The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0969-9988.htm

Application of machine learning in predicting construction project profit in Ghana using Support Vector Regression Algorithm (SVRA)

Received 9 August 2020
Revised 5 December 2020
24 January 2021
24 March 2021
Accepted 30 March 2021

Emmanuel Adinyira
Construction Technology and Management, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
Emmanuel Akoi-Gyebi Adjei
Building Technology, Accra Technical University, Accra, Ghana and Construction Technology and Management, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
Kofi Agyekum
Department of Construction Technology and Management, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, and
Frank Desmond Kofi Fugar
Construction Technology and Management, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Abstract

Purpose – Knowledge of the effect of various cash-flow factors on expected project profit is important to effectively manage productivity on construction projects. This study was conducted to develop and test the sensitivity of a Machine Learning Support Vector Regression Algorithm (SVRA) to predict construction project profit in Ghana.

Design/methodology/approach – The study relied on data from 150 institutional projects executed within the past five years (2014–2018) in developing the model. Eighty percent (80%) of the data from the 150 projects was used at hyperparameter selection and final training phases of the model development and the remaining 20% for model testing. Using MATLAB for Support Vector Regression, the parameters available for tuning were the epsilon values, the kernel scale, the box constraint and standardisations. The sensitivity index was computed to determine the degree to which the independent variables impact the dependent variable.

Findings – The developed model's predictions perfectly fitted the data and explained all the variability of the response data around its mean. Average predictive accuracy of 73.66% was achieved with all the variables on the different projects in validation. The developed SVR model was sensitive to labour and loan.

Originality/value – The developed SVRA combines variation, defective works and labour with other financial constraints, which have been the variables used in previous studies. It will aid contractors in predicting profit on completion at commencement and also provide information on the effect of changes to cash-flow factors on profit.

Keywords Cash flow, Construction project, Machine learning, Profit, Productivity, Support vector regression

Paper type Research paper

Introduction

The high level of uncertainties encountered in the construction industry worldwide renders most construction companies bankrupt (Majer *et al.*, 2020). Of all the uncertainties that contribute to the bankruptcy of construction companies, financial and budgetary factors play significant roles (Majer *et al.*, 2020; Arditi *et al.*, 2000). It is established that over 60% of

Engineering, Construction and Architectural Management
Vol. 28 No. 5, 2021
pp. 1491-1514
© Emerald Publishing Limited
0969-9988
DOI 10.1108/ECAM-08-2020-0618

construction company failures worldwide have their roots traced to financial issues, which affect their profitability status (Majer *et al.*, 2020). According to Zhu *et al.* (2019), most of these construction projects fail to achieve their expected profits because of the difficulties in managing certain risks in a more complex web of project organisation. Profitability is, therefore classified as one of the most obvious causes of insolvency in the construction industry (Arditi *et al.*, 2000).

Profitability essentially can be argued to mean a measure of productivity if improved productivity is viewed as an increase in value-added outputs with fewer resources (Snyman and Smallwood, 2017). When an organisation focuses more on increasing the amount of work it takes on without considering the profits, the organisation will often settle for smaller profit margins or secure work that is not profitable with the purpose of increasing its volume of work (Snyman and Smallwood, 2017). The issue of construction failure as a result of dwindling profit has been of interest to the scientific community. Construction is a very risky business (Lee, 2009) and operates within a competitive environment that usually compels contractors to introduce lower profit margins in their tender for projects to win bids. Unfortunately, the nature of construction projects that is largely characterised by constant environmental changes, pressures to maintain schedules and to reduce costs, among other things, creates a significant challenge to managing projects (Cheng *et al.*, 2015). As a matter of importance, there is a need to put in place measures that can provide construction companies with accurate information on critical factors like possible profit status (Zhu *et al.*, 2019). Unfortunately, this is not the case because, in practice, contractors are usually unwilling to conduct further profit performance predictions due to time and cash restrictions (Zhu *et al.*, 2019). Of these two factors, as revealed by Zhu *et al.* (2019), cash has been identified as a critical factor that imposes a significant influence on project profitability (Jiang *et al.*, 2011).

Currently, in Ghana, because construction companies operate in highly competitive environments and cannot survive without effective management, they have resorted to the introduction of low-profit margins in tender bids to compete within the industry (Adjei *et al.*, 2018). This decision by such companies in Ghana in the long term affect their liquidity, which then leads to the failure of projects and subsequent bankruptcy of such companies (Adjei *et al.*, 2018). Poor cash-flow control tends to cause project failure for contractors, as a result of the liquidity shortage for supporting their daily activities (Khosrowshahi and Kaka, 2007). Due to this, Mohsin *et al.* (2014) posit that reliably predicting cash flow over various phases of a project is highly desirable because the project manager will be in a better position to curb the potential financial issues that can hinder the success of the project. Cash-flow forecasting, therefore, is an important tool to evaluate the distribution of expenditure and revenues of projects.

Several researchers have developed models that have the potential to aid contractors to plan their cash flows (El-Kholy, 2014; Jiang *et al.*, 2011). Authors who have developed most of these models have employed more advanced forecasting algorithms like Kalman filters, Artificial Neural Networks (ANN) and Support Vector Machines (SVM) (Sapankevych and Sankar, 2009). Unfortunately, Mohsin *et al.* (2014) reckon that these models are either project-specific or do not address the consequences of failure to meet the minimum funds required for cash outflow. This notwithstanding, Cheng *et al.* (2015) has been of the view that since SVM has been utilised in several applications in construction engineering, it can be a promising tool for cash-flow problems.

This study has become necessary because apart from productivity, two basic threats that affect the financial future of construction organisations are profitability and cash flow (Peterson, 2009). With the issue of productivity and cash flow well reported in the international literature, there is the need to focus on how contractors within a typical developing country setting like Ghana can set up an appropriate decision support system that can guide construction companies (especially during the bidding process) to maximise their profits. Since the impact of cash flow on the profitability of contractors is well established, this study implores cash-flow factors to develop a model that can predict profit on construction projects in Ghana through Support Vector Regression Algorithm (SVRA). One of the key contributions of this study stems from the fact that the SVRA model developed contributes its quota to the state-of-the-art by potentially filling the profit measurement gap with the incorporation of variation, defective works and labour to financial constraints which have been the variables used in previous studies. The study begins with an introduction to the central theme of the study, followed by a critical review of related literature. The research methodology is then presented, after which the results are presented and discussed, and then finally, a conclusion is drawn.

Literature review

This section presents a comparative review of literature related to the subject under investigation. It begins with a general review of literature on the construction industry and profit challenges, and then narrows down to cash flow as a profit challenge and discusses the factors that affect it. The section continues with a review of literature on various theories of profits and then proposes a conceptual framework that relates the discussed theories of profit to construction project profit. The section concludes with a review of various predictive models, with emphasis laid on Support Vector Regression in Machine Learning.

The construction industry and profit challenges

Akintoye and Skitmore (1991) defined profitability as a percentage of the profit of turnover (POT) or returns on capital investment (ROI). It is a residual of sales revenue once all costs (including interest payment on debt) have been deducted (Achenef, 2016). Malik (2011) classified profitability as being one of the most important objectives of financial management because a business that is not profitable cannot survive. Wright (1977) considered profitability as a function of three factors, that is sales volume (work done), the capital investment needed to support business and the margin of profit earned. The level of profitability of any construction company depends on the demand for building works from both the private and public sectors (Lee, 2009). This means that a reduction in turnover generally reduces the level of profitability (Lee, 2009). Key challenges confronting the profitability of construction firms include the size of the company (Akintola and Skitmore, 1991); the extent of mechanisation (Hillebrandt, 2000); the extent of subcontracting (Reich, 2000); the extent of labour-only subcontracting (Langford and Chan, 1987); the volume of sales and profit margins (Hillbrandt *et al.*, 1995); gearing factor and turnover of capital employed (Wright, 1977); property and construction cycles (Hillbrandt, 2000); competition (Akintoye and Skitmore, 1991) and macro-economic factors (Lee, 2009). A few of these challenges are discussed in the further paragraphs.

Size of the company. Though there are mixed reactions regarding the influence of firm size on profitability, studies related to the construction industry have confirmed that the size of a construction company has a positive correlation with its profit ratio to an optimum level (Akintola and Skitmore, 1991). Hillbrandt (2000) was of the view that this assertion may be true because larger construction companies had the potential resources to heavily invest in construction activities, unlike the smaller ones. Spedding (1977) was also of the view that such large companies are more efficient and well organised in their managerial strategies compared to the smaller ones. Lee (2009) supported these assertions and added that when it comes to estimating, pricing and production, larger companies are very consistent. Though there is evidence in the literature that suggests that large contractors are consistent in maintaining their profitability levels as compared to smaller ones, the volume of size has yet to be analysed (Lee, 2009).

Competition. The construction industry is very competitive by its nature. The growing importance of competitiveness in the industry depicts it as a predictor of business performance (Oyewobi *et al.*, 2014). The competitive environment necessitates that construction organisations are more strategic and proactive to increase their chances of survival (Oyewobi *et al.*, 2014). According to the standard economic theory, increased competition among firms results in lower profitability (Lee, 2009). The competitiveness of the construction industry motivates contractors to bid for jobs using low mark-ups to provide a high chance of job acquisition (Halim, 2014). When this happens, firms with insufficient profit margins may go bankrupt, thereby resulting in their tendencies to collapse. A similar scenario was reported by Mahamid (2011), who indicated that in the USA, insufficient profit caused about 27% of construction firms to fail. It must be noted that the unfamiliarity of the risk involved in this competition that normally occurs during bidding in the construction industry has contributed to low profitability levels in the industry (Yean *et al.*, 2005).

