ON THE PROBLEMS OF MODELLING AND RELIABILITY ASSESSMENT OF CONSTRUCTION PROJECTS DURATION

S. Biruk, P. Jaskowski

•

Lublin University of Technology, Department of Civil Engineering
and Architecture, Lublin, Poland

e-mail: s.biruk@pollub.pl, p.jaskowski@pollub.pl

ABSTRACT

Construction projects are subject to risk. There are a number of methods that allow the planner to consider the effect of random occurrences on the project performance and to assess the chances of meeting the deadlines defined by the contract. PERT belongs to the most popular methods as it assumes a simple approach to estimating the distribution parameters of random variables (task durations) based on the experience of the planner.

The paper summarises PERT's assumptions on the type and parameters of task duration distributions, task duration independence, and the approach to the analysis of the network model in the function of time. The effects of these assumptions on the project makespan estimate are then examined and illustrated by an example.

1 INTRODUCTION

Duration of construction projects, as well as duration of particular tasks the project scope may be broken down into, is affected by a variety of occurrences whose frequency and impact depend on the project-specific, contractor-specific and location-specific conditions (e.g. Biruk and Jaskowski (2008), Dawood (1997), Jaworski and Biruk (2000), Jaskowski et al. (2010), Nasir et al. (2003), Schatteman et al. (2008)).

Table 1 lists top ten risk factors of the greatest mean impact, greatest mean frequency of occurrence and greatest mean importance (impact times frequency), defined on the basis of a survey among chartered engineers employed by construction companies in Poland.

Standard production rates, being often the basis for planning duration of construction processes, are usually expressed by single values – medians. To determine process duration distribution types and parameters, a considerable number of time measurements would be necessary to make the results statistically sound. This might be too costly, time consuming and in some cases unjustified as, due to the unique character of construction projects and processes, statistical data from the past may be of little use in the future.

Many models have been proposed to describe and predict activity / project durations or work produktivity on the basis of risk analysis. According to the way of describing the risk factors impact on activity duration, two groups of methods can be distinguished: quantitative and qualitative. The qualitative models use a verbal description of the impact. The quantitative models base on analytical or numerical relations; there exist simple analytical (e.g. Neil and Knack (1984), Woodward (2003), Ovarain and Popescu (2001), Jergeas and McTague (2002)), neural network (e.g. Kog et al. (1999), Chua et al. (1997), Zayed and Halpin (2005), Shi (1999), AbouRizk et al. (2001), Sonmez and Rowings (1998)), Bayesian belief network (Nasir et al. (2003)), fuzzy set (e.g. Lee and Halpin (2003), Perera and Imriya 2003)), regression (e.g. Hanna and Gunduz (2005), Jaselskis and Ashley (1991)) and simulation models (e.g. Dawood (1997), Schatteman et al. (2008)) to choose from. Most of the quantitative models assume that particular factors affect the processes independently. However, no model is considered to be superior as providing more reliable solutions than the other models. This is so because there are no reliable methods of comparing the results obtained by means of these models. Moreover, no comparative studies on the ease of application of these models in practice have been conducted.

Table 1. Ranking of risk factors, based on local experts' opinion survey

Rank	Hierarchy of factors according to:		
	impact (i)	frequency of occurence (f)	importance ($i \cdot f$)
1	Contractor's cash flow problems	Winter affecting structural works, facade works and external works	Winter weather affecting structural works, facade works and external works
2	Delay of preceding works (Delays in subcontractors' works)	Precipitation affecting structural works, facade works and external works	Precipitation affecting structural works, facade works and external works
3	Winter affecting structural works, facade works and external works	Mistakes and discrepancies in design documents	Delay of preceding works (Delays in subcontractors' works)
4	Precipitation affecting structural works, facade works and external works	Shortage of skilled labour	Shortage of skilled labour
5	Unforeseen ground conditions causing change of substructure works scope and quantity	Variations of works (scope and quantity) due to design changes	Mistakes and discrepancies in design documents
6	Client's low speed of decision-making	Delay of preceding works (Delays in subcontractors' works)	Client's low speed of decision-making
7	Inexperienced / unreliable subcontractors	Demotivating remuneration system	Variations of works (scope and quantity) due to design changes
8	Poor site management and supervision	Client's change of requirements	Demotivating remuneration system
9	Work stoppage according to inspection agencies order	Difficulty with finding subcontractors	Client's change of requirements
10	Delay with design (if delivered in packages)	Client's low speed of decision-making	Difficulty with finding subcontractors

The existing models find application in predicting duration of particular construction processes, groups of works and whole projects and than enable project timing and scheduling.

