Case Study

Project Schedule Forecasting for Skyscrapers

Yongkui Li[1]; Keyu Lu[2]; and Yujie Lu, A.M.ASCE[3]

Abstract: An accurate schedule estimation is critical for megaprojects such as skyscrapers, which have features such as high economic values and wide societal awareness. An inaccurate schedule may have serious consequences such as going seriously overbudget or developing a poor public image. However, the accuracy of using traditional methods to predict a skyscraper's schedule is low because of the limitation of available information at the early stage, large uncertainties, complex influencing factors, and their coupling effects on the project schedule. To improve the accuracy of schedule estimation, this paper establishes a revised case-based reasoning (CBR) model to estimate the schedule for skyscrapers. The CBR model comprises three steps, including (1) identify seven key influencing factors, (2) retrieve and rank candidate cases according to their similarity to the target case, and (3) revise selected cases based on multiple regression analysis (MRA). The model was then tested by using 33 skyscrapers in China from the last decade. The result shows that the estimated error in this model (4.83%) is significantly lower than that in the traditional models (9–33%). This result justifies the application of extending CBR in the estimation of project schedule for megaprojects. It also provides a reliable and accurate scheduling tool to help owners better allocate and manage resources in the early stage of a megaproject. **DOI:** 10.1061/(ASCE)ME.1943-5479.0000498. *© 2016 American Society of Civil Engineers.*

Author keywords: Case-based reasoning (CBR); Schedule estimation; Skyscrapers; Project schedule; Case revision.

Introduction

The United Nations predicts that the world's population is expected to increase by 30% by 2050 and 75% of the population will live in cities (Merrill and Gray 2012). Such a rapid population explosion demands innovative ways to supply more living and working spaces. Skyscrapers, also called super-high-rise buildings, have become an important alternative in the urbanization process to increase vertical space and to accommodate more people. Therefore, it has become a favorite way in the last decades of accommodating a rapid expansion of population. Skyscrapers have a variety of definitions in terms of their height. For instance, 165 m was the definition in the United States (CTBUH 2011), compared to 100 m in China (NSPRC 2000). This paper adopts the international standard given by the Council on Tall Buildings and Urban Habitat (CTBUH) and defines the height of skyscrapers as building at or above 200 m.

During last decades, skyscrapers have boomed around the world, especially in megacities. During 1960 and 2000, the total number of skyscrapers (completed) reached 224. This number had grown nearly four times and to hit 826 during 2000–2015. China has the largest number of skyscrapers with the quickest growth rate in the world. As shown in Fig. 1, the number of China's completed skyscrapers has increased during last 7 years except the years 2009 and 2012 (Merrill and Gray 2012). In 2013 and 2014, the number of skyscrapers in China exceeded that in the rest of world (ROW) (CTBUH 2015). Following by this trend, the boom of skyscrapers will continue in China. By the end of 2016, the total number of skyscrapers will reach 800, four times more than those in the United States (Le and Li 2013). Therefore, China was selected as the context of this study and its result can provide significant implications for over a half of the skyscrapers existing worldwide.

Schedule estimation is a significant factor in determining the feasibility of a project (Jin et al. 2014), especially for one with huge investment and great strategic significance. Meanwhile, schedule estimation also serves as crucial evidence to allocate and control project resources. An accurate schedule can help better distribute construction resources, such as money, machinery, and materials, in a more efficient way. For example, a project team can arrange the project finance and cashflow in advance according to the project schedule. On the other hand, skyscrapers are commonly regarded as a symbol of urban development and economic status, so a misjudged schedule of skyscrapers may have an influence beyond the project itself to extend to its neighboring zones or overall society. For example, a skyscraper is a landmark in a city, promoting the intensity of land utilization and attracting more enterprises to the zone. Therefore, completion of a skyscraper may influence the process of constructing surrounding facilities and infrastructures such as an underground railway, or the city image.

However, project owners are unable to accurately estimate a skyscraper's schedule using existing methods. Possible reasons that challenge the schedule estimation include four aspects:

- Complex and compounding factors that constantly change the project schedule, such as uncertain subsurface geotechnical conditions that affect the deep foundation of the skyscraper, or weather and climate conditions that affect crane operations and facade installation at the high level;
- Deviation between prevailing assumptions that construct the schedule estimation and the project actual situation. In the early stage of a project, only limited information is available (Merrill and Gray 2012). As a result, schedule planners need to make their own assumptions that unfortunately are shown to be different from the real scenario happening on the construction site,
- Extremely high and varying risks mean that compared with traditional projects, the construction process of skyscrapers has

[1]Professor, School of Economics and Management, Tongji Univ., 1239 Siping Rd., Shanghai 200092, China. E-mail: y.k.lee@126.com

[2]Research Assistant, School of Economics and Management, Tongji Univ., Shanghai 200092, China. E-mail: 393495541@qq.com

[3]Assistant Professor, Dept. of Building, School of Design and Environment, 4 Architecture Dr., National Univ. of Singapore, Singapore 117566 (corresponding author). ORCID: http://orcid.org/0000-0002-9585-0192. E-mail: luy@nus.edu.sg

Note. This manuscript was submitted on February 5, 2016; approved on September 6, 2016; published online on November 7, 2016. Discussion period open until April 7, 2017; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Management in Engineering***, © ASCE, ISSN 0742-597X.

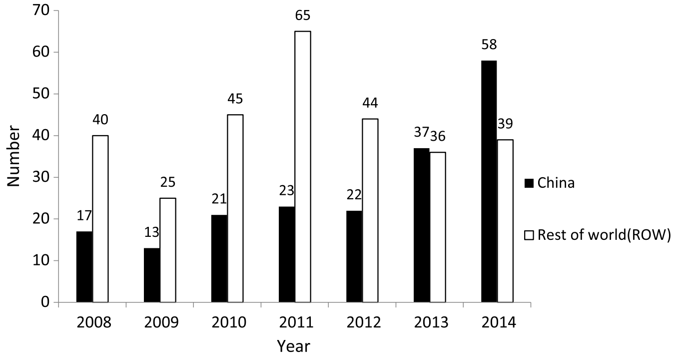

Fig. 1. Number of completed skyscrapers in China and ROW

more risks, such as long-term financing risks, macroeconomic risks, and political risks; and

- The correlation and coupling effects intertwined among the preceding factors makes the schedule estimation of skyscrapers even more complex and unpredictable.

These mentioned factors reduce the accuracy of a project's schedule. Consequently, an inaccurate estimation may affect a project's comprehensive performance, particularly in weakening the control of project schedule, quality, and investment objectives while increasing the risks of dispute among project stakeholders (Jin et al. 2016). Therefore, in order to avoid these problems, this paper aims to develop a case-based reasoning (CBR) model that is capable of predicting a skyscraper's schedule at an early stage.

The study is organized as follows. The next section reviews current methods of project schedule estimation, schedule influencing factors, and case-based reasoning, followed by the section that introduces the framework of CBR model, including database establishment, case retrieval, and case revision. Thereafter, the results, validation, and discussion of the CBR model are presented. The last section concludes the study and highlights the knowledge contribution.

Literature Review

Existing Methods and Influencing Factors to Estimate Project Schedules

An accurate schedule in the early stage of a construction project is significant for the project's success (Koo et al. 2010). Traditional estimating methods have been developed for the long-term and include the critical path method (CPM), program evaluation and review technique (PERT), bar charts, and line of balance (LOB) (Yamin and Harmelink 2001). Each tool has its own characteristics and applicable scope. For instance, LOB is typically used for repetitive activities, and CPM is suited for a project with complex relationships (Yamin and Harmelink 2001). The data sources for the preceding methods are commonly based on expert experiences and estimation (Kim and Kang 2004), or data analytics from historical data (Lin et al. 2011). In China, the quota method is widely applied. That method calculates the average duration of projects based on certain production technology and natural conditions. However, these methods have not considered project risk factors.

