

MODELLING THE STOCHASTIC DEPENDENCE UNDERLYING CONSTRUCTION COST AND DURATION

Xue XIAO^{1, 2}, Fan WANG^{1*}, Heng LI¹, Martin SKITMORE²

¹*Department of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong*

²*Science and Engineering Faculty, Queensland University of Technology, Brisbane, Australia*

Received 15 May 2018; accepted 28 June 2018

Abstract. Construction cost and duration are two critical project indicators. It is acknowledged that these two indicators are closely dependent and highly uncertain due to various common factors and limited data for explanatory model calibration. However, the stochastic dependence underlying construction cost and duration is usually ignored and the subsequent probabilistic analysis can be misleading. In response, this study develops a Nataf distribution model of building cost and duration, in which the uncertainties of total cost, unit cost, and duration are respectively quantified by univariate distribution fitting, while their stochastic dependence is inferred by maximum likelihood estimation. This method is applied to the costs and durations of 77 China residential building projects completed between 2011 and 2016. The goodness of fit test illustrates that the data conform well to the developed model. The conditional distributions of cost and duration are then derived and the corresponding conditional expectations and variances are given. The results provide the distribution of building costs for a desired duration and the expected duration given a budget. This, together with the ability to update probabilities when new project information is available, confirms the potential of the proposed model to benefit precontract decision making from a risk perspective.

Keywords: probabilistic modelling, cost-duration, stochastic dependence, Nataf distribution.

Introduction

Construction cost and duration are two critical project indicators. It is widely considered that construction cost and duration are interrelated (Bromilow 1969; Žujo *et al.* 2017). Several studies indicate that construction cost can be used to estimate the construction time (Bromilow 1969; Ogunsemi, Jagboro 2006) while some works explore the influence of duration delays on cost overrun (Huo *et al.* 2018; Lo *et al.* 2006). Despite the close relationship between construction cost and duration, they are usually deemed as two contradictory objectives that need to be traded off in project management. A planned budget largely determines a reasonable construction time through its impact on resource supply. In other words, completion within a certain period usually requires sufficient budget. From the owner's perspective, determining the planned budget demands a good understanding of the dependence between construction cost and duration. Some planners may prioritize financial and budgetary factors in early project planning while others would consider completion on time of paramount importance. Understanding this

dependency between construction cost and duration help planners to select reasonable duration according to the budget limit and make reasonable decision according to their risk preference.

The closeness of construction cost and time relation has been proved by numerous researches. Several studies have examined the causes of cost and duration overrun in different regions, including Vietnam (Kim *et al.* 2018), Nigeria (Dada, Jagboro 2007), Hong Kong (Huo *et al.* 2018; Lo *et al.* 2006), and Malaysia (Ismail *et al.* 2014; Shehu *et al.* 2014a, 2014b). Construction delay has been identified as one of the major reasons for construction cost overrun (Flyvbjerg *et al.* 2004). On the other hand, schedule delay often occurs due to an insufficient budget for covering the extra costs originating from unpredicted events such as bad weather or design changes. Some owners prioritize financial and budgetary factors in early project planning while others consider completion on time of paramount importance.

*Corresponding author. E-mail: wang_fan@hust.edu.cn

Several studies have attempted to address the prediction of construction cost and duration by using both qualitative and quantitative methods. This has resulted in several factors being identified as the major causes of construction duration and cost overruns, including ineffective early project planning, changes in design or scope, and late payments (Famiyeh *et al.* 2017; Larsen *et al.* 2015; Mulla, Waghmare 2015; Shehu *et al.* 2014a). Studies of the relationship between construction duration and cost began with a simple linear model (Fulkerson 1961), followed later by the fitting of several non-linear models, such as the quadratic (Deckro *et al.* 1995) and exponential (Žujo *et al.* 2017). A representative model developed by Bromilow (1969) in Australia has had its predictive ability tested and improved in different countries (Chan 1999; Kaka, Price 1991; Ogunsemi, Jagboro 2006; Ng *et al.* 2001; Yeong 1994). Optimization approaches based on these models have been proposed to address the cost-duration tradeoff problem; these methods include the genetic algorithm (Agdas *et al.* 2017; Koo *et al.* 2015; Zheng *et al.* 2004), artificial bee colony algorithms (Tran *et al.* 2015), and particle swarm optimization (Aminbakhsh, Sonmez 2016) and critical path (Alavipour, Arditi 2018).

Various factors contribute to construction performance, including project complexity, financial considerations, political conditions, and market environments. Some are difficult to quantify because of the limited amount of data available (particularly at the early stage of construction), while other factors are difficult to incorporate due to limited knowledge of their explanatory relations, such as project communication and team hierarchical diversity (Sanchez *et al.* 2017). Consequently, predicting construction cost and duration is surrounded by considerable uncertainty; even the performance of projects with an identical design still differ because of these uncertainties (Pinto, Morris 2004).

In minimizing the risk of unrealistic project planning, therefore, it is critical to capture the uncertainty in construction cost and duration. This is difficult to do by deterministic methods as they only provide point-value estimates (Kim, Reinschmidt 2009). As shown in Figure 1, due to the uncertainty of the underlying factors, the relationship between cost and duration cannot be represented by a deterministic function; rather, for a given duration (e.g. d in Figure 1), the corresponding cost (c in Figure 1)

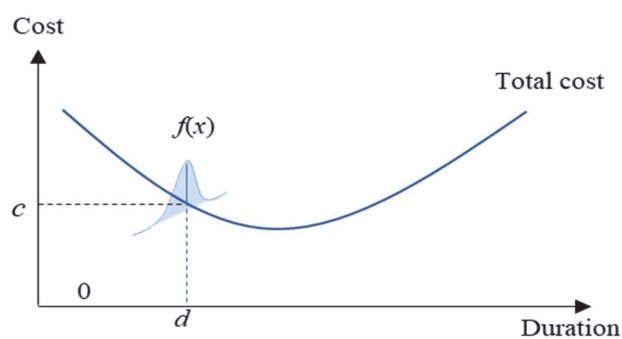


Figure 1. Construction duration-cost relationship

is uncertain and better characterized by a *distribution* of probable points ($f(x)$ in Figure 1).

In contrast, the probabilistic approach provides a *distributive* estimate rather than point estimate, which is potentially more useful (Skitmore 2001). The fitness of several general distribution types such as Weibull (Nassar *et al.* 2005) and uniform (Fine, Hackemer 1970; Grinyer, Whittaker 1973; Van Cauwelaert, Heynig 1979), Gamma (Friedman 1956), and lognormal (Skitmore 1991) have been explored for construction cost or duration. Other studies investigate the distribution of duration under different weather conditions (Lee *et al.* 2009), project and contract types (Irfan *et al.* 2011), and logic types (Wang 2005). Based on predefined distributions, several probabilistic methods such as Monte Carlo simulation (MCS), the program evaluation and review technique (PERT) and agent-based simulation have been widely used to address the uncertainty in construction duration (Ahuja, Nandakumar 1985; Ballesteros-Pérez 2017; Erol *et al.* 2017; Farshchian *et al.* 2017; Karabulut 2017; Moret, Einstein 2016). Based on assumptions of precedence relationships between activities, these methods can model the distribution of scheduled duration (Wang 2005). However, a common assumption underlying probabilistic simulation is the independence between variables (Irfan *et al.* 2011), whereas construction duration and cost are closely dependent (Žujo *et al.* 2017), raising concerns over the accuracy of the results obtained under the independence assumption (Isidore, Back 2002). Therefore, there is a need for a probabilistic method considering the stochastic dependence between construction cost and duration. The significance of cost and time dependency modelling is prominent within the context of early project decision making. Many construction companies consider the cost and time are main objectives for project management. Quantifying the changes of the cost caused by the deviation of the duration from a probabilistic perspective helps the planner to better select the duration and prepare for the possible crashing.

This study employs the Nataf distribution to capture the uncertainty of construction duration and cost and their stochastic dependence (Žujo *et al.* 2017). The Nataf distribution is widely used to describe dependent uncertainties. Unlike other multivariate distribution models, the Nataf distribution is more flexible in the choice of univariate normal or non-normal distributions. Moreover, the stochastic dependence is modeled independent of the uncertainty characterization of each random variable, which simplifies the procedures of model development. Several studies have employed the Nataf distribution for construction reliability analysis (Chen *et al.* 2015).

