FORECASTING MACHINE WORKING TIMES IN OFF-SITE CONSTRUCTION: A MACHINE LEARNING BENCHMARK FOR SUPPORTING PRODUCTION PLANNING

Andrea Sbaragli[1], Xiang Xie[2], John Hedley [2] and Mohamad Kassem[2]
[1]School of Architecture, Technology and Engineering, University of Brighton, Brighton, BN2 4AT, UK
[2]Department of Engineering, Newcastle University, Newcastle, NE1 7RU, UK

Abstract

Off-site construction relocates building processes to controlled manufacturing environments, enhancing productivity but increasing the risk of cost overruns due to the non-repetitive nature of module fabrication. This work presents a machine learning benchmark for a cyber-physical system that digitises manufacturing operations by capturing machine working times alongside design and production features. The benchmark spans tree-based and support vector regressors, with the best-performing model embedded in an industrial dashboard for data-driven scheduling. Validated in a UK steel frame manufacturer across eight months of production data, Random Forest achieved 10.84% median absolute percentage error on 420 frames, enhancing production planning and scheduling.

Introduction

Off-site manufacturing has increasingly been adopted within construction by drawing on industrial production and assembly systems methodologies. Approaches such as prefabrication and modular construction involve manufacturing building components in controlled factory environments before transporting them to the site for installation (Fan et al., 2024). Although this paradigm shift improves project planning, overall productivity, and delivery time, several threats may disrupt production cycles (Mehdipoor and Iordanova, 2021). Construction projects present a wide range of variability, arising from geographical conditions, budgetary constraints, intended building functions, and client personalisation. These drivers may lead to design–production misalignment and prevent manufacturing standardisation. Delays and inaccuracies in translating Building Information Models (BIM) into manufacturing processes can generate bottlenecks (e.g., reworks) that negatively impact profit margins and expose projects to penalties arising from missed deadlines (Jowett et al., 2024). These environments require an expanded digital footprint to enable continuous monitoring of industrial Key Performance Indicators (KPIs) and to support their control within desired efficiency levels (Saif et al., 2024). The fourth industrial revolution vastly digitised and automated manufacturing environments, exploiting several enabling technologies. While robotics and automation have progressively replaced human involvement in repetitive and hazardous tasks, workers continue to play a pivotal role in manufacturing value creation, performing operations where automation remains either technically unfeasible or economically unviable. Yet monitoring the human factor poses particularly significant challenges. Unlike machine processes, human behaviour is inherently variable, shaped by cognitive states, fatigue, skill levels, and environmental conditions, making it difficult to model and standardise. Compounding this, human-related data are often noisy, context-dependent, and subject to rapid fluctuations, further hindering reliable and continuous assessment (Sbaragli and Nardello, 2025). Several cyber-physical systems (CPS) have been developed to support decision-making in industrial processes by leveraging multiple architectural layers. Acquisition layers enable the automated collection of data from shop-floor resources, while cyber layers process these data streams and apply machine learning algorithms to compute Key Performance Indicators (KPIs) (Thiede et al., 2016; Alsakka et al., 2024). Before model deployment, various algorithmic structures and configurations are systematically benchmarked to identify the most performing solution for the specific use case (Sbaragli and Nardello, 2025; Albanese et al., 2022). These industrial metrics are subsequently integrated into dashboards that automate key tasks for manufacturing stakeholders and enhance the operational visibility of industrial processes (Thiede et al., 2016). Among the various industrial applications, CPS have been extensively investigated to support production planning and scheduling (Brunelli et al., 2019). In this context, product-related industrial and design features are used as inputs to machine learning models to predict manufacturing processing times (Çakıt and Dağdeviren, 2023). Despite providing a robust foundation for improving production planning and scheduling operations, these data-driven applications suffer from two main limitations. First, these methodologies offer proof-of-concept solutions, where acquisition layer designs are not discussed, affecting successful deployments in production environments (Çakıt and Dağdeviren, 2023; Masoumi and Bond, 2025). Second, the evaluation of cyber layers and related machine learning models is often constrained to a limited set of indicators, resulting in incomplete evaluations and discussion of their weaknesses (Masoumi and Bond, 2025; Çakıt and Dağdeviren, 2023). Based on these limitations, this work proposes a CPS to support production planning in human-centric off-site construction environments, with the core contribution being the systematic benchmarking of different machine learning regression models to accurately forecast semi-automated machine working times. The remainder of this work is organised as follows. A literature review of studies on forecasting resources' working time in production environments is presented. This is followed by the presentation of our developed CPS, its evaluation, and the discussion of findings. The last Section ends this work by outlining the conclusions and further research opportunities.

