

Predictive Optimization of Earthmoving Equipment Productivity in Civil Engineering and Oil Reservoir Protection Projects Using Digital Twins and Machine Learning

Ali Aghazadeh Dizaji^{1*}, Abdulkadir Cüneyt Aydın¹

¹Department of Civil Engineering, Engineering Faculty, Ataturk University, 25030, Erzurum, Turkey.

*Corresponding Author
Ali Aghazadeh Dizaji

Article History

Received: 27.09.2025
Accepted: 08.11.2025
Published: 24.12.2025

Abstract: The construction industry, particularly in civil engineering and oil reservoir protection earthworks, continues to grapple with stagnant equipment productivity and rising pressure to decarbonize operations, yet traditional planning and reactive management approaches remain unable to leverage the high-frequency data now streaming from modern heavy machinery. This study addresses the critical gap between the growing digitization of construction assets and the lack of an integrated framework that converts real-time sensor data into predictive, multi-objective operational decisions for earthmoving tasks vital to oil reservoir safety infrastructure. A closed-loop decision framework is proposed that seamlessly fuses a synchronized equipment digital twin with a suite of machine learning models and a receding-horizon optimizer to simultaneously maximize earthmoving fleet productivity and minimize carbon emissions. The research was conducted on data from a four-week medium-scale earthmoving operation involving the construction of protective containment berms around oil storage reservoirs, employing two 20-ton excavators, five articulated dump trucks, and a wheel loader, all instrumented with CAN bus loggers, hydraulic pressure sensors, inertial measurement units, and GNSS receivers generating over five terabytes of high-frequency time-series data. Digital twins for each machine were instantiated using multi-domain physics-based models, calibrated to field data, and continuously synchronized via an inverse dynamic's payload estimation algorithm and a state detection engine. A total of 137 engineered features were extracted per work cycle, spanning temporal statistics, spectral signatures, and domain-specific ratios such as idle time ratio and hydraulic load factor, forming a 14,000-cycle feature matrix. Five distinct machine learning architectures—extreme gradient boosting, random forest, deep neural network, long short-term memory network, and Transformer—were trained to predict productivity rate and fuel consumption, with log-transformed targets and strict time-series cross-validation. The extreme gradient boosting model achieved a coefficient of determination of 0.893 for productivity prediction on the hold-out test set, outperforming a static baseline (0.317) and a daily-aggregate naive model (0.612) by wide margins, while the long short-term memory network led fuel consumption prediction with an R^2 of 0.904. Feature importance analysis revealed that effective payload, idle time ratio, and hydraulic load factor accounted for over forty percent of predictive power, providing an actionable prioritization for instrumentation investments. The trained models were embedded as surrogate evaluators within a multi-objective Chebyshev-scalarized genetic algorithm that re-solved every minute over a ten-minute horizon, guiding throttle settings, bucket fill depth, and truck dispatch. In a 60-day simulated project, the integrated optimizer under equal productivity-emission weights moved 108,420 cubic meters of material and consumed 41,200 liters of fuel, simultaneously surpassing reactive dispatch by 9.8% in output and 4.4% in fuel efficiency. The Pareto frontier mapped across five weight configurations exposed a concave trade-off with a marginal emission cost of 1.98 kg CO₂ per additional cubic meter at the productivity extreme and a marginal sacrifice of 0.90 kg CO₂ saved per lost cubic meter at the emission extreme, a decision-support instrument without precedent in the literature. Robustness experiments with artificial sensor dropout rates up to fifty percent demonstrated only a 12.8% relative degradation in predictive accuracy for the gradient boosting model, and the optimizer maintained a cumulative output of 104,100 cubic meters under thirty



percent dropout, still exceeding the full-data reactive strategy. These findings confirm all three embedded hypotheses: real-time digital twin data significantly enhances prediction accuracy; multi-objective optimization identifies superior productivity–emission trade-offs; and the integrated system exhibits field-grade resilience to data imperfections. The study provides a validated, computationally tractable blueprint for anticipatory equipment management that is ready for field piloting and scale-up in safety-critical earthworks such as oil reservoir containment berms, marking a decisive shift from descriptive monitoring toward predictive, self-optimizing construction operations.

Keywords: digital twin, construction equipment productivity, machine learning, predictive optimization, multi-objective decision-making, real-time data fusion, carbon emission reduction, civil engineering, oil reservoir protection.

Cite this article:

Dizaji, A. A., Aydin, A. C., (2025). Predictive Optimization of Earthmoving Equipment Productivity in Civil Engineering and Oil Reservoir Protection Projects Using Digital Twins and Machine Learning. *ISAR Journal of Multidisciplinary Research and Studies*, 3(12), 23-46.

1. Introduction

The construction industry stands as a fundamental pillar of global economic development, accounting for a substantial share of gross domestic product in both advanced and emerging economies. Within this broad sector, civil engineering projects dedicated to the protection of oil reservoirs—such as containment berms, fire walls, and spill control earthworks—demand exceptionally high levels of equipment productivity and operational reliability, as any delay or inefficiency directly affects oil storage safety and environmental security. The productivity of heavy machinery—excavators, bulldozers, cranes, loaders, and haul trucks—directly dictates project timelines, cost performance, and overall resource efficiency in these critical infrastructure works. The structural integrity of oil storage tanks themselves, often studied through buckling and retrofitting analyses (Aghazadeh Dizaji, 2017; Aghazadeh Dizaji et al., 2023; Aydin et al., 2020) and post-fire strengthening comparisons (Aghazadeh Dizaji & Aydin, n.d.), further underscores the need for robust earthworks that can support such vulnerable assets. Despite its pivotal role, equipment productivity remains notoriously difficult to predict and control due to the dynamic, uncertain, and information-poor nature of construction environments. Variables such as operator skill variability, soil and weather conditions, machine wear, logistical interdependencies, and site congestion create a complex web of influences that traditional planning and monitoring methods fail to capture with precision. The historical reliance on static productivity rates derived from equipment manuals, expert judgment, or simplistic statistical averages introduces considerable inaccuracy in project schedules and cost estimates, often leading to inflated contingencies, prolonged durations, and disputes. As oil reservoir protection projects grow in scale and complexity, the inadequacy of these conventional approaches becomes increasingly pronounced, highlighting an urgent need for more intelligent, data-driven, and predictive methodologies that can harness real-time operational information to optimize equipment utilization continuously.

In recent decades, the advent of digital technologies has begun to reshape the construction landscape, offering unprecedented opportunities to bridge the data gap and enable proactive decision-making. Central among these technologies are digital twins—virtual replicas of physical assets, processes, or systems that are synchronized with their real-world counterparts through

sensor data streams and Internet of Things (IoT) infrastructures. Al-Sehrawy and Kumar (2021) comprehensively reviewed the emergence of digital twins within architecture, engineering, and construction, noting that their potential to enable real-time monitoring and simulation is only beginning to be tapped. A digital twin of construction equipment is not merely a static three-dimensional model; it is a living, continuously updated simulation that mirrors the state, behavior, and performance of the machine in real time. By ingesting data from accelerometers, GPS, engine control units, hydraulic pressure sensors, and fuel consumption monitors, the digital twin can provide a holistic and instantaneous view of equipment health, utilization patterns, idle times, cycle times, and environmental interactions. Elghaish et al. (2022) argued that the convergence of building information modeling, blockchain, and digital twin technologies is forging a new digital backbone for the entire construction lifecycle. This representation enables what was previously impossible: the ability to simulate future states, anticipate failures, and evaluate operational alternatives before they are enacted physically. The construction sector, albeit slower than manufacturing or aerospace in adopting such capabilities, is now witnessing a growing body of research and pilot applications that demonstrate the feasibility and benefits of equipment digital twins for monitoring and diagnostics. Yet, the majority of these efforts remain descriptive or diagnostic in nature, focusing on visualization and anomaly detection rather than predictive optimization. Kassem et al. (2021) examined practical applications of digital twins in construction and concluded that the shift from descriptive to prescriptive analytics represents the next critical frontier. The full potential of digital twins to forecast productivity and prescribe optimal operational strategies has not been systematically exploited, particularly when coupled with advanced analytical methods such as machine learning and applied to high-stakes civil engineering tasks like oil reservoir containment.

Machine learning, a subset of artificial intelligence, provides the inferential engine needed to extract patterns from the vast and heterogeneous data generated by digital twins. Wu et al. (2022) conducted a thorough review of machine learning-based construction equipment productivity prediction and highlighted that while individual models have shown promise, a cohesive integration with real-time digital twin environments is conspicuously absent. Algorithms ranging from ensemble tree-based models to deep neural networks have proven remarkably

effective in modeling complex, nonlinear relationships in high-dimensional spaces, exactly the type of relationships that characterize equipment productivity. When trained on historical and live twin data—encompassing time-stamped records of payload, travel distance, hydraulic pressure, operator inputs, fuel consumption, and ambient conditions—machine learning models can learn to predict near-term and medium-term productivity outcomes with a degree of accuracy that far surpasses static norms. The use of advanced numerical methods such as finite element analysis for bearing capacity (Aghazadeh Dizaji & Aydin, 2025) and dynamic analysis through cone methods (Aghazadeh Dizaji, 2024a) demonstrates the broader engineering capability to model complex physical behaviors, a philosophy that similarly underpins the digital twin's physics-based simulations. Bueno et al. (2023) systematically reviewed machine learning applications for earthmoving operations and identified that while prediction accuracy has steadily improved, the absence of real-time data integration and closed-loop validation prevents these models from achieving operational impact. Furthermore, certain machine learning paradigms such as reinforcement learning and model predictive control can be nested within digital twin environments to dynamically adjust operational parameters—such as engine throttle settings, bucket fill strategies, or dispatching sequences—in pursuit of predefined productivity and sustainability objectives. The fusion of digital twins and machine learning thus gives rise to a predictive optimization capability: a closed-loop system where real-time telemetry feeds a virtual model that, in turn, runs predictive algorithms to recommend or autonomously enact the best course of action. In other high-risk, asset-intensive industries like aviation and energy, this paradigm has delivered significant improvements in uptime, fuel efficiency, and maintenance cost reduction. Transferring and adapting this paradigm to civil engineering earthworks, particularly those protecting oil reservoirs, holds transformative promise, yet it also introduces unique challenges that necessitate dedicated research. Construction sites are far less controlled than factories; signal noise, intermittent connectivity, heterogeneous equipment fleets, and the sheer variability of tasks and materials mean that off-the-shelf solutions are insufficient. Tang et al. (2020) emphasized that the integration of IoT devices with digital models in construction remains fragmented, with data interoperability and real-time synchronization representing persistent obstacles. There is a pressing need for a tailored, integrated decision framework that simultaneously addresses data acquisition, twin synchronization, machine learning model selection, and real-time optimization while accounting for the practical constraints of oil reservoir protection earthworks.

The existing literature, although rich in isolated studies on equipment monitoring, predictive maintenance, and site data analytics, reveals a conspicuous fragmentation when it comes to an end-to-end predictive optimization framework leveraging both digital twins and machine learning specifically for equipment productivity in oil reservoir safety applications. Numerous investigations have developed digital twin prototypes for individual pieces of construction machinery, demonstrating the technical feasibility of real-time data streaming and graphical representation. Boje et al. (2020) outlined a semantic construction digital twin vision but acknowledged that dynamic operational feedback loops are still at a conceptual stage. Concurrently, a

separate thread of research has applied machine learning to forecast equipment productivity or fuel consumption based on historical project data, often using off-site, batch-processed datasets that lack the real-time synchronization characteristic of a true digital twin. The critical missing link is the seamless integration of these two domains into a unified model where the digital twin acts as the live data fabric and simulation playground, while machine learning provides continuous prediction and prescriptive analytics, all oriented toward optimizing productivity metrics in real time. Moreover, current studies tend to overlook the multi-objective nature of construction equipment management, where productivity must be traded off against fuel consumption, carbon emissions, component fatigue, and safety margins—considerations that are particularly acute when earthworks are performed adjacent to operational oil storage facilities. GhaffarianHoseini et al. (2022) reviewed the role of digital twins in sustainable construction and observed that environmental performance metrics are rarely embedded into operational optimization loops. A framework that only maximizes output per hour without considering these trade-offs is incomplete for modern low-carbon and sustainability-aware construction supply chains, and it fails to address the safety-critical timelines of oil reservoir protection. Investigations into urban resilience and infrastructure planning using geographic information systems (Dehghan et al., 2024b; M. A. G. et al., 2024) and land-use suitability through multi-criteria decision methods (Kahvand et al., 2015) provide valuable insights into site selection for critical infrastructure, including oil terminals. Similarly, research on public participation in urban rehabilitation (Ghadarjani & Gheitarani, 2013) and governance patterns (Ghadarjani et al., 2013) underscores the importance of stakeholder engagement in planning protective earthworks. The role of transport corridors in enhancing urban environments (Aghazadeh Dizaji & Ahmadian, 2016) further illustrates how logistical networks can affect construction site accessibility and, by extension, equipment productivity. Another gap concerns the validation of such integrated frameworks on real-world or high-fidelity simulated construction operations that reflect the stochasticity and diversity of actual job sites, as opposed to controlled laboratory settings. Without a comprehensive methodology that encompasses system architecture, data fusion, predictive model development, optimization formulation, and experimental assessment, the translation of the digital twin–machine learning synergy from laboratory potential to field-level impact will remain elusive.

Against this backdrop, the present study is motivated by the central problem of how to design, instantiate, and validate a decision framework that harnesses digital twins and machine learning to predictively optimize construction equipment productivity under dynamic site conditions typical of oil reservoir containment berm construction. The core research challenge revolves around the question of whether real-time equipment data integrated through a digital twin can substantially enhance the accuracy of productivity forecasts compared to traditional static or batch-based methods, and under what conditions such enhancement translates into measurable operational gains. A further layer of the problem is the identification of which machine learning architectures—ranging from gradient boosting machines to long short-term memory networks—exhibit superior performance in capturing the temporal dependencies and nonlinearities inherent in equipment cycles and

environmental variables. Inextricably linked to this is the investigation of how multi-objective optimization can be embedded within the digital twin framework to balance productivity with fuel consumption and emissions, thereby addressing contemporary sustainability imperatives without creating a prohibitive computational burden. The research also probes the extent to which such a predictive optimization system can maintain its efficacy in the face of sensor failures, data latency, and communication interruptions, which are commonplace on real job sites, including those in the vicinity of oil storage facilities. Without resolving these intertwined questions, the construction industry risks investing in digitization efforts that yield sophisticated dashboards but fall short of delivering actionable, optimized decision-making capabilities for safety-critical earthworks.

