The Assessment of Applying Chaos Theory for Daily Traffic Estimation

Abbas Mahmoudabadi
Department of Industrial Engineering
Payam-e-Noor University
Tehran, Iran

Saeed Andalibi
High Institute of Education and Research in Planning and Management
Tehran, Iran

Abstract

Road traffic volumes in intercity roads are generally estimated using probability functions, statistical techniques of regression or meta-heuristic approaches such as artificial neural networks. Although, road traffic volumes may be dependent upon input variables of road geometrical design, weather conditions, weekend, national holiday and so on, but traffic volumes are also estimated by using pattern recognition techniques. The main question from the researchers, here, is to check the using chaotic pattern of daily traffic volume and the performance of chaos theory to estimate daily traffic volume. In this research work, the existing of chaotic behavior in daily traffic volume in intercity roads has been examined while the performance of chaos theory is discussed compared to probability functions. The ratio between the minimum and maximum of daily traffic volume is considered as chaos factor. Data, gathered through two installed automatic traffic counters over a year, have been used in analytical process. Results show that daily traffic volumes have chaotic behavior with defined twenty-four hour time span, but applying the principle of chaos theory is not appropriate method for estimating daily traffic volumes comparing to a good fitness probability function.

Keywords
Chaos Theory, Traffic Volume Estimation, Probability Function, Pattern Recognition

1. Introduction
Estimating traffic volume is one of the most important concerns for traffic managers, in particular for improving enforcement and planning traffic management. Collecting data and predicting traffic patterns are usually performed for pavement design, fuel-tax revenue projection and highway planning. But monitoring activities, necessary for accurate annual average daily traffic (AADT) estimates, are expensive in terms of costs and personnel (Ricardo R., 2012). Traffic estimation is not only a main issue for tactical purposes of transportation such as enforcement (Ashok, 2002), but also it is highly related to road accidents' frequency and their severities (Ardekani, 1995). On the other hand, transport responsibilities mainly focus on predicting road traffic volumes in order to make road safety plans over the network regarding to improving enforcement and implementing road safety measures.

In the era of using intelligent transport systems, established in all over the world to improve the efficiency of transportation, traffic estimation has become more and more critical issue in road transport management (Caggiani, 2012). Estimating road traffic volumes has become more important while it is considered to determine origin-destination matrix as a basic stage of transportation planning (Caggiani, 2012).

There are many models and methodologies for estimating traffic volumes in the literature. The main techniques including Regression Analysis (Faghri & Hua, 1995), Ward's Minimum-Variance method of clustering (Sharma & Werner, 1981), Clustering-based methods (Zhao, et al., 2004), and some heuristic methods of Genetic Algorithms (Lingras, 2001) and Artificial Neural Networks (Faghri & Hua, 1995, Lingras, 2001, Mahmoudabadi & Fakharian, 2010) have been implemented for road traffic estimation. Simulation techniques (Juran et al., 2009) and data mining (Gecchele et al., 2011) which are mainly developed through using the probability distribution functions are also applied for traffic estimation. Variation in day-to-day traffic volumes is a very important factor in the process of estimating daily traffic (Bodle, 1976). 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 may be a useful technique to develop a chaotic pattern for estimating road traffic volumes. In addition, determining the performance of using chaos theory may or may not convince researchers to develop chaotic patterns for road traffic estimation.

According to all of the above mentioned, the basic principle of this research work is to consider daily traffic volume as a chaotic variable and develop an appropriate model to predict that. While an appropriate chaos factor related to daily traffic volume is defined, the presence of chaos, as a major concern for developing methodology, will be checked. Although the dependency upon initial condition is another concern to develop chaotic patterns (Mahmoudabadi and Seyedhosseini, 2012), but traffic volumes are being estimated in short-time prediction for traffic management. Therefore, it is assumed that, daily traffic volume is estimated in short-time prediction and the above concept may be useful for researchers, those are interested to know whether the chaos theory is a proper technique to estimate daily traffic or not.

This paper is articulated into seven sections. After introduction a short description of chaos theory is discussed followed by methodology definition in which details of application of chaos theory for daily traffic estimation are presented. Details on case studies and analytical process are demonstrated following the above sections presented in sections four and five, respectively. Discussions on rating the performances of three approaches are presented in section six followed by conclusion and some recommendations for further researches in the last section.

