

RESEARCH

Open Access



Stochastic forecasting of project streams for construction project portfolio management

Alireza Shojaei* and Ian Flood

Abstract

Background: Construction companies typically work on many projects simultaneously each with its own objectives and resource demands. Consequently, a key managerial function is to allocate financial, equipment, and human resources between these concurrent projects in a way that satisfies the individual project constraints while optimizing the company's overall objectives.

Project portfolio management is concerned with managing multiple projects to accomplish strategic goals. The main concentrations of research so far have been project selection, prioritization and alignment of a portfolio with strategic goals among a pool of awarded projects. The objective of this research is to develop a model capable of generating representative future streams of projects that can be used to assist in strategic planning and portfolio management.

Methods: This paper reports on an on-going research project aimed at developing and validating a stochastic model of streams of uncertain and unknown future projects and to demonstrate the significance and implications of such uncertainties on project portfolio and strategic planning. The scope of the research is Florida Department of Transportation (FDOT) projects. Records of the past 12 years letting information combined with a pool of candidate variables is analyzed to capture characteristics of the time series data and to determine any project characteristic correlations with macroeconomic factors.

Results: The performances of the various model components presented indicate the viability of an integrated project stream forecaster that predicts, within a simulation environment, the frequencies of projects and empirical distributions of project duration and cost.

Conclusion: The gap in the current body of knowledge is a lack of consideration of the effects of uncertainties associated with future projects, both known (but yet to be awarded to a contractor) and unknown (although statistically quantifiable). It is evident that companies should not just focus on current known projects but also uncertain and unknown future projects. Such a capability, looking into the future, is key to the effective medium and long-term strategic planning for the company. Contractors can, for example, use these stochastic streams of data to test different bidding strategies and to see how sensitive the performance of their portfolio is to changes in different market factors.

Keywords: Construction Project Portfolio Management, Stochastic Forecasting, Project Uncertainty, Strategic Planning, Time Series Modeling

* Correspondence: a.shojaei@ufl.edu
School of Construction Management, University of Florida, Gainesville, FL,
USA

Background

Project portfolio management

The success of a construction company is strongly impacted by its ability to strategically plan for and manage a stream of projects, many of which will overlap in time, and all of which are subject to uncertainty about their occurrence, scope and resource needs. This task can be broadly classified as project portfolio management (PPM). Cooper et al. (1997) describe PPM as “dealing with the coordination and control of multiple projects pursuing the same strategic goals and competing for the same resources, whereby managers prioritize among projects to achieve strategic benefit.” PPM is rooted in two complementary but independent tasks: (i) supporting investment decision making in terms of selecting project types and projects with the goal of optimizing return on investment and risk (Markowitz 1952); and (ii) allocating available resources across many different projects in a way that best meets the goals of those projects (such as contract deadlines and profitability) while managing risk (Pennypacker and Dye 2002).

Modern portfolio theory was introduced by Markowitz (1952) within a financial context. In his theory, a portfolio is defined as a set of financial assets and potential investments, which are used to select a set of investments that either maximize return on investment for a given risk or minimize risk for a given return on investment. Several years later, McFarlan (1981) introduced the concept of PPM in an information technology context. He suggested using projects as the elements of a portfolio (instead of investments) to better achieve an organization's objectives as well as reduce the overall risk that the organization encounters during execution of those projects.

The first definitions of project portfolios tended to be simple and fairly close to the financial portfolio definitions. For example, Archer and Ghasemzadeh (1999, 2004) propose a definition of project portfolio as “...a group of projects that are carried out under the sponsorship and/or management of a particular organization.” Dye and Pennypacker (1999) include the notion of fit to organizational strategy in their definition for project portfolio: “...a collection of projects that, in aggregate, make up an organization's investment strategy.” Githens (2002) adds the notion of program and fit organizational strategy in his definition: “...a collection of projects or programs that fit into an organizational strategy. Portfolios include the dimensions of market newness and technical innovativeness.” Project Management Institute (2013a; b) has defined the term portfolio in their standards as “a component collection of programs, projects, or operations managed as a group to achieve strategic objectives.”

PPM operates at the strategic level of decision making in the organization structure. It has different components such as defining, prioritizing, planning, managing and controlling the subparts of the project portfolio which are projects and programs, to better distribute available resources and address associated risks (Young and Conboy 2013). In other words, PPM is a continuous process which tries to align the management of all projects by continually examining and updating the selection and management of projects to increase the company's performance (Young and Conboy 2013).

However, while there is some agreement in the recent definitions of project portfolio, there is still much variation in the definition of PPM. Authors focus on different aspects of their definitions, and none of them are comprehensive.

For example, Project Management Institute (2013b) lists the PPM subprocesses and repeats its definition of portfolio in its definition of PPM as “...the coordinated management of one or more portfolios to achieve organizational strategies and objectives. It includes inter-related organizational processes by which an organization evaluates, selects, prioritizes, and allocates its limited internal resources to best accomplish organizational strategies consistent with its vision, mission, and values.” On the other hand, Dye and Pennypacker (1999) prefer to focus on the term ‘management’ and define project portfolio management as using management skills to satisfy an organization's investment strategy.

Some recent definitions emphasize the strategic alignment, for instance, Rajegopal et al. (2007) look at portfolio management as a tool to implement an organization's strategy. Levine (2005) similarly emphasizes the role of PPM in contributing to the overall success of the enterprise. Cooper et al. (2001) focus on the decision and revision processes in their definition of project portfolio management. This definition supports the view adopted in this paper that project portfolios are dynamic entities, which must continuously be monitored, analyzed and controlled to ensure that they are kept in line with the organizational goals. Finally, Turner and Müller (2003) take an alternative view by building on the notion of the project portfolio as an organization. They emphasize collective management of the projects to achieve better resource distribution among projects and reduce uncertainty. However, their definition of a project portfolio as an organization has not been widely accepted by the business and academic communities.

In spite of all the models that have been developed for assisting in the establishment of a project portfolio, allocating resources among the projects and examining the portfolio success, generally companies have not found that PPM modeling meets their expectations and, moreover, it does not appropriately address the dynamic

nature of project portfolios (Elonen and Artto 2003; Engwall and Jerbrant 2003).

Uncertainties in project portfolio management

The concept of uncertainty is very influential within the field of project portfolio management. Duncan (1972) and Daft (2009) demonstrated that changes in the business environment combined with projects with high complexity always result in an increase in uncertainty in parameters such as the number of projects, how rapidly and according to plan projects progress and changes in the economic conditions. This has led to an extensive literature on uncertainty and the ways to handle it in management.

The idea of uncertainty impact on organizations is well established in many disciplines from psychology to economics (Petit and Hobbs 2010). Environmental uncertainties and their relation to organizations are analogous to the state of a person with a shortage of critical information about the environment. Scott (2002) provides an example of the definition of environmental uncertainty as variability or the extent of predictability of the environment where work is executed. They also introduced some measures for uncertainty such as variability of inputs, the number of deviations in work process, and the number of changes in major products. In the project management context, uncertainty in a project is defined as the accuracy of predicting the variation of resource consumption, output, and work process (Dahlgren and Söderlund 2002). Uncertainty in a project can be seen as a variation from expected performance of the system under investigation.

