Journal of Transportation Technologies, 2014, 4, 256-266
Published Online July 2014 in SciRes. http://www.scirp.org/journal/jtts
http://dx.doi.org/10.4236/jtts.2014.43023

Real-Time Road Traffic Anomaly Detection

Jamal Raiyn[1], Tomer Toledo[2]

[1]Faculty of Exact Science, Computer Science Department, Al Qasemi Acedemic College, Baqa Al Qarbiah, Israel
[2]Faculty of Civil and Environmental Engineering, Department of Transportation Engineering, Technion, Haifa, Israel
Email: raiyn@qsm.ac.il, toledo@technion.ac.il

Received 14 February 2014; revised 12 March 2014; accepted 7 April 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

Many modeling approaches have been proposed to help forecast and detect incidents. Accident has received the most attention from researchers due to its impacts economically. The traffic congestion costs billions of dollars to economy. The main reasons of major percentage of traffic congestion are the incidents. Road accidents continue to increase in digital age. There are many reasons for road accidents. This paper will discuss and introduce new algorithm for road accident detection. Various forecast schemes have been proposed to manage the traffic data. In this paper we will introduce road accident detection scheme based on improved exponential moving average. The proposed traffic incident detection algorithm is based on the automatic exponential moving average scheme. The detection algorithm is based on analyzing the collected traffic flow parameters. The detection algorithm is based on analyzing the collected traffic flow parameters. In addition a real-time accident forecast model was developed based on short-term variation of traffic flow characteristics.

Keywords

Anomaly Traffic, Detection Scheme, Moving Average, Intelligent Transportation System

1. Introduction

The main reason in accidents on the highway can be divided into four categories such as the environment, traffic conditions, vehicles and drivers behavior. Many studies [1]-[4] showed that higher speeds did not lead to serious accidents. On the other hand, some studies showed that fatal accidents increased with high speed limits. Our analysis reveals that the major factor leading to an accident is not speed itself but the variation of speed. There are three basic strategies to relieve congestion [5]: The first strategy is to increase the transportation infrastructure. However this strategy is very expensive and can only be accomplished in the long term. The second strategy is to limit the traffic demand or make traveling more expensive, which will be strongly disapproved of by

How to cite this paper: Raiyn, J. and Toledo, T. (2014) Real-Time Road Traffic Anomaly Detection. *Journal of Transportation Technologies*, **4**, 256-266. http://dx.doi.org/10.4236/jtts.2014.43023

travelers. The third strategy is to focus on efficient and intelligent utilization of the existing transportation infrastructures. This strategy is a best trade-off and gains more and more attention. Currently, the Intelligent Transportation System (ITS) is the most promising approach to implementation of the third strategy. Various forecast schemes [6]-[9] have been proposed to manage the travel flow information. Meanwhile the robustness and accuracy of the exponential smoothing forecast are high and impressive. This paper reports on the performance of three moving average techniques in predicting average travel speeds up to 10 minutes ahead of time. The advantage of the exponential smoothing algorithm is simple. However its forecast precision is not high. If a high forecast precision is requested, it is necessary to consider the real-time information includes the non-conditions events. This paper introduces road accident detection scheme. Road accident detection scheme is focused on real-time information. The real-time information has been achieved to update the historical adaptive information.

To optimize the detection algorithm we have collected travel data by the mobile phone. For a successful forecast of traffic flow, it ought to apperceive the variety of environment and can adjust the parameters automatically. Furthermore it is important that the forecast model takes into consideration the abnormal conditions that occur in real-time [4] [10] [11].

The paper is organized as fellow: Section 2 describes the methodology of road accidents detection scheme. Section 3 and section 4 discuss the performance analysis of the proposed detection scheme and illustrate the simulation results.

2. Methodology

This section presents a methodology to detect road accidents based on travel time variations. We consider accident during peak periods (*i.e.*, morning or afternoon) and during non-peak periods. The observed traffic data consists of *normal* and *abnormal* (accident) travel data. The abnormal record is at least 30 km/h lower traffic speed than the average speed of all records at the same time on the same day of the week. The threshold of 30km/h is a symbolic value of the smallest speed change that people would consider "abnormal". Threshold determination depends on the travel observation data. Equation (1) will be used to forecast the accident scheme.

$$tt(t+1,k,acci) = \alpha \times tt(t,k,acci) + (1-\alpha) \times EMA(t,k,acci) \tag{1}$$

Alpha can be expressed as follows:

$$\alpha = \frac{1}{1+\left[\dfrac{Var(k)}{E(k)}\right]}$$

where $Var(k)$ is the variance of the expected number of crashes at the reference sites. $E(k)$ is the expected number of crashes at these reference sites.

