INCIDENT DETECTION ON THE BURLINGTON SKYWAY

THE MCMASTER INCIDENT DETECTION ALGORITHM

BY

EMILY S. A. PURCHASE

A Research Paper

Submitted to the Department of Geography

in Fulfilment of the Requirements

of Geography 4C6

McMaster University

April 1990

ABSTRACT

The McMaster Incident Detection Algorithm (MacAlg) automatically detects incidents on the Burlington Skyway for the Burlington Freeway Traffic Managment System (FTMS). This paper describes the calibration, testing and evaluation of functions of northbound stations 1 through 6. The testing and evaluation of the two weekly data sets is illustrated and discussed. Some of the resulting functions are recommended to the Burlingtion FTMS to evaluate how well the MacAlg detects incidents. This research compliments the work of Persaud, Hall and Hall (1989), who are developing and testing the logic of the MacAlg. The results of this paper contribute information to the further development and testing of the MacAlg's logic.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. F.L. Hall, for suggesting this research topic, and for providing guidance. I would also like to thank Don Cleghorn for the use of his Fortran programs, Lisa M. Hall for explaining how to use them, and Geof_Gray for the help with Quattro.

Special thanks to my parents who helped me develop my research and writing skills over the years, making this a more enjoyable reasearch project. Thanks also to my housemates for keeping the noise down to a dull roar.

TABLE OF CONTENTS

TABLE OF ILLUSTRATIONS

		Page
Figure A5.	Regression for Skyway NB4-1	34
Figure A6.	Regression for Skyway NB5-1	35
Figure A7.	Regression for Skyway NB5-2	36
Figure A8.	Regression for Skyway NB6-1	37
Figure A9.	Regression for Skyway NB6-2	38

INTRODUCTION

The purpose of this thesis is to calibrate functions for testing and evaluating the McMaster Incident Detection Algorithm (MacAlg). The MacAlg gathers information from the road detectors of the Freeway Traffic Management System (FTMS) on the Burlington Skyway. Six northbound stations, 1 through 6, have been put into operation by the Burlington FTMS this fall. These stations will be the focus of the research for the algorithm's calibration and evaluation.

Information is acquired by the algorithm from road detectors for each station independently. Each station is made up of two diamond-shaped induction loop detectors implanted under the pavement of a traffic lane, one upstream and one downstream, which detect the vehicle's presence or absence through magnetic induction. Three variables are recorded by the computer at the FTMS from magnetic induction. These three variables are flow, occupancy and speed. Flow relates to the number of vehicles that cross each loop in a thirty second interval. Occupancy is the percentage of time in a thirty second interval that vehicles remain over each of the two loops. Speed is the average in kilometers per hour that vehicles are going when they cross the station. With this information, the algorithm can detect incident congestion located upstream or downstream of a station. Persaud, Hall and Hall (1989) explain how these variables are combined within the logic of the MacAlg to improve incident detection in relation to using one variable. The California Algorithm, for example, uses only occupancy to detect incidents (Payne and Tignor, 1978). If one variable is not available because of insufficient data, the MacAlg can still detect incidents using the other two variables.

One strength of the MacAlg is the data screening logic it has for accounting for non-incident related changes in traffic (Persaud, Hall and Hall, 1989). Data screening tests are made in separate "if blocks." These "if blocks" screen out problems that affect the quality of the information the FTMS is gathering before the algorithm program decides whether or not an incident has occured. One possible drawback in using the screening tests is the inelasticity of boundaries made for flow, occupancy and speed resulting in the possible loss of valuable data beyond these boundaries.

When it detects an incident, the algorithm records an alarm. If the operator has not detected the incident, the alarm is considered false. In other words, the algorithm has detected a non-existent incident. This thesis is concerned with the procedures used to set thresholds on the information the algorithm reads so that the false alarms will be minimized and, eventually, will not occur.

The objective of this paper is to calibrate the information of the six previously unevaluated northbound stations to get the coefficients for traffic flow and the speed thresholds for congestion. The findings of this research are evaluated for recommendation to the Burlington FTMS to test and evaluate the algorith's ability to detect incidents.