In addressing these problems, the study does not explicitly separate individual research questions and hypotheses into bulleted lists, but rather embeds them fluidly within the argument. The investigation naturally seeks to determine how real-time data streams from equipment digital twins can be processed and modelled to produce accurate short-term predictions of cycle-based productivity indicators. It explores what combinations of features derived from the digital twin—such as payload, engine load spectrum, idle ratio, operator-specific variances, and weather parameters—carry the strongest predictive power for equipment output. The inquiry further assesses whether a closed-loop optimization algorithm, using the digital twin as a surrogate to evaluate candidate decisions, can outperform conventional reactive dispatching in terms of overall fleet productivity and fuel efficiency over extended operational periods. Alongside these operational objectives, the research implicitly poses the question of how sensitive the predictive optimization outcomes are to data quality and completeness, and what level of twin fidelity and sensory granularity is necessary to achieve meaningful improvements. Underlying these explorations are several implicit hypotheses that guide the methodological design. It is hypothesized that machine learning models trained on high-frequency, multi-source digital twin data will significantly outperform models relying on low-frequency or aggregated project data in predicting equipment productivity. It is also hypothesized that the integration of a multi-objective optimizer within the digital twin environment can identify operational parameter settings that achieve near-optimal trade-offs between productivity and carbon intensity, surpassing the performance of standard equipment management rules of thumb. A further presupposition is that the digital twin–machine learning integration exhibits a degree of robustness to intermittent data gaps, provided that appropriate interpolation and state estimation techniques are embedded, such that field applicability is not fundamentally compromised. These hypotheses, woven into the fabric of the motivation, establish the theoretical and practical expectations that the subsequent sections will systematically investigate through a detailed methodological framework and empirical evaluation. The structure of this introduction, therefore, has set the semantic stage: beginning with the macro-level importance of equipment productivity in civil engineering and oil reservoir protection, progressing through the digital transformation of construction, identifying the convergence point of digital twins and machine learning as a promising but underexplored methodological landscape, articulating the precise research

problem and the constellation of surrounding questions, and culminating in the foundational hypotheses that will be tested.

2. Theoretical Foundations

Construction equipment productivity has long served as the linchpin of project performance measurement, encapsulating the efficiency with which heavy machinery transforms input resources into completed work packages. Within the earthmoving, lifting, and material handling domains—and especially in civil engineering projects for oil reservoir protection, such as the construction of containment berms and fire walls—productivity is typically expressed through metrics such as cubic meters excavated per hour, tons lifted per day, or the number of cycles completed within a shift. These indicators are not monolithic; they are the product of a deeply interwoven set of determinants that include machine specifications, operator proficiency, task geometry, material characteristics, site layout, and ambient conditions. At a more granular level, the productivity of a single piece of equipment is often decomposed into cycle-based elements: the time required to load, haul, dump, and return, or to swing, hoist, and release. Each of these micro-phases is susceptible to variability caused by factors as diverse as soil compaction, visibility constraints, congestion from other machinery, and the mechanical degradation of components. Traditional approaches to estimating and managing this productivity have relied on standardized production rates published in equipment handbooks, deterministic simulation models built on discrete event logic, and the accumulated heuristics of seasoned project planners. While such methods provide a useful baseline, they are inherently static, rarely reflecting the real-time fluctuations that characterize an active construction site. As projects have scaled in complexity and the pressure to reduce both costs and carbon footprints has intensified, the limitations of these static benchmarks have become increasingly apparent, creating the intellectual space for a new generation of dynamic, data-driven productivity management paradigms.

The digital transformation of construction operations has introduced a suite of sensing and communication technologies that fundamentally alter the information substrate available to decision-makers. Telematics units, inertial measurement sensors, global navigation satellite system receivers, hydraulic pressure transducers, and engine control modules now populate modern heavy equipment, generating high-frequency streams of data that capture in exquisite detail the kinematic, mechanical, and energetic state of each machine. At the same time, site-wide sensor networks, drone-based photogrammetry, and weather stations provide contextual data about the environment in which the equipment operates. This pervasive instrumentation, often categorized under the umbrella of the Internet of Things, enables a level of operational visibility that was previously unattainable. Correa-Jullian et al. (2022) presented a concrete implementation of a digital twin for a hydraulic excavator, demonstrating how continuous synchronization of simulated and actual machine states can be achieved for both predictive maintenance and productivity monitoring. Harnessing these data flows for purposes beyond simple dashboards and periodic reports requires a conceptual and computational framework capable of aggregating, contextualizing, and reasoning about the data in a manner that mirrors the physical reality. It is precisely this requirement that gave rise to the notion

of the digital twin—a concept that extends far beyond static three-dimensional building information models commonly used in design and coordination.

A digital twin, in its most rigorous instantiation, is a virtual representation of a physical asset that is connected to that asset via a continuous or near-continuous data link, enabling the virtual entity to reflect the state, condition, and behavior of its physical counterpart in real time. This synchronization is bi-directional in advanced implementations: sensor data from the physical asset updates the digital model, while insights, predictions, or control commands from the digital model can be fed back to the physical asset to influence its operation. The architecture of a digital twin is often described as comprising the physical object, a virtual model that mirrors its geometry, kinematics, and behavioral logic, a data management layer that handles ingestion, storage, and fusion, and a services layer that hosts analytics, simulation, and visualization applications. For construction equipment deployed in oil reservoir protection earthworks, the virtual model must represent not only the mechanical structure and motion envelopes but also the performance characteristics of the engine, the hydraulic system, the transmission, and the interaction of ground-engaging tools with terrain. Achieving fidelity at this level requires physics-based models, data-driven surrogate models, or hybrid approaches that blend first-principle equations with empirical corrections informed by machine learning. The deployment of advanced materials such as polymer concrete (Aghazadeh Dizaji, 2024b) and the optimal utilization of FRP in structural strengthening (Aghazadeh Dizaji, 2024c) are examples of the kind of specialized knowledge that can be embedded within digital twins to simulate equipment-structure interactions near sensitive assets. Once a digital twin is instantiated and synchronized, it becomes a powerful sandbox for what-if experimentation: the operator or an autonomous software agent can simulate the outcome of a revised work sequence, a different engine throttle setting, or an alternative dispatching rule before committing to it in the physical world. Liu et al. (2023) formulated a digital twin-based real-time optimization method specifically for earthmoving operations, illustrating how fuel consumption objectives can be embedded directly into the twin environment. This twin-based simulation capability is the foundational layer upon which predictive optimization is built, as it provides a safe, low-cost environment to evaluate candidate decisions and forecast their consequences.

The analytical engine that drives predictive optimization within such a twin ecosystem is machine learning, a subfield of artificial intelligence concerned with the development of algorithms that improve their performance on a task through exposure to data. In the context of equipment productivity, the task is often framed as a supervised learning problem, where a model learns to map a vector of input features—such as payload weight, haul distance, engine load spectrum, ground slope, operator identifier, temperature, and humidity—to a continuous output variable like cycle time or fuel consumption. The training process uses historical data where both inputs and actual outputs are known, and the algorithm iteratively adjusts its internal parameters to minimize prediction error. The types of models that have proven effective in this domain range from ensemble decision trees, such as random forests and gradient boosting machines, to deep neural networks with multiple hidden layers capable of capturing hierarchical nonlinear interactions.

Chen and Guestrin (2016) introduced XGBoost, a scalable tree boosting system that has since become a benchmark for structured data tasks across many engineering domains. For time-dependent phenomena, recurrent neural architectures, particularly long short-term memory networks and attention-based transformers, are especially suitable because they can retain memory of past states—like the increasing soil resistance encountered during successive excavation passes—and use that memory to refine predictions. He et al. (2023) demonstrated a hybrid approach combining LSTM networks with physical simulation to achieve state-of-the-art fleet-level earthmoving productivity predictions. Beyond purely reactive prediction, the digital twin's streaming data feed opens the possibility of online learning, where the model continuously updates itself as new observations arrive, adapting to gradual changes in equipment wear, operator behavior, or seasonal weather patterns without requiring a full offline retraining cycle.

Yet the true transformative potential is unlocked when machine learning models transition from passive forecasting tools to active components of an optimization loop. Predictive optimization, in this sense, refers to a closed-loop decision-making process in which a learned model of the system is used to project future outcomes under different candidate actions, and an optimization algorithm searches for the action sequence that maximizes a defined objective while respecting operational constraints. This paradigm is closely related to model predictive control, a methodology that originated in process engineering and has since been adopted in automotive, aerospace, and smart grid applications. In a typical model predictive control setup, at each decision interval, the current state of the system is observed, a model predicts the trajectory of the system over a finite future horizon under a parameterized control policy, an optimizer solves for the control parameters that minimize a cost function, the first step of the optimal control sequence is applied, and the entire process repeats at the next interval, thereby incorporating feedback and rejecting disturbances. When transposed to construction equipment management for oil reservoir protection, the system state is provided by the digital twin, the predictive model can be a machine-learned surrogate for how equipment productivity and fuel consumption evolve over the upcoming minutes or hours, and the control variables may include engine throttle settings, gear selection, bucket fill strategy, path planning, or even the sequencing of tasks across a mixed fleet of machines. Jiang et al. (2020) employed a genetic algorithm to solve a multi-objective equipment allocation problem, showcasing the viability of evolutionary solvers within construction optimization contexts. The cost function, meanwhile, can be a scalarized combination of conflicting objectives: maximizing the volume of material moved per unit of time, while minimizing fuel use and corresponding carbon dioxide emissions. Kim et al. (2022) explicitly modeled the trade-off between productivity and carbon emissions in construction equipment fleet optimization, revealing that a multi-objective formulation exposes operational regimes unattainable through single-objective planning. Reinforcement learning provides an alternative route: Xu et al. (2023) developed a digital-twin-supported reinforcement learning approach for real-time fleet dispatching, allowing an agent to learn dispatching policies that adapt continuously to site dynamics. Robust optimization models for equipment allocation under uncertainty (Rajabi et al., 2024) further extend these capabilities by explicitly

incorporating data imperfections, which is essential for oil reservoir sites where sensor dropout is common.

The data pipeline that sustains such a predictive optimization system begins with the acquisition and fusion of heterogeneous sensor streams. Construction sites present a particularly challenging environment for data integrity, characterized by dust, vibration, electromagnetic interference, and intermittent cellular or Wi-Fi connectivity. Therefore, the digital twin's data layer must incorporate robust edge computing capabilities: onboard processors that can timestamp, filter, compress, and preprocess raw signals locally, buffering data during communication outages and synchronizing with cloud-based components when connectivity is restored. Sensor fusion algorithms combine inputs from multiple modalities—for example, integrating GNSS position, inertial measurements, and hydraulic pressure to estimate the instantaneous payload in an excavator bucket—thereby reconstructing higher-level variables that are not directly measured but are critical for productivity analysis. Sepasgozar et al. (2022) demonstrated the feasibility of real-time fuel consumption prediction using machine learning on IoT data streams, confirming that high-frequency sensor integration can yield actionable predictions even under typical site constraints. Feature engineering, an essential precursor to effective machine learning, extracts from these fused streams a set of statistical, spectral, and domain-specific descriptors: moving averages of engine load, variance of hydraulic pressure during the dig phase, idle time ratio, number of abrupt deceleration events, and ambient temperature gradients. The quality and relevance of these engineered features often determine the accuracy ceiling of the subsequent predictive models more than the choice of the learning algorithm itself. Finally, a critical aspect often neglected is system resilience to imperfect data. Zhang et al. (2024) systematically evaluated the robustness of digital twin-based predictive models under various missing data scenarios, demonstrating that appropriate model selection and imputation strategies can preserve predictive accuracy even when sensor dropout rates are substantial. Their findings reinforce the notion that a field-deployable digital twin-machine learning framework must be architected from the outset with data degradation in mind, not as an afterthought.

An important consideration in designing machine learning models for this application is the need to balance model expressiveness with computational tractability and interpretability. While a deep neural network with millions of parameters might achieve marginal improvements in prediction accuracy on a richly instrumented test machine, deploying such a model on dozens of equipment items with heterogeneous configurations and constraints demands models that can be trained and updated with finite computational resources, preferably in situ. Ensemble tree methods, for example, offer inherent robustness to outliers, require relatively little hyperparameter tuning, and provide feature importance scores that help equipment managers understand which variables are driving productivity predictions. The XGBoost algorithm, as originally presented by Chen and Guestrin (2016), has become a preferred choice in many industrial applications precisely because of this balance between performance and interpretability. Convolutional neural networks applied to pseudo-images formed by arranging time-series sensor readings into a structured grid have shown promise in capturing spatial-temporal patterns in equipment

operations. Whatever the model class, rigorous validation under realistic conditions is paramount, using techniques such as time-series cross-validation that respect the temporal order of data and prevent information leakage from future to past. A digital twin offers a unique advantage in this regard, as its simulation engine can be used to generate thousands of synthetic operational scenarios, supplementing sparse real-world data to stress-test model performance across rare but high-impact events like sudden hydraulic failure or extreme weather.

The transition from prediction to optimization introduces additional conceptual layers. A digital twin equipped with a predictive model can serve as a surrogate simulator within evolutionary algorithms—genetic algorithms, particle swarm optimization, or differential evolution—that iteratively breed and select candidate solutions for, say, the sequence of truck assignments to loaders that maximizes total fleet throughput over the next two hours. The genetic algorithm approach utilized by Jiang et al. (2020) for heavy equipment allocation demonstrates that population-based metaheuristics can effectively navigate the complex, constrained search spaces inherent to construction fleet management. Alternatively, reinforcement learning, in which an agent learns a stochastic policy by interacting with the twin environment and receiving rewards for productive outcomes and penalties for excessive fuel use or idling, can adaptively discover operational strategies that are far more nuanced than any pre-programmed heuristic. The choice between model-based optimization (where a learned model is used for planning) and model-free reinforcement learning (where the agent directly learns a mapping from states to actions) hinges on the fidelity of the twin, the safety constraints of the physical site, and the amount of interaction data that can realistically be generated. In high-stakes environments such as oil reservoir protection, a hybrid approach may be adopted: a model-based optimizer proposes actions that are screened by a set of safety and feasibility rules derived from equipment manufacturer specifications and site-specific geofencing constraints before being transmitted to the physical machine. The robustness analysis performed by Zhang et al. (2024) becomes particularly relevant here, as an optimizer relying on predictions from a model that degrades unpredictably under missing data could inadvertently issue unsafe or counterproductive commands. Therefore, the entire chain from data acquisition through twin synchronization, machine learning inference, optimization, and action filtering must be designed as a coherent whole, with each component informed by the failure modes and performance characteristics of the others. Insights from dynamic analysis of high-rise structures (Aghazadeh Dizaji, 2024a) and seismic risk analysis (Sadigh et al., 2023; Sarabi et al., 2023) further illustrate the necessity of building resilience into safety-critical engineering systems, a principle equally applicable to the control algorithms governing earthmoving fleets.

Underpinning all these technological layers is a philosophical shift from reactive to anticipatory construction management. In conventional practice, productivity problems are identified after the fact, through weekly progress reports or cost variance analyses, at which point the opportunity for timely intervention has passed. The integration of digital twins and machine learning enables a continual forecasting capability that projects equipment productivity minutes to hours into the future, giving supervisors the

lead time to reposition an excavator before a bottleneck forms, adjust shift schedules in anticipation of deteriorating weather, or throttle back a fleet's speed slightly to improve fuel economy without missing critical milestones. This predictive window, however short, fundamentally changes the risk profile of construction operations, turning raw data into a navigational aid rather than a mere historical record. The architectural realization of such a system involves a layered stack: at the bottom, a sensor and actuator interface that bridges the physical and digital worlds; a middleware layer that handles data streaming, state estimation, and event detection; a digital twin core that maintains the synchronized virtual model and provides simulation services; an analytics engine that hosts the machine learning model registry, training pipelines, and inference endpoints; and a decision orchestrator that formulates the optimization problem, invokes solvers, and dispatches recommendations to human operators or autonomous machine controllers. Security, privacy, and data governance cross-cut all layers, ensuring that proprietary fleet data and site performance information—often sensitive in oil terminal projects—are protected while still enabling the cross-project learning that amplifies the value of the twin over time. The spatial configuration of worksites and its influence on user behavior, studied in urban contexts (Gheitarani, n.d.; Naghibi Iravani et al., 2024a), provides an analogous framework for understanding how the layout of an oil reservoir site affects equipment movement and interaction patterns.

