International Journal of Science and Technology Volume 6 No. 8, August, 2017

Framework for Incorporating Risk in Project Scheduling Using Pert Paradigm

Oliver E. Charles-Owaba & **Adepeju A. Opaleye**
Department of Industrial and Production Engineering, University of Ibadan, Nigeria

ABSTRACT

Schedule time overrun, a perennial indicator of project failures, has posed intractable challenge to project managers. Though procedures for project scheduling are rife, the problem of accurately determining project times persists. Numerous unforeseen activity-related difficulties during execution have resulted in poor activity duration estimates. Consequently, risk analysis has now become an attractive component of project scheduling. However, existing definitions of risk are yet to provide a basis for quantifying observed data based risk values. In this paper, a theoretical framework for quantifying risk is suggested for activities scheduling using the definition that statistical risk is a "chance" that some deleterious events may occur against plan. Adopting the PERT paradigm, historical/field data-based features for estimating risk and risk-incorporated activity duration were proposed. Combining with a model for identifying risky activities, a procedure for deriving a set of risk-free and risk-incorporated project schedules for more effective time management in risky business environment was suggested, illustrated and evaluated.

Keywords: *Statistical Risk, Project Activity Scheduling, Risk Analysis, Risk-Incorporated Schedule.*

1. INTRODUCTION

Global competitive pressures are causing organizations to find ways to better meet the need of their customers, to reduce cost and to increase productivity. As such, managers saddled with responsibility of on-time completion adopt project-based planning techniques in their respective work system.

For a given set of activities of a specific project, the problem of predicting project duration has three aspects: network diagramming; estimation of the respective durations and identification of critical activities. Acceptable solution procedures for the first and last aspects have established that project duration is the sum of the respective durations of the critical activities (Azaron and Fatemi Ghomi 2008). Hence, the quality of a solution to the project duration prediction problem solely depends on the activity duration estimates as poor estimates of even one may adversely affect project completion time. Unfortunately, the task of attaining accurate estimate is made difficult by the high degree of uncertainty introduced by multiple random internal and external factors in project environments.

Program Evaluation and Review Technique (PERT) remains one of the most popular planning models to account for possible uncertainty in activity duration estimation. The authors proposed three point estimates: optimistic, most of the times and pessimistic values which are elicited by experts implicitly assuming the beta distribution is suitable for risk analysis. Using the pre-assumed probability $(\frac{1}{6}, \frac{4}{6}, \frac{1}{6})$ of observing each value, expected activity duration was determined as a weighted average. Since the model publication over six decades ago, it has been widely criticized based on its underlying assumptions.

Many authors have suggested technical improvements of the weighted average model. The first group of authors approve beta distribution but condemns the restrictions on its shape parameter (Shankar & Sireesha, 2010). To eliminate possible error associated with the use of beta distribution function, alternative distributions such as; triangular, lognormal, normal, weibull etc. are considered (Trietsch et al., 2012). It is however essential that the selected distribution to model duration data accurately reflect the properties of the data. A few other authors propose discretization method which simply represents a continuous random variable (such as activity duration) with a finite set of fractile points and associated probabilities. This exacerbate the duration prediction problem as fractile points and associated probability (weight) have to be pre-specified before an expert is asked to elicit values corresponding to them (Lau & Somarajan, 1995).

The above criticisms reveal that fundamentals of PERT's activity duration estimation postulation: that actually observed duration may agree with well thought out plan (most of the time estimate); fall below plan (optimistic) or exceed plan (pessimistic) is not theoretically untrue but the choice of statistical distributions, manner of estimating parameter values and derivation of duration from parameter values are unjustified (Khamooshi & Cioffi, 2013).

Outside the PERT paradigm, there are other attempts to improve duration prediction accuracy. Proactive scheduling considers intelligent insertion of time buffers at strategic point in the schedule (Davenport, et al., 2014). Here, buffer size and its location on the schedule are based on simulation study of possible risk scenarios. Cho (2006) considers developing a Bayesian based expert system which allows updating of project completion time as durations of completed activities are observed. This approach do not allow prediction of project completion time as duration changes as project progresses. When activity duration is vague, for example, in novel project situations, advocates of fuzzy theory recommend the use of fuzzy numbers for modeling activity durations rather than stochastic variables. Instead of probability distributions, these quantities make use of membership functions, based on possibility theory (Herroelen & Leus, 2005). Risk management researchers on the other hand (Creemers, et al., 2014), considers risk mitigation as an antidote for on time project completion. They propose indices for ranking activities or risk events requiring mitigation efforts. Existing methods are based on simulation study of possible risk impact on predicted project completion time.

