

Investigation and Comparative Analysis of Learning Curve Models on Construction Productivity: The Case of Caisson Fabrication Process

Panagiota Ralli¹, Antonios Panas², John-Paris Pantouvakis³, and Dimitrios Karagiannakidis⁴

¹Mechanical Engineer, M.Sc., Department of Projects Contracts & Procurement, Building Infrastructures S.A., Moudrou 12 str., 11146 Galatsi, Athens, Greece. E-mail: ralligiota@gmail.com

²Civil Engineer, Ph.D., Centre for Construction Innovation, National Technical University of Athens, Zografou Campus, Iroon Polytechniou 9 str., 15780 Zografou, Athens, Greece. E-mail: antpanas@gmail.com (corresponding author).

³Professor, Centre for Construction Innovation, National Technical University of Athens, Zografou Campus, Iroon Polytechniou 9 str., 15780 Zografou, Athens, Greece. E-mail: jpp@central.ntua.gr

⁴Civil Engineer, M.Sc., Aristotle University of Thessaloniki, Moudrou 30 str., 11146 Galatsi, Athens, Greece. E-mail: di.karagiannakidis@gmail.com

Project Management

Best Paper of EPPM 2019 Conference

Received February 28, 2020; revised April 22, 2020; accepted April 26, 2020

Available online May 24, 2020

Abstract: Learning curves in construction operations analysis is deemed as one of the main factors that determine the variation of on-site productivity and is always taken into account during the planning and estimation stage. This research attempts the assessment of learning curve models' suitability for the effective analysis of the learning phenomenon for construction operations that are fairly complicated concerning a floating caisson fabrication process for a large-scale marine project, using productivity data. This paper investigates the role of published learning curve models (i.e. Straight-line or Wright; Stanford "B"; Cubic; Piecewise or Stepwise; Exponential) by comparing their outcomes through the use of both unit and cumulative productivity data. There are two main research objectives: first, the model best fitting historical productivity data of construction activities that have been completed are investigated, while secondly, an attempt is made to determine which model better predicts future performance. The less actual construction data deviate from each model's yielded results, the better their suitability. In the case of unit data, the cubic model fits better historical data, while in the case of future predictions, the Stanford "B" model provides better results. Respectively, the Cubic model yields better results when using cumulative data on historical data and the Straight-line model predicts in a more reliable fashion future performance. Possible extensions could be developed in the area of future performance predictions, by adopting different data representation techniques (e.g. moving/exponential weighted average) or by including other (non-classic) learning curve models (e.g. DeJong, Knecht, hyperbolic models).

Keywords: Caisson, construction productivity, learning curve models, learning curves, marine projects, repetitive activities, statistical analysis.

Copyright © Association of Engineering, Project, and Production Management (EPPM-Association).
 DOI 10.2478/jeppm-2020-0024

1. Introduction

In project management, assessing and improving the productivity of construction activities is a fundamental challenge. The estimation of construction productivity takes into account several factors that reflect the managerial perspective and philosophy of the project personnel (Panas and Pantouvakis, 2010; Shan et al., 2011). One of the basic factors that affect productivity is the repetitive nature of construction activities taking place in construction projects. Therefore, it has been observed that when executing repetitive construction activities (e.g.

high-rise building construction) each subsequent production cycle may be improved from a productivity stance due to the learning phenomenon that is developed in relation to the resources that are deployed in the project (Thomas et al., 1986; Panas and Pantouvakis, 2014; Pellegrino and Costantino, 2018; Srour et al., 2018). In other words, the productivity of repetitive tasks is improved as the experience of the deployed crews is increased (Pellegrino et al., 2012). The required time (man-hours) for the completion of repetitive construction activities is decreased, as the repetitions increase, since (i)

the crews are familiarized with the nature of the works, (ii) the coordination of the mechanical equipment and the crews is improved, (iii) the project management discipline is enhanced, (iv) more efficient techniques and construction methods are implemented, (v) more effective logistics management methods are followed and (vi) the project scope is narrowed, thus limiting the need for additional corrective activities (Thomas et al., 1986; Zahran et al., 2016). Within that framework, the learning phenomenon or learning curve effect expresses the influence of the human factor, namely the contribution of the deployed crews' skill and experience in construction productivity.

Taking into account learning phenomena when studying construction productivity enhances the accuracy of time and cost management (Lutz et al., 1994; Ammar and Abdel-Maged, 2018), improves project control and programming (Pellegrino et al., 2012), as well as provides the required scientific evidence for claiming lost workhours (Thomas, 2009). The impact of the learning effect on labor productivity is graphically expressed by the learning curves, which are quantitatively analyzed by mathematical models (Pellegrino and Costantino, 2018) known as the learning curve models.

However, learning curve studies have also been criticized for having several limitations such as oversimplification of the construction process and implementation of a one-dimensional research approach where the analysis is based on a single learning model to interpret the actual data (Jarkas and Horner, 2011; Jarkas, 2016). More specifically, the vast majority of learning curve studies in construction uses the straight-line or Wright model, thus limiting the presented results' scope and possibly ignoring the effect of other learning parameters on the investigated construction process. In that view, this research intends to conduct a comparative analysis of five (5) established and widely acceptable learning curve models with the intent to interpret a relatively complex construction process relating to the realization of a large-scale marine infrastructure project. The purpose is the examination of each model's suitability to interpret historical productivity data and predict future performance, in order to provide project management executives with the necessary information to reach critical project decisions (e.g. increase/decrease of project resources deployment). It is, to the authors' best knowledge, the first research attempt to scrutinize thoroughly learning theory concepts' implementation of marine works from a productivity stance.

2. Literature Review

2.1. Learning Curve Theory

The theory of learning curves stems from the aircraft industry where T.P. Wright was the first one to implement it in 1936 (Jarkas, 2016; Ugulu and Allen, 2018), when he developed the so-called "Wright model" (Srouf et al., 2018). The latter predicted productivity improvements due to crew skills based on acquired experience (Jarkas, 2016).

In general, learning curves are used for the graphical representation of the period, the cost and/or the labour hours that are necessary in order to complete "sufficiently complex" construction operations (Everett and Farghal, 1994). The learning curve theory suggests that the required time (labour hours) for the production of a single

unit (e.g. a floor of a high-rise building) is incrementally decreasing as a percentage of the time that was demanded the production of the previous unit (UN, 1965; Jarkas and Horner, 2011). This percentage is called "learning rate" and is a characteristic variable for the extent of the learning phenomenon in single construction activity (Thomas et al., 1986). From a mathematical point of view, the learning rate coincides with the inclination of the learning curve. The learning phenomenon becomes more intense as the value of the learning rate is reduced, since each subsequent production cycle is a smaller percentage of the time required for the previous production cycle. For instance, when the learning rate equals 80%, then the required labour-hours for the production of a single unit is 20% less than the time needed for the production of the previous unit (Pellegrino et al., 2012). If an activity presents a learning rate equal to 100%, then no learning phenomenon is developed for that specific task (Lutz et al., 1994; Jarkas, 2016).