The PMI standard for portfolio management in spite of introducing the risk management concept at a portfolio level does not provide much information on how managers should handle uncertainty and risk within their portfolio. They only provide guidelines on categorizing different possible stages and processes plus naming some of the possible techniques available to handle uncertainties. The PMI only suggests monitoring risks and the performance of the project portfolio under the monitoring and control process group. The proposed framework by the PMI also includes monitoring changes in business strategy. This is an important task because when it occurs, it might result in a complete realignment of the portfolio. The mechanisms involved in this realignment are not specified other than starting the whole PPM process from the beginning. Also, ad-hoc disturbances to the ongoing and approved project portfolios are almost completely neglected.

This oversight is not due to the fact that the topic lacks interest or that authors assume a stable and predictable environment. Rather, it can probably be explained by the fact that the topic of PPM is fairly

young and that the researchers and academics preferred to focus on more pressing issues in this area. For many companies, the environment is unstable and the high level of uncertainty and unknowns resulting from the dynamic environment lead to a number of challenges. These challenges will be reviewed in the following section. New tools and techniques are required to help manage portfolios that exist within a continuously changing non-ergodic environment.

Martinsuo's (2013) review of empirical research on PPM noted that uncertainty and constant changes in company portfolios have a considerable impact on project portfolio performance. Furthermore, he proposes that further research is required into PPM as a continuous process of project selection, resource allocation optimization, sensing and adapting to changes within an uncertain dynamic environment.

Upcoming projects significantly affect the performance of a project portfolio (Araújo et al. 2010). The typical approach when a new project is added to the portfolio is to update the project portfolio's plans and to try to re-optimize everything. This is neither practical (requiring frequent updates to the plans) nor efficient. Additionally, it is very difficult to conduct due to the complex mathematical interface. In fact, a practical portfolio management scheme should enable a user to see different scenarios based on possible upcoming projects in a pipeline and incorporate possible impacts of significant uncertainties on a portfolio to facilitate a better understanding of the future options and likely best strategies. One solution to this issue would be to use a stochastic sampling of streams of uncertain upcoming projects operating within alternative business environments and model the resultant impact on portfolio performance. The idea is to plan proactively by taking into account statistical assessments of potential future opportunities and needs, as opposed to planning reactively (and therefore iteratively) to changing current circumstances.

Project portfolio management models

It is evident in the project portfolio literature that there is no single project portfolio management system that works for all companies. In fact, each company should customize their framework to best suit their situation (Florice and Miller 2003; Killen et al. 2007). For example, Dahlgren and Söderlund (2002) reviewed the project portfolio control mechanisms in four Swedish enterprises and found that different types of firms have different control mechanisms depending on the level of uncertainty and the extent of dependencies between projects. Danilovic and Sandkull (2002, 2005) studied the relationship between uncertainty and dependencies in multiple project situations. They claim that the sources of uncertainty in a new product development are the

organizational settings, the product architecture, and the project management. Providing a practical and comprehensive methodology to facilitate management and coordination of multiple projects in a company's portfolio is a challenging task. There are no appropriate analytical solutions available for dynamic scheduling and resource allocation of project portfolios in real time (Araúzo et al. 2010). Existing proposed mathematical models (such as those of Archer and Ghasemzadeh (1999), Browning and Yassine (2010), Carazo et al. (2010), and Engwall (2003)) cannot handle the complexity of real world challenges due to a limited consideration of significant uncertainties within their models and a lack of provision for dynamic and real-time analysis.

PPM frameworks are also critiqued for not considering all resource restrictions (such as time and interdependence) simultaneously, and for a lack of consideration of an organization's historical performance data which is necessary if a plan is to be based on the organization's capabilities (Henriksen and Traynor 1999; Martinsuo 2013). The main focus of PPM research was initially to improve organizational performance by introducing good practices to choose and prioritize projects and ensure that the right mix of projects was adopted. A recurring theme is the alignment of the projects with the organization's strategy. There is also extensive literature on project selection with a mathematical approach.

Table 1 shows a summary of the literature on approaches toward project portfolio management and compares the tackled problem, proposed solution, limitation and the gap in each model to help demonstrate the contribution of this research. It is clear that none of the presented models include unknown future project streams in their portfolio management methods and their planning horizon is limited to the known projects. However, it is repeatedly argued that upcoming projects significantly impact a portfolio's performance. The proposed model in this research is not a standalone portfolio management framework but should be considered as a supplementary component to current PPM frameworks. It can be used as an add-on to current PPM models to extend their horizon of planning and assist strategic planning by forecasting unknown future projects. In this research, it is not proposed that developed models are incorrect. Instead, it is argued they are incomplete, and the strategic horizon of portfolio planning can be extended by using the proposed method in this paper.

Research goals and objectives

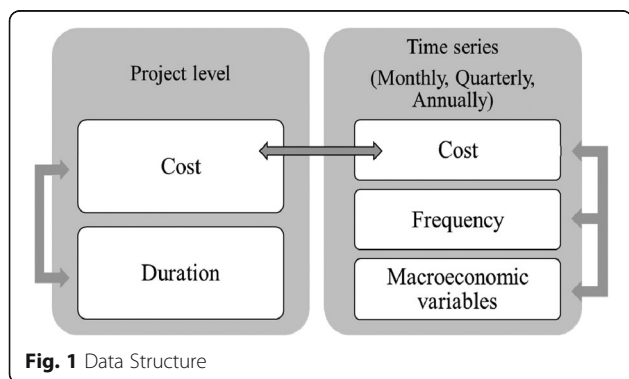
This paper reports on an on-going research project aimed at extending PPM to include streams of projects that advance far enough into the future to facilitate long and medium term strategic planning. These streams

Table 1 Summary of the literature on approaches toward project portfolio management

Author(s)	Research Problem	Solution	Limitations and Gap
Henriksen and Traynor (1999)	Project evaluation and selection in a portfolio	Developed a new algorithm with criteria of relevance, risk, reasonableness, and return	Limited application to research and development project evaluation and only focuses on project selection within a pool of known projects
Liu and Wang (2011)	Project selection and scheduling problems with time-dependent resource constraints	Developed an optimization model using constraint programming	Considers only known projects and financial objectives and lacks monitoring of resource utilization
Ghasemzadeh et al. (1999)	Selecting and scheduling an optimal project portfolio, based on the organization's objectives and constraints such as resource limitations and interdependence among projects	Developed a zero-one integer linear programming model	Lacks the ability of dynamic and real-time analysis. Only considers known projects. Lack of consideration of uncertainties in the model.
Archer and Ghasemzadeh (1999)	Selecting projects for a portfolio	A qualitative multistage framework is developed for selecting projects	Qualitative, only considers known projects and focuses on selecting projects.
Browning and Yassine (2010)	Performance of priority rules in Static resource constrained multi-project scheduling	Sensitivity analysis of priority rule method in different context by simulation	Deterministic, and only considers known projects.
Carazo et al. (2010)	Selection and scheduling of project portfolios from a set of candidate projects	A multi-objective binary programming model using a metaheuristic procedure based on Scatter Search	Only works for a pool of known projects.
Araúzo et al. (2010)	Dynamic scheduling of resources within a portfolio	Distributing resources by a multi-agent system through an auction mechanism	Model is limited to resource allocation optimization. It considers only known projects within the portfolio.
Pajares and López (2014)	The inclusion of resource allocation in PPM models	Suggests a research agenda for developing models for PPM including operative issues	A conceptual paper, showing that a proper PPM model should include operations management as well.

