

A REVIEW ON USING CHAOS THEORY TO PREDICT DAILY TRAFFIC

^{1} Sujit Kumar Rout, ²Biswajit Padhi*
^{1} Professor, Dept. of Civil Engineering, NIT BBSR,*
Asst. Professor Dept. of Civil Engineering, MIT, BBSR
^{1} sujitrout@thenalanda.com, padhibiswajit@gmail.com*

Abstract

Intercity road traffic volumes are typically calculated using probability functions, statistical methods, or meta-heuristic methods like artificial neural networks. Road traffic volumes are calculated using pattern recognition techniques since they depend on input variables such as the geometry of the road, the time of day or night, weekends or holidays, and other factors. The major goal of this research project is to evaluate how well chaos theory performs when used to estimate daily traffic volume utilising chaotic patterns. In this study, the daily traffic volume on intercity routes is analysed for chaotic behaviour, and the effectiveness of the chaos theory is reviewed and contrasted with probability functions. Data collected over the course of a year by installed automatic traffic counters were utilised in the analytical process to establish chaos factor as the ratio between the minimum and maximum daily traffic volume. The findings showed that daily traffic levels exhibit chaotic behaviour over a specified twenty-four-hour period. They also demonstrate that, for projecting daily traffic volume, the application of chaos theory is stronger than the uniform distribution function and inferior to the normal distribution function.

Keywords: Chaos theory, traffic volume estimation, probability function, pattern recognition tech-niques

1. Introduction

One of the most crucial considerations for traffic managers is estimating the volume of traffic, especially for bettering enforcement and management. For pavement design, fuel-tax revenue projections, and highway safety planning, data collection and daily traffic volume predictions are typically used. However, the expenses and personnel required for the monitoring efforts required for precise analytical average daily traffic (AADT) estimates are high [Rossi et al., 2012]. Traffic estimation is crucial for both tactical transportation needs like enforcement [Ashok, and Ben-Akiva, 2002] and for predicting the number and severity of road accidents [Ardekani, Hauer, and Jamei, 1995]. Transport authorities also put a lot of effort into estimating the volume of traffic on the

roads in order to develop network-wide plans for enhancing enforcement and putting in place road safety measures. Traffic estimation has grown to be an increasingly important issue in road transport management in the era of adopting intelligent transportation systems developed around the globe to increase transportation efficiency [Caggiani et al. 2012]. While determining the origin-destination matrix is thought of as a fundamental element of transport planning, estimating road traffic volumes has grown in importance [Caggiani et al. 2012].

Many models and methodologies utilized for estimating traffic volumes are observed in the literature. The main techniques including Regression Analysis [Faghri and Hua, 1995], Ward's Minimum-Variance method of clustering [Sharma and Werner, 1981], Clustering-based methods [Zhao, Li and Chow, 2004], and some heuristic methods of Genetic Algorithms [Lingras, 2001] and Artificial Neural Networks [Faghri and Hua, 1995, Lingras, 2001, Mahmoudabadi and Fakharian, 2010] have been implemented for road traffic estimation. Simulation techniques [Juran et al., 2009] and data mining [Gecchele et al. 2011] which are usually developed through using the probability distribution functions are also applied for traffic volume estimation. Variation in day-to-day traffic volumes is a very important factor in the process of estimating daily traffic [Bodley, 1976].

On one hand, looking at the methods or techniques utilized for traffic estimation in the previous studies in detail send a notation to researchers that traffic volumes follow dynamic patterns. On the other hand, chaos theory is known as a dynamic system, that will be discussed more in the second section, has been previously used for traffic or transport indicators. Therefore, it may be considered as a major issue in road traffic modeling and proposing a dynamic pattern based on the concept of chaos theory that may be a useful technique to develop a chaotic pattern for estimating road traffic volumes. Consequently,

determining the performance of using chaos theory may convince researchers to develop chaotic patterns for road traffic estimation. Following the above mentioned issues, the basic principle and the core of novelty of this research work is to consider daily traffic volume as a chaotic variable and develop an appropriate model to predict daily traffic volume. While an appropriate chaos factor related to daily traffic volume is defined in the proposed methodology, the presence of chaos as a major concern for developing methodology will also be checked. It is assumed that, daily traffic volume is estimated in short-time prediction and the above concept may be useful for researchers and those who are interested to know whether chaos

theory is a proper technique to estimate daily traffic or not. Depend-ency upon initial condition is also considered as another concern to develop chaotic patterns [Mahmoudabadi and Seyedhosseini, 2012], so traffic volumes are being estimated in short- time prediction for traffic management.