In principle, a learning curve, when plotted on logarithmic coordinates, is characterized by three sections, as shown in Fig. 1 (Thomas et al., 1986; Lutz et al., 1994; Couto and Teixeira, 2005). The first segment constitutes the operation-learning phase, during which the productivity is increased rapidly due to prior experience, as well as the crew's familiarization with the project's nature. The second segment represents that routine-acquiring phase, during which incremental productivity enhancement is achieved through the improvement of construction methods and organization. The third and last segment is the "typical" or "standard" production phase, during which the learning rate is constant and no further learning phenomena are observed. Point x_{p1} marks the completion of the "prior experience phase" and point x_{p2} coincides with the beginning of the "leveling off" phase. The latter is also called the "standard production point" as no significant further improvement of productivity is observed beyond that point (Thomas et al., 1986). In Fig. 1, unit X is produced after a specific amount ("Y") of cumulative cost, man-hours or time.

A fundamental prerequisite for the learning curve theory application is to execute repetitive, non-interrupted and non-differentiated activities, without delays due to shortages in materials delivery or specific guidelines (Thomas et al., 1986; Pellegrino et al., 2012).

In terms of the applied analytical tools, most published research in learning curve productivity analysis adopts the statistical approach for the elaboration of field data (Thomas et al., 1986; Everett and Farghal, 1997; Couto and Teixeira, 2005; Pellegrino et al., 2012; Ammar and Samy, 2015; Srouf et al., 2016), thus the same approach has been implemented in the current research as well.

Learning curve theory has been applied for modeling a variety of construction activities, such as (i) reinforced concrete buildings construction (Couto and Teixeira, 2005; Pellegrino et al., 2012), (ii) off-site fabrication of pre-cast concrete piles (Hinze and Olbina, 2009), (iii) rebar steel and formwork installation (Jarkas, 2010; Jarkas and Horner, 2011; Nguyen and Nguyen, 2013, Khanh and Kim, 2014), (iv) cell-shaped concrete caissons construction (Panas and Pantouvakis, 2014), (v) gas pipelines construction (Ammar and Samy, 2015) and (vi) megaprojects construction (Everett and Farghal, 1997; Naresh and Jahren, 1999).

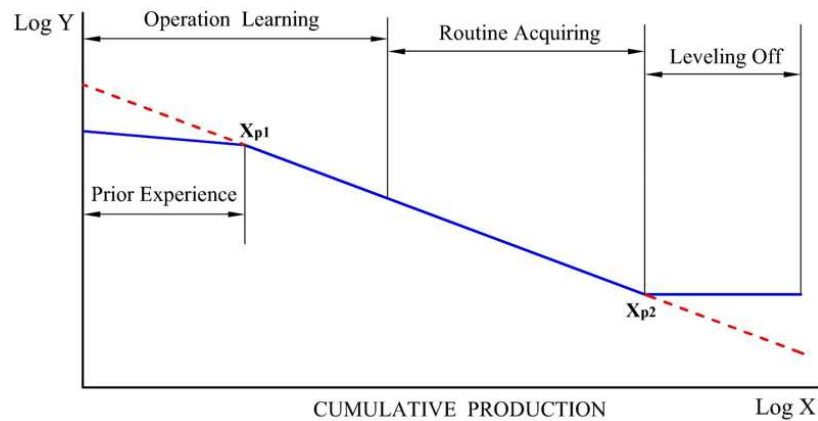


Fig. 1. Hypothetical Learning Curve (adopted from Thomas et al., 1986)

Regardless of the specific applications, the greatest value of learning curves for managers and construction engineers is their predictive capability of future productivity, rather than the assessment of historical productivity data (Farghal and Everett, 1997).

2.2. Data Representation

The researchers, applying learning curve theory, have to represent learning curve data, using one of the following four techniques: i) unit data, ii) cumulative average data, iii) moving average and iv) exponentially weighted average (Everett and Farghal, 1997; Mályusz and Pém, 2013; Ammar and Samy, 2015). A more detailed description follows in the next paragraph.

Unit data represents the time or cost required to produce a unit (e.g. building floor) or to finalize a specific production process in relation to the unit number or cycle number. They replicate the exact result of a specific repetitive activity (Everett and Farghal, 1997; Ammar and Samy, 2015). Cumulative average data represents the average time or cost required to produce or complete a given number of units or cycles (Everett and Farghal, 1997). It is computed as a fraction having as the numerator the total construction time or cost for a given number of units (or cycles) and as the denominator the completed units (or cycles) (Hinze and Olbina, 2009; Panas and Pantouvakis, 2014). A different version of the cumulative average is the moving average, which takes into account only recent data in the analysis (Everett and Farghal, 1997). The time frame for the data inclusion depends on the analyst's preference, leading either to the unit or cumulative data on each marginal situation (Mályusz and Pém, 2013). The integration of the most recent data and the previous average into one calculation results in the estimation of the exponentially weighted average. (Everett and Farghal, 1997).

In the construction industry, most researchers use the cumulative average technique for developing prediction models. However, it should be noted that the same dataset is used for representing data from all presented techniques.. For more detailed information about data representation, the reader is referred to Everett and Farghal (1997) and Mályusz and Pém (2013).

2.3. Learning Curve Models

The learning curve phenomenon is studied through the use of specific mathematical models, which interpret the variation of productivity in relation to critical factors such

as the number of units. These models quantify productivity improvements due to the execution of repetitious construction processes by either predicting or measuring performance (Jarkas, 2016).

Although different learning curve models have been presented in published pertinent research (Arditi et al., 2001; Wong et al., 2007; Srouf et al., 2016), the majority of researchers (Thomas et al., 1986; Everett and Farghal, 1994; Lutz et al., 1994; Couto and Teixeira, 2005; Ammar and Samy, 2015; Lee et al., 2015; Jarkas, 2016; Ammar and Abdel-Maged, 2018) refer to five main types of learning curve models as follows: (a) Straight-Line or Wright, (b) Stanford "B", (c) Cubic, (d) Piecewise or Stepwise and (e) Exponential models. These models are graphically represented in Fig. 2 and their detailed description follows in the next paragraphs.

2.3.1. Straight-Line model (or Wright model)

The Straight-line model was developed in 1936 by Wright and its purpose was the identification of cost affecting factors in the aircraft manufacturing sector (Hijazi et al., 1992). Its name is derived from the fact that a plot on a logarithmic scale forms a straight line (Ammar and Abdel-Maged, 2018; Lee et al., 2015). Due to its simplicity and its ability to provide acceptable precision (Srouf et al., 2018), it is most often selected by construction practitioners (Hinze and Olbina, 2009; Jarkas, 2016). The Straight-line model comes in two versions which are differentiated on whether they use the unit or cumulative average data. The mathematical equation is as follows (Glock et al., 2019):

$$Y = A * X^{-n} \quad (1)$$

where: Y= unit or cumulative average cost, man-hours or time to complete the X_{th} unit; A= cost, man-hours or time required for the first unit; X=unit number; n=slope of the logarithmic curve which takes values from zero to one. The Eq. (1) in its logarithmic format is presented below:

$$\log Y = \log A - n * \log X \quad (2)$$

The learning rate "L" (expressed as a percentage) can be derived from the slope of the logarithmic form as follow (Srouf et al., 2016):

$$L = 2^{-n} \text{ or } n = -\frac{\log L}{\log 2} \quad (3)$$

The main conceptual assumption in this model is that the learning rate does not fluctuate at all during the execution of the activity (Thomas et al., 1986).