2. Chaos Theory

The concept of chaos theory was firstly introduced by Edvard Lorenz in 1963 (Gleick, 1987). Chaos theory is known as a non-linear system, while in a chaotic nonlinear system, it may happen that small differences in the initial conditions produce very great ones in the final outputs (Xh et al., 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 information processing (Mingjun and Huanwen, 2004). Chaos theory has been widely applied on various fields of science particularly in the area of traffic flow theory, commonly in short-time prediction, because of the existence of the property of "sensitive dependence upon initial condition" (Lawrence et al., 2003).

In transportation systems, legal and social constraints may bind behavior, allowing researchers to more accurately predict human actions and their affect on system evolution (Frazier and Kockelman, 2004 and Sugihara and May, 1990). In addition, it has been proved that some traffic measures such as the number of road accidents can be predicted under the property of chaotic behavior (Mahmoudabadi and Seyedhosseini, 2012). The largest Lyapunov exponent, usually indicated as λ_{max} in formulation, is the clearest measure to prove the existence and to quantify chaos in a dynamical system or time series. Calculating the largest Lyapunov exponent is a more common technique to determine the presence of chaos which 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. 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.

$$\lambda_{max} = \frac{1}{N\Delta t}\sum_{t=0}^{N-1} Ln\left(\frac{|S(t+\Delta t) - S'(t+\Delta t)|}{|S(t) - S'(t)|}\right) \tag{1}$$

The well-known equation of logistic map is commonly applied to generate chaotic variables (Sugihara and May, 1990 and Trepaniera et al., 2009), particularly in traffic behavior, because of good adeptness to time series. The logistic map equation is defined as equation (2) (Lawrence et al., 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 Huanwen, 2004).

$$P(t+1) = K \times P(t) \times (1 - P(t)) \tag{2}$$

3. Methodology Definition

In this research work, traffic volume prediction is done following some stages. Experimental data which are 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 process. Therefore, at the first stage, a filtering process is done in which 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. Defining a chaos factor is on the next stage in which a chaotic variable is defined in closed interval [0 1] (Mahmoudabadi & Seyedhosseini, 2012). According to daily traffic volumes (TV), equation (3) is used to define chaos factor in the proposed model. It will be a useful technique to define chaos factor in which results are easily compared to the other methods. For each part of time, known as time period (t) represents a day, daily chaos factor (CF) is updated using logistic map equation modified as equation (4). Following the above steps, estimated daily traffic volume (ETV) for time (t) will be obtained using equation (5), where the maximum (MaxTVs) and minimum Min(TVs) of traffic volumes have been calculated through using experimental data.

$$CF(t) = \frac{TV(t) - MinTVs}{MaxTVs - MinTVs} \tag{3}$$

$$CF(t+1) = K \times CF(t) \times (1 - CF(t)) \tag{4}$$

$$ETV(t+1) = MinTVs + (MaxTVs - MinTVs) \times CF(t+1) \tag{5}$$

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

4. Case Study and Experimental Data

Two intercity roads of Qazvin-Rasht and Tabriz-Soufian have been selected as case studies in order to check the validity of proposed procedure. Qazvin-Rasht is currently a divided road, 130 km length and two lanes in each direction. Tabriz-Soufian is also divided, but 50 km length with two lanes in each direction. The difference between the above roads is regarding to the existence of U-Turns which causes that the definition of roads are officially be different regarding to constructive regulations. There are some U-Turn accesses in Tabriz-Soufian road, so Qazvin-Rasht and Tabriz-Soufian are known Freeway and Express way, respectively. Data over a year started from 21-March 2012 to 20-March 2013 have been collected through electronic devices installed for the purpose of traffic counting. They have been gathered in daily traffic volume pattern while those with less than 20 hours of gathered data have been filtered for analytical process.