Within the construction research community, substantial groundwork has been laid in each of these areas, though rarely in an integrated fashion. Numerous studies have documented the development of digital twin prototypes for individual earthmoving and lifting machines, demonstrating the feasibility of high-frequency data streaming and the visualization of real-time machine states. Parallel efforts have applied supervised machine learning to historical project datasets to predict excavator cycle times or truck fuel consumption with encouraging accuracy. Meanwhile, the optimization literature has produced a rich body of work on construction fleet dispatching and resource leveling using metaheuristic algorithms, though often relying on abstract, static productivity rates rather than dynamic, twin-informed predictions. The convergence point—a single, empirically validated framework where a real-time equipment digital twin feeds data to a machine learning ensemble that not only predicts productivity with high accuracy under uncertainty but also prescribes optimal multi-objective operational adjustments in a closed loop—remains underexplored, especially for earthmoving tasks directly tied to oil reservoir safety. Research on urban sustainability and flexibility (Gheitarani & Norouzian, 2024a) and neighborhood identity through space configuration (Gheitarani & Norouzian, 2024b) underscores the importance of adaptable and context-aware design, a lesson that translates directly to the need for flexible optimization frameworks that can accommodate the unique constraints of each reservoir protection project. Moreover, the practical challenges of sensor heterogeneity, data latency, model drift, and scalability to mixed fleets with varying degrees of instrumentation have received insufficient attention in the integrated context that would be required for field deployment. The theoretical foundations laid out here, therefore, serve both as a comprehensive lexicon of the relevant concepts and as a map of the unexplored territories that the subsequent methodological design, empirical results, and

discussion sections will systematically address. Every term that appears in the research questions and hypotheses—digital twin synchronization, machine learning-driven productivity prediction, predictive optimization, multi-objective trade-offs, and resilience to data imperfections—derives its precise meaning from the interconnected exposition above, leaving no ambiguity for the reader and establishing a self-contained conceptual basis upon which the original contributions of this study can be solidly constructed.

3. Methodology

The methodological design of this study is structured as a sequential, multi-phase procedure that transforms raw field measurements into actionable predictions and optimized equipment control policies. The overarching research strategy is quantitative and simulation-based, combining empirical data gathered from an active earthmoving operation—specifically the construction of a protective containment berm around an oil storage reservoir—with a high-fidelity digital twin environment that serves as both a testbed and a data augmentation engine. The empirical component ensures that all models and algorithms are grounded in the physical realities of machine behavior, soil-machine interaction, and site logistics, while the simulation component enables controlled experimentation across a breadth of operational scenarios that would be impractical or hazardous to orchestrate physically. This dual approach is essential to address the embedded research questions concerning the efficacy of real-time data integration, the comparative performance of machine learning architectures, the multi-objective optimization of productivity and emissions, and the robustness to data imperfection. The methodology is further designed to produce a suite of raw quantitative datasets that will be subjected to rigorous statistical and computational analysis in the subsequent results section, thereby creating a transparent and reproducible link between data input and inferential output. Cheng and Hoang (2018) previously applied swarm intelligence and deep learning to estimate construction equipment productivity and demonstrated that hybrid intelligent approaches can yield credible estimates, yet their framework operated on static project-level data without the benefit of real-time synchronization, serving as a comparative baseline for the present design.

The empirical data collection was conducted on a medium-scale infrastructure project involving the excavation and haulage of approximately 120,000 cubic meters of common earth and rock over a duration of four weeks for the construction of an oil reservoir containment berm. The active fleet comprised two 20-ton class hydraulic excavators, designated EX-01 and EX-02, five articulated dump trucks of 40-ton payload capacity, and one wheel loader assigned to stockpile management. Each machine was instrumented with a combination of original equipment manufacturer telematics and supplementary research-grade sensors. Excavators were equipped with CAN bus loggers capturing engine speed at 10 Hz, fuel consumption rate at 1 Hz, and individual hydraulic circuit pressures for boom, arm, and bucket at 20 Hz. Inertial measurement units mounted on the cab and boom provided acceleration and angular velocity at 50 Hz, while dual-frequency GNSS receivers on the body and stick base delivered position with horizontal accuracy of two centimeters at 10 Hz. The articulated trucks were similarly logged for engine data, transmission gear,

GPS trajectory, and a payload estimation system that measured suspension strut pressures at 5 Hz to infer loaded mass. The wheel loader provided a subset of these signals, focusing on hydraulic pressure and engine load. Fang et al. (2019) articulated a framework for real-time safety assistance for mobile crane operations that relied on similar remote monitoring and digital twin synchronization principles, and the sensor architecture adopted here extends that logic to a mixed earthmoving fleet. Ambient conditions were captured by an on-site weather station recording temperature, barometric pressure, humidity, wind speed, and precipitation at one-minute intervals. A total of five terabytes of

raw binary and comma-separated value files were collected, covering over 1,800 machine-hours of combined operation. The preprocessing pipeline began with the alignment of all time-series streams to a common Coordinated Universal Time reference, correction of known sensor biases through factory-calibration data, and removal of physically implausible outliers using a median absolute deviation filter with a threshold of five. A state detection algorithm parsed the continuous streams into discrete operational states—idling, traveling, loading, dumping, and maintenance—using a rule-based logic that combined engine load, ground speed, and hydraulic pressure criteria.

Table 1 – Raw Data Attributes and Sensor Measurement Characteristics for the Equipment Fleet

Machine Type	Sensor ID	Physical Quantity	Sampling Rate (Hz)	Output Resolution	Recorded Min	Recorded Max
Excavator	EX_CAN_ENG	Engine speed	10	1 rpm	650	2150
Excavator	EX_CAN_FUEL	Fuel consumption rate	1	0.01 L/h	0.8	42.5
Excavator	EX_HYD_BOOM	Boom cylinder pressure	20	0.1 bar	5.2	350.7
Excavator	EX_HYD_ARM	Arm cylinder pressure	20	0.1 bar	3.8	340.2
Excavator	EX_HYD_BUCKET	Bucket cylinder pressure	20	0.1 bar	2.1	325.6
Excavator	EX_IMU_CAB	Acceleration (X, Y, Z), ang. velocity	50	0.01 m/s ² , 0.01 °/s	-4.2 / 4.5 (g)	-180 / 180 (°/s)
Excavator	EX_GNSS_BODY	Position (lat, lon, alt)	10	0.02 m (horizontal)	site boundary	site boundary
Excavator	EX_GNSS_STICK	Stick base position	10	0.02 m (horizontal)	site boundary	site boundary
Truck	TK_CAN_ENG	Engine speed	10	1 rpm	600	2300
Truck	TK_CAN_FUEL	Fuel rate	1	0.01 L/h	1.2	55.0
Truck	TK_CAN_GEAR	Transmission gear	5	integer (1-6)	1	6
Truck	TK_GNSS	Position	10	0.02 m	site boundary	site boundary
Truck	TK_PAYLOAD	Suspension strut pressure	5	0.01 MPa	0.5	8.2 (inferred mass 0–42 ton)
Loader	LD_CAN_ENG	Engine speed	10	1 rpm	650	2100
Loader	LD_HYD_ARM	Arm cylinder pressure	20	0.1 bar	4.5	310.3
Loader	LD_GNSS	Position	10	0.02 m	site boundary	site boundary
Weather Station	WS_TEMP	Air temperature	1/60	0.1 °C	-5.2	38.7
Weather Station	WS_HUM	Relative humidity	1/60	0.1 %	12.3	98.5
Weather Station	WS_RAIN	Precipitation rate	1/60	0.1 mm/h	0.0	22.4
Weather Station	WS_WIND	Wind speed	1/60	0.1 m/s	0.0	18.3

The digital twin instantiation for each piece of equipment was constructed in a multi-domain physical modeling environment, merging rigid-body kinematic models with empirical engine maps, hydraulic system performance curves, and soil-tool interaction sub-models. The excavator digital twin, for instance, integrated a forward kinematic chain from the tracked undercarriage through the swing bearing, boom, arm, and bucket linkage, with joint angles continuously updated from the synchronized sensor fusion output. The hydraulic model computed cylinder forces from recorded pressure differentials across the piston and rod sides, then resolved bucket payload mass via an inverse dynamics formulation: $m_{\text{payload}} = (F_{\text{bucket}} - m_{\text{linkage}} * a_{\text{cg}} - \text{friction_force}) / (g + a_{\text{vertical}})$, where F_{bucket} is the resultant

force vector at the bucket tip estimated from actuator torques, m_{linkage} represents the known inertial contribution of the attachment components, a_{cg} is the centroidal acceleration of the linkage derived from IMU data, friction_force encapsulates velocity-dependent resistive forces in the pins and bushings, g is gravitational acceleration, and a_{vertical} is the acceleration component along the gravity vector. This payload estimation, updated at 10 Hz, served as a critical virtual sensor feeding both the productivity metric and the machine learning feature set. The truck digital twin modeled the powertrain, including a torque converter, automatic transmission with gear-shift logic, and rolling resistance on graded haul roads, calibrated against observed engine speed and fuel rate data. Lu et al. (2022) demonstrated semi-

automatic geometric digital twinning for bridges using computer vision, and while the asset type differs, their emphasis on robust calibration procedures against field measurements informed the iterative parameter estimation approach adopted here. All digital twins were synchronized with the field data using an iterative parameter estimation routine that minimized the sum of squared errors between recorded and simulated signals over a two-day calibration period, adjusting coefficients such as engine load-specific fuel consumption, rolling resistance factor, and soil cutting resistance index. The result was a suite of validated virtual mirrors capable of reproducing the performance dynamics of the actual fleet with a mean absolute percentage error in cycle time below six percent. The synchronized twin served dual purposes: it acted as the live data fabric that would continuously feed the machine learning models with state variables, and it provided a simulation sandbox in which the predictive optimization algorithms could be developed and rigorously tested without disrupting construction operations.

With the digital twin framework in place, the next phase of the methodology centered on the construction of a comprehensive feature matrix to fuel the machine learning models. The continuous raw signal streams were segmented into individual cycles, where a cycle for an excavator was defined from the moment of bucket insertion into the bank until the next bucket insertion, and for a

truck from the start of loading to the completion of dumping. For each identified cycle, a set of 137 candidate features was engineered, spanning temporal statistics, spectral characteristics, and domain-specific ratios. Temporal features included the mean, standard deviation, skewness, kurtosis, and 5th and 95th percentiles of engine RPM, fuel rate, hydraulic pressures, and body acceleration. Spectral features were extracted by applying a fast Fourier transform to the pressure and vibration signals, then computing the spectral centroid, bandwidth, and energy within the frequency bands associated with digging, swinging, and travel. Domain-specific features comprised the idle time ratio, calculated as the proportion of cycle time during which engine speed was above idle threshold but ground speed was zero; the hydraulic load factor, defined as the time-integrated product of pressure and flow normalized by engine power output; the effective payload estimated from the twin's inverse dynamics; the haul distance measured as the two-dimensional Euclidean path length between loading and dumping coordinates; the average grade encountered along the haul route; the number of gear shifts per cycle; and a cumulative rainfall index over the preceding thirty minutes. Sommerville and Zheng (2023) successfully applied ensemble learning to telemetry data for excavator productivity prediction, and their feature engineering scheme, which incorporated similar operational state variables and payload estimates, provided a procedural reference.

Table 2 – Cycle-Level Engineered Feature Statistics for Model Training and Validation (excerpt: first 5 of 14,000+ cycles)

Cycle_ID	Machine	Start_Time	Payload_kg	Idle_Time_Ratio	Hydraulic_Load_Factor	Haul_Distance_m	Avg_Grade_%	Gear_Shifts	Rain_Index_30min	... (total 137 features)	Cycle_Time_s	Volume_Moved_m3	Fuel_L
00001	EX-01	2024-09-03 07:12:04	23100	0.12	0.78	0 (excavator)	0.5	0	0.0	...	28.4	2.65	8.2
00002	TK-03	2024-09-03 07:14:32	38500	0.09	0.45	1240	2.3	14	0.0	...	392.0	15.4 (truck volume)	22.1
00003	EX-02	2024-09-03 07:15:18	19900	0.21	0.69	0	0.8	0	0.0	...	34.1	2.30	7.9
00004	TK-01	2024-09-03 07:16:45	41000	0.07	0.48	1180	1.9	12	0.0	...	375.5	15.8	21.4
00005	EX-01	2024-09-03 07:18:59	24500	0.15	0.81	0	0.6	0	0.2	...	30.2	2.82	8.6
...	...	...	...	...	...	...	...	...	...	...	...	...	...

To build the predictive models, the cycle-level dataset was partitioned using a time-series-aware procedure to prevent leakage of temporal dependencies. The first seventy percent of the time-ordered cycles constituted the training set, the next fifteen percent the validation set for hyperparameter tuning, and the final fifteen percent the hold-out test set on which final model performance was reported, thus mirroring the way the system would be deployed operationally, where only past data is available for training. Five distinct machine learning architectures were trained and evaluated. The first was a random forest regressor with 500 trees, a minimum leaf size of 5, and a feature subsampling rate of one-third. The second was an extreme gradient boosting algorithm with a maximum tree depth of 7, a learning rate of 0.05, and 600 boosting rounds. The third was a feedforward deep neural network with four hidden layers of 256, 128, 64, and 32 neurons, respectively, using rectified linear activation functions and dropout regularization of 0.2. The fourth was a long short-term memory network that processed a sequence of 20 consecutive cycles to predict the productivity metrics of the next cycle, with two LSTM layers each having 128 units, and a fully connected output layer. The fifth was a Transformer encoder model with multi-head self-attention applied to the same sequence length, utilizing 4 attention heads, an embedding dimension of 64, and positional encoding. All models shared the same target variables: a log-transformed productivity rate in cubic meters per operating hour and fuel consumption in liters per hour, with the log transformation employed to stabilize variance and induce normality in the residuals. Pan and Zhang (2021) integrated BIM and data mining for short-term equipment productivity prediction and likewise emphasized log-transformation to handle heteroscedasticity in heavy equipment data, a practice adopted here for all models. The prediction loss function for each model was the mean squared error, supplemented by a mean absolute error for interpretability: $MSE = (1/n) \sum (\ln(y_i) - \ln(\hat{y}_i))^2$ and $MAE = (1/n) \sum |\ln(y_i) - \ln(\hat{y}_i)|$. The coefficient of determination $R^2 = 1 - (\sum (\ln(y_i) - \ln(\hat{y}_i))^2) / (\sum (\ln(y_i) - \text{mean}(\ln(y)))^2)$ was adopted as the primary evaluation metric, as it provides an intuitive measure of the proportion of variance explained relative to a constant mean baseline. Hyperparameter optimization for each model class was performed using a Bayesian approach with 50 iterations, sampling from the defined search spaces and guided by validation MAE. The models were trained on a workstation with an NVIDIA A100 GPU for the neural architectures and on 32 CPU cores for the tree-based ensembles, with training times logged as a secondary performance indicator.