Table 2 Candidate variables and sources

Candidate variables	Source
GROSS DOMESTIC PRODUCTS (GDP)	U.S. Bureau of Economic Analysis
GDP IMPLICIT PRICE DEFLATOR	U.S. Bureau of Economic Analysis
INFLATION RATE	World Bank
CONSUMER PRICE INDEX	U.S. Bureau of Labor Statistics
NATIONAL HIGHWAY COST INDEX (NHCCI)	U.S. Department of Transportation
FDOT'S ANNUAL BUDGET	Florida department of transportation
FDOT'S PRODUCT BUDGET	Florida department of transportation
FEDERAL FUNDS RATE	Federal Reserve Systems
UNEMPLOYMENT RATE	U.S. Bureau of Labor Statistics
FLORIDA UNEMPLOYMENT RATE	U.S. Bureau of Labor Statistics
NUMBER OF EMPLOYEES IN CONSTRUCTION	U.S. Bureau of Labor Statistics
NUMBER OF EMPLOYEES IN CONSTRUCTION IN FL	U.S. Bureau of Labor Statistics
AVERAGE WEEKLY HOURS	U.S. Bureau of Labor Statistics
PRIME LOAN RATE	Federal Reserve System
BUILDING PERMITS	U.S. Bureau of Census
MONEY SUPPLY	Federal Reserve System
AVERAGE HOURLY EARNINGS	U.S. Bureau of Labor Statistics
EMPLOYMENT COST INDEX (ECI) CIVILIAN	U.S. Bureau of Labor Statistics
DOW JONES INDUSTRIAL AVERAGE	Yahoo Finance
CRUDE OIL PRICE	U.S. Energy Information Administration
BRENT OIL PRICE	U.S. Energy Information Administration
PRODUCER PRICE INDEX	U.S. Bureau of Labor Statistics
HOUSINGS STARTS	U.S. Bureau of Census
CONSTRUCTION SPENDING	U.S. Census Bureau



would allow PPM to be used for strategic project selection and resource planning taking into account a potential future stream of known projects and unknown projects. Known future projects are those that have been announced but have not yet been awarded to a contractor. Unknown projects are those that have not been announced (they may still be in the design process or may not have even been conceived) but can be modeled as a statistical expectancy based on historic data. The streams will be developed using stochastic techniques that are statistically representative of what may happen. A statistically significant sample of these streams can then be filtered through a company's bidding success model and the output from this then used to optimize strategic planning, taking into account uncertainty and variance in the future market. The sensitivity of the optimality of a plan to changes in key market parameters can also be tested, and appropriate contingencies for such events thereby established.

Methods

The focus of research so far has been the selection and prioritization of projects among a pool of known projects, ignoring the opportunities and needs of unknown future projects. It is by definition a short term planning strategy and has no guarantee of satisfying a company's longer term goals. The current horizon of strategic planning which covers selecting the projects for bidding and planning for their contractual needs and necessary resources for execution are limited to the extent of advertised projects in the market.

In the proposed approach, it is argued that future unknown projects can be represented statistically and by

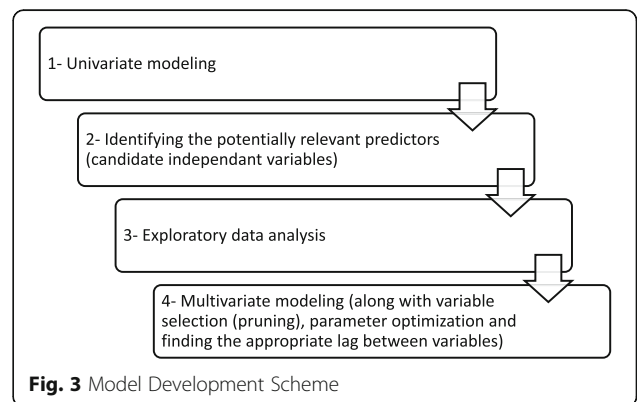


Table 3 Correlation for duration versus cost

		Duration	Cost
Duration	Pearson correlation	1	0.662
	Significance		0.000
Cost	Pearson correlation	0.662	1
	Significance	0.000	

bringing those into the strategic planning process companies can devise more appropriate medium and long term strategies. In the construction context, where projects are advertised by clients and won by companies by bidding for them or other procurement methods, it is fundamentally wrong to use a company’s past and current portfolio to predict unknown future projects. The correct approach is to use the historical data from the market to forecast upcoming projects (all the available projects in the future).

A preliminary study is underway developing, validating and testing a project stream generator for design-bid-build highway construction projects let by the Florida Department of Transportation (FDOT). The outputs from the generator are those parameters most critical to a company, namely the occurrence and letting date of a project, its expected duration, and its expected cost.

The main data for this study were obtained from FDOT’s historical project lettings database covering

12 years (from 2003 to 2015). The database contains 2,816 design-bid-build project letting reports. Based on the letting date, the frequency of projects per month, per quarter and annually are calculated as secondary variables. A pool of candidate variables including the macro-economics metrics and construction indices were compiled from the related sources and literature (Shahandashti and Ashuri 2016), then analyzed to capture characteristics of the time series data and to determine any project characteristic correlations with macroeconomic factors to facilitate modeling representative future streams of projects. Variables that did not significantly impact the project stream would then be cut from the list. The candidate variables were compiled into monthly, quarterly and annual groupings (if available) to be analyzed with the primary data in order to assess the relationships across different timeframes. Table 2 provides the list of the variables considered and their source.

Figure 1 presents the data structure and the investigated connection between variables. Note, when using the cumulative dataset for a period the duration variable is lost since it is meaningless to sum the duration of the projects in a given month or quarter. However, the cumulative datasets provide the project frequency for different timeframes. Cost is the common variable between the two dataset perspectives. The cost variable can be used as an anchor to link the duration and the frequency of the projects.