LITERATURE REVIEW

The focus of the research project is to calibrate functions for testing and evaluating the McMaster Incident Detection Algorithm (MacAlg). This review discusses literature that relates to incident detection and the MacAlg. First, Freeway Traffic Management Systems (FTMS's) are discussed, and the need for automatic incident detection. Second, previous approaches to automatic incident detection, specifically the currently used California algorithm, are described. Third, the MacAlg is compared to the California algorithm. Finally, the issue of currently on-going calibration and testing for the MacAlg is discussed as it directly relates to the research project.

Freeway Traffic Management Systems were introduced to North American freeways in the early 1960's (Gall, 1988, p. 1). One early system in the United States was the Los Angeles Area Freeway Surveillance and Control Project in California (LAAFSCP) while the Texas Transportation Institute (TTI) studied a different FTMS in its state (Cook and Cleveland, 1974, pp. 2-3). The first Canadian example of FTMS is the Queen Elizabeth Way (QEW) freeway surveillance and control system demonstration project which became operational in Mississauga, Ontario, in July of 1975 (Case and Williams, 1978, p. 84).

Case and Williams (1978) explain several reasons why FTMS's should be used on freeway systems (p. 84):

1. The continuous increase in traffic on the freeway system,
2. The appearance of congestion on the freeway system,
3. The high cost of constructing or reconstructing freeways,
4. Public aversion to more or bigger urban freeways, and
5. Favorable results from similar projects in the United States.

The purpose of the FTMS is to detect congestion occurring on the freeway systems so that any obstructions may be removed and traffic flow returns to a free-flowing pattern as soon as possible.

One of the most important aspects of the FTMS is automatic incident detection. A traffic incident is defined as a random, unpredictable event that temporarily reduces the capacity of a section of freeway significantly below normal (Dudek and Messer, 1974, p. 12). The purpose of the incident detector is to automatically detect these incidents and alert the FTMS operator to the obstruction of traffic. In the experimental warning system used on the Gulf Freeway in Houston Texas, the detector monitored the freeway for the presence of congestion which creates a growing traffic queue uptream of an incident (Dudek and Messer 1973, p. 1). Data were collected by "double-loop detectors... positioned on each lane.... Traffic flow data from detectors are transmitted to [a]... computer located in the surveillance and control center" (Dudek and Messer, 1973, p. 4).

The logic of automatic incident detection is found in the computer algorithm which monitors traffic through the detector stations. Algorithms vary, as well as the theories behind them. Gall (1988) suggests that the

> underlying concept of most automatic incident detection algorithms [arises from] roots in the hydrodynamic analogy of traffic flow, where the flow of traffic is seen as like the flow of a compressible fluid, rather than as a flow of discrete particles (p. 2).

The algorithm must discriminate between congested and uncongested patterns in traffic flow and "provide indications of the probable presence of freeway incidents by processing electromagnetic surveillance data" (Payne and Tignor, 1978, p. 30).

As the detection of incidents is based upon probable interpretations of patterns of congested and uncongested traffic, previous algorithms have generated a high level of false alarms. Payne and Tignor (1978) claim that the California algorithm is superior to previous algorithms (p. 31). It has three functions, rather than one, that are used in combination to distinguish patterns of traffic incident-related data (Payne and Tignor, 1978, p. 31). Occupancy data are averaged for all the lanes at a single detector station over a one minute interval and are then compared with the occupancies at adjacent (upstream or downstream) stations. When an incident occurs between two stations, capacity of traffic is reduced at the site, creating a high occupancy traffic backup at the upstream detector and a low occupancy cleared region at the downstream detector (Payne and Tignor, 1978, p. 32). If the incident has created a significant difference in average occupancies between the two stations, the three occupancy functions, which must be triggered in sequence, will exceed the calibrated thresholds and an incident is declared (Gall, 1988, p. 4).

There are at least two weaknesses of the California algorithm. One is the inelasticity of the thresholds, which results in undetected incidents occuring above the threshold and the detection of non-incidents (false alarms) below the threshold. Another major weakness in the California algorithm has to do with detector malfunction (Gall, 1988, p. 6). Because the algorithm detects incidents by comparing occupancy at adjacent detector stations, if one detector fails, incidents cannot be detected by the detector upstream and the detector downstream of this station (Persaud, Hall and Hall, 1989, p. 20). This problem is solved by using an algorithm that detects incidents from each detector separately. The MacAlg is such an algorithm.

