



PREDICTING THE DAILY TRAFFIC VOLUME FROM HOURLY TRAFFIC DATA USING ARTIFICIAL NEURAL NETWORK

*M.S.A Siddiquee**, *S. Hoque*[†]

Abstract: The prediction of traffic volume over time is very important to control the flow of traffic on a road network. Traffic count is usually averaged over time to predict for the larger time domain. This paper aims at finding the detail variation of a systematic survey of hourly traffic volume data over a time of four years along the North Bengal corridor of Bangladesh (at Jamuna toll collection point) and its equivalent numerical model by using a Artificial Neural Network. The Neural Network is trained with the intermittent data of 13 weeks over four years and the missing data is interpreted with quite reasonable accuracy (12.67% MAE) with this ANN model. The ANN model captured the variety of trends of the traffic data very accurately as has been depicted in the paper.

Key words: *artificial neural network, traffic data prediction, estimation*

Received: May 22, 2014

DOI: 10.14311/NNW.2017.27.015

Revised and accepted: July 2, 2017

1. Introduction

This paper aims at estimating the Daily Traffic (DT) from the short counts on selected rural highway of Bangladesh. There are four major highway corridors in Bangladesh. The Jamuna Multipurpose Bridge is connecting the Northern part and southern part of Bangladesh. As Jamuna Multipurpose Bridge (JMB) is a tolled bridge, the toll collecting agencies and Jamuna Multipurpose Bridge Authority (JMBA) are systematically preserving daily flow data since its opening in June 1998. Gaining access to the data base will give the opportunity to establish the basic flow parameters for at least one of the major corridors of Bangladesh. If these parameters can be established, they can also give a basic idea about these parameters in other regions of the country as well. Prediction of flow for design year is also important for the geometric and structural design of highway. For reliable assessment of average annual traffic as well as future demand, it is desirable that traffic counts are undertaken throughout the year to take account of seasonal

*Mohammed Saiful Alam Siddiquee – Corresponding author; King Abdulaziz University, Civil Engineering Department, Jeddah, Kingdom of Saudi Arabia, E-mail: msa.sid@gmail.com, msiddiquee@kau.edu.sa

[†]Shamsul Hoque; Bangladesh University of Engineering and Technology (BUET), Civil Engineering Department, Bangladesh, E-mail: shoque@ce.buet.ac.bd

variations in traffic flows. In the absence of long duration count there are two approaches for the estimation of Annual Average Daily Traffic (AADT) from short counts [1–3, 13, 14], either by using regression model or by applying previously established expansion factors for the similar facilities [8]. Several studies [9, 10, 15, 16, 19, 20] carried out in developed countries have shown that if continuous flow data is available for a particular road segment of a highway, model could be built considering recursive pattern of traffic flow in order to predict AADT from short count (i.e. less 24 hours) data. In India [13] attempt is also made to establish trend in traffic volumes pertaining to daily, weekly, monthly and yearly flow patterns and to use these particulars for forecasting of traffic volumes. Though, these studies were carried out in developed and developing country like India many years ago, so far no such effort in this subject is made in Bangladesh. One of the main reasons is stated earlier that is lack of long duration round the clock flow data. Moreover, to make a time series model in the context of Bangladesh is somewhat complex and intricate in nature as compared to that of developed countries for the following reasons:

- (1) The shifting nature of the month of Ramadan and Eid festivals greatly affect the traffic flow pattern and makes it difficult to predict monthly expansion factors. But in the western world, main festivals like Christmas, New year or even summer vacations does not shift compared to the calendar year.
- (2) During the monsoon season, substantial portion of bulk freight is transported through waterways. Since the duration of monsoon is not constant, it is difficult to develop a similar yearly flow pattern.
- (3) Unusual events such as hartals (political strikes), transport strikes etc., which greatly affect normal traffic flow, are more frequent in Bangladesh than in developed countries.
- (4) Natural calamities like heavy rainfall, long duration flood is also external parameter which needs to be considered while analyzing flow pattern in Bangladesh.