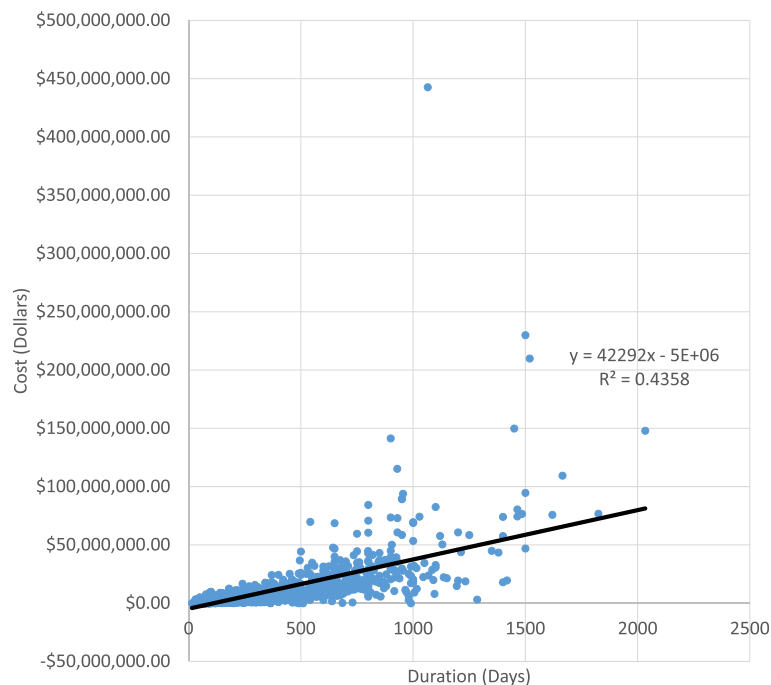
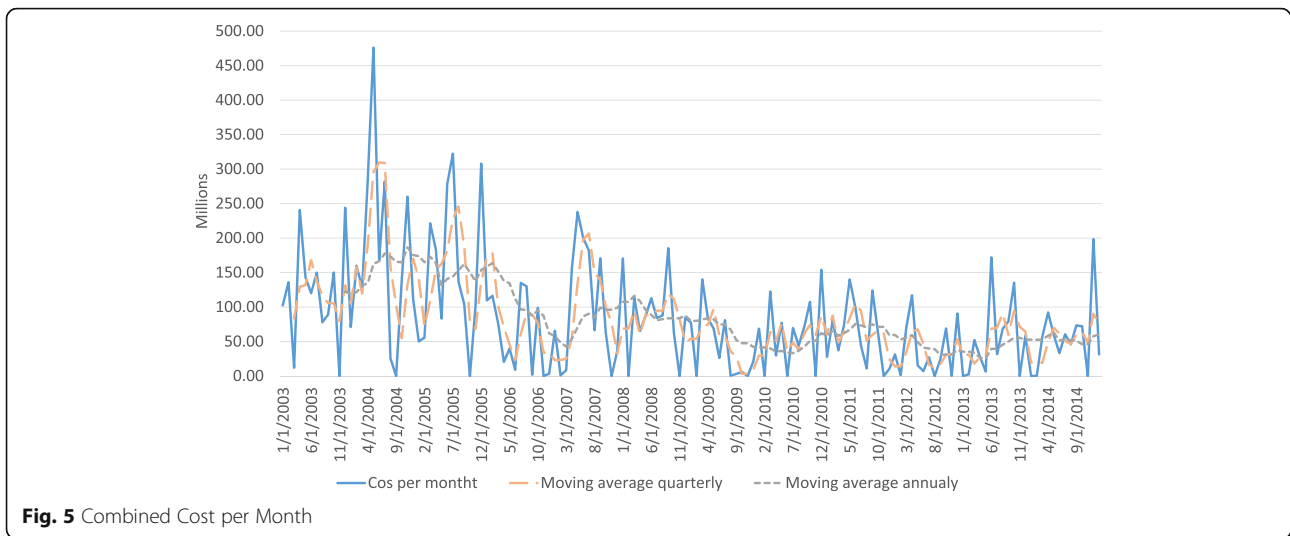


Fig. 4 Scatter Plot Showing Relationship between Duration and Cost

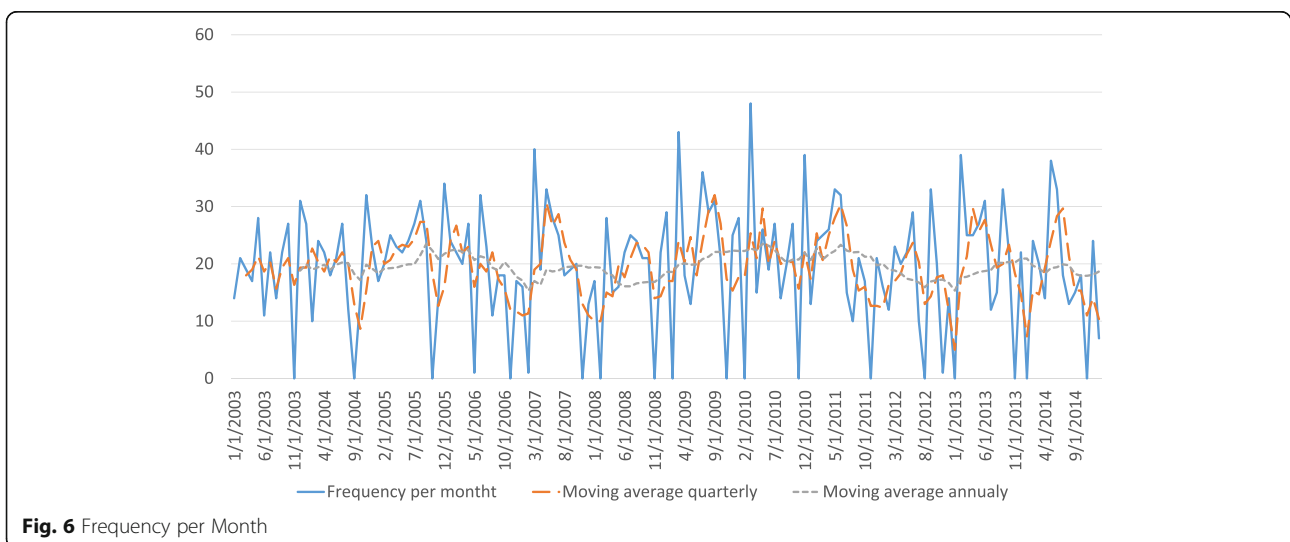


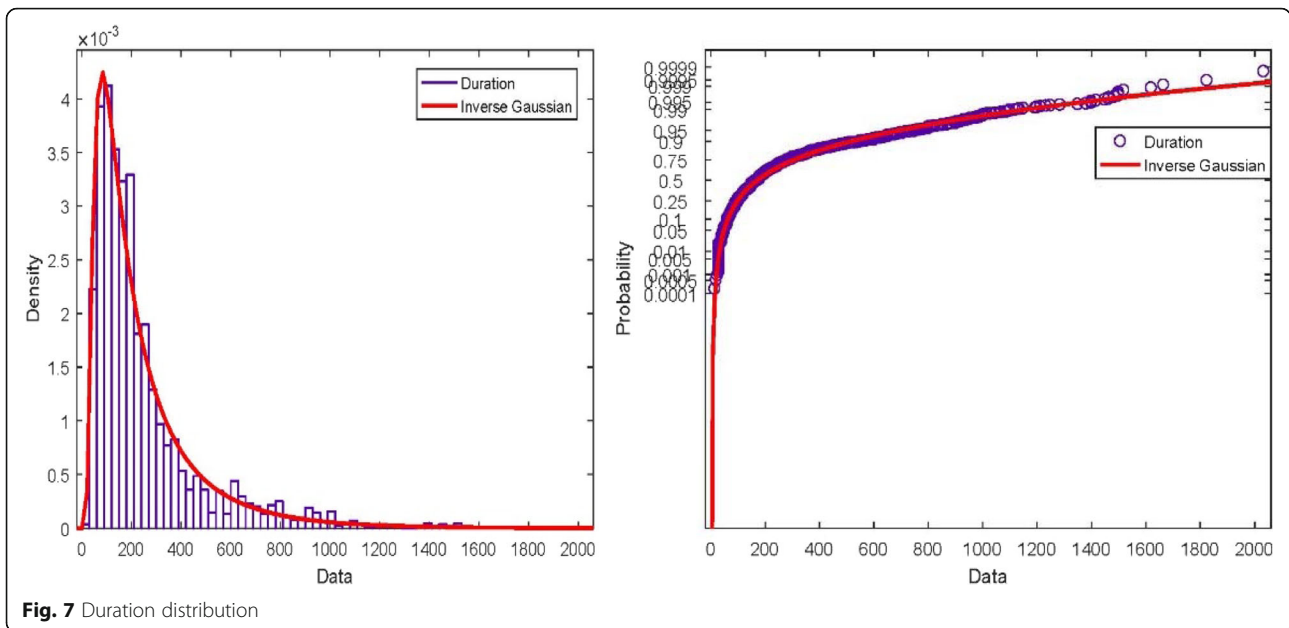
The sequence of generating information in the proposed model is shown in Fig. 2. The first step is to forecast the number of projects (frequency) for the chosen time span. Next, sampling from project cost distribution should take place. At each point in time, the number of samples from the distribution is based on the number of projects forecasted in the previous step. Finally, the same process applies to the duration distribution while the possible correlation between cost and duration should be considered in the sampling process.

