Tourism Management 52 (2016) 369–379

Contents lists available at ScienceDirect

Tourism Management

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

Case study

Using a Grey–Markov model optimized by Cuckoo search algorithm to forecast the annual foreign tourist arrivals to China

Xu Sun [a], Wangshu Sun [a], Jianzhou Wang [a, *], Yixin Zhang [b], Yining Gao [a]

[a] School of Statistics, Dongbei University of Finance and Economics, Dalian, 116025, China
[b] School of Mathematics and Statistics, Lanzhou University, Lanzhou, 73000, China

HIGHLIGHTS

- Proposing a novel optimized Markov-chain grey model.
- Applying the optimal input subset method and Cuckoo search algorithm.
- Improving the forecasting accuracy of traditional Markov-chain grey model.
- Providing a novel efficient forecasting method for the trend of tourism demand.

ARTICLE INFO

Article history:
Received 25 November 2014
Received in revised form
3 July 2015
Accepted 7 July 2015
Available online xxx

Keywords:
Forecast
China
Tourism demand
Optimal input subset
Cuckoo search algorithm

ABSTRACT

With the rapid development of the international tourism industry, it has been a challenge to forecast the variability in the international tourism market since the 2008 global financial crisis. In this paper, a novel CMCSGM(1, 1) forecasting model is proposed to address how forecasting precision is affected by the volatility of the tourism market. The Markov-chain grey model is adopted for its emphasis on the small-sample observations and exponential distribution samples. Additionally, the optimal input subset method and the Cuckoo search optimization algorithm are applied to improve the performance of the Markov-chain grey model. The experimental study of the forecasting of the annual foreign tourist arrivals to China indicates that the proposed CMCSGM(1, 1) model is considerably more efficient and accurate than the conventional MCGM(1, 1) models.

1. Introduce

China has not been immune to the significant growth of international tourism over the past two decades, and indeed today is a major generator and recipient of tourists. Because of the crucial contributions of the international tourism industry to the nation's economic development through trade performance and service industry employment, it is highly important for Chinese tourism agencies (e.g., government bodies and the private sector) to understand the trends affecting annual foreign tourism arrivals, which can help them to propose appropriate strategies or effectively allocate resources to boost the tourism industry further. Hence, for basic information and tourism planning, it is necessary to develop an efficient method to forecast the trends of international tourism demand accurately.

Based on the international tourism market share and the data published by China National Tourism Administration, in 2013, Korea was the most major tourist market for China, whereas Japan, Russia, U.S., Malaysia, Mongolia, Philippines and Singapore occupied the second to eighth positions, respectively. In recent years, with certain emergences and developments in emerging tourist destinations, the world's international tourism industry has become increasingly competitive. Under an increasingly fierce competitive environment, Chinese inbound tourism demand growth in these source markets exhibited greater volatility and a growing discrepancy with the others. Particularly after the global financial crisis, the growth rates of the tourism markets have varied immensely from year to year. The numbers of travellers from the Philippines and Mongolia rapidly reverted to prior numbers and maintained growth thereafter. The number of Mongolian travellers expanded to a post crisis one million from 0.7 million before the crisis; the Philippines is approximately the same. However, after recovering from the global financial crisis, the U.S., Malaysian and

* Corresponding author.
E-mail address: xusun@illinois.edu (J. Wang).

http://dx.doi.org/10.1016/j.tourman.2015.07.005
0261-5177/

Singaporean markets apparently did not rebound to vigorous growth; on the contrary, these markets' rapid growth momentum has stalled. The Korean, Japanese and Russian markets did not return to the pre-crisis level, which has resulted in a loss of more than 0.7 million to one million arrivals to China in volume terms from each market. Therefore, for these markets, recovery is a slow process and may take more than a year to return to or approach the pre-crisis growth rate.

It is exceptionally difficult to establish tourism plans for these various volatile tourist markets with different variation characteristics. In this paper we propose an efficient forecasting model possessing wide application and precise predictions.

Over the last several decades, numerous methods of international tourism demand forecasting have been developed. These methods can be generally classified into causal econometric models and univariate time series models. Causal econometric models attempt to identify and establish relations between tourism demand (as measured by tourist arrivals at a destination) and a set of explanatory factors such as income and exchange rate. Time series approaches are dominated by a naïve, moving average, exponential smoothing and Box–Jenkins models, which have also been applied extensively to identify stochastic components for forecasting future trends of tourism demand. However, causal econometric models are limited by incomplete knowledge of the hypothesized causal structure, and most of the time series models with high-accuracy need a large number of data for forecasting. Therefore, it is often difficult to implement these methods for small-sample data because of high cost considerations.

Successfully forecasting Chinese inbound tourism demand not only needs to find ways of reducing the fluctuation caused by the random interferences of uncertain factors but also needs to overcome the limitation of insufficient data (in fact, the foreign arrival tourism data are available from 1997).

