Statistical Analysis of Time Headways on an Urban Arterial

Abeer Khudhur Jameel **Estabraq Nashat Ezat** **Ali Kadhim Ibrahim**
Al-Mustansiriya University / College of Engineering

Abstract

This Paper aims to apply a statistical method in the analysis of the time headway on urban arterial. Palestine Street is the road that selected for this application. Survey of vehicle volume and time headway is done to gather the needed data. Chi-Square is the Statistical test that used to check the goodness of fit between the observed headway data and the predicted headway data that computed from a 2nd order polynomial Model, which is built after checking of several types of models. The results give a goodness of fit model which have a significant P value.

Key words: Time Headway, Traffic Flow rate, arterial, Chi square, goodness of fit

الخلاصة

يهدف هذا البحث الى تطبيق الطرق الأحصائية في تحليل قيم الفاصلة الزمنية في الطرق الشريانية في مدينة بغداد. وللتطبيق تم اختيار شارع فلسطين كدراسة حالة. تم جمع البيانات المطلوبة والتي شملت الحجم المروري وزمن الفاصلة الزمنية بين مركبتين متتالية على الطرق الشريانية. حللت النتائج باختيار طريقة الفحص الاحصائي Chi-Square لفحص التطابق بين القيم الحقلية والقيم النظرية وجودة الموديل المقترح. النتائج أعطت مؤشر لمدى التطابق بين القيم الحقلية و الموقعية باختبار قيمة P.

الكلمات المرشدة: الفاصلة الزمنية بين مركبتين متتالية، الجريان المروري، طريق شرياني، Chi square, goodness of fit

1. Introduction

There are at least eight basic variables or measures used in describing traffic flow and several other stream characteristics are derived from these. The three primary variables are speed**,** volume, and density. Three other variables used in traffic flow analysis are headway, spacing, and occupancy (Khisty and Lall 1998).

Vehicle headway is a measure of the temporal space between two vehicles, and is defined as: the elapsed time between the arrival of the leading vehicle and the following vehicle at a designated test point. It is usually measured in seconds. Since the average of vehicle headways is the reciprocal of flow rate, vehicle headways represent microscopic measures of flows passing a point. To some extent, the minimum acceptable mean headway determines the roadway capacity (Garber and Hoel 1997).

The area of time headway distribution modeling is fundamental to many traffic flow-related studies such as capacity estimation, safety analysis, microscopic simulation, etc. Up to now, most of the existing models concern about the behavior of general vehicles on a single lane traffic stream or on multiple lanes traffic stream (Rattaphol and Dongjoo 2003)

Consequently, the distribution of headways has an effect on platoon formation and delays. In the 2000 Highway Capacity Manual (TRB 2000) the level of service on two-lane rural highways is approximated by the proportion of headways less than five seconds—thus making a connection between the headway distribution and the level of service. In addition, the headways play a crucial role in the analysis of unsignalized intersections and roundabouts. The characteristics of vehicle headways are accordingly quite relevant in the study of optimal traffic signal control. The mathematical analysis and simulation of traffic operations are based on theoretical models, which must be evaluated against the properties of real world data. More advanced methods have increased the need for reliable knowledge of the statistical properties of vehicle headways. (Luttinen 1996)

Accurate modeling and analysis of vehicle headway distribution helps traffic engineers to maximize roadway capacity and minimize vehicle delays. Furthermore, proper headway models can be applied to generate vehicle events in microscopic simulation models and to traffic safety analysis. All these indicate that vehicle headway distribution is fundamental for many traffic flow research and simulation issues (Guohui Zhang *etal* 2007).

Many headway models have been developed over the past decades. In general, these models can be classified into two categories: single statistical distribution models and mixed models of two or more distributions. The mixed models are more flexible to represent headways by decomposing them into following and free-following components, but the calibration process may be too complex for field application. In practice, selection of the most suitable headway distribution for a certain traffic condition remains an open issue. Also, most previous headway models focused only on general-purpose-lane traffic (Guohui Zhang *etal* 2007).

The availability of accurate estimates for the headway distribution for specific lanes and classes, during diverse periods of the day is of dominant importance for several applications. For instance, these estimates can be used for realistic multiclass multilane vehicle generation in microscopic traffic simulation tools, on-line lane specific capacity and person-car-equivalents estimation, and their consequent use in ITS applications, infrastructural design and planning., safety analysis, and gap acceptance behavior (Hoogendorn and Bovy 1999).

