THE IMPACT OF ARTIFICIAL INTELLIGENCE ON CONSTRUCTION COSTING PRACTICE

Kudirat Ayinla[1], Abdullahi Saka[2] Rafiu Seidu[3], and Upeksha Madanayake[4]

[1 & 2] *School of the Built Environment and Architecture, London South Bank University, Borough Road, London, SE1 0AA, UK*

[3] *School of Built Environment and Computing, Leeds Beckett University, Leeds, LS2 8AG, UK*

[4] *School of Construction Management and Engineering, University of Reading, Chancellors Way, Reading, RG6 6AG, UK*

Cost estimation is a crucial process in the construction sector as the efficiency of the overall project cost serves as one metric in determining project success. Prevailing traditional approach suffers from human subjectivity and bias which affect accuracy. With the development and adoption of Artificial Intelligence (AI) such as the use of machine learning (ML) and deep learning (DL) algorithms, the construction industry is experiencing brisk technological change and new ways of working, particularly in terms of cost predictions and estimations. However, the application of AI is still in its infancy and the industry still prioritises traditional cost modelling approaches in determining early estimates. This research explores the application of the various ML methods for costing and assesses their usage and application in the costing practice via an exploratory critical review. Findings indicate that ML algorithms would improve the accuracy and efficiency of costing practice but cannot replace the professionals and data availability.

Keywords: AI; artificial neural network; cost estimating; machine learning

INTRODUCTION

Every construction project is faced with a series of risks and uncertainties at the early design stages and throughout the project's life cycle. One of the major contributors to this challenge is cost estimation inaccuracies (Alqahtani and Whyte 2016), whereby lack of accurate cost data leads to failure to realise set project objectives. Therefore, cost is a major criterion in decision-making throughout the life cycle of a project (Juszczyk 2017, Elhegazy *et al.,* 2022). Cost estimation is argued as one of the most important preliminary processes in a construction project (Elfaki *et al.,* 2014) and it involves the prediction of the cost required to perform the work within project the scope. At the early stage of a project, where the scope of the project is uncertain and involves lots of ambiguities, it is particularly challenging to obtain input data for cost estimation. Therefore, the impact of overestimating or underestimating may lead to resource allocation challenges and cost overrun (Hashemi *et al.,* 2020). The accuracy and comprehensiveness of cost estimation at this stage are perceived as delicate issues that can be easily influenced by various parameters. Meanwhile, this crucial task of

[1] ayinlak@lsbu.ac.uk

Ayinla, K, Saka, A, Seidu, R and Madanayake, U (2023) The Impact of Artificial Intelligence on Construction Costing Practice *In:* Tutesigensi, A and Neilson, C J (Eds) *Proceedings of the 39th Annual ARCOM Conference,* 4-6 September 2023, University of Leeds, Leeds, UK, Association of Researchers in Construction Management, 65-74.

predicting cost and generating accurate estimates has in the past depended on the expertise of human professionals, despite the notion that an expert is prone to subjectivity and unconscious bias which can influence the result (Elfaki *et al.,* 2014, Elmousalami 2021). Traditionally, this process depends on the know-how of an estimator and assumptions are made based on experience and comparison (Bodendorf *et al.,* 2021). Therefore, it is argued that the level of accuracy required in the estimating process is largely impossible to achieve manually (Elfaki *et al.,* 2014).

The importance of the use of intelligent techniques in dealing with the challenges faced with cost estimating in the construction sector has been overemphasised since intelligent costing methods have the potential to significantly reduce effort and time (Bodendorf *et al.,* 2021). The implementation of machine learning (ML), a subgroup of artificial intelligence (AI) in the construction industry is transforming project delivery and redefining tasks executed by construction professionals and has the potential to shape construction delivery processes (Xu *et al.,* 2021). ML focuses on mimicking human intelligence and is generally described as the ability of a computer to learn without being explicitly programmed because of extracting patterns from historical data (Akinosho *et al.,* 2020). This aspect of AI investigates the work and composition of algorithms which can take advantage of and create assumptions about data thus enabling computers to make decisions, recognize speech, and visualize in 3D (Ford 2015). There has been a growing interest in ML research particularly deep learning (DL), a branch of ML, due to its capabilities in automating construction processes and improving productivity and performance (Xu *et al.,* 2021), with some studies focused on how construction processes can benefit from digitisation and AI (Adesi *et al.,* 2018).