2.1. Section Mutual Influence

In the real-time forecasting we take into consideration the effect of the upstream (UP) and downstream (DS) as illustrates in Equation (2).

$$tt(t+1,k) = tt^{H}(t+1,k) + \gamma_1 \times \text{desired} + \gamma_2 \times \text{UP} + \gamma_3 \times \text{DS} \tag{2}$$

where

$$\text{desired} = \left[tt^{M}(t,k) - tt^{H}(t,k)\right]$$

$$\text{upstearm} = \left[tt^{M}(t,k-1) - tt^{H}(t,k-1)\right]$$

$$\text{downstream} = \left[tt^{M}(t,k+1) - tt^{H}(t,k+1)\right]$$

$$tt_{\text{abnormal}}(t,k) = tt^{-}(t-1,k) + tt^{+}(t+1,k)$$

$$\delta = \Delta\left(tt\left(t,k\right) - tt\left(t+1,k\right)\right)$$

k is the desired section, $(k - 1)$ is the upstream section, $(k + 1)$ is the downstream section.

Figure 1 and **Figure 2** illustrate the abnormal condition in the up and down stream.

2.2. Accident Detection Strategy

The performance of an incident detection system is determined on two levels: data collection and data processing. Data collection refers to the detection/sense/surveillance technologies that are used to obtain traffic flow data. Data processing refers to the algorithms used for detecting and classifying incidents through analyzing the traffic parameters from detectors or sensors for the purpose of alerting observers of the occurrence, severity, and location of an incident. The hybrid of data collection strategies and data processing methodologies results in a variety of solutions for incident detection. The main task of the proposed accident detection (AD) algorithm is to identify and distinguish different traffic modes in **Table 1**. It depends on an upstream occupation increase and a downstream occupation decrease at the level of loop detector where an incident happened. This algorithm compares a value of a traffic flow parameter with a known value. The algorithm trusts that an upstream occupation will increase and downstream occupation will decrease where an incident happened. In traffic incident detection, a time sequence is used to describe a traffic state. When a current measured value is deviated from the output of the algorithm seriously, the algorithm will think that an incident has occurred. The time sequence analytic algorithms include a moving average algorithm, an exponential smoothing algorithm.

- The accident characterized by temporal variation of speed at fixed road section (location) expressed as the coefficient of variation in speed.
- The spatial variation of speed along road sections expressed as the difference in speed between upstream and downstream location (Q).

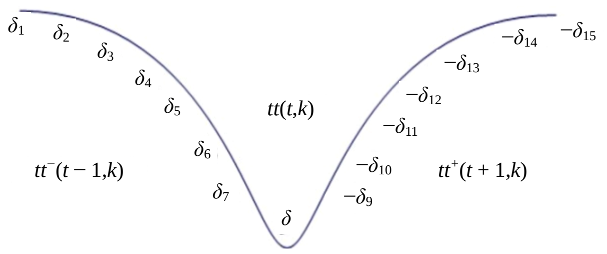

Figure 1. Variation of the travel time speed.

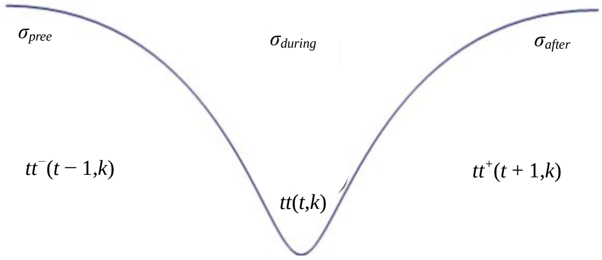

Figure 2. Accident characteristics-stdev.

Table 1. Optimized parameters in AD/NAD.

	AD	no AD
γ	0.9993	0.5346
β_1	0.9081	0.2215
β_2	0.9834	0.1138
β_3	0.8591	0.2315
β_4	0.9993	0.4643

$$Q = \left| t\tilde{t}\left(t, s1\right) - t\tilde{t}\left(t, s2\right) \right| \tag{3}$$

where $t\tilde{t}\left(t, s1\right)$, $t\tilde{t}\left(t, s2\right)$ average speeds computed over period of t upstream and downstream of a road sections, respectively (km/h).