The procedure used to develop the model to forecast the frequency of the projects is shown in Fig. 3. The purpose of this procedure is to look for characteristics of data, to capture them in the model's projections, and then to check to see if the model reproduces those characteristics. The univariate model was adopted as a benchmark against which the more complex multivariate

models were compared to see if they improve forecast accuracy and provide more insights into better ways of modeling this problem.

The first step is modeling the main variables through univariate modeling methods such as autoregressive modeling (AR), moving averages (MA), autoregressive moving average (ARMA), and exponential smoothing. More sophisticated approaches such as artificial neural networks can also be implemented considering the availability of the necessary data size. After establishing a benchmark, potentially relevant predictors were identified to populate a pool of candidate independent variables based on a literature review and cognitive theories. This brings in the environmental uncertainties into the forecast with the aim of improving the accuracy of the simulation. These variables will not necessarily have a causal relationship with the main variables; the only concern here is to be helpful in forecasting the dependent





variable. This paper reports the results of this research up to this stage.

The second step is exploratory data analysis. It starts with a graphical comparison of the independent and dependent variables such as scatterplots of pairs of variables. Relevant test include Pearson correlation, unit root (stationary or non-stationary test), Granger causality (helpful for short term forecasting), and cointegration (suitable for long term forecasting) tests.

The last step is to choose a set of multivariate modeling approaches based on the results of the exploratory data analysis and investigate whether including explanatory variables and models that are more complex can

improve the accuracy of the forecast. The range of the models should test for linear and non-linear relationships based on the results of the previous step along with variable selection (pruning), parameter optimization and finding the appropriate lag between variables. The authors suggest using two different variable selection methods and comparing the results to provide further explanatory insight into each variable's importance. The first method is univariate feature selection using the Granger causality test as an identifier of the appropriate explanatory variables. The second and more robust method is using recursive feature elimination with a greedy optimization algorithm. This iterative method

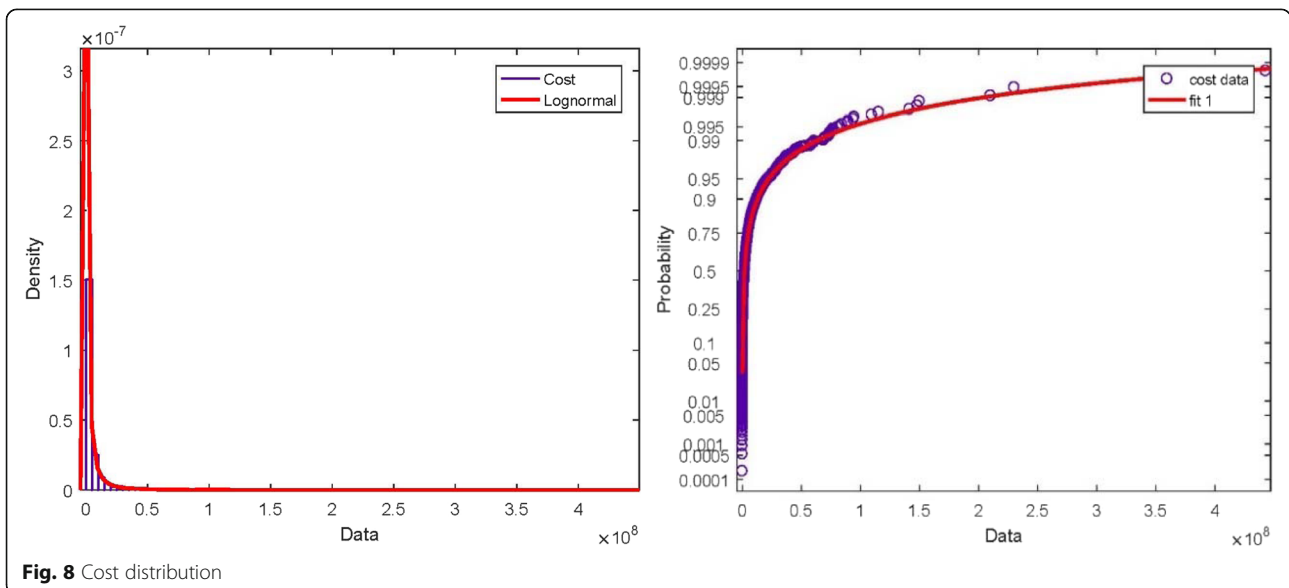


Table 4 Summary of ADF test on project frequency series

	Lag	t-statistic	P-value
ADF with intercept and trend	11	-3.15	0.100
ADF with intercept	11	-3.12	0.027
ADF without intercept and trend	12	-0.43	0.52

builds models and separates best and worst variables at each step. This process continues until all the variables have been considered. The result is the ranking of the variables based on their order of elimination. It is crucial to embed a cross validation method within the recursive variable selection method to avoid overfitting.

Results and Discussion

Table 3 presents the result of a Pearson correlation test between the duration and cost at the project level. It is evident that there is a relatively strong linear relationship between the two variables, which should be taken into consideration while sampling from the cost and duration distribution.

Figure 4 shows the scatter plot of duration and cost of the projects, illustrating visually the strong association between the two variables as summarized by the best-fit line. The R^2 value for the complete set of data is 0.4358, which implies a weak linear relationship between cost and duration, however, it can also be seen from the figure that there are several data points that are clearly outliers. These introduce a high amount of error that drastically reduces the R^2 value. For example, removing the top 16 outliers from the 2,816 data points increases the R^2 value to 0.7.

Figure 5 shows the accumulative cost per month plot along with quarterly and annually moving average for the 12-years period 2003–2015, demonstrating a decreasing trend in dollar value of the projects in the more recent years.

Figure 6 plots the project frequency for each month during the 12 years period (2003–2015). It also

represents the quarterly and annual moving average. The trend appears relatively constant other than a slight increase in the latter years. Considering the fall in the combined monthly project budget (Fig. 5) supplemented with the slight increase in the number of projects in a month it can be concluded that the budget for the individual projects has on average decreased.