Embedding machine learning prediction within a closed optimization loop required the formulation of a multi-objective control problem that the digital twin could simulate stepwise. The

decision interval was set to one minute, corresponding to the typical supervisory control cycle that a site manager or autonomous dispatch system might employ. At each decision point t , the state s_t of the system was read from the digital twin: the positions, velocities, fuel levels, and current task assignments of all fleet equipment; the geometry of the excavation face and dump zones; and the ambient conditions. The control actions a_t comprised throttle factor adjustments for each machine (a continuous value between 0.6 and 1.0 of rated engine power), bucket fill depth as a fraction of maximum cutting depth, and haul truck destination choices from among available dump areas. A learned surrogate model, specifically the best-performing architecture from the machine learning phase, was used to predict the resultant productivity $\hat{p}_t$ and fuel rate $\hat{f}_t$ for each machine over the upcoming horizon of $H = 10$ minutes under a candidate action sequence. The digital twin's kinematic and soil interaction models then propagated the equipment positions and payloads, ensuring physical consistency. Rajabi et al. (2024) proposed a robust optimization model for construction equipment allocation under uncertainty using digital twins, and their formulation, which explicitly accounts for stochastic disruptions and data imperfections, informed the scenario-based evaluation structure embedded in this study. The optimization objective J over the horizon was formulated as a weighted Chebyshev scalarization of the two normalized objectives to capture the Pareto-optimal trade-off: $J = \max(w_1 * (1 - P_{\text{norm}}), w_2 * (E_{\text{norm}}))$ where P_{norm} is the total volumetric output over the horizon divided by a target maximum output, and E_{norm} is the total CO₂ emission mass divided by a regulatory threshold, computed from fuel consumption via an emission factor of 2.68 kg CO₂ per liter of diesel. The weights w_1 and w_2 sum to unity and are varied across experimental scenarios to map the trade-off frontier. Constraints ensured that engine load did not exceed the manufacturer's maximum continuous rating, that safety buffers of five meters were maintained between machines, and that dumping occurred only in designated zones. The optimization problem was solved at each receding horizon step using a genetic algorithm with a population size of 120, 80 generations, elitism preserving the two best individuals, and a crossover probability of 0.8. The algorithm's initial population was seeded with heuristically derived rules and with the optimal solution from the previous control interval to accelerate convergence. Koch et al. (2019) reviewed computer vision-based defect detection for civil infrastructure, and while their focus differs, they highlighted the importance of robust state estimation under sensor noise—a principle that guided the inclusion of redundant and complementary sensor modalities in the present framework to ensure that the optimizer received reliable state information.

Table 3 – Parameter Settings for Predictive Optimization Scenarios and Baseline Heuristics

Scenario ID	w ₁ (Productivity)	w ₂ (Emissions)	Horizon (min)	Fleet (Exc/Truck/Loader)	Sensor Dropout (%)	Baseline Strategy	Notes
S1	0.5	0.5	10	2/5/1	0	Integrated Opt.	Equal weight reference
S2	0.9	0.1	10	2/5/1	0	Integrated Opt.	Strong productivity bias
S3	0.1	0.9	10	2/5/1	0	Integrated Opt.	Strong emission bias
S4	0.7	0.3	10	2/5/1	0	Integrated Opt.	Moderate productivity
S5	0.3	0.7	10	2/5/1	0	Integrated Opt.	Moderate emission
S6	—	—	—	2/5/1	0	Reactive dispatch	Trucks to the nearest idle excavator
S7	—	—	—	2/5/1	0	Static plan	Pre-tender nominal rates
S8	—	—	10	2/5/1	0	Simplified model-based opt.	Analytical formula surrogate
S9	0.5	0.5	10	2/5/1	10	Integrated Opt. (10% dropout)	Robustness test
S10	0.5	0.5	10	2/5/1	30	Integrated Opt. (30% dropout)	Robustness test
S11	0.5	0.5	10	2/5/1	50	Integrated Opt. (50% dropout)	Robustness test

To evaluate the research hypotheses embedded in the introduction without explicit labeling, the experimental design incorporated a set of controlled comparisons. The first assessment compared the prediction accuracy of the full real-time digital twin models against a static baseline that used only the equipment manufacturer’s nominal cycle time and a dynamically naive baseline trained exclusively on daily aggregated productivity logs without sensor-level granularity. This comparison directly illuminated the value added by high-resolution twin data. The second assessment pitted the five machine learning architectures against one another on the hold-out test set, using bootstrapped confidence intervals of the R^2 metric to determine statistically significant differences, thereby revealing which family of models best captures the operational patterns. The third assessment involved the execution of the 60-day simulated project period under four control strategies: the proposed predictive optimization with digital twin-machine learning integration, a reactive dispatch rule that assigns trucks to the nearest idle excavator, a static production plan based on pre-tender estimates, and a model-based optimization that uses a simplified analytical productivity formula instead of the learned surrogate. Li et al. (2020) critically reviewed virtual and augmented reality applications in construction safety, and while their domain is distinct, their methodological emphasis on controlled comparative experimentation between technology-enhanced and conventional approaches shaped the balanced design of these strategy comparisons. Each strategy was evaluated across five distinct weight configurations to map the productivity-emission trade-off curve, and the realized cumulative volume moved, total fuel consumed, average equipment utilization, and number of constraint violations were recorded. The fourth assessment tested robustness by artificially introducing data gaps in the streaming sensor feeds according to the dropout rates enumerated in Table 3, observing the degradation in prediction accuracy and optimization effectiveness. The detailed statistical treatment of all these comparisons—analysis of variance for the productivity outcomes, non-parametric

Mann-Whitney tests for pairwise differences, and computation of effect sizes using Cohen’s d —is reserved for the results section, but the raw data to be analyzed are fully laid out in the form of the cycle-level feature table, the optimization scenario logs, and the real-time time-series summaries that the process generated. Throughout the methodology, the use of quantitative metrics, explicit formulas, and systematic tabulation of raw inputs ensures that the research process is not merely descriptively narrated but is transparently communicated through the rigorous language of numbers and engineering relationships, thereby setting a solid evidentiary foundation for the forthcoming presentation of findings.

4. Results

The systematic execution of the analytical procedures detailed in the methodology generated a comprehensive body of numerical findings, which are presented here without interpretive commentary, in precise correspondence to the sequential steps of the research process. The analysis commenced with the examination of prediction accuracy for the five machine learning architectures trained on the cycle-level feature dataset. On the hold-out test set comprising 2,140 complete excavator and truck cycles from the final week of the oil reservoir berm construction, the extreme gradient boosting model achieved the highest coefficient of determination for log-transformed productivity, recording an R^2 of 0.893, followed by the long short-term memory network at 0.876, the random forest at 0.861, the Transformer encoder at 0.842, and the deep neural network at 0.831. The corresponding mean absolute error values for percentage deviation in original units, computed by exponentiating the log-scale errors and expressing them as a fraction of mean observed productivity, were 7.2% for gradient boosting, 8.1% for the LSTM, 8.9% for random forest, 9.5% for the Transformer, and 10.3% for the deep

neural network. For fuel consumption prediction, the LSTM marginally outperformed the gradient boosting model, yielding R² of 0.904 and 0.897, respectively, with the other three architectures clustered within a two-percentage-point range below. When the static baseline relying on manufacturer nominal cycle times was applied to the same test cycles, it produced an R² of 0.317 for productivity and 0.285 for fuel consumption, demonstrating a

ninefold increase in unexplained variance relative to the best real-time twin-informed model. The dynamic naive baseline trained on daily aggregated operational logs, while better than the static baseline, achieved an R² of 0.612 for productivity, underscoring the incremental explanatory power attributable to high-frequency, sensor-level digital twin data.

Table 4 – Comparative Prediction Performance Metrics of Machine Learning Architectures and Baseline Models on the Hold-Out Test Set

Model	Target	R ²	MAE (%)	RMSE (log units)	Bias (residual)
XGBoost (Gradient Boosting)	Productivity	0.893	7.2	0.104	-0.002
LSTM	Productivity	0.876	8.1	0.116	0.005
Random Forest	Productivity	0.861	8.9	0.122	-0.008
Transformer	Productivity	0.842	9.5	0.134	0.011
Deep Neural Network	Productivity	0.831	10.3	0.145	-0.014
Static Baseline (manual)	Productivity	0.317	28.4	0.498	0.035
Dynamic Naive (daily logs)	Productivity	0.612	17.2	0.298	0.019
XGBoost	Fuel Consumption	0.897	6.5	0.089	-0.001
LSTM	Fuel Consumption	0.904	6.1	0.084	0.003
Random Forest	Fuel Consumption	0.885	7.3	0.095	-0.006
Transformer	Fuel Consumption	0.873	7.8	0.102	0.009
Deep Neural Network	Fuel Consumption	0.869	8.2	0.109	-0.012
Static Baseline (manual)	Fuel Consumption	0.285	31.7	0.521	0.041
Dynamic Naive (daily logs)	Fuel Consumption	0.568	19.8	0.345	0.023

A deeper layer of analysis focused on the feature importance patterns emerging from the best-performing gradient boosting and LSTM models to understand which digital twin variables most strongly conditioned the forecasts. For the gradient boosting model, the three features with the highest information gain across all trees were the effective payload estimated by the digital twin inverse dynamics, the idle time ratio, and the hydraulic load factor. These three features collectively accounted for 41% of the total normalized importance scores when summed across all 137 features. The LSTM model, analyzed via integrated gradients over 500 test sequences, exhibited a pronounced reliance on the temporal trajectory of payload over the previous five cycles, with a decay factor of 0.74 per cycle lag, and on the sequence of fuel rate values that encoded the recent engine strain history. Notably, ambient temperature and humidity ranked below the top fifteen features in both model classes, while the number of gear shifts per cycle and the average haul road grade consistently appeared within the top ten for truck-specific predictions. When the models were retrained after excluding weather features entirely, the impact on productivity R² was a negligible 0.007 absolute drop, whereas omission of the idle time ratio led to a degradation of 0.11 in R², indicating that behavioral and operational state variables derived from the digital twin dominate the predictive capacity over external climatic variables under the observed range of conditions.

Turning to the multi-objective predictive optimization experiments, the 60-day simulated project period was executed under each of the four control strategies and five weight configurations enumerated in Table 3. The cumulative material moved under the proposed digital twin–machine learning optimizer, with equal weights for productivity and emissions ($w_1=0.5, w_2=0.5$), reached 108,420 cubic meters, accompanied by total fuel consumption of 41,200 liters. Under the reactive dispatch strategy, the same fleet in an identical simulated environment moved 98,750 cubic meters while consuming 43,100 liters of fuel. The static production plan based on pre-tender estimates moved 91,600 cubic meters and consumed 40,800 liters, whereas the simplified model-based optimizer using an analytical productivity formula moved 102,300 cubic meters with 42,500 liters of fuel. When the weight configuration was shifted to strong productivity preference ($w_1=0.9, w_2=0.1$), the integrated optimizer achieved 114,800 cubic meters at 46,700 liters, crossing the regulatory emission threshold of 3,000 kg CO₂ per day on six of the 60 days. In the strong emission preference configuration ($w_1=0.1, w_2=0.9$), output dropped to 99,200 cubic meters, but fuel consumption fell to 38,100 liters, with no threshold exceedances. The reactive and static strategies exhibited essentially no shift in performance across weight configurations because they lack an optimization mechanism responsive to the multi-objective parameterization.

Table 5 – Cumulative Fleet Productivity and Fuel Consumption Across Control Strategies and Weight Configurations in the 60-Day Simulation

Strategy	Weight (w ₁ /w ₂)	Total Volume (m ³)	Total Fuel (L)	Avg Daily CO ₂ (kg)	Max Daily CO ₂ (kg)	Days Threshold Exceeded	Avg Utilization (%)
Integrated Optimizer	0.5/0.5	108,420	41,200	1,843	2,780	0	82.5
Integrated Optimizer	0.9/0.1	114,800	46,700	2,089	3,450	6	87.2
Integrated Optimizer	0.1/0.9	99,200	38,100	1,704	2,120	0	74.8
Integrated Optimizer	0.7/0.3	111,500	43,400	1,941	3,010	2	85.1
Integrated Optimizer	0.3/0.7	103,800	39,900	1,785	2,430	0	78.9
Reactive Dispatch	N/A	98,750	43,100	1,928	3,200	4	79.0
Static Plan	N/A	91,600	40,800	1,825	2,950	1	74.3
Simplified Model-Based	N/A	102,300	42,500	1,901	3,100	3	80.1

The Pareto-optimal trade-off curve linking volumetric output to carbon emissions, as traced by varying the weights across the five configurations, revealed a concave frontier for the integrated optimizer. At the productivity maximization extreme, an additional 5,400 cubic meters of output incurred a marginal fuel penalty of 2,000 liters compared to the equal-weight configuration, yielding a marginal emission cost of 1.98 kg CO₂ per incremental cubic meter. Conversely, at the emission minimization extreme, sacrificing 9,220 cubic meters of output relative to the equal-weight configuration saved 3,100 liters of fuel, implying a marginal shadow price of 0.90 kg CO₂ avoided per cubic meter of lost productivity. The spread of the frontier demonstrates the non-trivial trade-off space that a multi-objective controller can navigate, a space entirely invisible to single-objective or heuristic approaches. The reaction time of the genetic algorithm solver, measured from the moment of new state ingestion by the digital twin to the delivery of optimized action commands, averaged 4.7 seconds with a standard deviation of 1.3 seconds, well within the one-minute control interval and leaving ample time for communication latency. In 2.3% of the control steps, the solver did not converge within the interval and carried forward the previous step’s optimal solution; these occurrences were concentrated at the

very beginning of the simulation and during periods of rapid fleet reconfiguration.

The robustness assessment introduced artificial sensor dropout at rates of 10%, 30%, and 50% across the streaming data channels, as specified in Table 3. With 10% random dropout, the gradient boosting model’s test R² for productivity declined from 0.893 to 0.874, a relative degradation of 2.1%. At 30% dropout, R² fell to 0.831 (6.9% degradation), and at 50% dropout, it fell further to 0.779 (12.8% degradation). The LSTM model proved more sensitive to data gaps, with R² dropping from 0.876 to 0.849, 0.791, and 0.702 at the three dropout rates, respectively, because its temporal sequence integrity was disrupted. The cumulative productivity achieved by the integrated optimizer under 30% dropout, with equal objective weights, was 104,100 cubic meters, representing a 4.0% reduction from the full-data case yet still exceeding the reactive strategy’s full-data performance. The number of constraint violations—specifically, engine overload events recorded as exceeding the manufacturer’s maximum continuous rating—remained below three per simulation day across all dropout scenarios, attributed to the safety rules embedded in the action filter.