2.3. Incident-Influence Traffic Data

An incident occurring on section i within time interval t is considered to have a significant impact on traffic when traffic measurements from the upstream and downstream stations satisfy the following conditions:

1) The difference between upstream speed *si*, *t* and downstream speed *si* + 1, *t* is greater than the threshold value;

2) The ratio of the difference between the upstream and downstream speeds to the upstream speed (*si*, *t* – *si* + 1, *t*/*si*, *t*, is greater than the threshold value;

3) The ratio of the difference between the upstream and downstream speeds to the downstream speed (*si*, *t* − *si* + 1, *t*)/*si* + 1, *t* is greater than the threshold value.

The abnormal record shows that at least 30 km/h lower traffic speed than the average speed of all records at the same time on the same day of the week. The threshold of 30 km/h is a symbolic value of the smallest speed change that people would consider "abnormal". The vehicle speed starts to decrease in upstream however the speed in downstream starts to increase.

When an incident occurs between stations *k and k* + 1, *the congestion causes a clear difference between* the occupancies of the upstream and the downstream stations as illustrates **Figure 3**.

$$\frac{tt\left(k,t\right) - tt\left(k+1,t\right)}{tt\left(k,t\right)} > \text{threshold} \tag{4}$$

$$\frac{tt\left(k,t\right)) - tt\left(k+1,t\right)}{tt\left(k+1,t\right)} > \text{threshold} \tag{5}$$

$$\text{Mean}\left(\text{accidents}\right) = \frac{1}{N}\sum_{i=1}^{n}\left(\mu - \sigma_i\right) \tag{6}$$

σ standard deviation, N number of the acidents.

2.4. Real-Time Accident Detection

The travel time forecast model considers the incident and non-incident conditions. We make different between:

- Accident during peak time (morning/afternoon);
- Accident during regular time;
- Heavy accident;
- Light accident.

The accident is cleared at current time t in section s, the duration is known and the speed is considered to be 30 km reduced of the average speed.

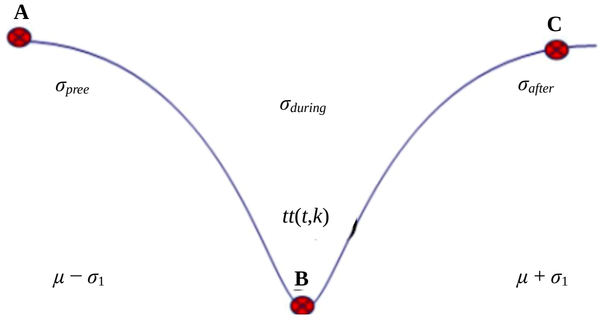

Figure 3. Accident characteristics of StDev.

$$tt(t+1,k) = tt^H(t+1,k) + \gamma \times (P_t) \times (tt_t^M - tt_t^H)$$

$$P_t = P(\text{accident})_t = \frac{1}{1+e^{-\upsilon_t}}, \quad \upsilon_t = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

$$x_1 = \frac{(\sigma_t - \sigma_t^H)}{\sigma_t^H}, \quad x_2 = \frac{(tt_t - tt_t^H)}{\sigma_t^H},$$

$$x_3 = \frac{(tt_t - tt_t^H)}{\sigma_t^H} - \frac{(tt_{t-1} - tt_{t-1}^H)}{\sigma_{t-1}^H}, \quad x_4 = \frac{(\sigma_t - \sigma_t^H)}{\sigma_t^H} - \frac{(\sigma_{t-1} - \sigma_{t-1}^H)}{\sigma_{t-1}^H}$$

where X denotes the vector of predictor variables. β is the vector of coefficient associated with the predictor variables. and can be computed according to the binary logit model. υ_t is the logit link function (which is a linear combination of the predictor variables).

