CASH FLOW FORECASTING WITH RISK CONSIDERATION USING BAYESIAN BELIEF NETWORKS (BBNs)

Mostafa KHANZADI\textsuperscript{a}, Ehsan ESHTEHARDIAN\textsuperscript{b}, Mahdiyar MOKHLESPOUR ESFAHANI\textsuperscript{a}

\textsuperscript{a}Department of Civil Engineering, Iran University of Science & Technology, Tehran, Iran
\textsuperscript{b}Department of Architecture & Art, Tarbiat Modares University, Tehran, Iran

Received 24 Apr 2017; accepted 02 Aug 2017

Abstract. Cash-flow management is very important for contractors given that inadequate cash resources typically are the main causes for bankruptcy of construction companies. In comparison to most other industries, the construction industry is severely plagued by risk, and the success of construction projects usually depends on valuating all risks. However, conventional methods suggested by extant research on cash flow forecasting do not consider comprehensive identification of risk factors, interactions between the factors, and simultaneous occurrences of the factors. This study introduced a simple and appropriate probabilistic cash flow forecasting model using Bayesian Belief Networks (BBNs) to avoid bankruptcy of contractors by considering influence diagrams and risk factors that affect a project. Workability and reliability of the proposed approach was tested on an important building construction project in Iran as a real case study, and the results indicated that the model performed well.

Keywords: probabilistic cash flow, construction risk factors, Bayesian Belief Networks (BBNs), forecasting.

Introduction

The importance of and need for cash-flow management is a matter of consensus among researchers and practitioners (Navon 1996). Cash is the most important resource for construction companies (Park \textit{et al.} 2005). For several years, the construction industry has incurred a proportionally high bankruptcy rate when compared to other industries (Hwee, Tiong 2002); more than 60\% of the contractors are bankrupt owing to economic factors (Russell 1991). With respect to a contractor, cash flow consists of a schedule of income and payments throughout a construction project (Son \textit{et al.} 2006). Additionally, a cumulative (net) contractor cash flow diagram is obtained by subtracting costs from incomes. The advantages of cash flow forecasting includes helping contractors select contracts that can be financed using available resources and in making provisions during difficult times to avoid a liquidity crisis (Kaka, Price 1991). It is necessary to make cash flow forecasts at all phases of the construction process (Navon 1996). For example, the tendering stage could include forecasting the amount of capital required, the amount of interest that needs to be paid to support an overdraft, and the evaluation of different tendering strategies (Kaka 1996; Hwee, Tiong 2002). Furthermore, an accurate cash flow forecasting profile and understanding of the risk factors affecting the profile are very essential elements in each project (Hwee, Tiong 2002). Thus, a quick and simple technique is required to accurately estimate the financial requirements of a contractor contract (Kaka, Price 1991). Most construction projects are plagued by risks. Additionally, when the risks of the projects are not considered, it is not possible to deliver the objectives on time, within the budget, or with suitable quality. Although several methods were used to forecast cash flow, they generally fail to consider uncertainties, comprehensive identification of risk factors, interactional mechanisms between the factors, and simultaneous occurrence of the factors. Hence, it is necessary to directly consider uncertainties and risks to provide a more probabilistic cash flow, and this was not examined by previous studies. Moreover, conventional methods offered by extant cash flow research involve other problems. For example, statistical models that use algebraic formulations and polynomial regressions (S-curves) are often unable to provide precise predictions by using standard models, and each researcher has his/her own approach, which is often not realistic (Kenley 2003). Additionally, fuzzy system software products are complex from the perspective of end users (Zhao \textit{et al.} 2015) and involve computational complexity (Dikmen \textit{et al.} 2007). Furthermore, models based on artificial neural networks (ANN) are empirical and require the resolution of many methodological issues (Tu 1996). Specifically, connection weights are not easy
to understand, and it typically involves a network without any constraint conditions (Wang, Elhag 2007).

In contrast, Bayesian Belief Networks (BBNs) offer a method for the analysis of data and expert knowledge particularly in fields involving considerable uncertainty (Uusitalo 2007). This method is more convenient for modelling complexities as it describes cause-effect relationships among variables through graphical models (Kim et al. 2009). The BBNs method includes several advantages that distinguish this approach from other methods. First, the graphical display of the BBNs method helps gain insights into relationships between variables involved in the process (Kim et al. 2009; Wu 2010) and combines diverse types of evidence including both subjective beliefs and objective data (Fenton, Neil 2012). It provides considerable flexibility in its capacity to accept inputs and provide outputs (McCabe et al. 1998; Kim et al. 2009). Additionally, it offers predictions using experts’ judgments when historical data are insufficient (Kim et al. 2009). Furthermore, the BBNs method can be easily combined with decision analytic tools to aid management (Jensen, Nielson 2007; Uusitalo 2007) and can update its probability based on Bayes theory (Christian 2004).

The aim of the present study included producing probabilistic cash flow using the BBNs method by considering construction risk factors. Hence, the application of the proposed model can help contractors in making appropriate provisions at critical times to avoid a liquidity crisis. Additionally, the step-by-step description of the modelling process through a case study can help researchers in understanding the complex relationships between risk factors and the manner in which they affect cash flow management. The aforementioned points will be described in detail in the following sections.

1. Cash flow forecasting and risk factor consideration review

Risk management is the art and science of anticipating and planning for future uncertain events (Alarcon et al. 2010). The construction industry is potentially maximal plagued by risk (Tah, Carr 2000). Additionally, risk management is an important concept for every construction contractor (Jannadi, Almishari 2003), and objectives cannot be delivered on time, on budget, or with suitable quality results if risk management is not considered in the construction process. A reliable cash-flow prediction should take into consideration the effect of risk factors, but past research has seldom conducted risk analysis of project cash flows (Yu et al. 2017).