Researchers have also considered risk factors such as weather, labor productivity, soil conditions, and speed of transportation (Lu et al. 2014), and put forward uncertainty-based scheduling tools, which include probabilistic network evaluation technique (PNET), narrow reliability bounds (NRB), Monte Carlo simulation (MCS), and so on (Okmen and Oztas 2008). These methods can provide a reliable estimation range for a project schedule, but they neglect the relevance and influence between two connecting activities (Wang and Demsetz 2000). Recently, scholars established estimation models that can combine project scheduling with advanced data-mining and information systems. For example, Kim et al. (2014) built a knowledge-based information system to estimate the cost and schedule of projects. In that study, the system calculated the project schedule based on mathematical modeling between cost, schedule, labor productivity, and supply of material. In addition, Hong et al. (2011) developed a simulation-based schedule estimation model for core wall construction. The proposed model only considered labor and equipment resources, but did not consider other factors such as subsurface conditions. However, those aforementioned methods have a low accuracy of estimating the schedule of skyscrapers because of this building type's structural complexity and the restricted information available in the front and planning stages.

In addition to project estimation methods, identifying factors that influence a project's schedule is equally vital to determine a project's schedule. So far, many scholars have studied influencing factors for project scheduling from different perspectives. Assaf et al. (1995) outlined 73 causes of delay, such as change orders, finance crises, and legal disputes, which were determined based on a questionnaire carried out to analyze their severity, importance, and frequency. For example, owners thought the shortage of labor as the first important delay factor, while contractors considered the delay in progress payments by owners as the first important factors. Sambasivan and Soon (2007) identified and ranked influencing factors according to Malaysian projects as follows: improper planning, contractor's poor site management, inadequate client's finance and payments for completed work, problems with subcontractors, labor supply, lack of communication among stakeholders, and so on. Kaming et al. (2010) undertook a questionnaire and factor-analysis technique to analyze the factors of schedule delay and cost overrun. The result showed that labor productivity, design change, weather, and inadequate planning were predominant causes. Ogunlana and Promkuntong (1996) concluded that the main reason for schedule delays was human factors, including communication effectiveness among stakeholders and the qualification of laborers. In addition, the supply of construction materials could also be an acute factor, especially during the boom of construction and real estate such as during the years 1988 and 1992. Sidwell (1984) recognized that schedule management was affected mostly by the level of project management. In terms of project lifecycle, Mulholland and Christian (1999) divided schedule risks into four parts: design, procedure, construction, and project management. Based on this framework, Mulholland and Christian (1999) refined and elaborated each of these dimensions. For example, design risks included the experience of designers and design change. Procedure risks included the selection of and speed of suppliers. Construction risks included the pattern of contracts.

These mentioned studies have promoted research on project schedules for typical buildings and infrastructure projects. However, three limitations exist when applying them for skyscrapers: (1) most of these methods have only considered several influencing factors in the schedule estimation, such as material supply, but have not considered the complexity of skyscrapers, such as the structural challenges, subsurface conditions, and so on; ignorance of such influencing factors makes the schedule unrealistic; (2) past methods are primarily based on the individual project and their internal procedure sequences, but ignore the comparison with similar projects. For instance, the schedule of the targeted project can refer to the schedule from a similar one that has equivalent area, floor, and types of structure, and so on; and (3) past methods estimate the

J. Manage. Eng., 2017, 33(3): 05016023

total project schedule using the bottom-up approach by aggregating the project schedule from individual working packages. However, skyscrapers that have countless and interwoven working packages are difficult to be broken down into calculative units and, therefore the conventional methods are not applicable.

To overcome these limitations, project scheduling for skyscrapers needs to consider more influencing factors and past experiences of similar cases from the project's overall perspective rather than individual components. However, for any skyscraper, it is difficult to gain enough information and expertise to produce the project schedule that meets the requirements. Therefore, this study provides a model that can predict the schedule estimation method based on the complexity of skyscrapers and the experiences from similar cases.

Case-Based Reasoning and Its Application to Project Scheduling

Case-based reasoning (CBR) is a continuous dynamic learning process that originated from artificial intelligence. It solved new problems by reusing experience of similar previous cases. CBR was first conceived in 1982 by Schank (1983) in his book *Dynamic memory*. In 1983, Kolodner (1983) realized the aforementioned idea in the proposed CYRUS system, which was regarded as the first system of CBR (Aamodt and Plaza 1994). In 1985, Kolodner first proposed the jargon CBR into the publication, debuting the foundation of the academic discussion of CBR.

The process of CBR is similar to a human's thinking pattern; that is, when people face new problems, they will automatically recall historical experiences, amend past solutions, and create new solutions. In 1994, Aamodt and Plaza regulated the four phases of CBR model—retrieve, reuse, revise, retain, also called the CBR-cycle (4R) (Aamodt and Plaza 1994). The detailed process is as follows: (1) retrieve the cases according to their similarity to the target case, (2) reuse the solutions of retrieved cases to deal with the target cases, (3) if deviations exist between retrieved cases and the target cases, revise retrieved cases to create a new solution, and (4) retain the new cases and solution in the database for future use (Kim and Kim 2010).

CBR has been commonly used in many disciplines to help identify suitable cases and provide decision-making support. For instance, CBR is most applicable in a field that has high similarity and rich experiences but ambiguous rules, such as the medical domain (Ping et al. 2015), fault diagnostics (Lin et al. 2009), intelligent decision support systems (Koo et al. 2010), and electronic information services (Thomasson et al. 2006). Particularly, CBR has also been utilized as a decision-making tool in the construction field in order to utilize the experience from previous projects when planning a new project (Jin et al. 2014). For instance, Duverlie and Castelain (1999) applied CBR to the cost estimation in the project design stage. In 2001, Chua et al. (2001) developed a decision-making model to provide bidding suggestions for contractors in different situations. In 2006, Ozorhon et al. (2006) established a CBR model to demonstrate how experiences of competitors in international markets may be used by contractors in order to support international market selection decision. In 2011, Goh and Chua (2010) applied CBR to identify potential construction hazards utilizing past experience in the form of past hazards.

However, only a few studies have applied CBR in project schedule estimation. Jin et al. (2016) established a CBR model to estimate the preliminary duration of multihousing projects in Korea. The model chose the k-nearest neighbor (KNN) method to retrieve cases and revised the cases by regression-based algorithms. The test showed an accurate estimation with the error rate of 5.73%, which was better than the traditional one of 9.15% (Jin et al. 2016). That study initiated a milestone for CBR-based project scheduling estimation but has two limitations when applying it to the skyscrapers. First, the selected cases of the study are relatively simple and small scale, so the CBR process and conclusions may not fully replicable for complex projects, such as skyscrapers. Second, the selected cases were limited to residential buildings, and thus did not include other types and functions of buildings, such as commercial buildings, offices, and hotels. Therefore, this present study seeks to improve the model and overcome these limitations by proposing a revised CBR model that can estimate the schedule for large-scale and multifunctional skyscrapers at an early stage.

Framework of the CBR Schedule Forecasting Model

The CBR-based schedule forecasting model comprises of three steps: data preprocess, case retrieval and ranking, and case revision, as shown in Fig. 2. The data preprocess stage aims to identify the most important influencing factors. The case retrieval stage is designed to retrieve the cases with high similarity to the target case. Case retrieval is the key process in the CBR cycle as its result lays the foundation of the following processes (Ji et al. 2010). Using different retrieval methods can significantly impact the performance of CBR results. So far, three approaches have been widely used, including KNN, inductive retrieved method, and knowledge-based retrieval methods. Among them, KNN is the most simple and the most widely used method, especially suitable for situations with a small number of cases. This study, therefore, applies KNN to retrieve similar cases since this study has small numbers of cases in the case base. The KNN process includes two parts: determining the weight of attributes (or influencing factors) and calculating the similarity of cases. Then the similar cases can be identified according to the similarity score. The third stage is the case revision, which aims to compensate for the deviations between retrieved cases and target cases and create a suitable solution for the target cases. This stage involves two steps: (1) using principle component analysis (PCA) to choose the most important factors, and (2) use multiple regression analysis (MRA) to revise selected cases. Three stages are further explained in detail in the following sections.