This paper introduces a probability model based on the Nataf distribution to capture the dependent uncertainty between cost and duration. This model is developed and tested on a dataset comprising the total cost (TC), unit cost (UC), and duration (D) of 77 China residential building projects. UC (total final contract cost/total floor area) is included here because it was also deemed a key indicator of construction performance (Chan, A. P. C., Chan, A. P. L.

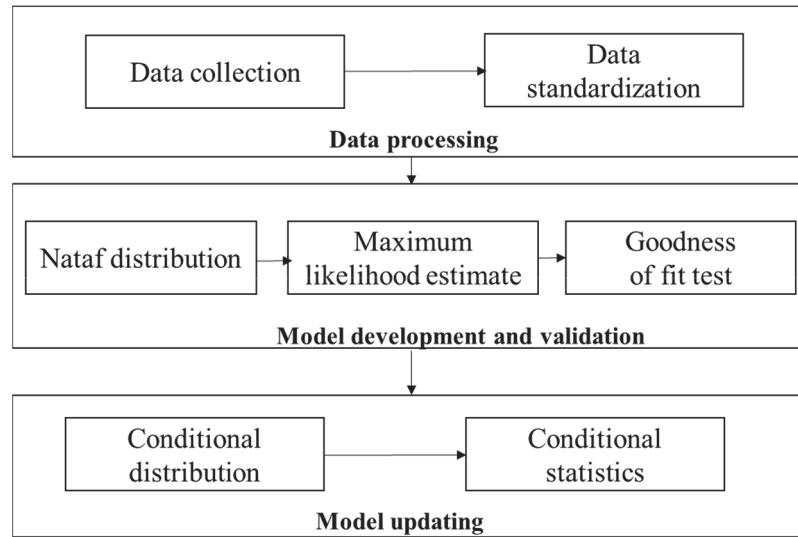


Figure 2. Methodology framework

2004). Univariate and multivariate analysis of TC , UC and D are carried out and the conditional distribution based on the multivariate distribution model is derived, and the method of updating of distributions given new observations is then illustrated. Finally, the potential application of the proposed model is presented and its implication for early stage project management is discussed.

1. Methodology

This research aims to model the stochastic relationship between building cost and duration. A multistep process is proposed (Figure 2), which includes data processing, model development and validation, and model updating. Each step is presented in detail in the subsequent paragraphs.

1.1. Data source and processing

The data collected in this study was from a professional database in the construction industry in China, which covers various types of construction buildings. The detailed information of each project is stored in a structured way that each project sample consists of several fundamental construction attributes, such as project type, storey, total floor area, bidding, decoration standard, delivery method. This study focuses on one major type of project, i.e., apartment building. Moreover, cost and schedule overrun is caused by some errors, omissions, or changes (Love *et al.* 2012). Therefore, the data samples corresponding to projects with cost or schedule overruns are removed to minimize the influence of unpredicted factors. In other words, this study focuses on the cost-duration relationship that a ‘normal’ project is likely to possess.

To eliminate the bias caused by the inconsistency in commence date and location, the project cost data was adjusted to the identical point of time using the price indices of construction from the National Bureau of Statistics of China. This index is an artificial statistical data used to es-

timate price fluctuation by assuming 100 as the price at a certain point in time (Koo *et al.* 2010).

1.2. Model development and validation

Investigating the stochastic relationship between cost and duration requires the modeling of their joint probability distribution. This study employs the Nataf distribution to capture the dependency uncertainty of building cost and duration. The major advantage of this model comes in terms of its flexibility to incorporate various univariate distributions while preserving the correlations between variables. The model is useful because the uncertainties of building cost and duration are usually non-Gaussian and correlated. The model parameter (including the univariate distribution) is estimated by the maximum likelihood method. Finally, univariate and multivariate goodness of fit tests are conducted to examine whether the uncertainties of cost and duration and their stochastic dependence are well represented by the developed model. The following explains the method used in this study.

The Nataf distribution is a popular model used to approximate the joint probability distribution of the input random variables. The model is flexible and efficient since it can incorporate various marginal distributions and can be easily generalized to higher dimensions. The stochastic dependence is modeled via a joint Gaussian distribution after mapping the random variables from the original space onto standard normal space (Chen *et al.* 2015).

Suppose $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is an n -dimensional random vector. Assume that the univariate probability density function (PDF) and cumulative distribution function (CDF) of each element $X_i, i = 1, \dots, n$ are $f_i(x_i)$ and $F_i(x_i)$, respectively. A transformation $\mathbf{Y} = T(\mathbf{X}): \mathbb{R}^n \rightarrow \mathbb{R}^n$ is used to map each element X_i independently onto standard normal random space:

$$y_i = \Phi^{-1}[F_i(x_i)], \quad (1)$$

where $\Phi^{-1}(\cdot)$ is the inverse of a univariate standard normal CDF denoted as $\Phi(\cdot)$; y_i is a standard normal random variable. By applying the differentiating rule, the joint PDF and CDF of \mathbf{X} are:

$$f_{\mathbf{X}}(\mathbf{x}) = \phi_n(\mathbf{y}; \boldsymbol{\rho}_0) \prod_{i=1}^n \frac{f_i(x_i)}{\phi(y_i)}; \tag{2}$$

$$F_{\mathbf{X}}(\mathbf{x}) = \Phi_n(\mathbf{y}; \boldsymbol{\rho}_0), \tag{3}$$

where $\phi_n(\cdot; \boldsymbol{\rho}_0)$ and $\Phi_n(\cdot; \boldsymbol{\rho}_0)$ are the PDF and CDF of an n -variate standard normal distribution respectively, and $\boldsymbol{\rho}_0$ is its correlation matrix. The joint probability distribution of \mathbf{X} with the PDF given by Eqn (2) is referred to as the Nataf distribution and $\boldsymbol{\rho}_0$ is termed the fictive correlation matrix acting as the Nataf distribution model parameter (Higham 2002).

Maximum likelihood estimation (MLE) is extensively used to infer parameters of statistical models from a set of observations. The basic idea is to maximize the likelihood of the statistical model given the observations with respect to the model parameters. For the Nataf distribution model, there are two types of parameters: (1) the parameter of the univariate distribution of each random variable; and (2) the fictive correlation matrix describing the dependence strength between different random variables. Estimating the Type 1 parameters for a univariate distribution is straightforward. For the Type 2 parameters, the inference by MLE where the log-likelihood of the model is:

$$\log f_{\mathbf{X}}(\mathbf{x}) = \log \phi_n(\mathbf{y}; \boldsymbol{\rho}_0) + \sum_{i=1}^n f_i(x_i) - \sum_{i=1}^n \phi(y_i). \tag{4}$$

Given specified marginal distributions, both $\sum_{i=1}^n f_i(x_i)$ and $\sum_{i=1}^n \phi(y_i)$ are independent of $\boldsymbol{\rho}_0$. Thus the maximization of $\log f_{\mathbf{X}}(\mathbf{x})$ with respect to $\boldsymbol{\rho}_0$ is equivalent to the maximization of $\log \phi_n(\mathbf{y}; \boldsymbol{\rho}_0)$ with respect to $\boldsymbol{\rho}_0$. In other words, the estimator of $\boldsymbol{\rho}_0$ is the estimator of the correlation matrix of an n -variate standard normal distribution, which is:

$$\rho_{0,ij} = \frac{\sum_{k=1}^m (\hat{Y}_{k,i} - \bar{Y}_i)(\hat{Y}_{k,j} - \bar{Y}_j)}{\sqrt{\sum_{k=1}^m (\hat{Y}_{k,i} - \bar{Y}_i)^2 \sum_{k=1}^m (\hat{Y}_{k,j} - \bar{Y}_j)^2}}, \tag{5}$$

where $\hat{\mathbf{Y}}_i = (\hat{Y}_{1,i}, \dots, \hat{Y}_{m,i})^T$ and $\hat{\mathbf{Y}}_j = (\hat{Y}_{1,j}, \dots, \hat{Y}_{m,j})^T$ are transformed from the m observations of the i -th and j -th random variable by Eqn (4) respectively; \bar{Y}_i and \bar{Y}_j denote the sample mean of $\hat{\mathbf{Y}}_i$ and $\hat{\mathbf{Y}}_j$, respectively.