Gearing factor and turnover of capital employed. The gearing factor measures the quantum of investment made against the volume of sales or work done (Wright, 1977). The gearing ratio is an important measure of the stability of a company since it is considered when raising external capital (Tunji *et al.*, 2015). If a company is already highly geared, it might find it extremely difficult to raise additional fund as the would-be lender may take a closer look at its structure and believe that the company might not be able to settle the debts as at when due (Tunji *et al.*, 2015). This is because it may already be exposed to so many creditors. The effect of having excess gearing is that, such a company would have to accumulate the higher amount of profit before interest and tax to be able to meet the demand for interest payment. Improving the effectiveness with which firms use their capital employed can improve the profitability of the firm (Lee, 2009). The turnover of capital employed in any construction business is largely determined by how that firm invests its capital in fixed assets, among other things (Lee, 2009). As a matter of fact, the lesser the investment expenditure made in, for instance, fixed assets to support a given level of activity, the more profitable that firm will be (Lee, 2009). When a construction company exhibits a high gearing ratio for loans, its profitability and chances of survival in the long term are greatly impacted (Lee, 2009).

Cash flow as a profit challenge and factors affecting it

Though there is a close relationship between cash flow and profit, the two are not the same. Cash flow represents the cash inflows and outflows from the business. Subtracting cash outflows from the cash inflows results in the net cash flow (Hofstrand, 2013). Profitability, on the other hand, represents the income and expenses of a business (Hofstrand, 2013). Subtracting the expenses from the income either results in a profit or a loss. Despite this difference, making a profit generates cash flow (Tracy and Tracy, 2015). This, therefore, makes it imperative for any company to have sufficient cash to pay its creditors, suppliers, employees, etc. (Omopariola *et al.*, 2020). According to Pansegrauw (2019), whilst revenue is vanity, cash flow is sanity and cash is king, and it constitutes an essential construction company resource. Omopariola *et al.* (2020) stressed that the significance of cash flow relates to cash as the basic source of every successful construction project. As a matter of fact, a stable financial performance that is achieved through a productive cash-flow analysis assists construction organisations to exploit their investment opportunities (Seo *et al.*, 2018). The contractor sees cash flow as very important because it reflects a project's financial performance prior to the completion of the contract and the settling of final accounts (Usman *et al.*, 2016).

Construction is undertaken within a risky environment and subjected to significant uncertainties that include the need for capital, delays in client payments and varying interest rates between the contract end time and final payment (Barbosa and Pimentel, 2010).

During high-interest rates and inflation levels periods, cash-flow forecasting becomes more important, and this makes it an ideal tool to evaluate the distribution of expenditure and revenues for projects concerning project time (Mohsin *et al.*, 2014). The non-availability or inadequacy of a structure to manage cash-flow results in liquidity problems that affect working capital without warning, and this defies the sustainability of projects.

Numerous factors have been identified from the literature to have an impact on cash flow (Olatunji, 2010). Studies such as those of Buertey and Adjei-Kumi (2012), Liu *et al.* (2009) and Odeyinka *et al.* (2013), have particularly identified several factors that impact the cash flow of construction activities from the perspective of developing countries. Table 1 presents cash flow factors gleaned from the extensive review of related literature. For lack of space, only a few of the critical factors that affect the cash flow of projects have been discussed. Other equally critical ones are summarised in Table 1.

In a study by Asante (2014), it was revealed that most Ghanaian contractors are not sufficiently capitally equipped, and this influences their loan acquisition possibilities in addition to the contractor's ability to deliver projects successfully. Contractors thus extremely depend upon outsourced capital from suppliers, subcontractors and advance payments from clients (Asante, 2014). It becomes increasingly difficult under such a situation to persuade creditors and potential lenders of the short-term inadequacy of cash. The unusual remedy is, therefore, to issue a loan at a high capital cost to cover the risk posed (Gundecha, 2013). This increase in loan cost affects cash flow on projects negatively and further causes greater operating cost that reduces profit (Bolek and Wiliński, 2012). Contractors, thus, strive to take the heavy daily construction expenses that often comprise huge sums of money

Sn	Cash flow factors	Sources
1	Progress payment duration	RICS (2011)
2	Progress payment conditions	RICS (2011)
3	Advance payment	Asante (2014), Su and Lucko (2015)
4	Percentage of retention	Kaka and Cheetham (1997), Hughes *et al.* (2000), Park *et al.* (2005), Liu *et al.* (2009)
5	Time of releasing the retention	Park *et al.* (2005), Liu *et al.* (2009)
6	Limit of retention	RICS (2011)
7	Repayment of loan	Bolek and Wiliński (2012)
8	Withholding tax	Buertey and Adjei-Kumi (2012)
9	Payment of creditors	Asante (2014)
10	Overwork measurement	Liu *et al.* (2009), Buertey and Adjei-Kumi (2012), Gundecha (2013), Aziz (2013)
11	Under work measurement	Liu *et al.* (2009), Buertey and Adjei-Kumi (2012)
12	Material cost	Park *et al.* (2005)
13	Wages of labour and staff	Park *et al.* (2005)
14	Plant and equipment cost	Park *et al.* (2005)
15	Bank interest rate	Barbosa and Pimentel (2010)
16	Sub-contracting	Lee (2009)
17	Delay of making payments to suppliers	Harris and McCaffer (2005)
18	Price variation	Harris and McCaffer (2005), Gundecha (2013)
19	Work execution errors	Long (2015)
20	Replacement of defective work	Hughes *et al.* (2000)
21	Variation of works	Mohsin *et al.* (2014)
22	Claims	Seppälä (2005), Long (2015), Usman *et al.* (2016)
23	Phasing of projects	RICS (2011)
24	Overheads	Lee (2009)
25	Number of projects being executed by a contractor	Buertey and Adjei-Kumi (2012)

Table 1. Construction project cash flow factors

coupled with payment delay (Sambasivan and Soon, 2007; El-Kholy, 2014). This problem has been identified in other studies that iterated that a toxic blend of cash-flow reliance on bank loans, high-interest rates paid, and cash-flow mismanagement is the prime source of business failure in developing countries (Enshassi *et al.*, 2009; Barbosa and Pimentel, 2010; Jiang *et al.*, 2011).

Duration has also been identified as one of the requirements on which projects are awarded. Olatunji (2010) defined project duration as the time taken to execute the project tasks from the inception of the site to the delivery of the project. Because duration plays a major role in the cash flow of a project, studies have concluded that timeliness that regulate the completion of the contract on the scheduled date are measures of contractors' quality performance that eliminates or minimises delays (Yasamis *et al.*, 2002; Olatunji, 2010). Cost elements relating to labour, materials and equipment constituting the greater percentage of project costs are assumed to be a fixed proportion of the total cost over the project's duration (Park *et al.*, 2005). However, changes in the duration of contracts occur because of a variation order or change of contract conditions, and this should be accompanied by an adjustment of weights of cost categories. Duration is, therefore considered to be an important factor that affects the cash flow of construction companies (Haseeb *et al.*, 2011; Owalabi *et al.*, 2014).

In addition to the duration of projects, the margin of profit has also been identified to play a critical role in the determination of the cash flow of any project. Its role is critical because it gauges the financial strength of a business (Mahamid, 2011). Avetisyan *et al.* (2020) defined profit of margin as the ratio of profit earned to sales receipts (costs) over a period. This means that the profit margin measures the amount of profit that a firm accrues from sales of a product or service. There is a tendency for less capital to be locked up in a contract when the margin of profit is higher (Avetisyan *et al.*, 2020). On the other hand, a greater amount of capital is locked up when the margin of profit is less (Mahamid, 2011). With the construction industry being associated with low-profit margins due to competition, contractors are encouraged to rely on cash flow as a mechanism to generate profit. A company must, therefore, operate efficiently to recover not only the cost of production but to provide compensation for clients in exchange for risk acceptance to generate a sizeable profit margin (Avetisyan *et al.*, 2020; Asante, 2014).

Payment delay is also recognised as a serious challenge confronting the construction industry of many countries (Ramachandra and Rotimi, 2011). This issue consequently exerts a financial burden on contractors who may not have large capital as the clients hence impacts on contractors' cash flow negatively in meeting the financial commitment to complete scheduled works. Available credit facilities could ease the financial burden of contractors and cover payment delay, thereby enhancing the cash flow on projects and the company at large. Asante (2014) reported that the most important aspect of cash-flow management is to avoid extended cash shortages that are caused by having too great a gap between cash inflows and outflows. He further cited empirical evidence by Jaafar and Abdul-Aziz (2005) that smaller capital is used by contractors to start construction firms compared to other businesses. This can be associated with the advance payment received from clients and the credit influence of suppliers as well as the capital support given by subcontractors, although they are paid when contractors receive payment from clients.

An equally important factor that affects the cash flow of companies is credit. Crediting is a delayed payment arrangement through which a buyer takes possession of something and pays later. It enhances the cash flow of contractors in the capital-intensive construction industry, where contractors are paid after executing works successfully and certified. Harris and McCaffer (2005) suggested that the delay of one-week is normal for labour payment while a period of three to six-weeks is also normal for plant hirers and material suppliers. The time-lapse indicated above is said to be a normal trading arrangement, and any period further than this range undermines the commercial confidence in a company.