However, some studies have considered risk in cash flows. For example, Hwee and Tiong (2002) presented a computer-based model that had cash flow forecasting ability and examined the impact of five risk factors (duration, over and under measurement risk (during work progress), variation risk, and material cost variances) on a project’s cash flows. They used internal rate of return (IRR) and capital requirements to evaluate the cash flow performance of firms but did not consider comprehensive risk factors. Internal rate of return (IRR) and capital requirements were used to evaluate the performance of cash flows, but comprehensive risk factors related to cash flows could not be considered. Poh and Tah (2006) modelled risk impacts on construction tasks by developing an integrated duration–cost influence network. This network allowed for cash out forecasting by increasing the time and cost of a task. However, they investigated only major sources of risk, as opposed to a complete list of potential risk events, and did not fully explain the relationships among risks. Cui et al. (2010) used the system dynamics (SD) method to manage cash flows when uncertainty exists. They considered the dynamics and feedback of cash flows, which involved selecting six feedback loops in project cash flows. However, their model focused primarily on cash flow management strategies rather than forecasting, and it could not consider a complete list of potential risk events. El razeq et al. (2014) presented a method for net cash flow prediction that depended on applying risk factors affecting the cash flow process. Although they used primavera risk analysis to simulate probabilistic cash flows after applying risk factors to the cash flow S-curve, their use of Monte Carlo simulations was a problem, because they could not examine interactions between risk factors. Namazi et al. (2016) tested Jensen’s assertion by identifying the relative significance of the factors influencing free cash flow (FCF) risks via an artificial neural network (ANN). They used seven independent variables relating to FCF risks, including debt policy, ownership concentration, ownership level, managerial ownership, state ownership, firm size, and profitability. Their results showed that the most significant factors affecting FCF were profitability, debt policy, and firm size, in that order. They acknowledged that the model had limitations arising from the artificial neural networks techniques employed. Yu et al. (2017) proposed an algorithm to evaluate uncertainty in cash flows and overdraft requirements for projects that had overlapping activities with fuzzy duration. They conducted risk analysis for cash flows and overdrafts with the appropriate α segments. They found that including overlap tended to incur a higher amount of overdraft, while shortening the project completion time led to greater cash flow uncertainty.

As can be seen, past research has suggested risks and interactions related to cash flows, but it has been limited by the modelling methods used. To address these limitations, this study introduces the Bayesian belief networks (BBNs) method to describe a complete set of effective risk factors and their relationships with cash flows. In addition to the simplicity of the modelling, it is possible to produce more tangible probabilistic cash flows.

In general, the main cash flow forecasting methods that were used in extant studies can be divided into the following 5 categories: 1) Statistical models that use alge-
braic formulations and polynomial regressions; 2) Fuzzy Logic (FL) theory; 3) Artificial Neural Networks (ANN); 4) Weights of cost categories; and 5) System Dynamics (SD).

Kenley (2003) provided a comprehensive review on cash flow prediction models using algebraic formulations and polynomial regressions. Extensively used studies included Bromilow and Henderson (1977), Hudson (1978), Berdicevsky (1978), Peer (1982), Berny and Howes (1982), Tucker (1986, 1988), Kenley and Wilson (1989), Miskawi (1989), Khosrowshahi (1991), and Betts and Gunner (1993). A statistical model using algebraic formulations and polynomial regressions (S-curves method) is the simplest method to forecast project cash flow (Touran et al. 2004; Jarrah et al. 2007). However, it includes several important disadvantages as follows: most researchers experience considerable difficulty in predicting the future from standard models. It was not possible to identify standardized curves to represent cash flow forecast owing to the unique nature of the projects. Additionally, each researcher’s model involved its own approach and was not actually realistic (Kenley 2003). Furthermore, in this method, expert judgments could not be efficiently used, and sometimes the numerical coefficient included in a formula was not applicable to all projects.

Another way to forecast cash flow involves using fuzzy logic (FL) theory that is usually based on CPM (Critical Path Method) and PERT (Program Evaluation and Review Technique) (Prade 1979; Chanas, Kamburowska 1981; Maravas, Pantouvakis 2012). In CPM, the duration and cost of an activity are assumed as deterministic, although they are often uncertain in reality (Zamorri et al. 2009; Maravas, Pantouvakis 2012). Maravas and Pantouvakis (2012) developed fuzzy logic models to forecast project cash flow that considered intermediate possibility levels. However, fuzzy system software products are complex from the perspective of end users (Zhao et al. 2015) and involve computational complexity (Dikmen et al. 2007). Furthermore, the risks and their mutual interactions with each other were not considered in these models.

The use of Artificial Neural Networks (ANN) is another approach to forecast cash flow. Cheng and Tsai (2010) presented an evolutionary fuzzy hybrid neural network to achieve a performance that exceeded those of previous studies that used a singular, linear Neural Network (NN). However, the major disadvantages of ANN include the following: a) the model is empirical and includes several methodological issues that should be resolved (Tu 1996); b) designing an ANN and determining its parameters involves subjectivity, and thus different results could be achieved for the same problem (Wang, Elhag 2007); c) it is not easy to describe the connection weights involved (Wang, Elhag 2007); d) typically, ANN is a network without any constraint conditions (Wang, Elhag 2007). Alternatively, the cash flow forecasting model uses weights for cost categories. This method was pioneered by Fondahl and Bacarreza (1972) and Ashley and Teicholz (1977) (Park et al. 2005). They presented fixed weights of cost categories in which each cost element was assumed as a fixed percentage of total cost with respect to the project duration. Hence, Park et al. (2005) adopted moving weights of cost categories in a budget to solve problems in models adopted by previous studies. Their model considered time lags for the cost and earned value, and it could be used in the construction phase. Additionally, relative weights of the different cost categories were changed over the project duration. However, but the model had two disadvantages. The first disadvantage was that the model was extremely dependent on the planning of cost and earned value. The second disadvantage involved obtaining reliable variables at the jobsite level (Park et al. 2005). Furthermore, it should be noted that the risks and the interaction between the risks were not considered in this method.