Design considerations and performance evaluation of highways is not possible using either incorrect or assumed traffic data. This may render the road networks to be inadequately facilitated or over designed. This problem arises due to unavailability of long duration continuous traffic data upon which a reliable pattern analysis can be done and expansion-factors or equations can be determined. But the so far collected and preserved traffic flow data on Jamuna bridge by the toll operators is providing a good opportunity for the transportation planners to effectively utilize these data for traffic flow analyses. Moreover, some other sources of short and long counts data have also been discovered during the study. Therefore, this study avails fair amount of quality traffic data for traffic flow characteristics analyses and expansion of short counts in the North Bengal corridor. The data for the short hourly counts of traffic data are surveyed by a team of consultants from Bangladesh University of Engineering and Technology (BUET) over a period of four years but at a rate of four weeks per annum.

Due to the sparse form of short count data from the survey of BUET consultants, an Artificial Neural Network is employed to interpolate and extrapolate the short count data whenever necessary. Using these predicted data, another robust Artificial Neural Network of ADT data is trained, which can predict ADT reasonably.

2. Source of data

There are four major highway corridors in Bangladesh:

- (1) Dhaka-Chittagong Highway,
- (2) Dhaka-Sylhet Highway,
- (3) Dhaka-North Bengal Corridor,
- (4) Dhaka-Aricha Road to South Bengal.

Different corridors exhibit different characteristics because of the variation in their economic activities. For example, if a corridor is full of mineral resources, the flow pattern, directional distribution, traffic composition would be different from those of an agriculturally resourceful corridor. All these considerations are to play their roles in flow pattern analysis and the development of the framework model. However, in either case, for an efficient analysis of traffic flow characteristics on rural highways, large quantity of reliable data is required.

Before the commencement of Jamuna Multipurpose Bridge operation, long duration reliable traffic data were not available. Jamuna Multipurpose Bridge was opened for public in June 1998. Since then, its operators have collected and preserved daily traffic data in an orderly fashion. Considering that it has taken around six months for the corridor to be stabilized, seven years of daily traffic data is available on Jamuna Bridge. Moreover, Bangladesh University of Engineering and Technology had conducted thirteen weekly surveys on Jamuna Bridge, where hourly traffic was also counted. And also, the first tolled road in the country 'Nalka-Hatikamrul-Bonpara New Road' is also situated in the North Bengal corridor. Taking all these into account, this is rather an easy task to select the corridor for this study.

Data were collected [18] in three ways, e.g., (1) Inductive loop-detector based AUTO counting system (lane based detectors are installed flushing with the road surface) – it gives total vehicle count; (2) CCTV based NN image technique (for incident detection and flow count); (3) Manual vehicular entry by the toll booth operators along with CCTV image (for revenue/income cross checking) – it gives classified vehicle count.

Now, the selected corridor should not only possess long duration traffic data, but also the quality of those data has to be high. Although, traffic generation and flow is a random event, but research has shown that, traffic flow in a particular corridor maintains definite flow pattern and also the same is repetitive in nature. Fig. 1 shows a typical hourly traffic flow pattern on Jamuna Multipurpose Bridge.

From the Fig. 1, it can be seen that, the 13 weeks covered during the surveys repeat the similar nature of curve. This undoubtedly proves the quality of the