The measure of project schedule reliability level, R, is the probability of the project's being completed no later than at the contractually agreed completion date, t:

$$R = P(T \leq t). \tag{1}$$

Due to the complexity of production processes in construction, the organization structure of the project team may be variable (Jaskowski (2008)). Harmonizing the work of all project participants at the planning stage is thus a complex task. To allow for all constraints and conditions, the planner would have to solve complex mathematical problems (Biruk and Jaskowski (2008), Jaworski and Biruk (2000)). In the practice of construction, simplified, but still reliable enough models are in demand.

The first attempt to allow for risks in project planning was made by the inventors of *PERT* (Program Evaluation and Review Technique). In spite of far going simplifications that inevitably affect reliability of results, the method stays popular in project management.

2 PERT ASSUMPTIONS ON DISTRIBUTION TYPE AND PARAMETERS OF TASK DURATION RANDOM VARIABLE

The authors of *PERT* assumed that the duration of a process (i.e. a task of a network model) is a random variable of beta distribution. The probability density function of a standardized beta distribution ($x \in [0,1]$) with parameters $\alpha > -1$ and $\beta > -1$ is:

$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+2)} x^{\alpha}(1-x)^{\beta}, & 0 < x < 1, \\ 0 & \text{for other } x. \end{cases} \tag{2}$$

The shape of the probability distribution function depends on the values of the shape parameters α and β and the relation between them. As observed in real life, the function representing a production process distribution is usually unsymmetrical and positively skewed. Another assumption concerns the standard deviation (Littlefield and Randolph (1987)): it is postulated that it equals one sixth of the range of the variable (as for the normal distribution), which implies that $\alpha + \beta = 4$. This assumption is difficult to accept, and its only justification seems to be this three-sigma empirical rule.

The distribution parameters of a process duration are determined on the basis of three estimates, given by a group of experts or a planner considering project risk analysis. These estimates are: optimistic (t_a), pessimistic (t_b), and most likely duration (t_m) that is considered to be the mode of the process duration distribution.

Applying a linear transformation $T = t_a + (t_b - t_a)X$ to the random variable of density function described by Equation 2, one obtains the following formulas that describe expected value and standard deviation of the process duration *T*:

$$E(T) = \mu = \frac{t_a + 4t_m + t_b}{6}, \tag{3}$$

$$D(T) = \sigma = \frac{t_b - t_a}{6}. \tag{4}$$

Figure 1 presents probability density function plots of three variables of beta distribution. All of them positively skewed ($0 \le m \le 0{,}5$), but of considerably different standard deviations. All of them could be approximations of the actual distribution of a process duration (MacCrimmon and Ryavec (1964)). The curve marked as D_1 corresponds to the distribution assumed by *PERT*. Its expected value is $\mu_1 = 1/6(4m+1)$, and standard deviation is $\sigma_1 = 1/6$. The curve D_2 represents a distribution close to a uniform distribution ($\mu_2 \approx 0{,}5$, $\sigma_2 \approx \sqrt{1/12}$). The parameters of the distribution D_3 are $\mu_3 \approx m$ and $\sigma_3 \approx 0$. For D_3, the maximum absolute error of the mean relative to the estimate as used by *PERT* is (MacCrimmon and Ryavec (1964)):

$$max\left\{\left|\frac{1}{6}(4m+1) - \frac{1}{2}\right|, \left|\frac{1}{6}(4m+1) - m\right|\right\} = \frac{1}{3}(1-2m),$$

and the maximum absolute error of the standard deviation is (MacCrimmon and Ryavec (1964)):

$$max\left\{\left|\sqrt{\frac{1}{12}} - \frac{1}{6}\right|, \left|0 - \frac{1}{6}\right|\right\} = \frac{1}{6}.$$