While machine and deep learning are arguably in their infancy in the construction sector (Hong *et al.,* 2020, Xu *et al.,* 2021, Ang 2022), they present opportunities in addressing challenges with early cost prediction (Ang 2022). AI techniques are now being considered as a key solution in handling the ambiguity and challenges with cost estimation in construction projects (Elhegazy *et al.,* 2022). Therefore, this study explores the application of ML in relation to the practice of construction cost estimating throughout a project lifecycle and critically analyses the impact of ML techniques on practices of cost estimating and the challenges and opportunities presented. The outcome of this study will provide necessary information on the current use of various AI techniques and their application in costing practice in the construction sector and provide information on strategic future directions on these practices and how to harness emerging opportunities that AI presents.

METHOD

This study employs an exploratory research approach through systematic identification of publications on the theme of Artificial Intelligence techniques for costing practice in the construction industry. Search queries on AI techniques ('Machine learning', OR 'Artificial intelligence' OR 'Artificial Neural Network'), costing practice ('Costing' OR 'Estimating' OR 'Cost modelling'), and construction industry ('Construction Industry' OR 'AEC industry' OR Architecture Construction and Engineering*') were developed and searched on Scopus, Web of Science and Google scholar. This is because these databases house relevant publications and have been employed in similar studies. Inclusion criteria include publications in English language and are limited to 'Construction Building Technology' and 'Engineering civil' categories without year limitations and document type restrictions. The final publications were critically reviewed to identify and aggregate key findings and lesson learnt from these studies which forms the basis of the output of this research.

LITERATURE REVIEW

Evolution of Construction Cost Modelling and Estimating Practice

There are various methods and models for generating the cost of a product. The suitability of each model is often dependent on the type of project, the information required for completing the cost estimate, and the field of application. Particularly in the construction field, the use of these cost models could be a question of what phase the project is and what data is available at that point in time (Günaydin and Doğan 2004). Niazi *et al.,* 2006) gave a comprehensive classification system for construction cost modelling methods using a 2-category classification namely the quantitative and qualitative approaches (Figure 1). This classification simply groups cost modelling techniques based on the level of information required to generate an estimate. For instance, the parametric and analytical methods which require some form of computational analysis and requires a lower level of granularity to derive the cost falls into the quantitative category, while the more subjective method such as intuitive and analogical methods are grouped under the qualitative approach. The basis of use for these categories of cost modelling methods will be reviewed next.