A quantitative description of the distribution of the main variables and the correlation between parameters such as cost and duration is essential for the stochastic generation of the project stream and/or validating the results of the generator. Figure 7 shows the histogram and the corresponding fitted distribution for the duration of the FDOT projects as well as the empirical and theoretical cumulative distribution function. An Inverse Gaussian distribution with $\mu = 244.67$ and $\lambda = 273.93$ was found to provide the best fit using Akaike information criterion (AIC).

Figure 8 shows the histogram and the corresponding fitted distribution for the cost of the FDOT projects as well as the empirical and theoretical cumulative distribution function. A lognormal distribution with (mean log) $\mu = 14.413319$ and (standard deviation log) $\sigma = 1.524961$ was found to provide the best fit using AIC.

Frequency univariate modeling

In order to better understand the characteristics of project frequency series and capture them, a set of preliminary analysis was undertaken. An essential analysis is testing for stationarity and finding the order of differencing that makes the series stationary. A time series is stationary if its mean and variance evolve around constant values. In order to be able to implement many of the modeling tools, it is necessary to make sure that a series is stationary. However, if it is turn out that a series is not stationary, it can be transformed to a stationary form by different tools such as differencing. Differencing is the difference between consecutive data points and the order of differencing is the number of times a series must be differenced to make it stationary. An

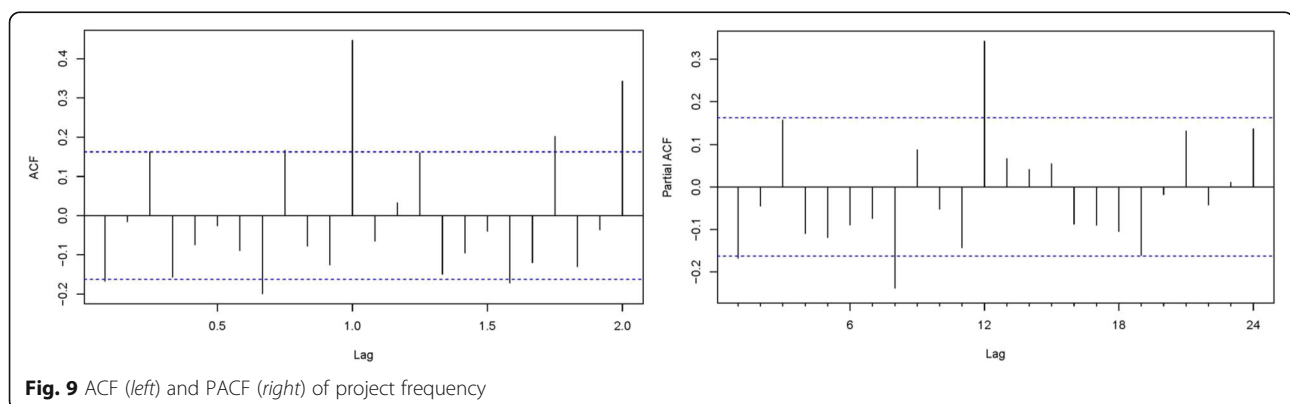


Table 5 Summary of the univariate models' performance

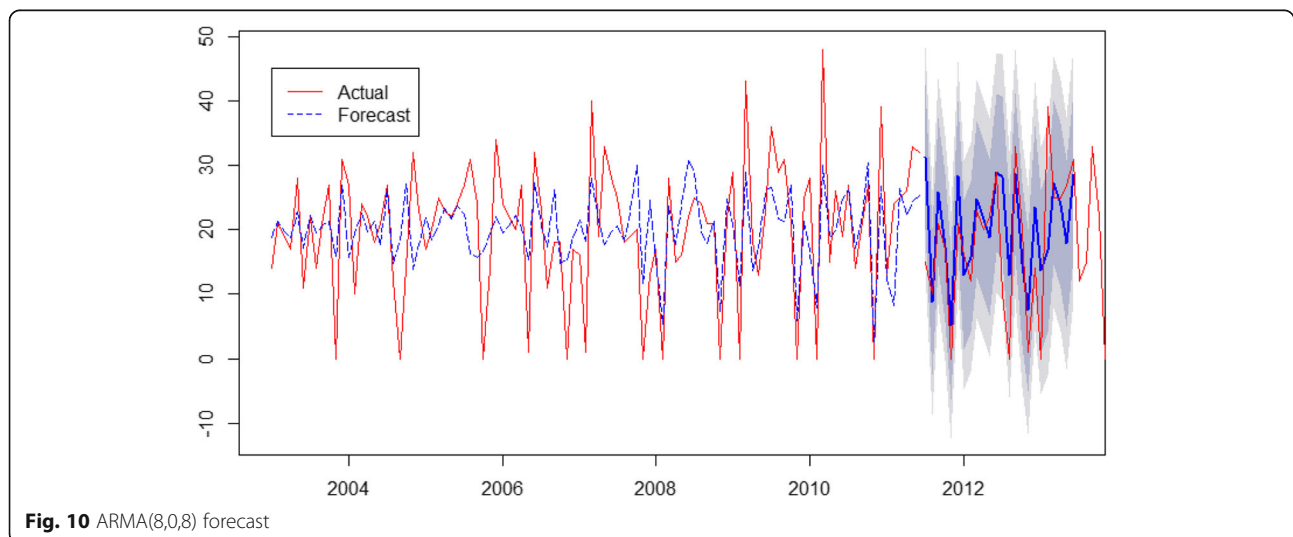
Model	Set	RMSE	MAE
Holt winter (additive seasonal)	Training set	8.10	6.58
	Test set	11.73	8.99
ARMA(8,8)	Training set	7.93	6.28
	Test set	8.82	6.84
ARIMA(0,0,1) (0,0,2) [12]	Training set	9.28	7.46
	Test set	10.22	8.15

Augmented Dickey-Fuller test (ADF) was conducted to test the stationarity of the data. There are three variations of the ADF test, all with the null hypothesis that a unit root is present in a time series sample (series is not stationary). If under any of the three variations the null hypothesis is rejected it can be inferred that the time series is stationary. Choosing the appropriate lag in ADF is critical. In this research, the appropriate lag is chosen based on the Akaike information criterion (AIC). A summary of the ADF results are presented in Table 4. It is evident that the null hypothesis can be rejected at 95% confidence level and it is concluded, therefore, that the frequency series is stationary.

Another essential analysis is provided by the Autocorrelation Function (ACF). Autocorrelation is the correlation between a time series and a delayed version of itself. ACF is helpful to find repeating patterns in the data. Correlogram is a figure that demonstrates the correlation between two series. Figure 9 (left) shows the ACF correlogram of the project frequency. The X-axis shows the lag (delay) in years, the Y-axis shows the correlation value and the dotted line shows the 5% significance boundary. We can see that lag 8 and lag

12 crossed the significance bounds. The result of the ACF correlogram shows that using a Moving Average (MA) model with lag 8 (the first lag with significant correlation) is appropriate. Partial Autocorrelation Function (PACF) is the autocorrelation function between a time series and its lagged version after removing any linear dependence on values with shorter lags. Figure 9 (right) shows the PACF correlogram of the project frequency. The X-axis shows the lag (delay) in months, the Y-axis shows the correlation value and the blue line shows the 5% significance boundary. We can see that lag 8 and lag 12 crossed the significance bounds. The result of the PACF correlogram shows that using an Autoregressive (AR) model with lag 8 (the first lag with significant correlation) is appropriate. Considering the results of ACF and PACF using an Autoregressive Moving Average (ARMA) model, which combines an AR and a MA model is appropriate.