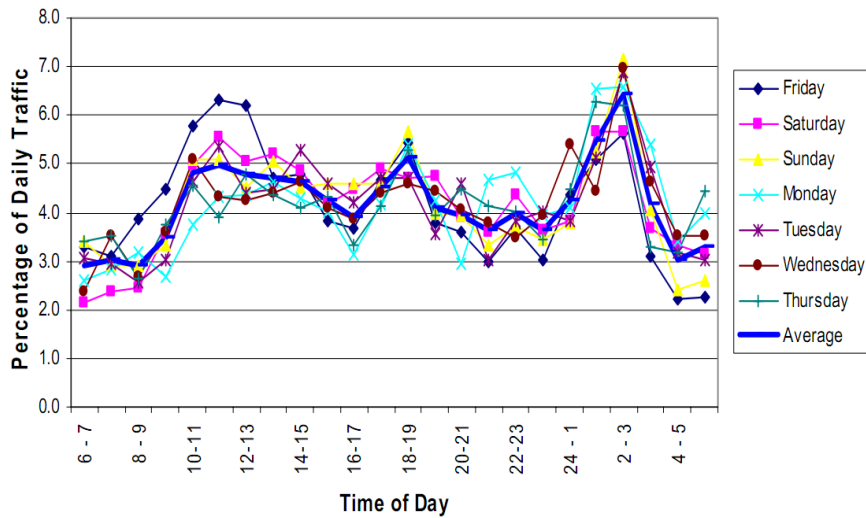


Fig. 1 Hourly Flow Variation on Jamuna Multipurpose Bridge (Data Source: averaged over 13 week’s surveys conducted by BUET).

data and only this type of quality data is worth to be used for flow characteristics analyses and predict the ADT at any future date. So this set of high quality data has been used here in this research to predict the variation of traffic flow with different time frames.

Traffic counts for the different days of the week show the same pattern which may seem unusual at the first look. But it is quite normal. The semi-trailers and trailers carry goods from the Northern part of the country to the Capital city (Dhaka) every day following a particular trend. The traffic flow at night is significantly higher as the semi-trailers and trailers carry goods from the Northern part of the country Capital city to the Capital city of the country at night only. This is controlled by metropolitan city regulations in order to alleviate the traffic congestion inside the city.

3. Artificial neural network

In urban areas of developing countries, like Bangladesh, where the same traffic facility is utilized by different vehicle types, it is indeed needed to induct techniques of artificial intelligence as an alternative approach for traffic parameters’ estimation [19]. Among existing approaches of artificial intelligence, the neural network (NN) approach is commonly used for traffic prediction problems [4–7, 11]. Use of Artificial Neural Networks (ANN) techniques is advantageous as it is not necessary to define relationship beforehand between causal input vectors and corresponding output vectors. Also, ANN techniques have the capability to simulate any degree of non-linear complexities of a system in a more direct way by error minimization procedure, although it takes very long time on a computer. NN approaches have also been used advantageously for short-term estimation and prediction of traffic pa-

rameters [7, 11, 12]. Back-propagation of NN for forecasting of traffic has been used by researchers [4, 7]). Back-propagation of NN was also used to make short-term forecasts of traffic flow, speed and occupancy in the Utrecht/Rotterdam/Hague region of the Netherlands [7]. A straightforward application of NN is possible for forecasting of traffic flows along a motorway link [10]. All these works strengthen the application of NN approach in estimation of traffic stream speed in a mixed traffic environment. For these various reasons, ANN is used in this research.

4. Methodology

It is already mentioned that daily traffic data collected by the toll collecting authority (JOMAC) is continuous over a time of 4 years [15]. But the hourly data is not continuous over this four years time. In total 13 weeks survey of hourly data were collected over a time of four years. So it is rather difficult to directly link this hourly data to the daily average data from toll collecting authority. With this end in view, a novel idea is introduced, an ANN is trained with limited data from the BUET consultant's survey. Firstly, the hourly data is trained with respect to its hour, day, month and year. After a successful training of the hourly data with the ANN, it is able to predict the missing hourly data on other hours of the missing days. Although the data used for training is not continuous data over the time, it can predict the missing data as well as the forecasted data.

The training of the ANN carried out by back-propagation technique. Usually the back-propagation can be performed by either batch model or individual set of data model. Here in this research, the later approach was adopted. The ANN structure had 4 input nodes, 24 hidden nodes and 2 output nodes. In the research, 80% of the data were used for training purpose, 20% of the data were used for validation.

5. Details of implementation

The Artificial Neural Network (ANN) Program used here in this paper is developed by the first author and available at [17]. The program uses a general form of back propagation algorithm to the network to the input and output data. In this paper, one ANN is used and the data structure is described below in Tab. I. The hourly traffic data is available in limited form over a time of 13 weeks spread over four years' time. The 13 weeks are as follows:

- 1st week of January, 1999,
- 3rd week of February, 1999,
- 3rd week of April, 1999,
- 3rd week of May, 1999,
- 1st week of January, 2000,
- 3rd week of February, 2000,
- 3rd week of March, 2000,
- 3rd week of May, 2000,
- 1st week of January, 2001,

3rd week of February, 2001,
 3rd week of March, 2001,
 3rd week of May, 2001,
 and last week of May, 2002.