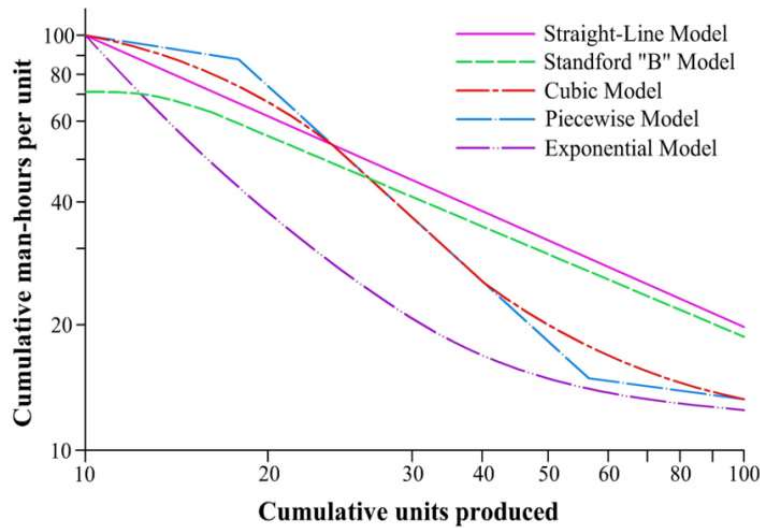


Fig. 2. Shape of Various Learning Curve Models (adopted from Thomas et al., 1986)

2.3.2. Stanford "B" model

The Stanford Research Institute proposed the Stanford "B" model in the 1940s. It was developed in order to incorporate the prior experience to the learning curve model especially for large-scale projects and its results have been corroborated in the development and improvement activities of Boeing 707 (Jaber, 2016). It is considered as a modified Straight-line model which includes a factor "B" to represent the amount of prior experience and diminishes the learning curve (Badiru, 1992; Srouf et al., 2016). The mathematical equation and its logarithmic form are as follows (Srouf et al., 2016):

$$Y = A * (X + B)^{-n} \tag{4}$$

$$\log Y = \log A - n * \log (X + B) \tag{5}$$

where: Y=unit or cumulative average cost, man-hours or time to complete the Xth unit; A=cost, man-hours or time required for the first unit; X=unit number; B=factor describing the crew's prior experience.

The Stanford "B" model is equivalent to the Straight-line model if the deployed crew is not experienced in executing the project in hand. The value of "B" fluctuates within the range of 0–10 (Gottlieb and Haugbølle, 2010; Mályusz and Pém, 2014). For inexperienced crews "B" is equal to zero and it can reach "B">4 depending on the crews' experience level (Thomas et al., 1986).

2.3.3. Cubic model

Carlson (1973) proposed the Cubic model and highlighted that the use of a multiple-curved slope can further enhance the Straight-line model (Hijazi et al., 1992). In this model, the learning rate is a variable factor with its values differentiating based on previous experience and productivity levelling-off towards the finalization phase. The mathematical expression is as follows (Thomas et al., 1986).

$$\log Y = \log A - b * \log X + C * (\log X)^2 + D * (\log X)^3 \tag{6}$$

where: Y=unit or cumulative average cost, man-hours or time to complete the Xth unit; A=cost, man-hours or time required for the first unit; X=unit number; b=initial logarithmic slope at the first unit; C=quadratic factor; D=cubic factor.

The first derivative of Eq. (6) gives the learning elasticity "a" for the Cubic model which can be expressed as (Karaoz and Albeni, 2005):

$$\frac{dy}{dx} = a = b + 2 * C * (\log X) + 3 * D * (\log X)^2 \tag{7}$$

Factors "C" and "D" are estimated using the Eq. (6) and Eq. (7) and knowing another data point along the curve.

2.3.4. Piecewise or Stepwise model

The Piecewise model is approximately similar to a linear version of the Cubic model distinguished in three segments, for which the learning rate has a different constant value. If plotted on a log-graph, three linear parts appear with different slopes (Everett and Farghal, 1994).

These three distinct phases correspond to the three phases of the hypothetical learning curve (see Fig. 1). The calculation formula is presented below (Thomas et al., 1986).

$$\log Y = \log A - n_1 * \log X - n_2 * J_1 * (\log X - \log x_{p1}) - n_3 * J_2 * (\log X - \log x_{p2}) \tag{8}$$

where: Y=unit or cumulative average cost, man-hours or time to complete the Xth unit; A=cost, man-hours or time required for the first unit; X=unit number; n₁=slope of the first segment; n₂=additional slope of the second segment; J₁=1 when X>x_{p1}, 0 otherwise; n₃=additional slope of the third segment; J₂=1 when X>x_{p2}, 0 otherwise; x_{p1}=first point where the slope changes, usually in the operation-learning phase (see Fig. 1); x_{p2}=second point where the slope changes, the end of the routine-acquiring phase (see Fig. 1); Total slope = n₁+ n₂+ n₃

2.3.5. Exponential model

The Exponential model was developed by the Norwegian Building Research Institute in 1960 (U.N., 1965). The main rule is that a partial segment of the cost or time required per unit should be considered constant and the rest may be diminished by 50% after a constant number of iterations. The mathematical equation is as follows (Zahran et al., 2016):

$$Y_u = Y_{ult} + \frac{A - Y_{ult}}{2 * X/H} \tag{9}$$

where: Y_u =unit cost, man-hours or time to complete the X^{th} unit; Y_{ult} =ultimate man-hours per unit at the end of routine-acquiring phase; A =cost, man-hours or time required for the first unit; X =unit number; H =constant named "Halving Factor".

The minimum cost or productive time (expressed in man-hours) achieved at the final stage of the routine-acquiring phase (Y_{ult}) must be known along with constant "H" which represents a "Halving Factor", namely to determine what amount of units can be diminished by 50% through repetition. There has been no evidence of a cumulative data learning model (Thomas et al., 1986).

2.4. Caisson Construction Operations

In principle, floating caissons are pre-cast concrete box-shaped elements that are used in marine infrastructure projects and are constructed on floating dry docks (Panas and Pantouvakis, 2014). The slipforming construction method is used to construct them in a repetitive fashion. The slipforming method is used to construct high-rise structures such as silos (Zayed et al., 2008). In the general case, the slipform method starts with the "slipform assembling phase", goes on with the "slipform phase" and concludes with the "slipform dismantling phase" (see Fig. 3).

Sometimes, a situation can occur where the most suitable floating dock may not be available according to the project time schedule. As a result, the floating may not have sufficient bearing capacity to accommodate the finalization of the complete caisson at once on the floating dock. When that happens, the slipforming procedure is partly executed on the floating dock (Phase A) and then continued afloat (Phase B) up until the caisson has been fully constructed.

3. Methodology

3.1. Case Study Selection

A large-scale marine project in Greece has been selected as a testbed for the present study. The investigated scope regards the construction of thirty-four (34) caissons fabricated via the slipform technique over eight months (January 2012 to August 2012).

A number of reasons justified the selection of this project as the research case study: (a) caisson construction is an iterative and complex construction activity, hence fulfilling the prerequisites for the development of the learning phenomenon; (b) a sufficient number of data

points expressed in workhours/activity's output (more than 1,700 on-site measurements) to yield robust and reliable learning curves. This project was also partially incorporated in the research of Panas and Pantouvakis (2014, 2018).

3.2. Selected Activities

Each batch for the caissons construction includes two items. The slipforming starts on the floating barge and stops at a height of +9.00m, due to weight capacity restrictions (so-called "Phase A"). Then the floating dock is submersed, the caissons are berthed along the quay and concreting is completed at a final elevation of +19.70m (so-called "Phase B"). The slipforming equipment is dismantled and re-assembled for the construction of the next two caissons. From a total of nineteen (19) activities (Pantouvakis and Panas, 2013) which are required for the construction of a caisson, only six (6) of them were studied, as follows: (a) Assembling, (b) Dismantling, (c) Initial Concreting (Phase A and B), as well as (d) Slipform (Phase A and B).