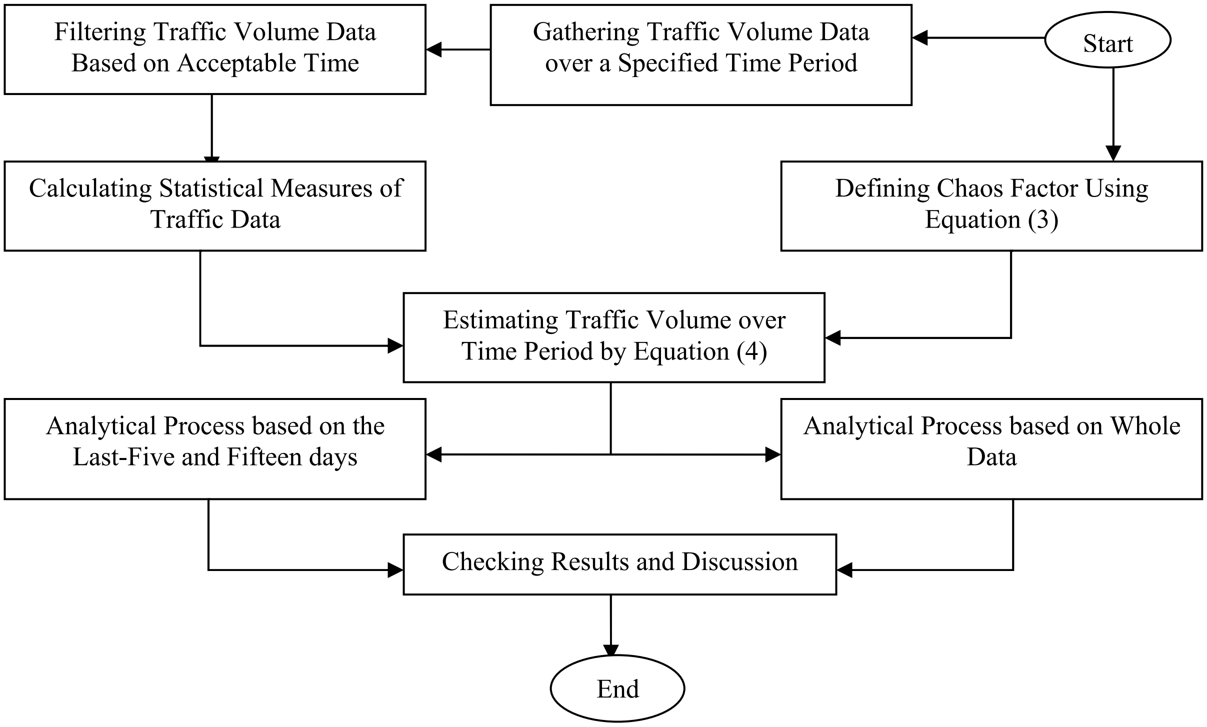

Figure 1: Overall view of the proposed methodology

5. Analytical Process

The first step of numerical process is to draw bar chart schematic pattern of daily traffic volume over time period. This can help decision makers to make an overall perspective for obtaining proper distribution functions. Consequently, 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 makers to develop better distribution functions for data sets. Daily traffic volumes for filtered data have been analyzed and statistical measures including minimum, maximum and standard deviation are presented in table 1.