Since it was first introduced by professor Deng in 1982, the grey model based on the grey system theory has been shown to be a concise and effective model to handle these uncertain systems with insufficient data. Another benefit is that the model can predict the trends of uncertain systems accurately based on small-sample observations. The grey forecasting model GM(1,1) is the most popular Grey model that utilizes the first-order differential equation to characterize the variable to be forecast. Because of the simplicity, this model has been successfully applied to various fields in modern times (Hsu, 2003; Lin & Yang, 2003; Zhou, 2006; Erdal Kayacan, Ulutas, & Kaynak, 2010; Ze Zhao, Wang, Zhao, & Su, 2012). Although the GM(1,1) had been widely studied, as a simple first-order one-variable grey model, the GM(1,1) may have worse curve-fitting effects for data showing great randomness (Cheng, Hsu, & Wu, 1997). Therefore, it is necessary to proffer new methods to improve its performance, particularly its forecasting accuracy. For example, Akay (2007) applied the Rolling Mechanism to improve upon the accuracy of the GM(1,1); Chang, Lai, and Yu (2005) proposed an available P value rolling Grey model and obtained greater result accuracy; Tzu-Li Tien (2009) presented the FGM(1,1) model, which includes the first-entry's messages. Chen, Ying, and Pan (2010) applied the nonlinear grey Bernoulli model (NGBM) to forecast tourist arrivals. Xie (2009) studied a Discrete Grey forecasting model.

Alternatively, as discussed in the literature, the MCGM(1,1) model, which combines the GM(1,1) and the Markov chain models, can yield better statistical analysis. By capturing the randomness that affects the performance of structures caused by uncertainties in applied stresses and the inherent uncertainty of the deterioration process, the MCGM(1,1) as a stochastic model has the advantage over the GM(1,1) in the significantly fluctuated sample data (Tan, C. L., & Lu, B. F.,1996; Lin, Su, & Hsu, 2001; G.Morcous & Lounis, 2005; Shen et al., 2005; YH. Lin & PL. Lee., 2007; Huang, He, & Cen, 2007; Li G.D., Yamaguchi, & Nagai, 2007a, b, Li, Wang, Masuda, & Nagai, 2011, 2013; Hsu, Liu, Yeh, & Hung, 2009; Kumar U& Jain VK. 2010; Mao Zhan-li & Sun Jin-hua, 2011; Xie, Yuan, & Yang, 2015). This combined model uses the GM(1,1) model to capture the basic trend of the original data, and further applies the Markov chain model to modify the residual error of the GM(1,1) through working on the discretized states based on a residual sequence. Because the Markov chain residual modified model was introduced, the random fluctuation, which affects the forecasting precision of GM(1,1), is reduced. Recently, there has been much new development in the pursuit of higher forecasting accuracy for the MCGM(1,1) model; for example, Li G.D. et al. (2007a, b, 2013) proposed the T-MCGM(1,1) model and the T-MC-RGM(1, 1) model as the improved versions of the MCGM(1,1), and the experimental results indicated that the MCGM(1,1) model has higher forecast precision and excellent applicability. As a mature model with strong performance on random fluctuating data, the MCGM(1,1) also has wide applications in tourism demand; for instance, Wang (2004) used it to forecast tourist arrivals to Taiwan from the United States, Hong Kong and Germany.

In recent years, Artificial Intelligence (AI) techniques have been frequently used to improve the optimization of the grey model and its modified models. Hsu (2003, 2009) proposed an improved grey forecasting model and applied an artificial intelligence technique that integrates the residual modification method based on an artificial neural network (ANN) sign estimation. Yao and Chi (2004) employed a Taguchi method to optimize the parameter settings to serve an energy management purpose. Chang and Tsai (2008) improved and constructed the grey model based on a support vector machine (SVM) approach. Wang and Hsu (2008) forecasted Taiwan's high technology industry output trends by combining the grey system and the genetic algorithms (GA). Ze Zhao et al. (2012) used a grey forecasting model optimized by a differential evolution algorithm to forecast the per capita annual net income of rural households in China. Jianzhou Wang, Zhu, Zhao, and Zhu et al. (2011), Zhilong Wang,Liu, Wu, and Wang et al. (2012, 2014) presented an optimal parameters estimation for the grey model based on a chaotic particle swarm optimization algorithm (CPSO).

Recently, several works on an alternative AI technique known as the Cuckoo search (CS) algorithm have been attracting considerable interests. The CS algorithm is a meta-heuristic algorithm developed by Yang & Deb, 2009. The CS algorithm exhibits several advantages over other optimization techniques, such as greater robustness, faster convergence, higher efficiency, fewer tuning parameters required (Gandomi, Yang, & Alavi, 2013a) and greater ease of implementation. There are certain studies on employing the Cuckoo search algorithm to solve different types of optimization problems in the literature (Bhargava V., Fateen, & Bonilla-Petriciolet, 2013; Gandomi, et al., 2013a, Gandomi, Yang, & Alavi, 2013b; Qin, S., Wang, & Sun, 2014; Bhandari AK., Singh, Kumar, & Singh, 2014); however, from our knowledge, there are no research papers published on the application of the CS algorithm in MCGM(1,1) model for improving the forecasting accuracy. This article will be the first research to apply the CS algorithm to optimize the MCGM model parameters and search the optimal input subset to solve the forecasting question regarding Chinese inbound tourism demand.

The remainder of this paper is organized as follows. Section 2 describes the conventional MCGM(1, 1) model, including the basic GM(1,1) model and the Markov-chain residual modified model. Section 3 presents a novel method including an optimal input subset of the GM(1,1) model and using the Cuckoo search algorithm to optimize the Markov chain. Section 4 reports the experimental results for tourism arrivals to China and the compared results. Finally, the conclusions and remarks are presented in Section 5.