The data is arranged in the following format seen in Tab. I to fit into the training module of the ANN program.

Year	INPUT			OUTPUT	
	Month	Day	Hours	Hourly traffic	Daily total traffic
1998	10	24	6	57	2185
1998	10	24	7	67	2185
...

Tab. I Data structure used for the training of ANN.

Using the above structured data, the training of back-propagation algorithm started. It needs quite a long time to properly train the data using the interface (Fig. 2). The learning rate and momentum of the training module of the program

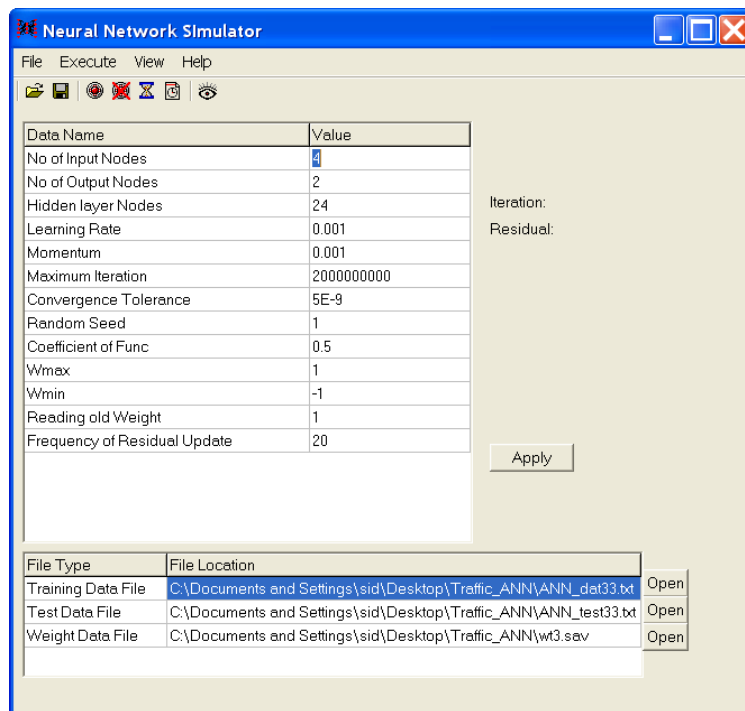


Fig. 2 Main interface of the ANN program used for the training and prediction of traffic data.

can be dynamically adjusted during the training session. Here 24 hidden nodes are estimated after running a node pruning algorithm, which suppressed the over training of the data. The training was stopped when the error norm reached a stable value, which did not change any more even after tweaking the parameters. The error was measured to be 12.67% MAE. Once the training is over, the program can be used to predict the values using existing data patterns.

6. Results and discussion

Once the training is over, a few prediction of traffic volume with hours, days, years are carried out to realize and understand the predicting capability of ANN. Figs. 3 and 4 show the hourly traffic volume against the hours of a day. It can be seen two segments of data from years 1999 and 2002 are selected for prediction. It is surprising to observe that the ANN model of the sporadic data is accurate enough to capture the trend of the data over the hours of traffic volume measurement. Figs. 5 and 6 show the variation of total traffic volume for that day over the hours of day. It is obvious that this should be a constant value over the hours as it is a single value (sum of all the hourly traffic volume). This has been a check to find out how the ANN captures the data pattern. It has been found that it is not exactly constant in both the cases, because of the influence of all other traffic data and the errors in the data due to the factors discussed before. Figs. 7 and 8 show the variation of daily traffic volume with the days of a week. It is found to be well predicted. In Fig. 8, the sudden spike in the data is automatically eliminated. Still a question remains, whether the spike in Fig. 8 is a real variation in the traffic flow data or a spurious variation in the data. The ANN gets the trend of the data pattern very effectively. But there is a fear of over-fitting if we force the ANN to predict smallest variation of the data by hand-tuning the parameters.