The goodness of fit of a statistical model measures the discrepancy between the proposed statistical model and the observations. This measure is commonly used in statistical hypothesis testing to test whether the data frequencies can be captured by a given distribution. The goodness of fit of a univariate distribution model is measured in

terms of the deviation between the theoretical CDF $F(x)$ from the proposed model and the empirical CDF $\hat{F}(x)$ from observations, which are respectively:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt; \tag{6}$$

$$\hat{F}(x) = \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{\hat{X}_i \leq x}, \tag{7}$$

where $f(x)$ is the theoretical PDF of X , $\hat{X}_i, i=1, \dots, m$ are m observations of X sorted in an ascending order, and $\mathbf{1}_E$ is the indicator function of event E .

The Anderson–Darling (A–D) statistic (A^2) has excellent power properties against a variety of alternatives (D’Agostino 2017). Therefore, in this work, the A–D test is taken to examine whether the data samples are drawn from the proposed univariate distribution model, where the A–D statistic (A^2) is defined as:

$$A^2 = -m - \frac{1}{m} \sum_{i=1}^m (2i-1) \left[\log F(\hat{X}_i) + \log(1 - F(\hat{X}_{m-i+1})) \right]. \tag{8}$$

For a multivariate distribution model, the PIT-based A–D test is employed to test its goodness of fit. Particularly, to examine whether the observations of a random vector \mathbf{X} are appropriately modeled by the Nataf distribution, a null hypothesis is defined as:

H_0 : \mathbf{X} conforms to the Nataf distribution.

The PIT of \mathbf{X} is a transformation $\mathbf{Z} = T(\mathbf{X}) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by:

$$z_i = P(X_i \leq x_i | X_1 = x_1, \dots, X_{i-1} = x_{i-1}) = F(x_i | x_1, \dots, x_{i-1}), \quad i = 1, \dots, n. \tag{9}$$

It has been proved that the elements of the random vector $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ are uniformly and independently distributed on $[0, 1]$ (Rosenblatt 1952). Therefore, the random variable $S = \sum_{i=1}^n [\Phi^{-1}(Z_i)]^2$ has a χ^2 distribution with n degree of freedom (DOF). In other words, if the random vector \mathbf{X} conforms to the Nataf distribution, the corresponding observations that transformed into S should conform to a χ_n^2 distribution. By defining the auxiliary null hypothesis:

H_0^* : S conforms to the χ_n^2 distribution.

The null hypothesis H_0 can then be tested by testing H_0^* . In other words, H_0 cannot be rejected if H_0^* cannot be rejected because H_0 implies H_0^* . The null hypothesis H_0^* can again be tested using the A–D statistic.

To visualize the fitness of the Nataf distribution, the theoretical CDF is plotted based on Eqn (2) whereas the multivariate empirical CDF is defined as:

$$\hat{F}_{\mathbf{X}}(\mathbf{x}) = \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{[(\hat{X}_{i,1} \leq x_1) \cap \dots \cap (\hat{X}_{i,n} \leq x_n)]}, \tag{10}$$

where \cap indicates the intersection of event sets.

1.3. Model updating

The established Nataf distribution model could be updated to infer the conditional distribution of some random variables given new observations of the others. In addition, the calculation of conditional statistics such as expectation and variance are introduced.

A potential application of the established Nataf distribution model is to infer the conditional distribution of some random variables given new observations of the others. Let $\mathbf{X} = (\mathbf{X}^{(1)}, \mathbf{X}^{(2)})$ where $\mathbf{X}^{(1)} = (X_1, \dots, X_k)$ and $\mathbf{X}^{(2)} = (X_{k+1}, \dots, X_n)$. The corresponding transformed standard normal random variable is then denoted as $\mathbf{Y} = (\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)})$ where $\mathbf{Y}^{(1)} = (Y_1, \dots, Y_k)$ and $\mathbf{Y}^{(2)} = (Y_{k+1}, \dots, Y_n)$. Suppose that the Nataf distribution has been established where the model parameter

$\boldsymbol{\rho}_0 = \begin{bmatrix} \rho_0^{11} & \rho_0^{12} \\ \rho_0^{21} & \rho_0^{22} \end{bmatrix}$, ρ_0^{11} , ρ_0^{22} , and $\rho_0^{12} = (\rho_0^{21})^T$ are the correlation matrix of $\mathbf{Y}^{(1)}$, of $\mathbf{Y}^{(2)}$, and between $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$, respectively. Without loss of generality, suppose that new observations of $\mathbf{X}^{(2)}$ is available. By definition, the PDF of the conditional distribution $f_{\mathbf{X}^{(1)}|\mathbf{X}^{(2)}}(\mathbf{x}^{(1)}|\mathbf{x}^{(2)})$ is:

$$f_{\mathbf{X}^{(1)}|\mathbf{X}^{(2)}}(\mathbf{x}^{(1)}|\mathbf{x}^{(2)}) = \frac{f_{\mathbf{X}}(\mathbf{x})}{f_{\mathbf{X}^{(2)}}(\mathbf{x}^{(2)})} = \frac{\phi_n(\mathbf{y}; \boldsymbol{\rho}_0)}{\phi_{n-k}(\mathbf{y}^{(2)}; \boldsymbol{\rho}_0^{22})} \prod_{i=1}^k \frac{f_i(x_i)}{\phi(y_i)} \quad (11)$$

while the CDF of the conditional distribution $F_{\mathbf{X}^{(1)}|\mathbf{X}^{(2)}}(\mathbf{x}^{(1)}|\mathbf{x}^{(2)})$ is:

$$F_{\mathbf{X}^{(1)}|\mathbf{X}^{(2)}}(\mathbf{x}^{(1)}|\mathbf{x}^{(2)}) = \frac{\int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_k} f_{\mathbf{X}}(u_1, \dots, u_k, x_{k+1}, \dots, x_n) du_1 \dots du_k}{f_{\mathbf{X}^{(2)}}(\mathbf{x}^{(2)})} = \frac{\int_{-\infty}^{y_1} \int_{-\infty}^{y_2} \dots \int_{-\infty}^{y_k} \phi_n(v_1, \dots, v_k, y_{k+1}, \dots, y_n; \boldsymbol{\rho}_0) dv_1 \dots dv_k}{\phi_{n-k}(y_{k+1}, \dots, y_n; \boldsymbol{\rho}_0^{22})} = \Phi_{\mathbf{Y}^{(1)}|\mathbf{Y}^{(2)}}(\mathbf{y}^{(1)}|\mathbf{y}^{(2)}), \quad (12)$$

which is the CDF of a multivariate normal distribution. In other words, the CDF of $(\mathbf{X}^{(1)}|\mathbf{X}^{(2)})$ is equivalent to the CDF of $(\mathbf{Y}^{(1)}|\mathbf{Y}^{(2)})$ which follows a k -variate normal distribution $N_k(\boldsymbol{\mu}^*, \boldsymbol{\rho}^*)$ with mean and covariance matrix respectively given by:

$$\boldsymbol{\mu}^* = \boldsymbol{\rho}_0^{12} (\boldsymbol{\rho}_0^{22})^{-1} \mathbf{y}^{(2)}; \quad (13)$$

$$\boldsymbol{\rho}^* = \boldsymbol{\rho}_0^{11} - \boldsymbol{\rho}_0^{12} (\boldsymbol{\rho}_0^{22})^{-1} \boldsymbol{\rho}_0^{21}. \quad (14)$$

For the computation of the marginal conditional CDFs in the PIT, we can simply apply Eqns (12)–(14) by setting $\mathbf{X}^{(1)} = X_i$ and $\mathbf{X}^{(2)} = (X_1, \dots, X_{i-1})$ and choosing the corresponding sub-matrix of $\boldsymbol{\rho}_0$ as the model parameter.

Based on the conditional distribution, the conditional statistical moments can be calculated. For example, the conditional expectation and variance are calculated as follows:

$$\mathbb{E}(X|\mathbf{e}) = \int_{-\infty}^{+\infty} xf(x|\mathbf{e})dx; \quad (15)$$

$$\mathbb{V}(X|\mathbf{e}) = \mathbb{E}\left((X - \mathbb{E}(X|\mathbf{e}))^2|\mathbf{e}\right), \quad (16)$$

where \mathbb{E} is the expectation operator, \mathbb{V} is the variance operator, \mathbf{e} denotes the observations of conditioning variables.

2. Data collected

Considering the consistency in the project attributes, this study finally collected 77 project samples of apartment buildings from the database. These projects were completed during 2011 to 2016 in China without occurrence of cost and schedule overrun. The project cost was adjusted using the price indices of construction from the National Bureau of Statistics of China to eliminate the effects of inflation.