Case Database Establishment and CBR Data Preprocessing

Case Database Establishment

A database must be set up prior to the establishment of a CBR model. The more integrated the case database is, the more accurate an estimation result the CBR can generate (Ji et al. 2010). The primary source of the case database is the Mega Project Case Study Center of China (MPCSC 2014). In addition, the study supplemented this data with additional case information from the skyscrapers' official websites and historical documents. In total, the study collected 33 cases in China with the following characteristics: (1) height above 200 m, (2) completion date during the years 1993–2015, and (3) region within mainland China, in addition to one case in Hong Kong and one case in Taiwan. The basic characteristics of all projects are summarized in Fig. 3.

Determine Attributes

The study needs to determine attributes that mostly influence the project schedule for skyscrapers. As a common practice, a literature review and analysis were used in this study to identify the attributes impacting project schedules (Kim and Kang 2004). After review, 32 attributes were summarized as the basis for data processing.

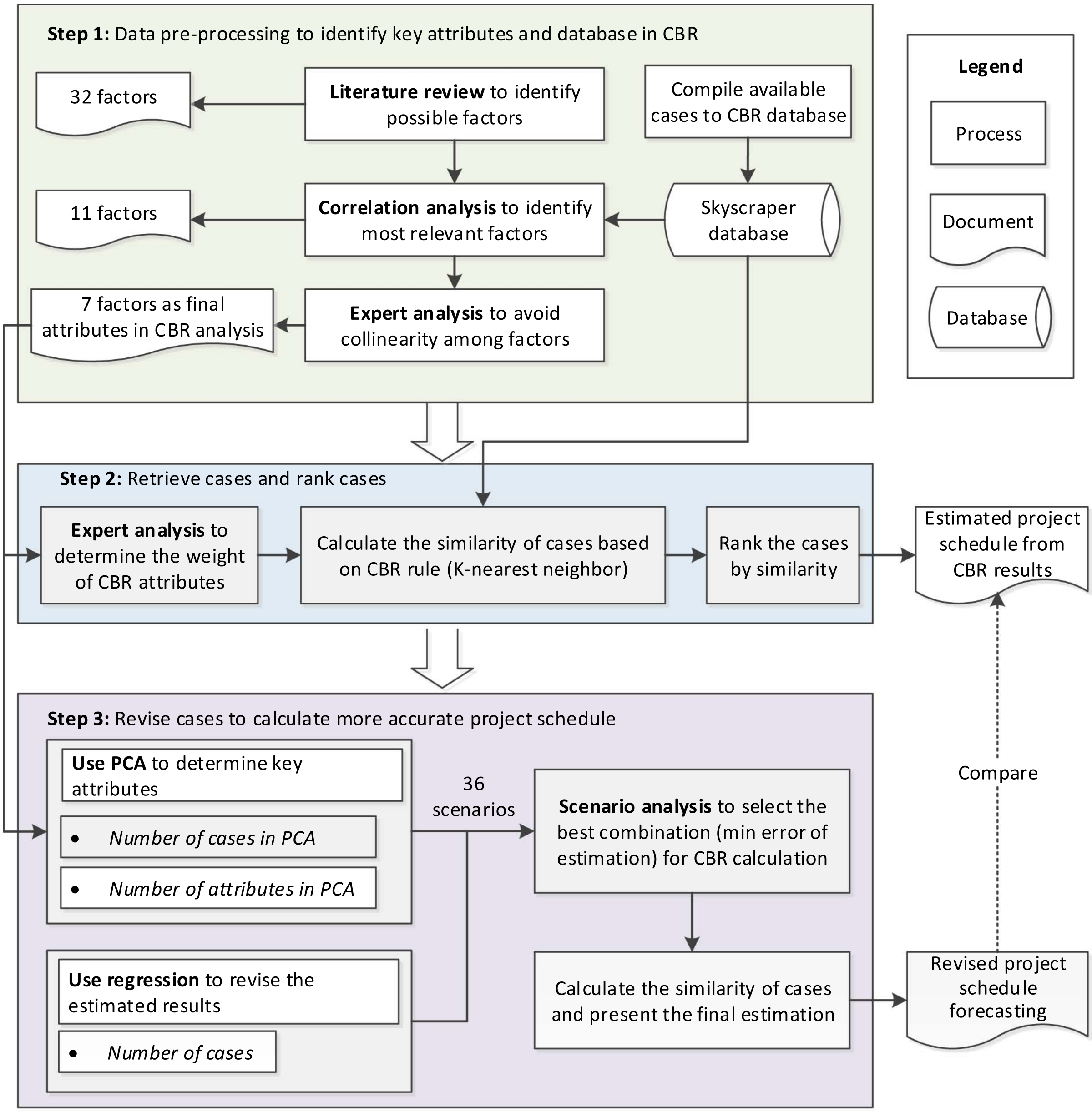

Fig. 2. Framework of CBR schedule forecasting model

The details of 32 attributes and their categories are presented in Table 1. Since these 32 attributes focus on different aspects of skyscrapers and their units are different, it is necessary to standardize the data into comparable values (Ji et al. 2010).

Then, the study applied correlation analysis to identify the most-relevant attributes to the project schedule based on all cases. The filtering criteria are defined as the bilateral probability of correlation coefficient, which was less than 0.05. Under these criteria, 11 attributes were identified. However, among them, one may correlate with one another, such as height and floors above ground. Therefore, two experts performed an independent check to eliminate overlapped attributes. To avoid collinearity, only one factor can be selected from A1, A2, A3, and A4. Similarly, only one factor can be selected between A5 andA6. Therefore, eight combinations were tested that include one factor from A1, A2, A3, and A4, and another one from A5 or A6 (Appendix I), then the group with minimal index values and lower collinearity effect was chosen. After discussion, seven attributes were finally selected, as given in Table 1. The selected attributed have also been compared to past studies to show their validity (Al-Momani 2000; Chan and Kumaraswamy 1999; CTBUH 2015; Jin et al. 2016; Kaming et al. 2010; Ogunlana and Promkuntong 1996), as compared in the last column of Table 1.

Case Retrieval and Ranking

Calculating Attribute Weights

Different methods exist to weight the attribute in CBR, such as analytic hierarchy process (AHP), genetic algorithms (Kim and Kim 2010), and expert analysis. Among them, the last one has been widely used in CBR studies like the CBR-based bidding decision-making model (Chua et al. 2001) and CBR-based construction procurement decision model (Luu et al. 2005). Therefore, expert interviews were selected to determine the weight that describes the importance of each factor. A total of four experts on construction