Literature Review

Production planning involves the systematic forecasting, coordination, and scheduling of activities, resources, and materials required to execute manufacturing processes efficiently (Göppert et al., 2024). Traditionally, operations research models, such as multi-objective programming and metaheuristic algorithms, have been extensively applied to support these decisions by iteratively identifying sub-optimal industrial configurations (Demartini et al., 2021). For instance, Auer et al. (2021) proposes a heuristic to minimise production throughput time under constraints such as resource capacities and set-up times. However, these approaches often rely on static parameters (e.g., standard cycle times), which limit their effectiveness in dynamic manufacturing environments. In this context, machine learning-powered CPS approaches are increasingly adopted to enhance the robustness and responsiveness of scheduling decisions. Recent advances demonstrate the effectiveness of supervised, unsupervised, and reinforcement learning techniques in forecasting key operational variables, such as machine utilisation and processing times (Göppert et al., 2024). By enabling data-driven and adaptive decision-making, machine learning methods effectively complement the limitations of traditional optimisation approaches. For instance, Çakıt and Dağdeviren (2023) designs a CPS benchmarking machine and deep learning models to predict cycle times of bus manufacturing operations. The trained dataset consists of several features, including the number of sub-assemblies, welding operations, and metal forming processes. The k-nearest neighbor emerges as the most performing classifier with a Root Mean Squared Error (RMSE) equal to 1.48, granting superior visibility into manufacturing operations before scheduling jobs into the shop floor. Despite these promising results, this methodology has some notable limitations. The acquisition layer does not discuss how the dataset features are automatically gathered and thus fed into computational models. Furthermore, some dataset features, such as environmental conditions (products' complexity and ergonomics patterns), are highly subjective as they depend on evaluators' perception as well as workers' experience and physiology. The algorithms' validation does not discuss percentage errors between actual and predicted working times, and falls short in validating the methodology with a production sequence. This analysis is pivotal to discussing how potential under- or overestimation may affect production environments and the related stakeholders' activity. Similarly, Masoumi and Bond (2025) validates an approach to optimizing and estimating the lead time of wooden furniture. The dataset maps chair products as a function of design (e.g., leg shapes) and industrial (e.g., machine configurations) features. Neural networks and tree-based regressors validate the CPS performances, showing a Mean Absolute Percentage Error (MAPE) equal to 1.63%. Despite the excellent performance in forecasting production times, this approach suffers from two main limitations. First, the authors do not discuss the design of the acquisition layer, limiting the deployment into manufacturing environments. Second, the investigation targets fully automated environments, neglecting the pivotal role that workers play in production environments. In particular, the intrinsic entropy of the human factor during task execution may challenge the learning capabilities and thus the classification and regression performances of cyber layers (Sbaragli et al., 2025). Besides traditional manufacturing industries, production planning digital systems have been leveraged in off-site construction environments to better schedule building phases and collaborate with project stakeholders (e.g., site installation). Aghajamali et al. (2022) proposes a CPS to estimate fitting and welding stations' working times in a Canadian construction specialist business. Dataset features relate to each resource involved in the productive value-creation. While fitting station targets the amount of bolted and cut parts, the welding process relies on the product's linear length. These datastreams are fed into a linear regressor and a fully connected neural network to estimate industrial resources cycle times. Despite the neural network's superior performance, the discussion solely evaluates the Mean Squared Error (MSE), falling short in outlining percentage errors between actual and predicted working times. Moreover, authors do not investigate the aggregate errors in a test production sequence in order to identify how potential overestimation may affect the sequential operations of other stakeholders. Data-driven approaches have also been developed to predict construction wall frames production cycle times (Alsakka et al., 2024). In particular, an acquisition layer digitizes the shop floor by leveraging MES and BIM datastreams and Internet of Things sensors such as cameras. This information is fed into neural networks and simulation models to predict working times in real-time. This CPS is validated with a run sequence of 25 assembly frames, showing an underestimation of 22 minutes compared to the actual cycle time. This approach massively outperforms a baseline model that relies on frames' linear meters. However, the sensors' cost, along with their potential occlusions, may limit the applicability of this human-centric solution in other off-site use cases. In summary, the reviewed CPS studies provide valuable methodologies and actionable metrics to support production planning. However, their applicability to off-site construction environments can be further enhanced from two key perspectives.