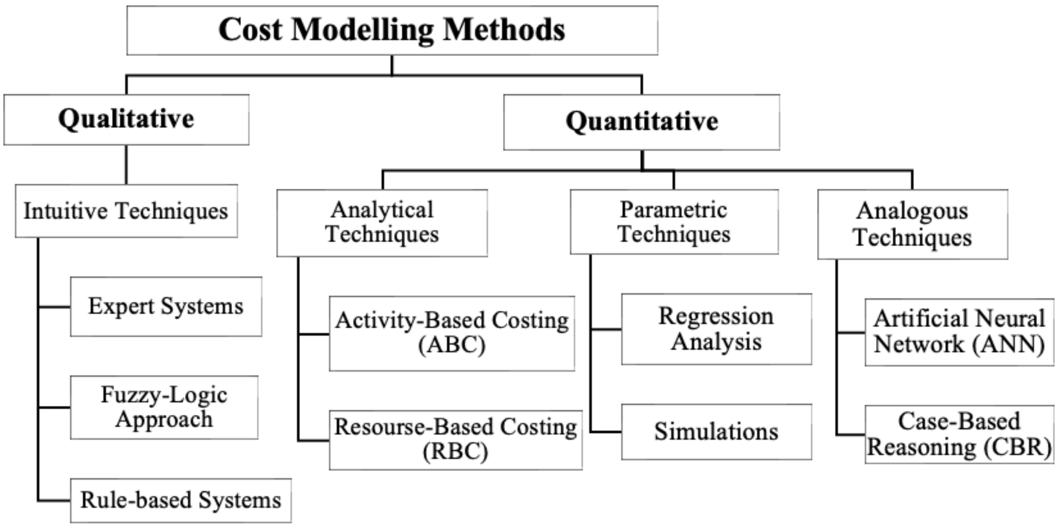

Figure 1: Classification of cost modelling techniques

Progression of the cost modelling practices.

The prediction of project cost received great interest during the early 1970s to the late 1980s in the construction industry because of the need for more accurate estimation due to the capital value of construction projects and the level of uncertainties involved throughout a project life cycle. Although few models were already in existence before this time, the early cost models were criticised for being less value-driven because of their failure to account for future uncertainties in construction and their inability to generate reliable cost estimates (Khosrowshahi and Kakat 1996, Yaman and Tas 2007). Accordingly, the historical development of cost-estimating techniques/tools have progressed from three stages in the construction industry: first, second and third generations, however given the recent technology drive in the construction sector, a fourth category is now necessary. As outlined in Table 1, the first-generation techniques are based majorly on building functional cost analysis approaches and consist of resource-based costing (RBC), also known as elemental-based costing (EBC) and activity-based costing (ABC) methods. The RBC method is the most widely used method in the construction sector and forms the basis of construction cost estimating guides for professionals as seen in the RICS new rules of measurements (RICS 2021). The ABC approach, on the other hand, presents an advancement to the RBC method in terms of its ability to accurately trace the cost/unit of products, thus improving the accuracy of the cost data. However, both these first-generation models (RBC and ABC) have limitations in the aspect of cost prediction (Khosrowshahi and Kakat 1996). This led to the development of other tools to fill this gap in the early 70s.

The parametric method is based on statistical approaches such as regression analysis (linear regression and multiple regression models) involving the use of historical data to predict cost. These methods became popular in the late 1970s and are described as second-generation methods. This method is considered powerful at the feasibility stage of a project due to its ease, and speed of application (Chou 2011). However, the accuracy of the method when relationships are non-linear has been an issue as they are augured to undermine the role of many variable cost drivers that influence a project cost. Third-generation cost modelling tools used for construction projects experienced usage at the beginning of the 1980s. These techniques are based on simulations and risk models such as the Monte Carlo simulation and are initiated as the field of project management continues to expand in the sector. The Monte Carlo simulation method significantly reduces the risk associated with cost estimates using a range estimating approach where estimators determine the minimum, most likely, and maximum possible cost (Chou *et al.,* 2009) and the probability of exceeding the 'most likely' estimate.

More novel approaches have since arrived after the simulation method because of the application of artificial intelligence and its implementation into cost modelling. Artificial intelligence (AI) approaches such as artificial neural networks (ANN), expert systems (ES) and case-based reasoning (CBR) models have been investigated since the late 1980s (Kim and Shim 2014) and classified as the fourth-generation models. These methods imitate the human brain function by learning from previous experience to predict cost and produce very accurate results as well as reducing the time for estimation. The recent technological advancement in the construction sector over the last decades has propagated the use and implementation of Building Information Modelling (BIM), which can be regarded as the fifth-generation approach. However, this method uses the first-generation RBC/EBC techniques and differs in that the quantity take-offs and estimation are now automated and supported by BIM applications (Wu *et al.,* 2014). Measurements are automatically taken off from digital models of a building, which traditionally has been very time-consuming. However, there is still a limitation to this method as BIM-based cost estimations are generated only based on the information in the BIM model while disregarding possible other external factors (Abioye *et al.,* 2021). More recently, there is ongoing research on the integration of the combination of AI and BIM to create more accurate estimates through the application of AI-based prediction models in the estimation and predictions of construction costs (Ang 2022). Therefore, the next section will focus on the progress with the use of AI-based approaches in cost-estimating practice and the benefits and challenges they present.