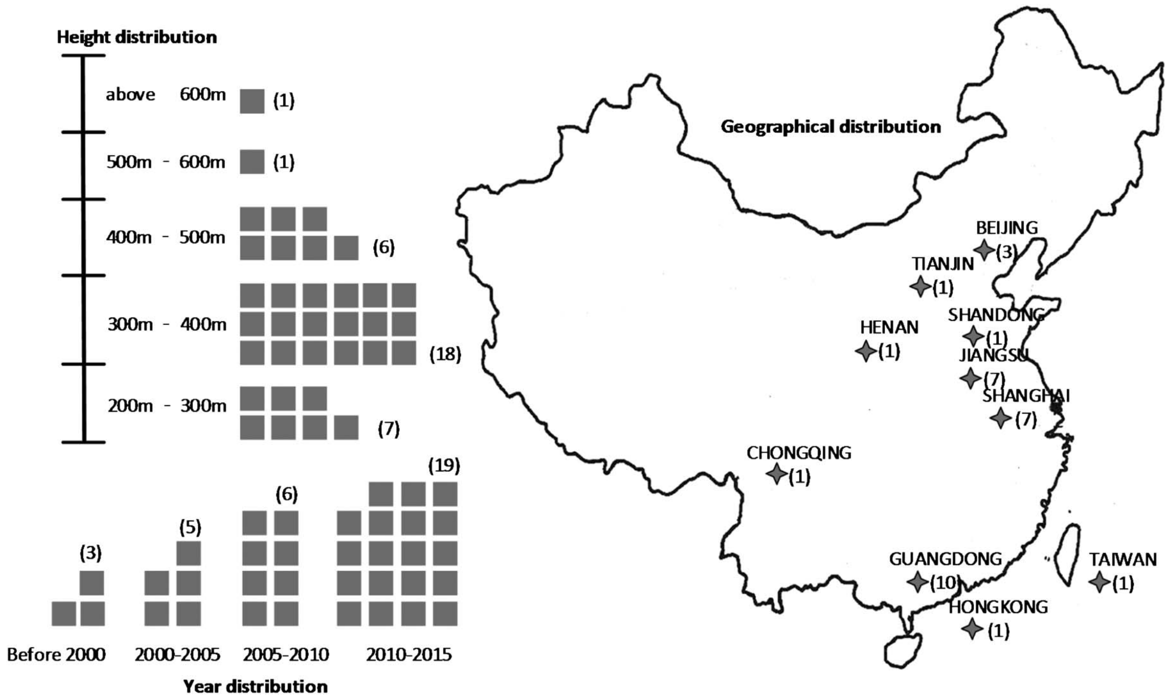

Fig. 3. Height, year, and location distribution of 33 selected skyscrapers in this study

Table 1. Selected Attributes and Corresponding Weights

Factor	Step 1[a]	Step 2	Step 3	Weight of factors[b]
Project-related	Height: ground to top (A1); **height: occupied floor (A2)(X1)**; floors above ground (A3); floors below ground (A5); tower GFA; gross floor areas; structural material; type of structure; **depths of excavation (A6) (X2)**; foundation; **excavation of earthwork (A7)(X3)**; functions	(A1), (A2), (A3), (A5), (A4), (A6), (A7)	(A2)[c], (A6)[d], (A7)[c]	X1 = 0.1355, X2 = 0.1694, X3 = 0.1864
Participants	Experience of constructors; type of owners; number of investors; **project cost (A8)(X4)**; experience of designers; experience of supervisors	(A8)	(A8)[e]	X4 = 0.1525
Condition of construction	Development of region; quantity of labor; **labor productivity (A9) (X5)**; building site categories; earthquake fortification intensity; annual freezing days (below 5°C); annual average rainfall; annual number of typhoons	(A9)	(A9)[f]	X5 = 0.1864
Economic	**Gross domestic product of the city (A10)(X6)**	(A10)	(A10)[g]	X6 = 0.0677
Skills factors	**Green building certificate (A11)(X7)**; application of building information modeling; top-down construction method; engineering quality award	(A11)	(A11)[h]	X7 = 0.1016
Hazard	Level of construction hazard	N/A	N/A	N/A

Note: Bold text represents the factors that were finally selected in the model.
[a]A total of 32 attributes were selected from Step 1 literature review, 11 attributes were selected from Step 2, correlation analysis, and 7 attributes were selected from Step 3, expert analysis.
[b]X1–X6 are numerical variables; X7 is an ordinary variable.
[c]Data from Chan and Kumaraswamy (1999).
[d]Data from Jin et al. (2016), which used the number of underground floors, similar to X2.
[e]Data from Ogunlana and Promkuntong (1996).
[f]Data from Kaming et al. (2010).
[g]Data from Al-Momani (2000), which used economic conditions, similar to X6.
[h]Data from CTBUH (2015).

and project management were interviewed, and they all had about 5–10 years of working experience on megaprojects. Among four experts, the first expert had years of experience in providing professional consultancy services for skyscraper construction; the second expert was a senior engineer and participated in construction of skyscrapers such as the Shanghai World Financial Center (492 m) and the Shenzhen Ping-An Building (600 m). The third one was a senior engineer with experience in complex high-rise building development. The fourth expert was an active researcher on megaproject scheduling estimation. A three-point scale weighting scheme was adopted from low to high impact, where one point indicates a factor that may be slightly important to deciding the schedule, and three points indicate that a factor is extremely important. After analysis, the weights for seven attributes are shown in the last column of Table 1.

Table 2. Estimation Results Based on Different Selected Referencing Cases

	Number of selected referencing cases							
	One case		Two cases		Three cases		Four cases	
Forecasting cases	Schedule (days)	ER (%)	Schedule (days)	ER (%)	Schedule (days)	ER (%)	Schedule (days)	ER (%)
Chongqing World Financial Center	1,740	10.72	1,918	1.59	1,854	4.85	1,854	4.86
Wuxi International Finance Square	1,460	5.01	1,703	10.78	1,753	14.04	1,725	12.23
Yuexiu Finance	1,580	10.96	1,491	4.70	1,461	2.61	1,514	6.31
Average	—	8.90	—	5.69	—	7.17	—	7.80

Calculating Case Similarity

The attributes are in different formats, such as numerical variables or ordinal variables. Correspondingly, two different calculation rules were proposed to calculate their similarity.

The similarity of continuous variable can be calculated as follows:

$$SW_i = 100 - 100 \times \frac{AV_{\text{test-case}} - \min(D_{AV})}{\max(D_{AV}) - \min(D_{AV})} \tag{1}$$

$$D_{AV_i} = |A_{\text{test-case}} - AV_{\text{retrieved-case}}|$$

where $AV_{\text{test-case}}$ = attribute data of test case; $AV_{\text{retrieved-case}}$ = attribute data of retrieved case (previous case); D_{AV} = difference between these two factors; and SW_i = similarity of the factor i.

As for ordinal variables, the similarity can be calculated as follows:

$$SW_i = \begin{cases} 100, & \text{if } AV_{\text{test-case}} - AV_{\text{retrieved-case}} = 0 \\ 0, & \text{else} \end{cases} \tag{2}$$

where the values of AV ranges from 1 to 3. Take green building certificates for instance, when the project has a platinum award, the value equals 1, a gold award equals 2, and no certificate is equal to 3.

The similarity of cases equals the sum of each factor's similarity

$$SC_j = \sum_{i=1}^{n} SW_i \tag{3}$$

where SC_j = similarity between test case and previous j case; and SW_i = similarity of the factor i.

Case Ranking

According to these given rules, all cases could be ranked in descending order of the case similarity. In CBR, the way of choosing training cases varies. Cho et al. (2010) used a ratio of testing case verse training case of 4:1, and Li and Sun (2011) selected a ratio of 7:3. Differently, Kim and Kim (2010) built a CBR model based on the total amount of collected data. In this study, the latter way was adopted due to the rationale that engineers in reality normally use all available cases to predict the schedule of a new skyscraper. The ranking result for the sample case Chongqing Finance is listed in Appendix II. Then, the case with highest similarity (SC_j) from the case database can be identified as the most similar case to the target case (Cho et al. 2010).

Error Rate

To show the accuracy of the model visually, the study used the error rate (ER) to measure the difference between actual schedule and estimated schedule and compared the results for different estimation methods (Cho et al. 2010)

$$ER = \frac{S_{\text{actual}} - S_{\text{estimated}}}{S_{\text{estimated}}} \tag{4}$$

where S_{actual} = actual schedule of the test case; and $S_{\text{estimated}}$ = estimated schedule of the test case.

Preliminary Results

The result of CBR model (unrevised) is listed in Table 2. During the calculation process, using a different number of retrieved cases will lead to different prediction results. For example, when 1–4 retrieved cases are used, the average error rate is 5.69–8.90%. The best result is $ER = 5.69\%$ when two similar cases were chosen. However, the ER of Wuxi International Finance Square (10.78%) showed large variances. Therefore, such deviations need to be decreased by revision.