Retention is also reported to be important when dealing with cash-flow issues. Retention is the cash retained by the employer from the contractor that protects the employer from a contractor's insolvency as well as ensuring that the contractor finishes the work as the contract stipulates (Hughes *et al.*, 2000). It provides clients of the fund to depend on when contractors fail to perform due to incompetence and bankruptcy. It also motivates contractors to undertake any minor outstanding works and the repair of defects after practical completion. Hughes *et al.* (2000) reported that most contractors are subjected to cash retention, and main contractors likewise withhold part of subcontractors' payments. Lower rates of retention are applied on large contracts since its application on the value will result in a large sum of cash. The percentage of retention is reported to be in the range of 1–15%, with 3% as median (Hughes *et al.*, 2000). An increase in retention rate from 2.5 to 5.0% would cause a corresponding increase in working capital needed from 2.61 to 4.05% of annual turnover (Kaka and Cheetham, 1997).

Theories of profit

Profit is the financial benefit realised when work is done, and also when the expenses incurred by a business are exceeded by the amount of revenue generated by the activities of that business (Avetisyan *et al.*, 2020). Mathematically, profit can be expressed as (Avetisyan *et al.*, 2020; Halim, 2014):

$$\text{sProfit} = \text{total revenue} - \text{total expenses} \tag{1}$$

Revenue is the overall amount of income received from work done or by sales of goods or services related to the primary operations of the company (Halim, 2014). It is one of the items required to continuously determine, control, and plan in connection with profit. Total expenses, on the other hand, is the total cost incurred in the construction or production of the goods or service (Avetisyan *et al.*, 2020).

The succeeding subsections present some of the notable theories on profit.

Modern theory of profit

This theory of profit also termed demand and supply of profit describe profit as the net income of an entrepreneur and an entrepreneur as a business enterprise itself (Marchal, 1951). Profit is regarded in this theory as the reward of an entrepreneur and overseen by the demand for and supply of an entrepreneur (Marchal, 1951). Demand for entrepreneurs mostly depends on the level of industrial development, elements of industrial uncertainty, the scale of production, and marginal revenue productivity of entrepreneurship. The expected profit is likely to be high when the level of industrial growth is high; the scale of production is large in addition to an increase in efficiency and productivity (Marchal, 1951).

Rent theory of profit

Walker's rent theory of profit considers profit as the rent of ability (Vercellone, 2008). This theory makes a comparative study between grades of land and the entrepreneur's abilities. Higher ability entrepreneurs earn profits as superior land earns rent (Chendroyaperumal, 2009). Just as there is the marginal or no rent land, similarly, there exists a marginal or no profit entrepreneur who earns only wages of management. Industries are managed and run by marginal entrepreneurs like marginal land. A land which at its margin earns no rent so as the marginal entrepreneur earns no profit.

Dynamic theory of profit

Clarks propounded the dynamic theory of profit and stated that profit is the difference between the price and the cost of the commodity (Bowman, 2014). It is the result of dynamic change. In a stationary state in which there are static economic conditions of demand and supply, no real or pure profit are realised as a surplus. The amount of capital invested, production methods, managerial organisation, technology and demand pattern continues to be constant in a stationary economy (Chendroyaperumal, 2009). Price, therefore, tends to equal average costs under competitive conditions, which results in zero surpluses. No pure profit is earned. However, there may be some frictional profits emerging due to frictions in the system and this cannot be regarded as real Profit (Bowman, 2014). Profits are earned because of changes or an increase in population, tastes, and wants. Also, profit is influenced by capital formation, advancement in technology, and changes in the form of business organisation (Chendroyaperumal, 2009).

Wage theory of profit

Taussig and Davenport's theory explains that profits are best regarded as a basic form of wages that accrue to the entrepreneur because of special ability (Bowman, 2014; Chendroyaperumal, 2009). The theory reasoned that there exists a close comparison between labourers and entrepreneurs in that, as wages are received by labourers for rendered services, so will entrepreneurs earn profit for a service (Bowman, 2014; Chendroyaperumal, 2009).

Managerial efficiency theory of profit

This theory, also known as the compensatory theory of profits, acknowledges that some firms are more efficient than others in terms of management of productive operations and meeting the needs of consumers successfully (McGuigan *et al.*, 2017). A firm with an average level of efficiency receives an average rate of return while that with higher managerial skills and production efficiency are required to be compensated by above-normal profits (McGuigan *et al.*, 2017; Chendroyaperumal, 2009; Sanyal, 2019).

Frictional theory of profit

This theory assumes that a normal proportion of profit exists, which is a return on capital paid to the owners of capital as repayment for saving and investment of funds rather than to spend all their income (McGuigan *et al.*, 2017; Sanyal, 2019). In a still economy where changes are not expected in demand or cost condition occurs, firms would be earning a normal rate of profit on capital and entrepreneurial capacity at equilibrium over the long run (Sanyal, 2019). Accrued economic profit under these conditions would not have ensued to a firm. The frictional theory of profit, therefore, describes the shocks or disorders that occasionally arise in an economy due to unexpected changes in demand for product or cost conditions which cause an imbalance condition (McGuigan *et al.*, 2017; Sanyal, 2019).

Monopoly theory of profit

This theory of profit is credited to the power monopoly firms enjoy (Chendroyaperumal, 2009). Products from these firms having the monopoly power are restricted and this is accompanied by prices higher than within perfect competition (McGuigan *et al.*, 2017; Sanyal, 2019). There exist robust barriers preventing the entry of new firms to add to firms under monopoly. Hence, monopolistic firms earn and continue to receive economic profits even in the long run. This power of monopoly may arise owing to sole control over some vital raw material needed to produce a product, from economies of scale, legal sanction or ownership patents and Government restrictions on the import of a commodity (McGuigan *et al.*, 2017; Sanyal, 2019).

Innovations theory of profit

This states that economic profit is generated from successful innovations introduced by entrepreneurs (Chendroyaperumal, 2009). The main function of an entrepreneur is to present innovations in the economy and gain profits as a reward for accomplishing the role (McGuigan *et al.*, 2017; Sanyal, 2019).

Risk and Uncertainty Bearing theory of profit

Risk and Uncertainty Bearing theory states that profits are the obligatory reward of an entrepreneur for accepting risk and uncertainty in a changing economy (McGuigan *et al.*, 2017; Chendroyaperumal, 2009; Sanyal, 2019). Profits occur because of future uncertainty; hence, entrepreneurs undertake production works under uncertain conditions (Sanyal, 2019).

Figure 1 conceptualises project profit in the context of this study by relating the above-discussed theories of profit to construction project profit.

Predictive models

Several studies have been undertaken for forecasting planning, and management concerning construction and other industries. In all these studies, mathematical models have been found to retain greater advantages of being practical, simple, fast and not requiring extensive information about projects (Khosrowshahi and Kaka, 2007). According to Adjei *et al.* (2019b), various comparative studies have been conducted between the traditional or conventional techniques of forecasting or predicting. Whereas regression models that basically predicts numbers abound in literature, there are also classification models that seek to predict groupings. While traditional regression procedures derive a function f(x) that has the least deviation between predicted and observed values for the training examples, machine learning regression techniques enable computer programs to automatically enhance performances of certain tasks through understanding (Pham and Afify, 2005).

Support Vector Machine Models (SVMMs) make use of machine learning regression techniques and are known to demonstrate high predictive and accuracy performance comparable to other regression techniques (Pham and Afify, 2005). SVMs are regressors that perform the regression and predict continuous ordered variables using algorithms

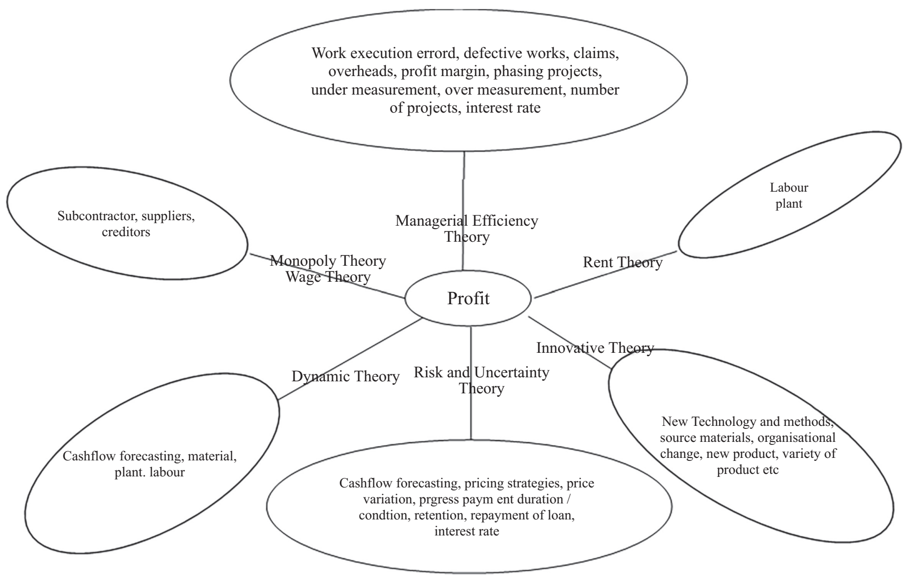

Figure 1.
Relating theories of profit to construction project profit

(Awad and Khanna, 2015). Advantages of Support Vector Regression (SVR) includes: the capability to model non-linear relationships, solves regression function by selecting only the necessary data (support vectors) resulting in a sparse solution, regression function is related to a quadratic problem that has a typical comprehensive solution, can be employed when there are few samples than variables called small and large problems, its computational complexity does not depend on the dimensionality of the input space and its excellent generalisation capability with high prediction accuracy. These advantages notwithstanding, in SVR, there is no structured method in the choice of kernel function and there is the need for the user to describe several free parameters as the overview performance of SVR models is dependent on the precise setting of free parameters. SVR raises a quadratic optimisation problem of the same training data set.