This paper is divided into seven sections. Af- ter introducing a short description about chaos theory, the paper is followed by methodology definition, where details of application of cha- os theory for daily traffic estimation are pre- sented. Details on case studies and analytical process are demonstrated after the above sec- tions, in sections 4 and 5, respectively. Discus- sions on rating the performances of three ap- proaches are presented in section 6, followed by conclusion and some recommendations for further researches in the last section.

2. Chaos Theory

The concept of chaos theory was firstly in- troduced by Edvard Lorenz in 1963 [Gleick, 1987]. Chaos theory is known as a non-linear system, while in a chaotic nonlinear system, small differences in the initial conditions oc- casionally produce very great ones in the fi- nal outputs [Xh, Yb and Zhang, 2002]. It has been studied within the engineering scientific and mathematical communities and found to be useful in many disciplines such as high- performance circuits and devices, collapse prevention of power systems and also infor- mation processing [Mingjun and Huanwen, 2004]. Chaos theory has also been widely ap- plied in various fields of science, particularly in the area of traffic flow theory, commonly in short-time prediction, because of the exist- ence of the property of “sensitive dependence upon initial condition” [Lawrence, Feng and Huang, 2003].

In transportation systems, legal and social constraints may bind behavior, allowing re- searchers to predict human actions and their effects on system evolution, more accurately [Frazier and Kockelman, 2004, Sugihara and May, 1990]. In safety planning programming, it has also been approved that traffic measure such as the number of road accidents can be predictable under the property of chaotic be- havior [Mahmoudabadi and Seyedhosseini, 2012].

The largest Lyapunov exponent, usually indi- cated as λ_{max} in the literature, is the clearest measure to prove the existence and to quan- tify chaos in a dynamic system or time series. Calculating the

largest Lyapunov exponent is a more common technique to determine the presence of chaos that measures the convergence or divergence of the nearby trajectories [Frazier and Kockelman, 2004]. As the system evolves, the sum of series of values (in each dimension) will converge or diverge. Lyapunov exponents measure the rate of convergence or divergence in each dimension. Thus, if the largest Lyapunov exponent exceeds zero, the system is chaotic [Frazier and Kockelman, 2004]. In other words, if the largest Lyapunov exponent is positive, it indicates that the system under investigation is sensitive to the initial condition and it is chaotic (Kiel and Elliott, 1996). Equation (1) is used to determine the largest Lyapunov exponent, where $S(t)$ and $S'(t)$ are the system situation and its nearest neighbor in time t , respectively. N is the number of available data points and Δt is time step or time span. Since the scientific background of chaos theory is related to mathematical science, readers interested to know more about chaos theory are recommended to refer to [Gleick, 1987]. In this research work, the nearest neighbor for each situation is considered as the situation on the previous time,

i.e. $S'(t) = S(t - \Delta t)$. If $\Delta t=1$ means that time step is considered as a day and $\Delta t=7$ means time step is a week.

$$\lambda = \frac{1}{N\Delta t} \sum_{(t=0)}^{(N-1)} \ln \left(\frac{|S(t+\Delta t) - S'(t+\Delta t)|}{|S(t) - S'(t)|} \right) \quad (1)$$

The well-known equation of logistic map is commonly applied to generate chaotic variables [Sugihara and May, 1990; Trepaniera, Leroux, and Marcellis-Warin, 2009], particularly in traffic behavior, because of good adeptness to time series. The logistic map equation is defined as equation (2) [Lawrence, Feng and Huang, 2003], in which $P(t)$ is the chaotic variable in time 't', and K , is the equation parameter which can be estimated by statistical techniques of mean square errors [Mahmoudabadi and Seyedhosseini, 2012] or pre-defined measure of 4 [Mingjun and Huan-wen, 2004].