The potential of using an ANN model to capture all the possible trends inside the traffic data, Fig. 9 is shown here. It plots the variation of hourly traffic data variation with its years of origin. In this way, the individual growth factor of each hourly traffic data can be seen here. It can be also observed that the hourly variation also changes with year. For the clarity of the Fig. 9, only up to 12 hours of data are plot.

The shifting nature of the traffic patterns due to various cultural, climatic and political scenarios can also be captured by this technique. ANN can identify any hidden patterns in any systematic natural data.

7. Conclusions

The following conclusions are deduced from this research.

- (1) The generated ANN model can capture the hourly traffic data patterns and predict the pattern quite reasonably.
- (2) The trained ANN model can also capture the hourly variation of daily traffic volume.

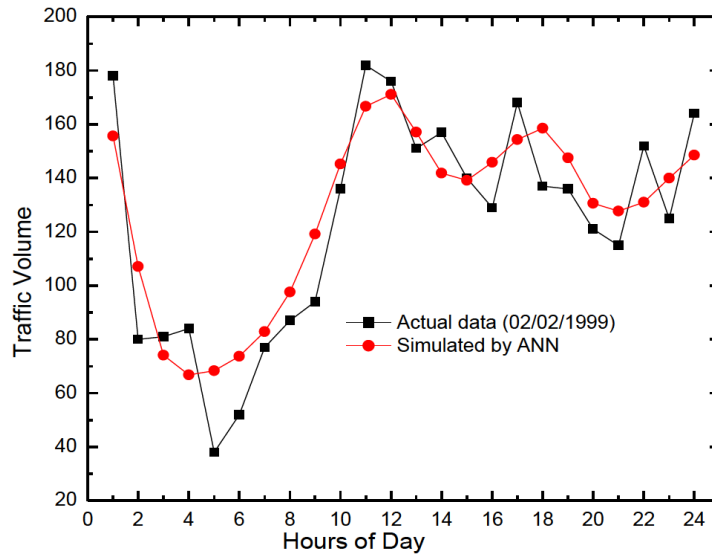


Fig. 3 Plot of hourly traffic volume versus hours of a day for the date (2nd February, 1999).

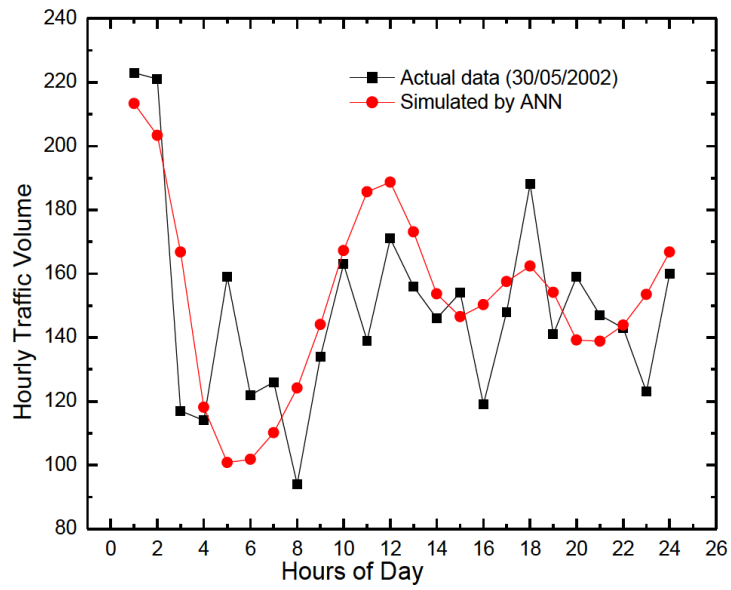


Fig. 4 Plot of hourly traffic volume versus hours of a day for the date (30th May, 2002).