Case Revision

Case revision refers to adjusting the deviation of the earlier estimates based on experts' experiences and knowledge. The current methods are primarily dependent on subjective judgment, which may lower the accuracy of the model's predictions (Begum et al. 2009). Alternatively, multiple regression analysis (MRA) has recently emerged as a promising method that provides a reliable way to adjust the original estimate based on the most-relevant variables. For example, Jin et al. (2012) provided a revised MRA. However, the application of this tool imposes a strict requirement that the number of cases should be greater than the number of factors (also called attributes). Given the limited number of cases, this study cannot directly apply the MRA method, but rather needs to reduce the number of attributes. To select the most-relevant attributes, this study used principle component analysis (PCA) to identify dominant factors from the seven factors used in the previous step. The technologies used to construct skyscrapers have changed rapidly over the years. To ensure the prediction is based on the current state-of-the-art technologies, the cases completed in recent 2 years, or a minimum of six recently completed cases, were used for the case revision process.

PCA to Select Dominant Factors

PCA is known as an orthogonal transformation that converts multiple variables into a set of principal components. The transformation process is defined in such a way that the sum of component score is calculated by the multiplication of two variables—the percentage of variation that represents the level of interpretation of the new factors to the original level and the score of original variable that represents the contribution towards the whole result (Ji et al. 2012). PCA has been applied in the following sequence. First, cases highly similar to the target case were selected. Second, PCA was applied to abstract the key attributes from these similar cases. Third, the selected attributes were fed into the MRA model. Among the seven attributed identified in the previous step, six numerical variables

Table 3. Rationale for Determining the Range of Uncertainty Variables

Uncertainty factors	Lower bound	Upper bound	Range
(A) Number of cases in the PCA	To determine the lower bound of (A): In order to conduct PCA, (A) must be equal to or greater than 8 because a maximum of 7 attributes exist in the PCA	To determine the upper bound of (A): When choosing more cases, the accuracy of the estimate will reduce. After several trials, the study decided to choose 12 as the upper limit of (A). The reason is that when the number of cases is 12, the error of estimate is larger than that when 11 cases are used	8–12
(B) Number of factors in the PCA (Factor_n)	To determine the lower bound of (B) and (C): To improve accuracy, the lower bound of (B) should be equal to or greater than 2	To determine the upper bound of (B) and (C): Take the similarity of previous cases into consideration, the upper bound of (C) is 6 because the cutting off point for the similar cases is six cases. Beyond six cases, the similarity to the target case will be low	2–4
(C) Number of cases in the multiple regression (Case_n)	Considering the same relationship as in the preceding row, that is $\text{Case}_n \gg \text{Factor}_n + 2$, the lower bound of (C) is 4	According to the upper bound of (C), the development of an appropriate MRA model requires to satisfy the condition that Case_n is greater than the Factor_n plus two ($\text{Case}_n \gg \text{Factor}_n + 2$)	4–6

and one ordinary attribute (X7, green building certificate) were selected as initial input variables in the PCA process.

MRA to Revise the Result

Followed by the PCA result, the filtered dominant attributes are then input to the MRA model to revise the CBR result. The revised estimate is calculated as follows:

$$Tr = C + k_1\text{factor}_1 + k_2\text{factor}_2 + \cdots + k_n\text{factor}_n \qquad (5)$$

where Tr = revised estimated schedule; C = constant; factor n = dominant factors chosen from PCA; and K_n = regression coefficient.

In sum, the preceding discussion outlines three uncertainty factors in Eq. (5), namely (1) the number of cases involved in PCA, (2) number of factors chosen from PCA, and (3) number of cases involved in MRA. The range of each factor and the reasons for selection are listed in Table 3.

Results

After considering three variances and their range of values, 36 possible combinations and results for Chongqing World Financial Center are listed in Table 4. The combination that has the lowest average ER (5.64%) were selected as the final result, in which three uncertainty factors can be determined—2 dominant attributes, 11 cases in the PCA, and 5 cases in the MRA model. In this scenario, the MRA formulation and case results are listed in Table 5. Project cost (X4) is significant to forecasting project schedule in the cases of Chongqing and Wuxi, while having a green building certificate (X7) is not a key determining variable.

To validate the accuracy of this model, these results were also compared with the results calculated by traditional methods. Four results are compared: (1) regression model that established a linear regression between 20 predictable factors—including number of stories above the ground, site location, total floor area, and project schedule (Koo et al. 2010); (2) industrywide quota estimation that used the national guidelines formulated by the Ministry of Construction in China (NSPRC 2000) to estimate the schedule of a building with fewer than 32 floors; (3) the original result of CBR model without revision; and (4) the result of the revised CBR model.

As shown in Table 6, the accuracy of the revised CBR model is the highest among all estimations. The average ER is 4.83%, which is 15.2% lower than the result calculated by the unrevised CBR model [i.e., $(5.7 - 4.83)/5.7 = 15.2\%$]. On the contrary, the results of traditional methods—both regression and quota estimation—are not accurate, especially the result of quota with ER as high as 35%. Regarding the estimation of project schedule, 5–10% of ER in the early stage is considered acceptable (Jin et al. 2016). So the results validate the applicability and reliability of using the CBR model to forecast project schedules for skyscrapers in the early stage of the construction process.

Discussion

As shown in the preceding sections, many variables have been examined during the CBR calculation and revision process. These variables may influence the final result from different perspectives. For instance, in the CBR revision process, the number of cases in PCA may impact on the result. Therefore, the robustness of the final result needs to be discussed and justified.

Decision Criterion and Robustness of Result

The selected combination used herein was based on the minimum ER on average in Table 4. However, from another point of view, the case revision process is also an MRA process, where the R^2 of the multiple regression is also a key indicator (Ji et al. 2012). Thus, it is worthy to compare the results of two different selection criteria: Option 1 with the minimum ER and Option 2 with the largest R^2 of MRA.

The average ER and R^2 of MRA for different combinations are listed in Table 7. The minimum ER in all combinations are 5.64, 6.54, 8.16, and 8.23%, respectively. The lowest and second lowest values of ER are 5.64 and 6.54%, while the two lowest values of R^2 are 0.62. Both results satisfy the requirement of the ER and R^2, so the difference between the two selections is negligible. In other words, the result of the revised MRA shown in Table 7 is robust regardless of the numbers of cases and selection between ER and R^2.

Another issue is when a new tallest building occurs, it will become an outlier and higher than the existing cases already collected. Therefore, this study calculated the scenario in which the average was calculated by excluding the largest ER in the testing cases. After removing an outlier (i.e., the highest ER) from each group, the predicted results are shown in the last row of Table 7, and this result could be more accurate. ER reduces to the range of 3.08–4.11%, and R^2 increases to the range between 0.67 and 0.69. So the final result is relatively robust, and it can only change a little no matter whether the lowest average ER criterion or apply the biggest R^2 has been applied.