System dynamics (SD) method is another technique that is used in cash flow forecasting. This method was first used by Cui et al. (2010) to manage cash flow. They considered the dynamics and feedbacks of cash flow that were not investigated in previous studies. This involved selecting six feedback loops in project cash flow, namely payment, owner’s retainage, subcontractor’s retainage, principal repayment, interest payment, and interest revenue loops. The system dynamics modelling consisted of three main modules including cash balance, material disbursement, and project operation modules. Likewise, three types of project cash flow management strategies were analysed in the model, namely, front-end loading, back-end loading, and optimal cash balance strategies. This model was introduced as a new approach to manage project cash flow. However, the model especially focused on cash flow management strategies instead of forecasting. Additionally, the model required customization of system parameters for a specific project, modification of the model equations, and also an unbounded software package (Cui et al. 2010). However, most of the fore-mentioned methods failed to consider comprehensive identifying risk factors, interactional mechanisms between these risk factors, and the simultaneous occurrence of these risk factors. Table 1 presents a brief summary of the features of each method.

BBNs are an increasingly popular method to model uncertain and complex domains and are used in data analysis and expert knowledge particularly in fields that are fraught with uncertainty (Uusitalo 2007). The BBN method includes several advantages that distinguish it from other methods. When compared to the ANN, Fuzzy logic (FL), and SD methods, the BBN method is more convenient to model complexities as it describes cause-effect relationships among variables through graphical models (Kim et al. 2009). It makes predictions when historical data are insufficient using experts’ judgments (Kim et al. 2009), and this cannot be performed by the two methods that use the weights of cost categories as well as statistical models using algebraic formulations.
and polynomial regressions. It can also update probabilities based on Bayes theory (Christian 2004). In contrast, statistical models use algebraic formulations and polynomial regressions that are not actually realistic owing to various approaches used by researchers (Kenley 2003). Unlike the ANN approach, BBN exhibits flexibility in its capacity to accept inputs, provides output (McCabe et al. 1998; Kim et al. 2009), overturns previous beliefs in the light of new findings, and combines diverse types of evidence including both subjective beliefs and objective data (Fenton, Neil 2012). It provides insights into relationships between variables in the process because of its graphical displays (Kim et al. 2009; Wu 2010), while other approaches (with the exception of the System dynamics method), do not account for relationships between variables. Finally, BBN can be easily combined with decision analytic tools to aid management (Jensen, Nielson 1998; Park et al. 2005). The calculations were complicated and time consuming when a lot of information is involved in the model.

2. Proposed modelling

This section first presents a brief description of the Bayesian Belief Networks (BBNs) method. This is followed by a description of the BBN-based cash flow forecasting model.

2.1. Bayesian Belief Networks (BBNs) method

BBNs are graphical representations of knowledge for reasoning under uncertainty and can be used at any stage of risk analysis (Sousa, Einstein 2012). This method has become increasingly popular in recent years to model complex systems. There are many methods for risk analysis such as artificial neural networks (ANNs), systems dynamics, and fuzzy systems. However, BBNs are an appropriate tool to represent risks as these methods are either too complicated to be used by practitioners or require abundant data (Cárdenas et al. 2014). Additionally, when compared with Monte Carlo simulations, BBNs can consider the interactions between each factor. A BBN is a directed acyclic graph (DAG) that is defined by a qualitative and a quantitative part. The qualitative part consists of a set of nodes that represent the system variables and their relationships. The quantitative part consists of probabilities and conditional probabilities that are used to make predictions about the system's behavior.

Table 1. Advantages and disadvantages of conventional methods of cash flow forecasting

<table>
<thead>
<tr>
<th>No.</th>
<th>Methods</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Statistical models that use algebraic formulations and polynomial regressions</td>
<td>It is the simplest method to forecast cash flow by using S-curves (Touran et al. 2004; Jarrah et al. 2007).</td>
<td>There is significant difficulty in predicting future from standard models, no standardized curves were identified to represent cash flow forecasts, and was not realistic as each researcher used their own approach (Kenley 2003).</td>
</tr>
<tr>
<td>2</td>
<td>Fuzzy Logic (FL) theory</td>
<td>It is possible to implement human knowledge and experience using natural language and linguistic terms (Sii et al. 2001). Additionally, intermediate possibility levels could be considered in the cash flow forecasting model (Maravas, Pantouvakis 2012).</td>
<td>Fuzzy system software products are complex from the perspective of end users (Zhao et al. 2015) and involve computational complexity (Dikmen et al. 2007). Furthermore, it is usually based on CPM and PERT and has problems (Prade 1979; Chanas, Kamburovski 1981; Maravas, Pantouvakis 2012), and the risks and their interactions were not considered.</td>
</tr>
<tr>
<td>3</td>
<td>Artificial Neural Networks (ANN)</td>
<td>It is not necessary to know the concrete functional relationship between outputs and inputs, and it is capable of modelling the data for multiple inputs and multiple outputs (Wang, Elhag 2007).</td>
<td>These models are empirical and many methodological issues remain to be resolved (Tu 1996). Another issue is the subjectivity in designing an ANN and determining its parameters, its connection weights cannot be explained easily, and it is usually a network without any constraint conditions (Wang, Elhag 2007).</td>
</tr>
<tr>
<td>4</td>
<td>Weights of cost categories</td>
<td>Park et al. (2005) resolved problems in previous models by considering time lags on the cost and earned value and changing relative weights of the different cost categories over the project duration. Additionally, it could be used in the construction phase (Park et al. 2005).</td>
<td>The model proposed by Park et al. (2005) is extremely dependent on the planning of cost and earned value and also involves obtaining reliable variables at the jobsite level (Park et al. 2005).</td>
</tr>
<tr>
<td>5</td>
<td>System Dynamics (SD)</td>
<td>The consideration of the dynamic nature of cash flow management was not investigated in previous studies (Cui et al. 2010).</td>
<td>Cui et al. (2010) presented the first system dynamics model of cash flow management. However, it was focused on cash flow management strategies as opposed to its forecasting. Furthermore, it required a customization of system parameters for a specific project, a need to modify model equations, and an unbounded software package (Cui et al. 2010). The calculations were complicated and time consuming when a lot of information is involved in the model.</td>
</tr>
</tbody>
</table>
a set of directed arcs between variables that represent the dependencies between the variables. The quantitative part consists of conditional probability distributions for each node with respect to the states of the influencing nodes (called as parent nodes) (Ale et al. 2014). Conditional probability tables (CPTs) associated with each node are used to define causal influence. Figure 1 shows an example of a simple Bayesian Network (BN).

