Dynamic Modeling of Multi-Factor Construction Productivity for Equipment-Intensive Activities

Nima Gerami Seresht[1], Ph.D. and
Aminah Robinson Fayek[2], Ph.D., P Eng., M.ASCE (corresponding author)

Abstract

Construction productivity is of major research interest within the construction domain. Since construction is a labor-intensive industry, previous research has often focused on construction labor productivity (CLP). However, equipment is the main driver of productivity for some construction activities, so called equipment-intensive activities. Existing models of activity-level productivity often predict a single-factor productivity measure —namely CLP—, yet determining multi-factor productivity, including labor, material, and equipment, provides more comprehensive predictions of productivity. Construction productivity models are often static in nature, or incapable of capturing the subjective uncertainty of some factors influencing productivity. Fuzzy system dynamics is an appropriate technique for modeling construction productivity, since it captures the dynamism of construction projects, and addresses the subjective and probabilistic uncertainty of factors influencing productivity. The contributions of this paper are threefold: identifying the key factors influencing the productivity of equipment-intensive activities, developing a predictive model of multi-factor productivity for equipment-intensive activities using fuzzy system dynamics technique; and developing an approach to reduce uncertainty overestimation in the simulation results of fuzzy system dynamics models.

[1] Post-doctoral fellow, Department of Civil & Environmental Engineering, 7-203 Donadeo Innovation Centre for Engineering, 9211 116 St NW, University of Alberta, Edmonton AB T6G 1H9, Canada.
[2] Tier 1 Canada Research Chair in Fuzzy Hybrid Decision Support Systems for Construction, NSERC Industrial Research Chair in Strategic Construction Modeling and Delivery, Ledcor Professor in Construction Engineering, Professor, Department of Civil & Environmental Engineering, 7-287 Donadeo Innovation Centre for Engineering, 9211 116 St NW, University of Alberta, Edmonton AB T6G 1H9, Canada.

Keywords

Construction productivity, system dynamics, fuzzy logic, construction equipment

Introduction

Construction productivity has been a major research interest within the construction management domain for some time. Previous research on construction productivity has either focused on identification of the factors influencing construction productivity or on the development of predictive models for construction productivity. Due to the fact that construction is a labor-intensive industry (Jarkas 2010), previous studies on the activity-level productivity have primarily focused on construction labor productivity (CLP) (Tsehayae and Fayek 2014, Naoum 2016, Tsehayae and Fayek 2016a, Mirhadi and Zayed 2016). However, construction equipment are now important resources in construction projects, and they are the drivers of productivity for some activities. Goodrum and Hass (2004) observed substantial long-term improvement in the productivity of the activities executed using equipment with significant technological advancements. Goodrum et al. (2010) developed a predictive model to measure the effect of equipment on construction productivity; this research confirms that technological advancements in construction equipment affect construction productivity. Ok and Sinha (2006) discussed that accurate prediction of the construction productivity for some activities (e.g., earthmoving operations) depends on the accurate prediction of the equipment production rate. These construction activities are identified as equipment-intensive activities, where equipment, rather than labor, is the driver of productivity. Accordingly, the factors influencing the productivity of equipment-intensive activities are different from the factors influencing the productivity of labor-intensive activities. Therefore, in order to model the productivity of equipment-intensive activities, the factors influencing the productivity of these activities must be identified.

Existing predictive models for construction productivity mostly focus on predicting CLP, which determines productivity of construction systems (i.e., construction projects or construction activities) using only one resource input (e.g., labor). Previous studies confirm that for determining productivity of construction systems, using the other resource inputs (i.e., equipment and material) in addition to labor, results in more comprehensive measures of productivity compared to CLP (Loosemore 2014). However, Carson and Abbott (2012) concluded that the construction industry suffers from a lack of predictive models that determine productivity of construction systems using such comprehensive productivity measures.

Existing predictive models of construction productivity were mostly developed using static techniques (e.g., artificial neural network (ANN) model by Heravi and Eslamdoost (2015); fuzzy rule-based system model by Tsehayae and Fayek 2016a), which means that they predict a single productivity value at a given point in time. However, due to the dynamic nature of construction projects, modeling techniques that are able to track project changes over time are more suitable for modeling construction productivity. Moreover, the factors influencing construction productivity are rarely independent from each other, and changes in certain factors can impact other factors (Mawdesley and Al-Jibouri 2009). Therefore, the cause and effect relationships between the factors influencing construction productivity need to be captured, along with their individual impact on productivity. The fuzzy system dynamics (FSD) technique, which integrates system dynamics (SD) with fuzzy logic, is an appropriate technique for modeling construction productivity. The SD component of the FSD technique captures the dynamism of construction projects and the relationships between the factors influencing construction productivity, while the fuzzy logic component addresses the subjective uncertainty of these factors.

This paper presents an FSD model of activity-level construction productivity that measures the productivity of equipment-intensive activities using the three resource inputs of construction activities (i.e., labor, equipment, and material). For this purpose, the factors influencing the productivity of equipment-intensive activities were first identified. Second, in order to increase the accuracy of the predictive model, the number of factors influencing construction productivity was reduced by feature selection. Third and fourth, the qualitative and quantitative FSD models of multi-factor productivity (MFP) were developed. Finally, the FSD model was validated using a case study of earthmoving operations on an actual construction project. This paper advances the state of the art in construction productivity modeling by identifying the key factors influencing the productivity of equipment-intensive activities, and by developing the FSD model of MFP for equipment-intensive activities. This FSD model simultaneously captures the dynamism of construction productivity (i.e., changes in productivity over time) and the cause and effect relationships between the factors influencing productivity, as well as the probabilistic and subjective uncertainties of the factors influencing construction productivity.

Literature Review

Construction productivity

In general, the productivity of a construction system (e.g., construction activity, construction project) can be calculated as the ratio of the inputs of the system (e.g., labor cost or person-hours) to its output (e.g., cubic meters of concrete placed). Talhouni (1990) introduces three different measures for construction productivity: (1) single factor productivity (SFP), which measures the productivity of construction systems using only one resource input (i.e., labor); (2) multi-factor productivity (MFP), which measures the productivity of construction systems using any combination of three resource inputs (i.e., labor, materials, and equipment); and (3) total factor productivity (TFP), which measures the productivity of construction systems using five resource inputs (i.e., labor, materials, equipment, energy, and capital). From the construction management perspective, construction productivity is often defined at the project level or the activity level, using two measures: construction labor productivity (CLP), which is a SFP measure that uses labor as the only input of productivity (Tsehayae and Fayek 2016a), or MFP, which uses any combination of the three inputs of productivity (i.e., labor, equipment, and material) (Eastman and Sacks 2008). Measuring the TFP at the project or activity levels can be inaccurate, due to the difficulties encountered in predicting the energy and capital inputs at the project or activity levels (Thomas et al. 1990, Loosemore 2014). Thus, MFP represents the most comprehensive measure of construction productivity at the project and activity levels. However, unlike other industries for which MFP measures of productivity are available, the construction industry suffers from a lack of predictive models for determining the MFP of construction systems (Carson and Abbott 2012).

Depending on which resource is the main driver of the productivity, construction activities can be grouped into two categories: labor-intensive activities, where labor is the main driver of productivity (e.g., electrical and mechanical activities) (Jarkas 2010), and equipment-intensive activities, where equipment is the main driver of productivity (e.g., earthmoving activities) (Ok and Sinha 2006). While the productivity of labor-intensive activities is mainly affected by CLP, the productivity of equipment-intensive activities is mainly affected by the production rate of the equipment used for the execution of the activity. There are numerous equipment-intensive activities in different types of construction projects, including earthmoving (Ok and Sinha 2006, Jabri and Zayed 2017), pavement construction (Choi and Ryu 2015), pile construction (Zayed and Halpin 2005), and tunneling (Shaheen et al. 2009). Since the resource that drives the productivity of labor-intensive and equipment-intensive activities is different, the factors that influence the productivity of these two types of activities are also different. However, previous research on construction productivity has failed to identify a comprehensive list of factors influencing the productivity of equipment-intensive activities. Moreover, traditionally the production rate of the equipment used for the execution of a given equipment-intensive activity is measured as the efficiency measure of the activity (Zayed and Halpin 2005, Shaheen et al. 2009, Jabri and Zayed 2017). Zayed and Halpin (2005) identified 12 factors that influence the production rate of a piling activity; they developed a statistical model using the linear regression method to predict the production rate of a piling activity in terms of number piles drilled per day. Shaheen et al. (2009) identified 11 factors that influence the production rate of a tunneling activity using a tunnel boring machine (TBM); they used an expert-driven fuzzy rule-based system (FRBS) to predict the production rate of the activity based on these 11 factors. Moreover, Shaheen et al. (2009) developed a discrete event simulation model to predict the total duration of the activity based on the production rate determined by the FRBS. Finally, Jabri and Zayed (2017) developed a predictive model for the production rate of an earthmoving operation using the agent-based modeling technique. The model developed by Jabri and Zayed (2017) predicts the production rate and total duration of an earthmoving operation based on the equipment and labor properties and environmental factors that affect the activity. The production rate is commonly calculated as the output per unit time. Although the prediction of production rate in the aforementioned studies facilitates the evaluation of the duration of equipment-intensive activities, the production rate does not provide comprehensive information regarding the resource inputs (i.e., labor, equipment, and material) and consequently the cost efficiency of these activities. Moreover, there are also a few predictive models developed for equipment-intensive activities that measure the CLP of these activities (Choi and Ryu 2015). Choi and Ryu (2015) identified nine factors that influence the productivity of highway pavement activities and developed a predictive model to measure the CLP of such activities using statistical methods. However, CLP is not an appropriate measure of productivity for equipment-intensive activities, since it does not provide any information regarding the resource input (equipment) that is the main driver of productivity for these activities. Therefore, there is a need to develop a predictive model for determining the MFP of equipment-intensive activities.

Finally, the existing predictive models of construction productivity are commonly developed using static modeling techniques, such as the ANN model developed by Heravi and Eslamdoost (2015). However, dynamic modeling techniques such as SD and FSD are more appropriate for modeling construction productivity, since construction systems are dynamic (i.e., changing over time) and their components interact with each other (Mawdesley and Al-Jiboury 2009, Alzraiee et al. 2015). Moreover, SD models of construction productivity (Mawdesley and Al-Jiboury 2009, Nasirzadeh and Nojedehi 2014) cannot capture the subjective uncertainty of the factors influencing productivity. Accordingly, Nojedehi and Nasirzadeh (2017) suggested that FSD is an appropriate technique for modeling construction productivity, since this technique captures the dynamism of construction systems and the interactions between the factors influencing productivity, while simultaneously representing the probabilistic and subjective uncertainty of these factors. Nojedehi and Nasirzadeh (2017) developed a predictive model of CLP using the FSD technique. Since their predictive model is developed for labor-intensive activities and predicts CLP, it is not an appropriate model for predicting the productivity of equipment-intensive activities. Accordingly, there is a need within the existing body of construction research to develop a predictive model of the MFP for equipment-intensive activities using FSD technique, which is addressed in this paper.