- **Acquisition layers design and ownership cost:** (Çakıt and Dağdeviren, 2023; Masoumi and Bond, 2025) do not address the design of acquisition layers for automated production data collection. Such functionality is essential for deploying production-ready CPS, as it enables the continuous integration of shop-floor data into algorithms to derive actionable metrics. Furthermore, acquisition systems should be cost-effective to facilitate their adoption in real-world operating environments (Alsakka et al., 2024).
- **Performance benchmarking & evaluation:** the benchmark of machine learning regressors is often based on a limited set of metrics, failing to highlight strengths and limitations in predicting work times. Also, these data-driven approaches are required to investigate how models' predictions affect manufacturing run sequence, potentially affecting other businesses' operations (e.g., site installation) (Masoumi and Bond, 2025; Çakıt and Dağdeviren, 2023).

Building on this discussion, this paper proposes a machine learning-powered CPS to address the identified limitations. The system comprises three interconnected components. First, an acquisition layer consisting of two cloud-based Extract, Transform, Load (ETL) pipelines automates the collection of real-time production data and product design workflows through a cost-effective approach. Second, a cyber layer pre-processes the collected data streams via feature extraction techniques and systematically benchmarks five machine learning regression models to forecast machine working times, including operators' manual material handling activities. Third, a web-based dashboard streamlines production scheduling by automatically dispatching selected frames to stand-alone machines and notifying manufacturing decision-makers via email with predicted working times, while providing key metrics to monitor operations and support integration with other business functions. The CPS is validated in a job shop of a UK-based off-site construction company that manufactures 16 steel frame types. The dataset contains more than 2000 frames manufactured during the first eight months of 2025. Among the benchmarked regressors, the Random Forest achieves the best performance, with a Median Absolute Percentage Error of 10.84% and a percentage residual error of 0.02% on a production order of 420 frames.

Cyber physical system

The digital system is developed to support decision-makers in monitoring and planning fast-paced human-centric off-site production environments. To meet these complementary goals, the CPS consists of four conceptual entities (Fig.1). The first layer targets any human-centric construction environment in which workers execute production cycles by interacting with a diverse set of resources, including machines and assembly stations. An acquisition layer digitizes process execution in these fast-paced, non-standardized settings, replacing reliance on plant managers' experience and lean methodologies. While products' design features and machine working times are acquired using cloud-based ETL pipelines, the Scanning App is an in-house built MES that records product and worker-dependent manufacturing events. These data streams, after specific pre-processing computations, are fed into the third system's entity, the cyber-physical layer. In this regard, machine learning regressors and time-driven algorithms forecast products' working times and derive industrial KPIs. These metrics flow in a dashboard, divided into four different areas. While the first three monitors manufacturing outputs and efficiencies of resources leveraging data coming from the MES (see Fig.1), the planning area presents an interactive user interface to support manufacturing stakeholders in scheduling production run sequences on the shop floor. This preliminary work investigates the feasibility of supporting production planning through the integration of machine learning algorithms within a cyber-physical system. The core contribution lies in benchmarking different regression models for predicting machine working times and in presenting the functionalities of the dashboard's planning area to support run sequence scheduling. The remaining CPS components, such as the MES, are beyond the scope of this paper and are therefore not discussed in further detail.

Acquisition Layer

The acquisition layer runs two parallel ETL pipelines to gather design data streams and machine working time from the shop floor (see Fig. 2). Starting with the first, the structural team publishes the modeled frames on Asite, where any stakeholder may track the product workflow. Any product is uniquely identified by an assembly reference and two supplementary files. While the PDF shows the assembly drawing useful to assist the shop floor task executions, the spreadsheet lists the coil parameters, stud length, and their machine-specific operations (e.g., dimples, notches, swages, etc.) along with a 3D frame position. Each frame also carries a design status, such as On Hold or Approved IFC, along with a revision identifier. A Databricks recursive pipeline (see green lines in Fig.2) continuously accesses all frames regardless of their design status, storing stud identifiers and required machine operations into a cloud database table. As a result, when a frame transitions from On Hold to Approved IFC, the corresponding database records are updated rather than newly created. Therefore, each frame and revision generates multiple entries, one per manufactured stud, with the assembly reference and stud identifier forming the composite primary key. The pipeline runs three times per day to ensure that any revision or status change is promptly reflected in the database. Only the latest revision of frames with an Approved IFC design status can be scheduled for production. Once the production is scheduled and running, the stand-alone machine processes frame-specific spreadsheets and records studs working times in TXT logs. The machine is network-connected and exposed to a static IP address that facilitates the data retrieval of the second pipeline (see blue