The MacAlg was created and is currently being developed by the Transportation Research Group at McMaster University (Persaud, Hall and Hall, 1989) as an alternative to previously developed algorithms. This single-station algorithm uses 30-second averages of speeds and flows as well as the occupancy variable used in the California algorithm. The interpretation of traffic patterns using these three varables is illustrated in Persaud and Hall's (1989) paper in which they apply the cusp-catastrophe theory developed by Gilchrist and Hall (6) to describe the relationship among three-dimensional traffic variables in the implication of incident detection.

The MacAlg was put on-line with the computer at the Ontario Ministry of Transportation FTMS in Burlington in 1988. Although the Burlington FTMS is using a California-type algorithm for incident detection, the MacAlg is being tested on the computer in the "background" (Persaud, Hall and Hall, 1989, p. 3). This means that the MacAlg test results do not go to the system operators when it detects an incident, but are filed on disk to be retrieved for comparison with the system operator's log of traffic incidents (Persaud, Hall and Hall, 1989, p. 3).

Persaud, Hall and Hall (1989) identify three main tasks in the calibration of functions for testing and evaluating the MacAlg (p. 8):

> 1. distinguishing between congested and uncongested flow-occupancy regions,
> 2. identifying a speed threshold to distinguish congested from uncongested speeds, and
> 3. establishing the duration of the persistance checks.

In selecting speed thresholds to distinguish congested from uncongested traffic, they compromise

> the detection rate with... the number of tolerable false alarms. This implies that not all "detectable" incidents will be detected. Since the detection rate is inversely proportional to the false alarm rate, some thresholds produce high detection rates at the cost of a high number of false alarms (Gall, 1988, p. 5).

Persaud, Hall and Hall (1989) found four strengths of the MacAlg. First, by using the combination of speed, flow, and occupancy at the same time, incident detection was increased (p. 19). This has an advantage over comparative Califronia-type algorithms. Second, unlike the California algorithm, the MacAlg's detection system is not as severely limited if a detector fails because it detects incidents from each detector station separately instead of comparatively (Persaud, Hall and Hall, 1989, p. 20). Third, the California algorithm detects an incident only at "critical occupancies," whereas the MacAlg is based on two other variables and can therefore detect incidents with a larger probability at occupancies that are less than critical (Persaud, Hall and Hall, 1989, p. 20). Finally, when an incident is detected by the California algorithm "incident detection is automatically suppressed at several stations surrounding a declared incident" (Persaud, Hall and

Hall, 1989, p. 21). In contrast, the MacAlg detects an incident at one station, sounds an incident alarm, and may sound a second alarm at an adjacent station if the congestion queue backs up that far, or to a further station (Persaud, Hall and Hall, 1989, pp. 20-1).

The current research is directly related to the work in the paper by Persaud, Hall and Hall (1989). It follows the methods and the Fortran programs they used, as well as using their results as a guideline for what results may be expected from this study. The 4C6 research also uses data gathered by the computer at the Burlington FTMS, but from the six previously unevaluated northbound detector stations on the Burlington Skyway.

DATA AQUISITION

As stated in the introduction, the purpose of this thesis is to calibrate functions for further testing and evaluation of the MacAlg. Although the MacAlg has been in operation on the Burlington Skyway since last summer, six northbound road detector stations (1 through 6) have been put into operation only this fall and therefore have not previously been tested.

Flow, occupancy, and speed data from the weeks of September 09, 1989 (890908) and December 01, 1989 (891201) from lanes 1 and 2 of the six stations are used for the calibrations. The Burlington FTMS collects for each lane two sets of data per minute, sixty minutes per hour, twenty-four hours per day, seven days a week or 20160 data sets per week. These data are saved on magnetic tapes in weekly summary files for later study. There is too much data on each summary file to store it in the personal VAX directory used to calibrate the data, so tapes are used.

As the summary file contains data on thiry stations, the "resky.for" (Resky) Fortran program reads the file, removes the desired data and puts them in the format required for this project. This program and the other programs used for calibration were created by Don Cleghorn, Master's Civil Engineering student at McMaster University. Instructions on the use of these programs were given by Lisa M. Hall, undergraduate Civil Engineering student at McMaster Univeristy.

A configuration file (Hall, 1989, p.7) is used to tell Resky to obtain data for each of the lanes separately because the summary file contains data for all lanes of the northbound and southbound stations on the Burlington Skyway. Separate lane files are created for calibration because, as the results indicate, there is no one constant function for all of the stations, so the calibrations must be done for each lane separately.