$$P(t+1) = K \times P(t) \times (1 - P(t)) \quad (2)$$

3. Methodology Definition

The proposed methodology is based on a prediction procedure following some stages. Experimental data collected by automated electronic systems are used for analytical process. In order to keep data accuracy, those gathered for more than 20 hours in a day, are used in the modeling and evaluation processes. Therefore, at the first stage, a filtering process is done

where data gathered for more than 20 hours a day are used in analytical process. Filtering stage is followed by calculating statistical measures of data in order to use for analytical process. It is necessary to define a proper chaos factor, where results would be easily compared to other methods, so defining a chaos factor, which is the main part of contribution of this paper, is on the next stage, where a chaotic variable is defined in closed interval [0 1]. According to daily traffic volumes (abbreviated as TV), equation (3) is used to define chaos factor in the proposed model. For each part of time, known as (time span) or time period (t) represents a day, daily chaos factor (CF) is updated using logistic map equation defined as equation (4). Following the above steps, estimated daily traffic volume (ETV) for time span (t) will be obtained by equation (5), where the maximum (MaxTVs) and minimum (MinTVs) of traffic volumes have been calculated by using experimental data.

$$CF(t) = (TV(t) - MinTVs) / (MaxTVs - MinTVs) \quad (3)$$

$$CF(t+1) = K \times CF(t) \times (1 - CF(t)) \quad (4)$$

$$ETV(t+1) = MinTVs + (MaxTVs - MinTVs) \times CF(t+1) \quad (5)$$

Two approaches are utilized for validation process, where criteria measures for accuracy are defined based on the whole data and the last five and fifteen days. Checking accuracy for the mentioned categories is related to checking the methodology performance in both long-term and short-term predictions. Checking the performance of proposed methodology and comparing results follows as the final stage. An overall view of the proposed methodology is demonstrated in figure 1.