The statistics of the data are shown in Table 1. The Pearson's correlation between TC and D is -0.4949 , UC and D is -0.3621 , TC and UC is 0.3112 , indicating that TC and D , UC , and D are negatively correlated while there is a positive correlation between TC and UC .

3. Results

This study includes three variables (UC , TC , D) indicating project performance, which are stochastic and interdependent. The target is their statistical modeling and unravelling their stochastic relationship. Univariate analysis and multivariate analysis are both included in this section.

3.1. Univariate analysis

To establish the model, the first step is to quantify the uncertainty of each random variable by determining the univariate distributions. A number of continuous distributions are examined, including the Beta, Burr, Weibull, Log-Logistic, Cauchy, Gumbel Max/Min, Gamma, Johnson SB, and Normal. The parameters of the distributions are estimated using the MLE method. The software package *Easyfit 5.5* is used for parameter estimation and the

Table 1. Statistics of project cost and duration

Project information	UC (¥/m ²)	TC (¥)	D (days)
Mean	1524.3	11357759.73	545.61
Variance	155906.01	4.83E+13	23925.24
Maximum Value	2546.63	29440351.52	1205
Minimum Value	861.17	2953483.15	284
No. of projects	77		

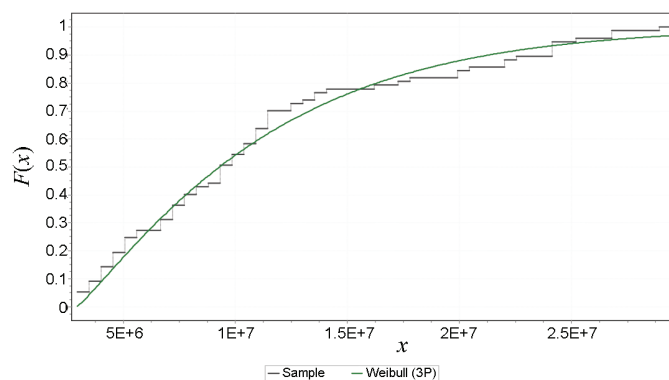
goodness of fit test of the univariate distributions. The null hypothesis H_0 that the observations follow the specified distribution cannot be rejected if the statistic A^2 is less than the critical value at the given significance level α ; otherwise, the null hypothesis is rejected.

Table 2 presents the hypothesis test results corresponding to the best-fit distributions for UP , TP , and D , respectively. The probability density functions (PDFs) and cumulative distribution functions (CDFs) of Weibull (3-parameter), Gamma (3-parameter), and Log-Logistic (3-parameter) distributions and the meaning of associated parameters are given in Table 3. Figure 3 compares the empirical distribution (stair step lines) with the best fit distribution for UP , TP , and D respectively. The empirical CDF

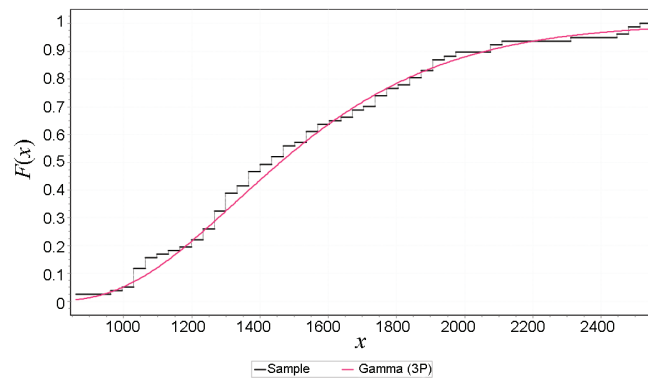
is generally in good agreement with the theoretical CDF, which indicates that the uncertainty of UP , TP , and D can be characterized by the theoretical CDFs.

3.2. Multivariate analysis

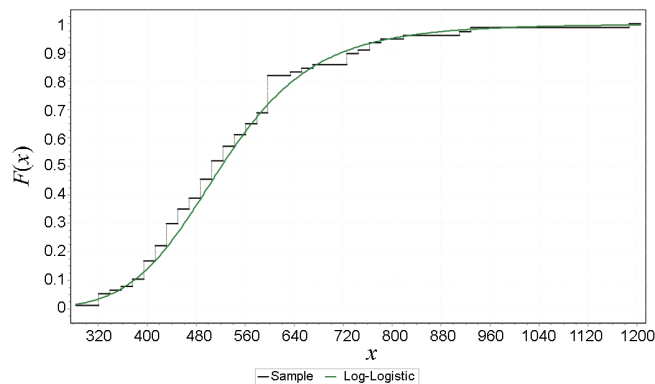
The next step in developing the Nataf distribution model is to estimate the fictive correlation matrix. The results are shown in Table 4. The fictive correlation matrix is close to, but deviates from, the correlation matrix associated with the original random vector. In general, the fictive correlation and original correlation are co-monotonic (i.e. fictive correlation strictly increases as the original correlation strictly increases, and *vice versa*) and the fictive correla-



(a) CDF of TC



(b) CDF of UC



(c) CDF of D

Figure 3. Best fit distributions of TC , UC and D

Table 2. Hypothesis test based on the A–D statistic

Variable	Distribution type	A–D statistic	α -value	Critical value	Reject?
TC	Weibull (3-parameter) $\alpha = 1.134$ $\beta = 8,804,609.80$ $\gamma = 2,918,564.67$	0.4835	0.2	1.3749	No
			0.1	1.9286	No
			0.05	2.5018	No
			0.02	3.2892	No
			0.01	3.9074	No
UC	Gamma (3-parameter) $\alpha = 4.07$ $\beta = 198.15$ $\gamma = 717.55$	0.2325	0.2	1.3749	No
			0.1	1.9286	No
			0.05	2.5018	No
			0.02	3.2892	No
			0.01	3.9074	No
D	Log-Logistic 3-parameter) $\alpha = 4.37$ $\beta = 332.56$ $\gamma = 185.01$	0.1938	0.2	1.3749	No
			0.1	1.9286	No
			0.05	2.5018	No
			0.02	3.2892	No
			0.01	3.9074	No

Table 3. Distribution functions

Distribution type	PDF and CDF
Log-logistic distribution (3-parameter)	$f(x) = \frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta} \right)^{\alpha-1} \left(1 + \left(\frac{x-\gamma}{\beta} \right)^{\alpha} \right)^{-2}$ $F(x) = \left(1 + \left(\frac{\beta}{x-\gamma} \right)^{\alpha} \right)^{-1}$
Gamma distribution (3-parameter)	$f(x) = \frac{(x-\gamma)^{\alpha-1}}{\beta^{\alpha} \Gamma(\alpha)} \exp\left(-\frac{(x-\gamma)}{\beta}\right)$ $F(x) = \frac{\Gamma_{(x-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)}$
Weibull distribution (3-parameter)	$f(x) = \frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta} \right)^{\alpha-1} \exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)$ $F(x) = 1 - \exp\left(-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}\right)$

Note: α is the shape parameter, β is the scale parameter and γ is the location parameter.

Table 4. Fictive correlation matrix of the Nataf distribution

Variable	TC	UC	D
TC	1	0.3090	-0.6169
UC	0.3090	1	-0.3985
D	-0.6169	-0.3985	1

tion is positive or negative if the original correlation is positive or negative. The deviation is largely determined by the normality of the univariate distribution of each random variable.