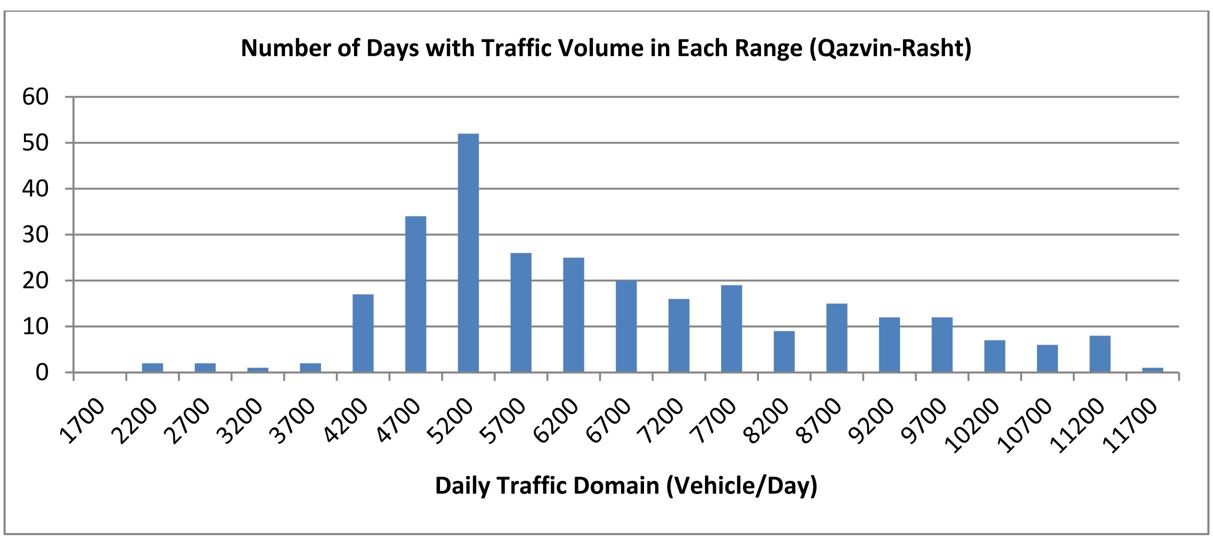

Figure 2: Bar Chart of Daily Traffic Volume for Qazvin-Rasht Freeway

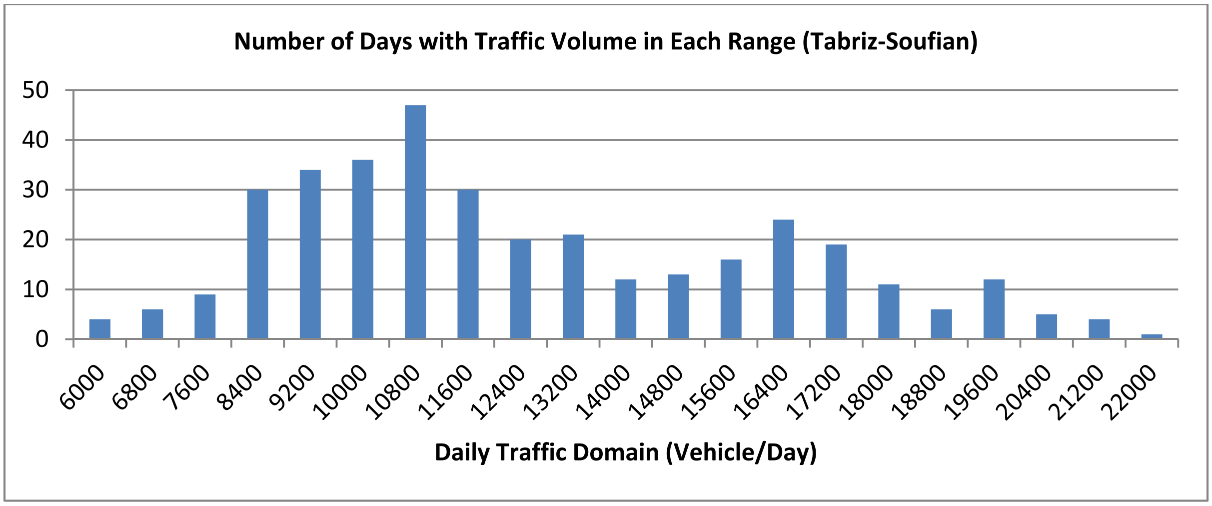

Figure 3: Bar Chart of Daily Traffic Volume for Tabriz-Soufian Express Way

Table 1: Statistical measures for daily traffic volume

Road	Number of records	Min	Max	Mean	Standard Deviation	Largest Lyapunov Exponent (×10-3)	
						Δt=1	Δt=7
Qazvin-Rasht	346	1765	41045	8599	6006	4.37	-5.96
Tabriz-Soufian	363	2250	22415	12217	3775	6.11	1.24

The second step of analytical process is to check the presence of chaos. The variation of chaos factors for two data sets 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.