Support vector regression in machine learning

The movements of the Fourth Industrial Revolution (4IR) as seen in literature can be classified into three distinct features: defined task, undefined task and improvement possibility to existing technologies/methods (Moon *et al.*, 2020). Whereas 4IR in terms of defined task, seeks to replace the human workforce, in the case of an undefined task, it is more about diverse and complex data sets (Moon *et al.*, 2020). It employs big data and machine learning techniques to effectively process and use the collection of unlimited data from numerous sources. It looks at how and what can be learned from data. Big data analytics pulls from existing information to look for emerging patterns that can help shape decision-making processes (Moon *et al.*, 2020). Machine learning, on the other hand, learns from existing data and provides the foundation required for a machine to teach itself. Big data analytics reveals patterns through classifications and sequence analysis, but machine learning takes this concept a step higher by using the same algorithms that big data analytics uses to automatically learn from the collected data (Moon *et al.*, 2020).

The basis of one of machine learning's most popular techniques, Support Vector Machines (SVMs), was developed by Vapnik and it is categorised under supervised learning (Vapnik, 1998). SVMs solve binary classification problems by formulating them as convex optimisation problems (Awad and Khanna, 2015). There are fundamentally two phases in supervised learning namely; learning and testing phases (Listiani, 2009). The learning phase is where training data is employed to develop a mathematical model to explain the relationship between variables. The significance of training in machine learning is to assist in effectively adjusting and optimising the model parameters with the use of the categorised quantitative dataset (Kolisetty and Rajput, 2020; Huang, 2007; Verma *et al.*, 2016). On the other hand, the test phase is when the developed model is employed to predict the results of the test data set.

SVR is the SVM tool applied in regression analysis, and it estimates a continuous-valued multivariate function (Huang, 2007). SVM possesses a firm mathematical basis on statistical learning theory and the key objective of the Vapnik–Chervonenkis (VC) theory is the characterisation of the generalised error instead of the error on specific data set (Verma *et al.*, 2016). This aids SVM to better generalise for unseen data. SVR tries to curtail the upper bound on the generalisation error based on the Structural Risk Minimisation (SRM) principle rather than Training Error Minimisation (TEM) (Awad and Khanna, 2015; Listiani, 2009).

SVR is generated from Support Vector Machine (SVM) through generalisation by introducing an ε-insensitive region around the function, termed the ε-tube. It is formulated as an optimisation problem with initially defining a convex ε-insensitive loss function to be minimised and recognising the flattest tube that includes the best of the training instances (Awad and Khanna, 2015). Convex optimisation has a unique local minimisation solution. The solution is achieved using appropriate numerical optimisation algorithms. A hyperplane is represented by support vectors in SVR which are training samples that lie outside the boundary of the tube.

SVM has three main attributes namely better generalisation capability, optimal global solution using optimisation theory, and Kernel functions for nonlinearity (Huang, 2007). Hyperparameters contribute immensely to the performance of SVM. It is therefore substantial to set good value for them. The SVM parameter C and ε together with kernel parameters k is known as hyperparameters, and optimise the performance generalisation. Various kernel functions tolerate adjustment concerning training data. The training data influence the optimal hyperparameter setting and usually set manually since there is no theory to delineate a good value (Huang, 2007).

The learning curve is required to appreciate the sensitivity of SVR relative to the size of the training data set. The finding of fitting values for the hyperparameters through several rounds of model development is a significant section in SVR. It requires more time when there is more data to crunch to develop and validate a model. This is because if the repetition required to discover a grid with a set of m cost values and n ε-values is $m \times n$ times, with an average computational effort of t, then estimated duration of $m \times n \times t$ will be required (Listiani, 2009). The size of processing data dictates the definite duration taken for each round. Hence, the computational time needs minimisation. The partial data set is enough to produce a model due to the dependency of SVR on support vectors. The sensitivity of SVR is performed through running numerous times with swelling sample numbers and observing the learning effect gained from additional data by measuring Root-Mean-Squared Error (RMSE) (Listiani, 2009).

Methodology

The methodology is presented and discussed under five key sub-sections to include data collection and processing, feature selection, support vector regression, model evaluation and model development and performance.

Data collection and processing

Historical data was used in the model development and validation. A total of 150 completed projects which were financed with internally generated or donor funds within the past five years (i.e. 2014–2018) formed the source of data for the study. The choice of the period was settled on due to the price stability within this period. Eighty percent (80%) of these projects were used in the development of the model while the remaining 20% were used in the validation. The study explored the various financing options available to contractors in project financing. This was motivated by the cost of capital on the various sources which erodes profit. Therefore various combinations of these options were investigated to minimise cost in the bid to maximise profit. A financing option consisting of contractor financing and loan was employed as this is the commonest financial combination for construction projects in Ghana. The variables used in this study were adopted from the findings of Adjei *et al.* (2018) which established the most significant factors influencing profit on construction projects in Ghana. Additionally, the cost contributions of the identified significant cash-flow factors established by Adjei *et al.* (2019a) were also applied to the completed projects to generate the dataset for the model development. This approach meant that no outliers were recorded or observed in the dataset. The dataset was subsequently organised to suit the SVR learning scheme in a spreadsheet.

Feature selection

Feature selection, also known as the variable selection, is selecting attributes in data that are utmost appropriate to the predictive modelling problem under study. It improves accuracy and efficiency and aims at reducing the dimensionality and noise in data sets (Ghojogh *et al.*, 2019; Roffo, 2018; Guajardo *et al.*, 2007). Feature selection facilitates visualisation of data, improves understanding of data (Guajardo *et al.*, 2007) and minimises the risk of overfitting (Hira and Gillies, 2015). Roffo (2018) stated that automatic feature selection, that is computational variable selection techniques such as forward selection, backward elimination and stepwise regression can be employed to remove inappropriate, redundant and noisy information from data to facilitate better performance in learning. However, the human operator sometimes defines the potentially useful features in many learning domains (Roffo, 2018). In line with this, the features employed in this model were selected from the work of Adjei *et al.* (2018) which used Principal Component Analysis (PCA) to establish the significant cash-flow factors impacting on profit. These factors (features) were wages of labour and staff, progress payment duration, bank interest rate and defective works. However, some of the projects utilised in this study had some differences in the initial and final contract sums, which established some variations; hence, variation was added to the selected features. According to Windapo (2017), suppliers discounts or credit facilities are offered to enhance the profit and cash flow of companies. In view of trying to predict profit from significant cash factors, the study subsequently included a credit to enhance the prediction of profit.

Support vector regression (SVR)

SVR goes through two phases namely: learning and testing phases. The generated dataset of 150 projects was employed in these phases. According to Listiani (2009) and Verma *et al.* (2016), a bigger proportion of data set is required in the learning phase than the testing phase; hence, one could use anything around 70–30% as a rule of thumb. The study, therefore, adopted an 80–20% data partition. Subsequently, data from the 150 projects obtained for the study was divided into 120 (80% for hyperparameter selection phase and final training phase) and 30 (20% for testing phase). The learning phase also goes through two steps namely: the hyperparameter selection and training. Three different kernels (linear, polynomial and radial basis function) were evaluated with the predictors of the generated dataset to establish a suitable kernel for the development. The linear kernel was established to be suitable as it resulted in the least RMSE among the others. At this stage, the suitable hyperparameters were also generated automatically and subsequently employed in the training of the dataset. The training is a process employed to enhance the prediction capabilities of the model. It can be regarded as trying to find the set of parameters for a specific function that can be used to make predictions given a set of independent variables. The accuracy of any model prediction most of the time depends on the training it goes through as well as the size of the training sample. In this study, a five-fold data partition was used to train the model. A lot of variables may satisfy the constraints hence, the training produced optimal model parameters. These are the minimum values for the respective parameters and were obtained after several iterations. Figures 2–4 present the hyperparameters selection and SVR learning scheme, the flowchart of model cross-validation and the steps in predicting values, respectively.

Model evaluation

This is performed with the validation data and at this stage, the performance of the model is critically examined. The performance of a model may not be satisfied after the initial training. This demands the altering of some parameters and the training and evaluation processes repeated until a good result is obtained. With the use of MATLAB for SVR, the parameters available for training were the epsilon values, the kernel scale, the box constraint and standardisation. The accuracy of the model, however, was always higher without standardisation. A margin of error was therefore allowed to avoid overfitting as an

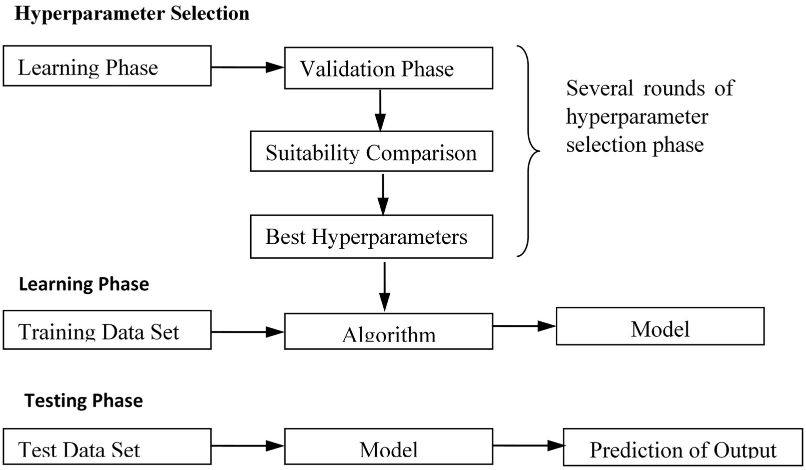

Figure 2.
Hyperparameters selection and SVR learning scheme

overfitting model strives hard to attain zero error. This makes the model too sensitive since it also fits outliers to produce a high variance and to prevent the model from missing relevant information about the relationship between the data. Therefore, an effort was made to attain a perfect model that was not too influenced and not too sensitive.