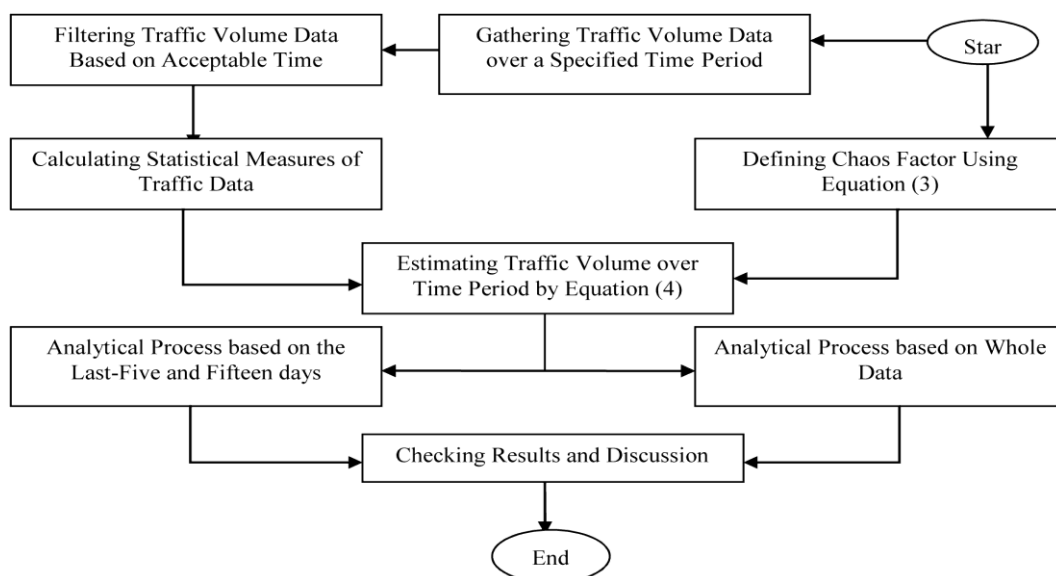


Figure 1. Overall view of the proposed methodology

4. Case Study and Experimental Data Two intercity roads of Qazvin-Rasht and Tabriz-Soufian

have been selected as case studies. Qazvin-Rasht is a freeway, 130 km length and two lanes in each direction. Tabriz-Soufian is a divided road, 50 km length with two lanes in each direction. The difference between the above roads is due to the existence of U-Turns. One of the main issues on selecting the above roads as case studies is to select roads in different topographic areas, where Qazvin-Rasht is located in the mountainous location and Tabriz-Soufian is located in a flat area. The other issue is data availability over a year, while reliable data have been collected by the selected system according to local experts' point of views. There are U-Turn accesses in Tabriz-Soufian road, so Qazvin-Rasht and Tabriz-Soufian are known Freeway and Express ways, respectively, according to national instructions. Automatic traffic counters are installed in the surface of pavement in several parts of the above selected roads which collect data on traffic volumes and also speed detection systems through induction loops. Systems detect the vehicles passing on the surfaces, recognize the axle load patterns and eventually categorize the vehicles according to the specified axle configurations. Installed automatic traffic counters are able to record traffic data whereas each record consists of date, time, location ID, axle configuration pattern and speed, which is automatically calculated according to the times when vehicles are detected at the beginning and end of installed loops, but the pack of data is transferred every 15 minutes which are appropriate to be studied in the various researches. Data over a year started from 21-March 2012 to 20-March 2013, being collected through electronic devices installed for the purpose of traffic counting transformed in the form of daily traffic volumes. The accuracy of traffic counting results is an essential concern, so data can be

collected in daily traffic volume pattern, while those with less than 20 hours of gathering data have not been used for analytical process. It is obvious that the accuracy of validation process will be improved if experimental daily traffic volumes of 24 hours are used in the process of validation, but the number of available data may hamper to achieve the best results. Therefore, the best decision has been made on using daily traffic volumes of available data with more than 20 hours of full collecting data.

5. Analytical Process

Drawing bar charts which represent a schematic pattern of frequency of daily traffic volume over a year is the first step of analytical process. This can help decision makers to imagine an overall perspective for obtaining proper distribution functions. Hence, two bar charts based on the minimum and maximum of traffic volumes in the selected roads are drawn in figures 2 and 3. In addition, statistical measures of two data sets will help decision mak-

ers to develop better distribution functions for all data sets. Daily traffic volumes have been analyzed and statistical measures including minimum, maximum and standard deviation are presented in table 1.

The second step of analytical process is to check the presence of chaos. According to this section of methodology definition, chaos fac-

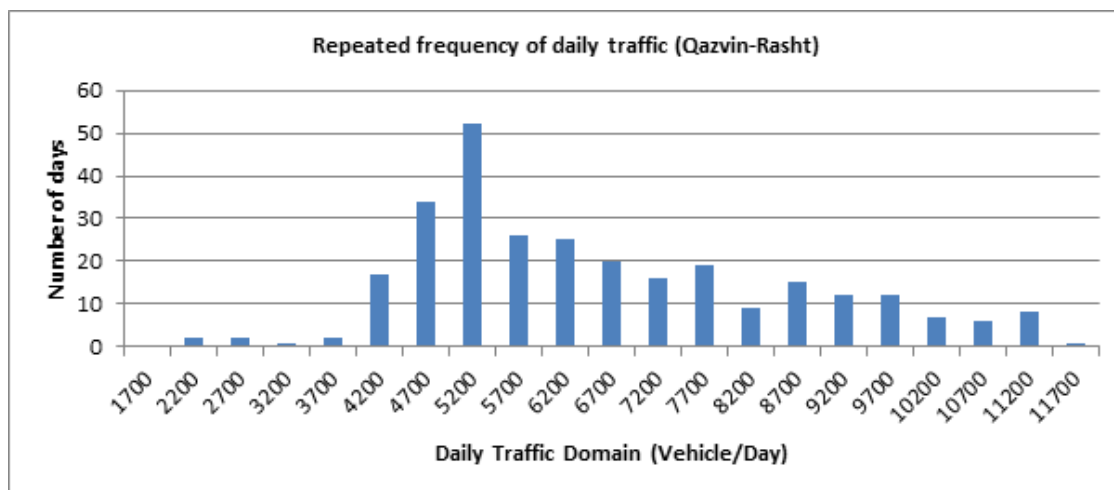


Figure 2. Bar Chart of the Frequency of Daily Traffic Volume for Qazvin-Rasht Freeway

