Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

REAL-TIME ACTIVITY DURATION EXTRACTION OF CRANE WORKS FOR DATA-DRIVEN DISCRETE EVENT SIMULATION

Manuel Jungmann
Lucian Ungureanu
Timo Hartmann

Civil Systems Engineering
Technische Universität Berlin
Gustav-Meyer-Allee 25
Berlin, 13355, GERMANY

Hector Posada
Rolando Chacon

Civil and Environmental Engineering
Universitat Politècnica de Catalunya
Carrer Jordi Girona 1-3
Barcelona, 08034, SPAIN

ABSTRACT

The construction industry is struggling with low productivity rates because of a low level of digitalization, dynamic interactions, and uncontrollable circumstances on sites, which make the planning process complex. Usage of the digital twin construction paradigm enables to facilitate construction management and leverage the sector's unexploited potential. This research addresses current shortcomings by real-time discrete event simulation. During crane operations, kinematic data were collected, which were classified by machine learning algorithms for activity recognition and duration extraction. Based on the identified durations, Goodness-of-Fit techniques determined suitable probability density functions. The resulting probability density functions were used as input parameters in stochastic discrete event simulations. It was shown that with enriched data collection, probability density functions have to be updated. The data-driven discrete event simulation facilitates decision-making processes by providing more reliable real-time information for the planning of upcoming construction works. Thus, data-based instead of experience-based management can be enabled.

1 INTRODUCTION

The construction field is one of the least digitalized sectors with flat or even falling productivity rates, although it is one of the biggest industries (Patteri and Stolton 2019). Heavy equipment especially suffers from low productivity (Slaton et al. 2020). Hence, analysis of construction equipment operations and determining activity durations is expedient for increasing overall productivity. Planning and scheduling of activities and availability of materials are the most influential attributes for construction labor productivity (Dixit and Sharma 2020). In general, a construction project consists of interconnected activities aimed at constructing a building within a specific time period. Thus far, it is common practice to use rigid methods like the Critical Path Method, which are unsuitable for daily construction management (Seppänen et al. 2014). Construction works are risk-sensitive and planning of sequences is complex as works are executed outdoors under uncontrollable conditions. Deviations from initial construction planning occur regularly and schedules have to be adjusted accordingly (Vahdatikhaki and Hammad 2014).

Digital technologies can simplify ongoing construction site management. Discrete event simulation (DES) is a method capable of modeling dynamic interactions with stochastics. The effects of different management decisions can be analyzed in a virtual environment. Probability density functions (PDFs) modeling is one of the most influential issues in these simulation studies, as the simulation model's reliability depends on input parameters (Song and Eldin 2012). Hitherto, static experience-based input parameters dating from the pre-construction phase are used and represent the major barrier for widespread usage of simulation in the architecture, engineering, and construction (AEC) field (Behzadan et al. 2015; Abbasi et al. 2020). However, advances in technology enable the collection of real-time performance data during construction execution and facilitate automatic activity recognition. Activity durations can be extracted for data-driven stochastic DES. Research for real-time data collection and its automated linkage to DES is required (Alvanchi et al. 2021). This can enable superior construction site management and decision-making in (near) real-time (Sherafat et al. 2020).

To provide the demonstration of DES's potential for dynamic planning of ongoing construction processes, collected kinematic real-time data of crane operations were analyzed for determining operation durations as input parameters. Machine learning (ML) algorithms were used for data classification to recognize operations automatically and extract operation durations. Goodness-of-Fit methods were applied for finding suitable PDFs as input parameters for stochastic DES. By collecting real-time data continuously, the PDFs need to be updated. Thus, the decision-making within construction planning and scheduling can be improved by real-time DES. The paper is structured as follows: Chapter two outlines the literature background. Afterwards, the framework for collecting real-time performance data and their data mining for stochastic DES is introduced. The applicability of the approach is demonstrated by collecting kinematic data during crane operations for concrete works on a construction site in Barcelona, Spain. Finally, the results are discussed and future research directions are summarized.

2 LITERATURE BACKGROUND

2.1 Real-Time Data and Digital Twins

Within the simulation field, the collection of data is the most important aspect and one of the biggest challenges for solving real-world situations (Banks et al. 2010). On construction sites, conventional manual data collection strategies are still prevalent, although these methods are error-prone, labor-intensive, and time-consuming (Xue et al. 2021). Thus far, gathered data are handled too late and information about present situations on construction sites reaches project managers delayed. This impedes the quality of construction planning and execution (Cheng and Teizer 2013). With advancements in digitalization and automation, such as the Internet of Things (IoT), it is possible to collect data more simply and get access in real-time (Boje et al. 2020). Nowadays, there are several technologies available for real-time data collection, such as location or movement tracking, computer vision or audio.

Real-time data collection of construction equipment's movements enables to build a digital twin (DT) during the construction phase. There is no commonly agreed definition for the DT concept in the AEC industry yet. According to Brilakis et al. (2019), it is essential to state the intended purpose before the creation of a DT. The most existing definition approaches identify three integral parts for a DT: a physical part (a construction site or building), a digital part (the digital representation of the physical part), and bidirectional information exchange. Data are collected at the physical part to gain project status information and knowledge of the current situation. Thus, the digital part can be updated and processed accordingly. Inversely, the digital part enables improved management of the physical part by investigating management decisions in a virtual environment due to simulation. IoT technology facilitates continuous data and information exchange. On the one hand, real-time data collection eases construction sequence control by comparing as-performed process and as-built product information with as-planned process and as-designed product status. On the other hand, real-time data and gained knowledge enable improved project management of future construction works in a timely manner. The method for using a DT as a means for data-based holistic management of ongoing works is defined as digital twin construction (DTC). The DT concept is still in its early development stage in the AEC industry and efforts regarding real-time data collection for creating a DT are rarely applied, especially within the construction phase (Sacks et al. 2020), although research has shown the effectiveness of using real-time data for decision support (Makarov et al. 2021).