The goodness of fit of the developed multivariate distribution model is then examined. A critical step in the goodness of fit test of the Nataf distribution model is the PIT, which involves the sequential computation of the marginal conditional CDFs of the random vector. The result of the A–D test for goodness of fit is shown in Table 5. This indicates that the auxiliary hypothesis H_0^* (statistical S conforms to the χ_n^2 distribution) is accepted, which means that the Nataf distribution can be used to model the trivariate distribution of the random vector. The histogram of S compared with the PDF of χ_3^2 distribution is shown in Figure 4 and the empirical cumulative distribution of S compared with the CDF of χ_3^2 distribution is

Table 5. A–D test for H_0^*

Variable	Distribution type	A–D statistic	α -value	Critical value	Reject?
S	χ_3^2 distribution	0.30545	0.2	1.375	No
			0.1	1.929	No
			0.05	2.502	No
			0.02	3.289	No
			0.01	3.907	No

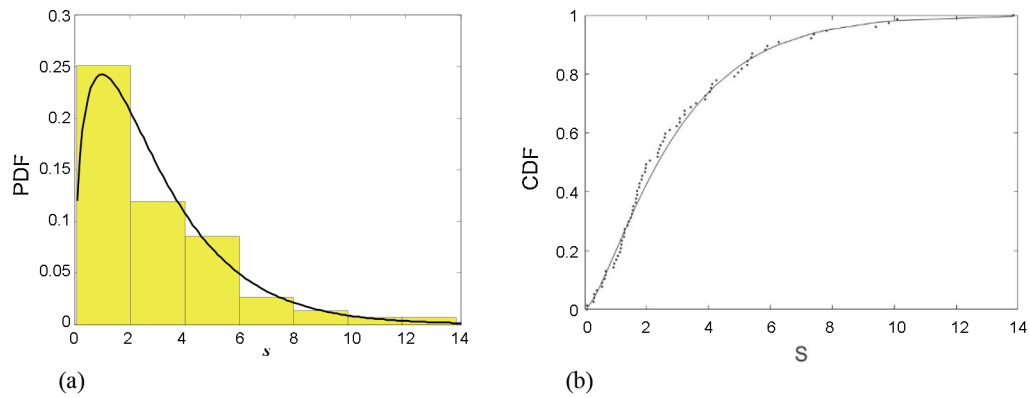


Figure 4. (a) Histogram of S compared with the PDF of χ_3^2 distribution, (b) P-P plot of S compared with the CDF of χ_3^2 distribution

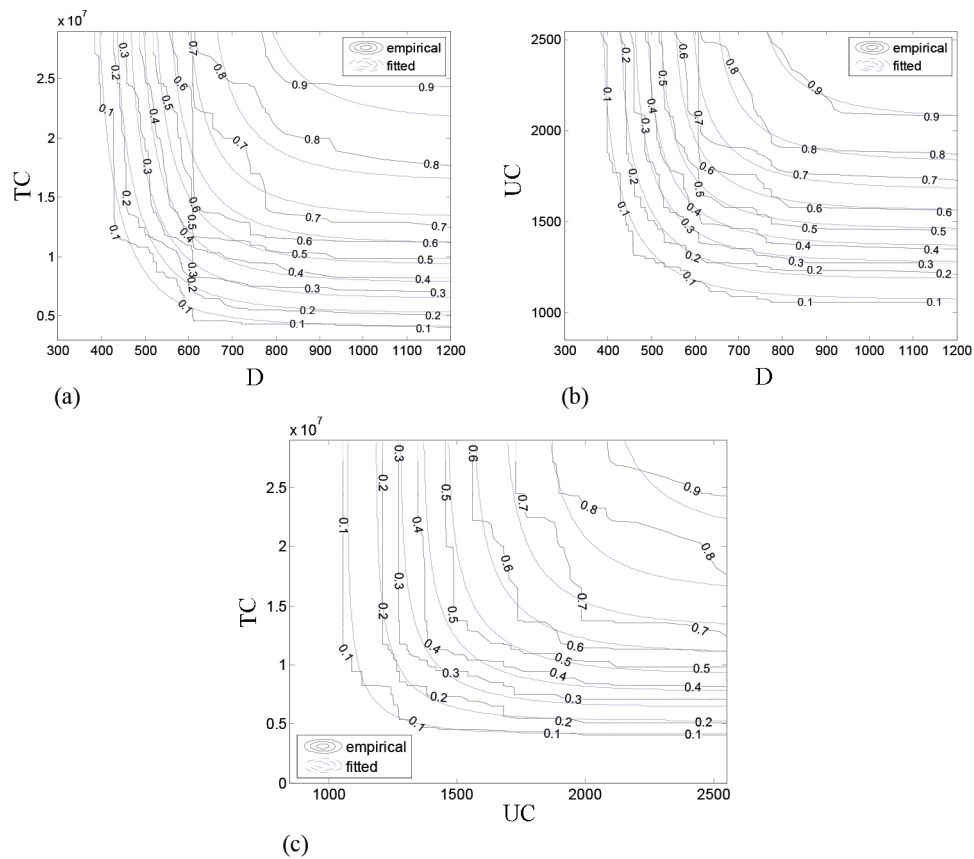


Figure 5. Comparison of contours of the empirical distribution and the Nataf distribution in dimensions: (a) TC - D dimension; (b) UC - D dimension, and (c) TC - UC dimension

shown in Figure 4(b). Figures 4(a) and 4(b) illustrate that the proposed model well fits the empirical data. Figure 5 compares the contours of the trivariate empirical distribution with that of the Nataf distribution in two dimensions. The empirical and fitted CDFs are generally consistent despite some small deviations.

4. Model application

An application of the established Nataf distribution model is to infer the conditional distribution of construction duration or cost when new information of the others is avail-

able. This section discusses the conditional distribution of cost or duration given new observations. In addition, the calculation of conditional expectation and variance of building cost and duration is illustrated. The potential of the proposed model for project management is discussed later.

4.1. Conditional distribution of TC or UC given D

This section discusses the conditional distribution of TC or UC given new values of D . Figures 6(a) and 6(b) present the PDF and CDF of the conditional distribution

of UC under different values of D respectively. Here, we take the 5th (355 days) and 95th (837 days) percentile of D for illustrative purpose. The unconditional distribution of UC is included as a reference. The mode of UC that corresponds to the maximum of unconditional PDF is ¥1322/m². If the duration equals to its 95th percentile, the mode of UC decreases to ¥1134/m² and the conditional CDF is above the unconditional CDF, which implies that a smaller value of UC is likely to occur. In contrast, as the duration is shortened to its 5th percentile, the mode of UC increases to ¥1619/m² and the conditional CDF is below the unconditional CDF. In other words, the probability of UC being less than a value is lower, which indicates that a larger value of UC is likely to appear. A similar pattern of influence is observed for conditional distribution of TC . When D equals its 5th and 95th percentile, the mode of TC deviates from its unconditional value of ¥4.22e6 to ¥1.50e7 and ¥3.15e6, respectively. The conditional CDFs of TC given in the 5th and 95th percentile of D are also below and above the unconditional CDF of TC . However, the deviation between the conditional and unconditional CDF lines for UC and TC are different. The following index is used to investigate the sensitivity of UC and TC to D :

$$\delta = \left| \frac{P(X_2 < x_2 | X_1 = x_1) - P(X_2 < x_2)}{P(X_2 < x_2)} \right| \times 100\%, \quad (17)$$

here, we set x_2 to the median of its unconditional distribution (i.e., $P(X_2 < x_2) = 0.5$) and x_1 to its 5th and 95th percentile, respectively. Table 6 summarizes the sensitivity

of TC and UC to D . The conditional CDFs of TC apparently deviates more significantly from its unconditional CDF than the result from UC . This difference indicates that TC is more sensitive to D .

Table 6. The sensitivity of TC and UC to D

δ -value	$X_1 = D$	
	$x_1 = F_{X_1}^{-1}(0.05)$	$x_1 = F_{X_1}^{-1}(0.95)$
$X_2 = TC, x_2 = F_{X_2}^{-1}(0.5)$	80.4%	80.0%
$X_2 = UC, x_2 = F_{X_2}^{-1}(0.5)$	52.2%	52.7%

Note: $F^{-1}(\cdot)$ is the inverse function of $F(\cdot)$ and the subscript indicates the name of the variable.