In BBNs, it is possible to compute the probability of an effect (child node) on any variable in the model from the probability of a given cause (parent node) based on the Bayes theorem. Bayes’ rule can be expressed as follows:

\[
P(B|A) = \frac{P(A|B)P(B)}{P(A)}.
\]

where \(P(A)\) denotes the probability of \(A\), and \(P(A|B)\) denotes the probability of \(A\) given that \(B\) has occurred.

### 2.2. BBN-based cash flow forecasting model

The cash flow consists of two parts, namely cash outflow (cost) and cash inflow (income). The overall process of the probabilistic cash flow is a BBN-based cash flow forecasting model that is depicted in Figure 2. Initially, 8 parameters including contractor’s overhead, retention percentage, taxes and insurance, prepayment, defect liability period, the duration between the contractor statement that is submitted and paid, actual cost, and planned cost are determined as input data. This is followed by using input data to obtain actual and planned income. Hence, deterministic cash flow for planned and actual states is achieved by using cost and income. To generate risk-based (probabilistic) cash flow, a BBN model is applied with respect to the planned cost and income. In the BBN model, risk factors initially interact with each other and lead to increasing amounts of cost and time for each activity and final nodes are specified. Thus, probabilistic cost and income are determined, and this is followed by determining the risk-based (probabilistic) cash flow. Therefore, decision-making involves focusing on the probabilistic and deterministic cash flows.

#### 2.2.1. Cash outflow (Cost)

The risk modelling process to obtain the cash outflow for project activities involves the following four steps:

1. Risk identification;
2. Determination of cause-effect relationships for identified risks;
3. Allocation of conditional probability tables (CPTs) in BBN;
4. Determination of the increasing amount of time and cost of each activity and production of new Gantt charts (Using risk analysis with the application of a BBN).

#### 2.2.1.1. Risk identification

The following methods are usually used to identify risk factors: brainstorming, interviewing, exploiting Delphi technique, using standard checklists, asking an expert, diagramming techniques, SWOT analysis, and direct observation of risks in the project. In the present study, three methods, namely standard checklists, interviewing, and brainstorming, are used to identify risk factors. These risks are based on scientific articles that will be discussed in the next section by implementing the same with respect to a real case study.

#### 2.2.1.2. Determination of cause-effect relationships

In a Bayesian causal structure, a node demonstrates a risk factor, and an arrow between two risks indicates the relationship between them. In order to determine the structure of BBN for a project, cause–effect relationships among risk factors were determined through an expert survey. For this purpose, a questionnaire in a matrix form was prepared in which the list of risks in the left column represented the causes, and the same risks listed across the top represented the effects. The structure of the questionnaire was based on a study by Nasir et al. (2003). Following this, a number of experts with suitable knowledge and experience were asked to score the relationships such that 1 = No relationship; 2 = Weak relationship; 3 = Strong relationship; and 4 = Very strong relationship (Eshtehardian, Khodaverdi 2016). The cut-off value of 2 was then selected to ignore weak relationships. Hence, only the average causal relationships that exceeded 2 were retained, and the causal structure of the BBN was prepared.

#### 2.2.1.3. Allocation of conditional probability tables (CPTs) in BBN

A BBN is composed of two parts, namely the qualitative and quantitative parts. The completion of the qualitative phase and the formation of the causal Bayesian network structure are followed by the quantitative stage that begins with determining the dependencies between the nodes. These dependencies are probabilistic and are obtained by expert opinions that are termed as conditional probability tables (CPTs). Specifically, conditional probability tables are tables that represent the joint probability distribution of variables. For this purpose, the prior probabilities of risks with no parents were initially assigned, and this was followed by assigning the posterior probabilities of the child nodes due to the occurrence of the parent nodes. It should be mentioned that the all risks discussed in this study were considered with two binary state variables (yes and no).
2.2.1.4. Determination of the increasing amounts of time and cost of each activity and production of new Gantt charts (Using risk analysis with the application of a BBN)

To determine cash outflow, a Bayesian network is composed of three parts, namely the first nodes (risk factors), intermediate nodes and final nodes. Specifically, sets of risks (primary nodes) initially interact with each other and lead to several intermediate nodes. This is followed by considering the effects of intermediate nodes on the final nodes and the increasing cost, and duration of each activity. The increased cost percentage of each activity is influenced by the following three nodes: increased costs of material, equipment, and staff (as described in detail in Fig. 3). After assigning the conditional probability table to each node, the increased cost and time of each activity is calculated using Bayes theorem by considering all the risks and factors that affect it. Thus, the new Gantt chart is obtained. The forthcoming section describes the manner in which cash outflow modelling is implemented in a real case study in detail.