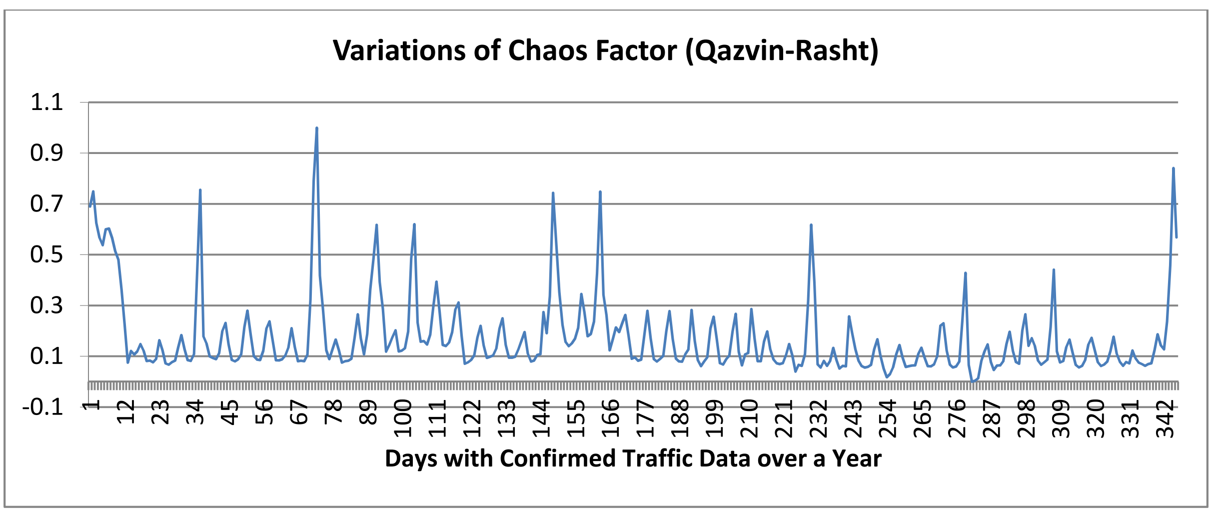

Figure 4: Variation of Chaos Factor for Qazvin-Rasht Freeway over a Year

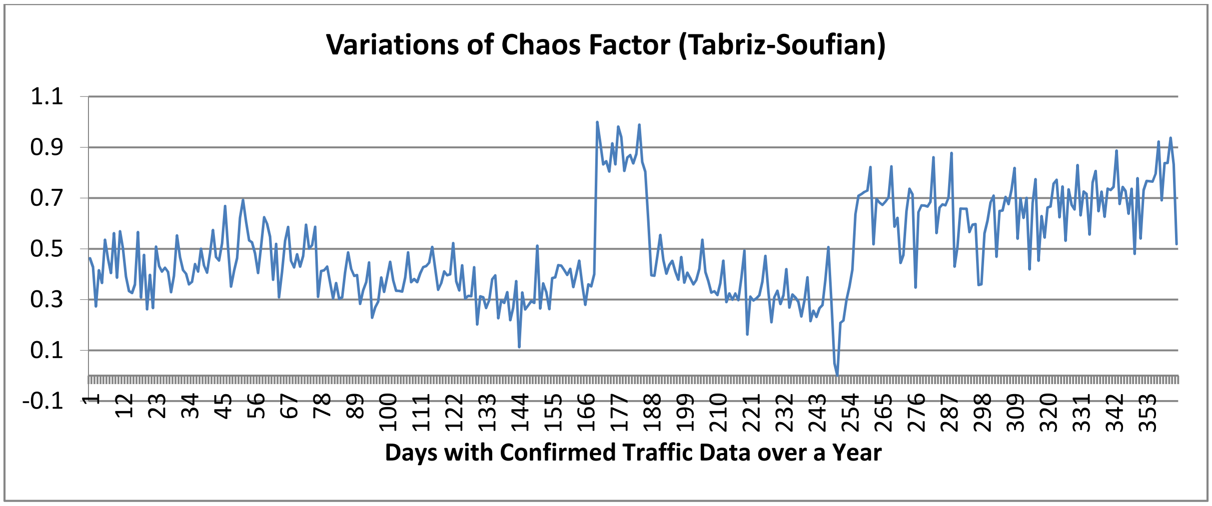

Figure 5: Variation of Chaos Factor for Tabriz-Soufian expressway over a Year

6. Discussion

In order to recognize the probability function, two histograms for traffic data over a year outlined in the previous section have been analyzed and it seems that two different distribution functions of normal and uniform may be compatible to traffic data sets. Therefore, two different probability functions of normal and uniform have been utilized to estimate traffic volume and compare results with the case of using chaos theory. Following that, two criteria measures of mean square errors and difference between the number of observation and model output in each domain (Mahmoudabadi & 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 corresponding to normal distribution function, in which the minimum of mean square errors between model outputs and observations are obtained. Chaos theory is located in the second rate, 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.

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 rate	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 5: difference between the number of observation and model output (Tabriz-Soufian)4

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

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 electronic traffic counter, have been analyzed for checking the performances of utilizing normal and uniform distribution functions versus applying chaos theory. The situation of traffic volume in each day regarding to 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 Lyaponuve exponent. Meanwhile, results show that the daily variation for road traffic may 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 reveals that the performances of three approaches are ordered as normal distribution function, chaos theory and eventually 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.

References

Ardekani S., Hauer E., Jamei B., Traffic Flow Theory, *Part 7: Traffic Impact Models*, pp. 1-24, 1995.