It is evident that inputs to existing activity duration prediction models are based on synthetic data; expert elicited or randomly generated, associated to PERT paradigm. At the time of PERT development for the Polaris weapon system problem, risk analysis in scheduling problem was unprecedented and PERT developers have no prior data on which to develop a more accurate approach. As such, they assumed time estimate may be elicited by experts who understand the performance of each project activity. Hence, the present methods can handle all cases when adequate historical data does not exist. However, with the adoption of project planning techniques in many work systems such as manufacturing, service, maintenance etc. where historical data situations exist, accuracy of project duration and maybe cost prediction may be improved with availability of historical data (Opaleye et al., 2017; Meyer and Visser, 2006).

Perhaps, the most attractive quantitative approach is the one suggested by Hall and Johnson (2003). They sought and applied field data on "most of the times" durations, "pessimistic" durations and the likelihood of overrun per activity in their effort to estimate new project duration but provide no theoretical basis for quantification in situations with secondary or historical project data. This may be associated with the apparent confusion by the introduction of two different quantitative definitions of risk by the management research community. Hall and Johnson (2003), Valerdi (2005) and Park (2007) define risk as the probability that some expected target value will be exceeded while others choose to define it as the product of the probability of risky event and the associated consequence (Corotis & Hammel, 2010; Pinto, 2012). Notice that both definitions mention probability without suggesting a clue on data-based quantification approach. The former definition appears more convenient for estimating activity over-run (Hall & Johnson, 2003) while the latter may be more useful for quantitative approach to two-dimensional ranking of risk factors/situations (Pinto, 2007); this provides information for risk mitigation planning.

In this study, a PERT-based theoretical framework for estimating the scheduling parameters using historical, secondary or field data is proposed**.** In particular, models for identifying project risky activities (bottleneck) and generating a set of risk-incorporated schedules are suggested and illustrated. In the following sections, the theoretical foundation for the proposed framework is presented and empirically illustrated. We also compared the performance of our method with the traditional PERT using Chebyshev model.

2. THEORETICAL CONSIDERATION

2.1. Assumptions

We commence by rephrasing the observations of Malcolm *et al* (1959) explicitly stated by Littlefield and Randolph (1987) as assumptions

Assumption 1.0: The duration of each project activity varies randomly due to unforeseen work conditions.

Assumption 2.0: There exist three distinct but mutually exclusive project work conditions which impact on the value of project activity duration.

Assumption 2.1: Stable work conditions which allow for normal project planning and execution. Natural disaster would not occur; labour unrest may occur but would be easily contained; most change factors would be anticipated and contingency plans put in place to effectively control them. It is a work situation in which organizations may regard as normal because most operational activities are in control. Periodic plans are based on such "all- things-being-almost-alright" conditions. Statistically, actual activity duration values experienced at instances of this type are those with the highest likelihood of occurrence. We tag such observations as "***most of the times***" values or the ***Mode***.

Assumption 2.2: Exceptionally favourable circumstances which are beyond human control. Some individuals may be very highly motivated naturally at random occasions, thereby breaking achievement records at such periods. The mood of the workforce may be aroused positively by some unforeseen natural favourable circumstances to influence organizational performance. In any such situations, the observed value of achievement is better than the "most of the times" observations. For activity duration, for example, it is the '***optimistic time***' values.

Assumption 2.3: There may exist exceptionally unfavourable but unforeseen work circumstances which project managers will want to mitigate. This may be caused by poorly defined requirements, design "drift" exceeding technical capabilities, accepting an unproven technology, limited availability of needed skills, funding limits, etc. Natural disasters too may strike unexpectedly or labour unrest become riotous far beyond expectation. The observed organizational performance is worse than that of the normal plan; it is '**pessimistic'** scenario. Activity durations observed at instances of this type are, naturally, longer than the mode or "most of the times" values.

Notice that these are all possible scenarios. Corporate annual plans are usually developed from very carefully thought-through information collected over the years, screened, modified and, possibly, verified. Statistically, such data are called 'most of the times' estimates. Once such a plan has been accepted, it becomes the personality of management: every effort is geared towards its attainment.

Now, consider a particular activity performed several times in the history of a corporate body. For the long time experience on the activity by management, duration may behave as planned; that is, exhibit "most of the times" values. However, when a deleterious unforeseen event occurs, this is the source of "pessimistic" duration values, the problem created for management by uncertainty. Of course, "optimistic" values too may occur because of the inconsistency of the human nature. These, however, pose no serious threats to management as they are usually handled through micro-management strategies at operation points. It is the aforementioned eventualities of uncertainty that are tackled in this study. In the following, an attempt is made to classify activity durations premised on the above assumptions with the hope of providing a theoretical basis for risk analysis in project scheduling and mitigation.

2.2. Definitions of Scheduling Parameters

At any moment of work, things may work as expected to the peace of all; better than expected, to their pleasant amazement; or worst than anticipated, a disappointment. It may not be otherwise. It therefore, appears reasonable to idealize that a performance datum point (observed duration) will belong to one and only one of these possibilities: as planned, better or worse than planned. Considering activity duration as performance measure, this principle will be adopted for classifying a population of durations as per the following definitions.