Fuzzy system dynamics

SD is a simulation methodology developed by Forrester (1961) for analyzing complex industrial systems. This modeling technique is able to model a dynamic system, in which the state of the system (e.g., construction productivity) changes over time and under the effect of different factors. Although SD models are able to capture the probabilistic uncertainties of real-world systems using the Monte Carlo simulation technique (Sterman 2000), these models cannot capture the non-probabilistic uncertainties (i.e., subjective, imprecise, or linguistically expressed information) of real-world systems. To address this limitation, Levary (1990) integrated SD with fuzzy logic and developed the fuzzy system dynamics (FSD) technique, which is capable of capturing deterministic values, as well as probabilistic and non-probabilistic uncertainties. Moreover, the fuzzy logic component of the FSD technique allows practitioners to evaluate subjective variables using linguistic terms, rather than precise numerical values.

FSD simulation models are developed through qualitative and quantitative modeling steps. First, the qualitative FSD model is developed by identifying and modeling the factors influencing the system, which are called system variables. Next, the quantitative FSD model is developed by developing fuzzy membership functions to represent the subjective system variables and defining the relationships between the system variables quantitatively. The fuzzy membership functions representing subjective system variables can be developed by one of the several approaches proposed in the literature, either by using data (e.g., fuzzy c-means (FCM) clustering approach) or by using expert knowledge (e.g., Saaty's priority approach). The relationships between system variables are defined by mathematical equations or by fuzzy rule-based systems (Khanzadi et al. 2012, Nasirzadeh et al. 2013). There are two types of relationships between system variables: hard relationships, where the mathematical form of the relationship is known, and soft relationships, where the mathematical form of the relationship is unknown (Coyle 2000). The hard relationships of FSD models are defined by mathematical equations, and fuzzy arithmetic operations are used in the mathematical equations that include subjective variables. The soft relationships of FSD models are defined either by mathematical equations developed statistically, if data are available to do so, or by FRBS developed by expert knowledge if data are not available (Khanzadi et al. 2012).

By implementing fuzzy arithmetic in the mathematical equations of the FSD models, the supports of the membership functions, which represent the simulation results, grow rapidly, producing a large amount of uncertainty (Tessem and Davidsen 1994). This phenomenon is called the overestimation of uncertainty, which reduces the ability of users to accurately predict the actual system output (e.g., actual productivity) based on the simulation results (Lin et al. 2011). The overestimation of uncertainty in the FSD models may be affected by various factors such as the number of parameters in the mathematical equations, number of time steps, membership functions of the inputs, and the method of the fuzzy arithmetic implementation. Fuzzy arithmetic operations can be implemented using one of the two following methods: the α-cut method, and the extension principle method, which uses different *t*-norms (Pedrycz and Gomide 2007). Implementing fuzzy arithmetic by the extension principle method using drastic product *t*-norm reduces the uncertainty overestimation in comparison to the α-cut method (Chang et al. 2006, Lin et al. 2011). However, in previous applications of FSD models in different construction areas such as risk analysis (Nasirzadeh et al. 2014), project contract administration (Khanzadi et al. 2012), and construction productivity (Nojedehi and Nasirzadeh 2017), fuzzy arithmetic is implemented only by the α-cut method due to its simplicity. As an example, in the results of the FSD model developed by Nojedehi and Nasirzadeh (2017), the α-cut method causes overestimation of uncertainty in the fuzzy number that represents labor productivity of concrete pouring activity, $[4.08, 29.37]\,\frac{\mathrm{m}^3}{\mathrm{month}}$, where the upper bound of the support is 620% larger than its lower bound. The large amount of uncertainty in the simulation results reduces the ability of users (e.g., construction practitioners) to accurately predict the actual productivity, project cost, and project duration based on the simulation results. In this paper, this limitation is addressed by implementing fuzzy arithmetic operations using the extension principle method with the min, algebraic product, Lukasiewicz, and drastic product t-norms, and selecting the most appropriate method to increase the accuracy of the simulation results, while simultaneously reducing the amount of uncertainty.

Construction Productivity Modeling Methodology

This section of the paper outlines the development of the FSD model of activity-level productivity for equipment-intensive activities; this process was accomplished in the following five steps: (1) identification of the factors influencing construction productivity, (2) reduction of the dimensionality of the factors by feature selection, (3) development of the qualitative FSD model, (4) development of the quantitative FSD model, and (5) validation of the full FSD model. These five steps are presented in Fig. 1.

In the first step, the factors influencing productivity of equipment-intensive activities were identified through a literature review. There are numerous studies available in the literature, which identify the factors influencing construction productivity at different levels of analysis (e.g., activity-level or project-level). In addition to micro-level factors (i.e., crew-level, activity-level, and project-level), macro-level factors (i.e., organizational-level, provincial-level, national-level, and global-level) may directly or indirectly influence construction productivity (Tsehayae and Fayek 2014). However, since the project-level and macro-level factors are static (i.e., constant) at the activity level, these factors are excluded from the FSD model presented in this paper. Thus, in

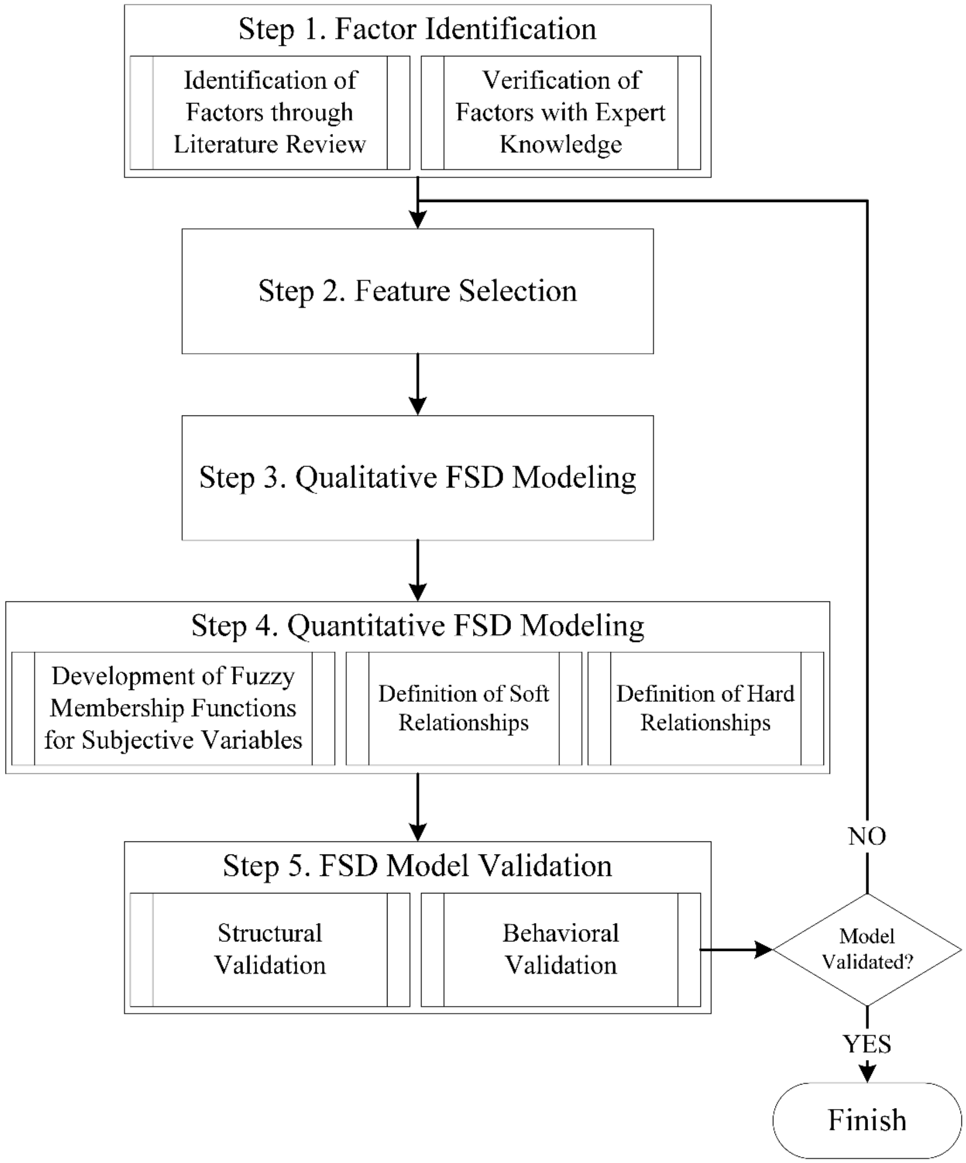

Fig. 1. Methodology for construction productivity modeling by FSD technique.

this paper, the crew-level and the activity-level factors that influence the productivity of equipment-intensive activities were identified through literature review. Next, the identified factors were verified by expert knowledge using interview surveys, which were administered to managerial personnel (i.e., general management, project management, project controls, and field engineers), and field personnel (i.e., labors/equipment operators and foremen) within a Canadian company active in the industrial construction sector. Fifteen project management surveys and 20 tradespeople surveys were collected and analyzed to verify the factors influencing construction productivity identified from the literature review. The respondents of the managerial personnel survey had an average of six years of experience in the construction industry and were involved in an average of six industrial pipeline projects. The respondents of the field personnel survey were most frequently union members (i.e., 95% of the respondents), who were involved in numerous projects with and average of 10 years of experience in industrial pipeline projects. The interview surveys assessed the impact of each factor on construction productivity using a seven point Likert scale, as suggested by Tsehayae and Fayek (2014) and Dai (2006). The scale used in the interview surveys had three levels of negative impact determined by negative impacts scores (i.e., "strongly negative" [-3], "negative" [-2], and "slightly negative" [-1]), one neutral point determined by zero (i.e., "no impact" [0]), and three levels of positive impact determined by positive impacts scores (i.e., "slightly positive" [+1], "positive" [+2], "strongly positive" [+3]). Table 1 presents an example of survey questions measuring the impact of the factors that affect construction productivity.