2.2 Machine Learning Classification for Activity Recognition

In the construction sector, an activity consists of several operations aimed at completing a physical component or at performing support services with resources (Halpin and Riggs 1992). As performance monitoring by location tracking of resources' operations helps to improve productivity on construction sites, the advent of automated activity recognition methods can be detected in recent years (Sherafat et al. 2020). For automated activity recognition, ML classifiers are used for analyzing collected data to categorize data points into different classes. ML algorithms can be grouped into supervised and unsupervised. Within supervised ML, data are labeled and algorithms look for patterns according to the labeling. For unsupervised ML, data are not labeled in advance and algorithms look automatically for patterns to distinguish between classes. Supervised algorithms lead to improved performance in equipment activity recognition (Golparvar-Fard et al. 2013). Nowadays, deep learning algorithms are applied to reach higher accuracies in different domains where feature extraction is complex. But for the application of deep learning, much larger and more comprehensive data sets (Langroodi et al. 2021) and high-performance computers are required (Li et al. 2018), which are rarely available on construction sites. Additionally, ML algorithms work well with structured data, such as tabulated data collected by movement sensors, and even outperformed deep learning algorithms in activity recognition studies (Baldominos et al. 2019; Lee et al. 2020).

In general, three possibilities for activity recognition of construction equipment can be distinguished based on different data collection technologies. The kinematic method works by collecting movement data with sensors (e.g. accelerometers, gyroscopes, etc.) mounted on construction equipment. The visual approach (e.g. 2D/3D cameras) records construction processes as images or videos. Audio-based methods (e.g. microphones) depend on the sound patterns of equipment during construction execution. In particular, the latter two methods are sensitive to external factors such as adjacent construction equipment or weather conditions, which can influence the data collection process and its quality.

The potential of activity recognition in the AEC industry is not fully exploited yet (Akhavian and Behzadan 2015). It is necessary to use automated activity recognition approaches to gain incremental knowledge about the construction sequence and to use this information for real-time decision-making (Sherafat et al. 2020).Construction worker activity recognition has been explored, but there are still challenges for construction equipment activity recognition (Slaton et al. 2020). Although cranes play a crucial role in construction execution regarding safety, costs, and durations by moving components on site, in a current state-of-the-art review (Sherafat et al. 2021) and in an overview (Langroodi et al. 2021) for activity recognition methods of construction resources, none of the listed past studies focused on cranes. Existing research focuses on crane handling for safety and maintenance, but not on crane usage and how to support construction management (Nakanishi et al. 2022).

2.3 Discrete Event Simulation

DES consists of interconnected activities as a sequence of events. The state of the system changes according to the occurrence of events. DES can be used to mimic construction operation logic and resource allocations through statistical analysis (Liu et al. 2015). Durations of activities can be provided as stochastic values, i.e. PDFs, and variate input values are generated accordingly. Thus, DES is capable of modeling dynamic interactions, uncertainties, and risks. As construction site conditions are changing continuously, DES has to be modified and applied according to current situations on site. To address this issue, collecting real-time data during construction works and gaining continuous knowledge regarding activity durations for DES input is needed (Alvanchi et al. 2021).

2.4 Research Gap

In recent years, many studies regarding automated activity recognition of construction equipment have been conducted, but a focus on progress monitoring and usage of the gained information for decision-making has been neglected (Nakanishi et al. 2022). There is a shortage of studies, which determined activity durations based on real-time performance data and used this knowledge to plan ongoing construction works according to DTC. Song and Eldin (2012) tracked real-time global positioning system (GPS) data of delivery trucks to gather information about truck delivery times for DES. To distinguish between different operations, Akhavian and Behzadan (2013) applied a multimodal data gathering approach in laboratory experiments. Activity durations are stated by mean and standard deviation. Vahdatikhaki and Hammad (2014) proposed a framework for near real-time simulation by analyzing location-tracked data. The collected data are averaged over time to determine parameters for normal distributions. Akhavian and Behzadan (2015) investigated the performance of a front-end loader by mounting inertial measurement units (IMUs) onto it and derived activity durations for DES. Kim et al. (2019) used vision-based analysis during earthmoving operations to determine productivity rates for simulation input. Wu and AbouRizk (2021) tested the update of simulation inputs but included expert opinions to determine PDFs and were limited to three conventional distributions, such as the uniform one. These data-based approaches showed improved quality in comparison to conventional, static methods. But, none of the studies updated suitable PDFs according to an increasing amount of real-time data for simulation of construction equipment operations according to the DT concept. As always only sample data can be used for determining activity durations, it is expedient to collect movement data continuously to cover the dynamic and uncertain nature of construction activities and update the PDFs accordingly. Hence, by having more reliable activity durations as input parameters, the model becomes more meaningful. Therefore, this research addresses the intersection of real-time data and DTs, machine learning for activity recognition, and stochastic DES to apply DTC.

