Expert Systems with Applications 39 (2012) 2247–2255

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A novel dynamic progress forecasting approach for construction projects

M. Chiao Lin [a,b], H. Ping Tseng [c,*], S. Ping Ho [c], D.L. Young [c]

[a] Division of Construction Engineering and Management, Department of Civil Engineering, National Taiwan Univ., No. 1, Roosevelt Rd., Sec. 4, Taipei, Taiwan
[b] Center of General Education, Tzu Chi Univ., Taiwan
[c] Department of Civil Engineering, National Taiwan Univ., No. 1, Roosevelt Rd., Sec. 4, Taipei, Taiwan

A R T I C L E   I N F O

Keywords:
Construction project
Forecast
Grey dynamic prediction model
Modified residual model
Polynomial

A B S T R A C T

In this paper, we propose a novel construction project progress forecasting approach which combines the grey dynamic prediction model and the residual modified model to forecast the current progress during the construction phase. Firstly, four typical S-curves simplified from various sigmoid curves are proposed and fitted to the grey dynamic prediction model. For higher prediction accuracy, three different residual modified models are taken to amend the initial prediction value which was derived from the above step. The mean absolute percentage error (MAPE) and standard deviation of the estimate of Y (SDY) are used to assess the accuracy of the composite results. The better residual modified prediction model is adopted to combine the grey dynamic prediction model to form the novel progress forecasting approach. Then, practical completed construction cases are provided for testing the prediction ability of the proposed progress forecasting approach. Results show that the forecasting approach proposed to forecast construction progress during construction phase is able to get better prediction accuracy almost within 10% whether typical S-curves or practical cases. The new approach relatively provides an accurate, simple and stable method for predicting construction progress in comparison with the previous traditional forecasting methods.

1. Introduction

Monitoring and controlling the progress of current construction project are always indispensable to project managers' mission. The common monitoring method of construction progress is to compare the difference of real performed progress against contract/scheduled progress of the project. Regrettably, it is most often too late for top managers to highlight and take action in the case that significant discrepancies of project's construction performance are found out or the real progress falls behind far from the schedule (Goldratt, 1977; Kaka, 1999). A critical fact of project failure is construction delay, which often results in construction claim (Yeo & Ning, 2002). Progress forecasting and monitoring are not only essential for the survival of any contractor during the construction phase, but also important for the performance of any client's manager. S-curves are usually taken as expression of construction progress and have become a requisite tool for engineering managers throughout their execution phase. They can be used to forecast the likely duration of the project once the contract price and cumulative expenditure are known (Tuker, 1988). It is not only applied to state current construction status but also to forecast the future of construction projects (Miskawi, 1989; Tuker, 1988). They can also be used to manage cash flow, current performance status, future necessary cost/duration, etc. for running projects (Barraza, Back, & Mata, 2000; Blyth & Kaka, 2006; Kaka, 1999). S-curves were developed for exhibiting statistically significant difference between successful (meeting or exceeding budget and schedule expectations) and less-than successful (not meeting budget and/or schedule expectations) two project categories with continuous variables (Russell, Jaselskis, & Lawrence, 1997). Furthermore, those curves were proposed to assess and monitor the performance of a project progresses to determine if it is tracking according to a successful project during project execution.

Many forecasting S-curve models had been created to offer project manager's exercise. The common methodology of those researches applied for producing S-curve forecasting models adopts the multiple linear regression technique (Blyth & Kaka, 2006; Kaka, 1999; Kaka & Price, 1993; Skitmore, 1998). However, the traditional regression models are limited in that they should gather a lot of history sample data and make many strict assumptions to distribution of samples. Every construction project is unique, and the individual project characteristics of a group may vary from situation to situation and could display diverse curve types. Additionally, different contract arrangements lead to different cash-flow profiles (Tuker, 1988). Even with the same contract, different managers could produce different cash-flow profiles. These reasons may be why the regression S-curve model taken to fit individual

* Corresponding author.
E-mail addresses: d93521019@ntu.edu.tw (M. Chiao Lin), hptserng@ntu.edu.tw (H. Ping Tserng), spingho@ntu.edu.tw (S. Ping Ho), dlyoung@ntu.edu.tw (D.L. Young).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.07.093

projects could not be well-fitted. On the other hand, numerous software packages have been developed and helpful to calculate the current performed progress based on the site physical construction activities, but they still could not forecast the progress of scheduled activities. Real progress reports were produced by site managers and required to be accurate and on time, but there are many drawbacks of site reports that include inaccurate data, misrepresentation of information, and untimely feedback, and those reports made managers unable to find out discrepancies instantaneously during projects' progress performance (Kaka, 1999). Furthermore, from the view of control and management, the progress control of projects should not solely be relied on reports generated from practitioners working on site because of a number of uncertain factors. It is necessary to develop detailed progress predictions for individual construction project.

