Title

Short-term traffic flow uncertainty prediction using an improved grey prediction model under different time intervals

Zhanguo Song[1,2] Xiao Qin[2]
1 School of Transportation, Southeast University, Nanjing, Jiangsu, China
2 Department of Civil and Environmental Engineering, University of Wisconsin-Miawaukee, Milwaukee, USA

Short-term traffic flow prediction, including level and uncertainty prediction, is a key component of proactive traffic control in the intelligent transportation systems (ITS). In particular, predicting traffic flow uncertainties can provide more valuable information for traffic managers to make reasonable decisions than predicting traffic levels. In this study, an improved grey prediction model was applied to predict traffic flow uncertainties by integrating both smooth pre-processing and background value construction with the traditional grey prediction model. In addition, a spectrum of data collection time intervals was tested to determine the model-specific time intervals for measuring the impact of time interval length on traffic forecasting.

Keywords

Intelligent transportation systems; short-term uncertainty prediction; data collection time interval; grey prediction model.

Introduction

Accurate prediction of short-term traffic flow is important for proactive traffic control systems [1]. Short-term traffic prediction includes both level prediction and uncertainty prediction; and uncertainty prediction of traffic flow can be more valuable for informed decision-making than level prediction [2]. Past literature shows that the studies of traffic uncertainty prediction primarily rely on linear stochastic time series models for confidence interval estimation, such as the Bootstrap method [3-4], the Stochastic Volatility method [5-6], and the GARCH model [7-8]. On the other hand, the grey theory proposed by Deng [9] offers an alternative approach to predicting traffic flow uncertainties. It is now one of the most widely used methods for constructing the interval range of time series data [10-12].

In the process of short-term traffic prediction, data collection interval serves as both the aggregation interval of traffic flow and the prediction horizon for one-step-ahead prediction [13]. The accuracy of traffic prediction results depends on data collection time intervals [14] and different applications require different data collection time intervals. Understanding the impact of data collection time interval on short-term traffic prediction can provide insights into model performance and application.

The objective of this study is to investigate the accuracy and validity of uncertainty prediction results under different time intervals by using an improved grey prediction model. Traffic flow data were collected from an urban freeway in Edmonton, Canada. As a comparison, the traditional grey prediction model and fuzzy information granulation method were estimated and compared with the improved grey prediction model. And two indicators, the mean kick-off percentage (KP) and width interval (WI), were used to measure model performance.

Methodology

The traditional grey prediction model can yield a satisfactory prediction accuracy when modeling a monotonic increasing (or decreasing) data sequence [15-17]. However, traffic flow data series are volatile, and it is necessary to improve the grey prediction model to accommodate the volatility. The improved grey prediction model extends the traditional grey model by adding two techniques: smooth pre-processing and background value construction. The modeling procedure is described as follows.

Step 1 Smooth pre-processing

Suppose that the traffic flow sequence is $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$. Set $M = \max\left\{x^{(0)}(t)\right\}$, $m = \min\left\{x^{(0)}(t)\right\}$, and the volatility amplitude can be calculated as *T*=*M*-*m*. The structure of smooth pre-processing is as below.

$$y^{(0)}(t) = \frac{\left[x^{(0)}(t)+T\right]+\left[x^{(0)}(t+1)+T\right]}{4}, t = 1,2,\cdots,n \tag{1}$$

Thus, the sequence $Y^{(0)} = (y^{(0)}(1), y^{(0)}(2), \dots, y^{(0)}(n-1))$ is called smoothness sequence. We can further obtain the accumulation sequence $Y^{(1)} = (y^{(1)}(1), y^{(1)}(2), \cdots, y^{(1)}(n))$ by using the basic first-order accumulated generating operation (1-AGO) structure [18-19]processing is defined as in Eq. (2).

$$y^{(1)}(t) = \sum_{i=1}^{t} y^{(0)}(i), i = 1,2,\cdots,t \tag{2}$$

Step 2 Background value construction

The background value sequence $Z^{(1)} = \left(z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n-1)\right)$ can be yielded based on the $Y^{(1)}$ by using the background value construction of three-parameter to alleviate the volatility by the Eq. (3).