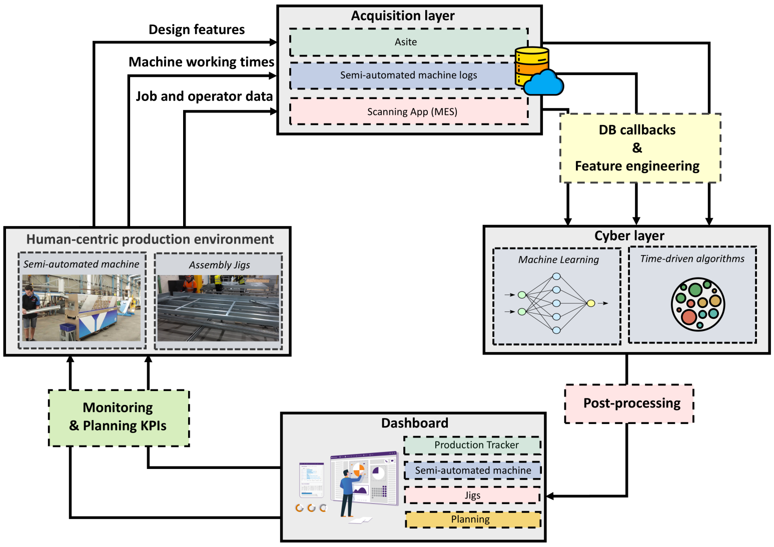


Figure 1: Cyber-physical system to monitor and plan human-centric offsite construction environments

lines in Fig.2). Production data are stored in a separate table but share the same composite primary key of the design flow. The machine pipeline time resolution is equal to 5 minutes, enabling future close-to-real-time shop floor operations monitoring.

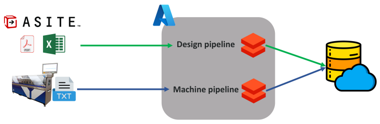


Figure 2: Acquisition Layer data flow

Cyber Layer - Machine Learning

Acquired datastreams are retrieved using database callbacks and subsequently joined on the shared composite primary key. Multiple machine learning models, commonly applied in industrial applications, are then trained on the aggregated dataset to forecast steel frames production times based on design features (Sbaragli et al., 2025). The dataset was partitioned into training and testing subsets using an 80%-20% split, and model performance was evaluated using 5-fold cross-validation. Models' hyperparameters are sub-optimized through the grid search algorithm. The following bullet point lists the evaluated models:

- **Baseline:** this model assumes a direct and linear relationship between frames' length and machine work time, as investigated by Alsakka et al. (2024). The training defines a coefficient as $minutes/millimeter$ that is used to infer process times in the test set.
- **Linear Regressor (LR):** a parametric model that assumes a linear mapping between predictors and the target variable. As LR requires no hyperparameter tuning, it represents a computationally efficient baseline against which to compare more flexible models (Aghajamali et al., 2022).
- **Support Random Vector (SVR):** this kernel-based method fits a regression function within an error tolerance. Its predictive behavior is governed by several parameters. While the regularization parameter C controls the trade-off between model complexity and tolerance to errors, the kernel coefficient γ determines the influence of individual training samples. Finally, the insensitive loss margin ϵ defines the tolerance zone around predictions where no penalty is applied (Çakıt and Dağdeviren, 2023).
- **Random Forest (RF):** an ensemble approach that develops multiple decision trees on bootstrapped samples and aggregates their predictions to reduce variance. Performance depends on parameters such as the number of trees, node-splitting thresholds, and feature selection strategy (Sbaragli et al., 2025).
- **Light Gradient Boosting Machine (Light GMB):** this boosting-based ensemble regressor builds trees sequentially, where each iteration corrects the residuals of the previous one. Key parameters include the number of boosting rounds, learning rate, boosting type (e.g., GBDT or DART), and the learning objective set for regression (Masoumi and Bond, 2025).

After having trained and tested these classifiers, the most performing model is used to support manufacturing engineers in production planning and scheduling. While the comparison benchmark of models is discussed in the Results Section, the following subsection outlines the benefit of leveraging this data-driven CPS in off-site manufacturing environments.

Dashboard - Planning Area

Building on the discussed cyber-layer capabilities for predicting machine work times, this final subsection discusses the dashboard's planning area. The frontend presents an interactive table that reflects the design workflow captured by the ETL pipeline. As shown in Fig. 3, the table maps assembly references (e.g., ARef) against key attributes including project, coil configuration, frame length (e.g., Len), and the number of screws to be fitted in the jig area. Since an assembly reference may go through multiple design revisions, the table displays only the latest revision for each, ensuring that engineers always work with the most up-to-date information. Each assembly entry also displays its current design status and scheduled production day. Manufacturing engineers can schedule a frame to the shop floor only when two conditions are met: i) the structural manager has set the design status to Approved IFC in Asite, and ii) the frame has not yet been assigned to a production date (i.e., the field is null in Fig.3). To support efficient navigation of large datasets, a filter button allows users to narrow the table view by attributes such as project, design status, and coil configuration. Once the desired eligible items have been selected for scheduling, the user clicks the Schedule Selected button, which triggers two back-end operations. First, the machine files for the selected frames are transferred via the File Transfer Protocol to a designated folder on the stand-alone machine, which the shop floor operator can then open and automatically load. Second, a spreadsheet summarising the expected production times and linear coil consumption for each configuration is automatically emailed to the relevant production stakeholders, such as manufacturing engineers and the production director. By automating both file transfer and the manual development of run sequences, the dashboard significantly reduces the non-value-added administrative burden on engineers, ultimately driving meaningful productivity enhancements across the production workflow.