4.2. Conditional distribution of UC and TC given D

This section explores the conditional joint distribution of UC and TC given the value of D . First, the CDF and PDF of unconditional joint distribution between TC and UC are presented in Figures 7(a) and 7(b). Figures 7(c) and 7(d) show the CDF and PDF of the conditional joint distribution of TC and UC given the 5th percentile of D and Figures 7(e) and 7(f) are the CDF and PDF of conditional joint distribution of TC and UC given the 95th percentile of D . The most probable values (summit of the PDF surface) corresponding to the unconditional joint distribu-

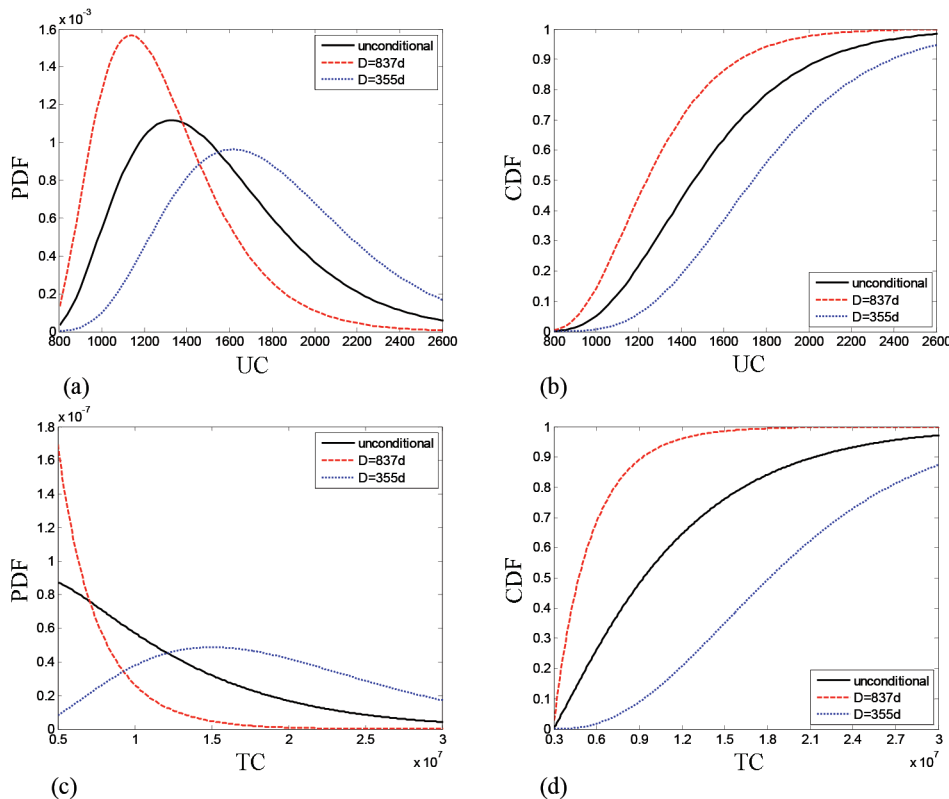


Figure 6. The conditional distribution of TC and UC given the value of D

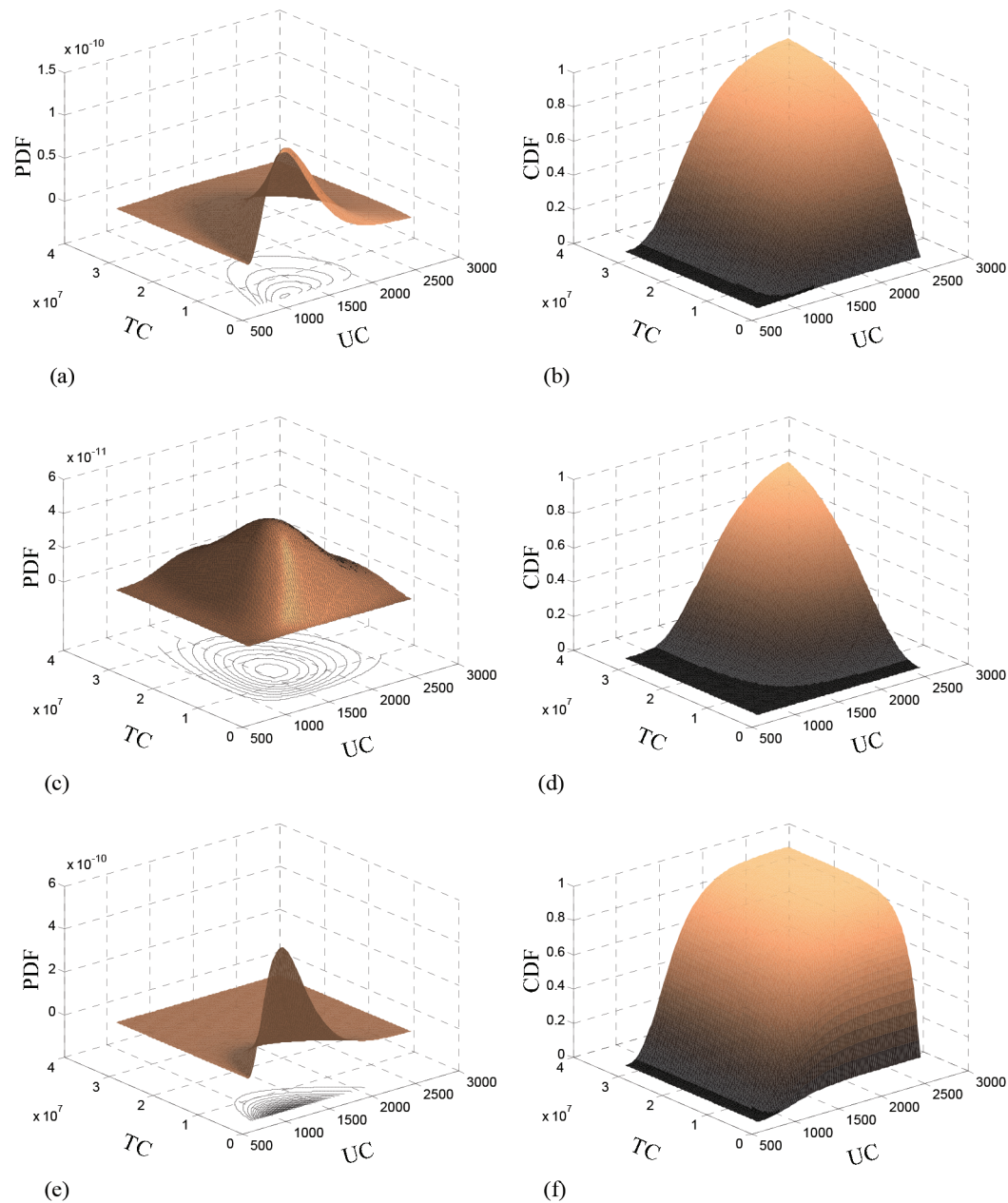


Figure 7. The unconditional and conditional joint distribution of TC and UC : (a) unconditional PDF; (b) unconditional CDF; (c) conditional PDF given the 5th percentile of D ; (d) conditional CDF given the 5th percentile of D ; (e) conditional PDF given the 95th percentile of D ; (f) conditional CDF given the 95th percentile of D

tion of UC and TC are ¥1178/m² and ¥3.52e6. Given the 5th and 95th percentile of D , the most probable values of both UC and TC become larger (¥1598/m² and ¥1.49e7) and smaller (¥1098/m² and ¥3.09e6), respectively, as shown in the contour below the PDF surface. The mode of the conditional joint distribution of UC and TC is different from that of univariate conditional distribution of UC or TC given the same value of D . For example, when D is set to its 5th percentile, the mode of UC or TC corresponding to the univariate conditional distribution is respectively ¥1619/m² or ¥1.50e7, while the mode of UC and TC becomes ¥1598/m² and ¥1.49e7, respectively, if they are jointly investigated. This difference implies the influence of the positive dependence between UC and TC on the distribution modeling.

4.3. Conditional distribution of D given UC and TC

Conversely, the possible duration under specific budgetary cost can be analyzed. Similarly, the conditional distribution of D is illustrated with UC and TC being their 5th ($UC = ¥997/m^2$, $TC = ¥3.56e6$) and 95th ($UC = ¥2274/m^2$, $TC = ¥2.61e7$) percentile. As illustrated in Figure 8, a limited budget can increase the likelihood of longer construction duration while sufficient financial support can decrease this likelihood. In particular, the mode of unconditional distribution of D is 484 days. If both UC and TC are set to their 5th percentile (or 95th percentile), the mode of D becomes to 675 days (or 385 days). This result quantifies the budgetary influence on construction duration.