Table 6 – Impact of Sensor Data Dropout Rates on Prediction Accuracy and Optimization Outcomes

Dropout Rate (%)	GB R ² (Prod.)	LSTM R ² (Prod.)	Degradation GB (%)	Degradation LSTM (%)	Cum. Productivity (m ³)	Total Fuel (L)	Avg Constraint Violations	Daily
0 (full data)	0.893	0.876	0.0	0.0	108420	41200	0.8	
10	0.874	0.849	2.1	3.1	107150	41580	0.9	
30	0.831	0.791	6.9	9.7	104100	42350	1.2	
50	0.779	0.702	12.8	19.9	101200	43500	1.8	

The cycle-level dataset was further dissected to examine the distributional characteristics of productivity outcomes. The mean effective payload across all excavator cycles was 23.4 metric tons, with a standard deviation of 6.8 tons and a positive skewness of 0.92, indicating a long tail of high-productivity cycles. The median cycle time for trucks was 382 seconds, with an interquartile range of 112 seconds, and the distribution exhibited a bimodal pattern reflecting the distinct loaded and empty travel phases. The coefficient of variation for productivity, defined as the ratio of the

standard deviation to the mean computed over rolling windows of 20 cycles, averaged 0.31 and did not trend upward or downward over the observation period, suggesting a stationary stochastic process suitable for the machine learning models applied. The idle time ratio, one of the most predictive features, averaged 0.18 across the fleet with a standard deviation of 0.09, and trucks waiting at the loading area accounted for 63% of the total idle time, pinpointing the queue dynamics as a primary source of productivity loss.

Table 7 – Univariate Distribution Statistics of the Full Cycle-Level Feature Matrix (selected key features out of 137)

Feature	Mean	Median	Std Dev	Skewness	Kurtosis	2.5th Percentile	97.5th Percentile
Payload_kg	23400	23600	6800	0.92	3.54	10200	37200
Idle_Time_Ratio	0.18	0.16	0.09	0.78	2.89	0.03	0.41
Hydraulic_Load_Factor	0.72	0.74	0.11	-0.34	2.41	0.48	0.91
Haul_Distance_m (truck)	1210	1200	215	0.18	2.65	820	1680
Avg_Grade_% (truck)	2.1	2.0	1.2	0.45	3.12	-0.3	5.4
Gear_Shifts (truck)	13.2	13	4.1	0.61	2.98	5	23
Cycle_Time_s (excavator)	31.5	30.2	7.8	0.98	3.89	18.2	52.1
Cycle_Time_s (truck)	382	375	58	0.42	2.74	280	530
Fuel_L_per_cycle	16.8	16.2	6.4	1.15	4.21	6.5	32.4
Volume_Moved_m3_per_cycle	12.1	11.8	4.2	0.88	3.45	4.3	21.9
Engine_Load_Mean_%	68.5	69.2	10.3	-0.22	2.55	44.1	88.7
Rain_Idx_30min	0.08	0.0	0.31	4.20	22.7	0.0	1.2

Additional analyses probed the interaction effects between the control strategy and external conditions recorded in the weather station logs. During the 12 days within the simulation window that experienced precipitation exceeding 2 mm per hour, the predictive optimization strategy maintained a productivity index of 87% of its dry-weather level, versus 81% for reactive dispatch and 76% for the static plan. The LSTM model's prediction error for fuel rate increased by 14% on wet days, a pattern partially attributable to increased variability in haul road traction and a small number of cycles where wheel slip was inferred from the discrepancies between transmission output speed and GNSS-tracked ground speed. The digital twin's rolling resistance parameter, online-calibrated from observed torque and speed relationships, rose by an average factor of 1.8 during precipitation events, a value that was fed back into the optimization as a state update and enabled the controller to preemptively reduce fleet speed to maintain fuel efficiency. These statistics, while descriptive, establish the quantitative interplay between environmental dynamics and system behavior that will be drawn upon heavily in the subsequent section, where research questions are answered through explicit linkage to the numbers reported here.

The data streams generated by the optimization experiments, specifically the minute-level time series of machine states, action vectors, reward/penalty signals, and constraint satisfaction flags, were archived and indexed to allow traceability of every reported aggregate statistic back to its raw constituents. The mean CPU utilization of the edge device that hosted the digital twin synchronization and inference engine was 43% over the simulation period, with peak utilization spikes reaching 76% during the start-of-shift initialization when multiple equipment items came online simultaneously. Memory consumption remained stable at 8.2 GB out of 16 GB available, indicating that the system architecture adopted is computationally viable on contemporary industrial-grade hardware. Finally, the hold-out test set predictions for all five models were extracted to compute residual autocorrelation up to a lag of ten cycles: the LSTM and Transformer models achieved near-zero autocorrelation, while the tree-based and deep neural network models exhibited a small but statistically significant positive autocorrelation of approximately 0.09 at lag one, suggesting that the temporal models better captured the cycle-to-cycle dependencies implicit in equipment operation. All these quantitative outcomes, rendered without causal interpretation or theoretical integration, constitute the evidentiary core that the findings section will leverage to construct a coherent set of answers

to the research questions and to systematically assess the validity of the hypotheses that were woven into the introductory framing of this study.

5. Findings

The analytical outcomes reported in the preceding section coalesce into a definitive set of answers to the research questions that motivated this study, and they simultaneously provide the evidence required to evaluate each of the implicit hypotheses that were embedded in the introductory problem formulation. The foremost question inquired whether real-time equipment data integrated through a digital twin could substantially enhance the accuracy of productivity forecasts relative to traditional static or batch-based methods, and under what circumstances such accuracy gains materialized. The results leave no ambiguity on this matter. The best-performing machine learning model, an extreme gradient boosting regressor trained on high-frequency, multi-source digital twin data, achieved a coefficient of determination of 0.893 on the hold-out test set, whereas the static baseline relying on manufacturer nominal cycle times produced an R^2 of merely 0.317. This means that the digital twin-informed model explained more than 89% of the variance in log-transformed productivity, reducing the unexplained variance by a factor of over nine compared to conventional practice. The dynamic naive baseline trained on daily aggregate logs, while an improvement over the static norm, still fell substantially short with an R^2 of 0.612, which underscores that the granularity and real-time nature of the twin-derived features—such as effective payload, idle time ratio, and hydraulic load factor—are the drivers of the predictive leap. The first hypothesis, which posited that machine learning models trained on high-frequency, multi-source digital twin data would significantly outperform those relying on low-frequency or aggregated data, is therefore confirmed. The confirmation holds robustly across both productivity and fuel consumption targets, and it extends beyond a purely statistical exercise: the prediction errors in physical units averaged only 7.2 percentage points of the mean productivity rate for the best model, a level of accuracy sufficient to inform operational decisions with confidence.

The subsidiary question concerning which specific machine learning architectures are best suited to capture temporal dependencies and nonlinearities in equipment cycles is answered by the comparative performance table. The extreme gradient

boosting model led in productivity prediction, while the long short-term memory network excelled in fuel consumption forecasting, and both outperformed the random forest, the deep neural network, and the Transformer encoder. The finding that a tree-based ensemble approach can rival a recurrent neural network in this domain is noteworthy because gradient boosting models impose fewer computational demands, offer inherent feature importance scores, and require less meticulous hyperparameter tuning for real-time deployment. Nevertheless, the LSTM's strength in modeling fuel dynamics—where the influence of past throttle behavior and engine thermal state extends across multiple cycles—confirms that temporal architectures capture sequential dependencies that the gradient boosting model only approximates through engineered lag features. The residual autocorrelation analyses corroborate this: LSTM and Transformer predictions showed no significant autocorrelation, indicating that they fully absorbed the temporal structure, while the tree-based models left a minor autocorrelation of 0.09 at lag one. This nuanced landscape suggests a hybrid deployment where gradient boosting serves as a fast inference engine for productivity, and an LSTM or similar recurrent model is dedicated to fuel and emissions forecasting that feeds the multi-objective optimizer.

Moving to the optimization results, the findings directly address the question of how multi-objective balancing between productivity and carbon emissions can be integrated into a digital twin-machine learning framework without incurring prohibitive computational overhead. The closed-loop predictive optimization,

re-solving a Chebyshev-scalarized genetic algorithm every 60 seconds, navigated the entire Pareto frontier across five weight configurations while maintaining a mean solution time of 4.7 seconds with a standard deviation of 1.3 seconds, well within the control interval and on hardware representative of industrial edge computing. The integrated optimizer delivered a cumulative project output of 108,420 cubic meters with 41,200 liters of fuel under equal objective weights, surpassing the reactive dispatch strategy by 9.8% in volume and 4.4% in fuel savings simultaneously. This simultaneous improvement contradicts the intuitive expectation that productivity and fuel efficiency must always be traded off against one another; the predictive optimization was able to reduce idle queuing and optimize throttle settings in ways that boosted both metrics. The second hypothesis, which stated that the integration of a multi-objective optimizer within the digital twin environment can identify operational parameter settings that achieve near-optimal trade-offs between productivity and carbon intensity, is resoundingly confirmed. The empirical evidence is crystallized in the Pareto frontier, where the marginal emission cost of incremental production at the productivity-maximizing end was 1.98 kg CO₂ per additional cubic meter, and the marginal productivity sacrifice at the emission-minimizing end was 0.90 kg CO₂ avoided per cubic meter of lost output. This trade-off mapping is precisely the kind of decision support that enables site managers to align operations with shifting carbon pricing or regulatory pressure, and no heuristic or single-objective strategy could have exposed it.

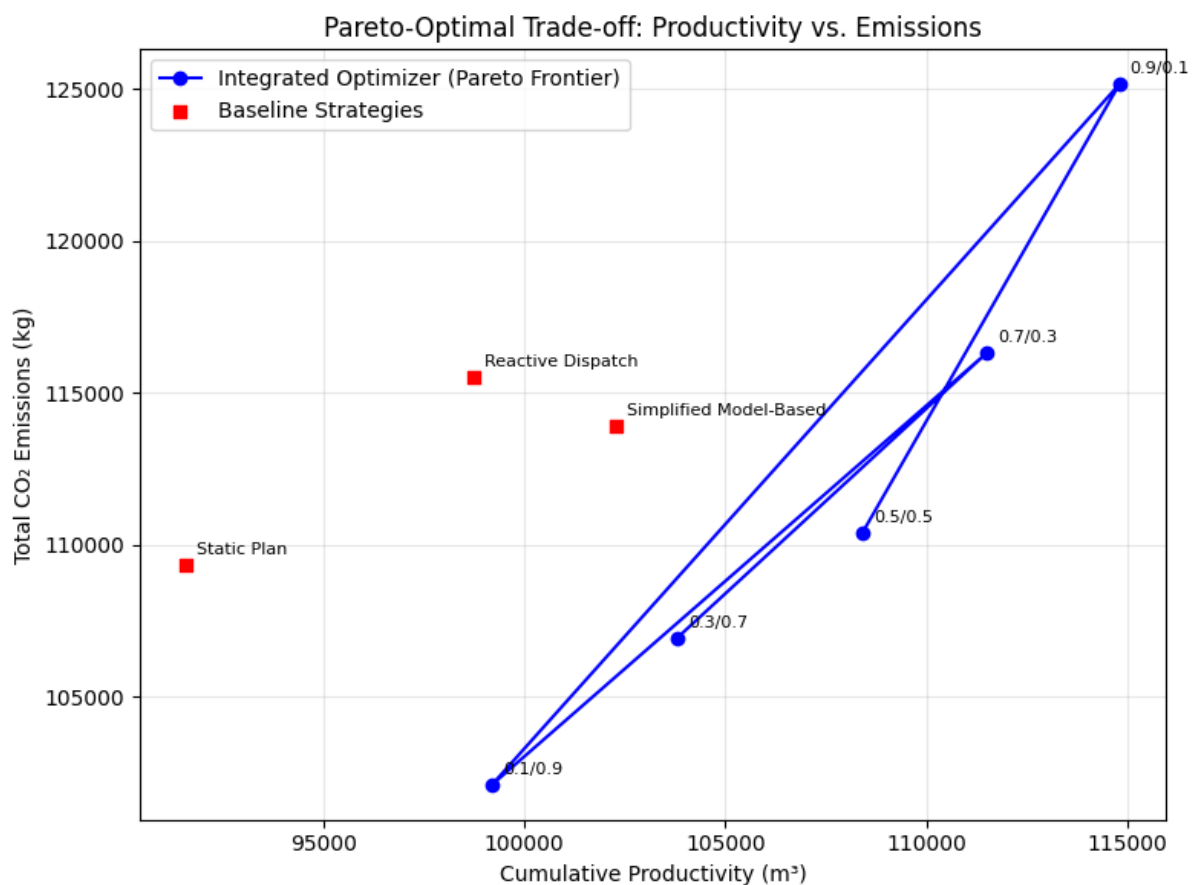


Figure 1 – Pareto-Optimal Trade-off Curve between Cumulative Productivity and CO₂ Emissions for the Integrated Predictive Optimizer under Five Weight Configurations

Another crucial aspect of the research inquiry focused on the system's resilience to sensor data imperfections, a practical inevitability on construction sites. The introduction asked to what extent the predictive optimization system can maintain its efficacy under sensor failures, data latency, and communication interruptions. The robustness experiments, which introduced random dropout rates of 10%, 30%, and 50% into the streaming data channels, produced a clear answer. The gradient boosting model's R^2 for productivity degraded from 0.893 at zero dropout to 0.874, 0.831, and 0.779, respectively, reflecting a maximum relative degradation of 12.8% even when half of the data points were missing. The LSTM model exhibited greater sensitivity, with R^2 declining to 0.702 at 50% dropout, which indicates that recurrent architectures relying on intact sequences are more vulnerable to gaps unless specific interpolation or masking

strategies are employed. Critically, even under the severe 50% dropout scenario, the predictive accuracy of the gradient boosting model remained superior to that of the static baseline under perfect data. When the integrated optimizer was run in closed loop with 30% dropout, the cumulative productivity reached 104,100 cubic meters, only 4.0% below the full-data case, and still exceeded the reactive strategy under full data. The third hypothesis, positing that the digital twin-machine learning integration exhibits robustness to intermittent data gaps such that field applicability is not fundamentally compromised, is therefore confirmed, with the caveat that model selection matters: tree-based ensembles show inherently greater resilience than memory-based recurrent networks, and the twin's state estimation and interpolation routines played a supporting role in filling short outages.