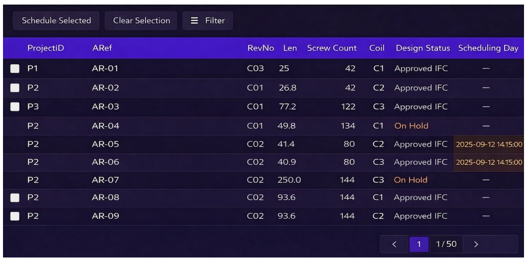


Figure 3: Interactive web-based application to support production scheduling

This novel data-driven methodology provides the business a key opportunity to strengthen its operational value creation. Indeed, the manufacturing team has greater operations visibility that enables better coordination and collaboration with other business areas:

- **Purchasing & Accounting:** the known a priori coil configurations consumption facilitates a just-in-time approach that minimizes stock occupancy and avoids cost overruns in low-rotating or project-specific material. At the same time, project cash flows, such as labor and material costs, are seamlessly integrated.
- **Site:** forecasted due dates of run sheets enable a data-driven definition of frames shipping windows, avoiding project delays and related penalties for missed deadlines.

Results

This section validates the CPS by feeding structural frames' design features into machine learning models to forecast semi-automated resource working times, as captured by the ETL pipeline. Notably, these times encompass material handling activities, which may challenge regressors' learning capabilities. Before discussing the performance of the benchmarked regressors, the following subsection discusses 1) the off-site use case manufacturing operations and 2) the obtained dataset by collecting 8 months of steel frame structural production.

Off-site use case: structural steel frame production

The discussed CPS is tested and validated in a UK-based off-site construction specialist. The monitored human-centric job shop manufactures several steel frame products involving different resources. Fig.4 outlines the off-site manufacturing operations, starting from the steel coil, which is distinguished by different configurations of depth and width (e.g., C1, C2, etc.). Once structural frames are modeled and set to Approved IFC, they are released into the shop floor by using the discussed interactive web-based dashboard. The manufacturing process is then carried out by a semi-automated machine, in which a skilled operator loads the machine file before execution and moves sequentially processed studs into work-in-progress (WIP) storage units. A jolly worker may join the process to facilitate the material handling of studs up to 6.4 meters long. Based on the final product's complexity, a structural steel frame may consist of up to 74 studs. Each processed stud is distinguished by a set of design features (e.g., dimples, swages, notches, length, weight, etc.) that affect the machine's working cycle times. However, ETL-captured working time logs, whose resolution and structure are dictated by the machine output rather than experimental design, also include material handling events. This may challenge machine learning regressors due to the intrinsic process variability (e.g., differences in stud size) and variability in human commitment and experience (Sbaragli and Nardello, 2025). The construction job shop follows a one-piece flow approach in normal operating conditions, and thus, WIP units are moved next to the available or less saturated assembly jig. Four jigs assemble steel frame products in parallel, leveraging up to 4 skilled operators each. Based on the designed assembly drawings, manual operations require stud fittings and screwing operations, as well as potential reworks to address some modelling and/or machine errors. Completed frames are classified into 16 products and thus follow separate production flows (see Fig.4). While Product IDs 1 and 2 are sent to the construction site for their installation, the remaining ones represent WIP modules for other in-plant job shops.

Based on this discussion on manufacturing operations, the following subsection details the acquired datasets from the semi-automated machine.