Table 1. Example of interview survey question.

Factors	Impact						
	Strongly negative	*Negative*	*Slightly negative*	*No impact*	*Slightly positive*	*Positive*	*Strongly positive*
The crew size is adequate for the task at hand	-3	-2	-1	0	1	2	3

Consequently, 72 crew-level and activity-level factors were identified through the literature review and were verified by expert knowledge to have either a negative or positive impact on construction productivity; these factors were grouped into seven categories based on their source (e.g., foreman-related factors, location-related factors, etc.). These factors were identified through the following previous studies: Zakeri et al. (1996), Goodrum and Hass (2004), Zayed and Halpin (2005), Ok and Sinha (2006), Mortaheb et al. (2007), Goodrum et al. (2010), Kannan (2011), Choi and Ryu (2015), and Sadeghpour and Andayesh (2015). Table 2 presents the 72 identified factors, their categories, and their average impact score on construction productivity; these factors are referred to as the system variables in the following steps.

Table 2. Crew- and activity-level factors influencing productivity of equipment-intensive activities.

Category	Factors (average impact score)
Crew-level factors	
Labor and crew	Crew size (+1.94), crew composition (+1.47), crew experience (+2.26), adequacy of crew (+1.94), crew makeup changes (-0.97), crew turnover rate (-1.66), number of languages spoken in the crew (-2.11), crew motivation (+2.37), level of interruptions and disruptions (-1.39), number of consecutive working days (-1.53), total daily overtime work (-1.47), crew skill level (+2.26), unscheduled breaks (-1.41), late arrival/early quit (-2.03), level of absenteeism (-1.82)
Material and consumables	Material availability (+2.03), waiting time for material (-1.77), material quality (-1.63), material storage practice (+1.86), pre-installation requirements (-1.06)
Equipment and tools	Number of equipment (+1.74), equipment breakdown frequency (-1.51), equipment breakdown downtime (-1.51), equipment maintenance frequency (-1.43), equipment maintenance downtime (-1.43), work equipment availability (+1.74), equipment delivery to working area (-1.80), appropriateness of equipment (+1.97), equipment ownership (+1.43), equipment production capacity (+1.97), equipment age (-1.29), equipment operator experience (+2.29), equipment operator education and trainings (+2.14), equipment operator skill level (+2.29), amplification of human energy (+1.97), level of control (+1.88), functional range (+1.71), equipment ergonomic design (+1.62), information feedback provision (+1.44), moving technology (+1.88), equipment warranty (+0.67), equipment specification (+1.97)
Foreman	Foreman experience (+2.11), change of foreman (-1.74), work planning skills (+2.14), leadership and supervisory skills (+2.14), coordination between labor and equipment operators (+2.15)
Activity-level factors	
Task characteristics	Task complexity (-1.15), total volume of work (+1.76), task repetitiveness (+1.38), out-of-sequence work (-1.24), problems with predecessors (-1.32), construction method (+1.93), task waste disposal (-0.79), rework frequency (contractor initiated) (-1.71), rework cost (contractor initiated) (-1.71), balance between labor and equipment (+1.91)
Location properties	Spaciousness of working area (+1.57), site restrictions (-1.13), soil type (-1.61), soil moisture (-1.61), groundwater level (-1.24), underground facilities (-1.24), hauling/delivery distance (-0.94)
Engineering/ instructions	Availability of drawings (+1.59), quality of drawings (+1.62), number of revisions on drawings (-1.24), design changes (-1.24), quality of specifications (+1.64), time to respond to RFIs (+1.41), time to do inspections (+1.35), rework frequency (design initiated) (-1.71), rework cost (design initiated) (-1.71)

Next, the number of system variables was reduced by feature selection to increase the accuracy of the predictive model for construction productivity (Ahmad and Pedrycz 2011). There are various methods for feature selection, out of which correlation-based feature selection (CFS) is the most common approach, due to its simplicity (Hall 1998). CFS reduces the dimensionality of the dataset by selecting the subset of the factors that have the highest Pearson correlation coefficient with the system output (e.g., productivity), and that have the lowest Pearson correlation coefficient with the other factors of the subset. For developing FRBS, Ahmad and Pedrycz (2011) proposed the use of wrapper methods for feature selection. Wrapper methods are based on evolutionary search methods (e.g., genetic algorithms [GAs]), which search for the subset of data where the FRBS has the highest accuracy (e.g., the lowest root mean square error). Feature selection was implemented using the following two approaches: CFS was applied to soft relationships that are defined by statistically-developed mathematical equations, and the wrapper method, using GA was applied to soft relationships that are defined by data-driven FRBS.

In the third step, the qualitative FSD model was developed by identifying two types of relationships between the system variables: soft relationships and hard relationships. Soft relationships were identified based on existing knowledge about real-world systems, which was acquired through a literature review and expert judgment, as suggested by Sterman (2000). The list of the factors that influence the productivity of equipment-intensive activities (refer to Table 2) was developed using literature review, as discussed previously; thus the soft relationships between these factors and MFP were confirmed by the literature. Moreover, these soft relationships were also verified by the expert knowledge obtained through the interview surveys, as discussed earlier. On the other hand, the hard relationships between the system variables were identified using the equations, which define the relationships. Equation 1 presents an example of the hard relationship between "crew size", "planned crew size," and "absenteeism":

$$Crew\ Size = Planned\ Crew\ Size - Absenteeism \quad \textbf{(1)}$$

In the fourth step, the quantitative FSD model was developed. First, the objective and subjective system variables were identified based on their scales of measure, where objective variables were evaluated using crisp numbers (e.g., 10 years of experience) and subjective variables were evaluated using subjective scales (e.g., *high* crew motivation) (Tsehayae and Fayek 2016b). Then, objective system variables were represented by crisp numbers, and fuzzy membership functions were developed to represent the subjective system variables. These fuzzy membership functions can be developed by one of several approaches proposed in the literature that use either data or expert knowledge. Fuzzy membership functions were developed by FCM clustering, which is a machine learning technique that is commonly used for developing fuzzy membership functions using data (Pedrycz 2013). FCM clustering was also used to develop the FRBS for defining the relationships between the system variables by projecting the clusters into the input space (e.g., the values of the factors influencing productivity) and the output space (e.g., the value of productivity) (Pedrycz 2013).

Next, the soft relationships of the system were defined quantitatively. The soft relationships were defined either by data-driven FRBS developed using FCM clustering (Gerami Seresht and Fayek 2015) or by mathematical equations developed using linear regression (Nasirzadeh et al. 2014). The performance of the two methods in defining the soft relationships of the system was evaluated using the root mean square error (RMSE); then, the method with the lowest RMSE was chosen for defining each relationship. FCM clustering and linear regression methods were implemented using 90% cross validation, which uses 90% of the data for training and 10% of the data for validation (i.e., measuring RMSE). Since the mathematical form of hard relationships was known, unlike soft relationships, these relationships were defined using mathematical equations. Fuzzy arithmetic was then used to solve both the soft relationships defined using mathematical equations as well as all the hard relationships, since they both contain subjective system variables. Fuzzy arithmetic operations were implemented by the α-cut method and the extension principle method using four common *t*-norms (min, algebraic product, Lukasiewicz and drastic product).

Finally, in the fifth step, the FSD model of construction productivity was validated using a case study of earthmoving operations. Since the common validation tests such as statistical hypothesis test are not appropriate for the validation of SD (and FSD) models (Forrester and Senge 1980), Barlas (1996) introduced two approaches for validation of the SD (and FSD) models: structure validity and behavior validity. The structural validation of the FSD model presented in this paper was determined using the dimensional consistency test and the structure verification test (Barlas 1996, Qudrat-Ullah and Seong 2010). The dimensional consistency test is a simple dimensional analysis of the mathematical equations of the FSD models that is appropriate for validation of hard relationships (Forrester and Senge 1980, Qudrat-Ullah and Seong 2010). On the other hand, soft relationships of FSD models can be validated by the structure verification test (Forrester and Senge 1980, Qudrat-Ullah and Seong 2010), which compares the structure of the model with the real-world system empirically using expert knowledge or theoretically using relevant literature. The behavioral validity of the FSD model was evaluated using the pattern verification test, as suggested by Barlas (1996). The pattern verification test compares the pattern of system results (e.g., number of peaks of the simulation results, frequency) to field data.

Fuzzy System Dynamics Model of Multi-factor Productivity for Equipment-intensive Activities

Seventy-two activity-level factors influencing the productivity of equipment-intensive activities, hereafter referred to as system variables, were identified through the literature review and were verified by expert knowledge. In order to increase the accuracy of the FSD model of construction productivity, the number of system variables was reduced by feature selection, as discussed in the methodology. Twenty-five system variables, divided into six categories (i.e., crew-related factor), were selected for the development of the FSD model, which are presented below in Table 3.

Table 3. System variables for FSD model of activity-level construction productivity.

Category	Factors
Equipment-related Factors	Number of Equipment, Equipment Capacity, Equipment Ownership, Equipment Functional Range, Operator Experience, Labour and Equipment Balance
Location-related Factors	Distance, Site Restrictions, Underground Facilities, Groundwater Level, Soil Type, Soil Moisture
Weather-related Factors	Gust Speed, Temperature, Total Precipitation
Task-related Factors	Daily Overtime Work, Total Work Volume
Crew-related Factors	Crew Experience, Crew Composition, Crew Size, Crew Motivation, Absenteeism, Foreman Experience
Material-related Factors	Material Pre-Installation Requirements, Material Quality

Once the system variables were selected, the qualitative FSD model of construction productivity was developed by identifying the relationships between the variables. As discussed in the methodology section, at the qualitative FSD modeling stage, the existence of these relationships is identified only. The soft relationships between the system variables and productivity were verified by the literature and expert knowledge. Moreover, the soft relationships between the system variables were identified by the researchers based on their knowledge about the real-world system (e.g., crew motivation and absenteeism). According to previous research, the soft relationships between system variables need to be identified by the modelers based on their knowledge about the real-world system; once the SD or FSD models are validated, the soft relationships between the system variables will be verified (Nojedehi and Nasirzadeh 2017, Ding et al. 2018). For presentation clarity, the qualitative FSD model of construction productivity presented in this paper is broken into two components: a stock and flow diagram, and a cause and effect diagram. Fig. 2 presents the stock and flow diagram that measures the MFP of the system using its three inputs (i.e., labor direct cost, equipment direct cost, and material direct cost), and it measures the total cost rate and the total activity direct cost using the MFP and the production rate of the activity.