3 METHOD

The developed framework for updating activity durations by kinematic real-time data for DES input can be retraced in Figure 1. At first, real-time raw data have to be collected during construction execution by mounting sensors on equipment. The collected data are send to an IoT platform, which construction managers can access. The data have to be prepared for further analysis. Duplicated or close timestamps have to be removed and missing instances have to be interpolated to get a continuous data set. The data are normalized to create a common scale among different data sources. The prepared raw data are used to calculate features by overlapping sliding windows. Sliding windows summarize data by calculating features for a certain number of instances. The sliding windows have to be labelled based on observed information for supervised learning. In the following, a principal component analysis (PCA) is applied to the feature data set. PCA is a process of calculating eigenvectors by feature inputs to reduce the extent of the data set. The eigenvectors are based on features' data. The variance can be set, i.e. the cumulative information content of the original data set, by calculating the axle variance of the eigenvectors.

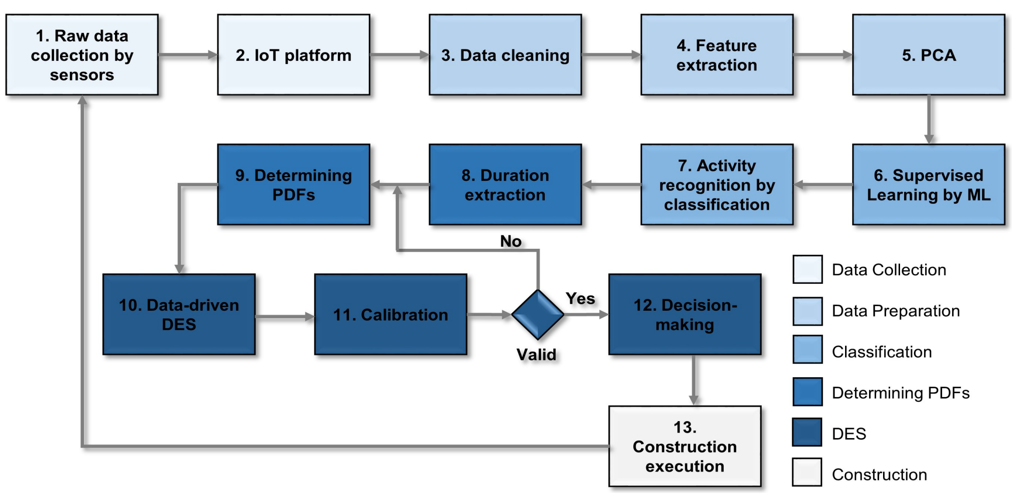

Figure 1: Framework for real-time DES.

The resulting data set by the PCA has to be used in a cross validation for supervised learning by different ML classifiers. Generally, there is no best ML classifier and it has to be tested which classifier performs best for the respective situation. The ML classifiers categorize the instances into different activities. The performance of different classifiers has to be compared and the classifier with the highest accuracy can be used for calculating the durations of each repetition of every operation. Each classified instance stands for one window. As the windows contain several data points from the original data set, the duration for each window is known according to the number of data points. The duration of each repetition of the operations can be calculated by the equation: *duration (seconds)* $= n * wl * ol$, with n = number of windows, wl = window length in seconds, and ol = percentage of overlap.

Heuristics are applied to exclude outliers such as singular windows, which do not represent reasonable classified data points. The different durations of each execution of every operation are used for applying Goodness-of-Fit techniques to determine suitable PDFs. Goodness-of-Fit is a statistical measure to test how closely the data can be represented. It can be distinguished between hypothesis tests and information criteria. Hypothesis tests calculate the difference between a data-based cumulative function and a possible PDF. Information criteria calculate the information loss of a possible PDF in comparison to input data. No PDF will exactly reproduce the input durations. It is advisable to apply several Goodness-of-Fit measures to a data set, as each Goodness-of-Fit approach has shortcomings and can be unreliable in some cases (Vincent 1998). If all tests prefer the same PDF, this PDF can be chosen as input. If there are different preferences among the tests, the results of each Goodness-of-Fit test are ranked and the PDF with the best overall ranking has to be chosen. When a PDF is determined by the Goodness-of-Fit methods, a null hypothesis test according to a certain significance level has to be applied to investigate whether the assumed PDF can be kept or whether the hypothesis has to be rejected. Therefore, a p value is calculated for each hypothesis test, which must lie over the in advanced determined significance level for not rejecting the null hypothesis.

Afterwards, the whole construction process has to be modeled within stochastic DES and the data-based PDFs resulting from Goodness-of-Fit analyses are used as activity duration inputs. Calibration has to be done by comparing DES's output with the real duration to validate the model. If the model is valid, the results of the DES enable decision-making regarding upcoming construction site activities. When the construction starts, further real-time data are collected and the whole procedure reruns. By enrichment of real-time data, the PDFs have to be updated.

4 APPLICATION

4.1 Demonstration Site

The applicability of the approach was demonstrated on a construction site in Barcelona, Spain, during shell construction by reinforced concrete for an office building. The walls and columns were poured by concrete buckets moved by a tower crane. The activity *concrete works* consisted of four iterative operations after the arrival and preparing of a concrete mixer truck: *Lifting a bucket down by crane, filling the bucket, lifting the bucket up by crane, and pouring concrete into the formwork* (Figure 2). If the truck was empty, it left and the next truck could take its position. The duration between leaving a concrete mixing plant and starting a curing process should not exceed a duration of 1.5 hours, as the quality of concrete decreases afterwards and it will become unusable (Lin et al. 2010). Therefore, an appropriate coordination of trucks' arrival times is needed. The trucks had a capacity of around 7,200 liters of concrete and the bucket of around 800 liters of concrete. Sensors were mounted on the crane hook during concrete works on the 7th floor. The data were collected in the afternoon of December 16^{th}, 2021, on the construction site during two ready-mixed concrete truck deliveries. In total, around 13,600 liters of concrete were poured. The weather was sunny with hardly any wind with a maximum of 2 m/s during measurements according to a closely located weather station. The Euclidean distance between the concrete truck and the working area was around 40 meters.