Definition 1: $\mathbf{A_i}$ is the **set** of activity durations observed at instances of the "exceptionally favourable circumstances" mentioned in assumption 2.2. They are actual activity durations collected at the "optimistic" instances of work periods.

Definition 2: $\mathbf{B_i}$ is the **set** of activity durations collected under the exceptionally unfavourable or "pessimistic" work circumstances (assumption 2.3).

Definition 3: $\mathbf{M_i}$ is the **set** of activities duration collected under the planned or "most of the times" expected project work circumstances (assumption 2.1).

An observed activity duration d_i may be according to the plan, better than the plan or worse i.e. a datum point, $\mathbf{d_i}$**,** may occur only in one performance situation, then sets A_i, B_i, and M_i are mutually exclusive. We formalise this as an assumption.

Assumption 3: $\mathbf{A_i \cap M_i \cap B_i . = \{\emptyset\}}$

Combining information from assumptions 2.1, 2.2, 2.3 and 3, the following inequalities hold:

$\text{Max}\ \{A_i\} < \text{Min}\ \{M_i\}$

$\text{Max}\ \{M_i\} < \text{Min}\ \{B_i\}$

Since $\text{Min}\{M_i\}$ is naturally less than or equal to $\text{Max}\{M_i\}$, then

$\text{Max}\ \{A_i\} < \text{Max}\ \{M_i\} < \text{Min}\ \{B_i\}$

(1.1)

This happens to be part information in the definition of statistical ***Mode*** which is hereby formally stated.

Definition 4: Given a finite population of activity durations, statistical **MODE**, $\mathbf{M_i}$, is a subset of historical data satisfying the following conditions:

1) the inequalities in (1.1); and
2) $P(M) \geq P(A)$;
3) $P(M) \geq P(B)$ with equality holding only for $A_i = M_i$, $B_i = M_i$.

where $P(A_i)$, $P(M_i)$ and $P(B_i)$ are the probabilities of observing subsets $\{A_i\}$, $\{B_i\}$, and $\{M_i\}$ respectively . From this definition, the problem of identifying sets $\{A_i\}$, $\{B_i\}$, and $\{M_i\}$ from historical data is reduced to that of finding the set,$\{M_i\}$. This can be done by any classification mean which satisfies definition (4) or relying on the findings of Scott(1979) and Wand (1997) that an optimal bin width, V, exists for a sample of data and that M can be found by data classification using the optimal bin width. Scott's function for the optimal bin width is given as follows:

$$V_i = 3.49\sigma_i/Z_i^{1/3} \qquad (1.2)$$

where σ is the standard deviation and Z is the size (population) of activity duration i.

Hence, for a large set of data, the mode corresponding to an optimal bin width can be determined using one of the several statistical software packages available. Alternatively, frequency count may be carried out to determine the set, M, as per **Definition 4 and** confirmed in Opaleye and Charles-Owab(2014).

2.3. Development of Probabilities Estimation Function

Based on assumptions 1 to 3 and Definitions 1 to 4, the following obvious but important observation for estimating empirical probabilities of observing PERT scheduling parameters is formally stated as a theorem without further proof.

THEOREM 1

Given a large set of historical durations of activity X_i, with size, Z_i and the existence of a unique mode, M_i as per Definition 1.4, X_i contains the distinct subsets $\{A_i\}$, $\{B_i\}$, and $\{M_i\}$ such that

$$Z_i = \alpha A_i + \alpha M_i + \alpha B_i \quad (1.3)$$

where "α" denotes "the number of elements".

The theorem implies that, as long as observed historical data have a unique mode, the pessimistic and optimistic subsets, $\{A_i\}$ and $\{B_i\}$ can be found by eye inspection or sorting using Definition 4. This useful information is hereby stated formally.

COROLLARY 1.1

Given Z durations for activity i and the subset $\{M_i\}$, the associated subsets, $\{A_i\}$, $\{B_i\}$ are identified by assigning values less than Min$\{M_i\}$ into $\{A_i\}$ and those greater than Max$\{M_i\}$ into $\{B_i\}$.

These results will be useful for developing simple functions for estimating the empirical probabilities that activity duration may or may not over-run. Accordingly, let d_i be a point in set X_i mentioned in Theorem 1 and αd_i, the number of times d_i occurs in X_i. Hence, the probability of observing d_i, $P(d_i)$ may be estimated with the following expression:

$$P(d_i) = \alpha d_i/Z_i \quad (1.4)$$

Similarly, the probabilities of observing the sets $\{A_i\}$, $\{B_i\}$, and $\{M_i\}$ may be estimated with the following:

$$P(A_i) = \alpha A_i/Z_i \quad (1.5)$$

$$P(M_i) = \alpha M_i/Z_i \quad (1.6)$$

$$P(B_i) = \alpha B_i/Z_i \quad (1.7)$$

Applying the principles of conditional probability, the probability of observing d_i as a member of {A}, {M} and {B} may be estimated with the following respective expressions:

$$P(d_i\ \varepsilon\ A_i) = P(d_i)/P(A_i) = (\alpha d_i/Z_i)/(\alpha A_i/Z_i) = \alpha d_i/\alpha A_i \quad (1.8)$$

$$P(d_i\ \varepsilon\ M_i) = P(d_i)/P(M_i) = (\alpha d_i/Z_i)/(\alpha M_i/Z_i) = \alpha d_i/\alpha M_i \quad (1.9)$$

$$P(d_i\ \varepsilon\ B_i) = P(d_i)/P(B_i) = (\alpha d_i/Z_i)/(\alpha B_i/Z_i) = \alpha d_i/\alpha B_i \quad (1.10)$$