2.5. Accident Probability

Based on statistical measurements of historical information and real information, the forecast model can estimate the occurrence of abnormal conditions without external information as express Equation (7) and Equation (8).

$$tt_{acc}^F(t+1,k) = EMA_{acc}^H(t+1,k) + \delta\left(tt_{acc}^M(t,k) - tt_{acc}^H(t,k)\right) \tag{7}$$

$$tt_{acc}^F(t+1,k) = EMA_{acc}^H(t+1,k) + \underbrace{\delta\left(tt_{acc}^M(t,k) - tt^H(t,k)\right)}_{\text{Corrections needed=epsilon}} \tag{8}$$

where,

$$\varepsilon = \left(\text{mean}\left(EMA_{\text{normal}}(k,t)\right) - \text{mean}\left(EMA_{\text{abnormal}}(k,t)\right)\right)$$

$$P_{\text{accident}}^{\text{hist}}\left(tt(t,k)\right) = P\left\{\sigma\left(tt_i(t,k)\right) > \delta_{th}(\text{hist})\right\} \tag{9}$$

$$P_{\text{accident}}^{\text{real}}\left(tt(t+1,k)\right) = P\left\{\sigma\left(tt_i^M(t,k)\right) > \max_i \sigma\left(tt_i^{\text{hist}}(t,k)\right)\right\} \tag{10}$$

where:

$$\delta_{th2}(M) > \delta_{th1}(\text{hist})$$

$$\sigma\left(tt^M(t,k)\right) > \sigma\left(tt^M(t-1,k)\right)$$

$$tt^M(t,k) > tt^H(t,k)$$

$$tt^M(t,k) > tt^H(t-1,k)$$

The Total number of the expected accident is expressed as following:

$$E\left(N_{\text{accident}}^M\right) = \left(1 - P_{\text{accident}}^{\text{real}}\right) \times n \tag{9}$$

2.6. Smoothed Parameter Optimization

To increase the exponential moving average forecast accuracy in real-time, the smoothed parameter alpha and gamma in Equation (4) should be optimized. **Figure 4** illustrated the value of the optimized smoothed parameter gamma in real-time accident conditions.

Figure 4 and **Figure 5** illustrate values of the optimized smoothed parameter gamma in real-time accident and non-accident conditions in highway. However **Figure 6** and **Figure 7** illustrate values of the optimized smoothed parameter gamma in real-time accident and non-accident conditions in urban road.

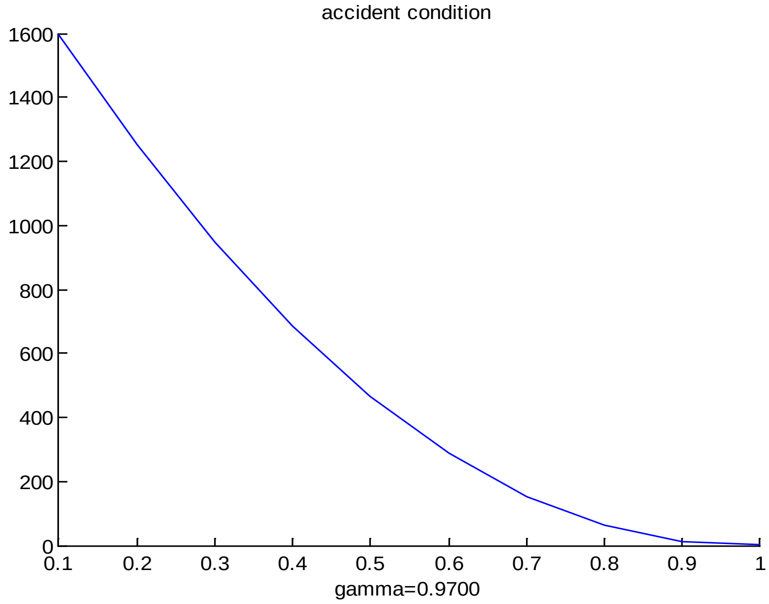

Figure 4. AC in highway.