Number of Cases and Forecast Accuracy

The number of cases in the CBR database is an important factor. The more cases recorded in the database, the more chances the target case can be matched within the case pool. Therefore, this

Table 4. Estimated Results of Multiple Regression Analysis (MRA) for 36 Combinations of Three Uncertainty Factors

		Four cases in MRA	Five cases in MRA		Six cases in MRA		
Numbers of cases in PCA	Projects	Two attributes in PCA (%)	Two attributes in PCA (%)	Three attributes in PCA (%)	Three attributes in PCA (%)	Three attributes in PCA (%)	Four attributes in PCA (%)
8	Zhengzhou[a]	20.02	8.73	4.40	15.44	0.70	19.24
	Guangsheng	4.63	12.45	1.51	9.12	14.32	14.61
	Moi City	11.51	9.26	23.25	9.16	28.03	21.10
	Leatop Plaza	4.26	1.24	5.28	15.23	10.14	54.39
	Deji Plaza	8.53	15.31	25.05	21.94	5.18	19.00
	Suning	4.39	1.59	1.68	5.01	3.52	4.52
	Modern	6.47	6.93	111.98	69.21	0.80	5.99
	Jinan	18.43	10.29	22.37	8.93	20.18	21.29
	Average	9.78	8.23	24.44	19.25	10.36	20.02
9	Zhengzhou	21.80	34.36	92.98	15.44	22.25	21.13
	Guangsheng	10.19	8.29	7.76	13.03	3.07	6.37
	Moi City	27.87	25.30	34.47	17.55	38.61	38.76
	Leatop Plaza	4.26	1.24	5.28	15.23	10.14	94.93
	Deji Plaza	8.53	15.31	25.05	21.94	5.18	19.00
	Suning	1.75	1.62	2.02	2.81	1.73	2.41
	Modern	6.47	6.93	111.98	69.21	0.80	9.27
	Jinan	15.55	2.69	22.37	3.99	20.18	21.29
	Average	12.05	11.97	37.74	19.90	12.75	26.65
10	Zhengzhou	20.95	22.44	37.42	17.46	19.53	21.13
	Guangsheng	17.20	4.74	7.76	4.18	3.07	16.76
	Moi City	27.87	25.30	28.28	17.55	14.32	27.11
	Leatop Plaza	4.26	1.24	5.28	15.23	10.14	94.93
	Deji Plaza	8.53	15.31	25.05	21.94	5.18	19.00
	Suning	44.72	14.40	12.99	2.54	1.28	15.16
	Modern	6.47	6.93	16.11	0.06	4.01	2.43
	Jinan	3.70	1.19	22.37	1.17	20.18	21.29
	Average	16.71	11.44	19.41	10.02	9.71	27.23
11	Zhengzhou	25.49	**1.82**	1.84	19.59	70.07	12.42
	Guangsheng	4.84	**3.22**	1.51	13.22	14.32	14.61
	Moi City	9.00	**6.79**	6.65	3.43	3.39	24.55
	Leatop Plaza	4.26	**1.24**	121.77	15.23	58.52	53.71
	Deji Plaza	23.58	**23.51**	25.05	4.55	5.68	42.15
	Suning	0.74	**1.59**	1.74	5.01	1.78	2.11
	Modern	4.78	**5.73**	112.43	3.06	0.80	9.27
	Jinan	3.70	**1.19**	22.37	1.17	20.18	21.29
	Average	9.55	**5.64**	36.67	8.16	21.84	22.51
12	Zhengzhou	12.30	55.51	53.12	21.43	32.38	27.16
	Guangsheng	3.48	2.31	1.53	14.33	14.92	28.61
	Moi City	9.00	6.79	6.65	3.43	3.39	24.55
	Leatop Plaza	1.12	6.48	21.15	13.80	38.25	53.71
	Deji Plaza	8.53	8.53	29.78	15.81	8.53	19.00
	Suning	3.43	4.76	5.73	4.90	2.41	4.34
	Modern	6.47	6.93	16.11	0.06	4.01	9.27
	Jinan	3.70	1.19	22.37	1.17	20.18	21.29
	Average	6.00	11.56	19.55	9.37	15.51	23.49

Note: Results are measured in the percentage (%) of error rate (ER); bold values represent the combination with lowest average ER.
[a]Zhengzhou refers to Zhengzhou Greenland Plaza; Guangsheng refers to Guangsheng International Plaza; Suning refers to Wuxi Suning Plaza; Modern refers to Modern Media Plaza; Jinan refers to Jinan Greenland Plaza.

Table 5. Revised MRA Formula for Test Cases

Project	MRA	Estimated schedule (days)	Actual schedule (days)	Error rate (%)
Chongqing[a]	$1{,}097.225 - 10.01X_4 + 43X_2$[b]	2,079	1,949	6.71
Wuxi	$863.515 + 15.43X_2 + 18.028X_4$	1,499	1,537	2.47
Yuexiu	$-33.157 + 12.001X_3 + 75.993X_2$	1,348	1,424	5.29

[a]Chongqing refers to World Financial Center; Wuxi refers to Wuxi International Finance Square; Yuexiu refers to Yuexiu Finance Plaza.
[b]X_2, X_3, and X_4 are described in Table 1.

study also tested the relationship between the forecasting accuracy and number of cases in the CBR database. According to the completion year, the number of completed skyscraper cases before the year 2012, 2013, and 2014 were 20, 26, and 30, respectively. The study then estimated the schedule in each year for the three targeted cases and the results are given in Fig. 4. As the candidate cases in the CBR database increased, the ER of three estimations became lower. Extended from this finding, the ER of 33 cases in the current CBR database is 4.83%, and the accuracy of the estimation would presumably become higher if more cases are recorded in the CBR database in the future.

Table 6. Results Calculated by Different Schedule Estimation Methods

Case	Regression		National quota[a]		CBR without revision[b]		CBR with revision		Actual schedule
	Days	ER (%)	Days	ER (%)	Days	ER (%)	Days	ER (%)	
Chongqing[c]	1,882	3.4	2,419	24.1	1,981	1.6	2,079	6.7	1,949
Wuxi	1,663	8.2	1,905	23.9	1,703	10.8	1,499	2.5	1,537
Yuexiu	1,192	16.3	2,162	51.8	1,491	4.7	1,348	5.3	1,424
Average	—	9.30	—	33.27	—	5.70	—	4.83	—

Note: Schedule is measured in days and the error rate (ER) measured in percentage (%).
[a]National quota is published to estimate the schedule for buildings within 32 floors above ground and 4 floors below ground; in this study, its estimated schedule is for reference and comparison only.
[b]Result is based on the selection of two similar cases.
[c]Chongqing refers to World Financial Center; Wuxi refers to Wuxi International Finance Square; Yuexiu refers to Yuexiu Finance Plaza.

Table 7. Comparison of the Estimation Schedule Calculated by Different Numbers of Cases

Project	Five cases in MRA, two attributes in PCA, and 11 cases in PCA		Four cases in MRA, two attributes in PCA, and 11 cases in PCA		Six cases in MRA, two attributes in PCA, and 11 cases in PCA		Five cases in MRA, two attributes in PCA, and eight cases in PCA	
	ER (%)	R^2	ER (%)	R^2	ER (%)	R^2	ER (%)	R^2
Zhengzhou[a]	1.82	0.60	1.45	0.70	19.59	0.03	8.73	0.94
Guangsheng	3.22	0.56	4.84	0.91	13.22	0.51	12.45	0.51
Moi City	6.79	0.79	9.00	0.77	3.43	0.76	9.26	0.03
Leatop Plaza	1.24	0.31	4.26	0.23	15.23	0.05	1.24	0.32
Deji Plaza	23.51	0.94	23.58	0.94	4.55	0.59	15.31	0.02
Suning	1.59	0.92	0.74	0.10	5.01	0.13	1.59	0.27
Modern	5.73	0.39	4.78	0.41	3.06	0.43	6.93	0.41
Jinan	1.19	0.47	3.70	0.93	1.17	0.45	10.29	0.28
Average	5.64	0.62	6.54	0.62	8.16	0.37	8.23	0.35
$\text{Average}_{\text{plus}}$[b]	3.08	0.67	4.11	0.69	6.53	0.41	7.21	0.39

[a]Zhengzhou refers to Zhengzhou Greenland Plaza; Guangsheng refers to Guangsheng International Plaza; Suning refers to Wuxi Suning Plaza; Modern refers to Modern Media Plaza; Jinan refers to Jinan Greenland Plaza.
[b]$\text{Average}_{\text{plus}}$ recalculated the result by removing the maximum ER value.