ARMA models are one of the most frequently used in forecasting univariate time series. The model is typically represented as ARMA(p,q), where p is the autoregressive order and q is the moving average order. The order of autoregressive and moving average is selected via autocorrelation correlogram and partial autocorrelation correlogram. Based on the results of autocorrelation and partial auto correlation an ARMA ($p=8, q=8$) is the right choice to model the project frequency series. Also a set of seasonal ARMA (8,8) models with possibility of differencing was fitted to the series and the best model selected by the AIC. Holt-Winters is a triple exponential smoothing method, which has two variations, the additive and multiplicative seasonal method. The multicaptive method is not appropriate for series with negative or zero values and as the project

**Fig. 10** ARMA(8,8) forecast

frequency series has zero values in some months, only the additive method is implemented. The data set was divided into two sections. The first section was used for training and contains 70% of the dataset, while the second section was used for testing and contains the remaining 30% of the dataset. Table 5 summarize the performance of the discussed univariate time series modeling for forecasting project frequency. It is evident that the ARMA (8,8) model outperforms the other methods.

Figure 10 shows the plot of the actual data and the fitted ARMA model. The predicted values are shown in blue, and the actual data is plotted in red. The gray area represents the prediction intervals for the test data set. The dark gray shows the 80% prediction interval, and light gray shows 95% prediction interval. Visual examination of Fig. 10 shows that the ARMA model better forecasts values after 2008. It is evident that the variance of the actual data increases as the time passes and the model better reproduces the data’s variance in later values. However, it can be observed that the predicted values’ variance is less than the actual data’s variance over the whole time span.

The performances of the various model components presented in this section indicate the viability of an integrated project stream forecaster that predicts, within a simulation environment, the frequencies of projects and empirical distributions of project duration and cost. Specifically, the generator will produce stochastic streams of unknown future FDOT projects.

According to the literature (Shahandashti and Ashuri 2016; Thomas Ng et al. 2000; Wong and Ng 2010), incorporating the interrelationships between identified macroeconomic factors and the main variables should further improve the ability of the generator to capture the essential characteristics of a project stream. Work is underway implementing this system using Monte-Carlo sampling for project stream generation. Statistical validation of the performance of the generator will be made for

all key variables and across multiple timeframes. Once validated, bidding strategy and bidding success filters will be added to the generator to produce streams of projects awarded to a contractor. Models will then be built for assessing the resource needs of the contractor over time (based on the simulated project streams) and optimizing company level resourcing strategies. The sensitivity of the optimality of these strategies will then be assessed in relation to uncertainty in the project streams, and changes in key economic factors.

As mentioned earlier, the proposed method is not a standalone portfolio management framework but should be considered a supplementary component to the current PPM frameworks that can extend the current PPM models’ horizon of planning. Figure 11 demonstrates how the proposed method can be implemented and how the research outcomes can be utilized. Historical data is used as the input to the model discussed in this research. Then, the number of projects with assigned cost and duration is forecast as the output from the model. Finally, the output of the proposed model along with the known projects (advertised projects) will be used as the input for the PPM models currently implemented by a company. For instance, the output of the model can be used as the input for the model developed by Liu and Wang (2011) or Archer and Ghasemzadeh (1999) to extend the horizon of strategic planning of those models.

Conclusion

This paper has outlined an extension to PPM that allows future streams of both known and unknown (but statistically quantifiable) projects to be taken into account in the strategic planning of a project portfolio. On-going work is concerned with developing, validating and testing a project stream generator based on FDOT historic data. This generator will work stochastically, producing samples of streams of future FDOT projects (in terms of time of occurrence, expected duration and expected cost) based on historical data and economic indicators.

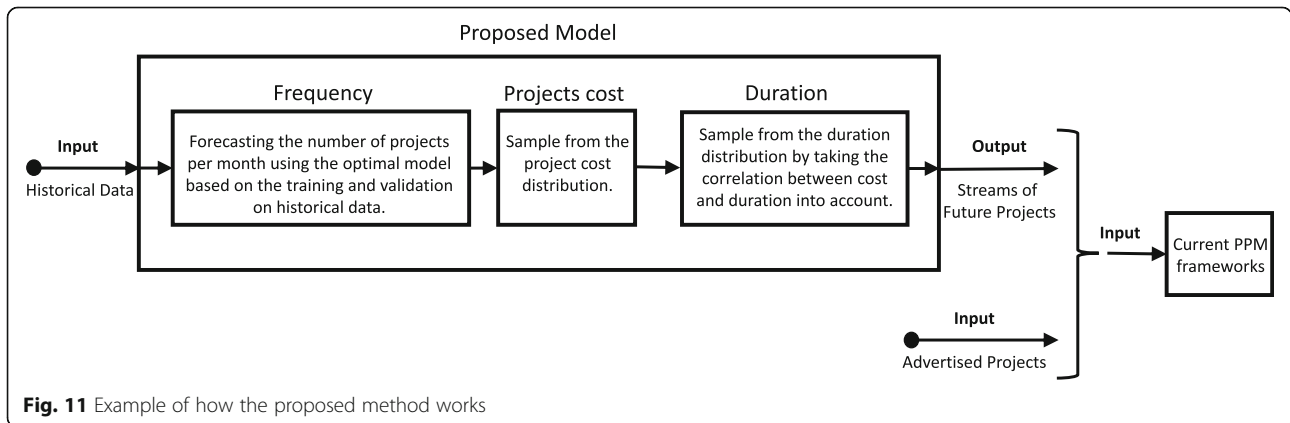


Fig. 11 Example of how the proposed method works

The next step of the project will be to combine the generator of streams of future projects with a sample portfolio to show the model capabilities and effectiveness. The complete framework will give users the opportunity to try different bidding and project selection strategies to see how these affect their future resource demands so they can plan ahead and find an optimal strategy and optimal resource distribution for the future. An ability to accurately predict future project streams, and to take into account uncertainties in these streams, may help reduce the extent of the continuous adjustments required to a company's portfolio plan, as well as help optimize the selection and management of the portfolio. The types of resources that may be considered are anything required by a business to conduct its operations, including human resources, financing, equipment, and production facilities.

Abbreviations

ACF: Autocorrelation Function; ADF: Augmented Dickey-Fuller; AIC: Akaike information criterion; AR: Autoregressive; ARMA: Autoregressive Moving Average; FDOT: Florida Department of Transportation; MA: Moving Average; PACF: Partial Autocorrelation Function; PMI: Project Management Institute; PPM: Project Portfolio Management

Acknowledgements

Not applicable.

Funding

This work was funded by the University of Florida graduate student fellowships program. The funding body did not play any role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

AS: He is the Corresponding Author. This paper is based on his Ph. D.'s dissertation. He has done the data collection, data analysis, and drafting the article under the supervision of Dr. Flood. Dr. IF: The conception of the work is by Dr. Flood. He has also done the critical revision of the article and final approval of the version to be published. Both authors read and approved the final manuscript.