The study focused on these activities because: (a) they are mainly labour-intensive and (b) they satisfy the criterion of being "sufficiently complex repetitive activities" (Thomas, 2009) which is the condition for the development of the learning phenomenon. Hence, only these activities were found complex enough, since the observed on-site productivity data for the rest did not present significant productivity variability. Lastly, for reasons of research completeness, the total construction process for each caisson comprising all nineteen activities (which represents a production cycle) was studied from a learning perspective.

3.3. Learning Curve Models Investigation

The five learning curve models were examined, so as to define the most optimum approach for the assessment of completed construction activities as well as to find the best solution for estimating future productivity values. The assessment criterion for the suitability is the deviation of the actual construction data from the predictions generated by each model. Although in construction settings it is often most convenient to use the cumulative average time, this research adopts both unit and cumulative average data, in order to provide a more robust research framework. Besides, cumulative average time data are more suitable for long-term planning, whereas unit data can be better used for weekly or daily planning (Everett and Farghal, 1997; Ammar and Samy, 2015).



Fig. 3. Floating caisson production cycle (Source: Panas & Pantouvakis, 2018)

3.3.1. Stage A: Assessment of best fit model with historical productivity data

The research approach is aligned with the suggestions of Thomas et al. (1986), which have been adopted by most pertinent research on learning curve theory (Everett and Farghal, 1994; Wong et al., 2007; Ammar and Samy, 2015; Ssour et al., 2016). There are three types of input data in the present research: (a) field-data, (b) empirical-data and (c) data that are estimated or calculated through explicit mathematical relationships or models. More specifically, the "A" and "Y_{ult}" values have been defined from the field data (except "A" for the Piecewise model). Factor "B" has been empirically set equal to 1.00. Point "x_{p1}" denotes the construction of the 5th caisson, whereas point "x_{p2}" is set after the construction of the caisson Nr. 26, because the productivity threshold is expected to be reached by then. MS Excel solver function (version 2010) has been used in conjunction with the least squares method, so as to determine the optimum learning curve model parameters and fitting curve.

Pearson's coefficient of determination (R²) is the preferred metric for each model's adjustment evaluation to historical productivity data since it is quite often used as a regression tool for the quantitative depiction of critical parameters in learning curve models. R² values fluctuate from 0 to 1.00 whereas the closer the R² values to 1.00, the better the correlation of the fitted data to the selected model.

3.3.2. Stage B: Assessment of best prediction model for future performance

The research methodology is the one that was developed and proposed by Everett and Farghal (1994) to assess the capability of estimating the expected productivity of scheduled activities. According to the research approach, the collected productivity data from the construction of the thirty-four (n=34) caissons are divided in half (m=n/2). The first seventeen caissons (m=17) become the "historical data", while the other seventeen caissons represent the future dataset. The least squares method was applied for the first seventeen (17) caissons in order to determine the optimum fitting curve, as well as the main model parameters. Pearson's coefficient of determination (R²₁₋₁₇) was calculated for the first half dataset and the estimated best-fit learning curves were extended for predictions within the range of 17th to 34th caisson.

However, since Pearson's coefficient of determination (R²) yields results that are acceptable only within the range

of the data which were used to plot the respective learning diagrams (Everett and Farghal, 1994), another metric was used for the evaluation of future performance. More specifically, the statistical metric E_f ("average percentage error") that was proposed by Everett and Farghal (1994) specifically for learning curve models was used, with its mathematical expression being as illustrated in Eq. (10):

$$E_f = \frac{\sum_{i=1}^k \frac{|y'_{m+i} - y_{m+i}|}{y_{m+i}}}{k} * 100 \tag{10}$$

where: m = the number of caissons to be fitted; k = the number of caissons to be predicted; y'_{m+i} = the value found on the extension of the best-fit curve; y_{m+i} = the actual measured values; E_f = average percentage error, which ranges from 0% indicating a perfect correlation between the extended best-fit curve and the actual data to large positive values indicating no correlation.

Essentially the statistical metric E_f expresses the difference (%) on average between the real dataset and the projected productivity estimates. The parameter determination method for the models is the same as the one in section 3.3.1. with the following exceptions: (a) the optimization is performed for the first seventeen (17) data points which represent the historical data and (b) the point "x_{p2}" (productivity threshold) for the Piecewise model is defined to be caisson Nr. 13 (i.e. it is the selection of the 26th caisson out of 34 caissons if expressed in percentage terms).

4. Results and Discussion

4.1. Stage A: Assessment of Best Fit Model with Historical Productivity Data

4.1.1. Unit data

Table 1 summarizes each model's unit data performance for the six activities, as well as the total caisson construction process.

There is a clear indication that the Cubic model yields the best fitting results for unit data relating to a) the total caisson construction process, b) Slipform Phase A and c) Initial Concreting Phase B. The Exponential model gives the least favorable adjustment, without being unacceptable, though, in absolute terms. Especially, for the total caisson construction process all investigated models have values of R²>0.92, thus denoting a satisfactory correlation with actual field data. These results enhance previous research and denote that the best fit model depends on the project's location and nature.

Table 1. Correlation of Learning Curve (LC) Models for Completed Activities with Unit Data

Activity	Pearson's coefficient of Determination (R ²) for LC Models				
	Straight-Line	Stanford "B"	Cubic	Piecewise	Exponential
Assembling	0.7979	0.7537	0.9213	0.8205	0.8259
Initial Concreting A	0.5937	0.5880	0.6269	0.5934	0.5491
Slipform A	0.9673	0.9612	0.9841	0.9695	0.9035
Initial Concreting B	0.8716	0.8622	0.8970	0.8730	0.8261
Slipform B	0.9543	0.9595	0.9390	0.9549	0.8887
Dismantling	0.3614	0.3429	0.4112	0.4142	0.3916
Total	0.9530	0.9327	0.9781	0.9573	0.9256

Fig. 4 illustrates the developed learning curves for the total caisson construction process. It should be noted that equivalent learning curves have been drafted for each activity, but are omitted for brevity reasons.

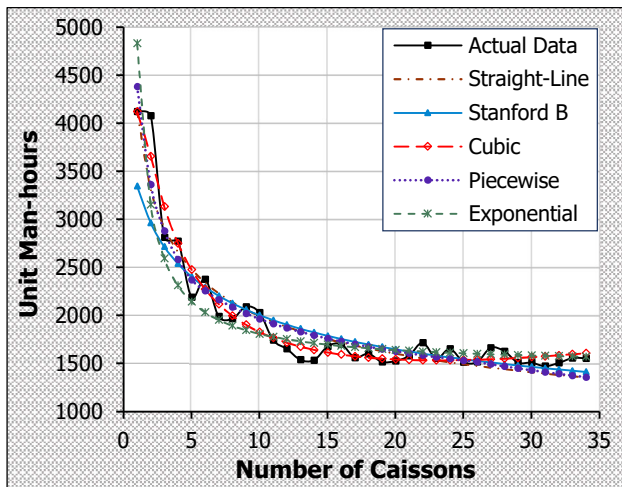


Fig. 4. Learning Curves for Historical Data (Unit)