In addition, the study also found diminishing marginal returns in that the forecasting accuracy increased disproportionately as number of cases increased. Specifically, the incremental increase of the forecasting accuracy became lower when cases were fed to the database at a later time. For example, the average ER of the three test cases decreased by 45% when increasing the number of cases from 20 to 26, and no change occurred when increasing the number of cases from 26 to 30. The possible reason behind this diminishing law is due to the single-pairing principle. It means that for each target case, the increase of CBR cases will increase the likelihood of finding the best pair for the targeted case. However, once the best pairing case was identified, the accuracy of the estimation will not be affected even if more additional projects are added into the CBR case database.

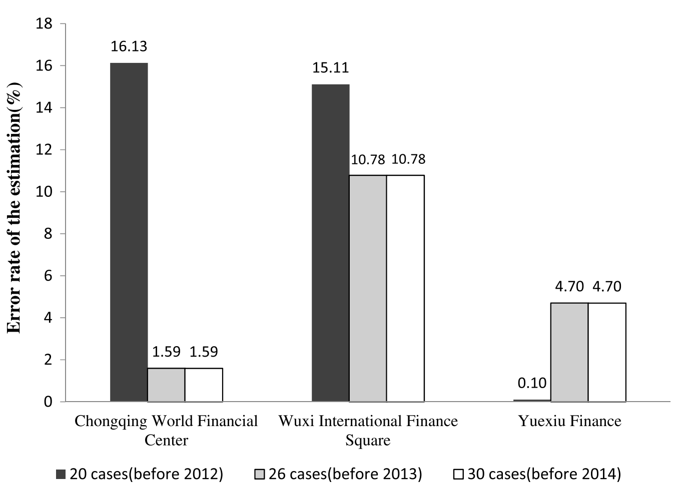

Fig. 4. Comparison of estimation accuracy (error rate, %) when using databases in different years

Conclusion

Accurate schedule estimation in the early stage of skyscraper construction plays an enormous role in supporting stakeholders' decisions. To better forecast a project's schedule, this study established a revised CBR model to predict project schedules for skyscrapers in considerations of the massive quantities of complex factors involved in the skyscrapers. Based on data collected in China, the ER of the revised CBR model is 4.83%, which satisfies the estimated standard and significantly lowers than the traditional method. The result of this study justifies the conclusion that the CBR is a promising tool for estimating project schedules for megaprojects such as skyscrapers. Though this study is based on the context of China, its contribution could be more generalized to skyscrapers around the globe.

Specifically, this study contributes three aspects to existing knowledge. First, this model offers distinct improvements over the traditional CBR technique. Compared with the traditional CBR that mostly depends on the subjective judgment or MRA, this study improves the CBR revision process by using a new set of calculation methods such as PCA, MRA, and scenario analysis (Table 4) to select the number of revision cases, ensuring the accuracy and efficiency of the revision process. Second, it expands the applicability of CBR in the domain of project scheduling estimation. The CBR results confirmed the strong potential and promising result of CBR in estimating construction duration, especially at the early stage of a project. Third, the CBR model developed in this study provides an expandable scheduling tool compared to existing scheduling tools (e.g., CPM, PERT, and bar charts), because the CBR model could be dynamically improved along with the richness of the retrievable cases database in the long term once more data become accessible. Such an adaptive feature is complementary to the existing body of knowledge on project scheduling methods.

The findings reveal several possible practical implications. First, as the symbol of a city, skyscrapers normally demand huge investments and long project periods; therefore, they face high economic and political risks. An accurate schedule forecasting method can effectively reduce such risks by narrowing down the contractual variations between the estimated and actual schedules. From a broader perspective, reduction of schedule variations could also avoid the risks and potentially negative impacts on the regional economy. Second, in the current practice of skyscraper construction, schedule estimation mainly depends on individual experts and their experience, meaning that a quantifiable and measurable schedule method is lacking. The proposed CBR model bridges this gap and provides a feasible estimation method for skyscrapers in the early stage of a project. Meanwhile, the CBR model is also of value to owners by providing them with a schedule decision support system. Owners can use this model to compare different project plans and make a reliable selection.

Although the CBR model can provide a satisfactory result, some limitations exist in this paper, which can be studied in the future. The first limitation is the number of cases in the CBR database. As demonstrated in an earlier section, an increased case numbers will boost the accuracy of the forecasting model. This can be done in the future when more data on skyscrapers are available to be collected. The second limitation is that the CBR revision considered both numerical and categorical variables, but did not consider descriptive information, such as the experience of constructors. Hence, further investigation can analyze the use of different formats of variables (such as continuous, ordinal, or categorical variables) so that all aspects of skyscrapers can be reflected in the model. Third, the expert interview was selected to determine the weight for different forecasting factors but this may induce inaccurate results due to bias of respondents. Future studies can use alternative weighting schemes, such as genetic algorithms, to determine the weight values of the attributes. Last but not least, a future study can automate this model and provide more-efficient and fast decision support functions for decision makers.

Appendix I. Eight Combinations of Pro-Process Attributes

Combination	A1 and A5	A2 and A5	A3 and A5	A4 and A5
R^2	0.699	0.704	0.699	0.717
Maximum conditional index value	18.910	18.917	18.675	18.389
Factor of VIF >5	None	None	None	A9
Combination	A1 and A6	A2 and A6	A3 and A6	A4 and A6
R^2	0.714	0.720	0.720	0.731
Maximum conditional index value	15.946	15.903	16.033	16.240
Factor of VIF >5	A9	A9	A9	A9

Note: A1 = height from ground to top; A2 = height from group to occupied floor; A3 = floor above ground; A4 = gross floor areas; A5 = floors below ground; A6 = depth of excavation; A9 = labour productivity.

Appendix II. Similarity Ranking List of the Case—Chongqing

Project name	Schedule (days)	Similarity
China World Trade Center Tower III	1,740	78.936129
Pearl river	2,097	78.426463
Leatop plaza	1774	76.665821
Modern media plaza	1,854	74.365379
Zhengzhou greenland plaza	2,010	74.180691
KK100	1,522	74.118444
The Pinnacle	1,675	71.641997
Tianjin world finance center	1,562	69.518441
Wuxi suning plaza	1,640	66.036657
Wheellock square	1,580	65.587034
Deji plaza	1,460	64.988661
Zifeng tower	1,946	62.885374
International commerce centre	1,999	61.935882
Moi city	1,509	59.853433
Jinan greenland plaza	1,424	56.719358
Guangzhou international finance center	1,409	55.209906
Shanghai world finance center	1,296	53.651769
CCTV headquarters	2,575	53.475409

Appendix II (*Continued.*)

Project name	Schedule (days)	Similarity
Longxi international hotel	1,256	52.027733
Jinmao tower	1,408	49.232698
Plaza66	1,705	48.81462
Beijing yintai centre	1,400	47.579486
SEG plaza	1,580	45.600156
Bank of China tower	1,058	44.40215
Tomorrow square	2,132	44.307313
Shimao international plaza	1,645	42.631853
Taipei 101	2,522	42.388967
Diwang mansion	1,104	35.845768
Canton tower	1,376	35.628412
Citic plaza	1,399	32.278775

Note: A1 = height from ground to top; A2 = height from group to occupied floor; A3 = floor above ground; A4 = gross floor areas; A5 = floors below ground; A6 = depth of excavation; A9 = labour productivity.

Acknowledgments

This study is financially supported by the National Natural Science Foundation of China (Project No. 71390523 and 71471136) and National University of Singapore (R-296-000-144-646). However, the views represented in this article are those of the authors only.

References

Aamodt, A., and Plaza, E. (1994). "Case-based reasoning—Foundational issues, methodological variations, and system approaches." *Ai Commun.*, 7(1), 39–59.

Al-Momani, H. A. (2000). "Construction delay: A quantitative analysis." *Int. J. Project Manage.*, 18(1), 51–59.

Assaf, S. A., Alkhalil, M., and Alhazmi, M. (1995). "Causes of delay in large building construction projects." *J. Manage. Eng.*, 10.1061/(ASCE)0742-597X(1995)11:2(45), 45–50.