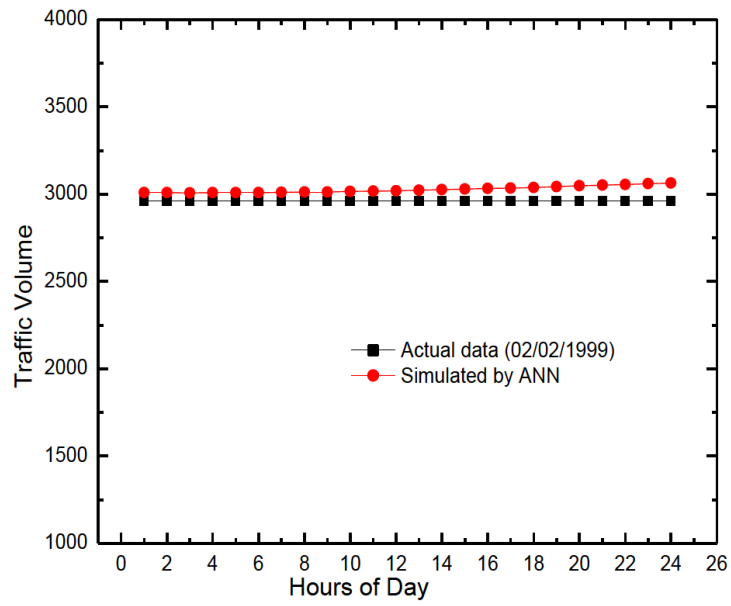


Fig. 5 Plot of daily traffic volume versus hours of a day for the date (2nd February, 1999).

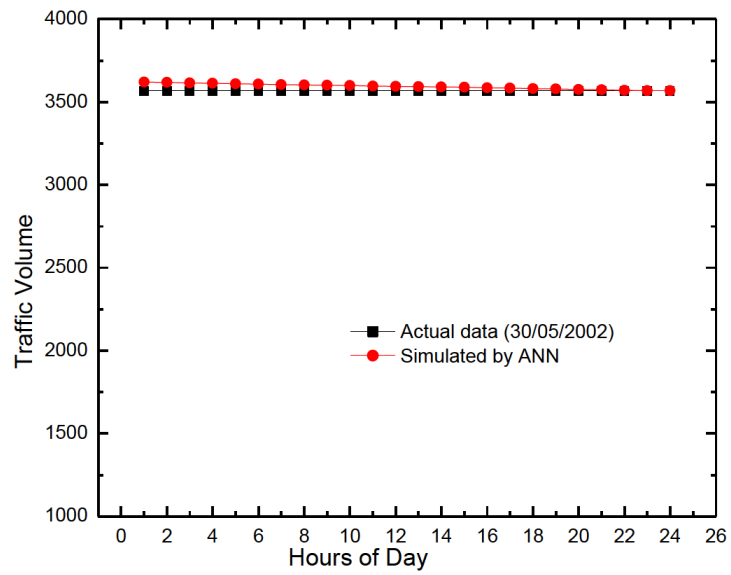


Fig. 6 Plot of daily traffic volume versus hours of a day for the date (30th May, 2002).

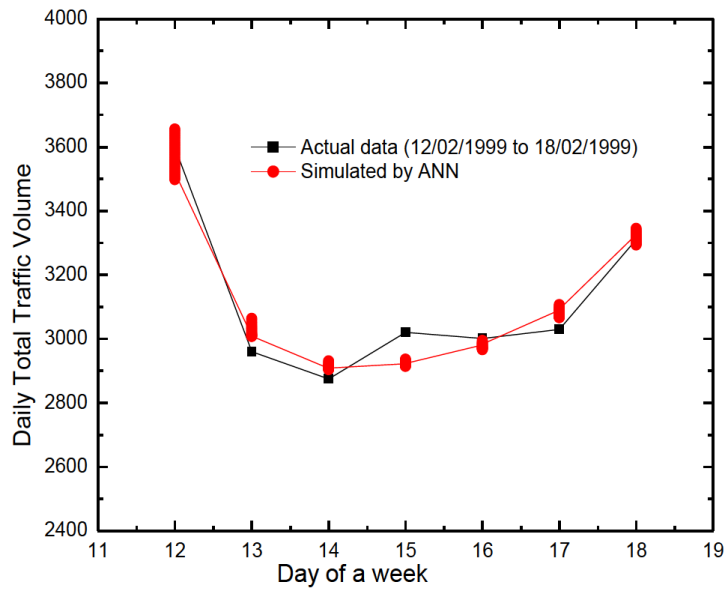


Fig. 7 Plot of daily total traffic volume versus day of a week (12 to 18 of February, 1999).