$$z^{(1)}(t) = \frac{y^{(1)}(t)+y^{(1)}(t-1)+y^{(1)}(t-2)}{3}, t = 3,4,\cdots,n \tag{3}$$

Step 3 Model parameter calibration

The basic form of improved grey prediction model is given as.

$$y^{(0)}(t) + az^{(1)}(t) = tb + c \tag{4}$$

Substituting Eq. (3) to Eq. (4), the improved grey prediction model is given as.

$$y^{(0)}(t) + \frac{1}{3}a\left(y^{(1)}(t) + y^{(1)}(t-1) + y^{(1)}(t-2)\right) = tb + c \tag{5}$$

Where a, b, c is coefficients of improved grey prediction model. And the coefficients are estimated by using the least squares estimate method [20].

Substituting $y^{(0)}(t) = y^{(1)}(t) - y^{(1)}(t-1)$to Eq. (5), we can obtain.

$$tb + c - \frac{1}{3}a\left(y^{(1)}(t) + y^{(1)}(t-1) + y^{(1)}(t-2)\right) = y^{(1)}(t) - y^{(1)}(t-1) \tag{6}$$

Then, the traffic flow prediction sequence can be yielded, which be referred to as $\hat{Y}^{(1)} = (\hat{y}^{(1)}(1), \dots, \hat{y}^{(1)}(t))$.

$$\begin{cases} \hat{y}^{(1)}(t) = \frac{3-a}{3+a}\hat{y}^{(1)}(t-1) - \frac{a}{3+a}\hat{y}^{(1)}(t-2) + \frac{3b}{3+a}t + \frac{3c}{3+a} \\ \hat{y}^{(1)}(1) = y^{(0)}(1) \\ \hat{y}^{(1)}(2) = y^{(0)}(2) \end{cases} \tag{7}$$

Step 4 Residual sequences generation

The residual prediction sequence is

$$R^{(0)} = \hat{Y}^{(1)} - Y^{(1)} \tag{8}$$

Step 5 Residual sequence partition and prediction

We use the line of $R^{(0)} = 0$ as the dividing line to divide the residual sequence into two groups. If $r^{(0)}(t) > 0$ (or $r^{(0)}(t) < 0$), then the residual sequence is named upper (or lower) residual sequence, and expresses as $R_{\mathrm{U}}^{(0)}$(or $R_{\mathrm{L}}^{(0)}$). The two groups of residual sequences are predicted by using the improved grey prediction model. The prediction interval of the residual is $\left[\hat{R}_{\mathrm{U}}^{(0)}(t) \quad \hat{R}_{\mathrm{L}}^{(0)}(t)\right]$.

Step 6 Uncertainty prediction results generation

The forecast interval of the proposed model is created as follows:

$$I = \left[\left(\hat{y}^{(1)}(t) - \hat{R}_{\mathrm{L}}^{(0)}(t)\right), \left(\hat{y}^{(1)}(t) + \hat{R}_{\mathrm{U}}^{(0)}(t)\right)\right] \tag{9}$$

Results

With the uncertainty prediction results, the validity of improved grey prediction model is demonstrated. It is found that all the generated prediction intervals vary by time of day. And the prediction interval range under different time interval is analyzed. The KP and WI under different time intervals are presented to compare models. Although the three models have different results, their performances are comparable when considering time interval as a key factor of forecasting performance. It is worth noting that the improved grey prediction model for interval prediction is not explicitly built upon statistical distributions, which implies that the confidence level cannot be associated with the prediction interval.

Conclusions

An improved grey prediction model was applied to quantify the uncertainty of short-term traffic flow. The uncertainty prediction results and model performance were calculated and analyzed for three different models under different time intervals. The main contributions of the study are: a) an improved grey prediction model that considers the volatility of traffic data; and b) the successful construction of uncertainty prediction results under different time intervals. However, the proposed model does not have an on-line forecasting capability and the KP and WI are not uniform performance measures. The future research will be focused on a) new adaptation mechanisms via which the proposed model could yield real-time prediction; and b) further investigation on uniform performance measures to evaluate the uncertainty prediction.