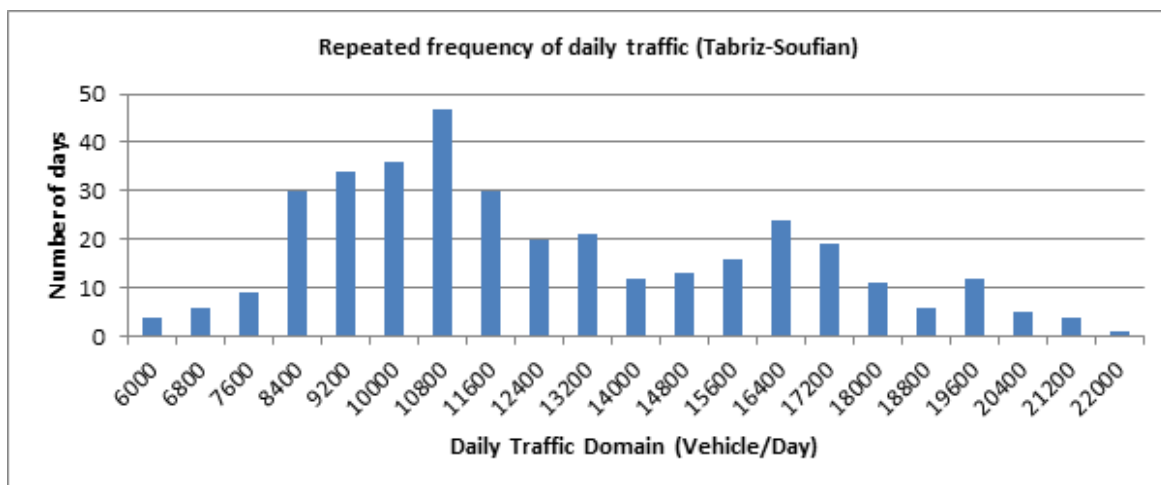


Figure 3. Bar Chart of the Frequency of Daily Traffic Volume for Tabriz-SoufianExpress Way

Table 1. Statistical measures for daily traffic volume

Road	Number of records	Min	Max	Mean	Standard Deviation	Largest Lyapunov Exponent ($\times 10^{-3}$)	
						$\Delta t=1$	$\Delta t=7$
Qazvin-Rasht	346	1765	41045	8599	6006	4.37	-5.96
Tabriz-Soufian	363	2250	22415	12217	3775	6.11	1.24

tor is the ratio between the maximum and the minimum of traffic volumes over a year. Equation (3) is used for calculating chaos factor. In this case the maximum of chaos factor is calculated as 1 for day 73 (with traffic volume 41045) and calculated as 0 for day 283 (with traffic volume 1765) in figure 4 corresponds to data gathered in Qazvin-Rasht. They have also been calculated for days 162 and 249 in figure 5 which corresponds to data gathered in Tabriz-Soufian. The variation of chaos factors for two data sets is drawn and shown in figures 4 and 5 corresponding to Qazvin-Rasht and Tabriz-Soufian, respectively. Numerical

results are tabulated in table 1. As the last columns of table 1 revealed, the behavior of daily traffic volumes for both roads are chaotic, but weekly traffic patterns are not. The main reason is that the weekly patterns are the same for road traffic volumes over a year, because the Largest Lyapunov Exponent for daily traffic patterns are greater than 0 but they are not for weekly traffic patterns.

6. Discussion

In order to recognize the probability function, two histograms for traffic data gathered over a year outlined in the previous section have