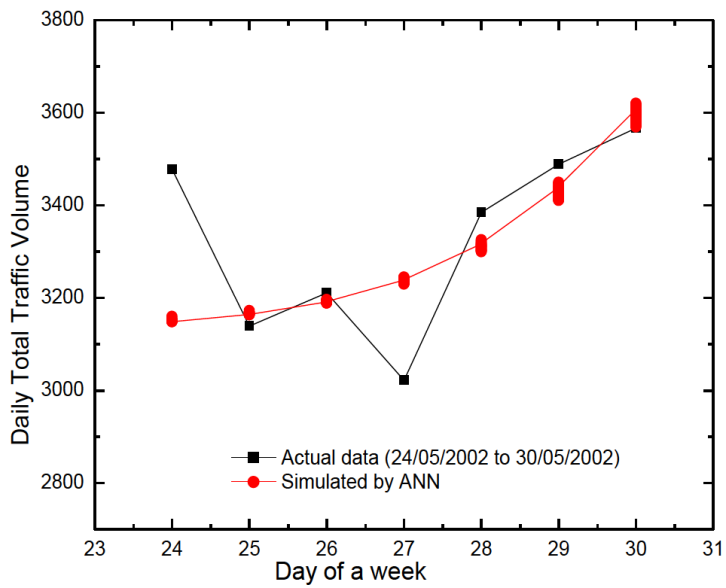


Fig. 8 Plot of daily traffic total volume versus day of a week (24 to 30 of May, 2002).

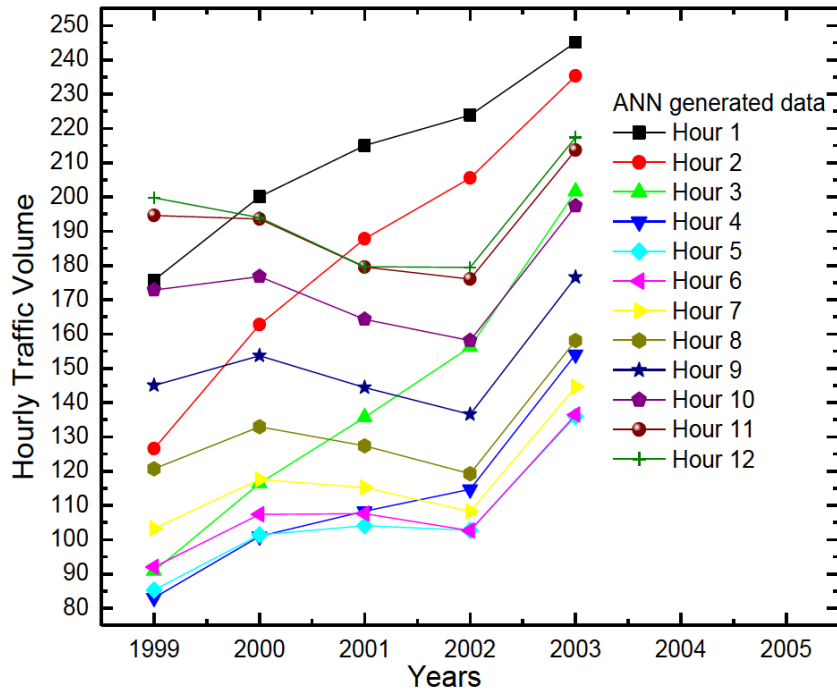


Fig. 9 ANN generated plot of hourly traffic volume versus Years (1999 to 2003).

- (3) The ANN model predicts the variation of daily total traffic volume over the days of a week quite successfully.
- (4) Above all it can predict the hourly growth rate of the traffic volume for the given total time (first six months).