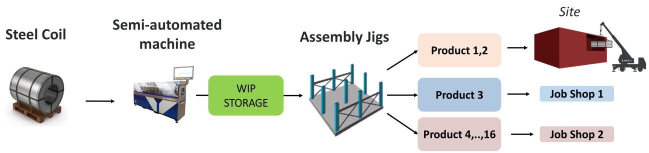


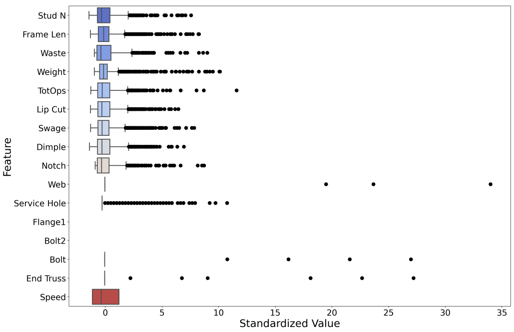
Figure 4: Human-centric production environment

Dataset

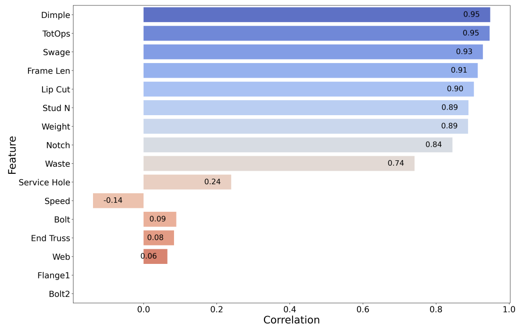
To validate the proposed methodology in forecasting machine processing times, 8 production months in 2025 (e.g., from January to August) are retrieved from the resource-specific database involving 5 construction projects across the United Kingdom. The stud-based dataset comprises 26500 samples, with individual work times ranging from 2 to 71.8 seconds. The average process time and standard deviation are equal to 20.4 and 14.2 seconds, respectively. This low-level dataset is grouped by assembly reference to achieve a frame-based representation, where one ARef may group 6 to 74 studs based on the final product complexity and application. This aggregation is implemented for two main aspects. First, the studs' working times are highly exposed to the intrinsic uncertainty of the human factor during task executions. The dedicated operator often pauses the machine to perform audit checks and move manufactured items into WIP storage units. Despite being cyclic, these events have high degrees of entropy, affecting the studs' working times and limiting the learning capabilities of any machine learning model. Second, frames might be assembled only if all studs are processed and stored in a specific location (see Fig. 4). The new compressed dataset includes 2099 steel frames having working times ranging from 0.2 to 26.4 minutes. The average process time and standard deviation are equal to 4.3 and 2.9 minutes, respectively. Once the machine dataset is compressed, a separate callback downloads and post-processes data from the design database table. Similarly to the machine dataset, frame features are obtained by summing or counting the studs descriptors while grouping on the ARef field. These frames' features include the total amount of required manufacturing operations (e.g., Lip Cut, Notch, Web, Service Hole, etc.), the aggregate dimensions (e.g., length and weight), and the count of grouped studs. Finally, the two datasets are joined together on ARef, the common and unique column. Fig. 5 (a) outlines the entire set of frame features, where waste and speed represent steel scrap and machine velocity in manufacturing a given product. It is worth noting that Stud N and TotOps represent the total amount of studs and machine operations on every frame, respectively. The remaining structural features are not detailed further due to the company non non-disclosure agreement. The job shop's production output is very customized to meet the customer and construction requirements of different healthcare projects. Fig.5 (a) streamlines features relative variability and distribution patterns through z-scores. The 16 different variables have standardized values ranging from 0 to approximately 35 and might be divided into different sub-groups. Structural information (e.g., Lip Cut, Swage, Dimple, Notch, etc.), dimensions (e.g., Weight and Frame Len), and Waste present compact distributions with upper quartiles below the standardized unit 5. However, many outlier reiterates the uniqueness of structural modules within projects. Less-exploited design specifications, such as Web, Bolt, and Service Hole, present very narrow distribution close to zero but with several distinct outliers at higher values. Flange1 and Bolt2 are not represented in this dataset. Finally, the Speed has no whiskers, underlying machine velocities equal to 20%, 50%, and 70% of the maximum. This set of features is ranked to reduce computational load and simplify the learning process for machine learning models. The data distributions are non-Gaussian, may lack linear relationships with machine work time, and include many outliers. Therefore, the Spearman correlation is applied to evaluate the strength and direction of monotonic associations. Fig. 5 (b) ranks the pair-wise correlation between each feature and the machine work time. It is worth noting that the Speed correlation captures a weak inverse relationship with the processing time. Feature selection employed a correlation threshold of 0.7 to minimize dimensionality for machine learning models. Therefore, the subsequent performance benchmark discussion excludes all descriptors with correlations below this cutoff, from Service Hole to Bolt2.