2.4 Risky Activities and Project

Notice that expression (1.7) is the probability that an observed value of the duration of activity i will exceed the planned or the mode. This is another way of saying that $P(B_i)$ is the probability of activity i having an over-run. Using the definition of risk by Park (2007), expression (1.7) is the Statistical Risk of activity i. With respect to project duration, the definition of Statistical Risk follows.

Definition 5: **Statistical Risk** is the probability that, relative to planned duration, actual activity duration may over-run.

From this definition, one may also define risky environment as follows; given a large set of historical data of an activity carried out in a well-defined area, say a nation, one may describe the project environment as risky or non-risky using the concept of risky or non-risky activities. Clearly, if no risky activity is found among a large set of activities carried out many times in the environment for a long period, then the project environment may be described as non-risky. It is risky environment, otherwise. Based on this, project risky environment is formally defined as follows:

Definition 6: A risky project environment is one in which risky project activities have been identified.

Definition 7: Degree of environmental risk for any particular project is the mean of the probabilities of the project critical activities given as

$$\sum_i \frac{P(B_i)}{c}, \quad where\ i = 1,2,\ldots.c(\ number\ of\ activities\ on\ the\ critical\ path)$$

In a risky project environment, activity duration may be estimated using empirical probabilities where historical data exist.

2.5 Estimation of Project Activity Duration

Depending on the degree of risk in an environment and availability of historical data, estimate of activity duration may be made using the PERT approach or the weighted average i.e. observed-data equivalent of the traditional PERT duration, (OPD_i) given as;

$$OPD_i = P(A_i)\,\Phi_i + P(M_i)\,\Omega_i + P(B_i)Б_i \qquad (1.11)$$

where the one-point estimate of the optimistic, mode and pessimistic durations are defined as the weighted average of all the elements in {A}, {M} and {B}. That is, one-point estimate of the optimistic duration is

$$\Phi_i = \sum_{k=1}^{a}\left(P(d_k \varepsilon A_i)d_k\right) \qquad (1.12)$$

where a, are the number of elements in set $\{A_i\}$; and $P(d_k \varepsilon A_i)$ is computed using expression (1.8). Similarly, the point estimates of Ω_i and $Б_i$ values is given as:

$$\mathbf{\Omega}_i = \sum_{k=1}^{m}\left(P(d_k \varepsilon M_i)d_k\right);\ \text{m is number of elements in } \{M\}$$
(1.13)

And

$$Б_i = \sum_{k=1}^{b}\left(P(d_k \varepsilon B_i)d_k\right);\ \text{b is number of elements in } \{B_i\}$$
(1.14)

Expression (1.11) may be used to estimate data-based risk incorporated project activity duration.

In other project risky environment where the overrun of activity duration may be high, the observed over-run, Δ_i, may be given by the expression, (Hall and Johnson, 2003):

$$\Delta_i = Max\{B_i\} - \mathbf{\Omega}_i \qquad (1.15)$$

applying the principle of statistical expectation, the value of the over-run of activity duration may be estimated with the following expression:

$$E(\Delta_i) = \text{Expected Over-run} = P(B_i)\times\Delta_i \qquad (1.16)$$

Thus, it may be possible to improve the accuracy of estimating the duration of each risky activity by simply adding the estimated over-run to the PERT "most of the times" or mode estimate. Referred to as **Risk Incorporated Duration** of activity i (RID_i) may be expressed as follows**:**

$$RID_i = \mathbf{\Omega}_i + P(B_i)\times\{\Delta_i\} \qquad (1.17)$$

Applying the above expression, duration of every risky project activity is estimated and the critical path analysis carried out for a risk incorporated schedule. Notice that this is a deviation from the PERT traditional approach of combining expert opinion and fixed probabilities $(\frac{1}{6}, \frac{4}{6}, \frac{1}{6})$ for estimating duration.

Another important observation stems from the definition of Statistical Risk. If an activity has a zero probability of Over-run, then the activity can be said to be not Statistically Risky; it is Risky, otherwise. This useful observation is summarized in the following theorem.