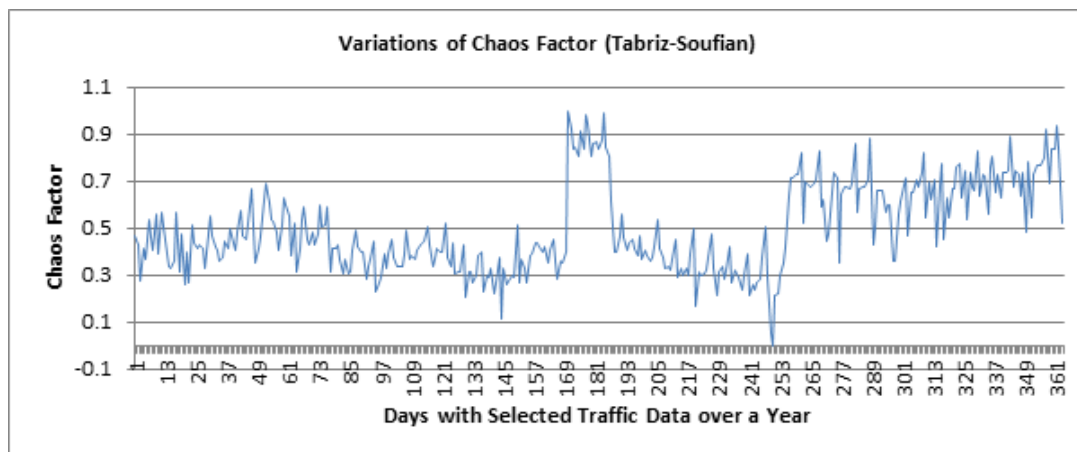
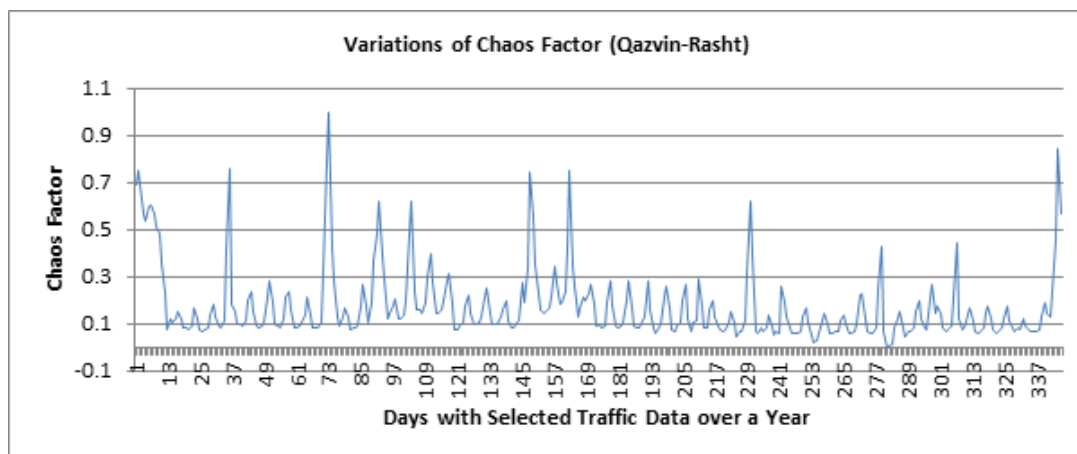


Figure 4. Variation of Chaos Factor for Qazvin-Rasht Freeway over a Year

Figure 5. Variation of Chaos Factor for Tabriz-Soufian Expressway over a Year

been analyzed and it seems that two different distribution functions of normal and uniform may be compatible to traffic data sets and needs to be checked. Therefore, two different probability functions of normal and uniform have been utilized to estimate traffic volume and compare results with using chaos theory applied in this research work. Following that, two criteria measures of mean square errors and difference between the number of observation and model output in each domain [Mahmoudabadi and Ghazizadeh, 2013] are considered for validation. Domains are defined for each data set regarding to minimum and maximum daily traffic volume over a year. The mean square errors corresponding to three utilized time prediction period are shown in table 2. The above periods used in analytical process show the number of daily traffic estimations in which 5 days, 15 days and the whole data set are used to calculate

mean square errors.

As shown in table 3, the best results are related to normal distribution function, where the minimum of mean square errors between

model outputs and observations are obtained. Chaos theory is rated in the second, in which criteria measures are better than uniform distribution function and weaker than normal one.

Results for all domains are also shown in tables 3 and 4. In this method, difference between observations in each domain is compared to the number of output model. Tables revealed that the number of differences in domains have the same conclusion on rating the performances of modeling, whereas the average differences between observations and models' outputs are set in the order of normal, chaos and uniform.

7. Summary and Conclusion

In order to assess the performance of applying chaos theory for daily traffic estimation, two main roads of Qazvin-Rasht and Tabriz-Soufian have been selected. Experimental daily traffic data, gathered through automatic traffic counter, have been analyzed for checking the performances of utilizing normal and uniform distribution functions versus applying

Table 2. Mean square errors of data and models' outputs ($\times 10^6$)

Time period of prediction	Qazvin- Rasht			Tabriz- Soufian		
	Normal	Uniform	Chaos	Normal	Uniform	Chaos
5-day estimation	267	659	309	31	126	38
15-day estimation	122	568	400	40	99	61
Total error	68	330	374	29	44	63
Performance rank	First	Third	Second	First	Third	Second

Table 3. Difference between the number of observation and model output
(Qazvin-Rasht)