References

- [1] Aashto A. *Policy on Geometric Design of Highway and Street*. USA: Amer Assn of State Hwy & Trans, 1990.
- [2] Barewood J.E. *Traffic Engineering Handbook*. 3rd ed. Institute of Traffic Engineers, Washington D.C.: USA, 1965.
- [3] Development of Geometric Design Standards. Road Material and Standards Study Bangladesh. Volume VIIA, Roads and Highways Department: GOB, 1994.
- [4] Dochy T., Danech-Pajouh M., Lechevallier Y. Short-term road traffic forecasting using neural networks. *RTS English Issue*. 1996, 11, pp. 73–82.
- [5] Dougherty M.S., Kirby H.R., Boyce D. Using neural networks to recognize, predict and model traffic. *Artificial Intelligence Applications to Traffic Engineering*. 1994, pp. 235–250.
- [6] Dougherty M.S., Kirby H.R., Boyle R.D. Using neural networks to recognize, predict and model traffic. In: M. Bielli, G. Ambrosino, M. Boero, eds. *Artificial Intelligence Applications to Traffic Engineering*, VSP, Utrecht, The Netherlands. 1994, pp. 233–250.
- [7] Dougherty M., Cobbett M. Short-term inter-urban traffic forecasts using neural networks. *International Journal of Forecasting*. 1997, 13(1), pp. 21–31, doi: [10.1016/S0169-2070\(96\)00697-8](https://doi.org/10.1016/S0169-2070(96)00697-8).

- [8] Garber N.J., Hoel L.A. *Traffic and Highway Engineering*. Revised 2nd ed. USA: PWS Publishing, 1990.
- [9] Jia Z.Z., Jin X.C., Yuan Z. Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent intersections. *Transportation Research Part C: Emerging Technologies*. 2014, 47(2), pp. 139–154, doi: [10.1016/j.trc.2014.06.011](https://doi.org/10.1016/j.trc.2014.06.011).
- [10] Kirby H.R., Watson S.M., Dougherty M.S. Should we use neural networks or statistical models for short term motorway traffic forecasting? *International Journal of Forecasting*, 1997, 13(1), pp. 43–50, doi: [10.1016/S0169-2070\(96\)00699-1](https://doi.org/10.1016/S0169-2070(96)00699-1).
- [11] Ledoux C. An urban traffic flow model integrating neural network. *Transportation Research*. 1997, 5(5), pp. 287–300, doi: [10.1016/S0968-090X\(97\)00015-6](https://doi.org/10.1016/S0968-090X(97)00015-6).
- [12] Matson T.M., Smith W.S., Hurd F.W. *Traffic Engineering*. McGraw-Hill, 1955.
- [13] Pavement Design Guide. Roads and Highways Department: GOB, 2005.
- [14] Pignataro L.J. *Traffic Engineering, Theory and Practice*. Englewood Cliffs, New Jersey, USA: Prentice-Hall Inc., 1973.
- [15] Rahman S.K. *Study of Vehicular Flow Pattern on Jamuna Multipurpose Bridge Access Road*. Dept. of Civil Engineering, MIST, 2002.
- [16] Robichaud K., Gordon M. An Assessment of Data Collection Techniques for Highway Agencies. British Columbia Ministry of Transportation, 2002.
- [17] Siddiquee, MSA. The ANN program used in this research [software], 2006. Available from: <http://teacher.buet.ac.bd/sid/download>.
- [18] Traffic Monitoring Guidelines, Federal Highway Administration, U.S. Department of Transport, 2001.
- [19] Yin H., Wong S.C., Xu J., Wong C.K. Urban traffic flow prediction using a fuzzy-neural approach. *Transportation Research Part C: Emerging Technologies*. 2002, 10(2), pp. 85–98, doi: [10.1016/S0968-090X\(01\)00004-3](https://doi.org/10.1016/S0968-090X(01)00004-3).
- [20] YUN S.Y., NAMKOONG, RHO J.H., SHIN S.-W., CHOI J.-U. A Performance Evaluation of Neural Network Models in Traffic Volume Forecasting. *Mathematical and Computer Modelling*. 1998, 27(9–11), pp. 293–310, doi: [10.1016/S0895-7177\(98\)00065-X](https://doi.org/10.1016/S0895-7177(98)00065-X).