Model development and performance

The linear kernel was subsequently used in developing the model. According to Deleen and Bonsu (2002), contractors mostly employ suppliers' credit as an external source of financing equipment, and this can be attributed to the initial high cost associated with equipment. Windapo (2017) also stated that suppliers credit facilities improve company profits and cash flow. It can, therefore, be inferred that projects cannot be executed successfully without credit, irrespective of duration, because of payment irregularities that impact cash flow. This necessitated the inclusion of credit to loan and contractor finance to develop the model. Hence, the finance options settled for the model developed in the study were contractor finance, loan and credit.

The Mean Absolute Percentage Error (MAPE) was employed to evaluate the performance of the developed model. Results obtained from the MAPE expression is interpreted using Table 2. Liu and Pan (2012) reported that MAPE is the most effective estimation index and formulates related evaluation standards. The sensitivity of a model is an important contemplation in appraising the performance of that model. The sensitivity analysis of a study deals with how the uncertainty in the output of a mathematical model can be divided and assigned to different sources of uncertainty in its inputs. Since the projects used in the model development and validation were donor and internally generated funded projects, it is presumed that the funds for the projects were readily available. Therefore, a delay is assumed to be minimised. It has been reported that the parameters of SVR, that is epsilon (ε), cost or error penalty (C) and the kernel type influence prediction performance (Priyadarshini *et al.*, 2011; Ustuner *et al.*, 2015). Additionally, the sensitivity index can be used to determine the degree to which the independent variables impact the dependent variable (Hamby, 1994). In achieving this, the prediction of a specific variable is determined at its maximum and minimum values, keeping the others constant. This index has a value between 0 and $\pm$ 1. The sensitivity of a variable is determined with the closeness of the index to 1 (Hamby, 1994). The study, therefore, adopted the sensitivity index to

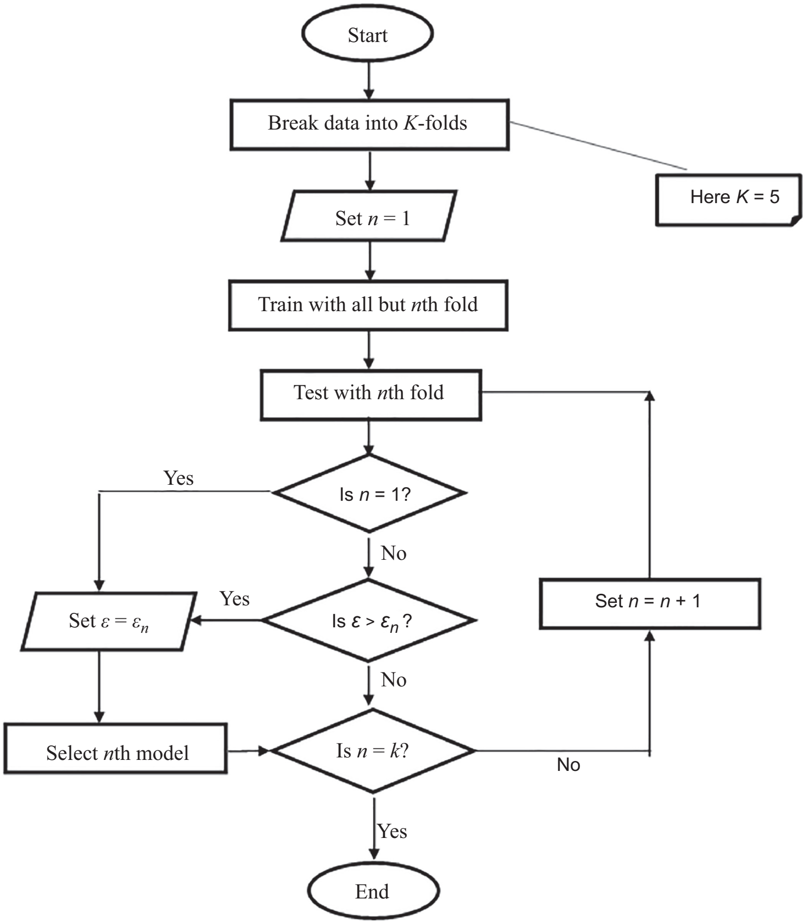

Figure 3.
Flowchart of model cross-validation

determine the degree to which the independent variables (cash-flow factors) impacted the dependent variable (profit).

Results

Completed institutional building projects with internally generated and donor funding were used for this study. The established cost contributions of the identified significant quantifiable cash-flow factors were subjected to 150 completed projects under study to determine the respective cost allocation to the projects. These consequently constituted the dataset for the model development and validation. It was identified that some of the projects had variations; hence, variation was included in the variables for model development and further predictions. Funds for the projects under study were presumed to be readily available; therefore, payment schedules were not unduly delayed. It can, therefore, be inferred that delay in the projects was minimised which subsequently had less influence on the variation and expected profit of the projects obtained.

Machine Learning Toolbox in MATLAB was used to develop the model, and this includes functions to perform various analyses on multivariate data. Support vector regression

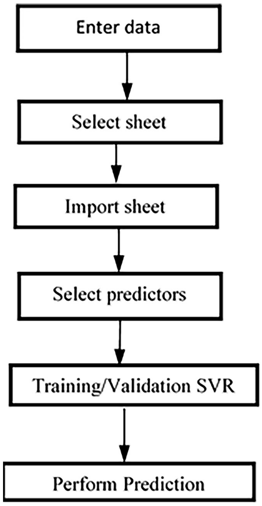

Figure 4. Steps in predicting values

MAPE	Interpretation
<10	Highly accurate forecasting
10–20	Good forecasting
20–50	Reasonable forecasting
>50	Inaccurate forecasting

Source(s): Liu and Pan (2012), José *et al.* (2013)

Table 2. Interpretation of MAPE values

analysis uses an optimal line, hyperplane drawn across the training data. This line is usually straight but might appear to be distorted when the number of dimensions constituting the hyperplane goes beyond three. The performances of the various kernels were assessed, and the linear kernel was settled on to be the best performing kernel for the model development as it generated the least RMSE.

The linear kernel was subsequently used in developing the model. In subjecting the financing option datasets to develop the model, the coefficients of the variables in the model shown in Table 3 were attained.

The coefficients generated and the independent variables are represented in $1 \times n$ and $n \times 1$ vectors or matrixes, respectively. The prediction which is a scalar (1×1) is the product of $1 \times n$ and $n \times 1$. Therefore, the general expression for the linear kernel which is the inner product plus a constant was thereafter performed to generate the model shown in equation (2).

$$\text{Profit} = \frac{1}{k}(\beta 1\text{ICS} + \beta 2\text{Vr} + \beta 3\text{Lb} + \beta 4\text{Icf} + \beta 5\text{Iln} + \beta 6\text{Icr} + \beta 7\text{LD} + \beta 8\text{DW} + b) \quad (2)$$

In assessing the measure of goodness of fit, a resulting value of 1 was obtained as the determinant (R^2) of the established model with a 1% margin of error. This indicates the closeness of the data to the hyperplane. This further explains that the model predictions perfectly fit the data and explains all the variability of the response data around its mean. Average predictive accuracy of 73.66% was achieved with all the variables on the different projects in validation. This result gives a MAPE value of less than 30% (26.18%). In reference to Table 2 as indicated by Liu and Pan (2012) and José *et al.* (2013), reasonable forecasting is said to be achieved when MAPE value range between 20 and 50%. Therefore, it can be inferred that; a reasonable prediction of profit is attained with this model.

The sensitivity of the model

From Table 4, it was identified that reasonable predictions were obtained with all the maximum and minimum values except for the minimum for labour which produces a good prediction when the sensitivity index was employed. Furthermore, the prediction results exhibited an upward trend of prediction from the maximum variable to the minimum. However, it was the reverse case for a loan. Labour and contractor finance generated SI values of −0.320 and −0.03, respectively as shown in Table 4. Upon introducing crediting, delay, and defective works separately, −0.002, −0.012, and −0.008 respectively, were the SI results obtained. Contrary to the trend of results, loans generated a positive SI value of 0.04. It can therefore be deduced that maximising this input variable whiles holding others constant demonstrated an enhancement in predicting profit. However, this is the reverse in the case of a loan. Although loans are associated with high interest paid on them, its acquisition improves the profit earnings.