The following paragraphs summarize the results of the fitting process for the other four activities and substantiate their differentiation as follows:

- **Assembling:** The Cubic model has, by far, the best performance amongst all the models. Slipforming equipment assembling is a complex activity, executed by highly specialized crews. Hence the Cubic model seems to better simulate the effect of prior experience, as well as productivity beyond the standard production point (Thomas et al.,1986).
- **Initial Concreting - Phase A:** All models yielded an average performance, with the Cubic model being slightly better. It seems that the initial concreting process is a labour-intensive but rather standardized activity, with no major margins for developing the learning phenomenon.
- **Slipform - Phase B:** All models yielded fairly satisfactory results. However, the best adjustment was achieved by the Stanford "B" model. Slipforming phase B is exactly the type of activity best simulated by the Stanford "B" model: complex (e.g. concreting taking place afloat etc.) and labour-intensive.
- **Dismantling:** The Piecewise model has provided the best adjustment, although the average performance

of all models was relatively low. The dismantling process is equipment-intensive (e.g. mainly use of mobile / tower cranes) with a minor margin for learning phenomenon development. In addition, it is an activity sensitive to a series of external productivity factors, such as adverse weather conditions, the intermittent flow of work, limited accessibility to the floating equipment, etc.

4.1.2. Cumulative average data

First, it must be clarified that the Exponential model is not included in the models evaluation because, as mentioned above, no such model for cumulative average data has been presented. Table 2 summarizes each model's cumulative data performance for the six activities, as well as the total caisson construction process. Fig. 5 illustrates the developed learning curves for the total caisson construction process.

The results indicate that the Cubic model has the best adjustment to historical productivity data for all studied activities. A thorough examination of Fig. 5 reveals that both the Cubic model best-fit curve, as well as the rest of the learning curves, almost coincide with the actual curve derived from the actual on-site data. In addition, the other three models yielded a Pearson's coefficient of determination (R^2) value very close to 1.00, denoting a satisfactory prediction capability for all learning curve models. For some activities (e.g. Initial Concreting A) the Stanford "B" model yields a better correlation than the Straight-line or Piecewise model. This is due to the fact that factor "B" equals to 1.00 for all activities, thus denoting the same level of experience for all deployed crews. The latter, of course, may not always be the case.

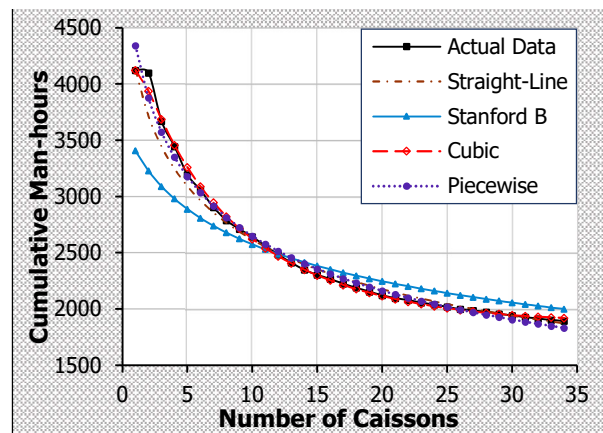


Fig. 5. Learning Curves for Historical Data (Cumulative)

Table 2. Correlation of Learning Curve (LC) Models for Completed Activities with Cumulative Average Data

Activity	Pearson's coefficient of Determination (R^2) for LC Models			
	Straight-Line	Stanford "B"	Cubic	Piecewise
Assembling	0.9761	0.9663	0.9957	0.9766
Initial Concreting A	0.9703	0.9747	0.9808	0.9742
Slipform A	0.9948	0.9970	0.9996	0.9977
Initial Concreting B	0.9875	0.9882	0.9959	0.9894
Slipform B	0.9980	0.9974	0.9991	0.9989
Dismantling	0.9423	0.9283	0.9843	0.9525
Total	0.9940	0.9890	0.9985	0.9941

4.1.3. Comparative analysis of learning curve models for completed activities

Taking a look at the results of Tables 1 and 2, there is a coincidence in the models' ranking for the total caisson construction process regarding both unit and cumulative average data, with the Cubic model providing the best results. This finding is in accordance with previous research (e.g. Thomas, 2009) since it has been observed that, in principle, cumulative average data provide better correlations than the unit data. Fig. 4 and Fig. 5 denote that learning curves based on cumulative average data provide a smoother graphical representation of the learning phenomenon. The same goes for the respective learning curves produced for each one of the six examined activities, but are omitted due to brevity reasons.

Comparing the Cubic and Piecewise model to the Straight-line, for the total caisson construction process yields a statistically insignificant difference of 2.51% and 0.43% respectively ($p < 5\%$) for unit data. Same goes for the cumulative average data, since the respective observed differences were 1.96% and 0.05% ($p < 5\%$). A similar trend is found for other activities as well. All the aforementioned findings corroborate the tendency of construction researchers and practitioners to adopt the

Straight-line model as a more "user-friendly" approach since that model (a) gives results of equal reliability to other learning curve models, (b) is simpler to apply since it requires fewer input parameters and assumptions from the engineer's perspective.

4.2. Stage B: Assessment of best prediction model for future performance

4.2.1. Unit data

Table 3 presents the respective Pearson ($R^2_{(1-17)}$) and $Ef_{(18-34)}$ statistical metrics, while Fig. 6 illustrates the learning curves that were derived from the fitting of the first seventeen (17) datasets, as well as their extended curves covering the total caisson construction process.

Stanford "B" model was found to be the best predictor of productivity rates for the total caisson construction process. The average Ef error percentage had the lowest value equal to $Ef_{(18-34)} = 5.7\%$, which denotes a very good correlation of the actual data and the extended curve. Generally, it is established that Stanford "B" model simulates complex construction processes since it was developed to integrate previous experience in the learning curve estimations.

Table 3. Results of Learning Curve (LC) Models for Future Performance Prediction with Unit Data

Activity	$R^2_{(1-17)}$ and $Ef_{(18-34)}$ Values for LC Models									
	Straight-Line		Stanford "B"		Cubic		Piecewise		Exponential	
	$R^2_{(1-17)}$	$Ef_{(18-34)}$ (%)	$R^2_{(1-17)}$	$Ef_{(18-34)}$ (%)	$R^2_{(1-17)}$	$Ef_{(18-34)}$ (%)	$R^2_{(1-17)}$	$Ef_{(18-34)}$ (%)	$R^2_{(1-17)}$	$Ef_{(18-34)}$ (%)
Assembling	0.8957	40.17	0.8971	32.57	0.9660	223.49	0.8960	47.79	0.8629	29.65
Initial Concreting A	0.6298	48.29	0.6469	48.01	0.6554	57.62	0.6291	49.58	0.5435	70.83
Slipform A	0.9662	7.60	0.9767	11.20	0.9815	5.61	0.9791	7.78	0.8922	8.15
Initial Concreting B	0.8757	27.12	0.8889	25.94	0.9016	30.66	0.8893	33.24	0.8079	42.27
Slipform B	0.9453	7.71	0.9503	7.41	0.9508	17.69	0.9504	7.99	0.8796	8.09
Dismantling	0.6891	25.57	0.6900	23.75	0.7298	50.10	0.6909	27.90	0.6360	38.25
Total	0.9629	11.33	0.9626	5.75	0.9767	17.42	0.9628	16.21	0.9096	6.87