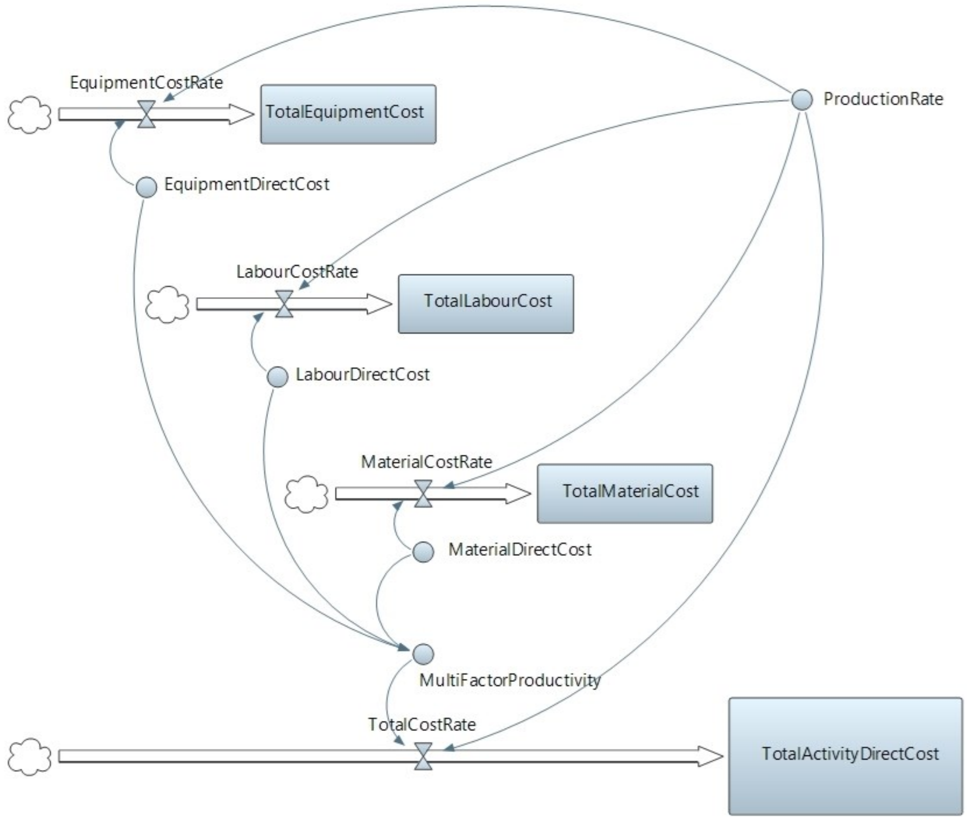

Fig. 2. Stock and flow diagram of qualitative FSD model of construction productivity.

There are four stock variables (i.e., representing accumulation in FSD models) in Fig. 2, which represent the cumulative costs of the three input resources “total equipment cost”, “total labor cost”, and “total material cost”, and the total direct cost of the activity “total activity direct cost”. There are four flow variables (i.e., representing the rate of increase/decrease in the stock variables of FSD models) in Fig. 2, which represent the daily cost of the three input resources (i.e., “equipment cost rate”, “labor cost rate”, and “material cost rate”) and the total daily direct cost of the activity (i.e., “total cost rate”). The MFP, the three inputs of MFP (i.e., “labor direct cost”, “equipment direct cost”, and “material direct cost”), and the “production rate” of the activity are presented as dynamic variables, where their values are determined by the cause and effect diagram presented in Fig. 3. In FSD models, the dynamic variables represent the variables that change in value due to their relationships with other variables. All relationships between the variables of the stock and flow diagram (represented by arrows in Fig. 2) are hard relationships. Fig. 3 presents the cause and effect diagram that measures the three inputs of MFP, and the production rate of the activity (inputs of the stock and flow diagram) using the system variables (refer to Table 3).

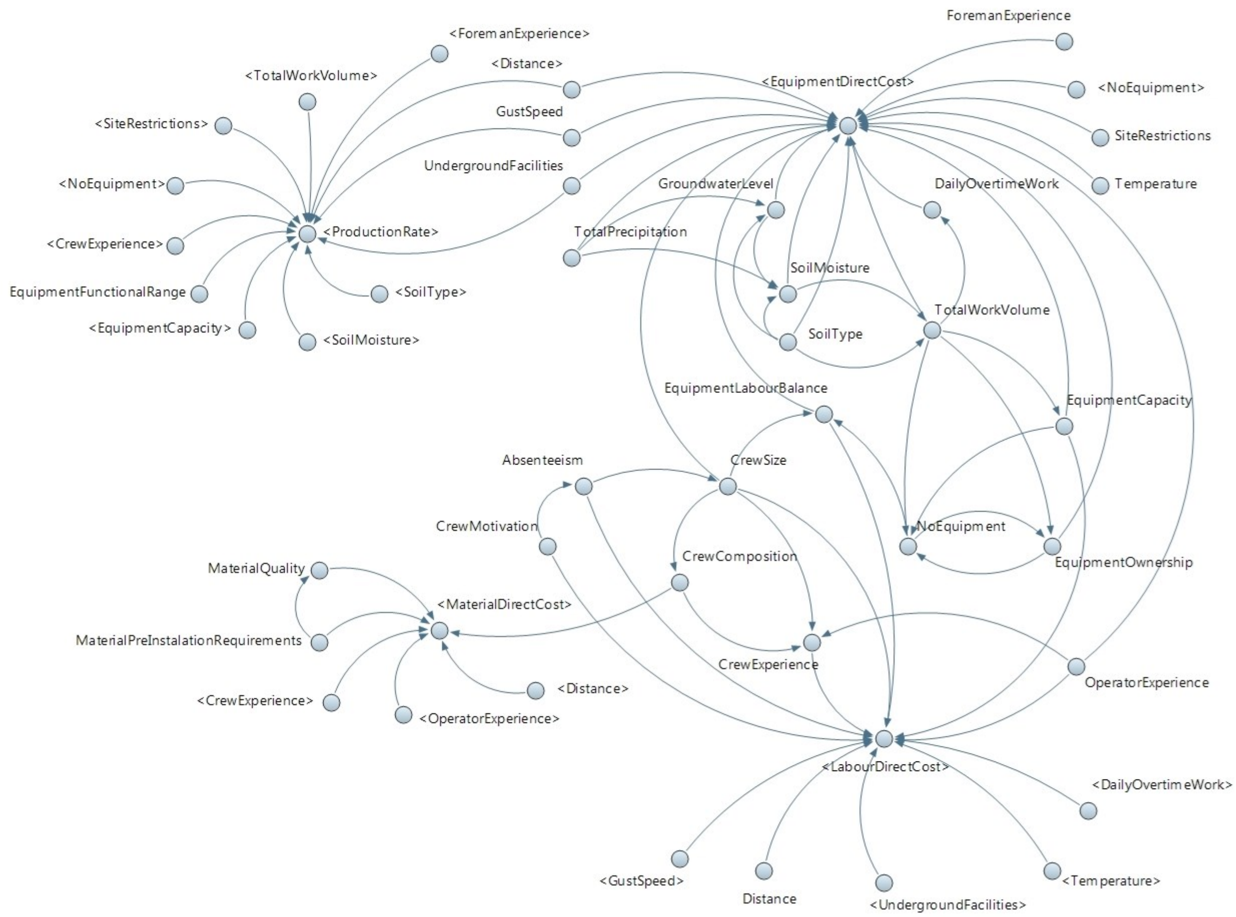

Fig. 3. Cause and effect diagram of qualitative FSD model of construction productivity.

The system variables that are selected for predicting the productivity of equipment-intensive activities (refer to Table 3) are presented in Fig. 3 as dynamic variables. These variables are used in the cause and effect diagram to predict the value of the three inputs of MFP (i.e., “labor direct cost”, “equipment direct cost”, and “material direct cost”), as well as the “production rate” of the activity. There are also two types of relationships that exist between the system variables in the cause and effect diagram: soft relationships, such as the relationship between “crew motivation” and “equipment direct cost”, and hard relationships, such as the relationship between “crew size” and “planned crew size” and “absenteeism”.

Next, in order to develop the quantitative FSD model of construction productivity, the objective and subjective system variables were identified. Referring to Table 3, there are 20 objective system variables and 5 subjective system variables. The subjective variables of the system include site restrictions, soil moisture, crew motivation, material quality, and material pre-installation requirements. Soil moisture can be also an objective system variable if it is measured numerically using soil tests; however, this factor is considered as a subjective system variable, since it may also be measured by subjective expert judgment if the test results are not available. Once the objective and subjective system variables were identified, the subjective system variables were represented by fuzzy membership functions. Each subjective variable was represented by five—as suggested by Pedrycz (2013)—triangular fuzzy membership functions, which are commonly used in engineering applications. As discussed in the methodology section, these fuzzy membership functions were developed using the FCM clustering technique, which is a data-driven technique for developing fuzzy membership functions. Fig. 4 shows the fuzzy membership functions developed for the representation of crew motivation, as an example.

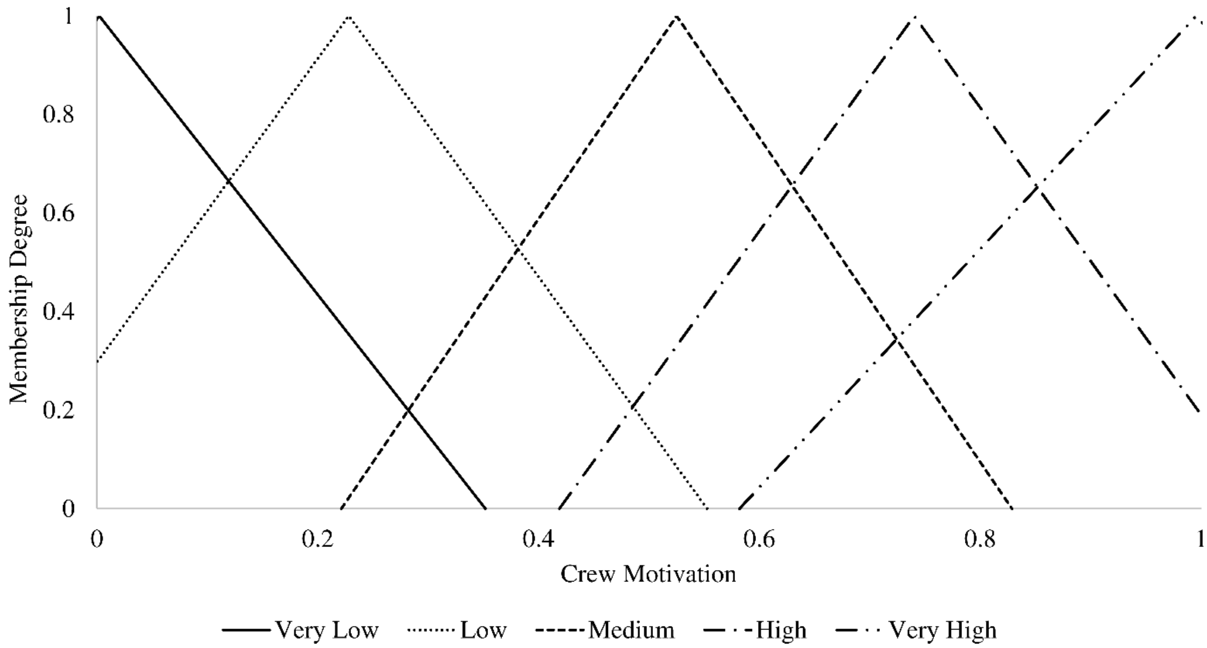

Fig. 4. Fuzzy membership functions for representing crew motivation.