2. Objectives of the Study

This research aims to analyze headways over a full range of volumes and to introduce a statistical methodology for the analysis of the headway data and for the evaluation of the proposed models. The main objective is to present proper statistical methods for the goodness of-fit tests and for the combination of significance probabilities.

3. Methodology

The following methodology has been adopted through the research undertaken

1- Selection of the suitable sites along some arterial streets in Baghdad City
2- Comprehensive survey to determine the peak and off- peak hours to collect the necessary field data during the selected observation intervals.
3- Analysis and reduction of the collected data

4- Selection of the proper statistical techniques for analyzing the headway results and Utilization of the reduced data in the development of the regression models using appropriate computer software.

4. Chi-square test (Vicki 2000)

One of the major problems in the headway studies has been the method for goodness-of-fit tests. The most commonly used test is the chi-square test, which is is a statistical test commonly used to compare observed data with data we would expect to obtain according to a specific hypothesis then you might want to know about the "goodness to fit" between the observed and expected. Were the deviations (differences between observed and expected) the result of chance, or were they due to other factors. How much deviation can occur before you, the investigator, must conclude that something other than chance is at work, causing the observed to differ from the expected. The chi-square test is always testing what scientists call the null hypothesis, which states that there is no significant difference between the expected and observed result.

The formula for calculating chi-square (X^2) is:

$$X^2 = \ (o\text{-}e)^2/e \ldots.1$$

That is, chi-square is the sum of the squared difference between observed (*o*) and the expected (*e*) data (or the deviation, *d*), divided by the expected data in all possible categories.

To calculate X^2 , first determine the number *expected* in each category. Then:

1. Determine degrees of freedom (df). Degrees of freedom can be calculated as the number of categories in the problem minus 1.

2. Determine a relative standard to serve as the basis for accepting or rejecting the hypothesis. The relative standard commonly used in biological research is $p > 0.05$. The p value is the *probability* that the deviation of the observed from that expected is due to chance alone (no other forces acting).

3. Refer to a chi-square distribution table (Table 1). Using the appropriate degrees of freedom, locate the value closest to your calculated chi-square in the table. Determine the closest (probability) value associated with your chi-square and degrees of freedom.

4.1 Step-by-Step Procedure for Testing Hypothesis and Calculating Chi-Square

1. State the hypothesis being tested and the predicted results. Gather the data by conducting the proper experiment (or, if working genetics problems, use the data provided in the problem).

2. Determine the expected numbers for each observational class. Remember to use numbers, not percentages.

3. Calculate X^2 using the formula. Complete all calculations to three significant digits. Round off your answer to two significant digits.

4. Use the chi-square distribution table to determine significance of the value. Determine degrees of freedom and locate the value in the appropriate column, Locate the value closest to your calculated ^{2} on that degrees of freedom *df* row, then move up the column to determine the p value.

5. State your conclusion in terms of your hypothesis. If the *p* value for the calculated X^2 is $p > 0.05$, accept your hypothesis. 'The deviation is small enough that chance alone accounts for it. A *p* value of 0.6, for example, means that there is a 60% probability that any deviation from expected is due to chance only. This is within the range of acceptable deviation. If the p value for the calculated ^{2} is $p < 0.05$, reject your hypothesis, and conclude that some factor other than chance is operating

for the deviation to be so great. For example, a p value of 0.01 means that there is only a 1% chance that this deviation is due to chance alone. Therefore, other factors must be involved.

The chi-square test will be used to test for the "goodness to fit" between observed and expected data from several laboratory investigations in this lab manual.

5. Site Selection

Past research found that the data should be collected from one site to cancel out the variability related to different sites, that one would use the minimum allowable headway from the observed data for the parameter of the shifted exponential distribution (Al-Ghamdi, 2001)

Palestine Street is the site that choice to meet the requirements of an interrupted flow facility. Three points are selected to collect data, the First is at the Mustansiriya squar, the 2nd at the Palestine Intersection and the third at Beirut Squar. This road has the following characteristics:

1. It is located in level terrain.
2. It carries a composite traffic volume of passenger car, truck, buses, pick up, coaster, and Microbuses.
3. It carries a large proportion of the daily traffic volume.
4. It is two-way, divided signalized urban arterial.