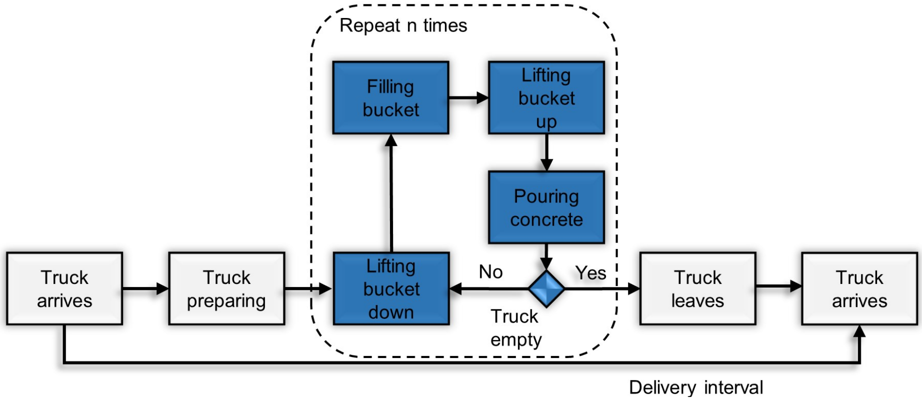

Figure 2: Work process pattern for concrete works.

4.2 Data Collection

For data collection, a three-axis IMU and location tracking sensors were mounted on a crane hook during concrete works, which could collect accelerometer, gyroscope, GPS, and height data. Thus far, a shortcoming was that data collection was rather done in an isolated manner by focusing on one technology source (Sacks et al. 2020; Kim et al. 2021). For this study, ten different raw data types were gathered and fused (X/Y/Z acceleration by IMU, X/Y/Z angular velocity by IMU, longitude/latitude/altitude by GPS tracker, and altitude by barometer). The device was connected to an ESP32 to gather sensor data and send it to an IoT platform. Each data point was saved with a related timestamp. As the second truck arrived delayed, around one hour of idle time occurred. This period was excluded in this study. A low sample rate was chosen for data collection in comparison to other equipment activity recognition studies with 50 to 100 Hz (Sherafat et al. 2020). If applying the DT concept in the future holistically by observing all resources on site and collecting data for each resource, an enormous amount of data will have to be handled. Thus, a focus on efficient usage of collected data is preferable. In human activity recognition, it has already been proven that a sampling rate of 1 Hz can save energy and still achieve high accuracy (Zheng et al. 2017). In total, 98:52 minutes of activity time were collected with 27,335 data points. The whole construction process was recorded on video for labeling data and validation.

4.3 Data Preparation

R 4.1.0 software was used for the study. The dataset was reduced to 4 Hz by deleting duplicated data points and interpolating missing ones. In the following, windows of a size of two seconds, including eight data points, were calculated with a 50 % overlap. The following time-domain features were extracted by statistical analysis of the raw data: minimum, maximum, mean, interquartile range (IQR), variance, and root-mean-square error. This resulted in 60 features in total. The windows were labelled manually by comparing the recorded video with the timestamps of the data points. Considering the PCA, a variance of 90 % was chosen, which resulted in a reduction down to 15 dimensions.

4.4 Classification

The performance of the following classifiers was compared according to a 10-fold cross validation on the whole data set for activity recognition: Naïve Bayes, Decision Tree, k-nearest neighbor (KNN), support vector machine (SVM), and Random Forest. The first four approaches are different basic ML techniques. Naïve Bayes is a probabilistic classifier based on the Bayes theorem and assigns data points to labels according to probability. Decision Tree uses simple decision rules to distinguish between labels at nodes and can be understood as a tree. KNN calculates the distance between data points and categorizes instances according to their neighbors. Within SVM, the instances are mapped into more-dimensional space according to the number of features and lines are used for splitting the data set for classification. Random Forest is an advanced version of the Decision Tree, as it consists of several uncorrelated Decision Trees and therefore reduces the possibility of overfitting. It has been proven that Random Forest is a suitable ensemble classifier for multi-dimensional data and complex problems (Ahmad et al. 2017) and performed best in several comparison studies of different classifiers for activity recognition (Baldominos et al. 2019; Lee et al. 2020).

For 10-fold cross validation, the folds were fixed so that all classifiers applied simulations with the same folds, although this would not be necessary for the Random Forest classifier. However, for the sake of comparison, cross validations for each classifier were executed in the same way. It could be detected that Random Forest achieved the highest accuracy among the investigated classifiers with 93.27 % (Table 1). This is an improvement of almost 20 % in comparison to Naïve Bayes and Decision Tree, 10 % better than KNN, and 7 % better than SVM. These results are plausible as the Random Forest algorithm is an enhancement of the Decision Tree algorithm, Considering the duration calculation and the fact that one window is equal to one second, the accuracy of Random Forest seems to be a suitable result.

Table 1: Classifier accuracy for 10-fold cross validation.

