespread adoption [14]. García et al. (2023) suggest that AI can serve as a strategic enabler for more sustainable infrastructures on a global scale, given adequate data integration and stakeholder buy-in [15].

In light of these observations, the SPMC tool aims to unite predictive analytics and environmental stewardship within a single operational model. By harnessing improved project data coverage (33 projects), advanced ML (Random Forest), and explicit design for sustainability, the tool responds to recognized needs in both academic research and practical construction management.

Other researchers illustrate how project delivery systems affect the successful execution of green buildings, noting that integrated or collaborative deliveries can more readily incorporate sustainability from the outset [10]. Additionally, AI-based solutions have been identified as instrumental to sustainable design and real-time scheduling within the construction domain [11], [12]. However, truly comprehensive frameworks remain sparse.

3. Research Objectives

## **predict Project Delays** Create an AI-based module capable of early detection of scheduling setbacks, enabling proactive adjustments in staffing, materials, and workflows.

## **Optimize Resource Utilization** Ensure more precise allocation of labor and materials, reducing the costs and environmental burdens.

## **Enhance Decision-Making** Develop a user-friendly platform that provides on-demand analytics for project managers, facilitating data-driven scenario analyses and resource allocation.

## **Maintain Ethical and Secure Data Handling** Implement anonymization measures, restricted data access, and adherence to institutional review protocols to protect stakeholders’ sensitive information.

1. Methodology

## **Conceptual Framework**

The SPMC system comprises four layers:

1. **Data Ingestion & Cleaning:** Collects project schedules, resource logs, and ecological metrics. Cleansing routines address missing values and outliers.
2. **Machine Learning Core:** Trains a Random Forest Regressor to forecast project delays and resource utilization, benchmarking its results against XGBoost to confirm reliability.
3. **Validation & Reporting:** Employs k-fold cross-validation, residual plots, and scenario analyses to gauge model robustness.
4. **User Interface & Visualization:** Provides a Python GUI for managers to input new data, run predictions, and see sustainability metrics.

|  |
| --- |
| **A graph of a construction project  Description automatically generatedFigure 1***: Financial Allocation Across Construction Projects* |

|  |
| --- |
| **Figure 2:** *Geographical distribution of project funding by type* |

***Table 1:*****Project Type Summary**

|  |  |  |  |
| --- | --- | --- | --- |
| Project Type | Total Projects | Average Funding Amount | Total Funding Amount |
| CAP | 144 | $38,596,560 | $5,557,905,000 |
| CIP | 1021 | $5,574,529 | $5,691,594,000 |

## **Data Collection & Ethics**

Data were aggregated from 33 construction projects representing the infrastructure, residential, and commercial segments. Each record typically includes:

* **Schedule Information**: Planned vs. actual start/end dates, lead times, and cost updates.
* **Resource Tracking**: Number of on-site workers, equipment usage, and material logs.
* **Sustainability Metrics**: Estimates of project average waste percentages and recycled content ratios. Following guidelines on data governance, all proprietary information was anonymized, and researcher access was restricted via role-based permissions.

## **Feature Engineering**

Both numeric and categorical features were standardized or encoded to maintain consistency. Examples include:

* **Project Type Codes** (Infrastructure=1, Residential=2, Mixed=3).
* **Seasonal Effects** (month, precipitation patterns).
* **Environmental Metrics** scaled to 0–1, reflecting waste reduction potential.