Begum, S., Ahmed, M. U., Funk, P., Xiong, N., and Von Scheele, B. (2009). "A case-based decision support system for individual stress diagnosis using fuzzy similarity matching." *Comput. Intell.*, 25(3), 180–195.

Chan, D. W. M., and Kumaraswamy, M. M. (1999). "Forecasting construction durations for public housing projects: A Hong Kong perspective." *Build. Environ.*, 34(5), 633–646.

Cho, S., Hong, H., and Ha, B. C. (2010). "A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the Mahalanobis distance: For bankruptcy prediction." *Expert Syst. Appl.*, 37(4), 3482–3488.

Chua, D. K. H., Li, D. Z., and Chan, W. T. (2001). "Case-based reasoning approach in bid decision making." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2001)127:1(35), 35–45.

CTBUH (Council on Tall Buildings and Urban Habitat). (2011). "The tallest 20 in 2020: Entering the era of the megatall." ⟨http://www.ctbuh.org/TallBuildings/HeightStatistics/BuildingsinNumbers/TheTallest20in2020/tabid/2926/language/en-US/Default.aspx⟩ (Nov. 11, 2015).

CTBUH (Council on Tall Buildings and Urban Habitat). (2015). "A tall building review." ⟨http://www.ctbuh.org/⟩ (Dec. 1, 2015).

Duverlie, P., and Castelain, J. M. (1999). "Cost estimation during design step: Parametric method versus case based reasoning method." *Int. J. Adv. Manuf. Technol.*, 15(12), 895–906.

Goh, Y. M., and Chua, D. K. H. (2010). "Case-based reasoning approach to construction safety hazard identification: Adaptation and utilization." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)CO.1943-7862.0000116, 170–178.

Hong, T., Cho, K., Hyun, C., and Han, S. (2011). "Simulation-based schedule estimation model for ACS-based core wall construction of high-rise building." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)CO.1943-7862.0000300, 393–402.

Ji, C., Hong, T., and Hyun, C. (2010). "CBR revision model for improving cost prediction accuracy in multifamily housing projects." *J. Manage. Eng.*, 10.1061/(ASCE)ME.1943-5479.0000018, 229–236.

Ji, S. H., Park, M., and Lee, H. S. (2012). "Case adaptation method of case-based reasoning for construction cost estimation in Korea." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)CO.1943-7862.0000409, 43–52.

Jin, R., Cho, K., Hyun, C., and Son, M. (2012). "MRA-based revised CBR model for cost prediction in the early stage of construction projects." *Exp. Syst. Appl.*, 39(5), 5214–5222.

Jin, R., Han, S., Hyun, C., and Kim, J. (2014). "Improving accuracy of early stage cost estimation by revising categorical variables in a case-based reasoning model." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)CO.1943-7862.0000863, 04014025.

Jin, R., Han, S., Hyun, C. T., and Cha, Y. (2016). "Application of case-based reasoning for estimating preliminary duration of building projects." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)CO.1943-7862.0001072, 04015082.

Kaming, P. F., Olomolaiye, P. O., Holt, G. D., and Harris, F. C. (2010). "Factors influencing construction time and cost overruns on high-rise projects in Indonesia." *Constr. Manage. Econ.*, 15(1), 83–94.

Kim, G., and Kang, K. (2004). "A study on predicting construction cost of apartment housing projects based on case based reasoning technique at the early project stage." *J. Archit. Inst. Korea*, 20(5), 83–92.

Kim, K. J., and Kim, K. (2010). "Preliminary cost estimation model using case-based reasoning and genetic algorithms." *J. Comput. Civ. Eng.*, 10.1061/(ASCE)CP.1943-5487.0000054, 499–505.

Kim, S., Maghiar, M., Li, L., Bai, Y., and Scott, J. (2014). "Developing a knowledge based information system (KISCCES) for construction cost estimating and scheduling." *Construction Research Congress 2014*, Construction Research Council, ASCE Construction Institute, Atlanta.

Kolodner, J. L. (1983). "Maintaining organization in a dynamic long-term memory." *Cognit. Sci.*., 7(4), 243–280.

Koo, C., Hong, T., Hyun, C., and Koo, K. (2010). "A CBR-based hybrid model for predicting a construction duration and cost based on project characteristics in multi-family housing projects." *Can. J. Civ. Eng.*, 37(5), 739–752.

Le, Y., and Li, Y. (2013). *Chinese Skyscraper Construction and Development Research Rep.*, China Architecture & Building Press, Beijing.

Li, H., and Sun, J. (2011). "Principal component case-based reasoning ensemble for business failure prediction." *Inform. Manage.*, 48(6), 220–227.

Lin, M. C., Tserng, H. P., Ho, S. P., and Young, D. L. (2011). "Developing a construction-duration model based on a historical dataset for building project." *J. Civ. Eng. Manage.*, 17(4), 529–539.

Lin, R.-H., Wang, Y.-T., Wu, C.-H., and Chuang, C.-L. (2009). "Developing a business failure prediction model via RST, GRA an CBR." *Exp. Syst. Appl.*, 36(2), 1593–1600.

Lu, Y., Cui, Q., Chen, X., and Baecher, G. (2014). "Model for schedule risk evaluation in subsurface explorations." *Transp. Res. Rec.*, 2433, 18–26.

Luu, D. T., Ng, S. T., and Chen, S. E. (2005). "Formulating procurement selection criteria through a case-based reasoning approach." *J Comput. Civ. Eng.*, 10.1061/(ASCE)0887-3801(2005)19:3(269), 269–276.

Merrill, E., and Gray, L. (2012). "A paradigm shift in the 21st century skyscraper." *Proc., Council on Tall Buildings and Urban Habitat. Asia Ascending: Age of the Sustainable Skyscraper: A collection of State-of-the-Art, Multi-Disciplinary Papers on Tall Buildings and Sustainable Cities—The CTBUH 9th World Congress*, China Civil Engineering Society and CTBUH, Dubai, 715–721.

MPCSC (Mega Projects Case Study and Data Center). (2014). "Research institute of complex engineering and management." ⟨http://www.mpcsc.org/⟩ (Aug. 1, 2015).

Mulholland, B., and Christian, J. (1999). "Risk assessment in construction schedules." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(1999)125:1(8), 8–15.

NSPRC (National Standard of the People's Republic of China). (2000). *Code for design of civil buildings*, China Planning Press, Beijing.

Ogunlana, S. O., and Promkuntong, K. (1996). "Construction delays in a fast-growing economy: Comparing Thailand with other economies." *Int. J. Project Manage.*, 14(1), 37–45.

Okmen, O., and Oztas, A. (2008). "Construction project network evaluation with correlated schedule risk analysis model." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2008)134:1(49), 49–63.

Ozorhon, B., Dikmen, I., and Birgonul, M. T. (2006). "Case-based reasoning model for international market selection." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2006)132:9(940), 940–948.

Ping, X. O., et al. (2015). "A multiple measurements case-based reasoning method for predicting recurrent status of liver cancer patients." *Comput. Ind.*, 69, 12–21.

Sambasivan, M., and Soon, Y. W. (2007). "Causes and effects of delays in Malaysian construction industry." *Int. J. Project Manage.*, 25(5), 517–526.

Schank, R. C. (1983). *Dynamic memory: A theory of reminding and learning in computers and people*, Cambridge University Press, Cambridge, U.K.

Sidwell, A. C. (1984). "The time performance of construction projects." *Archit. Sci. Rev.*, 27(4), 85–91.

Thomasson, B. J., Ratcliffe, M. B., and Thomas, L. A. (2006). "Improving the tutoring of software design using case-based reasoning." *Adv. Eng. Inform.*, 20(4), 351–362.

Wang, W. C., and Demsetz, L. A. (2000). "Application example for evaluating networks considering correlation." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2000)126:6(467), 467–474.

Yamin, R. A., and Harmelink, D. J. (2001). "Comparison of linear scheduling model (LSM) and critical path method (CPM)." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2001)127:5(374), 374–381.