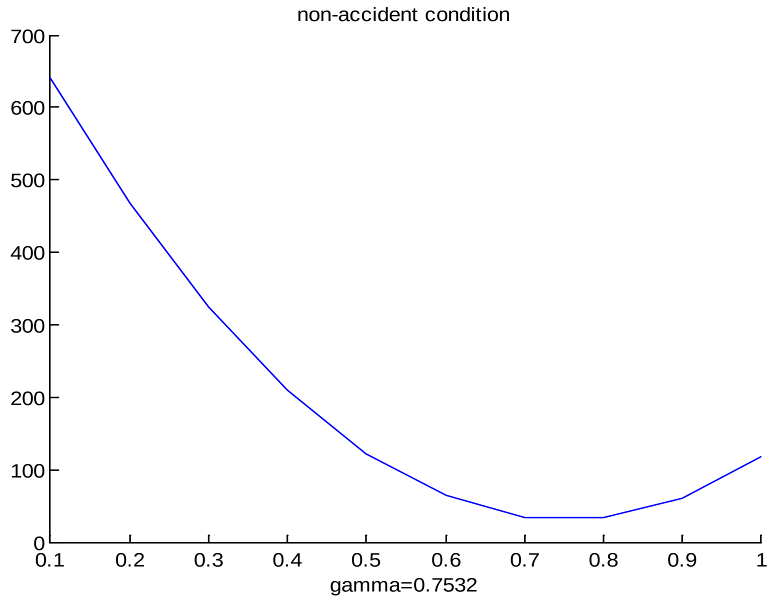

Figure 5. NAC in highway.

3. Performance Analysis

There are various measures of forecasting accuracy techniques proposed in the literature [5] [12]-[15]. The aim of this study is to evaluate forecast accuracy travel observations. The forecasting accuracy techniques are used to be able to select the most accurate forecast scheme. The forecasting performance of the various models and the measures of the predictive effectiveness was evaluated using various summary statistics. The comparing experiments are carried out under normal traffic condition and abnormal traffic condition to evaluate the performance of four main branches of forecasting models on direct travel time data obtained by license plate matching (LPM). The MAE is a measure of overall accuracy that gives an indication of the degree of spread, where all errors are assigned equal weights. The MSE is also a measure of overall accuracy that gives an indication of the degree of spread, but here large errors are given additional weight. It is the most common measure of forecasting accuracy.

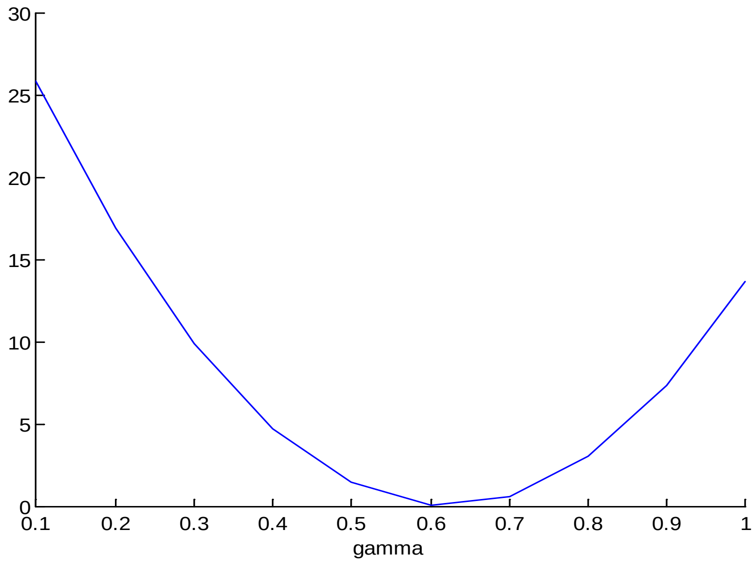

Figure 6. NAC in urban road.

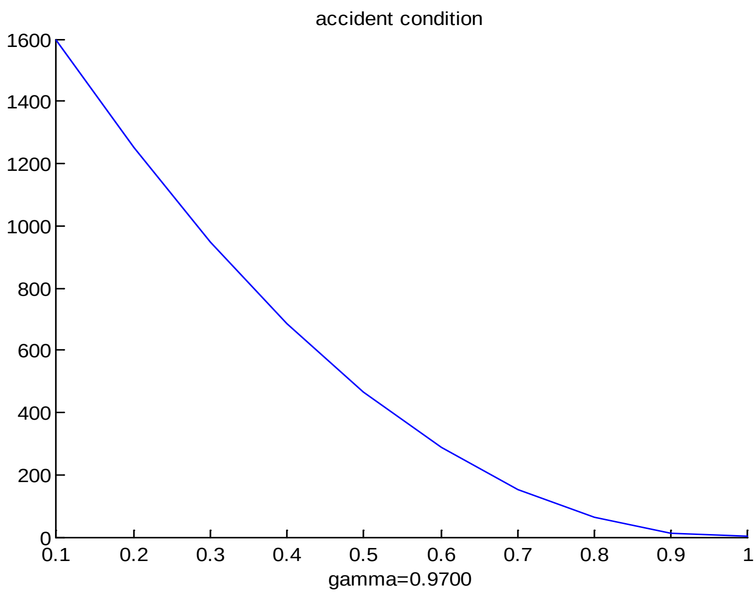

Figure 7. AC in urban road.