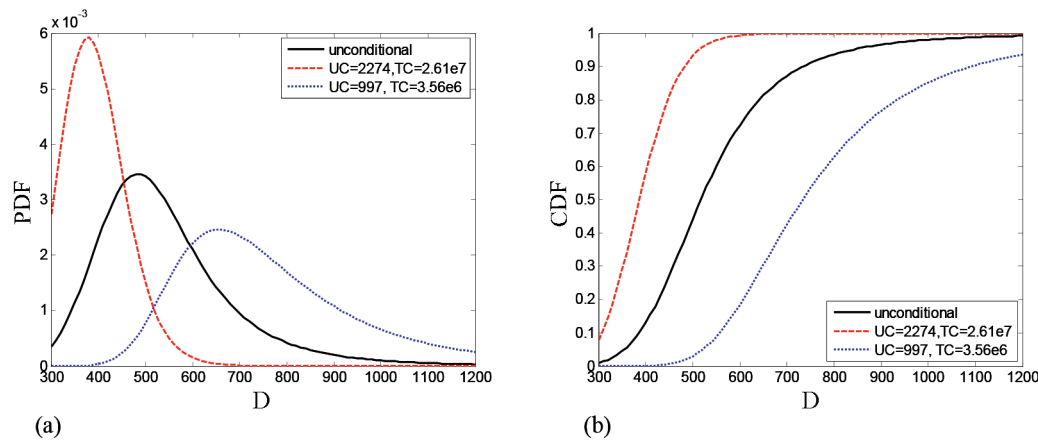


Figure 8. Conditional distribution of D given UC and TC

4.4. Conditional statistics

Based on the conditional distribution, the conditional statistics of construction duration and cost are calculated using Eqns (15) and (16). Table 7 tabulates some expectations and variances corresponding to the conditional distribution investigated above. The results illustrate that, with the increase of the cost, the expectation of duration decreases, which is consistent with the construction practice that shortening duration will cause an increase of the budget. However, this model provides a practical method to quantify this influence. For example, if the duration needs to be substantially shortened from 837 days to 355 days, the increase in total cost is expected to be ¥1.394e7. The model can therefore serve as an effective tool to support the owner and clients to estimate the changes of cost caused by shortening the construction time. The capability of the model not only results in reduced likelihood of an unreasonable schedule plan but also lowers the risk of an infeasible balance between budget and duration. As a result, owners can make rational decisions for project development and management with the aid of this model.

Conclusions

This paper proposes a probabilistic model based on the Nataf distribution to address the dependent uncertainties associated with building cost and duration. This was demonstrated and tested with a sample of 77 of residential buildings completed between 2011 and 2016 in China. The uncertainties of total cost, unit cost, and duration were quantified by univariate distribution fitting, while their stochastic dependence is inferred from the dataset by maximum likelihood estimation. The goodness of fit test indicates that the collected data conform well to the developed Nataf distribution. The conditional distribution of cost or duration given new observations was derived and the calculation of conditional expectation and variance is illustrated.

Based on the established distribution, the stochastic relationship between cost and duration is explored by investigating their respective conditional distribution given the 5th and 95th percentiles of the other variables. The results show how the distribution of TC and/or UC changes as the duration varies and *vice versa*. Moreover, it is illustrated

Table 7. Example results of conditional expectation and variance

Expectation	Variance
$\mathbb{E}(UC D=355) = 1797$	$\mathbb{V}(UC D=355) = 1.96e5$
$\mathbb{E}(UC D=837) = 1288$	$\mathbb{V}(UC D=837) = 8.44e4$
$\mathbb{E}(TC D=355) = 1.958e7$	$\mathbb{V}(TC D=355) = 8.0940e13$
$\mathbb{E}(TC D=837) = 5.6441e6$	$\mathbb{V}(TC D=837) = 7.9535e12$
$\mathbb{E}(D UC=997, TC = 3.56e6) = 795$	$\mathbb{V}(D UC=997, TC = 3.56e6) = 6.85e4$
$\mathbb{E}(D UC=2274, TC = 2.61e7) = 392$	$\mathbb{V}(D UC=2274, TC = 2.61e7) = 4.88e3$

how owners or clients can use the developed model to estimate the conditional expectation and variance of cost such that the risk of cost increasing due to shortening construction time can be assessed.

The multivariate distribution is established by modeling the uncertainty of each random variable and its stochastic dependence respectively. In general, the proposed model can update the uncertainty of project performance when new project information becomes available and it benefits on decision making in project management from a risk perspective. Quantifying the cost changes as a function of duration helps the planner to select the reasonable duration and prepare for the possible schedule crashing. Further research needs to be conducted by adding more explanatory factors to establish a more sophisticated model.

Funding

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant [number 51608399].

Author Contributions

Xue Xiao contributes to the conception and design of the study, acquisition of data, drafting and revision of the article. Fan Wang contributes to the conception and design of the study, analysis and interpretation of data, drafting and revision of the article. Heng Li and Martin Skitmore contribute to the revision of the article for important intellectual content.

References

- Agdas, D.; Warne, D. J.; Osio-Norgaard, J.; Masters, F. J. 2017. Utility of genetic algorithms for solving large-scale construction time-cost trade-off problems, *Journal of Computing in Civil Engineering* 32(1): 04017072. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000718](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000718)
- Ahuja, H. N.; Nandakumar, V. 1985. Simulation model to forecast project completion time, *Journal of Construction Engineering and Management* 111(4): 325–342. [https://doi.org/10.1061/\(ASCE\)0733-9364\(1985\)111:4\(325\)](https://doi.org/10.1061/(ASCE)0733-9364(1985)111:4(325))
- Alavipour, S. R.; Arditi, D. 2018. Optimizing financing cost in construction projects with fixed project duration, *Journal of Construction Engineering and Management* 144(4): 04018012. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001451](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001451)
- Aminbakhsh, S.; Sonmez, R. 2016. Discrete particle swarm optimization method for the large-scale discrete time–cost trade-off problem, *Expert Systems with Applications* 51: 177–185. <https://doi.org/10.1016/j.eswa.2015.12.041>
- Ballesteros-Pérez, P. 2017. M-PERT: Manual project-duration estimation technique for teaching scheduling basics, *Journal of Construction Engineering and Management* 143(9): 04017063. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001358](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001358)
- Bromilow, F. 1969. *Contract time performance expectations and the reality*. Paper presented at the Building Forum.
- Chan, A. P. 1999. Modelling building durations in Hong Kong, *Construction Management & Economics* 17(2): 189–196. <https://doi.org/10.1080/014461999371682>
- Chan, A. P. C.; Chan, A. P. L. 2004. Key performance indicators for measuring construction success, *Benchmarking: An International Journal* 11(2): 203–221. <https://doi.org/10.1108/14635770410532624>
- Chen, C.; Wu, W.; Zhang, B.; Sun, H. 2015. Correlated probabilistic load flow using a point estimate method with Nataf transformation, *International Journal of Electrical Power & Energy Systems* 65: 325–333. <https://doi.org/10.1016/j.ijepes.2014.10.035>
- D'Agostino, R. 2017. *Goodness-of-fit-techniques*. Routledge.
- Dada, J. O.; Jagboro, G. 2007. An evaluation of the impact of risk on project cost overrun in the Nigerian construction industry, *Journal of Financial Management of Property and Construction* 12(1): 37–44. <https://doi.org/10.1108/13664380780001092>
- Deckro, R. F.; Hebert, J. E.; Verdini, W. A.; Grimsrud, P. H.; Venkateshwar, S. 1995. Nonlinear time/cost tradeoff models in project management, *Computers & Industrial Engineering* 28(2): 219–229. [https://doi.org/10.1016/0360-8352\(94\)00199-W](https://doi.org/10.1016/0360-8352(94)00199-W)
- Erol, H.; Dikmen, I.; Birgonul, M. T. 2017. Measuring the impact of lean construction practices on project duration and variability: A simulation-based study on residential buildings, *Journal of Civil Engineering and Management* 23(2): 241–251. <https://doi.org/10.3846/13923730.2015.1068846>
- Famiyeh, S.; Amoatey, C. T.; Adaku, E.; Agbenohevi, C. S. 2017. Major causes of construction time and cost overruns: A case of selected educational sector projects in Ghana, *Journal of Engineering, Design and Technology* 15(2): 181–198. <https://doi.org/10.1108/JEDT-11-2015-0075>
- Farshchian, M. M.; Heravi, G.; AbouRizk, S. 2017. Optimizing the owner's scenarios for budget allocation in a portfolio of projects using agent-based simulation, *Journal of Construction Engineering and Management* 143(7): 04017022. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001315](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001315)
- Fine, B.; Hackemer, G. 1970. *Estimating and bidding strategy*. Costaine, Operational Research Department.
- Flyvbjerg, B.; Skamris Holm, M. K.; Buhl, S. L. 2004. What causes cost overrun in transport infrastructure projects?, *Transport Reviews* 24(1): 3–18. <https://doi.org/10.1080/0144164032000080494a>
- Friedman, L. 1956. A competitive-bidding strategy, *Operations Research* 4(1): 104–112. <https://doi.org/10.1287/opre.4.1.104>
- Fulkerson, D. R. 1961. A network flow computation for project cost curves, *Management Science* 7(2): 167–178. <https://doi.org/10.1287/mnsc.7.2.167>
- Grinyer, P. H.; Whittaker, J. D. 1973. Managerial judgement in a competitive bidding model, *Journal of the Operational Research Society* 24(2): 181–191. <https://doi.org/10.1057/jors.1973.36>
- Higham, N. J. 2002. *Accuracy and stability of numerical algorithms*. Vol. 80. SIAM. <https://doi.org/10.1137/1.9780898718027>
- Huo, T.; Ren, H.; Cai, W.; Shen, G. Q.; Liu, B.; Zhu, M.; Wu, H. 2018. Measurement and dependence analysis of cost overruns in megatransport infrastructure projects: Case study in Hong Kong, *Journal of Construction Engineering and Management* 144(3): 05018001. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001444](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001444)
- Irfan, M.; Khurshid, M. B.; Anastasopoulos, P.; Labi, S.; Moavenzadeh, F. 2011. Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type, and contract type, *International Journal of Project Management* 29(1): 78–92. <https://doi.org/10.1016/j.ijproman.2010.01.001>