Table 1: Progressive development of Cost Estimating Practices

Timeline	Methods	Strength	Weaknesses	Source
First Generation (Pre - 1960s)	Elemental-Based Costing Activity-Based Costing	Detailed breakdown on cost information Provides information on different levels of analysis. More accurate with ability to track product cost.	Lack of detailed consideration for risk and uncertainties. Not suitable for early development stage. Lacks process view. Not suitable for cost prediction	(Khosrowshahi and Kakat 1996, Akintoye and Fitzgerald 2000). (Günaydin and Doğan 2004, Ayachit *et al.,* 2014)
Second Generation (Early - Late 1970s)	Parametric	Easy to understand due to string mathematical basis. Good at prediction Speed of execution Considerably accurate	Over simplistic and undermine many variables. Inaccuracies when relationship is non-linear	(Khosrowshahi and Kakat 1996, Günaydin and Doğan 2004)
Third Generation (Early 1980s)	Probabilistic Method	Reduces risk with cost estimates	Need advanced user data quantity and quality. Mostly based on the assumption of triangular distribution	(Chou *et al.,* 2009, Chou 2011)
Fourth Generation (Late 1980s)	Network-Based Approaches	Needs less statistical training to perform prediction. Can detect non-linear relationship among variables. Accuracy due to capability developed by numerous training algorithms	Black box method without user understanding Difficult to explain the outcome. Requires large pool of data to be dependable.	(Juszczyk 2017, Elmousalami 2021)
Fifth Generation (Post 2000s)	Building Information Modelling	Speeds up traditional estimating process Ability to link cost information to building model	Based on the quality of the BIM model Cannot automatically identify missing or unmeasured elements	(Wu *et al.,* 2014)

Impact of ML Techniques on Costing Practice

To develop machines that can simulate human cognitive mechanisms (machines), different AI methods have been developed. AI systems utilise different inter-connected sensors for data collection through a process known as data fusion to integrate and detect possible inferences and characterisations from the data (Pan and Zhang 2021). Machine learning (ML) is concerned with the design of computer programs and algorithms that are capable of cognitive skills and capable of reaching decisions which are traditionally regarded as human skills (Abioye *et al.,* 2021, Xu *et al.,* 2021, Oluleye *et al.,* 2023). ML has 3 categories (i) supervised learning (ii) unsupervised learning (iii) reinforcement. Other studies have also used classifications based on (i) shallow learning which uses a single layer of neural network nodes (ii) deep learning which uses multiple layers to process large amounts of training data (Xu *et al.,* 2021). Deep learning (DL) being the current state-of-the-art in ML has been proven to provide more accurate predictions than conventional ML techniques.

There has been a growing interest in research on ML in the AEC industry which is partly because these technologies can play an important role in processing the 'huge' amount of data being generated and used by construction professionals during a project lifecycle and can be attributed to their potential impact on the cloud-based computing technologies used within the industry. The idea of AI technologies implementation in the AEC industry is to improve the industry's productivity and efficiency to support the complexities in factors such as varying roles, uncertainty in environmental hazards and others (Pan and Zhang 2021, Oluleye *et al.,* 2023). It is expected that the development of ML and DL will reshape the whole costing practice, however, the use of human-based methods is still common in the construction costing practice. ML techniques could be employed to solve classification or regression problems. Costing practice that deals with regression and ML algorithms such as ANN, Logistic Regression, Support Vector Regression (SVR), and Deep neural networks (DNN) can be leveraged. The following section presents a brief discussion of the algorithms and studies that have employed them for costing practice in the extant literature.