All surveys were achieved during daytime only in dry and clear conditions. Adverse weather conditions may cause variation in the usual traffic flow patterns.

During periods of holiday and weekends, the traffic flow characteristics on any highway change dramatically compared with typical weekday traffic conditions. Therefore, to eliminate any variance in the data attributable to this effect, data were collected only on weekdays.

6. Volume Conducting

Traffic volume was observed manually at the selected sections. The traffic counts used in this survey were collected of each segment at the midblock in 5-minute intervals for four and three arbitrary hours of observations in order to determine morning and afternoon peak and off-peak periods. A total of 16 observation hours were earned out on the two selected roads.

One observer with stop watch and field sheet was needed for each direction for the purpose of this survey.

Urban routes show very little variation in peak-hour traffic. In many urban areas, both the A.M. and P.M. peak periods extend for more than one hour.

Selection of the intervals data collection was based on the hourly variations in traffic volume. Use of 5- minute rates of flow provides enough points to obtain the traffic flow patterns.

7. Headway Conducting

Average headway time between each two following thorough vehicles is conducted. The most common method of collecting headway time data is the measurement of the time between two vehicles to cross a point. This time can be measured by manual method. This method requires a stopwatch.

Tables (1) and (2) present examples of the field Flow and Headway data. Tables (3) and (4) present overall flow data with its descriptive statics, while figures (1) and (2) observed headway at the selected arterial.

8. PCE Values for Traffic Volumes

Vehicles of different types require different amount of road space because of variations in size and performance. To allow for this in capacity measurements for roads, traffic volumes are expressed in passenger car unit.

The observed traffic volumes consist of five classes of vehicles namely the Passenger cars, bus, Coaster, Kia and truck.

To convert the traffic volume to passenger car, passenger car equivalents (PCE) values were taken as (2.0) for buses and trucks and 1.8 for condition of level terrain. These values were selected based on available Iraqi Specifications.

Tables (3) and (4) show traffic flow after conversion to passenger car for 5-minute interval with their descriptive statics

9. Building of the model

Several regression techniques are available for analyzing the relationships among many variables. Selecting the proper statistical tool depends on the nature and purpose of the investigation-adjusted coefficient of determination (R^2) is used as one of the main evaluation criteria in selecting the model that best fitted the field data. Tables (5) and (6) shows the comparison between the proposals models. The best model is polynomial order. This model is shown in Figures (3) and (4).

10 Headway time-vehicle flow Analysis

As shown in tables (3) and (4), arterial headway data covered the volume range from around 300 to 1900 vph for south direction and 300 to 2000 for direction. It was difficult to collect usable data for high traffic flows for this study.

The chi-square testing technique was used on the observed data for arterials. The results of the test are shown in tables (7) and (8).

The sample coefficient of variation (CV) is the proportion of sample standard deviation and sample mean. In distribution functions CV is the proportion of standard deviation to expectation. Figures (5) and (6) depicts CV values for the selected road. In this figure the CV values fall in the range of 0.05 to less than 0.3 over a range of flow rates from less than 500 to less than 2,000 vph, and shows that the CV is less than 1 in all samples. Therefore, the CV from this study is generally shorter than corresponding values from international research, indicating that a motorist in Iraq leaves a shorter headway from the car ahead than corresponding drivers in the developed world in which the past research was conducted. This finding may reflect the difference in traffic conditions, particularly driving behavior. Such differences may be attributable to the fact that driving behavior Iraq tends to be more aggressive

11 Conclusion

From the present research and after the analyzing of the observed headway time and flow data for the selected location, it was concluded the following:

1- The 2^{nd} order Polynomial model was best fitting the time headway-flow relationships for interrupted flow conditions.

2- The results of *P* value are greater than 0.05, this is within the range of acceptable deviation.

3- The data analysis in this study showed that driving behavior, in terms of time headway, in Baghdad is different from that in the developed world. That is, motorists in Baghdad follow one another more closely compared with drivers in developed countries. It should be said that the flow boundaries, as well as the results from the study, are valid for the collected data in Baghdad and should not necessarily reflect traffic conditions elsewhere.