Discussion

The study has been able to establish a novel profit prediction model using SVR with significant cash-flow factors as predictors. These predictors account for 100% of the variance in the profit generation as a result of a perfect correlation obtained from $R^2 = 1$. The approach adopted to generate the dataset for developing the model meant that there were no data outliers. Data outliers have a significant impact on the accuracy of the predictive model. However, the proposed model resulted in MAPE less than 30%, indicating a reasonable prediction. In Park and Kim's (2011) Bayesian profit predicting models of construction projects, technical factors were considered as predictors and these also accounted for at least

Table 3. Determined coefficients of independent variables of model

Variables	Symbol	Coefficients
Initial contract sum (Ics)	β_1	24498.22
Variation (Vr)	β_2	24798.70
Labour (Lb)	β_3	−13186.93
Interest on contractor finance (Icf)	β_4	−1300.02
Interest on loan (Iln)	β_5	−2351.68
Interest on credit (Icr)	β_6	−375.22
Loss due delay (Ld)	β_7	−1658.56
Defective works (Dw)	β_8	−1740.18
Bias (B)	B	−0.00207
Kernel scale (K)	K	172803.4

Table 4. Sensitivity index of respective variables

Variable	Labour	Loan	Contractor finance	Credit	Delay	Defective works
D_{max}	60.67	75.77	71.55	73.66	73.05	73.41
D_{min}	80.65	73.10	74.06	73.84	74.02	73.97
SI = D_{max}-D_{min} D_{max}	−0.329	0.04	−0.03	−0.002	−0.013	−0.008

76.9% of the variance in the dependent variable. However, this study failed to evaluate the sensitivity of the predictors. Additionally, El-Kholy (2014) developed a multi-objective fuzzy linear programming (FLP) model for cash-flow management that reduced the optimum value of final cash balance by a percentage between zero and 19% in FLP from the corresponding ideal value in the crisp linear programming (LP) model. This resulted in an optimum value of initial cash balance that increased between 6.4 and 15.6% in FLP from the corresponding ideal value in the LP. This study also failed to test for the sensitivity of the predictors. The multi-period dynamic model by Jiang (2012) permits operators to outline the cash-flow planning horizon, as well as forecast and maximise the cash balance for the planning horizon. Subsequently, the cash-flow model was established to be sensitive to a retained percentage, progress payment schemes. Huang (2007) further affirms the predictability and performance of SVR in the determination of the cost of complex products at the early phases of designing.

With reference to previous studies, the Choi *et al.* (2013) profit model considered labour productivity per construction worker whereas Jiang *et al.'s* (2011) model concentrated on financial market constraints such as long- and short-term loans from banks, gains on excess cash deposit and minimum cash reserves of a project. El-Kholy (2014) further modified Jiang *et al.'s* (2011) model and incorporated delay payment for only one period. The developed model in this study incorporated variables from El-Kholy's study with the inclusion of variation, initial contract sum and defective works.

Choi *et al.* (2013) employed predicted error sum of square (PRESS) and the sum of squared error (SSE) to evaluate the accuracy of the developed exponential profit model which considered labour productivity per construction worker. The study stated that, the closeness of PRESS to SSE indicates the model possesses significant predictability. However, if the difference is several times larger, then the model has an issue with validation. The values of PRESS and SSE obtained in the study were 3.265 and 2.819 respectively indicating a significant prediction. The accuracy of Park and Kim's (2011) model was evaluated with the differences of BRA model against the actual margin of profit ranged from 1.82 to 0.01%, while that of the multiple linear regression (MLR) ranged from 0.75 to 4.71%. The compared results clearly reflected the superiority of BRA over MLR. Bee-Hua's (2010) study also examined the combination of neural networks (NNs) and genetic algorithms (GAs) to forecast residential construction demand which resulted in reduction of average MAPE from about 6% to a mere 1%. However, the NNs and Gas models generated accurate forecasts, since their respective MAPE values consistently fell within the acceptable limit of 10%. This established an enhanced performance produced from the use of a hybrid technique. Additionally, Emsley *et al.'s* (2010) neural network model to predict total construction costs demonstrated the capability of the neural network to model the nonlinearity in the data. This was revealed in an average MAPE of 16.6% comparable to the traditional estimate which reported between 20.8 and 27.9%. More so, Hua and Pin's (2010) Boxjenkins model to predict the construction demand, price and productivity employed RMSE and MAPE to evaluate the model accuracy. All the developed models generated RMSE which were consistently smaller than the standard errors and MAPE also fell within the acceptable limit of 10% with demand model possessing the least.

The model developed in this study has the independent variables accounted for 100% variance explained in the dependent variable. It can be established that the SVR performs better in prediction accuracy than other techniques except for hybrid techniques. Although much work has not been conducted on SVR concerning productivity and profit, it can be inferred from these discussions that this study affirms the superiority of SVR in prediction and accuracy in predicting profit to enhance the productivity of the construction industry.

The utilisation of loans contributes to the early completion of projects for certification and onward payment. Therefore, this reduces capital lock-up should the project completion be delayed when funds are not available to finance for regular and prompt payments. It can be established that labour is sensitive to the prediction of profit compared to the other independent variables. Apart from the loan, it can be deduced that any upsurge of the variable will negatively impact on profit prediction. Sensitivity was also performed on variation since it is a variable employed to predict profit. This variable was employed in the development of the model because some of the projects used in the development and validation of the model had recorded significant variation. Since the coefficient of variation is positive, any addition or omission to a contract will impact on the profit positively or negatively respectively. About the sensitivity index results obtained, projects should be implemented with minimum labour, credit and contractor finance to enhance profit. However, financing with a loan can be encouraged to facilitate the early completion of projects to receive payment promptly to improve profit generation whenever the funds for payment are readily available.

Conclusion and recommendation

Construction productivity rate is the origin of the accurate estimating of the cost and time necessary to complete a project. Since enhanced productivity helps contractors and project owners to realise their profitability levels, productivity has been considered an important predictor of profitability. The threats posed by productivity to the profitability of construction organisations have called for several solutions from industry researchers across the globe. However, aside from productivity, cash flow has also been identified as a potential threat, with effective cash flow management considered as key for construction organisations to survive in the already competitive construction industry.

This study was conducted to develop and test the sensitivity of a Machine Learning Support Vector Regression Algorithm (SVRA) to predict construction project profit in Ghana. Historical data was used in the model development and validation. This data was obtained from 150 institutional projects that were executed within the past five years (i.e. 2014–2018). Eighty percent (80%) of the data from the 150 projects was used at hyperparameter selection and final training phases of the model development and the remaining 20% for testing the model. Data were analysed using MATLAB for Support Vector Regression. The findings from the study suggest that the predictions from the developed model perfectly fitted the data, thereby explaining all the variability of the response data around its mean. The findings further suggest that on average predictive accuracy of 73.66% was achieved with all the variables on the different projects in validation, with the developed SVR model being sensitive to labour and loan.

This study has led to the development of a Support Vector Regression Algorithm that will serve to prompt or warn contractors before the commencement and during the execution of projects about profit performance. There are significant theoretical and practical implications from the findings of this study. Theoretically, the study contributes to the ever-growing literature on the application of machine learning in predicting construction project profit. Literature has reported numerous models and techniques that have been developed and used in different contexts. The SVRA model developed in this study contributes its quota by potentially filling the profit measurement gap with the incorporation of variation, defective works, and labour to financial constraints, which have been the variables used in previous studies. Practically, the model should aid contractors to predict profit on completion before the commencement of projects. This will equip contractors seeking to enhance profit by helping them make informed and objective decisions towards maximisation of profit. It will provide contractors with relevant signals on the influence of any change in the significant variables used in the determination of profit because of the client (variations or payment issues), or the contractor (labour management).

Despite the useful outcome of the study, it was limited to only contractors who are well resourced and execute large volumes of works. It also only looked at quantifiable cash-flow factors and further considered a year duration of interest payment on the various sources of finance. It will be useful to conduct studies to evaluate the accuracy of the model over the long term when payment is unduly delayed with the incorporation of interest paid on delayed payment. Again, similar studies on projects undertaken by smaller contractors who are not well resourced could unearth other significant cash-flow factors from their perspective, affecting profit predictions by small contractors who are in the majority, especially when one considers construction industries in developing countries.

References

Achenef, T.K. (2016), "Factors affecting profitability: the case of Ethiopian selected private construction companies", A thesis submitted to, St. Mary's University School of Graduate Studies, Addis Ababa.

Adjei, E.A., Fugar, F.D.K., Adinyira, E., Edwards, D.J. and Pärn, E.A. (2018), "Exploring the significant cash flow factors influencing building projects profitability in Ghana", *International Journal of Construction Engineering and Management*, Vol. 7 No. 1, pp. 35-46.

Adjei, E.A., Fugar, F.D.K. and Adinyira, E. (2019a), "Establishing the cost contribution of significant cash flow factors impacting on building projects' profitability", in Marshall-Ponting, A., Ji, Y., Keraminiyage, K.P., Lee, A., Poppelreuter, T., Swan, W., Udeaja, C.E. and van Dijk, H. (Eds), *2019, Conference Proceedings of the 14th International Postgraduate Research Conference 2019: Contemporary and Future Directions in the Built Environment*, ISBN 9781912337309, University of Salford, Salford, pp. 25-37.

Adjei, E.A., Fugar, F.D.K. and Adinyira, E. (2019b), "A review of appropriate tool to predict profitability of building projects using established significant cash flow factors", in Marshall-Ponting, A., Ji, Y., Keraminiyage, K.P., Lee, A., Poppelreuter, T., Swan, W., Udeaja, C.E. and van Dijk, H. (Eds), *2019, Conference Proceedings of the 14th International Postgraduate Research Conference 2019: Contemporary and Future Directions in the Built Environment*, ISBN 9781912337309, University of Salford, Salford, pp. 38-48.

Akintoye, A. and Skitmore, M. (1991), "Profitability of UK construction contractors", *Construction Management and Economics*, Vol. 9 No. 4, pp. 311-325.

Arditi, D., Koksal, A. and Kale, S. (2000), "Business failures in the construction industry", *Engineering, Construction and Architectural Management*, Vol. 7, pp. 120-132.

Asante, J.A. (2014), "Financial distress related causes of project delays in the Ghanaian construction industry", Unpublished, available at: http://dspace.knust.edu.gh/bitstream/123456789/6465/1/ASANTE%20JOYCELINE%20ANNOA.pdf (Accessed July 2019).