Artificial Neural Network (ANN)

Artificial neural network (ANN) is one of the many algorithms of machine learning which models biological learning processes by computers (Hashemi *et al.,* 2020), and is commonly used for cost prediction/forecasting. ANN has been applied to mimic the human system of information processing and can predict the cost of relevant construction tasks. Its application in the construction sector is well documented in existing literature. For instance, Aibinu *et al.,* (2011) predicted the cost of pre-tender cost estimates using ANN algorithm. Similarly, Bala *et al.,* 2014) employed ANN for predicting the cost of institutional buildings. The method has also been used for life cycle cost modelling for buildings (Alqahtani and Whyte 2016) and for predicting the cost of composite flooring systems for multistorey buildings (Elhegazy *et al.,* 2022). The main limitation of ANN like any other ML algorithm is the need for adequate cost data from past projects to prevent overfitting of the ANN models (Elmousalami 2021). Furthermore, the 'black box' nature of ANN models makes it difficult for the stakeholders to understand how the predictions were made.

Case-Based Reasoning (CBR)

CBR is a technique that makes use of the information contained in past cases (i.e., previous projects) to generate cost estimates. This is a data mining technique which tends to remember information and uses the solution implemented for similar projects in solving new problems (Ji *et al.,* 2011). The best matching example like the project at hand is determined to cost of the new project (Niazi *et al.,* 2006). CBR has experienced wide application in construction costing practice such as for modelling the cost of new construction projects (An *et al.,* 2007, Ji *et al.,* 2011, Ahn *et al.,* 2020). Information on previous projects is usually stored in a database and the characteristics that match the specification of the new project (based on percentage similarity score) while taking note of changes in systems. There have been well-documented efforts in improving the result of CBR (Ahn *et al.,* 2020). However, there are still challenges with the attribution of weight values (Ji *et al.,* 2011).

Regression Algorithm (R)

There are different regression algorithms that could be employed for costing practice in the construction industry such as single learners and ensembles. Single-learner algorithms include linear regression, multiple linear regression, polynomial regression, decision tree regression and support vector regression. On the other hand, ensemble algorithms include random forest regression, gradient boosting regression and Bayesian regression. Ensemble algorithms do outperform single learners and single algorithm such as Support Vector Regression (SVR) has an advantage in self-learning and high performance in generalisation (Kim *et al.,* 2013). The limitations of these algorithms are linearity assumption, overfitting, underfitting, sensitivity to outliers, extrapolation, multicollinearity, and difficulty in handling categorical variables (Géron 2022).

Deep Neural Networks (DNNs)

Deep neural network (DNN) provides more depth to a standard neural network and are typically trained to model complex non-linear relationships through their ability to extract unique features (Akinosho *et al.,* 2020) Diverse DNN algorithms can be leveraged for costing practice in the construction industry. For instance, Convolutional Neural Network (CNN) can be employed for automated quantity estimation from images and blueprints, Recurrent Neural Networks (RNN) can be employed for time series analysis in construction cost forecasting, Generative Adversarial Networks (GAN) can generate synthetic cost data based on past historical data to improve accuracy; and lastly, transformer networks like ChatGPT which are large language model can be prompt with textual specification and project description to automate cost estimation (Saka *et al.,* 2023). However, the use of DNNs in costing practice is limited because of the data and expertise needed in deploying these models for costing. One instance found in published journals is Wang *et al.,* (2022)where it was used to determine the effects of economic factors on construction costs for public school projects. This technique, therefore, needs more practical applications in constriction costing practices.

CONCLUSIONS

The practice of construction cost estimating continues to be a major area of interest to construction researchers due to the importance of accurate cost estimation in decision-making for construction projects. Application of AI such as machine learning algorithms hold immeasurable benefits for costing practice because of their predictive strengths, capability to infer complex relationships and accuracy. However, the findings of this study revealed that the use of ML algorithms for costing practice is limited and majorly in the early estimating stage. There is limited deployment of these algorithms in tendering and construction phase of the project because of the unavailability of data to train the model and 'black box' nature of AI which affects trust.