2.2.2. Cash inflow (Income)

The process of obtaining the new Gantt chart (determining the cost and time for new activities) is followed by performing income modelling (cash inflow) for each activity. Six parameters were assumed to determine cash in-
The cash inflow (income) was calculated based on the approach proposed by Ghoddousi (2006) as shown in Eqs (2) to (7) expressed as follows:

Payment in the first month (prepayment):

$$ = d\times h. \quad (2)$$

Payment in the second month:

$$ = g1 \times (a-c) - b\times(g1\times(a-c)) - (d\times h / 2). \quad (3)$$

Payment in the third month:

$$ = g2 \times (a-c) - b\times(g2\times(a-c)) - (d\times h / 2). \quad (4)$$

Payment in the fourth month:

$$ = g3 \times (a-c) - b\times(g3\times(a-c)). \quad (5)$$

Payment after completion of an activity:

$$ = gn \times (a-c) - b\times(gn\times(a-c)) + (0.05\times h). \quad (6)$$

Final payment to the contractor after the defect liability period:

$$ = (0.05\times h), \quad (7)$$

where: \(g_i\) denotes the contractor’s incurred cost in the \(i^{th}\) month, \(n\) denotes the duration of the activity, and \(h\) denotes the contract cost of the activity.

flow. The mentioned value for each parameter was determined based on the prevalent amounts in Iran as follows:

1. Contractors’ overhead: It was assumed as 30% of a contract amount \((a = 1.30)\).
2. Retention percentage: It was assumed was 10% of a contract amount that was deducted from all monthly statements \((b = 0.10)\). Half of this amount was paid after the temporary delivery and the other 5% was paid at the defect liability period.
3. The amount of taxes and insurance: They were assumed as 12.83% of the contract amount \((c = 0.1283)\).
4. Prepayment: It was determined as 20% of the contract amount that was deducted for each activity in the first two statements \((d = 0.20)\). It was paid according to the general conditions of contracts in Iran.
5. Defect liability period: It was determined as 12 months after the completion of construction and temporary delivery \((e = 12\) months).
6. The duration between the contractor statement was submitted and paid by the client: It was assumed as 30 days \((f = 30\) days).
2.2.3. Probabilistic cash flow

The cash flow diagram represents cost and income curves simultaneously. A net cash flow diagram is achieved by subtracting costs from income. Thus, a cash flow diagram is obtained by using two charts of cost and income. Figure 4 shows typical diagrams of cash flow and net cash flow.

In the next section, the proposed probabilistic cash flow model is implemented in a real case study to illustrate the manner in which the methodology can be put into practice.

3. The case study

In this section, the methodology is implemented in a step-by-step manner in accordance with section 2.2 in an important construction project in Iran to provide a better understanding of the proposed model.

3.1. Introduction

The case study involved the service complex building in Pardis Technology Park (PTP) that belongs to the Iranian presidency. It is located near Pardis city in the Tehran province. The land area and building area of the project corresponded to 5046 square meters and 12200 square meters, respectively. The project consisted of five continuous blocks with the following facilities: hotels, restaurants, a sports complex, a fountain entrance, commercial shops, a kindergarten, a mosque, and a health clinic. This project was implemented to provide complete services to all the companies in the Technology Park that included knowledge-based companies and elite firms across the country. An excavation of approximately 7000 cubic meters was performed. Concrete works including foundations, retaining walls, and metal deck roofing corresponded to approximately 4000 cubic meters. The steel structure weighed approximately 1200 tons, and the metal deck roofing area included an area of 10500 square meters. The contract type corresponded to CM-at Risk with an original duration of 24 months and a total estimated cost of 9205 thousand dollars. The start date of the project was 6 July, 2014.

This case study was selected owing to operational complexity and sensitivity (as it was part of the presidential administration institution). In the present study, only the following three main structural activities were evaluated:
1. Excavation, foundation, and retaining walls;
2. Steel structure;
3. Metal deck roofing.

3.2. Procedure of the BBN-based cash flow forecasting model

3.2.1. Cash outflow

3.2.1.1. Risk identification

Risk identification is extremely important for every project. The types of risk elements vary with the types of construction projects (Paek et al. 1993). Therefore, it is not possible to achieve appropriate results if risk identification is not performed correctly. In this project, to identify risk factors in accordance with Section 2.2.1.1, a hybrid of three risk identification methods, namely standard checklists, interviewing, and brainstorming sessions, was applied. Due to the similarity of the case study (in terms of country and social status) with the research of Eshtehardian and Khodaverdi (2016), six experts with over 15 years of executive experience in construction projects were selected. Figure 5 lists expert specifications based on education and job experience. Following this, 37 risk factors were identified (primary and intermediate nodes) in the BBN. From the 37 risk factors, 10 risk factors were adopted from a study by Choudhry et al. (2014), six risk factors were adopted from a study by Dikmen et al. (2007), five risk factors were adopted from a study by Jarkas and Haupt (2015), five risk factors were adopted from a study by Kim et al. (2009), four risk factors were adopted from a study by Alireza et al. (2013), one risk factor was adopted from a study by Mahamid (2013), and six risk factors were incorporated from the expert judgments. The aforementioned studies were selected given that the countries examined in these studies (Pakistan, Qatar, and Palestine) were similar to Iran. Finally, the risk factors were divided into the following seven categories: financial, contractual, design, safety and health, management, construction, and external risk factors. Table 2 shows the identified risk factors with the extant studies relevant to the case study.