	Naïve Bayes	**Decision Tree**	**KNN**	**SVM**	**Random Forest**
Accuracy (%)	74.16	75.66	83.05	86.01	93.27

Another approach for evaluating the performance of classifiers is the use of confusion matrices by comparison of actual with predicted operation instances. The confusion matrix for the Random Forest as the classifier with the highest accuracy can be seen in Table 2 and it can be detected that there are no wrong classifications among the operations *Concrete Pouring* and *Filling*. This is due to the different altitudes of these operations. The wrong predictions occurred in relation to the operations *Lifting Up* and *Lifting Down*. The transition between these two operations was fluently, as, for instance, the concrete was poured into the formwork while the bucket was still lifted up by the crane. Thus, a clear distinction is challenging for the classifier.

Table 2: Confusion matrix for Random Forest.

		Prediction				
		Concrete pouring	Filling	Lifting down	Lifting up	Total
Actual	Concrete Pouring	1,333	0	30	76	1,439 (24.26 %)
	Filling	0	1,250	30	28	1,278 (21.54 %)
	Lifting Down	31	58	1,420	30	1,539 (25.94%)
	Lifting Up	76	17	23	1,560	1,676 (28.25%)
	Total	1,440	1,295	1,503	1,694	5,932

Precision and recall, which can be calculated by the confusion matrix, are performance metrics and were investigated to evaluate the quality of the classification approach. Precision is the fraction of relevant predicted instances among the positive predicted instances. Recall is the fraction of relevant predicted instances among all relevant instances in the data set. Precision and recall were calculated for each of the four operations for the Random Forest classifier (Table 3). The *Lifting Down* operation has the lowest value at 91.81 %, but there is no significant difference between recall and precision for each of the different operations. As the data set is balanced with ratios of 24.26 %, 21.54 %, 25.94 %, and 28.25 % for the respective operations (Table 2), these classifications can be seen as suitable. Overall, it can be detected that the results present a good balance between precision and recall.

The classifications of Random Forest were edited heuristically before calculating durations. For instance, if 15 windows were classified as *Concrete Pouring*, then two instances of *Lifting Up* occurred, followed by 23 instances of *Concrete Pouring*, they were merged to 40 instances of *Concrete pouring*. As the windows have a size of two seconds with 50 % overlap, the number of windows is multiplied by one second. By this approach, the duration of each repetition of every operation was calculated.

Table 3: Precision and recall for Random Forest.

Operation	Precision (%)	Recall (%)
Concrete pouring	92.70	92.63
Filling	93.75	95.07
Lifting down	94.33	91.81
Lifting up	92.00	93.32

4.5 Determining Probability Density Functions

To determine suitable PDFs, the following Goodness-of-Fit statistics were calculated: three hypothesis tests - Kolmogorov-Smirnov-Test (KS-Test), Cramér-von-Mises-Test (CvM-Test), and Anderson-Darling Test (AD-Test) - and two information criterion - Akaike information criterion (AIC) and Bayesian information criterion (BIC). The most frequently used Goodness-of-Fit methods are the KS-Test and the chi-square test, but the KS-Test leads to more precise results for continuous probability distributions (Massey 1951). The CvM-Test and AD-Test are calculated similarly, differing only by a different multiplier, and both are often more powerful than the KS-Test (Stephens 1986). In comparison to the KS-Test, the AD-Test emphasizes more risks by equal weighting of the body and tails of distributions. CvM-Test can be seen as a balanced approach between KS-Test and AD-Test. AIC and BIC are based on the Log-likelihood and are especially functional to prevent overfitting (Delignette-Muller and Dutang 2015).

The classified durations for the first concrete delivery and for both deliveries were used in Goodness-of-Fit analyses to investigate the update of PDFs. Hence, for each of the four operations, PDFs according to two different input sources were investigated. Each operation was repeated nine times for the first concrete delivery. For the second delivery, the operations were executed eight times, besides the operation *Lifting Down*, which was executed only seven times because the bucket was cleaned on the fifth floor after finishing execution for the working day. An example for the *Lifting Down* operation based on the classified durations for the whole data set is presented in Table 4 for a comparison of different, but not all investigated distributions. The Weibull distribution fits best according to the input data for all five Goodness-of-Fit methods.

Table 4: Comparison of Goodness-of-Fit results for Lifting Down.

	Normal	Log-normal	Logistic	Cauchy	Weibull
KS-Test	0.186	0.196	0.179	0.178	0.151
CvM-Test	0.064	0.077	0.061	0.110	0.053
AD-Test	0.356	0.432	0.351	0.677	0.306
AIC	123.32	124.15	123.86	129.22	122.91
BIC	124.87	125.70	125.41	130.76	124.46

A significance level of 0.05 was chosen and the calculated p values for each of the three hypothesis tests are above the significance level. Thus, the null hypothesis for the Weibull distribution was not rejected and the distribution can be used. The resulting PDFs for the operations according to Goodness-of-Fit analyses can be seen in Table 5. The execution of the construction works was done a bit faster during the second concrete delivery in comparison to the first truck, although it was getting darker as the second truck arrived around 06:00 pm. The shorter durations can be detected in the identified PDFs. It can be determined that it is possible to infer useful information from a small amount of data. This can aid in determining a suitable updating interval for the PDFs.

Table 5: Resulting PDFs.