THEOREM 2

Given a large but finite population of activity i durations, Z_i; collected from the same business environment, having a unique mode, M_i, 'i' is a statistically risky activity if and only if subset B_i is not empty, i.e. $\mathbf{B_i} \neq \{\emptyset\}$

PROOF:

By *combining expressions 1.3 to 1.7, it is easily shown that whenever $\mathbf{B_i} \neq \{\emptyset\}$, then*

$P(B_i) = \alpha B_i/Z_i > 0.0$; and when $B_i = \{\emptyset\}$; $P(B_i) = \alpha B_i/Z_i = 0.0$; which confirms the Theorem

It may therefore, be important to note that for an environment that have been observed over time in which all its activities are considered to be statistically not risky, such may also be considered as not risky. This is summarized in the following theorem.

THEOREM 3

Given that all critical activities for a specific project are not risky; $P(Bi) = \frac{\alpha Bi}{Zi} = 0.0\ \forall\, i$ *then from expression 1.17 becomes*

$$R\mathbf{ID_i} = \Omega_i \qquad (1.19)$$

i.e *risk-free schedule equals the risk incorporated schedule with no further proof.*

The three theorems also provide features for the use of field data similar to Hall and Johnson's (2003) approach. In collecting valid field data for risk analysis, questionnaire may be structured to extract the following information from experienced individuals who have worked or supervised extensively the activity in similar projects executed in identical environments:

- Total number of times activity was carried out, (Z_i);
- *“Most of the times”* observation of activity duration, Ω_i;
- Number of times the *“most of the times”* observation has been exceeded in the history of the activity, αB_i; and
- The worst case duration, $Max\{B_i\}$.
- Identify all possible causes of deviation from planned duration.

Notice that with such data, the earlier presented functions can be used to estimate the risk and possible overrun per activity in the proposed model.

3. RISK INCORPORATED PROJECT SCHEDULING PROCEDURE

Step 1: Identify the project activities and their pre-incidence relationships using information from the work breakdown structure and activities technical requirements.

Step 2: Collect set of historical/field data (durations) on each project activity.

Step 3: Determine the subset M_i by frequency count if feasible; else compute the optimal bin width of durations per activity using expression 1.2 and then find the associated subset M_i by classification and identify sets A_i and B_i based on definition (4).

Step 4: Compute the probabilities: P(A), P(M), P(B) for each activity using expressions 1.5 to 1.7.

Step 5: Estimate the one-point Mode per activity, one-point optimistic and pessimistic duration using expression 1.12 to 1.14;

Step 6: Then compute the observed- data equivalent of traditional PERT (OPD) and Risk Incorporated Duration (RID) using expressions 1.11 and 1.17 respectively.

Step 7: Identify the risky activities using information in Theorem 2 and causes of risk to further classify the risky activities into two categories: those with controllable causes (personnel-related issues, maintenance, material supplies, etc) and another with uncontrollable causes (inclement weather, natural disaster, militancy, labour unrest, etc) for treatment.

Step 8: Carry out the Critical Path Analysis using first, the Mode (risk-free Critical Path); second, PERT with data-based estimated probabilities (OPD) and the Risk Incorporated Duration (RID).

Step 9: Develop the Gantt charts for each of the critical paths in step 8 and determine the respective project durations.

4. PERFORMANCE EVALUATION

Performance of OPD, the traditional PERT approximations by Malcolm et al.,(1959) and the Risk Incorporated Duration (RID) in handling situations of possible schedule overrun was evaluated using the Chebyshev model (Rao et al.,2013). Chebyshev’s theorem states that for a wide class of probability distributions, at least $1-1/k^2$ of the distribution's values are within k standard deviations (σ) of the mean(μ) . Since the exact probability distribution of the project path is unknown but the variance of the path is known, the Chebychev Inequality could be used to get an estimate of the probability of project completion at time (X) as follows:

$$P(X) = \frac{X-\mu}{\sigma} \tag{1.20}$$

Where; μ and σ are mean and standard deviation of project completion time where estimated using the best distribution that describe each project activity.

5. APPLICATION AND DISCUSSION OF RESULTS

We applied the developed procedure using a type of building construction commonly found in West Africa. It is a 3-storey building with each floor having two 3-bedroom apartments i.e. a block of flats containing six separate three bedroom/residential flats. For its commercial value, it is common in most of the residential estates in Nigeria metropolitan cities. Our benevolent construction firm specialized in constructing this class of building. In order to collect data on each of these activities, we vet past and on-going project records, baseline/working schedule, minutes of meetings held at different time of the projects. Relevant management personnel were also interviewed, operations observed and durations of activities for on-going project recorded. However, detailed information about the causes of deviations was stone –walled.

Table 1.0 shows the building project activities’ description, labels and the immediate predecessors. Table 2.0 also contain set of historical and field data (durations) on each activity for 25 periods.