There is a need for quality and structured data to improve the accuracy of the ML models and adequate size to prevent overfitting and underfitting of the models. The study highlighted that although the current applications are limited, there are opportunities for deploying ANN, regression models, and DNNs for costing practice in the construction industry. These ML algorithms would improve the accuracy and efficiency of the costing and estimating practice but cannot replace the professionals. Domain knowledge is required in fitting of the model - determining the right predictors and understanding the data - explaining and deploying the models in practice. As such, it is important for professionals to leverage these algorithms and not see them as competitors and usurpers.

Lastly, this study is a critical review on the impact of AI on costing practice and is limited by the size of publications employed for the review. Further study would employ a systematic review to evaluate the impact of ML on costing practice whilst highlighting the challenges, and opportunities and providing a case validation.

REFERENCES

Abioye, S O, Oyedele, L O, Akanbi, L, Ajayi, A, Davila Delgado, J M, Bilal, M, Akinade, O O and Ahmed, A (2021) Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges, *Journal of Building Engineering*, **44**, 103299.

Adesi, M, Murphy, R and Kehily, D (2018) The role of digitisation in the strategic planning process of Irish quantity surveying practices. *In:* Gorse, C and Neilson, C J (Eds.), *Proceedings 34th Annual ARCOM Conference*, 3-5 September 2018, Queen's University, Belfast, UK. Association of Researchers in Construction Management, 250–259.

Ahn, S J, Han, S U and Al-Hussein, M (2020) Improvement of transportation cost estimation for prefabricated construction using geo-fence-based large-scale GPS data feature extraction and support vector regression, *Advanced Engineering Informatics*, **43**, 101012.

Aibinu, A, Dassanayake, D and Chau, T (2011) Use of artificial intelligence to predict the accuracy of pre-tender building cost estimate, *In*: *Management and Innovation for a Sustainable Built Environment*, Amsterdam: CIB.

Akinosho, T D, Oyedele, L O, Bilal, M, Ajayi, A O, Delgado, M D, Akinade, O O and Ahmed, A A (2020) Deep learning in the construction industry: A review of present status and future innovations, *Journal of Building Engineering*, **32**, 101827.

Akintoye, A and Fitzgerald, E (2000) A survey of current cost estimating practices in the UK, *Construction Management and Economics*, **18**(2), 161-172.

Alqahtani, A and Whyte, A (2016) Estimation of life-cycle costs of buildings: regression vs artificial neural network, *Built Environment Project and Asset Management*, **6**(1), 30-43.

An, S H, Kim, G H and Kang, K I (2007) A case-based reasoning cost estimating model using experience by analytic hierarchy process, *Building and Environment*, **42**(7), 2573-2579.

Ang, Y Q (2022) Data science and machine learning for quantity surveying: A brief perspective, *In*: *Proceedings of the PAQS Congress*, Singapore: Singapore Institute of Surveyors and Valuers, 5-16.

Ayachit, A C, Attarde, P M and Kulkarni, S (2014) Activity based costing in construction projects, *International Journal of Advanced Engineering Research and Studies*, **4**(1), 9-14.

Bala, K, Bustani, S A and Waziri, B S (2014) A computer-based cost prediction model for institutional building projects in Nigeria an artificial neural network approach, *Journal of Engineering, Design and Technology*, **12**(4), 518-529.

Bodendorf, F, Merkl, P and Franke, J (2021) Intelligent cost estimation by machine learning in supply management: A structured literature review, *Computers and Industrial Engineering*, **160**, 107601.

Chou, J S (2011) Cost simulation in an item-based project involving construction engineering and management, *International Journal of Project Management*, **29**(6), 706-717.