Operation	Data set: First truck	Data set: Both trucks
Filling	Cauchy (87.04, 5.64)	Logistic (80.42, 10.34)
Lifting Up	Logistic(98.68, 9.80)	Logistic (96.28, 8.29)
Concrete Pouring	Log-normal (4.30, 0.49)	Log-normal (4.17, 0.69)
Lifting Down	Weibull (10.71, 93.93)	Weibull (10.54, 95.54)

4.6 Data-Driven Discrete Event Simulation

In the following, the whole construction process was simulated by DES and the total construction duration was calculated. The only deviation from the real procedure was a just-in-time delivery of concrete instead of consideration of idle time. The results of the data-driven DES for the first truck and for both trucks were compared. To get reliable results, a Monte-Carlo simulation was applied and the DES was repeated 2,000 times. Before evaluating the results of the DES, outliers were removed by the boxplot approach:

$$\text{Lower outliers} < Q25 - 1.5 * IQR$$
$$\text{Upper outliers} > Q75 + 1.5 * IQR.$$

According to the boxplot outlier search, for the data-driven DES based on the first truck 147 outliers, 7.4 %, were detected and there were only 26 outliers for the data-driven DES based on two trucks. Finally, this results in the total construction durations in Table 6 and the density plots in Figure 3. The green density curve presents the durations based on the classified data set of the first truck and the blue curve shows the data-driven model based on the whole data set. The dotted red line displays the construction process's real duration of 5,932 seconds.

Table 6: Resulting durations from data-driven DES.

	Min (s)	Q25 (s)	Median (s)	Mean (s)	Q75 (s)	Max (s)
Data set: First truck (green)	5,416	5,999	6,162	6,181	6,336	6,979
Data set: Both trucks (blue)	5,215	5,755	5,931	5,950	6,120	6,713

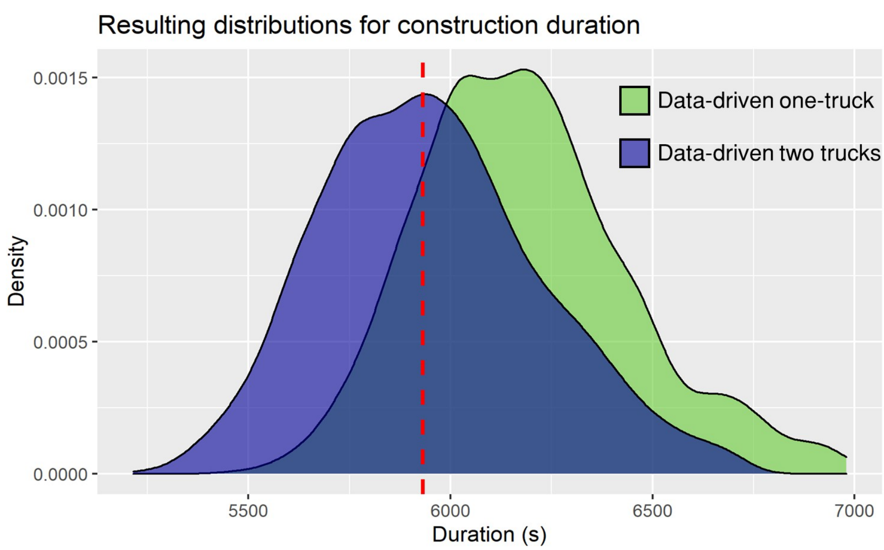

Figure 3: Density plots for total construction durations.

Figure 3 shows that the blue curve surrounds the real duration much better than the green curve. It can clearly be detected that the results based on the whole data set are closer to the real value. The median is only one second lower than the real duration, as the median based on the data set of the first truck is 230 seconds higher. The reason for this is that the execution of the operations during the second truck was a bit faster. This could be detected in the respective PDFs and in the results of the DES. The green curve is based only on a sample of the whole population and, thus, the results of the DES are inaccurate in comparison to the DES based on the whole data set. This emphasizes the need for DTC to update PDFs to model more reliable simulations.

4.7 Calibration

For calibration of the model, the Wilcoxon signed-rank test was applied to test for validity. The Wilcoxon signed-rank test is a hypothesis test that can compare a known value – the real duration – to the median of a data set – the results of the Monte-Carlo method – by ranking the results of the Monte-Carlo method in relation to the median. The Wilcoxon signed-rank test has greater statistical power in comparison to the widespread Student's t-test, as proven in several simulation studies (Blair and Higgins 1980). Within the Wilcoxon signed-rank test, a p-value is calculated. If the resulting p value is above the advanced determined significance level of 0.05, the null hypothesis cannot be rejected and the model is valid. The p value for the data set of one truck is 0.0 and for the data set of both trucks 0.14. Thus, it can be observed that only the DES model based on the whole data set is valid, as the model that used data only from the first truck has to be rejected according to the calibration.

5 CONCLUSION

The proposed framework helps to control construction activities and further improve decision-making based on real-time data-driven DES. For instance, the delivery period of upcoming concrete truck supplies can be adjusted or different construction activities can be coordinated to reduce idle times for resources. Hence, the developed approach helps to overcome the two most affecting attributes regarding construction productivity: planning and scheduling of activities and material deliveries (Dixit and Sharma 2020).

This research promotes the usage of real-time data during construction equipment's usage by advanced analytics technologies for leveraging digitalization and the DT concept in the AEC field. The proposed approach successfully shows how to apply the DT concept in the construction phase. Collected data are used for gaining project status information and knowledge about the current situation. Thus, an information exchange regarding as-built product and as-performed process can be achieved in real-time. As proven by the calibration, real-time data-driven DES enables more reliable results by updating PDFs continuously. The former approaches, by using static values or only simplified data-based activity durations, neglect the dynamics during construction execution. According to the central limit theorem, the results of the Monte-Carlo approach tend to a normal distribution, which can be used for planning similar projects in the future. Hence, to determine the most suitable final distribution, it is essential to combine the different distributions. From a practical point of view, the saving of collected DT data on a database will offer valuable knowledge for the planning of future construction projects. The different distributions can be adapted to the equipment, such as buckets with a higher capacity, if needed. After construction starts, further real-time data can be collected to update the PDFs according to the respective situations on site.