Next, the soft relationships between the system variables were defined quantitatively, either by data-driven FRBS —developed by the FCM clustering technique— or mathematical equations —developed by statistical techniques—, as discussed in the methodology section. Table 4 shows these soft relationships and the approach by which each soft relationship was defined.

Table 4. Soft relationships of FSD model of activity-level construction productivity.

Relationship Output	Relationship Inputs	Numerical Definition Approach
Equipment Direct Cost	Distance, Number of Equipment, Site Restrictions, Underground Facilities, Operator Experience, Equipment Ownership, Equipment Capacity, Daily Overtime Work, Total Work Volume, Soil Type, Soil Moisture, Groundwater Level, Total Precipitation, Temperature, Gust Speed, Foreman Experience, Labour and Equipment Balance, Crew Size	Linear Regression
Labor Direct Cost	Crew Motivation, Crew Size, Crew Experience, Absenteeism, Gust Speed, Distance, Underground Facilities, Temperature, Daily Overtime Work, Operator Experience, Equipment Capacity, Labour and Equipment Balance	Linear Regression
Material Direct Cost	Material Quality, Material Pre-Installation Requirements, Crew Experience, Crew Composition, Operator Experience, Distance	Linear Regression
Production Rate	Site Restrictions, Number of Equipment, Equipment Functional Range, Equipment Capacity, Soil moisture, soil Type, Gust Speed	Linear Regression

Number of Equipment	Equipment Ownership, Equipment Capacity, Total Volume of Work	FCM Clustering
Equipment Capacity	Total Volume of Work	FCM Clustering
Equipment Ownership	Number of Equipment, Total Volume of Work	FCM Clustering
Groundwater Level	Total Precipitation	FCM Clustering
Soil Moisture	Total Precipitation, Soil Type, Groundwater Level	FCM Clustering
Daily Overtime Work	Total Volume of Work	FCM Clustering
Total Work Volume	Soil Moisture, Soil Type	FCM Clustering
Crew Experience	Crew Size, Crew Composition, Operator Experience	FCM Clustering
Crew Composition	Crew Size	FCM Clustering
Absenteeism	Crew Motivation	FCM Clustering
Material Quality	Material Pre-Installation Requirements	FCM Clustering

As presented in Table 4, 11 soft relationships in the FSD model were defined by FRBS, and four of those relationships were defined by statistically-developed mathematical equations. Accordingly, in some cases, defining the soft relationships of FSD models using data-driven FRBS developed by FCM clustering can increase the accuracy of FSD models compared to using statistically-developed mathematical equations. However, neither of the two methods is universally the best approach for defining the soft relationships of the system. In order to simulate the FSD model and predict the productivity of any given equipment-intensive activity, the soft relationships of the system (presented in Table 4) were evaluated at each time step (i.e., daily). Once the soft relationships were defined, the hard relationships were defined quantitatively using mathematical equations, as discussed in the methodology. There are nine hard relationships in the FSD model, which were defined by the mathematical equations presented in Table 5.

In order to simulate the FSD model and predict the productivity of any given equipment-intensive activity, the mathematical equations presented in Table 5 were solved at each time step (i.e., daily).

Table 5. Hard relationships of FSD model of activity-level construction productivity.

Relationship Output	Mathematical Equation
Labor Cost Rate	$\text{Labor Cost Rate}\left(\frac{\$}{\text{day}}\right) = \text{Labor Direct Cost}\left(\frac{\$}{\text{units}}\right) \times \text{Production Rate}\left(\frac{\text{units}}{\text{day}}\right)$
Equipment Cost Rate	$\text{Equipment Cost Rate}\left(\frac{\$}{\text{day}}\right) = \text{Equipment Direct Cost}\left(\frac{\$}{\text{units}}\right) \times \text{Production Rate}\left(\frac{\text{units}}{\text{day}}\right)$
Material Cost Rate	$\text{Material Cost Rate}\left(\frac{\$}{\text{day}}\right) = \text{Material Direct Cost}\left(\frac{\$}{\text{units}}\right) \times \text{Production Rate}\left(\frac{\text{units}}{\text{day}}\right)$
Total Labor Cost*	$\text{Total Labor Cost (\$)} = \int \text{Labor Cost Rate}\left(\frac{\$}{\text{day}}\right).dt\ (\text{day})$
Total Equipment Cost*	$\text{Total Equipment Cost (\$)} = \int \text{Equipment Cost Rate}\left(\frac{\$}{\text{day}}\right).dt\ (\text{day})$
Total Material Cost*	$\text{Total Material Cost (\$)} = \int \text{Material Cost Rate}\left(\frac{\$}{\text{day}}\right).dt\ (\text{day})$
Multi Factor Productivity	$\text{Multi Factor Productivity}\left(\frac{\$}{\text{units}}\right) = \text{Labor Direct Cost}\left(\frac{\$}{\text{units}}\right) + \text{Equipment Direct Cost}\left(\frac{\$}{\text{units}}\right) + \text{Material Direct Cost}\left(\frac{\$}{\text{units}}\right)$
Labor and Equipment Balance**	$\text{Labor and Equipment Balance} = \frac{\text{Crew Size (Person)}}{\text{Number of Equipment (Count)}}$
Crew Size***	$\text{Crew Size (Person)} = \text{Planned Crew Size (Person)} - \text{Absenteeism(Person)}$

* dt stands for the time step's duration used for simulation of FSD model that is equal to one day in this paper.
** Number of equipment represents the number of equipment, which are working on the activity.
*** Planned crew size stands for the crew size that is specified for execution of the activity in planning phase, and absenteeism represent the number of absent crew members.

Model Validation and Construction Application

The FSD model of construction productivity was developed by integrating AnyLogic®, Matlab® software, and a Fuzzy Calculator class, which was developed in the Python programming language. AnyLogic® was used to develop the SD component of the model; and Matlab® and the Fuzzy Calculator class were used to develop the fuzzy components of the model. AnyLogic® calculates the results of the mathematical equations, in which all system variables are objective. The Fuzzy Calculator class calculates the results of the mathematical equations that include subjective system variables, and Matlab® calculates the results of the FRBS. The Fuzzy Calculator class was developed by the authors for implementing fuzzy arithmetic on triangular fuzzy numbers using the α-cut method and the extension principle method, the latter of which uses min, algebraic product, Lukasiewicz, and drastic product *t*-norms.

The FSD model was evaluated through structural and behavioral validation tests, as discussed in the methodology section. The structural validity of the FSD model was evaluated using the dimensional consistency test and the structure verification test. The dimensional consistency test is implemented by dimensional analysis of the mathematical equations, which defines the hard relationships of the system. Referring to Table 5, the dimensional consistency test determines if the units of measure on both sides of each equation are consistent or not. For example, in Equation 2, the unit of measure for the left side of the equation is $\left(\frac{\$}{\text{day}}\right)$, and the unit of measure for the right side of the equation is $\left(\frac{\$}{\text{units}}\right) \times \left(\frac{\text{units}}{\text{day}}\right) = \left(\frac{\$}{\text{day}}\right)$, which shows that Equation 2 has dimensional consistency.

$$\mathit{Labor\ Cost\ Rate}\ \left(\frac{\$}{\text{day}}\right) = \mathit{Labor\ Direct\ Cost}\ \left(\frac{\$}{\text{units}}\right) \times \mathit{Production\ Rate}\ \left(\frac{\text{units}}{\text{day}}\right) \quad \textbf{(2)}$$

The structure verification test was implemented by verifying the list of the system variables (i.e., factors influencing construction productivity) and the soft relationships of the system through expert knowledge, which was acquired by the interview surveys, as discussed in the methodology section. In order to evaluate the behavior validity of the FSD model, the model was implemented on a case study of earthmoving operations on a pipeline maintenance project in Alberta, Canada. This project included 79 work packages (i.e., digs), each of which includes the following activities: excavation, sandblasting, welding, coating, and backfilling. The case study presented in this paper is focused on the earthmoving activities (i.e., excavation and backfilling), which are executed by eight earthmoving crews. Field data were collected for these two equipment-intensive activities, excavation and backfilling, by documenting the value of the factors that influence construction productivity. Field data were also collected for the actual activity-level MFP of the two activities, measured in $\frac{\$}{m^3}$, using the daily costs of the input resources (i.e., labor, equipment, and material) measured in dollars and the daily quantity of work completed measured in cubic meters (i.e., volume of earth excavated or backfilled). Various sources were used for field data collection, including contract documents, project scorecards, project timesheets, and onsite observations by the researchers. Due to confidentiality constraints, all field data were normalized into the range of [0,1] using Equation 3.

$$V_{i,normalized} = \frac{V_i - \min(V_i)}{max(V_i) - min(V_i)}, \quad \textbf{(3)}$$

where $V_{i,normalized}$ stands for the normalized value of any system variable and V_i represents the original value of the system variable. In order to run the simulation model, the initial values of the system variables are entered, where the values of the objective system variables are entered as crisp numbers (e.g., four people for crew size), and the values of the subjective system variables are entered as linguistic terms, which are represented by fuzzy membership functions (e.g., *high* crew motivation). Table 6 presents the results of simulation for the MFP for earthmoving operations in a 30-day period and compares the results to the actual field data; Fig. 5 presents these results graphically.

Table 6. Simulation results and actual field data for MFP.