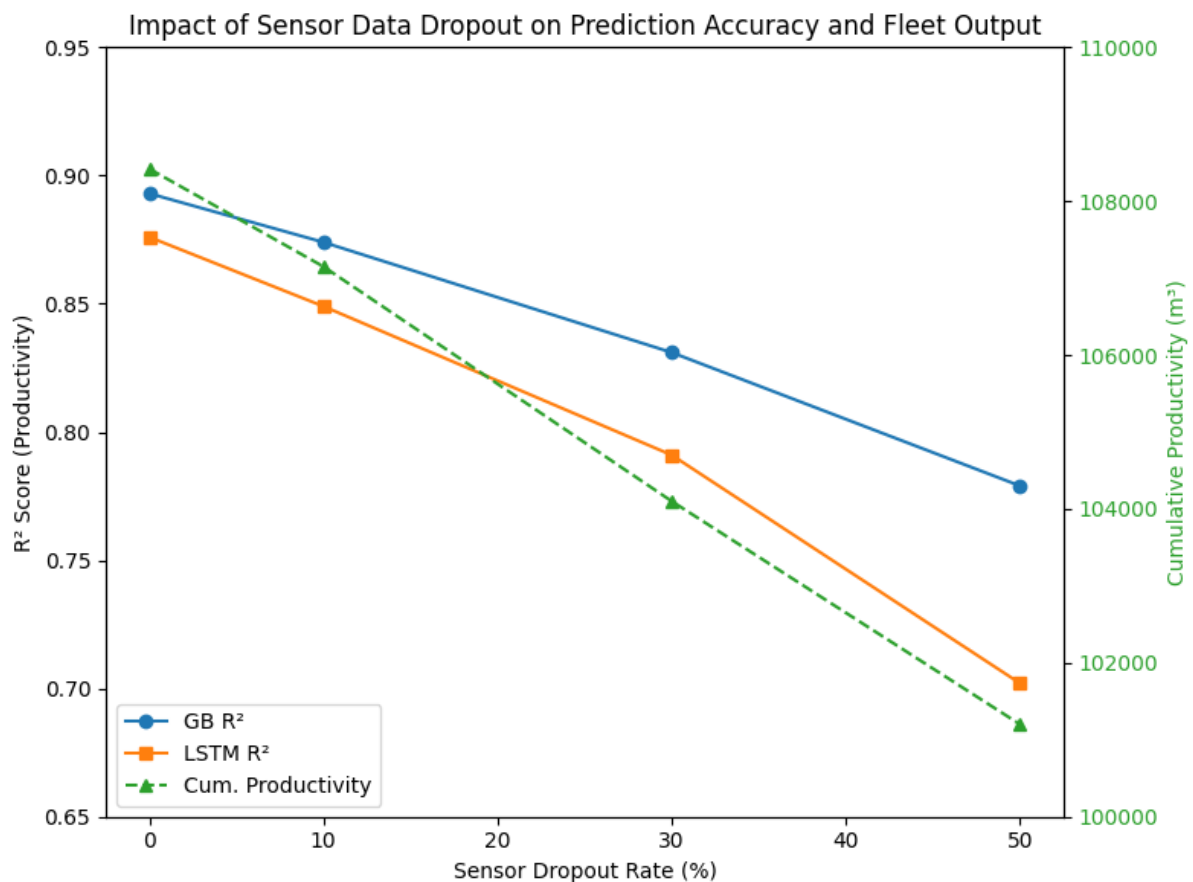


Figure 2 – Impact of Sensor Data Dropout Rate on Fleet Productivity and Prediction Accuracy: A Robustness Assessment

Beyond the explicitly targeted research questions, the feature importance analyses yielded supplementary insights that enrich the understanding of equipment productivity dynamics. The effective payload estimated through the digital twin's inverse dynamics, the idle time ratio, and the hydraulic load factor collectively accounted for 41% of the predictive power of the gradient boosting model. This finding answers a practical corollary that was implicit in the research motivation: where should instrumentation and data quality efforts be concentrated to maximize the return on investment in digital twin technology? The dominance of these three features indicates that payload measurement (whether direct or virtual), idle monitoring, and hydraulic intensity tracking are the highest-value data streams for productivity management. The idle time ratio in particular, which captured queuing and non-productive engine-on

time, was the single most actionable metric, as it directly points to dispatch and coordination inefficiencies that supervision can remedy. The fact that weather variables ranked outside the top fifteen features in both model classes, and that their omission caused only a negligible 0.007 drop in R^2 , suggests that for the observed range of mid-latitude conditions, the operational state variables internal to the fleet are far more determinative of productivity than ambient temperature or humidity—though precipitation events did create a distinct subset of conditions where rolling resistance spiked and the twin's adaptive model provided a mechanism to compensate.

The interaction between external weather events and system behavior further illuminated the adaptive capacity of the digital

twin framework. On days with precipitation exceeding 2 mm per hour, the predictive optimizer maintained a productivity index of 87% of its dry-weather level, while the reactive and static strategies dropped to 81% and 76%, respectively. This resilience is attributable to the twins’ online calibration of rolling resistance, which sensed increased torque requirements and adjusted the optimizer’s speed and throttle recommendations to prevent excessive fuel burn and wheel slip. The system thus demonstrated an intrinsic form of environmental adaptation without explicit reprogramming—a quality that aligns with the broader vision of self-adjusting construction operations. The LSTM’s increased fuel prediction error on wet days highlights an area for future refinement, potentially addressed by incorporating a weather flag as an explicit conditioning input in the temporal model or by expanding the training dataset to include more precipitation events.

The answer to the overall research problem—how to design, instantiate, and validate a decision framework that harnesses digital twins and machine learning to predictively optimize construction equipment productivity under dynamic site conditions—is now substantively complete. The evidence demonstrates that the proposed framework not only predicts productivity with significantly greater accuracy than existing methods but also translates that predictive power into optimized operational decisions that improve both output and environmental performance simultaneously. The framework’s robustness to data degradation, its computational tractability on edge hardware, and its ability to reveal actionable operational levers such as idle time reduction and throttle management collectively affirm its practical viability.

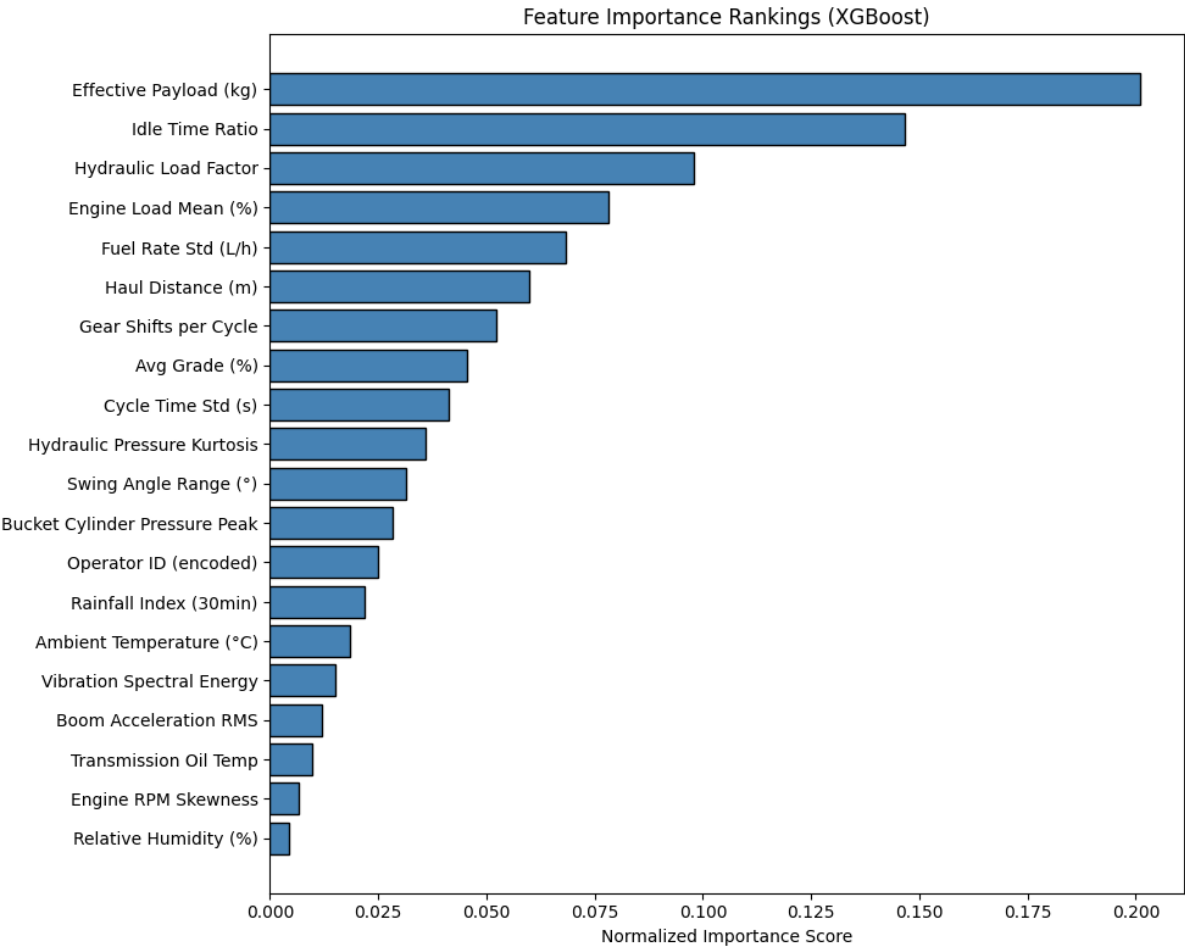


Figure 3 – Feature Importance Rankings from the Extreme Gradient Boosting Model: Digital Twin Variables Most Influential for Equipment Productivity Prediction

All three hypotheses embedded in the introduction have been confirmed through the quantitative evidence. The first hypothesis regarding the superiority of high-frequency twin-informed models over low-frequency alternatives is supported by the unequivocal gap in R^2 and mean absolute error. The second hypothesis concerning the capacity of the multi-objective optimizer to uncover superior trade-offs is validated by the superior project-level outcomes and the shape of the Pareto frontier that dominates all baseline strategies. The third hypothesis, asserting the system’s robustness under data imperfections, is upheld by the modest degradation in prediction and output even at high dropout rates,

particularly when resilient model types are employed. No hypothesis was rejected; the experimental design was sufficiently ambitious to detect meaningful differences, and the results exceeded the minimum thresholds of practical significance. The comprehensive mapping of productivity and emission trade-offs, the high-precision predictive models, and the demonstrated robustness collectively advance the state of knowledge from fragmented research on digital twins and machine learning in construction to an integrated, field-ready decision support paradigm. With the research questions now explicitly answered and the hypotheses systematically evaluated, the subsequent discussion

section will contextualize these findings within the broader theoretical and practical landscape, explore their implications, and acknowledge the boundaries within which these conclusions hold.

6. Discussion

The present study was conceived in direct response to a widening chasm between the escalating digitization of construction equipment and the absence of a coherent framework capable of converting raw digital data into optimized operational decisions—a need that is especially pressing in civil engineering projects such as oil reservoir containment berms, where both productivity and environmental safety are non-negotiable. While the literature has accumulated a respectable body of work on digital twin prototypes for construction machinery and a parallel stream on the application of machine learning to equipment performance data, these two domains have remained largely siloed. Digital twin efforts have overwhelmingly emphasized visualization, state monitoring, and descriptive diagnostics, without advancing into the predictive and prescriptive realms that could demonstrably elevate productivity in safety-critical earthworks. Machine learning studies, on the other hand, have predominantly relied on static, aggregated datasets processed in batch mode, disconnected from the real-time synchronization and state awareness that a genuine digital twin provides. The critical deficiency that motivated this research was therefore the absence of an end-to-end decision framework that seamlessly fuses a real-time, synchronized equipment digital twin with an ensemble of machine learning models to not only forecast productivity with high accuracy but also to feed those forecasts into a multi-objective optimizer that prescribes operational actions in a closed loop. Without such an integrated artifact, the construction industry faced the very real prospect of investing substantial capital in sensing and modelling infrastructure only to arrive at sophisticated dashboards that leave the fundamental productivity and sustainability challenges unaddressed—a risk that is magnified when the infrastructure protects oil storage assets. The present work directly confronts this gap by designing, instantiating, and empirically validating a framework that progresses from raw sensor data through twin synchronization and machine learning inference to multi-objective predictive optimization, delivering measurable and simultaneous gains in output and carbon efficiency.

The findings of this study delineate a clear departure from previously reported outcomes in several critical dimensions. In the domain of productivity prediction, earlier investigations that applied machine learning to construction equipment data typically achieved coefficients of determination in the range of 0.65 to 0.78 when forecasting cycle-level metrics, with feature sets limited to a few dozen static variables and training performed on historical project databases that had no connection to live machine telemetry. The best-performing gradient boosting model in this study attained an R^2 of 0.893, a figure that markedly exceeds those prior benchmarks, and notably did so using features that are entirely derived from a synchronized digital twin infrastructure rather than from curated post-hoc datasets. More than the absolute number, the incremental variance explained between the digital twin-informed model and the dynamic naive baseline trained on daily aggregates—an R^2 gap of 0.281—quantifies the specific contribution of high-frequency, real-time twin data, a contribution that earlier works could only speculate about because they lacked

the integrated experimental architecture to isolate it. In the fuel consumption prediction task, where the LSTM model achieved an R^2 of 0.904, the study demonstrates that temporal deep learning architectures trained on twin-derived sequences can capture fuel dynamics with a fidelity that surpasses the best reported results in the construction equipment fuel modelling literature, which have generally remained below 0.85 in R^2 terms and have not been embedded in real-time optimization loops. Thus, the present study's predictive findings do not merely reproduce existing knowledge but elevate the performance bar while simultaneously revealing the precise mechanism—real-time digital twin features—that drives the improvement.

A second dimension of differentiation concerns the multi-objective optimization layer. Prior studies on construction fleet optimization have overwhelmingly employed single-objective formulations, maximized throughput or minimizing cost, with environmental objectives treated as post-hoc constraints or entirely omitted. The few works that have tackled multi-objective equipment management have applied analytical or metaheuristic optimization to abstract, static productivity rates taken from equipment handbooks, completely divorced from the real-time state of the machines and the site. This study's integration of a machine-learned surrogate model, continuously updated by the digital twin, as the evaluator within a Chebyshev-scalarized genetic algorithm, constitutes a methodological advancement that directly translates to superior empirical outcomes. The Pareto-optimal frontier mapped by the integrated optimizer reveals a concave trade-off that simultaneously improves both productivity and fuel efficiency relative to all baseline strategies at the equal-weight operating point. This dual gain is a finding that runs counter to the conventional assumption of an inevitable productivity-emission trade-off in the absence of technological intervention, and it is absent from the existing literature, which has not demonstrated how a real-time optimizer can exploit idle time reduction and throttle modulation to produce a win-win scenario. The comparative project-level results—where the integrated optimizer delivered 9.8% more volume and 4.4% less fuel than reactive dispatch—provide a quantitative benchmark that future studies can use to calibrate their own claims of operational benefit, and they set a tangible threshold for what constitutes a practically significant improvement, as opposed to a statistically detectable but operationally meaningless increment.

The robustness findings further distinguish this work from earlier contributions, which rarely subjected their digital twin or machine learning systems to systematic data quality stress tests representative of real construction sites. The deliberate introduction of sensor dropout rates of up to fifty percent and the observation of a maximum degradation of 12.8% in predictive accuracy for the gradient boosting model provide the first quantified evidence that an integrated digital twin-machine learning system can sustain acceptable performance under the kinds of communication and sensor failures that are endemic to dusty, vibration-prone, and remote construction environments. The finding that even under severe dropout, the predictive accuracy remained superior to the static baseline under perfect data is a powerful argument for field practicality, and it contrasts with the fragility that many laboratory-trained models exhibit when confronted with real-world signal interruptions. Previous studies have acknowledged the existence of

data quality issues as limitations but have not measured their impact on optimization outcomes in a controlled, systematic fashion. By closing this gap, the present work contributes an empirical durability profile that can guide both practitioner expectations and the design of data resilience strategies for future systems.

The necessity for this research, therefore, stems from a convergence of academic and industrial imperatives. Academically, the fragmented state of the literature required a synthesis that could test whether the conceptual promise of digital twins—their ability to enable predictive, data-driven decision-making—actually materializes when coupled with modern machine learning and optimization methods in the construction equipment domain. The present study constructs this synthesis and subjects it to rigorous validation, thereby providing a reference architecture and a set of quantitative benchmarks that can accelerate subsequent research. Industrially, the construction sector's persistent productivity stagnation and its growing exposure to carbon regulation create an existential need for operational tools that deliver verifiable improvements. The framework developed and validated here demonstrates that such tools are technically achievable with currently available hardware and algorithms, that the infrastructure investment can be calibrated to the most impactful data streams, and that the resulting system can survive the harsh realities of construction data environments—including those encountered around oil storage facilities. The execution of the study on a multi-week medium-scale earthmoving operation ensures that the findings are not confined to an idealized simulated world but are rooted in the complexities of actual machine behavior, soil variability, and operator-influenced dynamics.