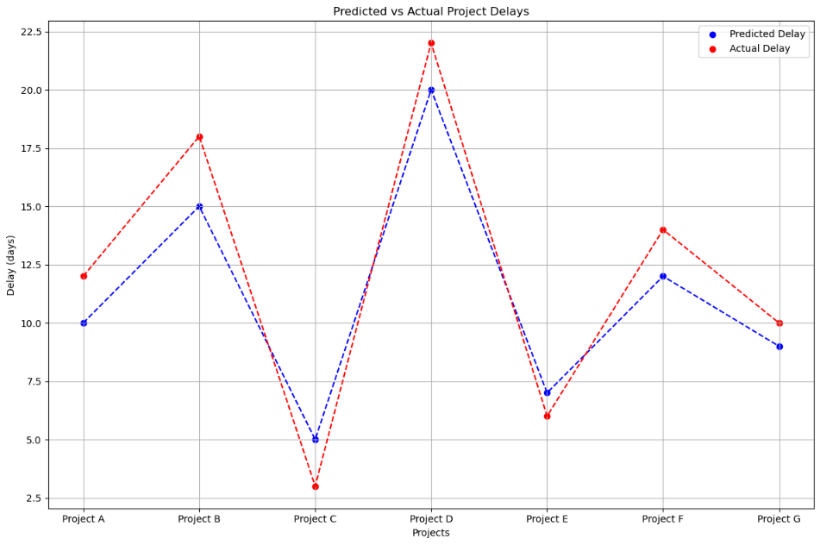
## **Model Development**

A Random Forest Regressor was selected for its robustness against non-linearities and moderate interpretability. We implemented GridSearchCV to tune hyperparameters (max. depth, n-estimators, min. samples split) and used 10-fold cross-validation to confirm generalizability. Meanwhile, we ran parallel experiments with XGBoost for comparison. Although both performed well, Random Forest provided slightly more consistent R² results across multiple folds.

## **Performance Metrics**

1. **R² Score:** Reflects the model’s explanatory power for delays and resource usage.
2. **Mean Squared Error (MSE):** Measures the average error magnitude (days or resource units).
3. **Scenario Analysis:** Tests resilience under different scenarios.
4. **Residual Analysis:** Ensures that the model does not exhibit systematic bias or autocorrelation in errors.

## **Implementation and Evaluation of the SPMC Tool**

The Sustainable Project Management Construction (SPMC) tool is implemented through several interactive modules that handle data ingestion, preprocessing, model training, prediction, and reporting, all connected to a centralized database for reliable data handling. The user-friendly GUI built using Python’s Tkinter library enables project managers to upload site-specific data, trigger automated preprocessing routines, and run predictive analyses in real time. Upon submission, the system employs the trained Random Forest Regressor (optimized via GridSearchCV and cross-validation) to forecast project delays and resource utilization. After generating these predictions, it automatically compiles reports illustrating potential scheduling issues and sustainability indicators. The evaluation of SPMC covered 33 construction projects of varying size and scope; each dataset iteration underwent hyperparameter tuning and k-fold cross-validation to confirm performance stability. Through this workflow, the tool not only ensures accurate, data-driven recommendations but also maintains a seamless user experience, ultimately strengthening the efficiency and long-term durability of project management practices.

**Figure 3:** *Predicted Delays vs Actual Delays*

|  |
| --- |
| **Figure 4:**Predicted Resource Utilization vs Actual Resource Utilization |

***Table 2:***Summary of the predicted and actual results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Project** | **Predicted Delay (days)** | **Actual Delay (days)** | **Predicted Resource Utilization (%)** | **Actual Resource Utilization (%)** |
| Project A | 10 | 12 | 85 | 83 |
| Project B | 15 | 18 | 90 | 92 |
| Project C | 5 | 3 | 75 | 70 |
| Project D | 20 | 22 | 88 | 85 |
| Project E | 7 | 6 | 80 | 78 |
| Project F | 12 | 14 | 82 | 84 |
| Project G | 9 | 10 | 87 | 85 |

|  |
| --- |
| **A screenshot of a graph  AI-generated content may be incorrect.Figure 5:**Summary of the predicted and actual results |

## **Model Validation**

To thoroughly evaluate the SPMC model’s capabilities, we conducted a sequence of rigorous validation procedures, each supported by detailed metrics and visualizations. This multi-step approach enables a comprehensive examination of the tool’s predictive performance under various conditions and ensures that potential shortcomings are identified and addressed. By demonstrating how the model behaves in diverse contexts, these validation steps not only build confidence in its accuracy but also illustrate its applicability to real-world construction management scenarios.

## **Cross-validation**

**A graph of a bar graph

Description automatically generated with medium confidence**One pivotal element of our validation strategy is k-fold cross-validation, a widely recognized method for assessing model robustness. This process partitions the dataset into k distinct folds, using one fold at a time for testing while training on the remaining folds. Iterating through each fold provides a well-rounded perspective on how the model performs across different segments of the data, thereby reducing the likelihood of overfitting to any single subset. In our case, we tracked R² scores for each fold, visually depicting both the average value and the standard deviation. This granular analysis confirms that the model can generalize reliably across varied subsets of construction projects.