Avetisyan, E.H., Tesha, D.N.G.A.K., Lello, D.S. and Mtitu, F.S. (2020), "Profit maximization strategies employed by the small and medium size building contractors in Dar-Es-Salaam, Tanzania", *International Journal of Engineering and Management Research*, Vol. 10 No. 1, doi: 10.31033/ijemr.10.1.17.

Awad, M. and Khanna, R. (2015), "Support vector regression", in *Efficient Learning Machines*, Apress, Berkeley, CA, pp. 67-80.

Aziz, R.F. (2013), "Optimising strategy for repetitive construction projects within multi-mode, resources", *Alexandria Engineering Journal*, Vol. 52, pp. 67-81, doi: 10.1016/j.aej.2012.11.003.

Barbosa, P.S.F. and Pimentel, P.R. (2010), "A linear programming model for cash flow management in the Brazilian Construction Industry", *Construction Management and Economics*, Vol. 19 No. 5, pp. 469-479, doi: 10.1080/01446193.2001.9709623.

Bee-Hua, G. (2010), "Evaluating the performance of combining neural networks and genetic algorithms to forecast construction demand: the case of the Singapore residential sector", *Construction Management and Economics, Volume*, Vol. 18 No. 2, pp. 209-217, doi: 10.1080/014461900370834.

Bolek, M. and Wiliński, W. (2012), "The influence of liquidity on profitability of polish construction sector companies", *Financial Internet Quarterly e-Finance*, Vol. 8 No. 1, pp. 38-52, available at: https://www.econstor.eu/bitstream/10419/66742/1/721357075.pdf.

Bowman, C. (2014), "Where does profit come from?", available at: https://ssrn.com/abstract=2484460 or http://dx.doi.org/10.2139/ssrn.2484460 (accessed 21 August 2014).

Buertey, J.I.T. and Adjei-Kumi, T. (2012), "Cash flow forecasting in the construction industry: the case of Ghana", *Pentvars Business Journal*, Vol. 6 No. 1, pp. 65-81, available at: http://pentvarsjournal.com/articles//january-2012/cashflow.pdf.

Chendroyaperumal, C. (2009), "Profit theories: modern and Indian", (Working Paper), available at: https://ssrn.com/abstract=1424237 (accessed December 2020).

Cheng, M.-Y., Hoang, N.-D. and Wu, Y.-W. (2015), "Cashflow prediction for construction project using a novel adaptive time-dependent least squares support vector machine inference model", *Journal of Civil Engineering and Management*, Vol. 21 No. 6, pp. 679-688.

Choi, K., Haqui, M., Lee, H., Cho, Y. and Kwak, Y. (2013), "Macroeconomic labour productivity and its impact on firm's profitability", Palgrave Macmillan, *Journal of the Operational Research Society*, Vol. 64, pp. 1258-1268, doi: 10.1057/jors.2012.157.

Deelen, L. and Bonsu, K.O. (2002), *Equipment Finance for Small Contractors in Public Work Programmes*, International Labour Organization, Geneva.

El-Kholy, A.M. (2014), "A multi-objective fuzzy linear programming model for cash flow managemen", *International Journal of Engineering Research and Applications*, Vol. 4 No. 8, pp. 152-163, (Version 3), available at: http://www.ijera.com/papers/Vol4_issue8/Version%203/W4803152163.pdf (accessed July 2018).

Emsley, M.W., Lowe, D.J., Duff, A.R., Anthony, H. and Hickson, A. (2010), "Data modelling and the application of a neural network approach to the prediction of total construction costs", *Construction Management and Economics*, Vol. 20 No. 6, pp. 465-472, doi: 10.1080/01446190210151050.

Enshassi, A., Mohamed, S. and Abushaban, S. (2009), "Factors affecting the performance of construction projects in the Gaza Strip", *Journal of Civil Engineering and Management*, Vol. 15 No. 3, pp. 269-280, doi: 10.3846/1392-3730.2009.15.269-280.

Ghojogh, B., Samad, M.N. and Mashhadi, S.A. (2019), "Feature selection and feature extraction in pattern analysis: a literature review", available at: https://www.researchgate.net/publication/332960460_Feature_Selection_and_Feature_Extraction_in_Pattern_Analysis_A_Literature_Review (accessed November 2020).

Guajardo, J., Weber, R. and Miranda, J. (2006), "A forecasting methodology using support vector regression and dynamic feature selection", *Journal of Information and Knowledge Management*, Vol. 5 No. 4, pp. 329-335, doi: 10.1142/S021964920600158.

Gundecha, M.M. (2013), "Study of factors affecting labour productivity at building construction project in the USA", *Web Survey*, available at: http://library.ndsu.edu/tools/dspace/load/?file=/repository/bitstream/handle/10365/22772/Gundecha_Mahesh.pdf?sequence=1 (accessed July 2017).

Halim, M.S.B.A. (2014), "Unrealistic profit experienced by Bumiputera entrepreneurs in the Malaysian construction industry", *Advances in Environmental Biology*, Vol. 8 No. 9, pp. 489-496, 1995-0756.

Hamby, D.M. (1994), "A review of techniques for parameter sensitivity analysis of environmental models", *Environmental Monitoring and Assessment*, Vol. 32, pp. 135-154, available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.224.7813&rep=rep1&type=pdf.

Harris, F. and McCaffer, R. (2005), *Modern Construction Management*, Blackwell Publishing, New Jersey, p. 118138.

Haseeb, M., Bibi, A. and Rabbani, W. (2011), "Problems of projects and effects of delays in the construction industry of Pakistan", *Australian Journal of Business and Management Research*, Vol. 1, pp. 41-50.

Hillebrandt, P.M. (2000), *Economic Theory and the Construction Industry*, Palgrave Macmillan, New York.

Hillebrandt, P.M., Cannon, J. and Lansley, S. (1995), *The Construction Company in and Out of Recession*, Macmillan, London.

Hira, Z.M. and Gillies, D.F. (2015), "A review of feature selection and feature extraction methods applied", *Advances in Bioinformatics*, Vol. 2015, pp. 1-13, 198363, doi: 10.1155/2015/198363.

Hofstrand, D. (2013), "Cash flow and profitability are not the same", *Ag Decision Maker, File C5-213*, available at: www.extension.iastate.edu/agdm (accessed 15 January 2021).

Hua, G.B. and Pin, T.H. (2010), "Forecasting construction industry demand, price and productivity in Singapore: the Boxjenkins approach", *Construction Management and Economics*, Vol. 18 No. 5, pp. 607-618, doi: 10.1080/014461900407419.

Huang, G. (2007), "Cost modelling based on support vector regression for complex products during the early design phases", Unpublished.

Hughes, W., Hillebrandt, P. and Murdoch, J. (2000), "The impact of contract duration on the cost of cash retention", *Construction Management and Economics*, Vol. 18, pp. 11-14, doi: 10.1080/014461900370906.

Jaafar, M. and Abdul-Aziz, A.R. (2005), "Resource-based view and critical success factors: a study on small and medium sized contracting enterprises (SMCEs) in Malaysia", *International Journal of Construction Management*, Vol. 5 No. 2, pp. 61-77.

Jiang, A. (2012), "Negotiating construction contracts through practical cash flow planning and analysis model", *International Journal of Construction Management*, Vol. 12 No. 2, pp. 23-33, doi: 10.1080/15623599.2012.10773188.

Jiang, A., Issa, R.R.A. and Malek, M. (2011), "Construction project cash flow planning using the Pareto optimality efficiency network model", *Journal of Civil Engineering and Management*, Vol. 17 No. 4, pp. 510-519, doi: 10.3846/13923730.2011.604537.

José, J.M.M., Alfonso, P.P., Albert, S.A. and Berta, C.B. (2013), "Using the R-MAPE index as a resistant measure of forecast accuracy", *Psicothema*, Vol. 25 No. 4, pp. 500-506, doi: 10.7334/psicothema2013.23.

Kaka, A.P. and Cheetham, D.W. (1997), "The effect of some tendering and payment strategies on contractors' financial performance", in *13th Annual ARCOM Conference*, pp. 15-17.

Khosrowshahi, F. and Kaka, A. (2007), "A decision support model for construction cash flow management", *Computer-Aided Civil and Infrastructure Engineering*, Vol. 22, pp. 527-539, available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8667.2007.00508.x/epdf?r3_referer=wolandtracking_action=preview_clickandshow_checkout=1andpurchase_referrer=onlinelibrary.wiley.comandpurchase_site_license=LICENSE_DENIED.

Kolisetty, V. and Rajput, D.S. (2020), "A review on the significance of machine learning for data analysis in big data", *Jordanian Journal of Computers and Information Technology (JJCIT)*, Vol. 06 No. 1, doi: 10.5455/jjcit.71-1564729835.

Langford, D.A. and Chan, K.K. (1987), "Labour only subcontracting – the changing position", *Construction Management and Economics*, Vol. 5 No. 1, pp. 31-33.

Lee, F.B. (2009), "Factors affecting the profitability of construction companies in Hong Kong", Unpublished, doi: 10.5353/th_b4727769 (accessed November 2017).

Listiani, M. (2009), "Support vector regression analysis for price prediction in a car leasing application", Unpublished, available at: https://www.ifis.uni-luebeck.de/~moeller/publist-sts-pw-and-m/source/papers/2009/list09.pdf (accessed November 2017).

Liu, S. and Pan, N. (2012), "'Construction firms' operation management performance measurement model", *Journal of Statistics and Management System*, Vol. 15 No. 6, pp. 627-642, doi: 10.1080/09720510.2012.10701645.