Simulation Time (day)	Simulation Results	Actual Field Data	Error $\lvert \text{simulation result} - \text{actual Field data} \rvert$
1	0.321	0.365	0.044
2	0.552	0.582	0.03
3	0.858	0.775	0.083
4	0.949	0.978	0.029
5	0.738	0.749	0.011
6	0.911	0.978	0.067
7	0.798	0.775	0.023
8	0.714	0.500	0.214

9	0.692	0.775	0.083
10	0.320	0.206	0.114
11	0.273	0.146	0.127
12	0.824	0.929	0.105
13	0.810	0.765	0.045
14	0.633	0.765	0.132
15	0.933	0.929	0.004
16	0.857	0.765	0.092
17	0.540	0.765	0.225
18	0.000	0.054	0.054
19	0.234	0.039	0.195
20	0.744	0.926	0.182
21	0.873	0.926	0.053
22	0.873	0.912	0.039
23	0.988	0.912	0.076
24	0.942	0.912	0.03
25	0.551	0.504	0.047
26	0.630	0.450	0.18
27	0.823	1.000	0.177
28	0.949	1.000	0.051
29	0.898	1.000	0.102
30	0.903	0.894	0.009

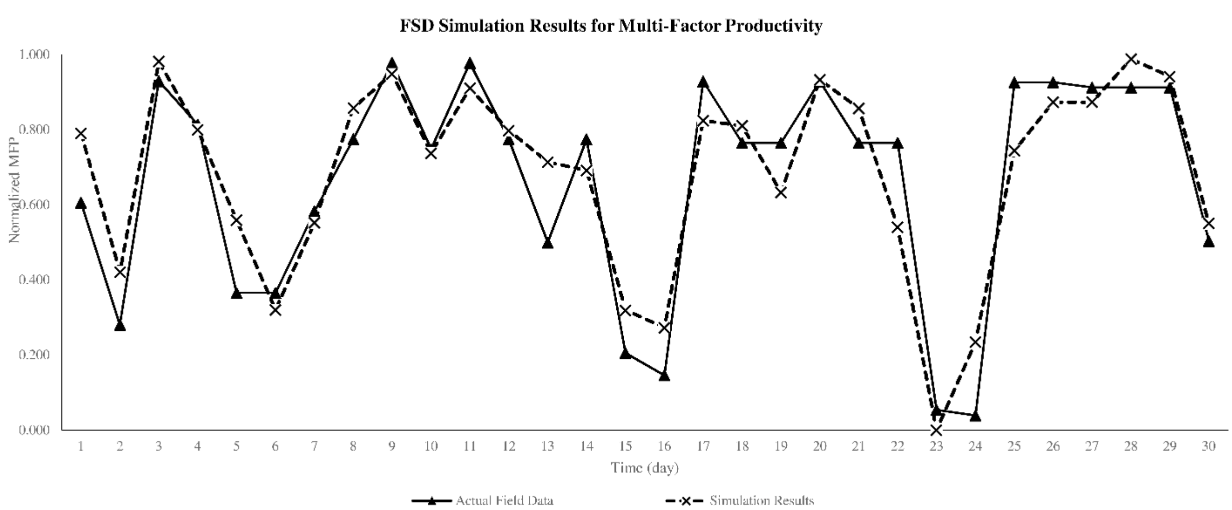

Fig. 5. Simulation results for MFP in comparison to actual field data.

The *y*-axis in Fig. 5 shows the normalized value of the MFP of the earthmoving operations, and the *x*-axis shows the duration of earthmoving operations measured in days. The simulation results can be presented as fuzzy numbers or defuzzified values. Defuzzification is the process of converting a fuzzy number to a crisp number. In order to present the simulation results as fuzzy numbers, the results need to be presented at each time step. Representing the simulation results as fuzzy numbers is not appropriate for the pattern verification test, since this test compares changes in the results over the simulation time to the actual field data. The simulation results presented in Fig. 5 are the defuzzified values of MFP for the earthmoving operations, which are defuzzified using the using center of area method (COA). Referring to Fig. 5, behavioral validity of the FSD model may be evaluated by the pattern verification test, which shows the following: the trends in the actual MFP values (i.e., an increase or decrease of productivity between any two consecutive points) are predicted correctly by the simulation results in 70% of cases (refer to Table 6); and the turning points in the actual MFP values (i.e., the points in which the trend of productivity changes) are predicted correctly by the simulation results in 70% of cases (refer to Table 6). Finally, the RMSE of the simulation results is 0.11, which is calculated using Equation 4.

$$\mathrm{RMSE} = \sqrt{\frac{\sum(\text{simulation result} - \text{actual field data})^2}{\mathrm{n}}}. \quad \textbf{(4)}$$

In addition, the normalized root mean square error (NRMSE) of the simulation results is 15%. The NRMSE compares the RMSE of the data to the average value of the actual field data using Equation 5.

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{\text{Mean(actual field data)}}. \quad \textbf{(5)}$$

Kleijnen (1995) introduced regression analysis of SD [or FSD] models as an appropriate approach for identification of the most significant factors in SD [or FSD] models. In this approach, the value of independent system variables (i.e., system variables that are not affected by any other system variables) are changed between the minimum and maximum values (i.e., [0,1] in this case study) and the effect of these factors on the simulation results is analyzed using regression analysis (Kleijnen 1995, Phan et al. 2018). Next, the significance of the factors' influence on the FSD model is identified based on the regression coefficient, where the system variable with the highest absolute value of regression coefficient has the most significant effect on the FSD model. The FSD model presented in this paper has 12 independent system variables (refer to Fig. 2 and Fig. 3): gust speed, total precipitation, temperature, soil type, underground facilities, site restrictions, distance, equipment operator experience, foreman experience, crew motivation, equipment functional range, and material pre-installation requirements. The regression analysis approach was implemented on these independent system variables, and the results of the analysis show that the most significant variables in the FSD model are: (1) crew motivation, which has a negative correlation with the simulation results; (2) equipment operator experience, which has a positive correlation with the simulation results; and, (3) gust speed, which has a positive correlation with the simulation results.

By implementing fuzzy arithmetic operations on the mathematical equations of the FSD model, the support of the resulting fuzzy numbers grows rapidly, which is interpreted as an overestimation of uncertainty. In general, an increase in the length of the support of a fuzzy number shows an increase in the amount of uncertainty represented by that fuzzy number. The overestimation of uncertainty in FSD models is affected by the chosen fuzzy arithmetic implementation method, which is used to solve the mathematical equations of the FSD model. Accordingly, the effect of fuzzy arithmetic implementation methods on the simulation results were evaluated to determine the most appropriate method. The results of the simulation for the "Total Cost Rate" of the activity were calculated using the α-cut method and using the extension principle method with the min, algebraic product, Lukasiewicz, and drastic product *t*-norms, as presented in Table 7.

Table 7. Simulation results and actual field data representing fuzzy number for total cost rate.

Sim. Time	***Min t-norm***		***Algebraic Product t-norm***		***Lukasiewicz t-norm***		***Drastic Product t-norm***		**Actual Field Data**
	Sim. Results *	**Support Length**	**Sim. Results***	**Support Length**	**Sim. Results***	**Support Length**	**Sim. Results***	**Support Length**	
1	0.069	0.154	0.064	0.154	0.062	0.125	0.062	0.125	0.049
2	0.254	0.295	0.249	0.295	0.247	0.232	0.247	0.231	0.261
3	0.075	0.181	0.070	0.181	0.070	0.159	0.069	0.159	0.027
4	0.069	0.154	0.064	0.154	0.062	0.125	0.062	0.125	0.029
5	0.089	0.207	0.084	0.207	0.083	0.184	0.083	0.184	0.027
6	0.382	0.321	0.379	0.321	0.375	0.217	0.375	0.217	0.417
7	0.023	0.070	0.019	0.070	0.018	0.055	0.018	0.055	0.050
8	0.043	0.130	0.039	0.130	0.038	0.115	0.038	0.115	0.054
9	0.165	0.207	0.162	0.207	0.158	0.139	0.158	0.138	0.074
10	0.184	0.222	0.180	0.222	0.177	0.153	0.177	0.152	0.089
11	0.333	0.307	0.329	0.307	0.326	0.217	0.326	0.217	0.424
12	0.154	0.234	0.149	0.234	0.146	0.186	0.146	0.186	0.127
13	0.147	0.216	0.142	0.216	0.139	0.165	0.139	0.165	0.120
14	0.165	0.249	0.160	0.249	0.158	0.202	0.158	0.202	0.134
15	0.177	0.242	0.173	0.242	0.170	0.187	0.170	0.187	0.140
16	0.203	0.249	0.198	0.249	0.195	0.187	0.195	0.187	0.155
17	0.203	0.249	0.198	0.249	0.195	0.187	0.195	0.187	0.155
18	0.206	0.250	0.201	0.250	0.198	0.186	0.198	0.186	0.149
19	0.208	0.254	0.204	0.254	0.201	0.192	0.201	0.191	0.144
20	0.222	0.245	0.218	0.245	0.215	0.169	0.215	0.169	0.156
21	0.205	0.218	0.202	0.218	0.199	0.113	0.198	0.110	0.138
22	0.203	0.249	0.198	0.249	0.195	0.187	0.195	0.187	0.120
23	0.629	0.393	0.626	0.393	0.622	0.235	0.621	0.233	0.544
24	0.259	0.277	0.255	0.277	0.252	0.201	0.252	0.202	0.174
25	0.280	0.275	0.276	0.275	0.273	0.188	0.273	0.188	0.183
26	0.238	0.261	0.234	0.261	0.231	0.187	0.231	0.186	0.132
27	0.069	0.173	0.064	0.173	0.064	0.153	0.064	0.153	0.381
28	0.294	0.285	0.290	0.285	0.287	0.198	0.286	0.198	0.183
29	0.254	0.295	0.249	0.295	0.247	0.232	0.246	0.231	0.120
30	0.320	0.315	0.316	0.315	0.313	0.236	0.313	0.236	0.173
RMSE	**0.0915**		**0.0898**		**0.0884**		**0.0883**		**-**

* **Sim. Results** stands for the defuzzified value of the simulation results using the *t-norm* that is presented in the first row.