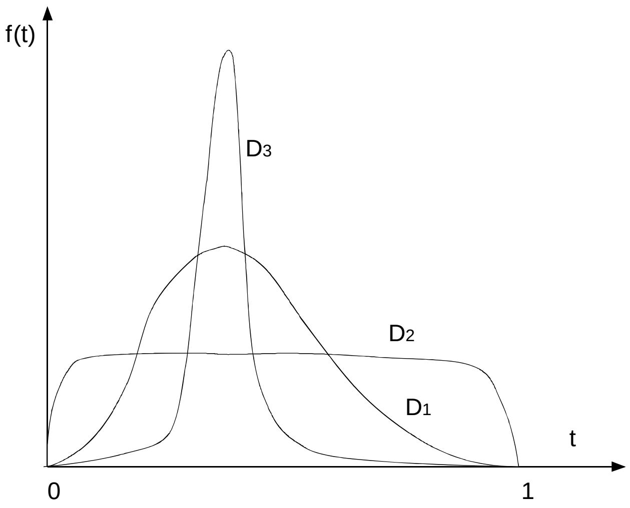

Figure 1. Examples of density functions of beta distribution defined on the interval [0 , 1]

Considering the fact that the actual distribution of a process duration may be significantly different from that assumed by *PERT*, the errors may propagate along the paths of the network model, increasing or decreasing the total error (Hon-Siang and Somarajan (1995)). Elimination of the mean's and standard deviation's errors can be achieved by increasing the number of estimates being predefined quantiles of the process duration. Usually, 7 quantiles are used, e.g. $T_{0.01}, T_{0.125}, T_{0.25}, T_{0.50}, T_{0.75}, T_{0.875}, T_{0.99}$. With so many estimates, one can determine the probability function's shape parameters α and β with precision by means of the least squares method. This approach gives more accurate results than the classic *PERT* three estimates approach because more input is available, and the input is considered to be more accurate: the experts giving estimates are reported to be more accurate with estimates closer to the mode or the mean than to the extreme values (Lichtenstein et al. (1982)).

Three quantiles are sufficient to calculate the parameters of a beta distribution. These could be e.g. $T_{0.05}, T_{0.50}, T_{0.95}$ (Cox (1995)), so no difficult to estimate extreme values are needed. Keefer and Verdini (1993) provided a numerical proof that, in most cases, following formulas are adequate for estimation of the mean and standard deviation of a beta distribution:

$$\mu = 0.630\, T_{0.50} + 0.185\left(T_{0.05} + T_{0.95}\right), \tag{5}$$

$$\sigma^2 = 0.630\left(T_{0.50} - \mu\right)^2 + 0.185\left(T_{0.05} - \mu\right)^2 + \left(T_{0.95} - \mu\right)^2 . \tag{6}$$

Estimates (5) and (6) by Pearson and Tukey (1965) generate smaller errors than estimates (3) and (4) of the classic *PERT*.

The experts' opinions on the pessimistic, optimistic and modal duration have a considerable impact on the error of distribution parameter estimates. If the input is to be given by a group of experts, it is advisable to use the median and not the mode of their opinions to reduce the error.

3 PERT ASSUMPTIONS OF THE ANALYSIS OF A NETWORK MODEL IN THE FUNCTION OF TIME

PERT assumes that the expected value of the project duration and its variance equal, respectively, the sum of expected durations and the sum of variances of the critical processes. However, this assumption is statistically sound only if the random variables being added are independent and if a process starts after only one of its predecessors has finished.

If a process start is conditioned by a number of predecessors' being completed, the distribution of the random variable of the event that represents the successor's start becomes a complex problem (described by e.g. Cox (1995) and Clark (1961)). Therefore, PERT networks are often analysed by means of the *Monte Carlo* simulation.

The problem is illustrated by the following example (network model presented in Figure 2) where the problem consists in estimating the early start of the event 4.