Often the square root of the MSE, RMSE, is considered, since the seriousness of the forecast error is then denoted in the same dimensions as the actual and forecast values themselves. Mean square percentage error (MSPE) is the relative measure that corresponds to the MSE. The more commonly used measure is the root mean square percentage error (RMSPE). Theil's Coefficient is another statistical measure of forecast accuracy. One specification of Theil's compares the accuracy of a forecast model to that of a naive model. A Theil's greater than 1.0 indicates that the forecast model is worse than the naïve model; a value less than 1.0 indicates that it is better. The closer U is to 0, the better the model.

4. Simulation Results

The travel observation data consists of *normal* and *abnormal* (accident) travel data. **Figure 8(a)** and **Figure 8(b)** illustrate the abnormal conditions in up and download stream in peak hours. However **Figure 8(c)** illustrates the abnormal condition in no peak hours.

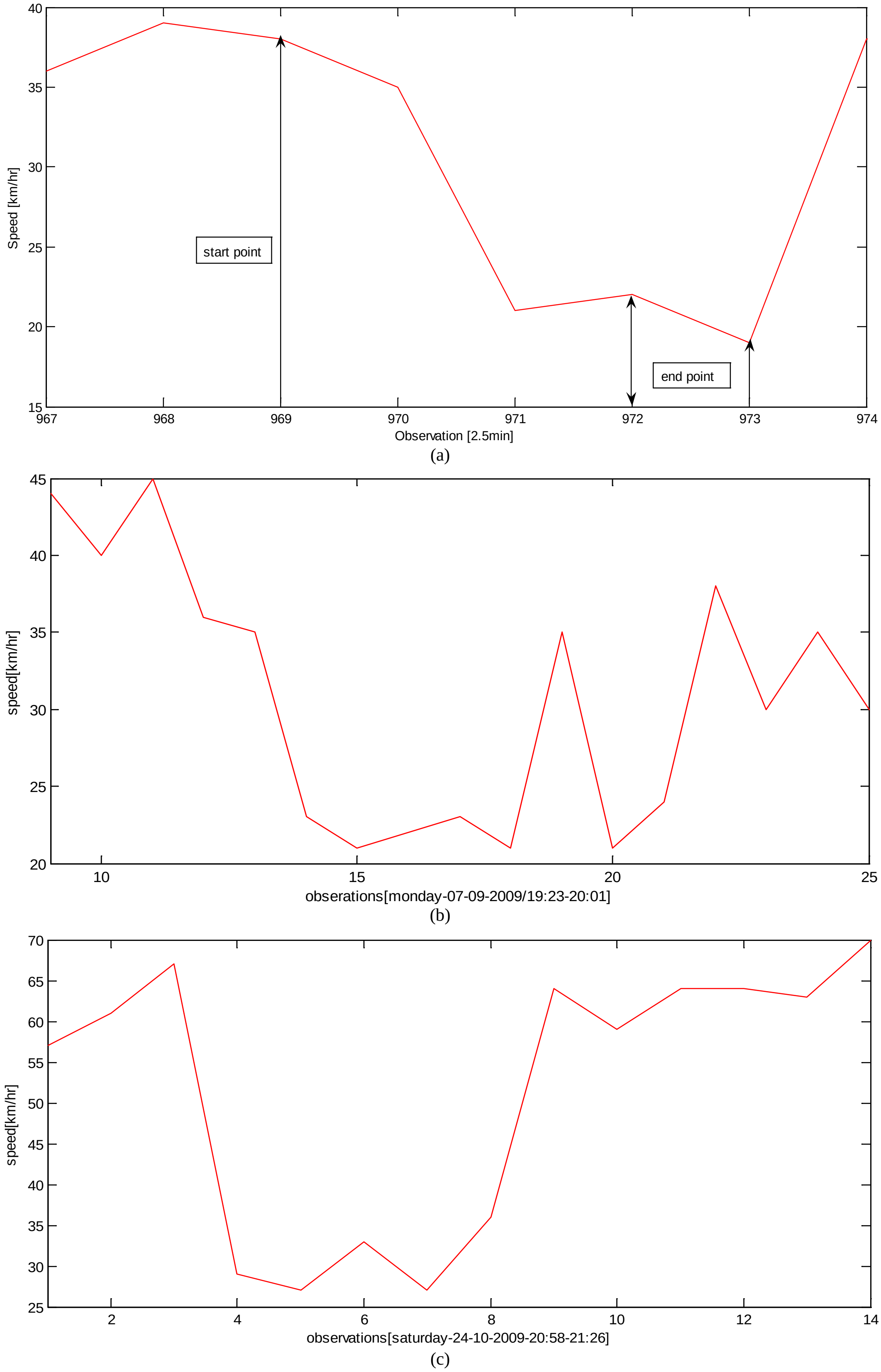

Figure 8. (a) Travel time variation in AC; (b) Travel time variation in AC; (c) Travel time variation in AC.

Table 2 and **Table 3** illustrate the performance analysis of exponential moving average scheme based on historical and real time forecasting. The comparison has been introduced based on accident and non accident conditions.

Table 4 describes the comparison of exponential moving average scheme based on sorted data that the difference between two neighbor observations is bigger than 5 km and 10 km. **Figure 9** illustrates the comparison between exponential moving average and improved exponential moving average.

Table 2. Hist vs. real-time in NAC.

Non-Accident Condition	Hist	Real
mean data	67.805	67.805
mean prediction	65.622	66.798
std data	17.809	17.809
std prediction	18.682	16.968
Observations with error over 5 km/hr	33.086	31.293
Observations with error over 10 km/hr	17.385	15.735
max abs. error	73.39	73.264