Ashok K. Ben-Akiva M.E. 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, 2002.

Bodle R., Evaluation of Rural Coverage Count Duration for Estimating Annual Average Daily Traffic. *Highway Research Record*, Vol. 199, pp. 67-77, 1967.

Caggiani L., Dell'Orco M., Marinelli M., Ottomanelli M., 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, 2012.

Faghri A., Hua J., Roadway Seasonal Classification Using Neural Networks, *Journal of Computing in Civil Engineering*, Vol. 9 No. 3, pp. 209-215, 1995.

Frazier C., Kockelman K.M., "Chaos Theory and Transportation Systems: Instructive Example", *Journal of the Transportation Research Board*, No. 1897, TRB, National Research Council, Washington, D.C. pp. 9–17. , 2004.

Gecchele G., Caprini A., Gastaldi M., Rossi R., Data mining methods for traffic monitoring data analysis. A case study,*Procedia Social and Behavioral Sciences*, Vol. 20, pp. 455-464, 2011.

Gleick J., Chaos: making a new science, 1987.

Juran, I., Prashker, J.N., Bekhor, S., Ishai, I. A dynamic traffic assignment model for the assessment of moving bottlenecks, *Transportation Research, Part C*, Vol. 17, No. 3, pp. 240-258, 2009.

Kiel L.D., Elliott E., "Chaos Theory in the Social Sciences, Foundations and Applications", University of Michigan, ISBN 0-472-08472-0, 1996).

Lawrence W.L., Feng Y.L., Huang Y.C., "Diagnosis of Freeway Traffic Incidents with Chaos Theory", *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 5, pp. 2025-2038, 2003.

Lingras P., Hierarchical Grouping Versus Kohonen Neural Networks, *Journal of Transportation Engineering,* Vol. 121, No. 4, pp. 364-368, 1995.

Lingras P., Statistical and Genetic Algorithms Classification of Highways, *Journal of Transportation Engineering*, Vol. 127, No. 3, pp. 237- 243, 2001.

Mahmoudabadi A., Fakharian S., Estimating missing traffic data using artificial neural network, *Proceeding of Second international conference on International Conference on Intelligent Network and Computing (ICINC 2010)*, November 26 - 28, 2010, Kuala Lumpur, Malaysia, Vol. 2, pp. 144-147.

Mahmoudabadi A. Ghazizadeh A., A hybrid PSO-Fuzzy Model for Determining the Category of 85^{th} Speed, *Journal of Optimization*, Volume 2013, Article ID 964262, 8 pages,

Mahmoudabadi A., Seyedhosseini S.M., "Application of Chaos Theory in Hazardous Materials Transportation", *International Journal of Transportation Engineering,* Vol. 1, No. 1, pp. 15-22, 2012.

Mahmoudabadi A., Seyedhosseini S.M., "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, 2012.

Mingjun J., Huanwen T., "Application of chaos in simulated annealing", *Chaos, Solitons & Fractals*, Vol. 21,pp. 933–941, 2004.

RossiR., GastaldiM., GeccheleG., KikuchiS., "Estimation of Annual Average Daily Truck Traffic Volume, Uncertainty treatment and data collection requirements", *Procedia - Social and Behavioral Sciences*, Vol. 54 , pp. 845 – 856, 2012.

Sharma S.C., Werner A., Improved Method of Grouping Province wide Permanent Traffic Counters, *Transportation Research Record*, Vol. 815, pp. 12-18, 1981.

Sugihara G., May R.M., "Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series", *Nature,* Vol. 344, pp. 734-741, 1990.

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

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

Zhao F., Li M.Y., Chow L., Alternatives for Estimating Seasonal factors on Rural and Urban Roads in Florida, *Research Office, Florida Department of Transportation*, 2004.

Biographical Notes

Abbas Mahmoudabadi is PhD in Industrial Engineering and works as a senior expert in road safety and public transportation in road maintenance and transportation organization. He has more than 23 journal or international conference papers and one book chapter published in the field of industrial engineering, transportation and road safety. In addition to his teaching experiences in universities, he has 19 years of executive experiences in road safety planning in developing countries. Some national transportation projects have been implemented under his supervisory responsibilities with the results of fatality reduction in intercity transportation.

Saeed Andalibi graduated in management science around 15 years ago and he currently works as an Iranian national expert in legislation. Some national regulations have been provided and approved under his supervisory roles in transportation, E-government and social aspects during the last decade.