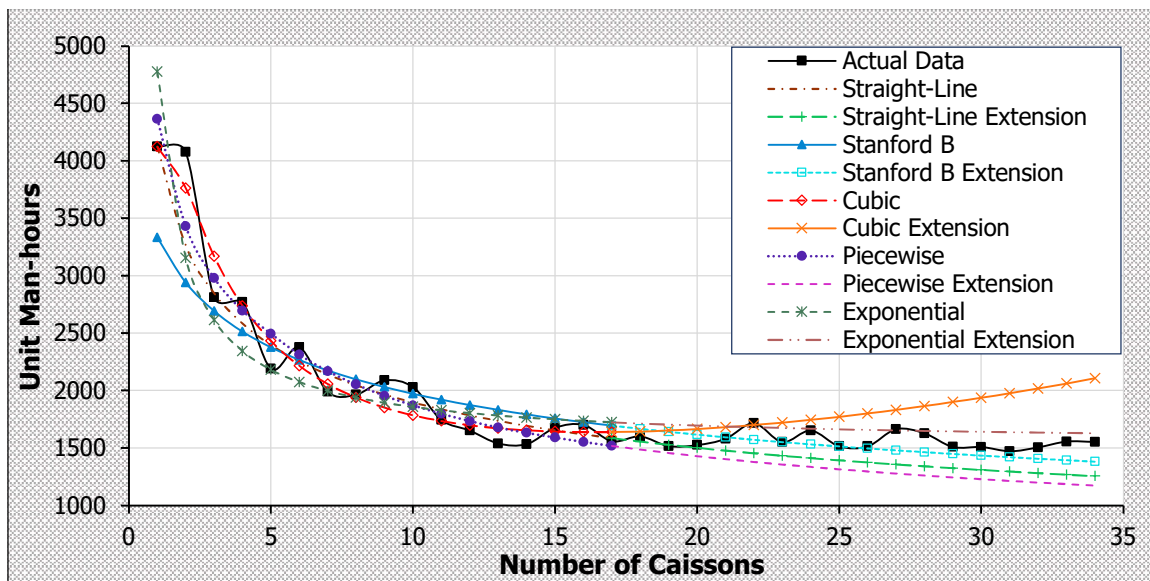


Fig. 6. Learning Curves and Learning Curves Extensions of Models for Unit Data

The following observations were made for the other activities, as follows:

- The Dismantling, Initial Concreting for all its phases and Slipform (for the so-called “Phase B”) were best simulated by the Stanford "B" model, with satisfactory correlation results. An exception is found for the Initial Concreting Phase A activity ($E_{f(18-34)}=48.01\%$), which is due to the reasons analytically presented in section 4.1.1. The differences in the models’ ranking for the other activities in relation to the total construction process is attributed to the different construction type and nature of each activity that generates a different simulation mechanism.
- The Assembling activity was better simulated by the Exponential model, while Initial Concreting Phase A was better represented by the Cubic model. Their differentiation is again related to their different construction scope.

4.2.2. Cumulative Average Data

Table 4 presents the Pearson ($R^2_{(1-17)}$) values and the statistical metrics $E_{f(18-34)}$, while Fig. 7 illustrates the optimum best-fit curves, as well as their extended curves for the total construction process. The Straight-line model was found to better predict future productivity values for the total caisson construction process. The $E_{f(18-34)}=2.55\%$ corroborates the very good correlation between the actual data and their extended predictions. This finding is in accordance with the respective results of Everett and Farghal (1994), who claimed that the Straight-line model estimated in the best possible way expected productivity in leaning-prone activities.

Table 4 shows that the Straight-line model was not found to be the best predictor in any of the other activities, with the exception of the Assembling activity. More specifically, Slipform Phase A&B as well as Initial Concreting B were best fitted by the Piecewise model, the Dismantling activity by the Stanford "B" and the Initial Concreting Phase A by the Cubic model.

Table 4. Results of Learning Curve (LC) Models for Future Performance Prediction with Cumulative Average Data

Activity	$R^2_{(1-17)}$ and $E_{f(18-34)}$ Values for LC Models							
	Straight-Line		Stanford "B"		Cubic		Piecewise	
	$R^2_{(1-17)}$	$E_{f(18-34)}$ (%)	$R^2_{(1-17)}$	$E_{f(18-34)}$ (%)	$R^2_{(1-17)}$	$E_{f(18-34)}$ (%)	$R^2_{(1-17)}$	$E_{f(18-34)}$ (%)
Assembling	0.9698	5.16	0.9786	25.67	0.9969	82.93	0.9749	19.79
Initial Concreting A	0.9313	7.11	0.9438	14.24	0.9534	1.05	0.9433	1.92
Slipform A	0.9901	7.90	0.9982	19.80	0.9993	4.24	0.9983	2.65
Initial Concreting B	0.9748	10.58	0.9853	24.55	0.9920	7.57	0.9842	6.94
Slipform B	0.9971	4.16	0.9982	14.26	0.9988	2.95	0.9986	0.69
Dismantling	0.9647	5.59	0.9749	2.88	0.9891	3.73	0.9744	13.70
Total	0.9903	2.55	0.9937	14.40	0.9979	10.73	0.9919	5.17

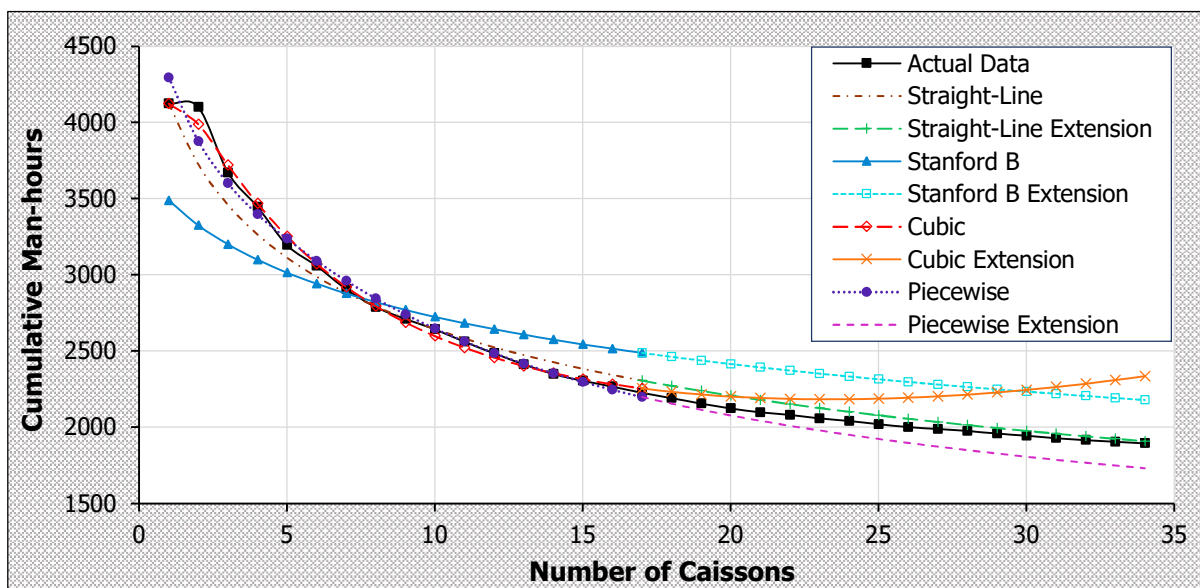


Fig. 7. Learning Curves and Learning Curves Extensions of Models for Cumulative Data

The differentiation of the results for the individual activities in relation to the respective results of the total construction process is attributed to a number of reasons as follows:

1. The statistical analysis for the prediction of future performance based on cumulative average data, does not correlate the learning rate of each activity to the learning rate of the total caisson construction process (Panas and Pantouvakis, 2014).
2. According to Everett and Farghal (1997), the prediction accuracy of future performance of individual activities is improved for about 20-40% of the activity and then is slightly differentiated. When average cumulative data is used, the yielded results become less accurate as the activity comes close to its end, making them an unreliable prediction metric.
3. The applied methodology for future predictions has only been validated for unit data (Everett and Farghal, 1994), thus it might present some inherent deficiencies in the case of cumulative average data for predicting individual activities' performance.