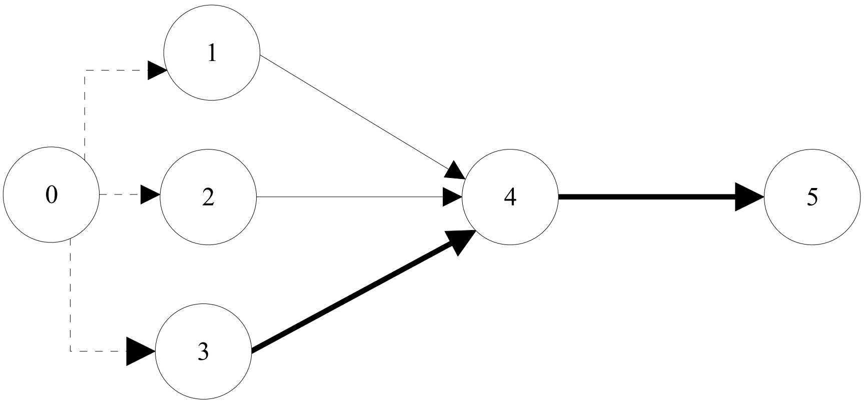

Figure 2. Example: predecessors of a process 4-5 (activity on arrow model)

In the example, the durations of processes were assumed to be random variables of beta-*PERT* distribution, and that their parameters were calculated on the basis of three estimates: optimistic, modal and pessimistic (respectively t_a, t_m, t_b):

- t_{1-4} : t_a =16, t_m =22, t_b =40,
- t_{2-4} : t_a =10, t_m =15, t_b =40,
- t_{3-4} : t_a =23, t_m =24, t_b =29.

Table 2 lists values of each process duration provided by means of a random number generator. Event 4 earliest occurrence is possible when all the predecessors of the process 4-5 have been completed. For 20% of possible cases, the moment of event 4 is not decided by the duration of the critical process 3-4.

Table 2. Generated process durations (example)

Run	t_{1-4} ($\mu = 24, \sigma = 4$)	t_{2-4} ($\mu = 18{,}33, \sigma = 5$)	t_{3-4} ($\mu = 24{,}67, \sigma = 1$)	Earliest occurrence of event 4
1	21,00	15,98	**24,30**	24,30
2	24,02	**27,52**	23,39	27,52
3	18,47	12,51	**23,04**	23,04
4	**30,18**	17,34	25,66	30,18
5	22,83	11,57	**26,35**	26,35
6	25,03	17,80	**26,24**	26,24
7	21,45	18,97	**24,65**	24,65
8	19,48	17,97	**25,42**	25,42
9	25,25	24,32	**25,31**	25,31
10	21,76	10,50	**25,93**	25,93
11	21,00	15,98	**24,30**	24,30

If the duration of all processes on a critical path is substantially greater than the duration of the processes on the other paths that connect the project network's start and finish nodes, the interaction of processes at network "sinks" (as illustrated in Figure 2) is not strong and *PERT* may provide accurate results (Jaworski and Biruk (2000)). In the other case, neglecting the analysis of non-critical paths may lead to serious underestimation of the project finish date.

The exact calculation of the probability of not exceeding the project's due date may be done by means of a formula:

$$P(T \le t) = 1 - \left[P(T_1 > t) + P(T_1 \le t \wedge T_2 > t) + \ldots + P(T_1 \le t \wedge T_2 \le t \wedge \ldots \wedge T_{n-1} \le t \wedge T_n > t)\right], \quad (7)$$

where T is the random variable of project duration and T_i is the random variable of duration of processes on the path i ($i = 1, 2, \ldots, n$).

Calculating the probabilities is a complex task. Therefore, practical application of Equation 7 is limited.

The *PERT*'s formulas for calculating distribution parameters of the project duration are correct only under assumption of independence of durations of critical processes. In real life, there are a number of factors, such as weather, that may affect a number of parallel processes in the same way, so the variables of process durations may be positively correlated. A positive correlation may occur between durations of processes executed by the same subcontractor. It is also possible that a negative correlation occurs – an example would be shifting limited resources from non-critical to critical tasks to assure that project due date is met, which may cause delays of non-critical processes.

Furthermore, according to the central limit (Lindeberg's) theorem, *PERT* assumes that the random variable that represents the project duration is of normal distribution as a sum of random variables being durations of critical tasks. The normal distribution would be an adequate approximation of the project duration distribution if the number of critical processes is large enough (more than 30, but smaller numbers as 20 or 10 are also accepted by practitioners). However, the accuracy of the probability estimation of meeting a project due date is conditioned by not only the number of critical processes, but also similarity of process duration distribution types.