Working times prediction: algorithms' performance benchmarking

The structural dataset has been validated using the five different regressors discussed in the cyber layer subsection. The training process involves five-fold cross-validation and grid-search hyperparameter optimisation. Regressors are benchmarked using Training Time (TrTime), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2), which provide standard error and goodness-of-fit measures, alongside Median Absolute Error (MdAE) and Median Absolute Percentage Error (MdAPE) as robust, scale-independent alternatives less sensitive to outliers. Tab.1 benchmarks the performance of the algorithms to forecast the semi-automated machine work times by having as input the features discussed in the previous section. The baseline approach that leverages frame length delivers the weakest performance. The high error levels (e.g., RMSE = 1.19, MAE = 0.78, MAPE = 19.09%) and a relatively low coefficient of determination (e.g., R^2 = 0.84) reflect the model's inability to capture the underlying data patterns and relationships. LR substantially improves these metrics, achieving the highest R^2 and competitive error values. RMSE and MAE are equal to 0.72 and 0.51, respectively. However, LR's percentage-based metrics (e.g., MAPE = 14.51%, MdAPE = 11.36%) remain higher compared to tree-based methods, suggesting reduced robustness when predicted working times are



(a) Z-score features distribution

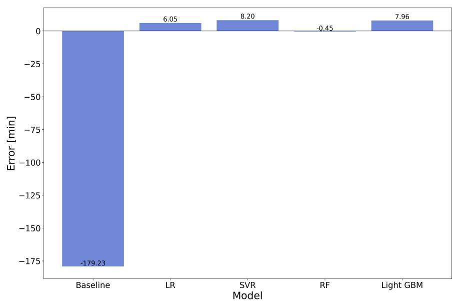


(b) Spearman Correlation

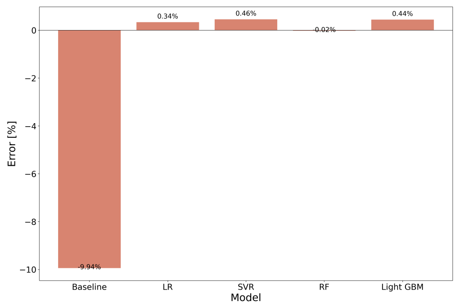
Figure 5: Dataset features distribution and ranking

Table 1: Performance comparison of regression models.

Regressor	TrTime [sec]	RMSE	MAE	MAPE	MdAE	MdAPE	R^2
Baseline	NA	1.19	0.78	19.09%	0.50	17.81%	0.84
LR	NA	0.72	0.51	14.51%	0.37	11.36%	0.94
SVR	8.2	0.88	0.55	17.26%	0.39	12.39%	0.92
RF	**44.0**	**0.85**	**0.52**	**13.79%**	**0.33**	**10.84%**	**0.92**
Light GBM	14.8	0.85	0.55	16.02%	0.37	11.92%	0.92



(a) Absolute Residuals



(b) Relative Residuals

Figure 6: Aggregate model performances on the test set

assessed relative to actual values. SVR performs moderately well, even though the RMSE of 0.88 and MAPE equal to 17.26% indicate less stability compared to LR and RF models. Light GBM achieves comparable performance to SVR (e.g., RMSE = 0.85 and MAE = 0.55), but MAPE and MdAPE equal to 16.02% and 11.92% show limitations in effective generalization. Within this benchmark, the RF regressor demonstrates the most balanced and robust performance metrics. Despite a training time equal to 44 seconds and slightly higher exposure to outliers compared to LR (e.g., RMSE = 0.85), RF yields the lowest MAPE (13.79%) and MdAPE (10.84%), indicating that its predictions are not only accurate but also less sensitive to variations in scale. The performance benchmark is further expanded by investigating how models' performances affect the run sequence of the test set (e.g., 420 frames). In this regard, Fig. 6 depicts the aggregate absolute and relative residuals after having summed the actual machine working times and the ones predicted by each algorithm. RF is the only model to achieve an aggregate absolute residual close to zero (0.45 minutes), whereas LR, SVR, and LightGBM display absolute and relative deviations ranging from 6.05 to 8.20 minutes and from 0.34% to 0.46%, respectively. The baseline remains the worst regressors, accounting for an underestimation of 179.23 minutes and 9.94%.

Operational implications

These results confirm that the RF regressor provides the most reliable and operationally robust performance among the tested models (Tab. 1 and Fig. 6). Crucially, the dashboard workflow automates production scheduling by generating optimised run sequences directly from the web application, eliminating the need for engineers to interact with multiple company software systems (e.g., Asite), manually download files, or transfer machine files to the stand-alone resource. Preliminary investigations on run sequences ranging from 20 to 30 structural frames demonstrate a 75% reduction in planning time. At the same time, machine learning-based predictions enhances visibility into manufacturing operations, allowing stakeholders to better anticipate workloads and coordinate downstream processes. By integrating this capability across departments, the system supports tighter alignment across key functions. For planning, it enables more realistic deadline scheduling against production constraints. For purchasing, it facilitates just-in-time material deliveries and reduces stock occupancy. For finance, it provides advance visibility into material consumption, supporting improved cost forecasting and budget control.