As mentioned above, there are inherent limitations to conventional statistical prediction methods such as requiring a large amount of data to analyze the characteristics of the system, ascertaining assumption regarding statistical distribution of data, and so on. Particularly, it is difficult that traditional statistical methods can not flexibly determine the model coefficients to form a sectional model to adapt any current situation. In contrast, the grey system theory (Deng, 1982) only requires a limited amount of discrete data to study the behavior of a system with incomplete information. In the grey system theory, there are three systems classified by the degree of information completed. A white system is defined as the case where information is fully known; while a black system is defined as the case where information is unknown. A system with partial information known and partial information unknown is defined a grey system. In the grey system, every series stands for one of factors, one of patterns, one of schemes, or one of behaviors, etc. (Deng, 2005). The grey dynamic prediction model, dynamic GM (1,1), kernel of grey theory system, is the first order grey model with one variable based on grey system theory, and has advantages in establishing a model with few data and incomplete information to realize the prediction of the system. It is not necessary to employ all data from the original series to construct a grey prediction model. The minimum number of data must be four in consecutive order without bypassing any data (Deng, 1986). The critical essential of GM (1,1) is the use of grey generating approach to settle the variation of the original data series. The accumulated generating operation, AGO, converts a series lacking obvious regularity into a monotonously increasing series to reduce the randomness of the series, and increase the smoothness of the series (Deng, 1989). The grey dynamic prediction model has advantages in establishing a sectional model *with recent few data to answer then situation,* and has taken the place of the original grey prediction model to form the core of grey prediction theory.

For the possibility of improving forecasting accuracy, Makridakis et al. (1982) proposed that the combination of forecasting methods performed well for most of the various types of series. According to accumulative literatures, Clemen (1989) pointed that forecast accuracy can be improved through the combination of multiple individual forecasting. There are many ways in which various hybrid models may be devised thorough previous research in the use of combined forecasts (Clemen, 1989; Fildes, Hibon, Makridakis, & Meade, 1998). Charles and Chase (2000) also argues that combining forecasts from different methods offers a powerful strategy for improving accuracy. A composite forecast is created by using different techniques independent of one another and combining the results.

In the paper, we proposed a novel approach which combines the grey dynamic prediction model GM (1,1) and a residual modified model to generate the expected progress prediction values. The algorithmic procedure is started with the four successive data from the start point to build a grey prediction model in order to get the next initial prediction value and four subsequent residuals of the model. Those residuals form a series and are used to build a residual modified model for getting the predictive error value of the above initial prediction value. The combination of the initial prediction value and the predictive error value is what the progress prediction value expected. The above calculating operation is continuously repeated until lately data series are run out. For higher prediction accuracy, three different residual modified models are taken to amend the initial prediction value derived from the grey dynamic prediction model. The standard deviation about the estimate of Y (SDY) and the mean absolute percentage error (MAPE) were used to assess the accuracy of the two models. The model that got better modified prediction value is adopted to combine the grey dynamic prediction model to form the novel progress forecasting approach. Finally, four practical completed construction cases are provided for testing the novel progress forecasting approach. Results show that the approach used to forecast construction progress during construction phase gets excellent prediction accuracy almost within 10% whether typical S-curves or practical cases. The contribution to this research is that the approach built with few data can be unique to construction project in more accuracy without reference to the program of site work.

2. Literatures review

Various mathematical formula forms for S-curves had been developed to forecast and monitor the progress status and cash flow of construction projects in past decades (Barraza et al., 2000; Berny & Howes, 1982; Blyth & Kaka, 2006; Hudson 1978; Kaka, 1999; Kenley & Wilson, 1986; Miskawi, 1989; Skitmore, 1998; Tuker, 1988). Most developed formulas were based on different classified groups to fit into the individual groups. Those classified and distributed criteria included type of project, duration of contract, type of procurement, and size of company (Kaka & Price, 1993). Kenley and Wilson (1986) argued that substantial variations existing between construction project and forecasts of individual construction project cash flows are invalid when derived from analysis of grouped data. They proposed an idiographic methodology to build individual construction project cash flows model based on the logit transformation approach, and had achieved good forecast results from the model on condition that the range of data points 10% and above 90% of S-curves were excluded. At the same time, they mentioned that every project is generally unique and should be modeled separately. Skitmore (1998) utilized three approaches, termed (1) analytic, (2) synthetic, (3) hybrid, in combination with six alternative models comprising (1) Hudson, (2) Kenley and Wilson, (3) Berny–Howes, (4) cumulative logistic, (5) cumulative normal, (6) cumulative lognormal, etc. to determine the best approach/model combination for the available database and forecast for future expenditure flows. The result also showed that the best hybrid models produce the most accurate ex-ante forecasts for 10–90% completed. He mentioned for the future work that the development of diagnostics, such as residual analysis, would be an important aid to future refinement in the modeling procedure. In contrast with deterministic cost flow models, Kaka (1999) developed a stochastic model based on historical data with logit transformation technique to allow users to incorporate variability and inaccuracy in forecasts and decision-making. Barraza et al. (2000) also developed stochastic S-curves which could provide probability distributions of budgeted cost and planned elapsed time for a given percentage of progress to evaluate cost and time variations by comparing the expected budgeted and planned duration. Blyth and Kaka (2006) proposed a model that standardized activities, and forecast the duration, cost and end dates of these activities based upon determined project characteristics. The model was used to