12 References

Al-Ghamdi Ali S., “Analysis of Time Headways on Urban Roads: Case Study from Riyadh”, *Journal of Transportation Engineering*, Vol. 127, No. 4, July/August 2001, pp. 289-294

Garber, N.J.,and Hoel, L.A.(1997).”Traffic and Highway Engineering.” 2nd Edition.

Guohui Zhang, Yinhai Wang, Heng Wei, Yanyan Chen, “Examining Headway Distribution Models Using Urban Freeway Loop Event Data TRB 2007 Annual Meeting CD-ROM, Paper revised from original submittal, http://www.uwstarlab.org/STARLab_Papers/2007_TRB

Hoogendorn S P and Bovy P H L, “Estimation and analysis of multilane multiclass headway distribution model”, European Transport Conference 1999, Traffic Management Safety and Intelligent Transport Systems.

Khisty, C., J. and Lall, B., K.”Transportation Engineering, An Introduction”, The United States of America: Prentice-Hall, Inc., 1998.

Luttinen R. Tapio, “Statistical Analysis of Vehicle Time Headways”, PHC thesis submitted to University of Technology Lahti Center (Neopoli, Lahti, Finland) , Otaniemi, 1996.

Rattaphol Pueboobpaphan and Dongjoo Park, “Characteristics Of Probe Vehicles’ Time Headway”, Journal of the Eastern Asia Society for Transportation Studies, Vol.5, October, 2003 , p 2065- 2076.

Transportation Research Board (TRB), Highway Capacity Manual. National Research Council. Washington, D.C. ,2000.

Vicki Sharp, "Statistics for the Social Sciences",2000

Table (1):An example of flow data (North Direction)

Time	Flow (PC)	Average Headway (sec)
11:30-11:35	204	1.5
11:35-11:40	201	1.6
11:40-11:45	218	1.4
11:45-11:50	220	1.4
11:50-11:55	213	1.3
11:55-12:00	214	1.2
12:00-12:05	226	.9
12:05-12:10	198	1.6
12:10-12:15	208	1.6
12:15-12:20	199	1.8
12:20-12:25	236	.8
12:25-12:30	263	.5

Table (2): An example of flow data (South Direction)

Time	Flow (PC)	Average Headway (sec)
11:30-11:35	192	1.6
11:35-11:40	195	1.5
11:40-11:45	185	1.8
11:45-11:50	195	1.7
11:50-11:55	197	1.7
11:55-12:00	195	1.7
12:00-12:05	190	1.6
12:05-12:10	215	.9
12:10-12:15	201	.9
12:15-12:20	216	.7
12:20-12:25	209	.8
12:25-12:30	189	2

Table (3) Descriptive statics of the Time Headway Data/ south approach

Hours of survey	N	Flow /hr.	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic		Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
1st hour	12	1900	.50	1.80	1.3000	.11146	.38612	.149
2nd hour	12	1890	.90	1.80	1.3500	.07638	.26458	.070
3rd hour	12	1697	.70	2.00	1.4083	.12995	.45017	.203
4th hour	12	589	3.30	4.20	3.5500	.08483	.29388	.086
5th hour	12	393	4.20	6.00	5.1667	.15635	.54160	.293
6th hour	12	312	5.50	6.00	5.7833	.05482	.18990	.036
7th hour	12	843	.50	2.00	1.1750	.14571	.50475	.255
8th hour	12	348	5.20	6.00	5.7500	.09170	.31766	.101

Table (4): Descriptive statics of the Time Headway Data / North approach

Hours of survey	N	Flow /hr	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic		Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
1st hour	12	2006	.90	1.60	1.3000	.07977	.27634	.076
2nd hour	12	1236	.60	1.80	1.3083	.10833	.37528	.141
3rd hour	12	1708	.70	2.10	1.4083	.14059	.48703	.237
4th hour	12	602	3.20	4.20	3.5333	.09073	.31431	.099
5th hour	12	498	4.20	6.00	5.1583	.15592	.54013	.292
6th hour	12	333	5.40	6.10	5.8000	.06276	.21742	.047
7th hour	12	1222	2.20	4.30	2.7000	.17233	.59696	.356
7th hour	12	862	.50	2.00	1.1750	.14571	.50475	.255