Table 1 Description of Project Work Breakdown

Code	Activity Description	Predecessor (s)
A	Mobilization	-
B	Site Clearing and setting out	-
C	Excavation trench and foundation	B
D	Reinforced concrete column	C
E	Floor beams, suspended slab and associated works for F (1)	D
F	Rough MEP piping for floors (1) & (2)	E
G	Floor beams, suspended slab and associated works for F(2)	F
H	Staircase	A,G
I	Internal and external walls	G
J	Roof and roof covering	G
K	Finish MEP installations	I
L	Plastering	I
M	Wall finishing	H,K,L
N	Ceiling finishing	I,J
O	Floor finishing	M,N
P	Fixtures and fittings	N
Q	Painting and decorating	P
R	Drainage, External works & demobilization	O,Q

Table 2 25-period historical/field data (durations) for 25-activities Building Project

K	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	36	5	5	12	43	41	16	38	30	21	7	14	14	21	17	29	18	32
2	17	7	4	10	44	40	13	38	33	33	8	27	25	17	7	17	20	36
3	27	10	5	12	43	37	16	38	40	12	5	14	34	21	17	29	20	32
4	27	7	8	12	55	39	7	40	39	18	12	22	23	21	22	18	27	32
5	20	7	8	12	36	40	11	43	33	19	13	27	24	24	18	34	23	35
6	20	6	7	6	36	39	9	39	24	27	11	18	27	20	20	23	16	25
7	28	4	7	7	39	39	11	35	30	17	20	19	29	11	17	16	21	31
8	31	11	4	12	32	42	12	46	28	16	18	22	20	28	22	16	17	34
9	37	9	7	10	34	40	14	41	34	21	9	23	22	11	17	10	22	50
10	22	6	5	8	58	37	10	42	27	21	16	13	25	23	18	17	18	15
11	31	4	6	6	57	42	14	40	35	21	12	17	24	15	12	25	18	35
12	20	7	8	12	39	41	10	36	45	16	14	18	24	27	18	16	18	43
13	18	9	6	9	46	43	18	43	31	18	19	21	30	25	21	21	25	52
14	35	5	6	11	41	36	16	43	31	15	6	17	27	18	14	16	20	28
15	27	6	7	12	41	46	13	42	30	16	4	26	33	29	20	19	17	15
16	23	8	5	12	42	40	16	38	31	25	4	19	28	26	28	8	14	38
17	26	6	4	7	39	41	16	44	25	27	12	22	25	20	25	25	16	55
18	15	5	7	7	28	36	11	45	32	25	7	27	28	16	14	15	19	45
19	31	12	8	12	59	40	18	44	30	29	7	22	23	29	23	16	20	30
20	28	4	7	8	53	38	12	36	24	19	20	14	25	29	25	19	21	34
21	29	8	4	6	36	47	12	47	30	21	19	21	23	18	22	26	16	48
22	19	2	7	12	29	36	12	49	31	21	4	17	32	25	20	13	21	43
23	24	4	5	5	40	35	8	37	22	25	8	27	33	26	23	10	30	23
24	20	8	6	8	52	36	16	39	33	33	5	24	35	21	14	14	27	40
25	26	4	3	12	42	35	11	41	20	24	5	28	43	11	23	10	20	56

Following steps 3 to 6, one-point mode per activity, one-point optimistic and pessimistic duration were computed using expression 1.12 to 1.14 and the observed- data equivalent of traditional PERT (OPD) and Risk Incorporated Duration (RID) determined using expressions 1.11 and 1.17 respectively.

Table 3 Results of the Parameter Estimation Functions

Activity Code	P(A)	P(M)	P(B)	Φi	Ωi	Бi	E(Δi)	OPD_i	RID_i	Status	Best fit fistribution
A	0.36	0.40	0.24	19.0	26.5	33.5	2.5	25.480	29	R	Beta
B	0.24	0.44	0.32	3.7	6.1	9.4	1.9	6.560	8	R	Weibull
C	0.00	0.56	0.44	0.0	4.9	7.4	1.4	5.960	6	R	Beta
D	0.00	0.56	0.44	0.0	7.7	12.0	1.9	9.600	10	R	Lognormal
E	0.28	0.48	0.24	33.0	41.6	55.7	4.2	42.560	46	R	Lognormal
F	0.36	0.52	0.12	36.2	40.3	45.3	0.8	39.440	41	R	Normal
G	0.20	0.48	0.32	8.8	12.2	16.5	1.9	12.880	14	R	Beta
H	0.40	0.60	0.00	37.4	43.3	0.0	0.0	40.960	43	NR	Beta
I	0.20	0.60	0.20	23.0	30.7	38.6	2.9	30.720	34	R	Uniform
J	0.32	0.36	0.32	16.0	20.9	28.0	3.9	21.600	25	R	Lognormal
K	0.00	0.52	0.48	0.0	6.1	15.5	6.7	10.600	13	R	Weibull
L	0.36	0.40	0.24	15.8	21.5	27.0	1.6	20.760	23	R	Normal
M	0.08	0.52	0.40	17.0	24.4	32.5	7.4	27.040	32	R	Lognormal
N	0.24	0.40	0.36	13.5	20.7	27.1	3.0	21.280	24	R	Weibull
O	0.40	0.44	0.16	13.7	20.0	30.0	2.1	19.080	22	R	Weibull
P	0.28	0.48	0.24	11.4	17.8	28.0	3.9	18.480	22	R	Lognormal
Q	0.40	0.44	0.16	16.8	20.6	27.3	1.5	20.160	22	R	Lognormal
R	0.20	0.48	0.32	21.2	34.1	49.0	7.0	36.280	41	R	Normal