The resulting accuracies of the ML classifiers are consistent in comparison to other kinematic equipment activity studies, which have a higher sample rate during data collection and did not apply a PCA (Akhavian and Behzadan 2015; Kim et al. 2018; Rashid and Louis 2019). The accuracy of this approach can still be improved by hypertuning of the algorithms in future studies, but the scope of this research lies in presenting the whole procedure and the validity of the model was proven by calibration. On the basis of this study, the usage of data-driven DES offers the possibility to improve aspects such as safety, costs, and resource efficiency. An in-depth investigation of different key performance indicators within DES is needed in the future. This would allow for addressing the decision-making process by considering the effects of management decisions more comprehensively. A limitation of this research is that data were collected only during optimal weather conditions. DES enables stochastic simulation of processes and consideration of uncertainties and risks. Thus, a deeper investigation of risk factors such as changing weather conditions or congestion during material deliveries would be expedient in future research.

ACKNOWLEDGMENTS

Part of this research has been conducted for the H2020 funded project ASHVIN, Grant Number: 958161. The authors are grateful to BIS structures, BCA, DRAGADOS, and FREO for providing access to construction site and for sharing project information.

REFERENCES

Abbasi, S., K. Taghizade, and E. Noorzai. 2020. "BIM-Based Combination of Tact Time and Discrete Event Simulation for Implementing Just in Time in Construction Scheduling under Constraints". *Journal of Construction Engineering and Management Volume* 146(12): 04020143:1-04020143:15.

Ahmad, M.W., M. Mourshed, and Y. Rezgui. 2017. "Trees vs Neurons: Comparison between Random Forest and ANN for High-Resolution Prediction of Building Energy Consumption". *Energy and Buildings* 147:77-89.

Akhavaian, R., and A. H. Behzadan. 2015. "Construction Equipment Activity Recognition for Simulation Input Modelling Using Mobile Sensors and Machine Learning Classifiers". *Advanced Engineering Informatics* 29:867-877.

Alvanchi, A., F. Baniassadi, M. Shahsavari, and H. Kashani. 2021. "Improving Material Logistics Plan in Road Construction Projects Using Discrete Event Simulation". *Engineering, Construction and Architectural Management* 28(10):3144-3162.

Banks, J., J. S. Carson II, B. L. Nelson, and D. M. Nicol. 2010. *Discrete-Event System Simulation.* 5th ed. Upper Saddle River, New Jersey: Pearson Education Inc..

Baldominos, A., A. Cervantes, Y. Saez, and P. Isasi. 2019. "A Comparison of Machine Learning and Deep Learning Techniques for Activity Recognition Using Mobile Devices". *Sensors* 19(3):521:1-521:21.

Behzadan, A. H., C. C. Menassa, and A. R. Pradhan. 2015. "Enabling Real Time Simulation of Architecture, Engineering, Construction, and Facility Management (AEC/FM) Systems: A Review of Formalism, Model Architecture, and Data Representation". *Journal of Information Technology in Construction* 20:1-23.

Blair, R. C. and J. J Higgins. 1980. "A Comparison of the Power of Wilcoxon's Rank-Sum Statistic to that of Student's t Statistic Under Various Nonnormal distributions". *Journal of Educational Statistics* 5(4):309-335.

Brilakis, I., Y. Pan, A. Borrmann, H.-G. Mayer, C. Vos, E. Pettinato, and S. Wagner. 2019. "Built Environment Digital Twinning, Report for the International Workshop on Built Environment Digital Twinning". Technical University of Munich.

Cheng, T., and J. Teizer. 2013. "Real-Time Resource Location Data Collection and Visualization Technology for Construction Safety and Activity Monitoring Applications". *Automation in Construction* 34:3-15.

Delignette-Muller, M. L. and C. Dutang. 2015. "Fitdistrplus: An R Package for Fitting Distributions". *Journal of Statistical Software* 64(4):1-34.

Dixit, S., and K. Sharma. 2020. "An Empirical Study of Major Factors Affecting Productivity of Construction Projects". In: *Emerging Trends in Civil Engineering*, edited by K. G. Babu, H. S. Rao, and Y. Amarnath, 121-129, Singapore: Springer.

Golparvar-Fard, M., A. Heydarian, and J. C. Niebles. 2013. "Vision-Based Action Recognition of Earthmoving Equipment Using Spatio-Temporal Features and Support Vector Machine Classifiers". *Advanced Engineering Informatics* 27(4):652-663.

Halpin, D. W. and L. S. Riggs. 1992. *Planning and Analysis of Construction Operations*. New York: John Wiley & Sons, Inc.

Kim, H., Y. Ham, W. Kim, S. Park, and H. Kim. 2019. "Vision-Based Nonintrusive Context Documentation for Earthmoving Productivity Simulation". *Automation in Construction*102:135-147.

Kim, J., S. Chi, and C. R. Ahn. 2021. "Hybrid Kinematic-Visual Sensing Approach for Activity Recognition of Construction Equipment". *Journal of Building Engineering* 44:102709:1-102709:10.

Langroodi, A. K., F. Vahdatikhaki, and A. Dorée. 2021. "Activity Recognition of Construction Equipment Using Fractional Random Forest". *Automation in Construction* 122:103465:1-103465:17.

Lee, Y.-C., M. Scarpiniti, and A. Uncini. 2020. "Advanced Sound Classifiers and Performance Analyses for Accurate Audio-Based Construction Project Monitoring". *Journal of Computing in Civil Engineering* 34(5):04020030:1-0402003:11.