Fig. 4. (a) Contractor’s cash flow; (b) Contractor’s net cash flow (De Marco 2011)

Fig. 5. Expert specifications: (a) Experts’ education; (b) Experts’ job experience
3.2.1.2. Determination of cause-effect relationships

The Bayesian network structure was obtained when the cause-effect relationships were formed. For this purpose, expert judgments were used (expert specifications were shown in Fig. 5) in accordance with Section 2.2.1.2. Hence, a 37*37 matrix based on a study by Nasir et al. (2003) was used. Finally, the causal Bayesian structure with 31 risk factors (primary nodes), 6 intermediate nodes (P1 to P6), and final nodes were obtained as shown in Figure 6.

3.2.1.3. Allocation of conditional probability tables (CPTs) in BBN

As mentioned in Section 2.2.1.3, the prior probabilities of risk factors with no parent were initially determined. This was followed by assigning child nodes based on the posterior probabilities for the occurrence of parent nodes using expert judgments. Hence, the six aforementioned experts were asked to participate in a joint meeting to assign conditional probability tables for each risk factor using the brainstorming method.

---

Table 2. Identified risk factors for the case study

<table>
<thead>
<tr>
<th>Risk category</th>
<th>Risk factor</th>
<th>Source of data</th>
<th>Risk symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial risks</td>
<td>Inflation</td>
<td>Alireza et al. (2013)</td>
<td>R1</td>
</tr>
<tr>
<td></td>
<td>Delay in payment process by the client</td>
<td>Jarkas and Haupt (2015)</td>
<td>R2</td>
</tr>
<tr>
<td></td>
<td>Financial difficulties of contractor</td>
<td>Kim et al. (2009)</td>
<td>R3</td>
</tr>
<tr>
<td></td>
<td>Economic disaster</td>
<td>Choudhry et al. (2014)</td>
<td>R4</td>
</tr>
<tr>
<td></td>
<td>Financial difficulties of owner</td>
<td>Kim et al. (2009)</td>
<td>R5</td>
</tr>
<tr>
<td></td>
<td>Escalation of material prices</td>
<td>Choudhry et al. (2014)</td>
<td>R6</td>
</tr>
<tr>
<td>Contractual risks</td>
<td>Contract specific problems</td>
<td>Yildiz et al. (2014)</td>
<td>R7</td>
</tr>
<tr>
<td></td>
<td>Disputes and claims</td>
<td>Choudhry et al. (2014)</td>
<td>P3</td>
</tr>
<tr>
<td>Design risks</td>
<td>Design changes</td>
<td>Choudhry et al. (2014)</td>
<td>R8</td>
</tr>
<tr>
<td></td>
<td>Errors and omissions in design drawings</td>
<td>Jarkas and Haupt (2015)</td>
<td>R9</td>
</tr>
<tr>
<td></td>
<td>Project complexity</td>
<td>Yildiz et al. (2014)</td>
<td>R10</td>
</tr>
<tr>
<td></td>
<td>Design problems</td>
<td>Yildiz et al. (2014)</td>
<td>R11</td>
</tr>
<tr>
<td>Safety and health risks</td>
<td>Force majeure</td>
<td>Alireza et al. (2013)</td>
<td>R12</td>
</tr>
<tr>
<td></td>
<td>Failure to implement safety measures</td>
<td>–</td>
<td>R13</td>
</tr>
<tr>
<td>Management risks</td>
<td>Lack of coordination between construction parties</td>
<td>Mahamid (2013)</td>
<td>R14</td>
</tr>
<tr>
<td></td>
<td>Incentive policies</td>
<td>–</td>
<td>R15</td>
</tr>
<tr>
<td></td>
<td>Increase staff motivation</td>
<td>–</td>
<td>R16</td>
</tr>
<tr>
<td></td>
<td>Poor site management and supervision (by consultant)</td>
<td>Choudhry et al. (2014)</td>
<td>R17</td>
</tr>
<tr>
<td></td>
<td>Incompetent subcontractors</td>
<td>Choudhry et al. (2014)</td>
<td>R18</td>
</tr>
<tr>
<td></td>
<td>Client’s incompetency</td>
<td>Yildiz et al. (2014)</td>
<td>R19</td>
</tr>
<tr>
<td>External risks</td>
<td>Political instability</td>
<td>Choudhry et al. (2014)</td>
<td>R20</td>
</tr>
<tr>
<td></td>
<td>Restricted access to the site</td>
<td>Jarkas and Haupt (2015)</td>
<td>R21</td>
</tr>
<tr>
<td>Construction risks</td>
<td>Inappropriate construction methods</td>
<td>Kim et al. (2009)</td>
<td>R22</td>
</tr>
<tr>
<td></td>
<td>Weather</td>
<td>Alireza et al. (2013)</td>
<td>R23</td>
</tr>
<tr>
<td></td>
<td>Lack of working space</td>
<td>Choudhry et al. (2014)</td>
<td>R24</td>
</tr>
<tr>
<td></td>
<td>Shortage of equipment</td>
<td>Kim et al. (2009)</td>
<td>R25</td>
</tr>
<tr>
<td></td>
<td>Difficulties in preparing access road</td>
<td>–</td>
<td>R26</td>
</tr>
<tr>
<td></td>
<td>Equipment damage</td>
<td>Choudhry et al. (2014)</td>
<td>R27</td>
</tr>
<tr>
<td></td>
<td>Accelerate to the project execution</td>
<td>–</td>
<td>R28</td>
</tr>
<tr>
<td></td>
<td>Failure to meet performance criteria</td>
<td>Alireza et al. (2013)</td>
<td>R29</td>
</tr>
<tr>
<td></td>
<td>Contractor’s lack of experience</td>
<td>Yildiz et al. (2014)</td>
<td>R30</td>
</tr>
<tr>
<td></td>
<td>Decrease in quality of work</td>
<td>Yildiz et al. (2014)</td>
<td>R31</td>
</tr>
<tr>
<td></td>
<td>Poor labour productivity</td>
<td>Jarkas and Haupt (2015)</td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td>Shortage in specified materials</td>
<td>Jarkas and Haupt (2015)</td>
<td>P2</td>
</tr>
<tr>
<td></td>
<td>Construction delays (slow speed in work execution)</td>
<td>Choudhry et al. (2014)</td>
<td>P4</td>
</tr>
<tr>
<td></td>
<td>Poor equipment productivity</td>
<td>–</td>
<td>P5</td>
</tr>
<tr>
<td></td>
<td>Defective works and reworks</td>
<td>Kim et al. (2009)</td>
<td>P6</td>
</tr>
</tbody>
</table>
Fig. 6. Bayesian causal networks of the project as a result of the evaluation of cause-effect relationships