3 EXAMPLE

Figure 3 presents a network model for the case study – a modernisation project of a partly two-storey, post and beam structured building. The number of critical processes is 15. Table 3 lists the tasks of the work breakdown structure together with their duration estimates.

The project expected duration and standard deviation calculated according to *PERT* are $\mu = 291.83, \sigma = 4.78$, respectively.

In order to verify the accuracy of the results, a Monte Carlo simulation. The project duration mean of 10000 simulations was $\mu = 291.54$, and the standard deviation was $\sigma = 5.53$.

Figure 4 compares the cumulative probability density functions: the one obtained in the course of simulations, and the one of a normal distribution and parameters established by PERT. The maximum error of duration estimate at the predefined reliability level is less than three working days, which is about 1% of the project duration. This accuracy level seems more than adequate for practical engineering applications.

3 CONCLUSIONS

PERT is a simple tool that supports planning projects carried out in random conditions, and, as such, often used in practice. The assumptions of *PERT* made it possible to reduce the complexity of network model analyses but, at the same time, affected the accuracy of time estimates of individual project events and the project as a whole. Understanding these assumptions allows the planner to interpret the results of *PERT* calculations and to prepare more reliable project programmes.

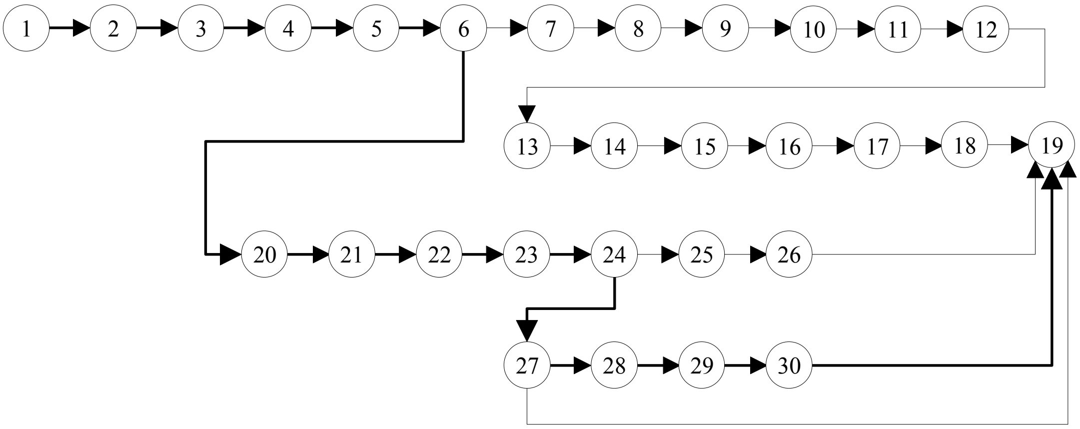

Figure 3. Case study project network model

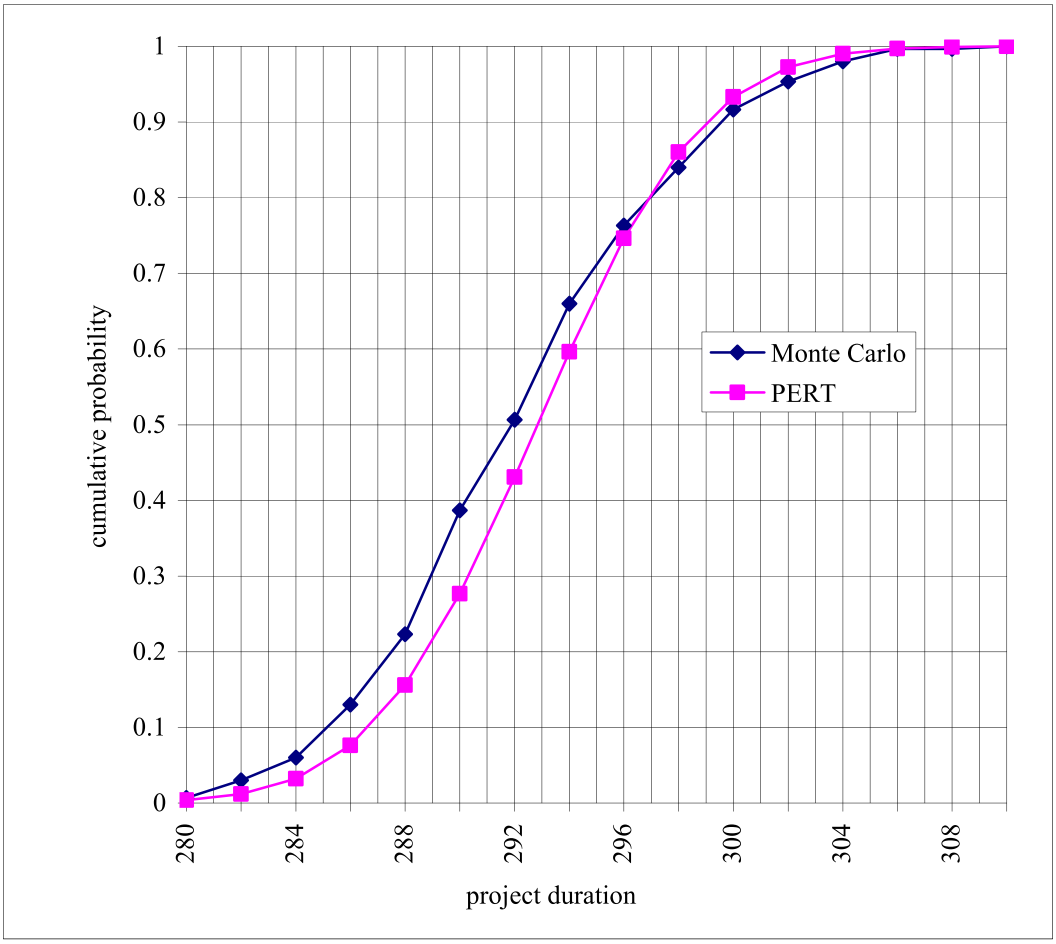

Figure 4. Cumulative probability distribution: simulation based and according to PERT

Table 3. Case study project tasks with estimates of optimistic, most likely and pessimistic durations, in working days