Competing interests

We do not have any competing interest. Springers' disclosure of potential conflict of interest forms are filled and signed by each author and presented in the following pages.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 December 2016 Accepted: 2 June 2017

Published online: 19 June 2017

References

Araújo, J. A., Pajares, J., & Lopez-Paredes, A. (2010). Simulating the dynamic scheduling of project portfolios. *Simulation Modelling Practice and Theory*, 18(10), 1428–41.

- Archer, N., & Ghasemzadeh, F. (1999). An integrated framework for project portfolio selection. *International Journal of Project Management*, 17(4), 207–16.
- Archer, N., & Ghasemzadeh, F. (2004). Project Portfolio Selection and Management. *The Wiley guide to managing projects*, 27, 237–55.
- Browning, T. R., & Yassine, A. A. (2010). Resource-constrained multi-project scheduling: Priority rule performance revisited. *International Journal of Production Economics, Elsevier*, 126(2), 212–28.
- Carazo, A. F., Gómez, T., Molina, J., Hernández-Díaz, A. G., Guerrero, F. M., & Caballero, R. (2010). Solving a comprehensive model for multiobjective project portfolio selection. *Computers and Operations Research*, 37(4), 630–9.
- Cooper, R. G., Edgett, S. J., & Kleinschmidt, E. J. (1997). Portfolio management in new product development: Lessons from the leaders-I. *Research-Technology Management*, 40(5), 16–28.
- Cooper, R., Edgett, S., & Kleinschmidt, E. (2001). Portfolio management for new product development: results of an industry practices study. *R and D Management*, 31(4), 361–80.
- Daft, R. L. (2009). *Organizational Theory and Design*. South-Western Cengage Learning (pp. 138–57).
- Dahlgren, J., & Söderlund, J. (2002). *Management control in multi-project organizations: a study of R&D companies*. at IRNOP V.
- Danilovic, M., & Sandkull, B. (2002). *Managing Complexity and Uncertainty in a Multiproject Environment*. IRNOP V, International Research Network on Organizing By Projects, Proceedings, IRNOP V, International Research Network on Organizing By Projects, Proceedings.
- Danilovic, M., & Sandkull, B. (2005). The use of dependence structure matrix and domain mapping matrix in managing uncertainty in multiple project situations. *International Journal of Project Management*, 23(3), 193–203.
- Duncan, R. B. (1972). *Characteristics of Organizational Environments and Perceived Environmental Uncertainty*. *Administrative Science Quarterly* (pp. 313–27).
- Dye, L. D., & Pennypacker, J. S. (1999). *Project portfolio management: selecting and prioritizing projects for competitive advantage*. Center for Business Practices.
- Elonen, S., & Arto, K. A. (2003). Problems in managing internal development projects in multi-project environments. *International Journal of Project Management*, 21(6), 395–402.
- Engwall, M. (2003). No project is an island: Linking projects to history and context. *Research Policy*, 32(5), 789–808.
- Engwall, M., & Jerbrant, A. (2003). "The resource allocation syndrome: The prime challenge of multi-project management?". *International Journal of Project Management*, 21(6), 403–9.
- Floricel, S., & Miller, R. (2003). An exploratory comparison of the management of innovation in the New and Old economies. *R&D Management*, 33(5), 501–25.
- Ghasemzadeh, F., Archer, N., & Iyogun, P. (1999). A zero-one model for project portfolio selection and scheduling. *Journal of the Operational Research Society*, 50(7), 745–55.
- Githens, G. D. (2002). *Programs, Portfolios, and Pipelines: How to Anticipate Executives' Strategic Questions*. *Managing Multiple Projects: Planning, Scheduling, and Allocating Resources for Competitive Advantage*, CRC Press (pp. 83–90).
- Henriksen, A. D., & Traynor, A. J. (1999). A practical r&d project-selection scoring tool. *IEEE Transactions on Engineering Management*, 46(2), 158–70.
- Killen, C. P., Hunt, R. A., & Kleinschmidt, E. J. (2007). *Managing the New Product Development Project Portfolio: A Review of the Literature and Empirical Evidence*. PICMET'07-2007 Portland International Conference on Management of Engineering & Technology, IEEE (pp. 1864–74).
- Levine, H. A. (2005). *Project portfolio management: A Practical Guide to Selecting Projects, Managing Portfolios, and Maximizing Benefit*. San Francisco: Jossey-Bass, John Wiley & Sons.
- Liu, S.-S., & Wang, C.-J. (2011). Optimizing project selection and scheduling problems with time-dependent resource constraints. *Automation in Construction*, 20(8), 1110–9.
- Markowitz, H. (1952). PORTFOLIO SELECTION. *The Journal of Finance*, 7(1), 77–91.
- Martinsuo, M. (2013). Project portfolio management in practice and in context. *International Journal of Project Management*, 31(6), 794–803.
- McFarlan, W. F. (1981). Portfolio approach to information systems. *Harvard Business Review*, 59(5), 142–50.
- Pajares, J., & López, A. (2014). New Methodological Approaches to Project Portfolio Management: The Role of Interactions within Projects and Portfolios. *Procedia-Social and Behavioral Sciences*, 119, 645–52.
- Pennypacker, J. S., & Dye, L. D. (2002). *Project Portfolio Management and Managing Multiple Projects: Two Sides of the Same Coin?* *Managing Multiple Projects: Planning, Scheduling, and Allocating Resources for Competitive Advantage*, CRC Press (pp. 1–10).

- Petit, Y., & Hobbs, B. (2010). Project portfolios in dynamic environments: Sources of uncertainty and sensing mechanisms. *Project Management Journal*, 41(4), 46–58.
- Project Management Institute. (2013a). *A Guide to the Project Management Body of Knowledge (PMBOK). Management*. Project Management Institute.
- Project Management Institute. (2013b). *The Standard for Portfolio Management. Book*. Project Management Institute.
- Rajegopal, S., Waller, J., & McGuin, P. (2007). *Project Portfolio Management: Leading the Corporate Vision*. Palgrave Macmillan.
- Scott, W. R. (2002). *Organizations: Rational, Natural, and Open Systems*. Prentice Hall.
- Shahandashti, S. M., & Ashuri, B. (2016). Highway Construction Cost Forecasting Using Vector Error Correction Models. *Journal of Management in Engineering*, 32(2), 4015040.
- Thomas Ng, S., Cheung, S. O., Martin Skitmore, R., Lam, K. C., & Wong, L. Y. (2000). Prediction of tender price index directional changes. *Construction Management and Economics*, 18(7), 843–52.
- Turner, J. R., & Müller, R. (2003). On the nature of the project as a temporary organization. *International Journal of Project Management*, 21(1), 1–8.
- Wong, J. M. W., & Ng, S. T. (2010). Forecasting construction tender price index in Hong Kong using vector error correction model. *Construction Management and Economics*, 28(12), 1255–68.
- Young, M., & Conboy, K. (2013). Contemporary project portfolio management: Reflections on the development of an Australian competency standard for project portfolio management. *International Journal of Project Management*, Elsevier Ltd, 31(8), 1089–100.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