Fig. 7. Final risk analysis of BBN networks to determine the increasing amounts of time and cost involved in each activity
3.2.1.4. Determination of the increasing amounts of time and cost of each activity and production of new Gantt charts (Using risk analysis with the application of a BBN)

The formation of cause-effect relationships and assignment CPTs was followed by the application of the Bayesian formulations described in section 2.1 to the causal networks using Netica software (2010). Thus, the occurrence probability of each node was calculated. Consequently, the increases in the amounts of time and cost for each activity were obtained using risk analysis of BBNs as shown in Figure 7. Table 3 shows the increases in the amounts of time and cost for each activity in the actual and probabilistic states.

Based on Table 3, the highest probabilities for the cost and duration intervals were considered as probabilistic results. Thus, percentage increases in the cost and duration for the excavation, foundation, and retaining wall activity were obtained as 11–20% and >60%, respectively. With respect to the steel structure activity, the percentage increase in the cost and duration corresponded to 0–10% and 31–60%, respectively. Additionally, the percentage increase in the cost and duration for the metal deck roofing activity corresponded to 0–10% and 0–30%, respectively. Hence, the probabilistic results indicated compliance with the actual data. Next, the sequence of activities for probabilistic Gantt charts in accordance with the planned chart was considered to create Gantt charts. Thus, the start date of the steel structure activity was considered as a milestone at the beginning of the third month. Additionally, the start date of the metal deck roofing activity was considered after the completion of the steel structure operation. Accordingly, the probabilistic Gantt charts consisted of two parts, namely limit 1 and limit 2, wherein the acceptable ranges were between the two limits. The charts for limit 2 and limit 1 were composed based on the least and maximum increases, respectively, in the amounts of time and cost for each activity in Table 3. All Gantt charts including probabilistic (limit 1 and limit 2), planned, and actual charts are shown in Figures 8 to 11.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Increasing amounts of cost and duration (%)</th>
<th>Probabilistic</th>
<th>Actual increasing amount (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cost and Duration intervals (%)</td>
<td>Probability of the interval (%)</td>
<td></td>
</tr>
<tr>
<td>Excavation, foundation, and retaining wall</td>
<td>0–10</td>
<td>19.00</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>11–20</td>
<td>68.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;20</td>
<td>12.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0–30</td>
<td>9.20</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>31–60</td>
<td>26.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;60</td>
<td>64.80</td>
<td></td>
</tr>
<tr>
<td>Steel structure</td>
<td>Cost</td>
<td>0–10</td>
<td>66.70</td>
</tr>
<tr>
<td></td>
<td>11–20</td>
<td>28.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;20</td>
<td>5.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0–30</td>
<td>6.40</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>31–60</td>
<td>58.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;60</td>
<td>35.30</td>
<td></td>
</tr>
<tr>
<td>Metal deck roofing</td>
<td>Cost</td>
<td>0–10</td>
<td>80.10</td>
</tr>
<tr>
<td></td>
<td>11–20</td>
<td>16.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;20</td>
<td>3.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0–30</td>
<td>52.60</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>31–60</td>
<td>29.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;60</td>
<td>17.90</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 8. Planned Gantt chart

Fig. 9. Probabilistic Gantt chart (Limit 1)

Fig. 10. Probabilistic Gantt chart (Limit 2)

Fig. 11. Actual Gantt chart
3.2.2 Cash inflow

Cash Inflow was obtained using Gantt charts and the formulae presented in Section 2.2.2. A Microsoft Excel spreadsheet was used for this purpose.

3.2.3 Cash flow

According to the Section 2.2.3, a net cash flow diagram was obtained by subtracting cost (cash outflow) from income (cash inflow). Figures 12, 13, and 14 show cash inflow, cash outflow, and net cash flow diagrams, respectively, for the three states of probabilistic (limit 1 & limit 2), actual, and planned values.

4. Discussion

Based on Figures 12 to 14, the following findings were obtained. The risks in the planned curves were not considered, and thus the results were very crude at best. Hence, more reliable intervals for decision-making were obtained by applying risks on the model.

The final actual cost corresponded to USD 1648000 and was located between USD 1558000 and 1711000 (limit 1 and limit 2). Additionally, the last actual incurred cost was in the thirteenth month. However, the probabilistic state predicted that it was in the period between the twelfth and fourteenth months. Therefore, the actual cost curve indicated favourable convergence with probabilistic charts (Fig. 12).