4.2.3. Comparative analysis of learning curve models for future performance prediction

The Cubic model presents the best $R^2_{(1-17)}$ value for unit data, but not the lowest average percentage error $E_{f(18-34)}$ (see Table 3). Although the $E_{f(18-34)}=17.42$ value is acceptable for the total caisson construction process, its extended prediction curve presents a significant increasing trend beyond the 25th caisson (see Fig. 6), which is magnified outside of the thirty-four caissons range. The same findings have been observed for Slipform Phase A&B, where the extended curves demonstrate a slight upward and downward trend respectively. At the same time, it is interesting, that the Cubic model was the best fitting model for Slipform Phase A ($E_{f(18-34)}=5.61$). This finding is again in accordance with published research which claims that Cubic model is poor future performance predictor (Everett and Farghal, 1994).

The Piecewise model for unit data presents a similar (but opposite) trend to the Cubic model for the total caisson construction process. The statistical metric $E_{f(18-34)}=16.21$ is acceptable but its extended curve presents a downward trend beyond the 25th caisson (see Fig. 6), which is magnified outside of the thirty-four caissons range. Same goes for Slipform Phase A&B, although its $E_{f(18-34)}$ values are generally acceptable.

In the case of cumulative average data for the total caisson construction process, the Cubic and Piecewise model presented similar trend to the unit data. Despite its fairly acceptable $E_{f(18-34)}$ metrics (see Table 4), their extended curves imply a significant upward and downward trend for the Cubic and Piecewise model respectively (see Fig. 7). The same observation is made for the six individual activities despite their satisfying $E_{f(18-34)}$ values.

As a final remark, it can be argued that the Cubic and Piecewise models are poor performance predictors, with the distinctive difference that the Piecewise model needs more production cycles to "present" the deviations in relation to the Cubic model. Both prediction curves possess a significant risk of producing non-realistic results if extended beyond the sample of the thirty-four caissons.

5. Conclusions

The conducted research demonstrated that the learning effect was intensely present in the studied project, which resulted in significant improvements in thirty-four (34) caissons construction productivity. More specifically, from the approx. 4.000 man-hours that were required for the first two caissons, the project was completed with average productivity of approx. 1.500 man-hours/caisson.

In the case of the total caisson construction process, all five (5) learning curve models were scrutinized for unit and cumulative historical productivity data and yielded a coefficient of $R^2 > 0.90$, which denotes a strong correlation to actual data. The Cubic model has proven to be the best performer, as far as its convergence to past data is concerned. The results coincide with the findings of other published research (Thomas et al., 1986; Everett and Farghal, 1994) and resulted in the Cubic model being more suitable in describing completed construction activities.

In the case of predicting future performance with unit data for the total caisson construction process, the Stanford "B" model gave the best predictions, while the Straight-line model yielded better results when average cumulative data were used. Our findings corroborate previous research (e.g. Everett and Farghal, 1994), by proving that linear models are deemed more suitable to predict future performance: the Straight-line model is the first linear model, while Stanford "B" model comes as a slight variation of the Straight-line model with some minor modifications in the first cycles due to prior experience.

During the analysis of the six remaining activities the models' results differentiated according to the studied scope, namely whether the caisson construction process was studied as a whole or based on individual activities.

The Straight-line model yielded acceptable results for all examined scenarios, thus corroborating its broad acceptability from construction practitioners.

Therefore, it is beyond doubt that the learning curve theory is an efficient and effective tool for assessing historical and predicting future productivity data in the case of caisson construction operations. The analysis becomes a useful tool for construction practitioners in order for them to plan and monitor both time schedules and cost estimations for their projects.

Possible research extensions could be developed in the area of future performance predictions, by adopting different data representation techniques such as a) moving average data and b) exponential weighted average. The research scope may be enhanced with the inclusion of other (non-classic) learning curve models (e.g. DeJong, Knecht, hyperbolic models), which were excluded from the current study due to brevity reasons. The enhancement of the already established historical project database with even more data covering similar activities is deemed necessary, so as to be able to structure a future performance prediction tool with inherent flexibility to simulate different work scenarios and feed project executives with valuable insights for informed decision making.

References

- Ammar, M. A. and Samy, M. (2015). Learning curve modelling of gas pipeline construction in Egypt.