Acknowledgement

This research is funded and supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX18_0151) and the China Scholarship Council (CSC).

References

[1] Papageorgiou M, Diakaki C, Dinopoulou V, Kotsialos A, Wang Y (2003) Review of road traffic control strategies. Proc. IEEE. 91(12): 2043-2067.
[2] Hugosson M. B (2005) Quantifying uncertainties in a national forecasting model. Transp. Res. A 39(6):500-547.
[3] Takemoto S., Yamaaka K., Nishikawa M., Yano Y., Y Takakura (2018) Bootstrap method based estimation of the minimum sample number for obtaining pharmacokinetic parameters in preclinical experiments. J Pharm Sci. 99(4): 2176-2184.
[4] Pei X, Sze NN, Wong, SC, Yao DY (2016) Bootstrap resampling approach to disaggregate analysis of road crashes in Hong Kong. Accident Anal Prev. S1(95):512-520.
[5] Tsekeris T and Stathopoulos A (2009) Short-Term Prediction of Urban Traffic Variability: Stochastic Volatility Modeling Approach. J Transp Eng-ASCE. 136(7):606-613.
[6] Zhang Y, Ali H, and Sun R (2014) A Stochastic Volatility Modeling Approach to Account for Uncertainties in Travel Time Reliability Forecasting. Transp Res Rec, 2442:62-70.
[7] Lin L, Handley, John C, Gu, YM (2018) Quantifying uncertainty in short-term traffic prediction and its application to optimal staffing plan development. Transp Res Part C: Emerging Technologies, 92:323-348.
[8] Ding C, Duan JX, Zhang YR, Wu XK, Yu GZ (2018) Using an ARIMA-GARCH Modeling Approach to Improve Subway Short-Term Ridership Forecasting Accounting for Dynamic Volatility. IEEE Trans Intel Trans Syst, 19(4)1-11.
[9] J.L. Deng (1982) Control problems of grey systems. Syst. Control Lett. 1(5): 288-294.
[10] Wang ZX, Hipel KW, Wang Q, He SW (2011) An optimized NGBM(1,1) model for forecasting the qualified discharge rate of industrial wastewater in china. Appl Math Model. 35(12):5524-5532.
[11] Morita H, Kase T, Tamura Y, Lwamoto S (1996) Interval prediction of annual maximum demand using grey dynamic model. Electrical Power & Systems. (18)7:409-413.
[12] Ye J, Dang YG, Ding S, Yang YJ (2019) Nonlinear Fractional Order Grey Model of Urban Traffic Flow Short-Term Prediction. J Clean Prod. 229:256-267
[13] Tang RT, Kanamori R, Yamamoto T (2019) Short-Term Urban Link Travel Time Prediction Using Dynamic Time Warping With Disaggregate Probe Data. IEEE Access. 7:98959-98970.
[14] Song ZG, Guo YY, Wu Yao, Ma J (2019) Short-term traffic speed prediction under different data collection time intervals using a SARIMA-SDGM hybrid prediction model. Plos one. 14(6):1-19.
[15] Mao S, Chen Y, Xiao XP (2012) City traffic flow prediction based on improved GM(1,1) model. J. Grey Syst. 24(4):337-346.
[16] Xia M, Wong WK (2014) A seasonal discrete grey forecasting model for fashion retailing. Knowl-Based Syst. 57:119-126.

[17] Xiao XP, Yang JW, Mao SH, Wen JH (2017) An improved seasonal rolling grey forecasting model using a cycle truncation accumulated generating operation for traffic flow. Appl Math Model. 51:386-404.

[18] Liu SF, Yang YJ, Xie NM, Forrest J (2016) New progress of grey system theory in the new millennium. Grey Systems Theory & Application. 6(1):2-31.

[19] Duan HM, Xiao XP (2019) A Multimode Dynamic Short-Term Traffic Flow Grey Prediction Model of High-Dimension Tensors. Complexity. Doi: 10.1155 / 2019 / 9162163.

[20] Mao SH, Xiao XP, Gao MY, Wang XP, He Q (2018) Nonlinear Fractional Order Grey Model of Urban Traffic Flow Short-Term Prediction. J Grey Syst. 30(4):1-17.