The simulation results presented in Table 7 show the following: the implementation of fuzzy arithmetic operations using the α-cut method and using the extension principle method with the min *t*-norm always return the same results (Elbarkouky et al. 2016); using the α-cut method and the extension principle method with the min *t*-norm return the largest defuzzified values of the simulation results, followed by the extension principle method with the algebraic product *t*-norm, Lukasiewicz *t*-norm, and drastic product *t*-norm, respectively; and finally, using the extension principle method with the drastic product *t*-norm has the lowest RMSE, followed by the extension principle method with the Lukasiewicz *t*-norm, algebraic product *t*-norm, and min *t*-norm (and the α-cut method), respectively. In order to compare the uncertainty overestimation caused by the fuzzy arithmetic implementation methods, the length of the support of the fuzzy number for "Total Cost Rate" is presented in Table 7, and it is shown graphically in Fig. 6. The length of the support of the fuzzy number for "Total Cost Rate" represents the level of uncertainty overestimation.

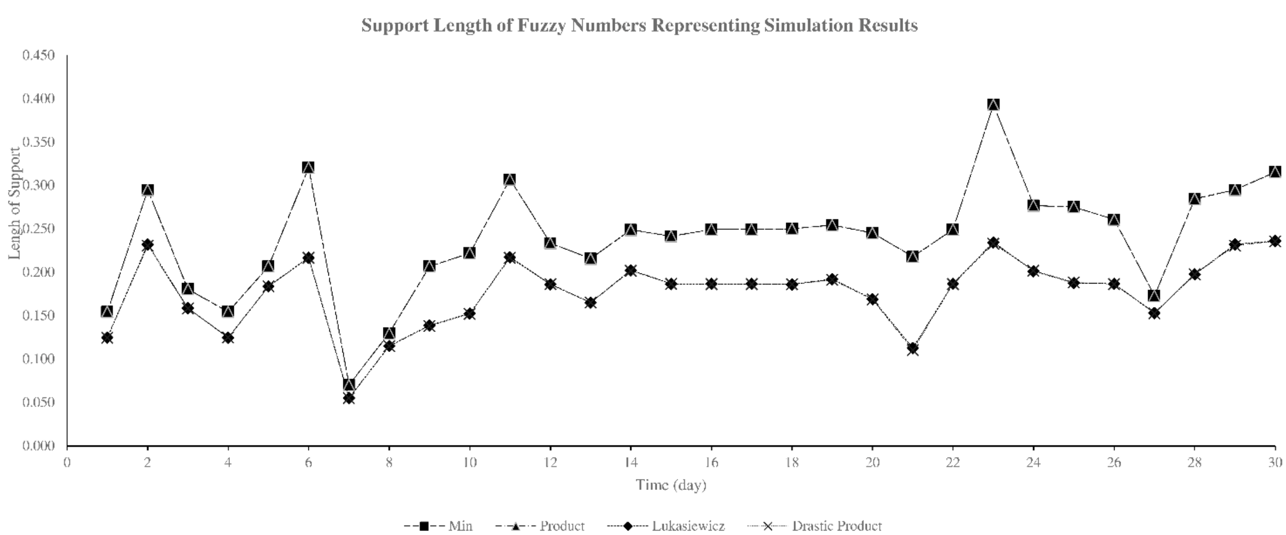

Fig. 6. Length of support of fuzzy numbers for total cost rate.

Referring to Table 7 and Fig. 6, a comparison of the length of the support of the fuzzy number for "Total Cost Rate" shows the following: the length of the support of the fuzzy number is always equal when using the α-cut method and when using the extension principle method with the min and algebraic product t-norms*;* and using the extension principle method with the drastic product t-norm returns a fuzzy number with the smallest length of the support, followed by the extension principle with the Lukasiewicz t-norm; and the other methods (i.e., using the α-cut method, using the extension principle method with the min and algebraic product t-norms) return a fuzzy number with the largest support length. Based on the fact that the extension principle method using the drastic product t-norm has both the lowest RMSE and the smallest uncertainty overestimation, this method was deemed to be the most appropriate method for fuzzy arithmetic implementation in the FSD model presented in this paper.

Discussion

The FSD model of construction productivity presented in this paper can be used to predict the MFP of equipment-intensive activities for construction projects. Accordingly, the FSD model can facilitate the construction planning process by allowing users to predict the productivity of construction activities for different execution plans prior to the execution phase. Users can change the system variables based on their execution plans (e.g., changing the crew size or number of equipment) and simulate the model to predict the productivity, and accordingly, they can select the most appropriate execution plan for the activity. The FSD model of productivity can predict the daily value of MFP, which provides more information about productivity, as compared to existing static productivity models, by allowing users to track changes in productivity over time. Moreover, this model allows construction planners to analyze the effect of each system variable (e.g., number of equipment) on construction productivity in order to optimize these variables. For the purpose of analysis, the system variable that is being analyzed must first be changed in the desirable range, while the other system variables are kept unchanged; once this is accomplished, the FSD model can then be simulated. Accordingly, the results of simulation represent the effect of the system variables that were changed in step 1 on construction productivity.

The FSD model of construction productivity presented in this paper is capable of capturing the probabilistic and non-probabilistic uncertainties of the system variables, as well as the deterministic values for the system variables. In order to capture these probabilistic uncertainties, the model allows users to represent variables with probabilistic distributions, such as the temperature in future projects. For capturing the non-probabilistic uncertainties of the system variables, the model allows users to determine the values of the subjective system variables using linguistic terms, which are represented by fuzzy membership functions, such as *high* crew motivation (refer to Fig. 4). Due to the fact that the case study presented in this paper was extracted from a previously executed construction project, the system variables do not exhibit any probabilistic uncertainty; accordingly, in the case study presented in this paper, the system variables are represented by either deterministic values or by fuzzy membership functions.

In comparison to the SD models of productivity developed by Nasirzadeh and Nojedehi (2013) and Mawdesley and Al-Jiboury (2009), the FSD model of productivity presented in this paper can increase the accuracy of productivity predictions by capturing the effect of subjective variables (e.g., crew motivation) on productivity, as well as allowing practitioners to evaluate these variables using linguistic terms rather than precise numerical values. In contrast to the FSD model developed by Nojedehi and Nasirzadeh (2017), which is for labor-intensive activities and predicting CLP, the predictive model presented in this paper predicts MFP, which is the appropriate measure of productivity for equipment-intensive activities. Moreover, the predictive model presented in this paper provides construction practitioners with information regarding the cost of the three input resources of an activity (equipment cost, labor cost, and material cost), while the predictive models of CLP provide this information for one input resource only (i.e., labor). Finally, the comparison of the two fuzzy arithmetic implementation methods (i.e., the α-cut method and the extension principle method) shows that the implementation of fuzzy arithmetic operations by the extension principle using drastic product *t*-norm reduces the overestimation of uncertainty in comparison to the α-cut method, while increasing the accuracy of the simulation results, in contrast to previously developed FSD models (e.g., Nojedehi and Nasirzadeh 2017, Khanzadi et al. 2012), which only employ the α-cut method. Reducing the uncertainty overestimation of the simulation results increases the ability of construction practitioners to accurately predict the actual productivity of an activity based on the simulation results.

The FSD model presented in this paper has a few limitations, which need to be addressed in future research. First, the computational approach used for implementing fuzzy arithmetic operations is only applicable to triangular fuzzy numbers; thus the FSD model is limited to the use of triangular fuzzy numbers for representing subjective system variables. Next, for defining the soft relationships of the FSD model, the accuracy of FCM clustering technique decreases as the number of input variables increases (i.e., high dimensionality of soft relationships). Accordingly, in this paper, for defining high dimensional soft relationships, statistically developed mathematical equations outperformed the FRBSs developed by FCM clustering in terms of accuracy. In the future, the accuracy of the FCM clustering technique for defining high dimensional soft relationships can be increased by developing a method to increase the weights of the output variables in comparison to the input variables. Finally, the FSD model of MFP presented in this paper has been developed using field data collected for earthmoving activities. In order to develop a generic model of MFP for different types of equipment-intensive activities, new field data for other types of equipment-intensive activities need to be collected, and the FSD model needs to be updated with the new field data.

Conclusions and Future Research

Construction productivity has long been a major research interest within the construction engineering domain. Due to the fact that construction is a labor-intensive industry, the majority of previous studies have been focused on construction labor productivity (CLP). However, with recent advancement in technology, construction equipment are now the main drivers of productivity for some construction activities, which are identified as equipment-intensive activities. Since the main driver of productivity for equipment-intensive and labor intensive activities are different, the factors influencing the productivity of these two activities are also different. Accordingly, the predictive models that have been developed for labor-intensive activities cannot predict the productivity of equipment-intensive activities accurately. This paper presents the list of 72 factors that influence the productivity of equipment-intensive activities identified through a literature review and verified by expert knowledge collected through interview surveys. It presents a predictive model of productivity for equipment-intensive activities using the FSD modeling technique. The FSD model presented in this paper predicts the MFP of equipment-intensive activities considering three inputs resources of these activities (i.e., labor, equipment, and material).

In this model, the subjective factors influencing construction productivity (e.g., crew motivation) are represented by fuzzy membership functions. Representation of subjective factors by fuzzy membership functions enhances the applicability of the predictive model by allowing practitioners to evaluate the value of subjective variables using linguistic terms (e.g., *high* crew motivation), rather than numerical values. Moreover, in this paper, the accuracy of FCM clustering and linear regression methods were compared for defining the soft relationships of the FSD model. Although the FCM clustering method has not been used for defining soft relationships in previous research, the results of comparison show that, in some cases, the use of the FCM clustering method can increase the accuracy of FSD models, as compared to the use of the linear regression method. However, neither of the two methods is universally the best method for defining the soft relationships of the system. Previous applications of the FSD modeling technique in construction show that the use of fuzzy arithmetic operations for solving the mathematical equations of the FSD model can cause overestimation of uncertainty in the fuzzy numbers representing the simulation results. In this paper, the two methods of fuzzy arithmetic implementation (i.e., the α-cut method and the extension principle method using min, algebraic product, Lukasiewicz, and drastic product *t*-norms) were evaluated in order to reduce the overestimation of uncertainty and increase the accuracy of the FSD model. Accordingly, the extension principle method using the drastic product *t*-norm was found to be the most appropriate method for implementing fuzzy arithmetic in the FSD model, since it has the highest accuracy in calculating the simulation results and the lowest level of uncertainty overestimation.

This paper contributes to construction productivity research by identifying the key factors that influence the productivity of equipment-intensive activities, and developing a predictive model of MFP for equipment-intensive activities using FSD technique. The MFP model presented in this paper provides practitioners with information regarding the cost of the three input resources of an activity, in contrast to existing models that predict CLP, which provide information for only one input resource. This paper also contributes to the application of FSD technique in construction research by developing an approach to reduce the uncertainty overestimation in the simulation results of FSD models. Reducing the uncertainty overestimation in the simulation results increases the ability of practitioners to accurately evaluate the actual system output (e.g., actual productivity) based on the simulation results.