Activity	Activity title	t_a	t_m	t_b
1–2	Removing floor finishings, plastering and wall claddings	57	62	70
2–3	Demolition of partition walls	9	12	17
3–4	Dismantling steel structures	5	6	8
4–5	Dismantling aluminium structures	1	2	3
5–6	Demolition of RC structure elements	5	6	8
6–7	Assembling supproting structures for AC units	3	4	6
7–8	Assembling precast concrete elements	2	3	4
8–9	Earthworks - trenches	6	8	11
9–10	RC foundations	3	4	5
10–11	RC stairs and plates	1	2	3
11–12	Substructure waterproofing	3	4	6
12–13	Backfill	2	3	4
13–14	Dismantling elements of roof cladding with gutters and downpipes	1	2	3
14–15	Roof cladding	2	3	4
15–16	Thermal insulation of external walls and substructure	55	59	65
16–17	Roof gutters and downpipes	10	13	16
17–18	External cladding	5	6	8
18–19	Landscaping works	1	2	3
6–20	Partition walls	5	6	8
20–21	Steel gates and doors, aluminium facades, partitions and doors	5	6	8
21–22	PVC windows	3	4	5
22–23	Plastering	34	37	43
23–24	Internal wall cladding	34	37	45
24–25	Underfloor insulation	3	4	5
25–26	Subfloors	1	2	3
26–19	Floor tiling	31	34	39
24–27	Painting	30	34	40
27–28	Suspended ceilings, plasterboard claddings and partitions	60	66	75
28–29	Internal doors	4	5	7
29–30	Assembly of awnings	1	2	3
30–19	Heater screens	2	3	4
27–19	Other floor finishes	42	48	56

ACKNOWLEDGMENTS

The paper is based on the research sponsored by the Polish Ministry of Science and Higher Education (grant N N506 254637).

REFERENCES

AbouRizk S., Knowles P., Hermann U. R. 2001. Estimating labor production rates for industrial construction activities. *Journal of Construction Engineering and Management* 127(6): 502-511.

Biruk S., Jaskowski P. 2008. Simulation modelling construction project with repetitive tasks using Petri nets theory. *Journal of Business Economics and Management* 9: 219-226.