Northbound station 1 lane 1 is used in this paper as an example to explain and discuss the steps of calibration. It does not necessarily reflect the results found in the other stations.

First the configuration file is set for Resky to gather the data for NB1-1. When it is run, Resky asks for the input summary file (eg. 890908), which is found in the scratch sub-directory created to contain the summary files, an output file name for NB1-1, and the time boundaries to be used. Data for the whole week were used.

Resky also provides output on the percent of bad data for flow, occupancy, and speed. Table 1 illustrates the percent of bad data within each file. Bad data has not remained constant for any lane between the two weeks. This difference is one reason why more than one summary file is used for calibrating, testing, and evaluation. For example, the data for NB3-1 and 2 are almost all bad in 890908 while the bad data in 891201 are much lower and comparable with other stations of that week. The same observation is made for NB 6-1 and 2. Except for Stations 3 and 6, bad speed data has increased between 890908 and 981201. Because of the percent of the high percent of bad data, Station 3 for 890908 is dropped from those stations whose calibratios are to be recommended for use of the MacAlg.

Resky is a valuable program because it illustrates which of the lanes or stations has remarkably high bad data percentages. For 890908, an unacceptable level of bad data is discovered in NB3-1&2, and more than half of the data in NB6-1&2 are bad. Although calibrations can be done on these specific lanes, their data results in unreliable calibrated functions. On the other hand, Resky has indicated a high level of good data for the other lanes which should result in acceptable functions. The Ministry of Transportation in Burlington has been tuning the road detectors which may account for the improvement in the data received from stations 3 and 6 between the two summary files.

Table 1. Percentage of Bad Data, Stations 1 Through 6, Lanes 1 and 2.

890908

Stations →	1		2		3		4		5		6	
Lanes →	1	2	1	2	1	2	1	2	.1	2	1	2
Variables ↓												
Flow	0.6	0.6	0.5	0.5	99.7	99.7	0.5	0.5	0.5	0.5	54.0	54.0
Occupancy	0.6	0.6	0.5	0.5	99.7	99.7	0.5	0.5	0.5	0.5	54.0	54.0
Speed	6	5.3	2.2	4.7	99.7	99.7	14.2	5.2	18.3	5.3	60.8	56.0

891201

Stations →	1		2		3		4		5		6	
Lanes →	1	2	1	2	1	2	1	2	1	2	1	2
Variables ↓												
Flow	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.2
Occupancy	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.2
Speed	22.3	8	27.3	7.3	22.7	12.9	19.2	7.5	22.9	7.8	22.3	8.1

CALIBRATION

"Cal.for" (Cal) is the Fortran program used for calibrating functions for testing and evaluating. It finds the parameters the MacAlg uses in incident detection (Hall, 1989, p. 8). In this study, Cal.for finds the critical speed for each lane, the flow-occupancy data scatter plot before and after the critical speed cut-off, and regression input for the equation (function) obtained through the Minitab program.

Critical Speed (Threshold)

Critical speed is defined as the lowest speed for uncongested operation. Critical speed is used to separate congested data from uncongested data in a lane because,

> In order to find a function relating volume [flow] and occupancy for uncongested operations at a station, it is necessary to eliminate the congested data from the data set. This is done on the basis of speed above which operations are always uncongested (Hall and Hall, 1989, p. 3).

There are several steps to find the critical speed using Cal. First, an initial speed of zero km/h is input. All data below this speed is deleted. For this new set of data, Cal calculates the mean and three standard deviations from the mean and produces a new critical speed. The new critical speed is input into Cal for the proceedure to be repeated until the critical speed is above the mean minus three standard deviations. This is obtained when the critical speed of the current set is the same as the one before it. For example, when the Resky file for NB1-1 is input into the critical speed option for Cal, the result is a critical speed of 57 km/h for 890908 and 80 km/h for

891201, a 23 km/h difference. It cannot be assumed that either speed is correct. However, it may be argued that 891201 has the correct critical speed because its good data points number 19688 out of 20160, approximately 1100 more points than the 8680 points in 890908. Obtaining a critical speed from other summary files for NB1-1 may reveal a more valid critical speed.