Table 4 Critical Path and Project Completion Time

S/N	Approach	Critical Path	Project Duration
1	Risk-free schedule	$B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow G \rightarrow I \rightarrow L \rightarrow M \rightarrow O \rightarrow R$	243days
2	ODB	$B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow G \rightarrow I \rightarrow L \rightarrow M \rightarrow O \rightarrow R$	251days
3	RID	$B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow G \rightarrow I \rightarrow L \rightarrow M \rightarrow O \rightarrow R$	276days
4	PERT $(\frac{1}{6}, \frac{4}{6}, \frac{1}{6})$	$B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow G \rightarrow I \rightarrow L \rightarrow M \rightarrow O \rightarrow R$	249

Based on the information in Theorem 2, only one of the project activities may be considered as not risky with $P(B_i) = 0.0$. All other activities with $P(B_i) > 0.0$ are classified as risky activities. Hence, the project is carried out in a risky environment. However, information regarding the causes of variation of these durations was stone-walled as such simulated data was used for further classification and analysis.

It is clear from the resulting critical paths and associated project durations presented on Table 4.0 that the use of the risk-free, observed- data equivalent of traditional PERT (OPD) and risk incorporated (RID) schedule may be capable of giving alternative project plans for judicious choice by management. The adoption of which depends on the degree of risk inherent in project environment. Suppose, theorem 3 holds for a project, then the risk-free schedule may be adopted and for work situations which are not subject to any apparent uncertainty but cannot be guaranteed to be risk-free, the OPD schedule is however proposed for project planning. The risk incorporated schedule is however intended to avoid project over-runs if a firm has been suffering from such results in the past.

Furthermore, from the illustrative case, it also appears feasible to estimate the empirical probabilities of risky future events using historical data. Theorem 1 and its Corollary provide a theoretical basis for an analyst to segregate activity historical data into optimistic (A), most of the times (M) and pessimistic (B) subsets while expression 3.10 allows identification of activities which pose risk to on time project completion. With the availability of a realistic set of historical/field data which includes possible causes for any deviation from plan, both theorems 1 and 2 allows the analyst to further segregate risky activities into those with controllable causes and those with uncontrollable causes. Controllable causes of activity risk are defined as risk causes within the control of project management personnel and thus, marked for mitigation or elimination based on the effectiveness of chosen risk management strategy. For activities with no historical observation in the pessimistic subset or whose causes of risk are controllable, a none risky (Theorem 3) work environment may be assumed and the risk-free duration adopted with a schedule of possible mitigation actions. However, the risk incorporated duration in the critical path analysis is intended to avoid frequent or uncontrollable project over-runs. This selective treatment is intended to reduce avoidable over-runs in cost and time by allowing space for totally unavoidable events on the schedule while intensifying effort to eliminate controllable risky events. This last but very important approach, however, is not part of the intent of this study and not adequately demonstrated in the example because of the limitation of information. Such activity classification may require established causes of observed deviation from plan per activity.

Table 5. Percentage Possible Project Overrun Accommodated Per Model

μ	σ	X	Nos. of Std. (Z)	P(X)	Traditional PERT	OPD	RID
					248	251	276
251	57	247.862	-0.055	0.478	**0.479**		
251	57	251	0	0.500		**0.5**	
251	57	254.137	0.055	0.522			
251	57	257.27	0.110	0.544			
251	57	260.412	0.165	0.566			
251	57	263.55	0.220	0.587			
251	57	266.687	0.275	0.608			
251	57	269.825	0.330	0.629			
251	57	272.962	0.385	0.650			
251	57	276.1	0.440	0.670			**0.669**

Evaluating performance of the selected models in handling situations of possible project duration overrun was evaluated using the Chebyshev's inequality (Rao et al., 2013). Table 5.0 shows the percentage possible overrun that is accommodated by each of the selected models (given that mean duration is that obtained using the expected value of the best fit distribution). From the table, it is shown that PERT underestimate the project duration given by the best fit distribution by -1.25% (under run) and it has a 47.9% chance of on-time project completion. This may be causes of incessant overrun of project completion time experienced by managers depending on the PERT approach for planning. Similarly, the Observed Data-Based duration (equivalent of the traditional PERT), ODB, can handle situations of 0% overrun with a 50% chance of on-time completion. Thus, it is a good alternative for PERT in a pure uncertain work environment which is neither guaranteed to be risk-free nor susceptible to any apparent risk. The risk incorporated duration (RID) schedules on the other hand can handle situations when overrun is up to +8.8% of plan. This seems robust to accommodate significant overrun due to uncontrollable risk consequence of a well-managed project.