max relative error	587.12	586.11
mean error	2.183	1.0076
mean abs. error	6.6768	5.472
mean relative error	12.238	10.562
root mean squared error	12.452	9.2418
root mean squared percent error (1)	26.42	23.514
root mean squared percent error (2)	18.364	13.63
Theil's coefficient	9.0011	6.6476
bias proportion	3.0737	1.1886
variance proportion	0.49122	0.82716
co-variance proportion	96.435	97.984

Table 3. Hist vs. real-time in AC.

Accident Condition	Hist	Real
mean data	79.234	79.234
mean prediction	75.324	78.981
std data	17.737	17.737
std prediction	22.993	16.673
Observations with error over 5 km/hr	42.206	40.281
Observations with error over 10 km/hr	26.006	22.356
max abs. error	93.492	81.288
max relative error	1181.7	4538.8
mean error	3.9104	0.25324
mean abs. error	10.588	7.4191
mean relative error	16.743	12.656
root mean squared error	20.505	12.118
root mean squared percent error (1)	39.798	32.049
root mean squared percent error (2)	25.88	15.294
Theil's coefficient	12.82	7.4844
bias proportion	3.6367	0.04367
variance proportion	6.57	0.77046
co-variance proportion	89.793	99.186

Table 4. Up- and downstream effect.

Real-time	EMA	Speed > 5 km	Speed > 10 km
mean data	69.276	58.275	56.729
mean prediction	55.884	49.321	49.265
std data	23.449	21.738	20.014
std prediction	22.77	4.1422	2.2321
Observations with error over 5 km/hr	92.415	85.318	83.302
Observations with error over 10 km/hr	78.546	71.246	67.392
max abs. error	86.853	69.984	64.842
max relative error	408.87	520.96	527.45
mean error	13.392	8.9542	7.4641
mean abs. error	15.299	19.812	17.62
mean relative error	25.257	38.407	35.075
root mean squared error	18.179	23.746	21.341
root mean squared percent error (1)	32.904	54.475	50.779
root mean squared percent error (2)	26.242	40.747	37.619
Theil's coefficient	13.619	21.26	19.495
bias proportion	54.27	14.219	12.233
variance proportion	0.13941	54.909	69.425
co-variance proportion	45.591	30.872	18.342

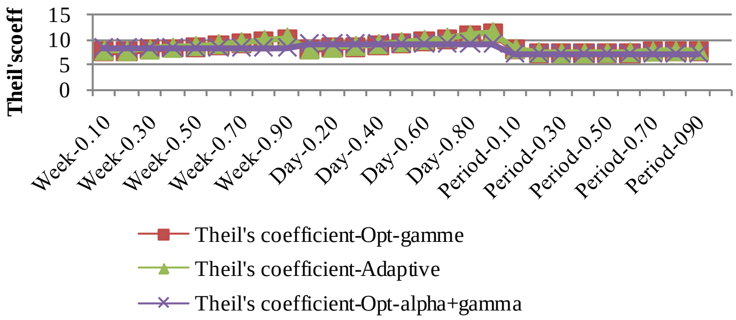

Figure 9. EMA vs. improved EMA.