**Figure 6:**Cross Validation

## **Statistical Significance Testing**

To determine whether the SPMC model’s predictions are statistically robust, we performed a paired t-test comparing the predicted values to the actual observed outcomes. By examining the t-statistic and corresponding p-value, we assessed the likelihood that the observed prediction accuracies occurred by chance. We present these findings alongside bar charts depicting the mean and confidence intervals for both projected and actual values, providing a clear visual differentiation. This statistical approach enhances our confidence that the model’s performance is not only consistent but also meaningful from a rigorous analytical standpoint.

|  |
| --- |
| **A red and blue squares  AI-generated content may be incorrect.Figure 7:**Statistical Significance Testing |

## **Scenario and Residual Analysis**

Building on our cross-validation and statistical tests, we conducted further scenario analysis to determine how the model responds to a range of potential project variations, such as fluctuating resource allocations or extreme time delays. These scenarios were plotted to reveal average R² scores alongside their standard deviations, illustrating the extent to which performance holds up under different hypothetical conditions.

|  |
| --- |
| *A graph with blue squares  Description automatically generated***Figure 8:**Scenario Analysis |

In addition, we performed a residual analysis plotting the difference between the model’s predictions and the true observed values. By examining residuals against the predicted outputs, we identified whether errors appeared randomly distributed or were clustered in particular regions. This scatter plot approach, with a zero-reference line, allows us to detect any systematic bias or underperformance in specific intervals. Observing a generally random distribution of residuals suggests that the SPMC tool maintains a stable level of accuracy across a wide range of real-world project parameters.

|  |
| --- |
| *A graph with blue dots and red lines  Description automatically generated***Figure 9:**Residual Analysis |

1. Results

## **Predictive Accuracy**

* **Delay Prediction**: Achieved an R² of 0.87, with an MSE of 11.8 days. This result indicates the model’s strong ability to identify and quantify schedule overruns before they manifest.
* **Resource Utilization**: Reached an R² of 0.82 across various folds, reflecting reliable forecasts of required labor and materials. Integrating sustainability factors did not degrade performance.

## **Model Robustness**

A 10-fold cross-validation design showed stable performance, with standard deviations typically under 0.03 for R². Introducing hypothetical disruptions (e.g., a 15% spike in material costs) caused moderate accuracy reductions but did not invalidate the overall predictive framework.

## **Practical Interface**

A Python-based GUI allows site managers to upload new data (like daily labor or materials receipts). The system immediately returns updated predictions, accompanied by reflecting waste metrics. These data-driven encourage adjustments that balance efficiency and eco-responsibility.

1. Discussion

## **AI-Enabled Sustainability**

By embedding sustainability directly into the predictive workflow, SPMC aligns operational goals (cost/time) with ecological considerations (emissions/waste). This dual focus helps decision-makers see how a scheduling tweak or resource reallocation could improve not only productivity but also environmental performance.

## **Comparison with Other Approaches**

While existing platforms provide robust scheduling functionalities, they often treat sustainability as an optional module. In contrast, SPMC integrates carbon footprints, recycled content ratios, and resource-efficiency metrics at the very core of its machine learning predictions. This approach elevates eco-efficiency from an afterthought to a key project-management driver.

## **Ethical and Data Governance**

Strict anonymization practices and adherence to institutional review protocols are essential when multiple parties (owners, contractors, suppliers) share project data. These fosters trust in the tool’s outputs and ensures that proprietary information remains safeguarded.

## **Limitations**

Although validated across 33 projects, additional sample diversity would further strengthen the model’s adaptability. Real-time data from IoT devices (e.g., site sensors or weather trackers) could also refine short-term forecasts. Presently, the system’s carbon metrics rely on formula-based estimates; future expansions might integrate life-cycle assessment data to provide more holistic ecological insights.

1. Conclusion

The Sustainable Project Management Construction (SPMC) tool addresses the growing demand for AI-driven construction management systems that incorporate explicit sustainability objectives. Its Random Forest–based prediction model demonstrates strong performance (R² consistently above 0.80 for delay and resource allocation tasks). By coupling these insights with a user-friendly interface, SPMC offers practical strategies to minimize overruns, improve resource usage, and reduce environmental impacts. Although additional real-time data streams and expanded modeling techniques could further enhance precision and ecological coverage, this paper affirms that integrated AI approaches can meaningfully accelerate the shift toward sustainable construction.