Critical speed varies station to station, lane to lane. Therefore, one critical speed cannot be set as a constant for all stations monitored by the algorithm. Table 2 illustrates the critical speeds for each of the twelve lanes for 890908 and 891201. Many of the critical speeds are in the mid-seventy km/h. Several lane speeds are identical or similar between weeks. Indentical critical speeds are observed for NB4-2 at 73 km/h and NB5-1 at 80km/h. Critical speeds within 1 km/h are observed at NB1-2, NB2-2, and NB4-1.

Table 2. Critical Speed

	Summary Files	
	890908	891201
Station-lanes		
NB1-1	57	80
NB1-2	78	76
NB2-1	-72	80
NB2-2	77	76
NB3-1	--	83
NB3-2	--	31
NB4-1	79	80
NB4-2	73	73
NB5-1	80	80
NB5-2	69	79
NB6-1	77	80
NB6-2	73	78

Critical speeds for 891201 (with exception of NB3-2) are between 73 and 83 km/h. Critical speed is low for NB3-2 on 891201 indicating a possible problem with the station detector because the percent of bad data decreased between 890908 and 891201 (Table 1). Other stations in 891201 have a higher percent of bad data than NB3-2 but have acceptable critical speeds while the critical speed for NB3-2 is very low, and istherefore unreliable.

It is important to obtain the correct approximation of the critical speed for each lane because the critical speed is used as an input for the flow-occupancy data plot and the regression function obtained through minitab, and the data points of the lower bound. If the critical speed is overestimated, the threshold is lower than it should be, including data points that are not incident points. An underestimation puts the critical speed too high and results in incident points falling below the threshold. The critical speed should fall in a place to obtain calibrated data that maximizes incident detection at a minimum false alarm rate. Theoretically, no false alarms should occur and all incidents should be detected.

Table 3 illustrates the speed histograms for frequency of speed for NB1-1 for 890908 and 891201. It illustrates where the critical speed is found for each week. The MacAlg considers all speeds including and above critical speed to be 100 percent uncongested, while the speeds below critical speed are considered not uncongested.

Notice that the critical speed is found at the base of the largest cluster of speed points, or frequency curve, on the histogram.

Table 3. Speed Histogram for NB1-1, 890908 and 891201

890908		891201	
Speed	Frequency	Speed	Frequency
35	0	58	0
36	0	59	0
37	0	60	0
38	0	61	2
39	0	62	0
40	2	63	0
41	0	64	1
42	0	65	0
43	0	66	1
44	10	67	1
45	0	68	1
46	0	69	0
47	0	70	0
48	0	71	2
49	0	72	3
50	1	73	7
51	0	74	4
52	0	75	0
53	7	76	9
54	0	77	7
55	2	78	15
56	0	79	9
o 57	0	o 80	10
58	3	81	20
59	0	82	25
60	3	83	33
61	4	84	142
62	0	85	50
63	2	86	60
64	2	87	80
65	3	88	95
66	2	89	131
67	4	90	184
68	4	91	901
69	4	92	321
70	4	93	410
71	3	94	548
72	8	95	839
73	21	96	962
74	9	97	789
75	9	98	879
76	9	99	2721
77	23	100	834
78	112	101	815
79	19	102	882
80	20	103	566
81	34	104	723
82	34	105	404
83	95	106	305
84	539	107	98
85	103	108	118
86	173	109	919
87	228	110	55
88	266	111	31
89	276	112	32
90	445	113	14
91	1432	114	7
92	349	115	57
93	395	116	5
94	388	117	8
95	322	118	1
96	350	119	0
97	170	120	2
98	253	121	119

* Critical speed

For all of the lanes, critical speed is found below this curve, just above three standard deviations from the mean. It should be noted that critical speed is not the mean speed, which for NB1-1 will be in the 90-110 km/h range.