Domain		Real	Normal		Uniform		Chaos	
Lower	Upper	No.	No.	Diff.	No.	Diff.	No.	Diff.
	<1700	0	31	31	0	0	14	14
1700	2200	2	4	2	4	2	15	13
2200	2700	2	4	2	5	3	9	7
2700	3200	1	17	16	5	4	8	7
3200	3700	2	12	10	3	1	11	9
3700	4200	17	12	5	5	12	9	8
4200	4700	34	7	27	8	26	6	28
4700	5200	52	13	39	4	48	4	48
5200	5700	26	11	15	4	22	3	23
5700	6200	25	10	15	4	21	2	23
6200	6700	20	15	5	6	14	5	15
6700	7200	16	12	4	7	9	4	12
7200	7700	19	6	13	3	16	4	15
7700	8200	9	8	1	4	5	4	5
8200	8700	15	15	0	4	11	4	11
8700	9200	12	19	7	3	9	5	7
9200	9700	12	13	1	5	7	6	6
9700	10200	7	10	3	5	2	0	7
10200	10700	6	9	3	2	4	3	3
10700	11200	8	13	5	3	5	6	2
>11200		1	11	10	4	3	224	223
Average				10.19		10.67		23.14

Table 4. Difference between the number of observation and model output
(Tabriz-Soufian)

Domain		Real	Normal		Uniform		Chaos	
Lower	Upper	No.	No.	Diff.	No.	Diff.	No.	Diff.
	<6000	4	13	9	65	61	108	104
6000	6800	6	15	9	11	5	9	3
6800	7600	9	13	4	13	4	13	4
7600	8400	30	12	18	13	17	9	21
8400	9200	34	22	12	13	21	7	27
9200	10000	36	18	18	10	26	8	28
10000	10800	47	30	17	14	33	9	38
10800	11600	30	28	2	18	12	9	21
11600	12400	20	29	9	24	4	6	14
12400	13200	21	27	6	16	5	13	8
13200	14000	12	28	16	15	3	12	0
14000	14800	13	32	19	15	2	10	3
14800	15600	16	21	5	13	3	7	9
15600	16400	24	27	3	13	11	12	12
16400	17200	19	17	2	13	6	12	7
17200	18000	11	9	2	11	0	12	1
18000	18800	6	6	0	7	1	17	11
18800	19600	12	6	6	20	8	11	1
19600	20400	5	4	1	27	22	10	5
20400	21200	4	3	1	9	5	18	14
>21200		1	0	1	14	13	51	50
Average				7.62		12.48		18.14

chaos theory. The ratio between the minimum and maximum traffic volumes over a year is considered as chaos factor. Following that the presence of chaos has been checked by calculating the largest Lyapunov exponent. At the first step, results showed that the daily variation for road traffic can be considered as chaotic pattern, but applying chaos theory is not an appropriate technique for estimating daily traffic volumes compared to normal distribution function. Applying three approaches of using normal and uniform distribution functions as well as the concept of chaos theory revealed that the performances of three approaches are ordered as normal distribution function, chaos theory and eventually uniform distribution function. The main conclusion of this research work lies on the topic that although daily traffic volumes may adapt with chaotic behavioral patterns but the best way is to check the reliable data probability functions when there is no input variables for modeling traffic volume recognition. In other words, the main result of this research work for researchers interested on using chaos theory for estimating daily traffic is to use proper distribution function instead of chaos theory, but because of existing dynamic patterns it would be better to be utilized instead of uniform distribution function.

Researchers, interested on studying in this topic, are recommended to search for another definition of chaos factors for road traffic patterns. The other parameters of traffic may be found to have chaotic patterns. They are also recommended to check the power of chaos-based techniques compare to the other statistical techniques such as regression models, while input parameters are available.

8. References

- Ardekani, S., Hauer, E. and Jamei, B. (1995) "Traffic flow theory, Part 7: Traffic impact models", pp. 1-24
- Ashok, K. and Ben-Akiva, M. E. (2002) "Estimation and prediction of time-dependent origin-destination flows with a stochastic mapping to path flows and link flows", *Transportation Science*, Vol. 36, pp. 184-198
- Bodle, R. (1967) "Evaluation of rural coverage count duration for estimating annual average daily traffic", *Highway Research Record*, Vol. 199, pp. 67-77
- Caggiani, L., Dell'Orco, M., Marinelli, M. and Ottomanelli, M. (2012) "A meta-heuristic dynamic traffic assignment model for O-D matrix estimation using aggregate data", *Procedia - Social and Behavioral Sciences*, Vol. 54, pp. 685 – 695