4. Conclusion

Analysis of the road incidents based on the speed variation is not robust enough to develop real-time forecast model. Because a speed observation can be zero when there is *no vehicle*, *or the system collects a wrong speed observation*, in this case, the computation of CVS can be done in many variations.

References

[1] Ronen, B., Coman, A. and Schragenheim, E. (2001) Peak Management. *International Journal of Production Research*, **39**, 3183-3193. http://dx.doi.org/10.1080/00207540110054588

[2] Tu, H., Van Lint, H. and Van Zuylen, H. (2008) The Effects of Traffic Accidents on Travel Time Reliability. *IEEE Conference on Intelligent Transportation Systems*, Beijing, 12-15 October 2008.

[3] Wang, Z. and Murray-Tuite, P. (2010) Modeling Incident-Related Traffic and Estimating Travel Time with a Cellular Automaton Model. *Proceedings of Transportation Research Board's* 89*th Annual Meeting CD-ROM*, 10-14 January 2010, Washington, DC.

[4] Wild, D. (1997) Short-Term Forecasting Based on a Transformation and Classification of Traffic Volume Time Series. *International Journal of Forecasting*, **13**, 63-72. http://dx.doi.org/10.1016/S0169-2070(96)00701-7

[5] Zheng, X. and Liu, M. (2009) An Overview of Accident Forecasting Methodologies. *Journal of Loss Prevention in the Process Industries*, **22**, 484-491. http://dx.doi.org/10.1016/j.jlp.2009.03.005

[6] Andrada-Felix, J. and Fernandez-Rodriguez, F. (2008) Improving Moving Average Trading Rules with Boosting and Statistical Learning Methods. *Journal of Forecasting*, **27**, 433-449. http://dx.doi.org/10.1002/for.1068

[7] Guin, A. (2006) Travel Time Prediction Using a Seasonal Autoregressive Integrated Moving Average Time Series Mode. *Proceedings of the IEEE Intelligent Transportation Systems Conference*, Toronto, 17-20 September 2006, 493-498.

[8] Lv, Y. and Tang, S. (2010) Real-time Highway Traffic Accident Prediction Based on the K-Nearest Neighbor Method. *International Conference on Measuring Technology and Mechatronics Automation*, Volume 3, 547-550.

[9] Xia, J (2010) Predicting Freeway Travel Time under Incident Condition. *Proceedings of Transportation Research Board's* 89*th Annual Meeting CD-ROM*, 10-14 January 2010, Washington, DC.

[10] Alger, M. (2004) Real-Time Traffic Monitoring Using Mobile Phone Data. *Proceedings of* 49*th European Study European Study Group with Industry*, Oxford, United Kingdom.

[11] Stephanedes, Y.J., Michalopoulos, P.G. and Plum, R.A. (1981) Improved Estimation of Traffic Flow for Real-Time Control. *Transportation Research Record*, **795**, 28-39.

[12] Jo, H., Lee, B., Na, Y.-C., Lee, H. and Oh, B. (2007) Prioritized Traffic Information Delivery Based on Historical Data Analysis. *Proceedings of the* 2007 *IEEE Intelligence Transportation Systems Conference*, Seattle, September 30-October 3 2007, 568-573.

[13] Karim, A. and Adeli, H. (2003) Fast Automatic Incident Detection on Urban and Rural Freeways Using Wavelet Energy Algorithm. *Journal of Transportation Engineering*, **129**, 57-68. http://dx.doi.org/10.1061/(ASCE)0733-947X(2003)129:1(57)

[14] Lee, H., Chowdhury, K.N. and Chang J. (2008) A New Travel Time Prediction Method for Intelligent Transportation Systems. Springer-Verlag, Berlin, 473-483.

[15] Xiaoqiang, Z., Ruimin, L. and Xinxin, Y. (2010) Incident Duration Model on Urban Freeways Based on Classification and Regression Tree. 2*nd International Conference on Intelligent Computation Technology and Automation*, *TRB* 2010 *Annual Meeting*, **2**, 526-528.

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.