Li, F., K. Shirahama, M. A. Nisar, L. Köping, and M. Grzegorzek. 2018. „Comparison of Feature Learning Methods for Human Recognition Using Wearable Sensors". *Sensors* 18(2):679:1-679:22.

Lin, P. C., J. Wang, S..H. Huang, and Y.-T. Wang. 2010. "Dispatching Ready Mixed Concrete Trucks under Demand Postponement and Weight Limit Regulation". *Automation in Construction* 19(6):798-807.

Liu, H., M. S. Altaf, Z. Lei, M. Lu, and M. Al-Hussein. 2015. "Automated Production Planning in Pannelized Construction Enabled by Integrating Discrete-Event Simulation in BIM". In *the Canadian Society for Civil Engineering 5th International/11th Specialty Conference, June 7-10, Vancouver, Canada, 8-10.*

Massey Jr., F. J. 1951. “The Kolomogorov-Smirnov Test for Goodness of Fit”. *Journal of the American statistical Association* 46(253):68-78.

Nakanishi, Y., T. Kaneta, and S. Nishino. 2022. “A Review of Monitoring Construction Equipment in Support of Construction Project Management”. *Frontiers in Built Environment* 7:632593:1-632593:11.

Patteri, D. and S. Stolton,. 2019. Digitizing the EU’s Construction Industry. http://eurac.tv/9Q2S, accessed March 2019.

Rashid, K. M., and J. Louis. 2019. “Time-Series Data Augmentation and Deep Learning for Construction Equipment Activity Recognition”. Advanced Engineering Informatics42:100944:1-100944:12.

Sacks, R., I. Brilakis, E. Pikas, H. S. Xie, and M. Girolami. 2020. “Construction with Digital Twin Information Systems”. *Data-Centric Engineering* 1:E14:1-E14:26.

Seppänen, O., J. Evinger, and C. Mouflard. 2014. “Effects of the Location-Based Management System on Production Rate and Productivity”. *Construction Management and Economics* 32(6):608-624.

Sherafat, B., C. R. Ahn, R. Akhavian, A. H. Behzadan, M. Golparvar-Fard, H. Kim, Y.-C. Lee, A. Rashidi, and E. R. Azar. 2020. “Automated Methods for Activity Recognition of Construction Workers and Equipment: State-of-the-Art Review”. *Journal of Construction Engineering and Management* 146(6):03120002:1-03120002:19.

Slaton, T., C. Hernandez, and R. Akhavian. 2020. “Construction Activity Recognition with Convolutional Recurrent Networks”. *Automation in Construction* 113:103138:1-103138:12.

Song, L., and N. N. Eldin. 2012. “Adaptive Real-Time Tracking and Simulation of Heavy Construction Operations for Look-Ahead Scheduling”. *Automation in Construction* 27:32-39.

Stephens, M. A. 1986. “Tests based on EDF Statistics”. *Goodness-of-Fit Techniques*, edited by R. B. D’Agostino and M. Q. Stephens, 97-194. New York: Marcel Dekker Inc.

Vahdatikhaki, F., and A. Hammad. 2014. “Framework for Near Real-Time Simulation of Earthmoving Projects Using Location Tracking Technologies”. *Automation in Construction* 42:50-67.

Vincent, S. 1998. “ Input Data Analysis”. In *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*, edited by J. Banks, 55-93, Hoboken: John Wiley & Sons, Inc.

Xue, J., X. Hou, and Y. Zeng. 2021. “Review of Image-Based 3D Reconstruction of Building for Automated Construction Progress Monitoring”. *Applied Scien*ces 11(17):7840:1-7840:24.

Zheng, L., D. Wu, X. Ruan, S. Wenig, A. Peng, B. Tang, H. Lu, H. Shi, and H. Zheng. 2017. “A Novel Energy-Efficient Approach for Human Activity Recognition”. *Sensors* 17(9):2064:1-2064:21.

AUTHOR BIOGRAPHIES

MANUEL JUNGMANN is a Research Assistant in the ASHVIN project and PhD candidate at the department for Civil Systems Engineering at the Technische Universität Berlin. His research interests include how to use real-time data for stochastic discrete event simulation of construction processes. His e-mail address is manuel.jungmann@tu-berlin.de.

LUCIAN UNGUREANU is a Research Assistant at the Technische Universität Berlin. In his role he is coordinating ASHVIN, an European research project that focuses on developing the required tools and methods to create digital twins for safer, more productive, and more resource efficient construction phases. His research interests are in understanding how certain engineering tasks can be automated to a high degree, but also how to meaningfully design and egineers civil systems that are easy to change. His e-mail address is l.ungureanu@tu-berlin.de

TIMO HARTMANN is Professor for Civil Systems Engineering at the Technische Universität Berlin. In this role he is managing the International Masters program Civil Systems Engineering teaching students advanced digital engineering methods. His research interests are in understanding how computational methods allow to formalize complex engineering knowledge. His e-mail address is timo.hartmann@tu-berlin.de.

HETOR POSADA is a PhD candidate in Construction Engineering at UPC. His main focus is on the development of automated pipelines of information between physical models and FEM simulations. His e-mail address is hector.posada@upc.edu.

ROLANDO CHACON is an Associate Professor at the department of Civil and Environmental Engineering at UPC. He is presently Chair person of the division of structural technology, construction and materials. Research ranges from structural analysis of steel structures to bridge monitoring and digital twins. His e-mail address is rolando.chacon@upc.edu.