Table (5) Model Summary and Parameter Estimates/South direction

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	b1	b2	b3
Linear	.411	3.485	1	5	.121	.019		
Logarithmic	.717	12.644	1	5	.016	.699		
Exponential	.367	2.899	1	5	.149	.006		
Logistic	.367	2.899	1	5	.149	.994	.000	
Power	.676	10.453	1	5	.023	.218	-.002	4.939E-6
Polynomial order2	.999	53.542	1	5	.001	98.893		
Growth	.367	2.899	1	5	.149	.006		

Table (6) Model Summary and Parameter Estimates/North Direction

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Linear	.860	1032.149	1	4	.000	6.592	-.028
Logarithmic	.852	79.468	1	4	.001	14.446	-2.430
Power	.703	16.291	1	4	.016	105.291	-.790
S	.577	5.451	1	4	.080	.505	41.491
Growth	.860	96.614	1	4	.001	2.172	-.010
Exponential	.911	96.614	1	4	.001	8.775	-.010
Logistic	.860	96.614	1	4	.001	.114	1.010

Table (7). Chi-Square Test for Headway data (South Direction)

Time Headway	Observation Time Headway			Theoretical Time Headway		Chi-Square
	Frequency	% Frequency	%Cumulative of Freq.	Frequency	%Cumulative of Freq.	
1	33	205	100	205.7496	100	0.002731
2	39	172	83.9	169.5064	82.3848	0.036683
3	30	133	64.9	135.3704	65.79376	0.041507
4	28	103	50.2	103.3416	50.22688	0.001129
5	30	75	36.6	73.42	35.68415	0.034002
6	45	45	22	45.6056	22.16558	0.008042
Total	205					**0.124094**

Table (8). Chi-Square Test for Headway data (North Direction)

Time Headway	Observation Time Headway			Theoretical Time Headway		Chi-Square
	Frequency	% Frequency	%Cumulative of Freq.	Frequency	%Cumulative of Freq.	
1	42	204	100	199.187	96.81039	0.116298
2	36	162	79.41176	163.644	79.53551	0.016516
3	38	126	61.76471	128.101	62.26063	0.034459
4	33	88	43.13725	92.558	44.98575	0.224458
5	28	55	26.96078	57.015	27.71087	0.071213
6	27	27	13.23529	21.472	10.43599	1.423192
	204					1.886136

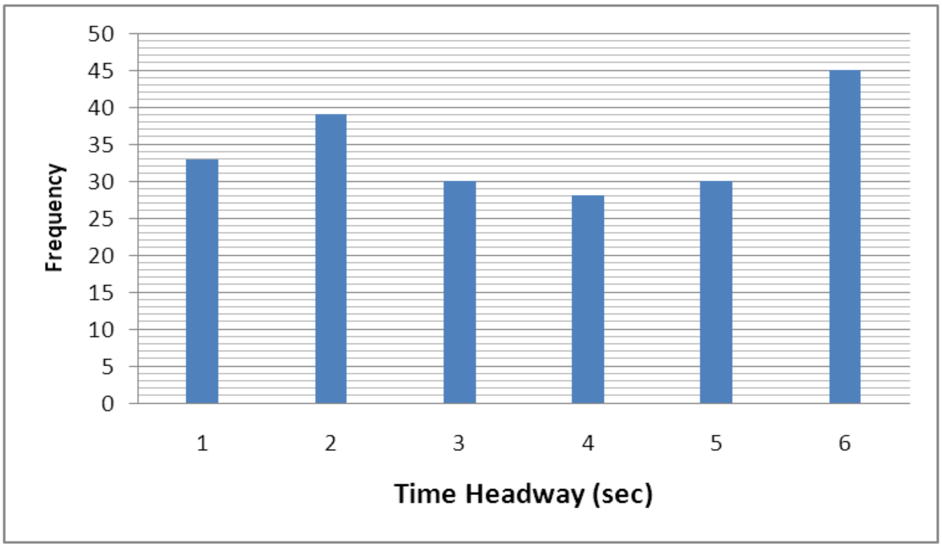

Figure (1) the observed time Headway (South direction)