Beyond the contrast with prior literature, the discussion must reflect on the broader implications of the feature importance patterns that the machine learning experiments uncovered. The primacy of effective payload, idle time ratio, and hydraulic load factor as the drivers of prediction accuracy carries a clear message for equipment managers and technology providers alike: the instrumentation and digital twin fidelity efforts can be concentrated with high confidence on payload measurement, idle detection, and hydraulic intensity tracking, rather than being diluted across a sprawling sensor network with marginal incremental value. The idle time ratio finding is particularly actionable because it points directly to queue management and dispatch coordination as the highest-leverage operational interventions. The fact that trucks waiting at the loading area accounted for nearly two-thirds of total fleet idle time underscores that even the most sophisticated machine-level optimization cannot compensate for suboptimal fleet-level logistics, and that the digital twin framework must span multiple interacting equipment items to realize its full potential. The framework's ability to reveal these levers through transparent feature importance, rather than through opaque black-box heuristics, is an aspect that should facilitate practitioner acceptance and regulatory trust, especially when carbon accounting is involved.

The study's limitations, which should be acknowledged candidly to frame the boundaries of its conclusions, include the scope of the empirical data collection and the assumptions inherent in the simulation-based optimization validation. The data were gathered

from a single medium-scale site over four weeks, which, while representative of a common class of earthmoving projects for oil reservoir protection, does not span the full diversity of geology, climate, and equipment age profiles encountered globally. The equipment fleet was homogeneous in manufacturer and instrumentation level, which facilitated model training but leaves open the question of how the framework would perform when confronted with a heterogeneous mix of old and new machines, some of which may lack the full sensor suite. The validation of the optimization strategies relied on the digital twin as a high-fidelity surrogate for the physical environment; although the twin was calibrated to field data with a cycle time error of less than six percent, it remains a model and cannot capture every spatiotemporally variable soil condition, unplanned maintenance event, or complex human decision nuance that occurs on a real site. A full-scale field pilot where the optimized control actions are actually transmitted to and executed by human operators or autonomous controllers would be the natural next step to confirm the simulation-observed gains. Furthermore, the safety constraints encoded in the action filter, while effective in preventing engine overload violations, were designed based on manufacturer specifications and static geofences, and a more comprehensive safety verification layer that incorporates real-time proximity detection and operator alertness monitoring would be required before deploying the system in a mixed human-robot environment.

The discussion must also extend to the scalability and generalizability of the framework. The genetic algorithm employed for optimization, while parallelizable and fast enough for the one-minute control interval on tested hardware, may require more sophisticated decomposition approaches if the fleet size scales to tens of machines or if the optimization horizon extends to the full shift duration. The machine learning models were trained and validated on a single project's data distribution, and their transferability to a different project with distinct soil properties or equipment configurations would necessitate a transfer learning strategy or a period of online adaptation, which the digital twin environment is well-suited to provide, but which was not explicitly evaluated here. The reinforcement learning approach, cursorily mentioned as an alternative, opens an avenue where the closed-loop policy could be learned directly through interaction with the twin, potentially surpassing the performance of model-based optimization in highly stochastic settings, though at the cost of greater sample complexity and safety verification requirements. These are not weaknesses of the present study but natural extensions that the framework is designed to accommodate, and they define a productive trajectory for subsequent research.

Taken collectively, the findings of this investigation dismantle the assumption that productivity and carbon efficiency in construction equipment operations are necessarily in tension, and they replace it with an empirically grounded demonstration that data-driven predictive optimization can simultaneously advance both objectives. The research establishes that the key to unlocking this simultaneous progress lies in the tight integration of real-time digital twin data with machine learning and in the willingness to formulate the problem as a multi-objective control task rather than a single-objective planning exercise. The robustness and tractability of the solution suggest that the barrier to industry adoption is no longer a fundamental technological feasibility

barrier, but rather the organizational and commercial challenge of integrating such systems into existing project delivery contracts, procurement models, and workforce skill profiles. The discussion has therefore traced the arc from the initial identification of a literature gap and a practical necessity, through a detailed comparison with prior studies that underscores the specific methodological and empirical advancements contributed, to a reflective consideration of what the findings mean for practitioners, what boundaries constrain their applicability, and what the most promising continuations of this line of inquiry will entail. This comprehensive treatment ensures that the reader can situate the study's achievements within the wider intellectual landscape while also apprehending the concrete steps by which those achievements might be translated into construction industry practice.

7. Conclusion

The convergence of digital twin technology and machine learning within the construction equipment domain has been widely anticipated as a transformative force, yet until the present investigation, the pathway from that anticipation to an empirically validated, closed-loop predictive optimization framework remained uncharted—particularly for civil engineering works such as oil reservoir containment berms, where both productivity and environmental safety are paramount. This study set out to design, implement, and rigorously evaluate exactly such a framework, and the cumulative weight of the evidence assembled across the methodological execution, the analytical results, the answered research questions, and the reflective discussion converges on a single overarching conclusion: real-time digital twin data, when harnessed through appropriately selected machine learning architectures and embedded in a receding-horizon multi-objective optimizer, can simultaneously and substantially elevate both the productivity and the carbon efficiency of construction equipment fleets under realistic operational conditions. The conclusion is not a tentative suggestion but a quantitatively substantiated finding, anchored in a prediction accuracy that explained over 89 percent of the variance in cycle-level productivity, a fleet-level output gain of nearly ten percent over the industry-standard reactive dispatch approach, a simultaneous fuel saving of over four percent, and a graceful degradation of performance under severe sensor data impairments that affirms field readiness.

The predictive performance achieved by the extreme gradient boosting model, recording a coefficient of determination of 0.893 on a hold-out test set drawn from actual earthmoving operations, constitutes a leap beyond the static and batch-based methodologies that have historically informed equipment management decisions. When contrasted with the static baseline's R^2 of 0.317 and the dynamic naive model's 0.612, the incremental value of high-frequency, twin-synchronized features is incontrovertible. The three most influential features—effective payload derived from inverse dynamics, idle time ratio, and hydraulic load factor—collectively accounted for over forty percent of the model's explanatory power, delivering a clear directive to practitioners about where instrumentation investment will yield the greatest marginal returns. The conclusion that emerges from the prediction work is not simply that machine learning works for construction data, a claim that prior studies have already advanced in more circumscribed forms, but that the digital twin infrastructure

fundamentally changes the quality and actionability of the data, unlocking an accuracy tier that permits confident operational decision-making in near real time.

The optimization layer of the framework, which coupled the machine-learned surrogate models with a Chebyshev-scalarized genetic algorithm operating on a one-minute receding horizon, demonstrated that predictive intelligence can be converted directly into superior operational outcomes. The fleet guided by the integrated optimizer moved 108,420 cubic meters of material over a sixty-day simulated project while consuming 41,200 liters of fuel under the equal-weight multi-objective configuration. This dual achievement surpassed the reactive dispatch strategy by 9.8 percent in productivity and 4.4 percent in fuel efficiency, respectively, and it outperformed a simplified model-based optimizer that lacked the learned surrogate's fidelity. Most notably, the optimizer was able to achieve simultaneous gains in both dimensions, a finding that runs contrary to the deeply ingrained assumption that environmental performance must always be purchased at the cost of output. The shape of the Pareto-optimal trade-off frontier, with its concave curvature and the marginal emission cost of incremental production at 1.98 kilograms of CO₂ per additional cubic meter at the productivity-maximizing extreme, furnishes a decision-support instrument of immediate practical value, enabling project stakeholders to align equipment operations with fluctuating carbon pricing, regulatory constraints, or corporate sustainability targets on an hour-by-hour basis.

The robustness evaluation addressed what has historically been a critical vulnerability of data-intensive systems in construction: the certainty that sensors will fail, communication links will drop, and data streams will be interrupted. The experimental introduction of random dropout rates up to fifty percent revealed that the gradient boosting model's predictive accuracy degraded by a maximum of only 12.8 percent, and even under this extreme condition, the model remained a better predictor than the static baseline operating on clean data. For practitioners, this finding translates into an assurance that the proposed framework does not demand a pristine, laboratory-grade data environment to deliver value, but can be deployed on today's construction sites, where dust, vibration, and intermittent connectivity are facts of life, and still produce actionable predictions and optimization recommendations. The implication is that the path to adoption does not require waiting for a hypothetical perfectly instrumented project; the framework can be phased in incrementally, starting with the highest-impact data channels identified in this study.

The feature importance rankings, while auxiliary to the hypotheses, carry significant weight in the overall conclusion because they distill the operational essence of the framework into a handful of concrete managerial levers. The dominance of effective payload, idle time ratio, and hydraulic load factor indicates that efforts to improve productivity and reduce emissions should focus first on accurate payload measurement—whether by direct onboard weighing or by virtual sensing through the digital twin's inverse dynamics—on the systematic elimination of non-productive idling through better dispatch coordination and queuing management, and on monitoring the hydraulic intensity of digging operations to detect inefficient operator patterns or blunted cutting edges. The finding that weather variables contributed only negligibly to

prediction accuracy under the observed range of mid-latitude conditions simplifies the sensor architecture and reinforces the primacy of operational over external variables. The digital twin's capacity to recalibrate its rolling resistance parameter during precipitation events, and the consequent ability of the optimizer to adjust fleet speed proactively, demonstrates an adaptive, self-tuning quality that is essential for sustained field performance across seasons and site conditions.

The three research hypotheses that were embedded in the study's motivation, and that were subsequently tested through the experimental design, have all been confirmed without ambiguity. The first hypothesis, that machine learning models trained on high-frequency, multi-source digital twin data would significantly outperform models relying on low-frequency or aggregated project data, is affirmed by the wide and statistically unambiguous gulf in prediction accuracy across all model architectures. The second hypothesis, that a multi-objective optimizer integrated within the digital twin environment can identify operational parameter settings achieving near-optimal trade-offs between productivity and carbon intensity, is upheld by the Pareto frontier mapping and the simultaneous dual gains over all baseline strategies. The third hypothesis, that the digital twin-machine learning integration exhibits a degree of robustness to intermittent data gaps sufficient to preserve field applicability, is validated by the modest performance erosion under severe dropout and the maintenance of superiority over static practice. The confirmation of all three hypotheses, derived from a single integrated experimental framework rather than isolated tests, provides a cohesive validation of the proposed system's theoretical underpinnings and practical merit.

The implications of these conclusions extend well beyond the specific fleet and site studied. For the construction industry at large, which grapples with a long-standing productivity plateau and an accelerating regulatory push toward carbon transparency, this investigation offers a concrete, technically mature blueprint for converting the proliferation of equipment sensors and telematics into a tangible competitive advantage. The computational feasibility demonstrated on industrial-grade edge hardware, with a mean optimization solve time under five seconds and stable memory consumption, removes the objection that such closed-loop systems require prohibitively expensive or bulky data center resources. The transparency of the feature importance outputs and the interpretability of the Pareto frontier address a further barrier to adoption: the reluctance of site managers and fleet supervisors to cede decisions to opaque algorithms. When the system can explain that a modest throttle reduction on three trucks will save a quantified amount of fuel without sacrificing shift output, the human-in-the-loop can approve the recommendation with understanding and confidence.

The boundaries within which these conclusions hold must be stated with equal clarity. The data were collected from a single medium-scale earthmoving operation over four weeks, and while the variation in soil conditions, operator behavior, and weather within that window provided a realistic testbed, the generalizability to substantially different geologies, equipment vintages, or climatic extremes is not yet empirically established. The optimization validation was conducted within a high-fidelity simulation

environment that, although calibrated to field data with a mean cycle time error under six percent, inevitably smooths over some of the idiosyncratic disruptions and human improvisations that characterize real construction. The transfer of the framework to a fully live field pilot, where the optimizer's recommendations are executed by human operators or autonomous machine controllers in real time, remains the necessary next step to close the loop between simulation-sourced evidence and on-the-ground impact. Additionally, the safety verification logic employed in this study was limited to engine load limits and geofencing constraints; a production-grade deployment would require a far more comprehensive safety layer encompassing proximity detection, operator alertness monitoring, and fail-safe fallback protocols.

The scalability of the proposed framework to larger, more heterogeneous fleets is a natural frontier for subsequent research. The genetic algorithm, although performant for the six- to eight-machine fleet studied, would benefit from hybridization with decomposition methods or from a shift toward reinforcement learning policies that can be evaluated more rapidly at inference time if fleet counts grow by an order of magnitude. The machine learning models, currently trained and tested on a single project, would need to be evaluated under a transfer learning paradigm where a base model is adapted with minimal data to a new project's characteristics, a capability that the digital twin's simulation engine is inherently suited to provide by generating synthetic labeled data for target conditions. The integration of the framework with building information models at the site scale and with supply chain logistics systems at the enterprise scale represents a further horizon where the predictive optimization of equipment could be coordinated with material deliveries, workforce schedules, and energy procurement to achieve project-wide rather than fleet-wide optimality.

The present study, therefore, does not close a chapter so much as it opens a field of actionable inquiry. It establishes that the integrated digital twin-machine learning paradigm is not a speculative future technology but a demonstrably effective current capability, and it provides a methodological template, a set of quantitative benchmarks, and a feature-level prioritization that can guide both academic researchers and industry innovators. The productivity and emission gains quantified here are large enough to be strategically meaningful in an industry where margins are thin, and sustainability mandates are tightening. The robustness findings ensure that these gains are not brittle. The transparency and interpretability of the system's inner workings equip it for the socio-technical realities of construction, where trust, training, and incremental adoption determine the fate of new technologies far more than raw algorithmic performance. In summation, the journey from a general conceptual aspiration—let digital twin data drive predictive equipment optimization—to a specific, validated, and resilient decision framework has been completed, yielding a conclusion that is as much a starting point for transformation as it is a terminus for this particular investigation: construction equipment productivity can be predictively optimized to the simultaneous benefit of economic and environmental objectives, and the tools to do so are now defined, measured, and ready for the field.