- International Journal of Construction Management*, 15(3), 229-238. doi: 10.1080/15623599.2015.1067346
- Ammar, M. A. and Abdel-Maged, A. F. (2018). Modeling of LOB scheduling with learning development effect. *International Journal of Construction Management*, 18(6), 517–526. doi: 10.1080/15623599.2017.1357350
- Arditi, D., Tokdemir, O. B., and Suh, K. (2001). Effect of learning on line-of-balance scheduling. *International Journal of Project Management*, 19(5), 265–277. doi: 10.1016/S0263-7863(99)00079-4
- Badiru, A. B. (1992). Computational survey of univariate and multivariate learning curve models. *IEEE Transactions on Engineering Management*, 39(2), 176–188.
- Carlson, J. (1973). Cubic Learning Curves: Precision tool for labor estimating. *Manufacturing Engineering and Management*, 67(11), 22–25.
- Couto, J. P. and Teixeira, J. C. (2005). Using linear model for learning curve effect on highrise floor construction. *Construction Management and Economics*, 23(4), 355–364. doi: 10.1080/01446190500040505
- Everett, J. G. and Farghal, S. (1994). Learning curve predictors for construction field operations. *Journal of Construction Engineering and Management*, 120(3), 603–616.
- Everett, J. G. and Farghal, S. (1997). Data representation for predicting performance with learning curves. *Journal of Construction Engineering and Management*, 123(1), 46–52.
- Farghal, S. and Everett, J. G. (1997). Learning curves: Accuracy in predicting future performance. *Journal of Construction Engineering and Management*, 123(1), 41–45.
- Glock, C. H., Grosse, E. H., Jaber, M. Y., and Smunt, T. L. (2019). Applications of learning curves in production and operations management: A systematic literature review. *Computers & Industrial Engineering*, 131, 422–441. doi: 10.1016/j.cie.2018.10.030
- Gottlieb, S. C. and Haugbølle, K. (2010). *The repetition effect in building and construction works*. Kobenhaven: Danish Building Research Institute.
- Hijazi, A. M., AbouRizk, S. M., and Halpin, D. W. (1992). Modeling and simulating learning development in construction. *Journal of Construction Engineering and Management*, 118(4), 685-700.
- Hinze, J. and Olbina, S. (2009). Empirical analysis of the learning curve principle in prestressed concrete piles. *Journal of Construction Engineering and Management*, 135(5), 425–431.
- Jaber, M. Y. (2016). *Learning curves: Theory, Models, and Applications*. Boca Raton, FL: Taylor & Francis Group.
- Jarkas, A. M. (2010). Critical investigation into the applicability of the learning curve theory to rebar fixing labor productivity. *Journal of Construction Engineering and Management*, 136(12), 1279–1288. doi 10.1061/(ASCE)CO.1943-7862.0000236
- Jarkas, A. M. (2016). Learning effect on labour productivity of repetitive concrete masonry blockwork: Fact or fable? *International Journal of Productivity and Performance Management*, 65(8), 1075-1090. doi 10.1108/IJPPM-08-2015-0119
- Jarkas, A. M. and Horner, M. (2011). Revisiting the applicability of learning curve theory to formwork labour productivity. *Construction Management and Economics*, 29(5), 483-493. doi: 10.1080/01446193.2011.562911
- Karaoz, M. and Albeni, M. (2005). Dynamic technological learning trends in Turkish manufacturing industries. *Technological Forecasting-Social Change*, 72(7), 866–885.
- Khanh, H. and Kim, S. (2014). Determining labor productivity diagram in high-rise building using straight-line model. *KSCCE Journal of Civil Engineering*, 18(4), 898–908.
- Lee, B., Lee, H., Park, M., and Kim, H. (2015). Influence factors of learning-curve effect in high-rise building constructions. *Journal of Construction Engineering and Management*, 141(8), 04015019.
- Lutz, J. D., Halpin, D. W., and Wilson, J. R. (1994). Simulation of learning development in repetitive construction. *Journal of Construction Engineering and Management*, 120(4), 753-773. doi: 10.1061/(ASCE)0733-9364(1994)120:4(753)
- Mályusz, L. and Pém, A. (2013). Prediction of the learning curve in roof insulation. *Automatization in Construction*, 36, 191–195. doi: 10.1016/j.autcon.2013.04.004
- Mályusz, L. and Pém, A. (2014). Predicting future performance by learning curves. *Procedia-Social and Behavioral Sciences*, 119, 368-376. doi: 10.1016/j.sbspro.2014.03.042
- Naresh, A. L. and Jahren, C. T. (1999). Learning outcomes from construction simulation modeling. *Civil Engineering and Environmental Systems*, 16(2), 129-144. doi: 10.1080/02630259908970257
- Nguyen, L. D. and Nguyen, H. T. (2013). Relationship between building floor and construction labor productivity: A case of structural work. *Engineering, Construction and Architectural Management*, 20(6), 563–575. doi: 10.1108/ECAM-03-2012-0034
- Panas, A. and Pantouvakis, J. P. (2010). Evaluating research methodology in construction productivity studies. *The Built and Human Environment Review*, 3(1), 63-85.
- Panas, A. and Pantouvakis, J. P. (2014). Simulation-based and statistical analysis of the learning effect in floating caisson construction operations. *Journal of Construction Engineering and Management*, 140(1), 04013033.
- Panas, A. and Pantouvakis, J. P. (2018). On the use of learning curves for the estimation of construction productivity. *International Journal of Construction Management*. doi: 10.1080/15623599.2017.1326302
- Pantouvakis, J. P. and Panas, A. (2013). Computer simulation and analysis framework for floating caisson construction operations. *Automation in Construction*, 36, 196-207. doi: 10.1016/j.autcon.2013.04.003
- Pellegrino, R. and Costantino, N. (2018). An empirical investigation of the learning effect in concrete operations. *Engineering, Construction and Architectural Management*, 25(3), 342-357. doi: 10.1108/ECAM-02-2017-0036
- Pellegrino, R., Costantino, N., Pietroforte, R., and Sancilio, S. (2012). Construction of multistorey concrete structures in Italy: patterns of productivity and learning curves. *Construction Management and Economics*, 30(2), 103-115. doi: 10.1080/01446193.2012.660776
- Shan, Y., Goodrum, P. M., Zhai, D., Haas, C., and Caldas, C. H. (2011). The impact of management practices on mechanical construction productivity. *Construction Management and Economics*, 29(3), 305–316.

- Srouf, F. J., Kiomjian, D., and Srouf, I. M. (2016). Learning curves in construction: a critical review and new model. *Journal of Construction Engineering and Management*, 142:06015004.
- Srouf, F. J., Kiomjian, D., and Srouf, I. M. (2018). Automating the use of learning curve models in construction task duration estimates. *Journal of Construction Engineering and Management*, 144(7): 04018055.
- Thomas, H. R. (2009). Construction learning curves. *Practice Periodical on Structural Design and Construction*, 14(1), 14-20.
- Thomas, H. R., Mathews, C. T., and Ward, J. G. (1986). Learning curve models of construction productivity. *Journal of Construction Engineering and Management*, 112(2), 245-259. doi: 10.1061/(ASCE)1084-0680(2009)14:1(14)
- Ugulu, A. R. and Allen, S. (2018). Using the learning curve theory in the investigation of on-site craft gangs' blockwork construction productivity. *Built Environment Project and Asset Management*, 8(3), 267-280. doi: 10.1108/BEPAM-09-2017-0067
- United Nations Committee on Housing Building and Planning (1965). *Effect of repetition on building operations and processes on site*. United Nations, New York.
- Wong, P. S., On Cheung, S., and Hardcastle, C. (2007). Embodying learning effect in performance prediction. *Journal of Construction Engineering and Management*, 133, 474-482. doi: 10.1061/(ASCE)0733-9364(2007)133:6(474)
- Zahran, K., Nour, M., and Hosny, O. (2016). The effect of learning on line of balance scheduling: Obstacles and potentials. *International Journal of Engineering Science and Computing*, 6(4), 3831-3841. doi: 10.4010/2016.889
- Zayed, T., Sharifi, M. R., Baciou, S., and Amer, M. (2008). Slip-form application to concrete structures. *Journal of Construction Engineering and Management*, 134, 157-168.



Panagiota Ralli is a Mechanical Engineer with a M.Sc. in Engineering Project Management from Hellenic Open University. Over 20 years of experience in the design, planning and construction of complex infrastructure projects (Olympic Stadiums, Opera Houses & Concert Halls, Public Hospitals, Aircraft Maintenance & Repair Facilities). Research interests in the optimization of construction productivity and project management. Currently working as a Mechanical Engineer at the sole constructor of the Greek Public Infrastructure Buildings, named as Building Infrastructures S.A., within the Department of Projects Contracts & Procurement.



Antonios Panas is a Civil Engineer, M.Sc., with a Ph.D. in Construction Engineering & Management. Assistant Professor of N.T.U.A. and Adjunct Lecturer of Hellenic Open University. Research interests in construction productivity and simulation. Currently working on large-scale marine construction

projects as senior project manager. More information available at panas.view.gr.



Management Association. More information available at jpp.view.gr.

John-Paris Pantouvakis is a Professor in the Department of Construction Engineering and Management, Faculty of Civil Engineering, National Technical Univ. of Athens. Heads the Centre for Construction Innovation laboratory and is an active member of the International Project



Dimitrios Karagiannakidis is a Civil Engineer with a M.Sc in Transport & Logistics Management from Aristotle University of Thessaloniki. Big experience on construction of complex infrastructure and renewable energy sources projects (Transadriatic Pipeline and Kafireas WF). Research interests in sustainable mobility and construction productivity. Special experience at Transport and Logistics issues. Currently working in a multinational wind energy company as Logistics Coordinator for Southern-East Europe.