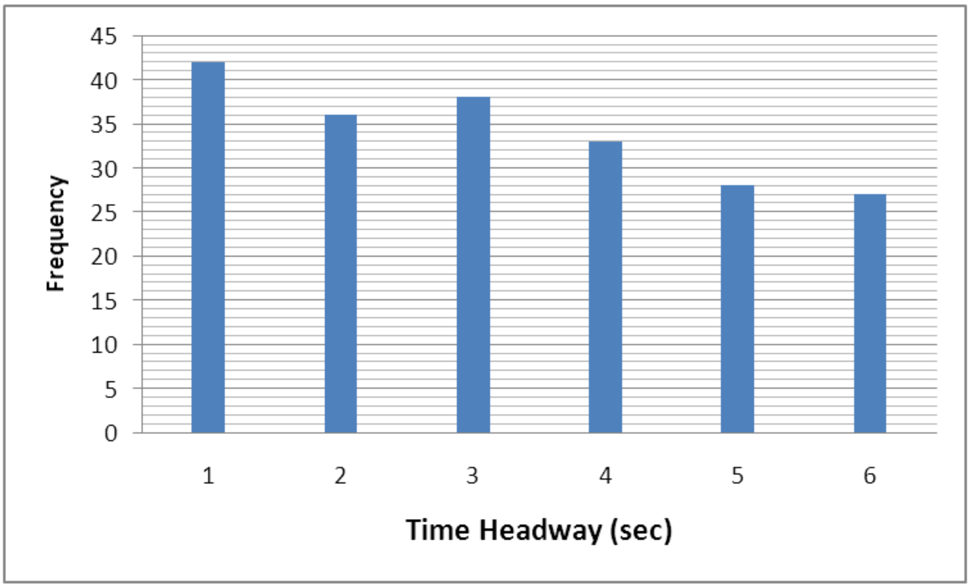

Figure (2) the observed time Headway (North direction)

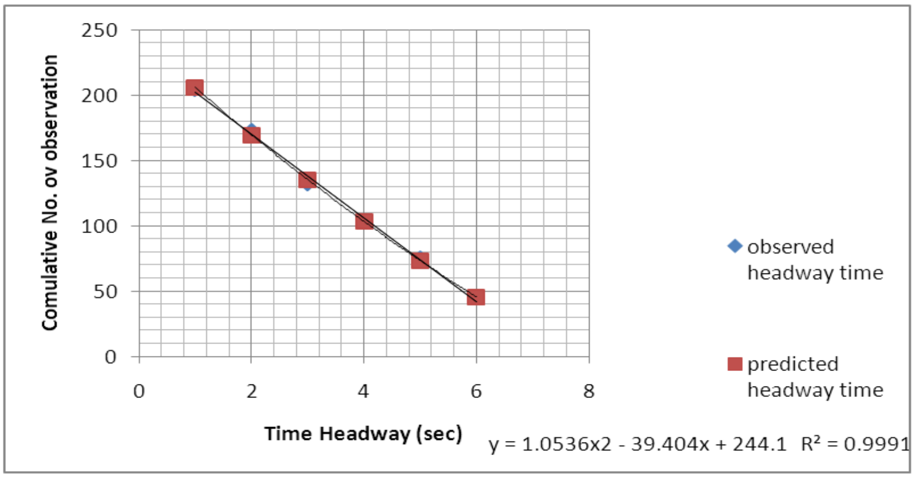

Figure (3) the predicted model of headway time (South direction)

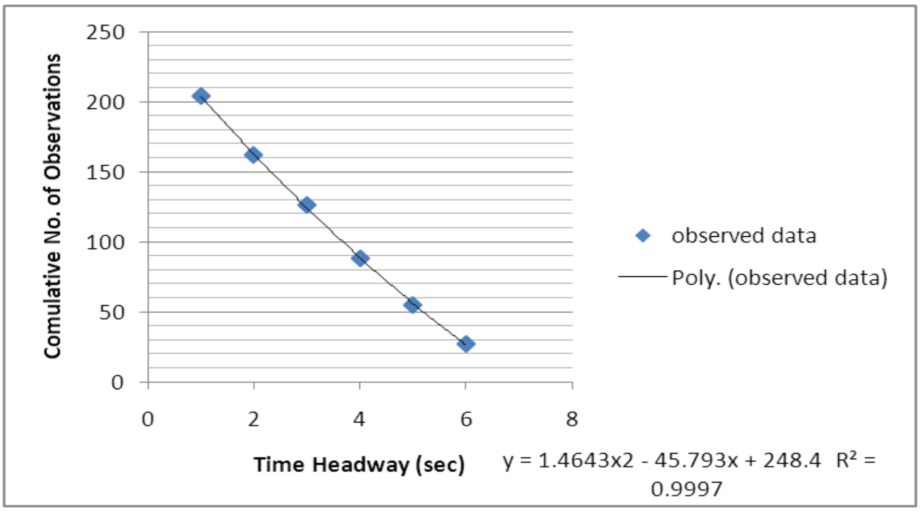

Figure (4) the predicted model of headway time (North direction)