- Faghri, A. and Hua, J. (1995) "Roadway seasonal classification using neural networks", Journal of Computing in Civil Engineering, Vol. 9, No. 3, pp. 209-215
- Frazier, C. and Kockelman, K. M. (2004) "Chaos theory and transportation systems: Instructive example", Journal of Transportation Research Board, No. 1897, TRB, National Research Council, Washington, D.C., pp. 9-17
- Gecchele, G., Caprini, A., Gastaldi, M. and Rossi, R. (2011) "Data mining methods for traffic monitoring data analysis. A case study", Procedia Social and Behavioral Sciences, Vol.20, pp. 455-464
- Gleick, J. (1987) "Chaos: making a new science", Copyright Vintage publishing company.
- Juran, I., Prashker, J.N., Bekhor, S. and Ishai, I. (2009) "A dynamic traffic assignment model for the assessment of moving bottlenecks", Transportation Research, Part C, Vol. 17, No.3, pp. 240-258
- Kiel, L. D. and Elliott, E. (1996) "Chaos theory in the social sciences, foundations and applications", University of Michigan.
- Lawrence, W. L., Feng, Y. L. and Huang, Y. C. (2003) "Diagnosis of freeway traffic incidents with chaos theory", Journal of Eastern Asia Society for Transportation Studies, Vol.5, pp. 2025-2038
- Lingras, P. (1995) "Hierarchical grouping versus Kohonen neural networks", Journal of Transportation Engineering, Vol. 121, No. 4, pp. 364-368
- Lingras, P. (2001) "Statistical and genetic algorithms classification of highways", Journal of Transportation Engineering, Vol. 127, No. 3, pp. 237- 243
- Mahmoudabadi, A. and Fakharian, S. (2010) "Estimating missing traffic data using artificial neural network", Proceedings of Second International Conference on Intelligent Network and Computing (ICINC 2010), November 26 - 28, 2010, Kuala Lumpur, Malaysia, Vol. 2, pp. 144-147
- Mahmoudabadi, A. and Ghazizadeh, A. (2013) "A hybrid PSO-fuzzy model for de-

termining the category of 85th speed”, Journal of Optimization, Volume 2013, Article ID 964262

- Mahmoudabadi A. and Seyedhosseini S. M. (2012) “Application of chaos theory in hazardous materials transportation”, International Journal of Transportation Engineering, Vol. 1, No. 1, pp. 15-22

- Mahmoudabadi, A. and Seyedhosseini, S. M. (2012) “Time-risk tradeoff of hazmat routing problem in emergency situation”, Proceedings of the 2012 International Conference on Industrial Engineering and Operations Management, Istanbul, Turkey, July 3 – 6, 2012, pp. 344-351

- Mingjun, J. and Huanwen, T. (2004) “Application of chaos in simulated annealing”, Chaos, Solitons & Fractals, Vol. 21, pp. 933–941

- Rossi, R., Gastaldi, M., Gecchele, G. and Kikuchi, S. (2012) “Estimation of annual average daily truck traffic volume, uncertainty treatment and data collection requirements”, Procedia - Social and Behavioral Sciences, Vol. 54, pp. 845 – 856

- Sharma, S. C. and Werner, A. (1981) “Improved method of grouping province wide permanent traffic counters”, Transportation Research Record, Vol. 815, pp. 12-18

- Sugihara, G. and May, R. M. (1990) “Non-linear forecasting as a way of distinguishing chaos from measurement error in time series”, Nature, Vol. 344, pp. 734-741

- Trepaniera, M., Leroux, M. H. and Marcellis-Warin, N. (2009) “Cross-analysis of hazmat road accidents using multiple databases”, Accident Analysis and Prevention, Vol. 41, pp. 1192–1198

- Xh, Y., Yb, Y. and Zhang, Y. C. (2002) “A hybrid chaotic genetic algorithm for short-term hydro system scheduling”, Mathematical Computer Simulation Vol. 59, pp. 319–27

- Zhao, F., Li, M.Y. and Chow, L. (2004) “Alternatives for estimating seasonal factors on rural and urban roads in Florida”, Research Office, Florida Department of Transportation