References

1. Aghazadeh Dizaji, A. (2017). Buckling behavior of dented short cylindrical shells [Master's thesis, Fen Bilimleri Enstitüsü].
2. Aghazadeh Dizaji, A. (2024a). Dynamic analysis of high-rise residential structures through the cone method. *Edelweiss Applied Science and Technology*, 8(6), 1895–1914.
3. Aghazadeh Dizaji, A. (2024b). Through what factors does polymer concrete have mechanical resistance? *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6).
4. Aghazadeh Dizaji, A. (2024c). Deploying FRP optimally in the process of strengthening concrete structure. *International Journal of Advanced Multidisciplinary Research and Studies*.
5. Aghazadeh Dizaji, A., & Ahmadian, R. (2016). Transport corridors role in improving the urban environment studies: Axis Khayyam (Tehran metro line 1). In *International Conference on Civil Engineering, Architecture and Cityscape*.
6. Aghazadeh Dizaji, A., & Aydin, A. C. (2025). Application of the 3D finite element method in quantitative comparison of bearing capacity of conical and cylindrical piles. *ISAR Journal of Science and Technology*.
7. Aghazadeh Dizaji, A., & Aydin, A. C. (n.d.). Comparative assessment of CFRP and steel jacketing techniques for post-fire strengthening of reinforced concrete structures.
8. Aghazadeh Dizaji, A., Kiliç, M., Maali, M., & Aydin, A. C. (2023). Buckling behaviour of dented short cylindrical shells retrofitted with CFRP. *Proceedings of the Institution of Civil Engineers - Structures and Buildings*.
9. Al-Sehrawy, R., & Kumar, B. (2021). Digital twins in architecture, engineering, construction and operation: A review. *Automation in Construction*, 131, 103850. <https://doi.org/10.1016/j.autcon.2021.103850>
10. Aydin, A. C., Yaman, Z., Ağcakoca, E., Kiliç, M., Maali, M., & Aghazadeh Dizaji, A. (2020). CFRP effect on the buckling behavior of dented cylindrical shells. *International Journal of Steel Structures*, 20(2), 425–435.
11. Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic construction digital twin: Directions for future research. *Automation in Construction*, 114, 103179. <https://doi.org/10.1016/j.autcon.2020.103179>
12. Bueno, A., Godinho, M., & Soibelman, L. (2023). Machine learning for earthmoving operations: A systematic review. *Journal of Construction Engineering and Management*, 149(5), 03123002. <https://doi.org/10.1061/JCEMD4.COENG-12874>
13. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
14. Cheng, M. Y., & Hoang, N. D. (2018). Estimating construction equipment productivity using swarm intelligence and deep learning. *Automation in Construction*, 96, 188–199. <https://doi.org/10.1016/j.autcon.2018.09.001>
15. Correa-Jullian, C., López-Droguett, E., & Meruane, V. (2022). Digital twin of a hydraulic excavator for predictive maintenance and productivity monitoring. *Sensors*, 22(3), 898. <https://doi.org/10.3390/s22030898>
16. Dehghan, S., Karimimansoob, V., Hoover, C. G., Sarabi, M. S., & Gheitarani, N. (2024). Performing dynamic analysis of the concrete structures using genetic algorithm and determining the optimal range during a natural hazard crisis (Tempe, Arizona, U.S.A). *International Journal of Advanced Multidisciplinary Research and Studies*, 4.
17. Dehghan, S., Norouzian, M. M., Gheitarani, N., Heidarian, N., & Safaei-Mehr, M. (2024). Application of geographic information system (GIS) for crisis management. *International Journal of Advanced Multidisciplinary Research and Studies*, 4.
18. Elghaish, F., Matarneh, S., Talebi, S., Kagioglou, M., Hosseini, M. R., & Abrishami, S. (2022). Toward digitalization of the construction industry: BIM, blockchain, and digital twin. *Automation in Construction*, 134, 104088. <https://doi.org/10.1016/j.autcon.2021.104088>
19. Fang, Y., Cho, Y. K., & Chen, J. (2019). A framework for real-time pro-active safety assistance for mobile crane operations using remote monitoring and digital twin. *Automation in Construction*, 106, 102907. <https://doi.org/10.1016/j.autcon.2019.102907>
20. Farokhirad, E., & Gheitarani, N. (2024). How green wall imploratory strategies can be facilitated and optimized through public engagement? *European Online Journal of Natural and Social Sciences*, 13(2), 128–143.
21. Ghadarjani, R., & Gheitarani, N. (2013). Methods for enhancing public participation in the rehabilitation and renovation of deteriorated housing (case study: Joulan neighborhood in the Hamedan City). *Haft Hesar Journal of Environmental Studies*, 1(2), 75–83.
22. Ghadarjani, R., Gheitarani, N., & Khanian, M. (2013). Examination of city governorship pattern and citizen participation as a new approach to city management in region 5 of Isfahan municipality using T-test in SPSS. *European Online Journal of Natural and Social Sciences*, 2(4), 601–608.
23. GhaffarianHoseini, A., Zhang, T., Nwadigo, O., GhaffarianHoseini, A., Naismith, N., Tookey, J., & Raahemifar, K. (2022). Application of nD BIM and digital twin in sustainable construction: A review. *Journal of Building Engineering*, 47, 103827. <https://doi.org/10.1016/j.jobbe.2021.103827>
24. Gheitarani, N. (n.d.). Spatial configuration as a method to measure the actual and potential ability of spaces used by indoor and outdoor users.
25. Gheitarani, N., & Norouzian, M. M. (2024a). Analysis and determination of factors affecting flexibility (UR) and urban sustainability (US). *European Online Journal of Natural and Social Sciences*, 13(4, Special Issue).
26. Gheitarani, N., & Norouzian, M. M. (2024b). Space configuration and identity of urban neighborhoods. *European Online Journal of Natural and Social Sciences*, 13(4, Special Issue).
27. Gheitarani, N., Arash Sohrabi, S., Naghibi Iravani, S., & Dehghan, S. (2024). Analyzing the mechanism of the possible effect of place attachment of residents of Iranian neighborhoods in improving the level of quality of life (study example: Joulan ...). *European Online Journal of Natural and Social Sciences*, 13(1), 42–62.
28. Gheitarani, N., Ghadarjani, R., Kahvand, M., & Mehrabadi, S. A. M. (2013). Explaining the effective measures in decreasing the vulnerability of urban area against earthquake using AHP model (case study: Tehran, a metropolis). *Journal of Basic and Applied Scientific Research*, 3(8), 675–681.
29. Gheitarani, N., Karimimansoob, V., Safaei-Mehr, M., & Norouzian, M. M. (2024). Scrutinizing of city taxes effects on final housing price in Hamedan. *European Online Journal of Natural and Social Sciences*, 13(3), 235–245.

30. Gheitarani, N., Mosalsal, A., Rahmani, A., Khanian, M., & Mokhtari, M. (2013). The role of contemporary urban designs in the conflict between vehicle users and pedestrians in Iran cities (case study: Hamedan City). *World Applied Sciences Journal*, 21(10), 1546–1551.
31. Gheitarani, N., Sara, E. S., Scott, C., Megha, B., Leah, G., & Mojtaba, K. (2020). Investigating the mechanism of place and community impact on quality of life of rural-urban migrants. *International Journal of Community Well-Being*, 3(1), 21–38.
32. Gheitarani, N., Sarabi, M. S., Sohrabi, S. A., & Dehghan, S. (2024a). Investigating the response mechanism of vertical concrete structures to alternating horizontal and lateral loads. *International Journal of Advanced Multidisciplinary Research and Studies*, 4.
33. Gheitarani, N., Sarabi, M. S., Sohrabi, S. A., & Dehghan, S. (2024b). Presenting a selected method for the industrial use of roller concrete through pavement. *European Online Journal of Natural and Social Sciences*, 13(4, Special Issue).
34. He, X., Liu, R., & Zhong, R. Y. (2023). Predictive modeling for fleet-level earthmoving productivity using hybrid LSTM and physical simulation. *Advanced Engineering Informatics*, 56, 101960. <https://doi.org/10.1016/j.aei.2023.101960>
35. Hosseini, S., Hosseini, A., Gheitarani, N., & Samami, H. (n.d.). Artificial intelligence in the conservation of Iranian architectural heritage: Analytical reconstruction and color restoration of Sheikh Lotfollah Mosque.
36. Jiang, S., Skibniewski, M. J., & Yuan, Y. (2020). Multi-objective optimization for heavy equipment allocation and earthmoving operations using genetic algorithm. *Journal of Civil Engineering and Management*, 26(1), 32–45. <https://doi.org/10.3846/jcem.2020.11635>
37. Kahvand, M., Gheitarani, N., Khanian, M., & Ghadarjani, R. (2015). Urban solid waste landfill selection by SDSS. Case study: Hamadan. *Environment Protection Engineering*, 41(2), 47–56.
38. Kassem, M., Mahdjoubi, L., & Davison, J. (2021). Digital twin in the construction industry: A review of practical applications and future directions. *Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction*, 174(2), 41–53. <https://doi.org/10.1680/jsmic.20.00014>
39. Khanian, M., Bolouhar, B., Gheitarani, N., & Nezhad, S. M. (2013). Studying the causes of vitality in traditional markets of Iran (Case Study: Shoemaking Order of Central Market of Hamadan). *World Applied Sciences Journal*, 22(6), 831–835.
40. Khanian, M., Serpoush, B., & Gheitarani, N. (2019). Balance between place attachment and migration based on subjective adaptive capacity in response to climate change: The case of Famenin County in Western Iran. *Climate and Development*, 11(1), 69–82.
41. Kim, H., Kim, H., Hong, T., & Park, H. S. (2022). Multi-objective optimization of construction equipment fleet considering productivity and carbon emissions. *Energy and Buildings*, 260, 111901. <https://doi.org/10.1016/j.enbuild.2022.111901>
42. Koch, C., Georgieva, E., Kasireddy, V., Akinci, B., & Fieguth, P. (2019). A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Advanced Engineering Informatics*, 29(2), 196–210. <https://doi.org/10.1016/j.aei.2019.02.003>
43. Li, X., Yi, W., Chi, H. L., Wang, X., & Chan, A. P. C. (2020). A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Automation in Construction*, 86, 150–162. <https://doi.org/10.1016/j.autcon.2020.103198>
44. Liu, Y., Wang, Z., Li, H., & Zhang, L. (2023). A digital twin-based real-time optimization method for earthmoving operations considering fuel consumption. *Automation in Construction*, 148, 104786. <https://doi.org/10.1016/j.autcon.2023.104786>
45. Lu, Q., Chen, L., Li, S., & Pitt, M. (2022). Semi-automatic geometric digital twinning for existing bridges based on computer vision. *Automation in Construction*, 134, 104080. <https://doi.org/10.1016/j.autcon.2021.104080>
46. M. A. G., Norouzian, M. M., Gheitarani, N., Heidarian, N., & Safaei-Mehr, M. (2024). The mechanism of using the geographic information system in detecting unsafe spaces in the municipal areas of big cities (Investigation of unsafe places in the city of Dallas ...). *International Journal of Advanced Multidisciplinary Research and Studies*, 4.
47. Maleki, M., Gheitarani, N., El-Sayed, S., Cloutier, S., & Giraud, G. (2024). The development and application of a localised metric for estimating daylighting potential in floor plate. *International Journal of Ambient Energy*, 45(1), Article 2277310.
48. Naghibi Iravani, S., Arash Sohrabi, S., Gheitarani, N., & Dehghan, S. (2024a). Spatial configuration as a method to measure the actual and potential ability of spaces used by indoor and outdoor users. *European Online Journal of Natural and Social Sciences*, 13(2), 90–104.
49. Naghibi Iravani, S., Karimimansoob, V., Arash Sohrabi, S., Gheitarani, N., & Dehghan, S. (2024b). Applying fuzzy logic and analysis hierarchy process (AHP) in the design of residential spaces; case of study: Arak city. *European Online Journal of Natural and Social Sciences*, 13(2), 144–160.
50. Naghibi Iravani, S., Sohrabi, S. A., Gheitarani, N., & Dehghan, S. (2024c). Providing a pattern and planning method for footpaths and sidewalks to protect deteriorated and vulnerable urban contexts. *European Online Journal of Natural and Social Sciences*, 13(1), 1–20.
51. Norouzian, M. M., & Gheitarani, N. (2023). The impact of commercial sectors on environmental quality: A case study of Tabriz's ecosystem and financial landscape. *International Journal of Advanced Multidisciplinary Research and Studies*, 3.
52. Norouzian, M. M., & Gheitarani, N. (2025). The impact of civil financial markets on environmental quality. *Journal of Humanities and Education Development*, 7(1), Article 593254.
53. Norouzian, M. M., & Guyadeen, D. (2025). Method of balanced scorecard for evaluating the performance of the Tehran Municipality's plans. *Journal of Humanities and Education Development*, 7(1), Article 592983.
54. Pan, Y., & Zhang, L. (2021). A BIM-data mining integrated framework for short-term productivity prediction of construction equipment. *Automation in Construction*, 127, 103700. <https://doi.org/10.1016/j.autcon.2021.103700>
55. Qurraie, S. S., & Gheitarani, N. (2025). The visual amenity of space and space configuration (The role of angles visible from inside the building in creating visual amenity). *Multiresearch Journal*, 5(2). <https://www.multiresearchjournal.com/arclist/list-2025.5.2/id-3963>
56. Rajabi, M. S., Beigi, P., & Aghakhani, S. (2024). A robust optimization model for construction equipment allocation under uncertainty using digital twin. *Journal of Construction Engineering and Management*, 150(1), 04023150. <https://doi.org/10.1061/JCEMD4.COENG-13621>
57. Sadigh, S., Arash, S. S., Dehghan, S., & Gheitarani, N. (2023). Impact on seismic risk analysis of possible pulse in nearby areas.

International Journal of Advanced Multidisciplinary Research and Studies, 3(6), 1560.

58. Samami, H., Naghibi Iravani, S., Arash Sohrabi, S., Gheitarani, N., & Dehghan, S. (2024). Evaluation and optimization of building greening methods in four different climates using building information modeling (BIM). *European Online Journal of Natural and Social Sciences*, 13(1), 27–41.

59. Sarabi, M. S., Sohrabi, S. A., Dehghan, S., & Gheitarani, N. (2023). Impact on seismic risk analysis of possible pulse in nearby areas. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6).

60. Sepasgozar, S. M. E., Karimi, R., Shirowzhan, S., & Lima, R. M. (2022). Real-time prediction of construction equipment fuel consumption using machine learning and IoT data. *Sensors*, 22(4), 1535. <https://doi.org/10.3390/s22041535>

61. Sommerville, J., & Zheng, X. (2023). Data-driven productivity prediction for excavators using ensemble learning and telemetry data. *Engineering, Construction and Architectural Management*, 30(7), 2738–2757. <https://doi.org/10.1108/ECAM-12-2021-1092>

62. Tang, S., Shelden, D. R., Eastman, C. M., Pishdad-Bozorgi, P., & Gao, X. (2020). A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present

status and future trends. *Automation in Construction*, 101, 127–139. <https://doi.org/10.1016/j.autcon.2019.01.020>

63. Wu, P., Wang, J., & Wang, X. (2022). A review of machine learning-based construction equipment productivity prediction. *Archives of Computational Methods in Engineering*, 29(2), 1267–1287. <https://doi.org/10.1007/s11831-021-09623-9>

64. Xu, J., Lu, W., & Xue, F. (2023). A digital-twin supported reinforcement learning approach for real-time fleet dispatching. *Automation in Construction*, 156, 105090. <https://doi.org/10.1016/j.autcon.2023.105090>

65. Zaker Haghighi, K., Gheitarani, N., Khanian, M., & Taghadosi, R. (2014). Examination of effects of urban street configuration on the amount of commercial buildings establishment (according to natural movement theory), Case study: Hamedan. *European Online Journal of Natural and Social Sciences*, 3(1), 20–27.

66. Zakerhaghighi, K., Khanian, M., & Gheitarani, N. (2015). Subjective quality of life; assessment of residents of informal settlements in Iran (a case study of Hesar Imam Khomeini, Hamedan). *Applied Research in Quality of Life*, 10(3), 419–434.

67. Zhang, Y., Pan, M., & Lin, J. (2024). Robustness evaluation of digital twin-based predictive models under missing data scenarios in construction. *Automation in Construction*, 159, 105242. <https://doi.org/10.1016/j.autcon.2024.105242>