- Isidore, L. J.; Back, W. E. 2002. Multiple simulation analysis for probabilistic cost and schedule integration, *Journal of Construction Engineering and Management* 128(3): 211–219. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2002\)128:3\(211\)](https://doi.org/10.1061/(ASCE)0733-9364(2002)128:3(211))
- Ismail, I.; Memon, A. H.; Rahman, I. A. 2014. Expert opinion on risk level for factors affecting time and cost overrun along the project lifecycle in Malaysian construction projects, *International Journal of Construction Technology and Management* 1(2): 10–15.
- Kaka, A.; Price, A. D. 1991. Relationship between value and duration of construction projects, *Construction Management and Economics* 9(4): 383–400. <https://doi.org/10.1080/01446199100000030>
- Karabulut, M. 2017. Application of Monte Carlo simulation and PERT/CPM techniques in planning of construction projects: A case study, *Periodicals of Engineering and Natural Sciences (PEN)* 5(3). <https://doi.org/10.21533/pen.v5i3.152>
- Kim, B.-C.; Reinschmidt, K. F. 2009. Probabilistic forecasting of project duration using Bayesian inference and the beta distribution, *Journal of Construction Engineering and Management* 135(3): 178–186. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2009\)135:3\(178\)](https://doi.org/10.1061/(ASCE)0733-9364(2009)135:3(178))
- Kim, S.-Y.; Tuan, K. N.; Lee, J. D.; Pham, H.; Luu, V. T. 2018. Cost overrun factor analysis for hospital projects in Vietnam, *KSCCE Journal of Civil Engineering* 22(1): 1–11. <https://doi.org/10.1007/s12205-017-0947-5>
- Koo, C.; Hong, T.; Hyun, C.; Koo, K. 2010. A CBR-based hybrid model for predicting a construction duration and cost based on project characteristics in multi-family housing projects, *Canadian Journal of Civil Engineering* 37(5): 739–752. <https://doi.org/10.1139/L10-007>
- Koo, C.; Hong, T.; Kim, S. 2015. An integrated multi-objective optimization model for solving the construction time-cost trade-off problem, *Journal of Civil Engineering and Management* 21(3): 323–333. <https://doi.org/10.3846/13923730.2013.802733>
- Larsen, J. K.; Shen, G. Q.; Lindhard, S. M.; Brunoe, T. D. 2015. Factors affecting schedule delay, cost overrun, and quality level in public construction projects, *Journal of Management in Engineering* 32(1): 04015032. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000391](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000391)
- Lee, H.-s.; Shin, J.-w.; Park, M.; Ryu, H.-G. 2009. Probabilistic duration estimation model for high-rise structural work, *Journal of Construction Engineering and Management* 135(12): 1289–1298. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000105](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000105)
- Lo, T. Y.; Fung, I. W.; Tung, K. C. 2006. Construction delays in Hong Kong civil engineering projects, *Journal of Construction Engineering and Management* 132(6): 636–649. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2006\)132:6\(636\)](https://doi.org/10.1061/(ASCE)0733-9364(2006)132:6(636))
- Love, P. E.; Edwards, D. J.; Irani, Z. 2012. Moving beyond optimism bias and strategic misrepresentation: An explanation for social infrastructure project cost overruns, *IEEE Transactions on Engineering Management* 59(4): 560–571. <https://doi.org/10.1109/TEM.2011.2163628>
- Moret, Y.; Einstein, H. H. 2016. Construction cost and duration uncertainty model: application to high-speed rail line project, *Journal of Construction Engineering and Management* 142(10): 05016010. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001161](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001161)
- Mulla, S. S.; Waghmare, A. P. 2015. A study of factors caused for time & cost overruns in construction project & their remedial measures, *International Journal of Engineering Research and Applications* 5(1): 48–53.
- Nassar, K. M.; Gunnarsson, H. G.; Hegab, M. Y. 2005. Using Weibull analysis for evaluation of cost and schedule performance, *Journal of Construction Engineering and Management* 131(12): 1257–1262. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2005\)131:12\(1257\)](https://doi.org/10.1061/(ASCE)0733-9364(2005)131:12(1257))
- Ng, S. T.; Mak, M. M. Y.; Skitmore, R. M.; Lam, K. C.; Varnam, M. 2001. The predictive ability of Bromilow's timecost model, *Construction Management and Economics* 19(2): 165–173. <https://doi.org/10.1080/01446190150505090>
- Ogunsemi, D. R.; Jagboro, G. O. 2006. Time-cost model for building projects in Nigeria, *Construction Management and Economics* 24(3): 253–258. <https://doi.org/10.1080/01446190500521041>
- Pinto, J. K.; Morris, P. W. 2004. *The Wiley guide to managing projects*. Hoboken, NJ: Wiley.
- Rosenblatt, M. 1952. Remarks on a multivariate transformation, *The Annals of Mathematical Statistics* 23(3): 470–472. <https://doi.org/10.1214/aoms/1177729394>
- Sanchez, O. P.; Terlizzi, M. A.; de Moraes, H. R. d. O. C. 2017. Cost and time project management success factors for information systems development projects, *International Journal of Project Management* 35(8): 1608–1626. <https://doi.org/10.1016/j.ijproman.2017.09.007>
- Shehu, Z.; Endut, I. R.; Akintoye, A. 2014a. Factors contributing to project time and hence cost overrun in the Malaysian construction industry, *Journal of Financial Management of Property and Construction* 19(1): 55–75. <https://doi.org/10.1108/JFMPC-04-2013-0009>
- Shehu, Z.; Endut, I. R.; Akintoye, A.; Holt, G. D. 2014b. Cost overrun in the Malaysian construction industry projects: A deeper insight, *International Journal of Project Management* 32(8): 1471–1480. <https://doi.org/10.1016/j.ijproman.2014.04.004>
- Skitmore, M. 1991. The construction contract bidder homogeneity assumption: An empirical test, *Construction Management and Economics* 9(5): 403–429.
- Skitmore, M. 2001. Raftery curves for tender price forecasting: Empirical probabilities and pooling, *Financial Management of Property and Construction* 6(3): 141–154.
- Tran, D.-H.; Cheng, M.-Y.; Cao, M.-T. 2015. Hybrid multiple objective artificial bee colony with differential evolution for the time–cost–quality tradeoff problem, *Knowledge-Based Systems* 74: 176–186. <https://doi.org/10.1016/j.knosys.2014.11.018>
- Van Cauwelaert, F.; Heynig, E. 1979. Correction of bidding errors: the Belgian solution, *Journal of the Construction Division* 105(1): 13–23.
- Wang, W.-C. 2005. Impact of soft logic on the probabilistic duration of construction projects, *International Journal of Project Management* 23(8): 600–610. <https://doi.org/10.1016/j.ijproman.2005.05.008>
- Yeong, C. M. 1994. *Time and cost performance of building contracts in Australia and Malaysia*. University of South Australia.
- Zheng, D. X.; Ng, S. T.; Kumaraswamy, M. M. 2004. Applying a genetic algorithm-based multiobjective approach for time-cost optimization, *Journal of Construction Engineering and Management* 130(2): 168–176. [https://doi.org/10.1061/\(ASCE\)0733-9364\(2004\)130:2\(168\)](https://doi.org/10.1061/(ASCE)0733-9364(2004)130:2(168))
- Žujo, V.; Car-Pušić, D.; Žileska-Pančovska, V.; Čeček, M. 2017. Time and cost interdependence in water supply system construction projects, *Technological and Economic Development of Economy* 23(6): 895–914. <https://doi.org/10.3846/20294913.2015.1071292>