Chua D.K.H., Kog Y.C., Loh P.K., Jaselskis E.J. 1997. Model for construction budget performance-neural network approach. *Journal of Construction Engineering and Management* 123(3): 214-222.

Clark C.E. 1961. The greatest of finite set of random variables**.** *Operations Research* 9: 145-162.

Cox A.A.A. 1995. Simple normal approximation to the completion time distribution for a PERT network. *International Journal of Project Management* 13: 265-270.

Dawood N. 1998. Estimating project and activity duration: a risk management approach using network analysis. *Construction Manngement and Economics* 16: 41-48.

Hanna A.S., Gunduz M. 2005. Early warning signs for distressed projects. *Canadian Journal of Civil Engineering* 32(5): 796-802.

Hon-Siang L., Somarajan C. 1995. A proposal on improved procedures for estimating task-time distribution in PERT. *European Journal of Operational Research* 85: 39-52.

Jaworski K.M., Biruk S. 2000. A model of construction project based on Petri nets theory. *Archives of Civil Engineering* XLVI: 71-82.

Jaselskis E.J., Ashley D.B. 1991. Optimal allocation of project management resources for achieving success. *Journal of Construction Engineering and Management* 117(2): 321-340.

Jaskowski P. 2008. Designing the structure of a construction project operating system using evolutionary algorithm. *Archives of Civil Engineering* LIV: 371-394.

Jaskowski P**.,** Biruk S., Bucoń R. 2010. Assessing contractor selection criteria weights with fuzzy AHP method application in group decision environment. *Automation in Construction* 19: 120-126.

Jergeas G., McTague R. 2002. Construction productivity: an auditing and measuring tool. *AACE International Transactions* CS91-CS104.

Keefer D.L., Verdini W.A. 1993. Better estimation of PERT activity time parameters. *Management Science* 39: 1086-1091.

Kog Y.C., Chua D.K.H., Loh P.K., Jaselskis E.J. 1999. Key determinants for construction schedule performance. *International Journal of Project Management* 17(6): 351-359.

Lee S., Halpin D.W. 2003. Predictive tool for estimating accident risk. *Journal of Construction Engineering and Management* 129(4): 431-436.

Lichtenstein S., Fischhoff B., Philips L. 1982. *Calibration of probabilities: the state of the art to 1980*. In: Kahneman D., Slovic P., Tversky A. (ed.). *Judgement under uncertainty: heuristics and biases*. New York: Cambridge University Press.

Littlefield T.K., Randolph P.H. 1987. Reply an answer to Sasieni's question on PERT times. *Management Science* 33: 1357–1359.

MacCrimmon K. R., Ryavec C. 1964. An analytical study of PERT assumptions. *Operational Research* 12: 16-37.

Nasir D., McCabe B., Hartono L. 2003. Evaluating risk in construction-schedule model (ERIC-S): construction schedule risk model. *Journal of Construction Engineering and Management* 129(5):518-527.

Neil J.M., Knack L.E. 1984. Predicting Productivity. *AACE International Transactions* H.3.1-H.3.8.

Ovararin N., Popescu C. 2001. Field factors affecting masonry productivity. *AACE International Transactions* ES91-ES97.

Pearson E.S., Tukey J.W. 1965. Approximate means and standard deviations based on distances between percentage points of frequency curves. *Biometrics* 52: 533-546.

Perera A., Imriya K. 2003. Knowledge-based system for construction cost control. *AACE International Transactions* IT101-108.

Schatteman D., Herroelen W., Van de Vonder S., Boone A. 2008. Methodology for integrated risk management and proactive scheduling of construction projects. *Journal of Construction Engineering and Management* 134(11): 885-893.

Shi J. J. 1999. A neural network based system for predicting earthmoving production. *Construction Manngement and Economics* 17: 463-471.

Sonmez R., Rowings J.E. 1998. Construction labor productivity modeling with neural networks. *Journal of Construction Engineering and Management* 124(6): 498-504.

Woodward C. 2003. Project productivity analysis: what is 1.0? *AACE International Transactions* ES11-ES15.

Zayed T.M., Halpin D.W. 2005. Pile Construction Productivity Assessement. *Journal of Construction Engineering and Management* 131(6): 705-714.