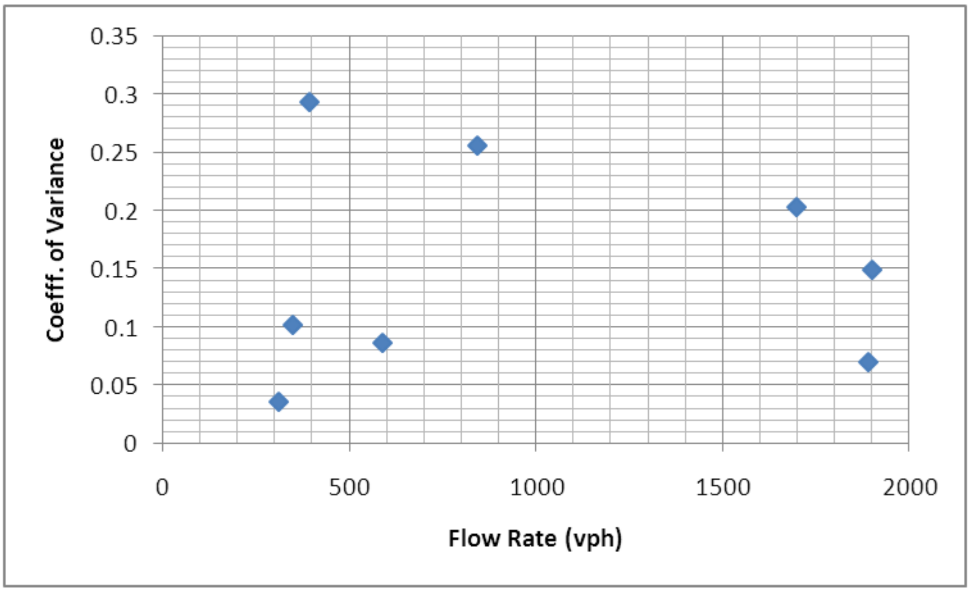

Figure (5) Coefficient of variation of time headway versus flow rate / South Direction

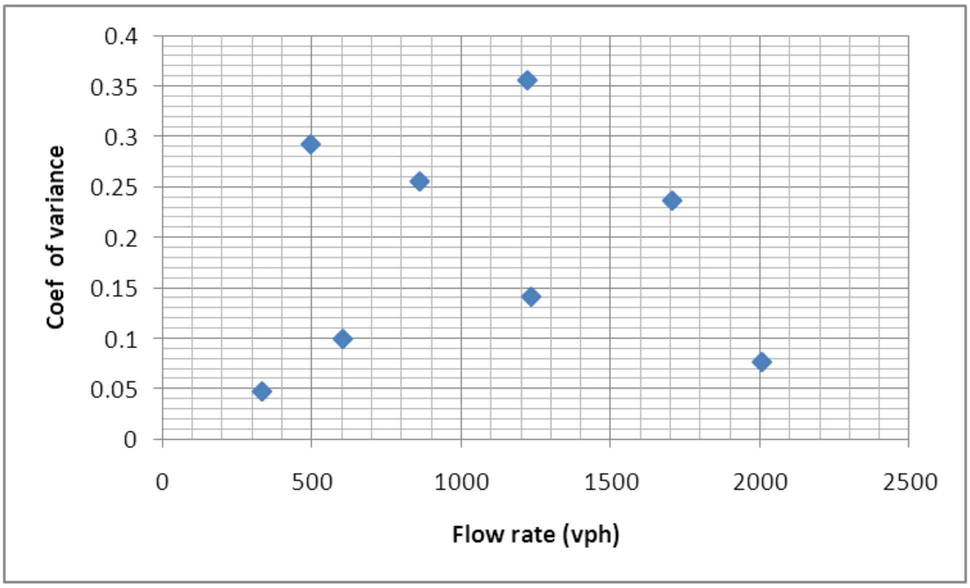

Figure (6) Coefficient of variation of time headway versus flow rate / North Direction