Liu, Y., Zayed, T. and Li, S. (2009), "Cash flow analysis of construction projects", *2nd International/8th Construction Specialty Conference*, pp. 1-8, available at: https://pdfs.semanticscholar.org/72b1/55e3c8543315719e6517c2e549560f7fa55b.pdf (accessed July 2017).

Long, R.J. (2015), *Contract Scheduling Provisions*, Long International, Littleton, CO.

Mahamid, I. (2011), “Causes of contractors’ failure: contractors’ view”, *2nd International Conference on Construction and Project Management IPEDR*, Vol. 15. ISBN: 0047735945.

Majer, R., Ellingerová, H. and Gašparik, J. (2020), “Methods for the calculation of the lost profit in construction contracts”, *Buildings*, Vol. 10, pp. 1-13.

Malik, H. (2011), “Determinants of insurance companies profitability: an analysis of insurance sector of Pakistan”, *Academic Research International*, Vol. 1 No. 3, pp. 315-321.

Marchal, J. (1951), “The construction of a new theory of profit”, *The American Economic Review*, Vol. 41 No. 4, pp. 549-565.

McGuigan, J.R., Moyer, R.C. and Harris, F.H. (2017), *Managerial Economics: Applications, Strategy, and Tactics*, 14th ed., South-Western, Cengage Learning, London, ISBN-13: 978-1-4390-7923-2.

Mohsin, M.A., Alnuaimi, A. and Tobi, S.A. (2014), “Contractual implications of cash flow on owner and contractor in villa construction projects”, *International Journal of Research in Engineering and Technology*, Vol. 3 No. 4, pp. 442-447, doi: 10.15623/ijret.2014.0304079.

Moon, S., Ham, N., Kim, S., Hou, L., Kim, J. and Kim, J. (2020), “Fourth industrialisation-oriented offsite construction: case study of an application to an irregular commercial building”, *Engineering Construction and Architectural Management*, Vol. 27 No. 9, pp. 2271-2286, doi: 10.1108/ECAM-07-2018-0312.

Odeyinka, H.A., Lowe, J. and Kaka, A.P. (2013), “Artificial neural network cost flow risk assessment model”, *Construction Management and Economics*, Vol. 31 No. 5, pp. 423-439, doi: 10.1080/01446193.2013.802363.

Olatunji, A.A. (2010), “Influences on construction project delivery time”, A Thesis Submitted at, The Nelson Mandela Metropolitan University, Port Elizabeth.

Omopariola, E.D., Windapo, A., Edwards, D.J. and Thwala, W.D. (2020), “Contractors’ perceptions of the effects of cash flow on contruction projects”, *Journal of Engineering, Design and Technology*, Vol. 18 No. 2, pp. 308-325.

Owalabi, J.D., Amusan, L.M., Oloke, C.O., Olusanya, O. and Tunji-Olayeni, P. (2014), “Causes and effects of delay on project construction delivery time”, *International Journal of Education and Research*, Vol. 2 No. 4, pp. 197-208, available at: https://www.ijern.com/journal/April-2014/19.pdf (accessed July 2017).

Oyewobi, L.O., Windapo, A.O. and Cattell, K.S. (2014), *Competitiveness of Construction Organisations in South Africa*, Construction Research Congress, ASCE, Atlanta, pp. 2063-2073.

Pansegrauw, M. (2019), “Revenue is vanity, profit is sanity but cash is king”, available at: https://www.procurementexpress.com/blog/podcast/revenue-is-vanity-profit-is-sanity-but-cash-is-king/ (accessed 21 January 2021).

Park, S. and Kim, S. (2011), “A strategy Bayesian Model to predict profit of construction projects”, *Architectural Research*, Vol. 13 No. 3, pp. 49-56.

Park, H.K., Han, S.H. and Russell, J.S. (2005), “Cash flow forecasting model for general contractors using moving weights of cost categories”, *Journal of Management in Engineering*, Vol. 21 No. 4, pp. 164-172, doi: 10.1061/(ASCE)0742-597X(2005)21:4(164).

Peterson, S.J. (2009), *Construction Accounting and Financial Management*, 2nd ed., Pearson, London.

Pham, D.T. and Afify, A.A. (2005), “Machine-learning techniques and their applications in manufacturing”, *Journal of Engineering Manufacture*, Vol. 219 No. 5, pp. 395-412, doi: 10.1243/095440505X32274.

Priyadarshini, D., Acharya, M. and Mishra, A.P. (2011), “Link load prediction using support vector regression and optimisation”, *International Journal of Computer Applications*, Vol. 24 No. 7, pp. 22-25.

Ramachandra, T. and James Rotimi, J.O.B. (2011), “The nature of payment problems in the New Zealand construction industry”, *Australasian Journal of Construction Economics and Building*, Vol. 11 No. 2, pp. 22-33, doi: 10.5130/AJCEB.v11i2.2171.

Reich, R. (2000), *The Wealth of Nations*, The Modern Library United State of America, New York.

RICS (2011), "Cash flow forecasting", RICS Practice Standards, UK 1st ed., Guidance note, ISBN 978 1 84219 687 8.

Roffo, G. (2018), "Feature selection library (MATLAB Toolbox)", available at: https://arxiv.org/pdf/1607.01327v5.pdf.

Sambasivan, M. and Soon, Y.W. (2007), "Causes and effects of delays in Malaysian construction industry", *International Journal of Project Management*, Vol. 25, pp. 517-526, doi: 10.1016/j.ijproman.2006.11.007.

Sanyal, R. (2019), "Profit theory 1", Book Chapter In Book: Planned Economy and growth Publisher: S Chand, available at: https://www.researchgate.net/publication/335023002_Profit_Theory_1.

Sapankevych, N. and Sankar, R. (2009), "Time series prediction using support vector machines: a survey", *IEEE Computational Intelligence Magazine*, Vol. 4 No. 2, pp. 24-38.

Seo, K., Soh, J. and Sharma, A. (2018), "Do financial constraints affect the sensitivity of investment to cash flow? New evidence from franchised restaurant firms", *Tourism Economics*, Vol. 24 No. 6, pp. 645-661.

Seppälä, C.R. (2005), *Contractors Claims Under the FIDIC Contracts for Major Works*, International Construction Contracts and Dispute Resolution Co-Hosted by ICC and FIDIC In partnership with The Cairo Regional Centre for International Commercial Arbitration, Unpublished.

Snyman, T. and Smallwood, J. (2017), "Improving productivity in the business of construction", *Procedia Engineering*, Vol. 182, pp. 651-657.

Spedding, A.H. (1977), "Trends and prices in primary school building", PhD thesis, University of Leeds, Leeds.

Su, Y. and Lucko, G. (2015), "Synthetic cash flow model with singularity functions for unbalanced bidding scenarios", *Construction Management and Economics*, Vol. 33 No. 1, p. 3554, doi: 10.1080/01446193.2015.1012527.

Tracy, J.A. and Tracy, T. (2015), "The relationship between cashflow and profit in business", available at: https://www.dummies.com/business/accounting/the-relationship-between-cash-flow-and-profit-in-business-2/ (accessed 15 January 2021).

Tunji, S.T., Adebayo, O.S. and Tolulope, O.O. (2015), "Impact of gearing on performance of companies", *Arabian Journal of Business and Management Review*, Vol. 3 No. 1, pp. 68-80.

Usman, N., Sani, A. and Tukur, I.R. (2016), "Exploring the use of financial capacity as apredictor of construction company corporate performance: evidence from South Africa", *International Journal of Research and Development Organisation*, Vol. 2 No. 2, pp. 2455-2466.

Ustuner, M., Sanli, F.B. and Dixon, B. (2015), "Application of support vector machines for land use classification using high-resolution RapidEye images: a sensitivity analysis", *European Journal of Remote Sensing*, Vol. 48, pp. 403-422, doi: 10.5721/EuJRS20154823.

Vapnik, V. (1998), "The support vector method of function estimation", in *Nonlinear Modeling*, Springer, Boston, MA, pp. 55-85.

Vercellone, C. (2008), *The New Articulation of Wages, Rent and Profit in Cognitive Capitalism*, The Art of Rent, Queen Mary University School of Business and Management, London, ⟨halshs-00265584⟩.

Verma, N., Das, S. and Srivastava, N. (2016), "Multiple kernel support vector regression with higher norm in option pricing", *International Journal of Applied Engineering Research*, Vol. 11, No. 5, pp. 3168-3174.

Windapo, A. (2017), "Determinants of the sustainability and growth of construction contractors", pp. 1-13, doi: 10.20944/preprints201707.0091.v1.

Wright, M.G. (1977), "Profit and competition: profitability", *Building Technology and Management*, s.l., s.n, pp. 4-6.

Yasamis, F., Arditi, D. and Mohammadi, J. (2002), "Assessing contractor quality performance", *Construction Management and Economics*, Vol. 20 No. 3, pp. 211-223, doi: 10.1080/01446190110113693.

Yean, F., Ling, Y. and Liu, M. (2005), "Factors considered by successful and profitable contractors in mark-upsize decision in Singapore", *Building and Environment*, Vol. 40, pp. 557-1565, doi: 10.1016/j.buildenv.2004.12.001.

Zhu, F., Hu, H., Xu, F. and Tang, N. (2019), "Predicting profit performance of international construction projects", *Proceedings of the 2019 IEEE IEEM Conference*, pp. 445-449.

Corresponding author
Emmanuel Adinyira can be contacted at: eadinyira.cap@knust.edu.gh

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**