6. CONCLUSION

In this subsection, a non-parametric framework based on the paradigm of PERT was illustrated to incorporate risk in project scheduling. Using 25-period historical data of 18-activity project, statistical distribution free model for estimating PERT's optimistic, most of the times, pessimistic duration and the associated probabilities using historical data has been established. It is also shown that a framework for generating a set of project risk-free and environmental risk incorporated schedules is feasible. Other features (Theorems 1 and 2) provide information for selecting risky activities for closer scrutiny in project planning and execution as well as determine their respective chances of contributing to project overrun.However, to effectively apply, valid historical or field data are required. It is hoped that the suggested simple-to-apply framework for risk analysis will spur project managers to consider regular record keeping of activities' durations, their deviation from plan and all possible physical causes a worthwhile policy.

RECOMMENDATION FOR FUTURE STUDIES

The suggested risk-incorporated schedule may provide some information for developing risk mitigation plan provided inventory of all possible causes of deviation from past project plans are identified through the history of the associated business or creative thinking. This aspect which is beyond the scope of this paper is part of our on-going research.

REFERENCES

Azaron, A., & Fatemi Ghomi, S. M. T. (2008). Lower bound for the mean project completion time in dynamic PERT networks. European Journal of Operational Research, 186(1), 120–127.

Corotis, R. B., & Hammel, E. M. (2010). Multi-attribute aspects for risk assessment of natural hazards. International Journal of Risk Assessment and Management, 14(6), 437–458.

Cho, S. (2006), An exploratory project expert system for eliciting correlation coefficient and sequential updating of duration estimation, Expert Systems with Applications 30 (2006) 553–560.Creemers, S., Demeulemeester, E., & Van de Vonder, S. (2014). A new approach for quantitative risk analysis. Annals of Operations Research, 213(1), 27–65.

Davenport, A. J., Gefflot, C., & Beck, J. C. (2014). Slack-Based Techniques for Robust Schedules. In Sixth European Conference on Planning (pp. 43–49).

Hall, E., and & Johnson, J. (2003). Risk Analysis, Integrated Project Management (1st ed.). New Jersey: Pearson Education, Upper saddle, New Jersey.

Herroelen, W., & Leus, R. (2005). Project scheduling under uncertainty: Survey and research potentials. European Journal of Operational Research, 165(2), 289–306.

Khamooshi, H. and, & Cioffi, D. F. (2013). Uncertainty in Task-Duration and Cost Estimates: A Fusion of Probabilistic Forecasts and Deterministic Scheduling. Journal of Construction Engineering and Management, 139(May), 488–497.

Lau, H.-S., & Somarajan, C. (1995). A proposal on improved procedures for estimating task-time distributions in PERT. European Journal of Operational Research, 85(1), 39–52.

Littlefield, J. T. K., & Randolph, P. H. (1987). Reply—An Answer to Sasieni’s Question on PERT Times. Management Science, 33(10), 1357–1359.

Malcolm DG, Rooseboom JH, Clark CE FW. Application of a technique for research and development program evaluation. Oper Res. 1959;7:646–69.

Opaleye, A. A., & Charles-Owaba, O. . (2014). Mode Behaviour in Relation To Bin Size and Data Distribution. International Journal of Mathematics and Statistics Studeis, 2(2), 86–93.

Opaleye, A. A., Charles-owaba, O. E., & Bender, B. (2017). Relevance Of Historical-Data Based Activity Scheduling And Risk Mitigation Model, 6(3), 728–732.

P.H. Meyer and J.K. Visser. (2006). Improving Project Schedule Estimates Using Historical Data And Simulation. South Africa SA Journal of Industrial Engineering, 17(1), 27–37.

Park, C. S. (2007). Contemporary Engineering Economics (4th ed.). Pearson Prentice Hall, Upper Saddle River, NJ 07458.

Pinto, J. K. (2012). Project management: Achieving the Competitive Advantage (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Rao, K.B.,Anoop, M.B and Iyer, N.R (2013), Application of Chebyshev and Markov-Type Inequalities In Structural Engineering, RT&A, 8(28),pp1-20

Scott, D. W. (1979). On optimal and data-based histograms. Biometrika, 66(3), 605–610. http://doi.org/10.1093/biomet/66.3.605

Shankar, N. R., Rao, K. S. N., & Sireesha, V. (2010). Estimating the Mean and Variance of Activity Duration in PERT. International Mathematical Forum, 5(18), 861–868.

Trietsch, D., Mazmanyan, L., Gevorgyan, L., & Baker, K. R. (2012). Modeling activity times by the Parkinson distribution with a lognormal core: Theory and validation. European Journal of Operational Research, 216(2), 386–396.

Valerdi, R. (2005). The Constructive Systems Engineering Cost Model (COSYSMO). University of Southern California.

Wand, M. P. (1997). Data-Based Choice of Histogram Bin Width. The American Statistician, 51(1), 59–64.