In the future, this study will be extended by developing an FSD model for the activity-level MFP of labor-intensive activities. Moreover, the FSD model of project-level MFP will be developed as an integration of the two activity-level FSD models of MFP, equipment-intensive and labor-intensive, as well as by including the factors influencing construction productivity at the project level. The soft relationships of the FSD model were defined either by mathematical equations developed using linear regression or by FRBS developed using FCM clustering, the latter of which is a machine learning technique. In order to increase the accuracy of the FSD models in future studies, other machine learning techniques, such neuro-fuzzy systems and ANNs, will be evaluated for the purpose of defining the soft relationships between the system variables.

Data Availability Statement

All data generated or analyzed during the study are included in the submitted article or supplemental materials files.

Acknowledgements

This research is funded by the Natural Sciences and Engineering Research Council of Canada Industrial Research Chair in Strategic Construction Modeling and Delivery (NSERC IRCPJ 428226–15), which is held by Dr. A. Robinson Fayek. The authors gratefully acknowledge the support and data provided by the industry partners, construction companies, and all personnel who participated in this study. The authors also thank Mr. Ming kit Yau, who diligently helped in the data entry process.

References

Ahmad, S. S. S., and Pedrycz, W. (2011). "Feature and Instance selection via cooperative PSO." *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on,* IEEE, 2127-2132.

Alzraiee, H., Zayed, T., and Moselhi, O. (2015). "Dynamic planning of construction activities using hybrid simulation." *Automat. Constr.,* 49 176-192.

Barlas, Y. (1996). "Formal aspects of model validity and validation in system dynamics." *Syst. Dynam. Rev.,* 12(3), 183-210.

Carson, C., and Abbott, M. (2012). "A review of productivity analysis of the New Zealand construction industry." *Construction Economics and Building,* 12(3), 1-15.

Chang, P., Pai, P., Lin, K., and Wu, M. (2006). "Applying fuzzy arithmetic to the system dynamics for the customer–producer–employment model." *Int. J. Syst. Sci.,* 37(10), 673-698.

Choi, J., and Ryu, H. (2015). "Statistical analysis of construction productivity for highway pavement operations." *KSCE J. Civ. Eng.,* 19(5), 1193-1202.

Coyle, G. (2000). "Qualitative and quantitative modelling in system dynamics: some research questions." *Sys. Dynam. Rev.,*16 (3), 225.

Dai, J. (2006). “A latent analysis and prototype system to manage jobsite factors impacting construction labor productivity.” Ph.D. thesis, Univ. of Kentucky, Kentucky, KY.

Ding, Z., Zhu, M., Tam, V. W., Yi, G., & Tran, C. N. (2018). “A system dynamics-based environmental benefit assessment model of construction waste reduction management at the design and construction stages.” *J. Clean. Prod.*, 176, 676-692.

Eastman, C. M., and Sacks, R. (2008). "Relative productivity in the AEC industries in the United States for on-site and off-site activities." *J. Constr. Eng. Manage.,* 134(7), 517-526.

Elbarkouky, M. M., Fayek, A. Robinson, Siraj, N. B., and Sadeghi, N. (2016). "Fuzzy arithmetic risk analysis approach to determine construction project contingency." *J. Constr. Eng. Manage.,* 142(12), 04016070Forrester, J. W. (1961). "Industry dynamics." *Cambridge, MA, USA*.

Gerami Seresht, N., and Fayek, A. Robinson. (2015). "Application of fuzzy logic integrated with system dynamics in construction modeling. *Proc., of the 5th International Construction Specialty Conference of the Canadian Society for Civil Engineering,* 050-01-050-10.

Goodrum, P. M., and Haas, C. T. (2004). "Long-term impact of equipment technology on labor productivity in the US construction industry at the activity level." *J. Constr. Eng. Manage.,* 130(1), 124-133.

Goodrum, P. M., Haas, C. T., Caldas, C., Zhai, D., Yeiser, J., and Homm, D. (2010). "Model to predict the impact of a technology on construction productivity." *J. Constr. Eng. Manage.,* 137(9), 678-688.

Hall, M. A. (1998). “Correlation-based feature selection for machine learning.” Ph.D. thesis, Univ. of Waikato, Hamilton, New Zealand.

Heravi, G., and Eslamdoost, E. (2015). "Applying Artificial Neural Networks for Measuring and Predicting Construction-Labor Productivity." *J. Constr. Eng. Manage.,* 141(10), 04015032.

Jabri, A., and Zayed, T. (2017). "Agent-based modeling and simulation of earthmoving operations." *Automat. Constr.,* 81 210-223.

Jarkas, A. M. (2010). "Critical investigation into the applicability of the learning curve theory to rebar fixing labor productivity." *J. Constr. Eng. Manage.,* 136(12), 1279-1288.

Kannan, G. (2011). "Field studies in construction equipment economics and productivity." *J. Constr. Eng. Manage.,* 137(10), 823-828.

Khanzadi, M., Nasirzadeh, F., and Alipour, M. (2012). "Integrating system dynamics and fuzzy logic modeling to determine concession period in BOT projects." *Automat. Constr.,* 22 368-376.

Levary, R. R. (1990). "Systems dynamics with fuzzy logic." *Int .J. Syst. Sci.,* 21(8), 1701-1707.

Lin, K., Wen, W., Chou, C., Jen, C., and Hung, K. (2011). "Applying fuzzy GERT with approximate fuzzy arithmetic based on the weakest t-norm operations to evaluate repairable reliability." *Appl. Math. Model.,* 35(11), 5314-5325.

Loosemore, M. (2014). "Improving construction productivity: a subcontractor's perspective." *Engineering, Construction and Architectural Management,* 21(3), 245-260.

Mawdesley, M. J., and Al-Jibouri, S. (2009). "Modelling construction project productivity using systems dynamics approach." *International Journal of Productivity and Performance Management,* 59(1), 18-36.

Mirahadi, F., and Zayed, T. (2016). "Simulation-based construction productivity forecast using neural-network-driven fuzzy reasoning." *Automat. Constr.,* 65 102-115.

Mortaheb, M., Ruwanpura, J., Dehghan, R., and Khoramshahi, F. (2007) "Major factors influencing construction productivity in industrial congested sites." *Proceedings of the Annual Conference of the Canadian Society for Civil Engineering,* Canadian Society for Civil Engineering Yellowknife, 1111-1120.

Naoum, S. G. (2016). "Factors influencing labor productivity on construction sites. A state-of-the-art literature review and a survey." *International Journal of Productivity and Performance Management,* 65(3).

Nasirzadeh, F., and Nojedehi, P. (2013). "Dynamic modeling of labor productivity in construction projects." *Int. J. Project Manage.,* 31(6), 903-911.

Nasirzadeh, F., Khanzadi, M., Afshar, A., and Howick, S. (2013). "Modeling quality management in construction projects." *Int. J. Civ. Eng.,* 11(1), 14-22.

Nasirzadeh, F., Khanzadi, M., and Rezaie, M. (2014). "Dynamic modeling of the quantitative risk allocation in construction projects." *Int. J. Project Manage.,* 32(3), 442-451.

Nojedehi, P., and Nasirzadeh, F. (2017). "A hybrid simulation approach to model and improve construction labor productivity." *KSCE Journal of Civil Engineering*, 21(5), 1516-1524.

Ok, S. C., and Sinha, S. K. (2006). "Construction equipment productivity estimation using artificial neural network model." *Constr. Manage. Econ.,* 24(10), 1029-1044.

Pedrycz, W. (2013). Granular computing: analysis and design of intelligent systems. CRC press, Boca Raton, FL, USA.

Pedrycz, W., and Gomide, F. (2007). *Fuzzy systems engineering: toward human-centric computing.* John Wiley & Sons.

Qudrat-Ullah, H., & Seong, B. S. (2010). "How to do structural validity of a system dynamics type simulation model: the case of an energy policy model." *Energy policy*, 38(5), 2216-2224.

Sadeghpour, F., and Andayesh, M. (2015). "The constructs of site layout modeling: an overview." *Can. J. Civil Eng.,* 42(3), 199-212.

Senge, P. M., and Forrester, J. W. (1980). "Tests for building confidence in system dynamics models." *System Dynamics, TIMS Studies in Management Sciences*, 14 209-228.

Shaheen, A. A., Fayek, A. Robinson, & AbouRizk, S. M. (2009). "Methodology for integrating fuzzy expert systems and discrete event simulation in construction engineering." *Can. J. Civil Eng.*, 36(9), 1478-1490.

Sterman, J. D. (2000). *Business dynamics: systems thinking and modeling for a complex world.* Irwin/McGraw-Hill Boston, USA.

Talhouni, B. T. (1990). Measurement and Analysis of Construction Labour Productivity." Ph.D. thesis, Univ. of Dundee, Dundee, Scotland.

Tessem, B., and Davidsen, P. I. (1994). "Fuzzy system dynamics: an approach to vague and qualitative variables in simulation." *System Dynamics Review,* 10(1), 49-62.

Thomas, H. R., Maloney, W. F., Horner, R. M. W., Smith, G. R., Handa, V. K., and Sanders, S. R. (1990). "Modeling construction labor productivity." *J. Constr. Eng. Manage.,* 116(4), 705-726.

Tsehayae, A. A., and Fayek, A. Robinson (2014). "Identification and comparative analysis of key parameters influencing construction labour productivity in building and industrial projects." *Can. J. Civil Eng.,* 41(10), 878-891.

Tsehayae, A. A., and Fayek, A. Robinson (2016a). "Developing and Optimizing Context-Specific Fuzzy Inference System-Based Construction Labor Productivity Models." *J. Constr. Eng. Manage.,* 142(7), 04016017.

Tsehayae, A. A., and Fayek, A. Robinson (2016b). "System model for analyzing construction labour productivity." *Construction Innovation,* 16(2), 203-228.

Zadeh, L. A. (1965). "Fuzzy sets." *Information and Control,* 8(3), 338-353.

Zakeri, M., Olomolaiye, P. O., Holt, G. D., and Harris, F. C. (1996). "A survey of constraints on Iranian construction operatives' productivity." *Construction Management & Economics,* 14(5), 417-426.

Zayed, T. M., and Halpin, D. W. (2005). "Productivity and cost regression models for pile construction." *J. Constr. Eng. Manage.,* 131(7), 779-789.