A detailed examination of Figure 13 indicated that the deviation of probabilistic cash inflow charts from the actual chart was low. It only occurred between the tenth to twelfth months owing to the lack of liberalization with respect to the retention of the subcontractors (after the temporary delivery). The subcontractor had defective works following this. However, the final actual payment (USD 1725000) was evidently very close to the range of probabilistic values (USD 1734000 to 1904000).

As indicated by Figure 14, the probabilistic net cash flow curves exhibited a favourable convergence with the actual curve. The worst actual incurred cost corresponded to USD 273000 in the fourth month in which it was clearly located in the range of USD 217000 and 291000 (among probabilistic values). Conversely, in the planned curve, the estimated cost corresponded to USD 147000, and this was approximately half the actual value. Furthermore, the predicted most critical time corresponded to the period between the third and fifth months in the probabilistic state, and this was fully compatible with reality. In contrast, in the planned curve, the most critical time was considered to occur between the fifth and seventh months, and thus this result could be misleading.

The lack of full compliance probabilistic curves with actual data that occurred from the eleventh month onward was due to the lack of liberalization of the subcontractor’s retention. Therefore, the lack of retention liberalization (after the temporary delivery) was considered as a new risk in the BBN structure to accommodate the probabilistic curves with reality, and modelling was performed again. It was then observed that the actual amounts were properly between the probabilistic limits (limit 1 and 2), as shown in Figure 15.

Therefore, the comparison between probabilistic and actual states with respect to the planned value validated the suitability of the proposed methodology.
The importance of and the need for cash-flow management is a matter of consensus among researchers and practitioners. Cash is the most important resource for construction companies. Hence, cash flow forecasting helps contractors to select the contracts that can be financed by the available resources and can be beneficial to make provisions during difficult periods to avoid a liquidity crisis. The construction industry is debatably the industry that is most plagued by risk, and the success of construction projects usually depends on the consideration of all risks involved. However, conventional methods presented in previous studies fail to consider the comprehensive identification of risk factors, interactional mechanisms between these risk factors, and their simultaneous occurrence. Hence, this study investigated the application of BBNs for the first time to reduce the aforementioned issues by generating probabilistic cash flow. The results of this study include the following:

- A case study was considered that involved the assessment of three main structural activities of an important construction project in Iran with respect to the relevant risk factors. The evaluation of cause-effect relationships between the risk factors led to the development of the project’s BBNs. Thus, the increasing percentages of cost and duration for each activity that converged with the actual data were obtained. Increasing interval amounts of cost for the excavation, foundation and retaining wall activity, steel structure activity, and metal deck roofing activity corresponded to 11–20%, 0–10%, and 0–10%, respectively. Additionally, increasing interval amounts of duration for the excavation, foundation, and retaining wall activity, steel structure activity, and metal deck roofing activity were corresponded to > 60%, 31–60%, and > 60%, respectively.
- The increases in the amounts of cost and duration of each activity were determined to obtain probabilistic cash outflow charts (cost) with two limits (limit 1 and limit 2). The acceptable ranges were between the two limits. The determination of the input data and usage of the presented formulae were followed by the prediction of cash inflow charts (income). Hence, probabilistic net cash flow charts were obtained by subtracting cost from income.
- The worst actual incurred cost corresponded to USD 273000 in the fourth month, and this was clearly located in the range of USD 217000 to 291000 (among the probabilistic limits). The most critical time ranges in the planned curve were located between the fifth and seventh month. Nevertheless, it was anticipated that this range would be located between the third and fifth months, and this was entirely consistent with reality.
- The lack of full compliance probabilistic curves with actual data that occurred from the eleventh month was due to the lack of liberalization of subcontractor’s retention. Hence, the lack of retention liberalization (after the temporary delivery) was considered as a new risk in the BBN structure, and the modelling was performed again. It was then observed that the actual amounts were properly between the probabilistic limits.

Thus, the results showed the accuracy of the methodology and its favourable convergence with realistic conditions. This study was able to illustrate how to impose a complete set of risks and their interactions on cash flows. The simplicity of the modelling was possible to produce more tangible probabilistic cash flows. Additionally, the causal Bayesian Network represented the most important nodes (risks), increasing duration and cost of each activity. Then, probabilistic cash outflows were calculated. Using the input data and developed mathematical formulations, the probabilistic cash inflows and, finally, probabilistic net cash flows were obtained. The proposed approach could lead to proper management of projects (in terms of time and cost) and updates across the entire project lifecycle.

In future, a similar study could be performed for specific types of construction projects, e.g., utility projects, highway and dam construction projects. Another approach would be to replace the discrete model used in this study with the continuous BBNs model to better illustrate causal BBNs.

**References**


Ghoddousi, P. 2006. Planning and control of civil engineering projects. Tehran: Iran University of Science and Technology Press.


Mostafa KHANZADI. Is an Associate Professor of Civil Engineering at Iran University of Science & Technology, Iran. His main research interests include construction project management, risk management, sustainable development and concrete technology. He has published over 100 technical articles.

Ehsan ESHTEHARDIAN. Is an Assistant Professor at Tarbiat Modares University, Iran. He is a professional engineer in project control and planning with more than 10 years of experience in the construction sector. His main research interests include project management, risk management, ICT in construction (artificial intelligence techniques, decision support system) and building information modelling.

Mahdiyar MOKHLESPOUR ESFAHANI. He received his Master of Engineering degree from Iran University of Science & Technology. He is a Project Planner & Controller and Civil Engineer at the City of Tehran, Iran. His research interests include risk management, project cost estimation, sustainable development and artificial intelligence techniques.