Conclusions & Further Research

Off-site construction environments tend to suffer from cost overruns and constantly delayed deadlines due to poor operations management and a lack of standardization across projects. Therefore, construction factories are required to leverage Industry 4.0 enabling technologies to support production planning and achieve a more reliable operations visibility. This work proposes a machine-learning-powered CPS that supports manufacturing stakeholders in forecasting machine cycle times and automating run sequences on the shop floor. While two ETL pipelines acquire design features and machine work times, the cyber layer benchmarks 5 different algorithms, establishing the RF as the most performing model to forecast frame production times. This tree-based regressor is embedded in an interactive web-based dashboard, enabling manufacturing stakeholders to forecast production times before scheduling run sequences into the shop floor. This CPS has been validated in a UK-based steel frame job shop production using 8 months of data in 2025. Further research should validate this methodology in the jigs area and use metaheuristic algorithms to optimise production scheduling given a set of constraints.

Acknowledgments

This work has been supported by Innovate UK under the Knowledge Partnership Program with grant number 10099000.

References

Aghajamali, K., Suliman, A., Alattas, A., and Lei, Z. (2022). Data-driven cycle time prediction of fitting and welding stations in steel fabrication. Modular and offsite construction (MOC) summit proceedings, pages 122–129.

Albanese, A., Nardello, M., Fiacco, G., and Brunelli, D. (2022). Tiny machine learning for high accuracy product quality inspection. IEEE Sensors Journal, 23(2):1575–1583.

Alsakka, F., Yu, H., El-Chami, I., Hamzeh, F., and Al-Hussein, M. (2024). Digital twin for production estimation, scheduling and real-time monitoring in off-site construction. Computers & Industrial Engineering, 191:110173.

Auer, P., Dósa, G., Dulai, T., Fügenschuh, A., Näser, P., Ortner, R., and Werner-Stark, Á. (2021). A new heuristic and an exact approach for a production planning problem. Central European journal of operations research, 29(3):1079–1113.

Brunelli, L., Masiero, C., Tosato, D., Beghi, A., and Susto, G. A. (2019). Deep learning-based production forecasting in manufacturing: a packaging equipment case study. Procedia Manufacturing, 38:248–255.

Çakıt, E. and Dağdeviren, M. (2023). Comparative analysis of machine learning algorithms for predicting standard time in a manufacturing environment. AI EDAM, 37:e2.

Demartini, M., Tonelli, F., Pacella, M., Papadia, G., et al. (2021). A review of production planning models: Emerging features and limitations compared to practical implementation. Procedia CIRP, 104:588–593.

Fan, J., Chen, L., and Chen, K. (2024). Digitalizing industrialized construction projects: Status quo and future development. Applied Sciences, 14(13):5456.

Göppert, A., Kaven, L., Baum, J., Melnychuk, O., and Schmitt, R. (2024). Machine learning for online scheduling in manufacturing: A systematic literature review. Procedia CIRP, 130:154–160.

Jowett, B., Edwards, D. J., and Kassem, M. (2024). Field bim and mobile bim technologies: a requirements taxonomy and its interactions with construction management functions. Construction Innovation, 24(1):134–163.

Masoumi, A. and Bond, B. H. (2025). Machine learning-based prediction of processing time in furniture manufacturing to estimate lead time and pricing. European Journal of Wood and Wood Products, 83(1):35.

Mehdipoor, A. and Iordanova, I. (2021). Systematic literature review on the combination of digital fabrication, bim and off-site manufacturing in construction—a research road map. In Canadian Society of Civil Engineering Annual Conference, pages 283–296. Springer.

Saif, W., RazaviAlavi, S., and Kassem, M. (2024). Construction digital twin: a taxonomy and analysis of the application-technology-data triad. Automation in Construction, 167:105715.

Sbaragli, A., Ghafoorpoor, P. Y., Thiede, S., and Pilati, F. (2025). A cyber-physical architecture to monitor human-centric reconfigurable manufacturing systems. Journal of intelligent manufacturing, pages 1–23.

Sbaragli, A. and Nardello, M. (2025). From sensing to segmentation: Transformer-based worker activity recognition for industrial assembly. IEEE Access, 13:207475–207487.

Thiede, S., Juraschek, M., and Herrmann, C. (2016). Implementing cyber-physical production systems in learning factories. Procedia Cirp, 54:7–12.