Chou, J S, Yang, I T and Chong, W K (2009) Probabilistic simulation for developing likelihood distribution of engineering project cost, *Automation in Construction*, **18**(5), 570-577.

Elfaki, A O, Alatawi, S and Abushandi, E (2014) Using intelligent techniques in construction project cost estimation: 10-Year survey, *Advances in Civil Engineering*, 107926.

Elhegazy, H, Chakraborty, D, Elzarka, H, Ebid, A M, Mahdi, I M, Aboul Haggag, S Y and Abdel Rashid, I (2022) artificial intelligence for developing accurate preliminary cost estimates for composite flooring systems of multi-storey buildings, *Journal of Asian Architecture and Building Engineering*, **21**(1), 120-132.

Elmousalami, H H (2021) Comparison of artificial intelligence techniques for project conceptual cost prediction: A case study and comparative analysis, *IEEE Transactions on Engineering Management*, **68**(1), 183-196.

Géron, A (2022) *Hands-on Machine Learning with Scikit-Learn, Keras and Tensorflow 3rd Edition*, New York: O'Reilly Media, Inc.

Günaydin, H M and Doğan, S Z (2004) A neural network approach for early cost estimation of structural systems of buildings, *International Journal of Project Management*, **22**(7), 595-602.

Hashemi, S T, Ebadati, O M and Kaur, H (2020) Cost estimation and prediction in construction projects: A systematic review on machine learning techniques, *SN Applied Sciences 2*, 1703.

Hong, T, Wang, Z, Luo, X and Zhang, W (2020) State-of-the-art on research and applications of machine learning in the building life cycle, *Energy and Buildings*, **212**, 109831.

Ji, S H, Park, M and Lee, H S (2011) Cost estimation model for building projects using case-based reasoning, *Canadian Journal of Civil Engineering*, **38**(5), 570-581.

Juszczyk, M (2017) The challenges of nonparametric cost estimation of construction work with the use of artificial intelligence tools, Procedia Engineering, **196**, 415-422.

Khosrowshahi, U K F and Kakat, A P (1996) Estimation of Project Total Cost and Duration for Housing Projects in the UK, *Building and Environment*, **31**(4), 373-383.

Niazi, A, Dai, J S, Balabani, S and Seneviratne, L (2006) Product cost estimation: Technique classification and methodology review, *Journal of Manufacturing Science and Engineering*, **128**(2).

Oluleye, B I, Chan, D W M and Antwi-Afari, P (2023) Adopting Artificial Intelligence for enhancing the implementation of systemic circularity in the construction industry: A critical review, *Sustainable Production and Consumption*, **35**, 509-524.

Pan, Y and Zhang, L (2021) A BIM-data mining integrated digital twin framework for advanced project management, *Automation in Construction*, **124**, 103564.

RICS (2021) *New Rules of Measurement: Detailed Measurement for Building Works 2nd Edition*, London: Royal Institution of Chartered Surveyors (RICS).

Saka, A B, Oyedele, L O, Akanbi, L A, Ganiyu, S A, Chan, D W M and Bello, S A (2023) Conversational artificial intelligence in the AEC industry: A review of present status, challenges and opportunities, *Advanced Engineering Informatics*, **55**, 101869.

Wang, R, Asghari, V, Cheung, C M, Hsu, S C and Lee, C J (2022) Assessing effects of economic factors on construction cost estimation using deep neural networks, *Automation in Construction*, **134**, 104080.

Wu, S, Wood, G and Ginige, K (2014) A technical review of BIM based cost estimating in UK quantity surveying practice, standards and tools, *Journal of Information Technology in Construction*, **19**, 534-563.

Xu, Y, Zhou, Y, Sekula, P and Ding, L (2021) Machine learning in construction: From shallow to deep learning, *Developments in the Built Environment*, **6**, 100045.

Yaman, H and Tas, E (2007) A building cost estimation model based on functional elements, *ITU Journal of Faculty of Architecture*, **4**(1), 73-87.