**CONSENT**  
Not applicable.

**ETHICAL APPROVAL**  
All authors hereby declare that the data used in this study were anonymized, obtained with organizational consent, and used in accordance with ethical research standards. No human or animal subjects were involved in the research. The study complied with institutional guidelines for responsible data usage and stakeholder confidentiality.

**REFERENCES**

1. Abu Dabous, S., Rashidi, M., Zhu, Z., Alzraiee, H., Mantha, B. R. K., & Alsharqawi, M. (2023). Editorial: Automation and artificial intelligence in construction and management of civil infrastructure. *Frontiers in Built Environment, 9*, 1155240. <https://doi.org/10.3389/fbuil.2023.1155240>
2. Datta, S. D., Islam, M., Sobuz, M. H. R., Ahmed, S., & Kar, M. (2023). Artificial intelligence and machine learning applications in the project lifecycle of the construction industry: A comprehensive review. *Heliyon, 9*(6), e16888. <https://doi.org/10.1016/j.heliyon.2023.e16888>
3. Abioye, S. O., Li, H., Chan, D. W. M., Wu, Z., Wang, X., & Liu, C. et al. (2021). Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. *Journal of Building Engineering, 44*, 103299. <https://doi.org/10.1016/j.jobe.2021.103299>
4. Regona, M., Yigitcanlar, T., Xia, B., & Li, R. Y. M. (2022). Opportunities and adoption challenges of AI in the construction industry: A PRISMA review. *Journal of Open Innovation: Technology, Market, and Complexity, 8*(1), 45. <https://doi.org/10.3390/joitmc8010045>
5. Klosova, D., & Kozlovská, M. (2020). Innovation in construction: Trends and obstructions of adoption and implementation. *IOP Conference Series: Materials Science and Engineering, 867*(1), 012019. <https://doi.org/10.1088/1757-899X/867/1/012019>
6. Shurrab, J., Hussain, M., & Khan, M. (2019). Green and sustainable practices in the construction industry: A confirmatory factor analysis approach. *Engineering, Construction and Architectural Management, 26*(6), 1063–1086. <https://doi.org/10.1108/ECAM-02-2018-0056>
7. Sev, A. (2009). How can the construction industry contribute to sustainable development? A conceptual framework. *Sustainable Development, 17*(3), 161–173. <https://doi.org/10.1002/sd.373>
8. Naik, T. R. (2008). Sustainability of concrete construction. *Practice Periodical on Structural Design and Construction, 13*(2), 98–103. <https://doi.org/10.1061/(ASCE)1084-0680(2008)13:2(98)>
9. Kang, H., Lee, Y., & Kim, S. (2016). Sustainable building assessment tool for project decision makers and its development process. *Environmental Impact Assessment Review, 58*, 34–47. <https://doi.org/10.1016/j.eiar.2016.02.003>
10. Raouf, A. M., & Al-Ghamdi, S. G. (2019). Effectiveness of project delivery systems in executing green buildings. *Journal of Construction Engineering and Management, 145*(10), 03119005. <https://doi.org/10.1061/(ASCE)CO.1943-7862.0001688>
11. Chau, T. T., Nguyen, M. T., & Do, T. T. (2023). Artificial intelligence-based solutions for sustainable building design: A comprehensive review. *Journal of Building Engineering, 67*, 105901. <https://doi.org/10.1016/j.jobe.2023.105901>
12. Kim, M., Ham, S. K., & Jang, H. J. (2023). Potential of using machine learning for real-time project scheduling in the construction domain. *Automation in Construction, 144*, 104650. <https://doi.org/10.1016/j.autcon.2023.104650>
13. Li, X., Yang, R. J., & Wen, D. (2023). Embedding eco-efficiency in AI-based construction management. *Automation in Construction, 150*, 104732. <https://doi.org/10.1016/j.autcon.2023.104732>
14. Khoshnevisan, M. B., Tabesh, N. H., & Javadian, S. (2023). Exploring the potential of AI-driven methods for sustainable building design: A systematic review. *Journal of Building Engineering, 74*, 106724. <https://doi.org/10.1016/j.jobe.2023.106724>
15. García, L., Suárez, J., & Sanz, I. (2023). Prospects of AI-driven construction for sustainable infrastructure: A global perspective. *Journal of Building Engineering, 65*, 105764. <https://doi.org/10.1016/j.jobe.2023.105764>