In recommending critical speed thresholds as parameters used by the MacAlg for the Burlington FTMS, there are several responses according to the results found in this study. Table 4 illustrates the critical speeds recommended by this research to the Burlington FTMS for use by the MacAlg. For 890908, critical speeds were immediately thrown out due to a high percent of bad speed data for NB3-1 and 2, and NB6-1 and 2. NB2-1 is also rejected because it is a negative critical speed. The critical speed for NB3-2 for 891201 is rejected as it has been inferrd that the detector station is malfunctioning. The critical speeds for NB4-2 and NB5-1 are the same for both weeks and therefore have satisfactory critical speeds for parameters. Concerning the stations NB1-2, NB2-2, and NB4-1, with critical speeds differing by one km/h of each other, one critical speed will be refected according to the percentage of bad speed data illustrated in Table 1. According to this criterion, the speeds for 891201 are rejected for the three stations. Although the bad speed data is lower in 890908 than 891201 for NB5-2, 891201 is chosen because it is closer to the range of chosen critical speeds than 890908. It is difficult to determine which of the two NB1-1 critical speeds is correct. Although the 890908 critical speed has a much lower percent of bad speed data, its critical speed is very low while the speed of 891201 is in the 70 to 80 km/h range that the other critical speeds are chosen from. A look at the flow-occupancy function for each of these weeks illustrates that the criterion speed for 891201 is the better choice. Although there is a problem with the critical speed data for NB1-1 890908, it is still used as an example and to investigate why this speed is so low, considering that it could not be detected from the Resky data.

Table 4. Critical Speeds Considered for Recommendation to the Burlington FTMS for Testing and Evaluation.

	Summary Files		
	890908	891201	Recommended Speeds
Station-lanes			
NB1-1	57	80	80
NB1-2	78	76	78
NB2-1	--	80	80
NB2-2	77	76	77
NB3-1	--	83	83
NB3-2	--	--	--
NB4-1	79	80	79
NB4-2	73	73	73
NB5-1	80	80	80
NB5-2	69	79	79
NB6-1	--	80	80
NB6-2	--	78	78

The Flow-Occupancy Function

Another option in Cal plots the flow versus occupancy variables to obtain a scatter plot of the points for regression. The critical speed found earlier is used as input to remove the data considered to be congested thus far. The points remaining are those at or above critical speed.

Tables 5 and 6 are the flow-occupancy plots before and after critical speed cut-off for NB1-1 890908. After the critical speed cut is taken, 69 points are cut off of Table 6. At 7 percent occupancy, one data point is taken off the bottom of the column at flow (volume) 240. This is an isolated data point which fell below the critical speed of 57 km/h therefore not fitting into the scatter plot of uncongested data points which are used to find the regression equation. As the equation describes the lower bound, the data used to find it should not include isolated data points below the critical speed that would throw the equation off.

The critical speed of 57 km/h was puzzlingly low considering only that its percent of bad data is lower than those of other stations with higher critical speeds. In Tables 5 and 6 it is observed that the data points are all clustered between zero and 10 percent occupancy. This formation of points is not typical of the other stations with low percent bad data in which the points spread out to the 17 and 19 percent occupancy range and may be up to thirteen rows per column (Table 7). This specific cluster group is sparce in data points which may result in a lower critical speed because of lack of data points. The problem of NB1-1 890908 is further illustrated in the lower bound calculations.

The data represented by Table 6 are input into Minitab to find the regression equation for each lane. The function used to find this equation is a quadratic polynomial with the form, y = C + Ax + Bx↑2 (Hall and Hall, 1989, p. 7). Table 8 contains the equations for each of the 12 lanes for both weeks.

Table 5. Flow-Occupancy Scatter Plot for NB1-1, 890908, at Critical Speed 0 km/h

Total Number of Data Points: 20159
Number of Good Points: 8680

Critical Points: 7621

Flow	Occupancy 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
3480	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3360	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3240	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3120	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2760	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2640	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2520	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2400	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2280	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2160	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2040	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1920	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1800	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1680	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1560	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
1440	0	0	0	0	0	0	0	0	8	5	1	0	0	0	0	0	0	0	0	0
1320	0	0	0	0	0	0	0	21	5	3	0	0	0	0	0	0	0	0	0	0
1200	0	0	0	0	0	0	28	26	3	0	0	0	0	0	0	0	0	0	0	0
1080	0	0	0	0	0	44	76	14	1	0	0	0	0	0	0	0	0	0	0	0
960	0	0	0	0	20	214	25	2	1	0	0	0	0	0	0	0	0	0	0	0
840	0	0	0	4	279	80	11	0	1	0	0	0	0	0	0	0	0	0	0	0
720	0	0	2	198	286	10	2	0	0	0	0	0	0	0	0	0	0	0	0	0
600	0	0	27	540	29	7	0	1	0	0	0	0	0	0	0	0	0	0	0	0
480	0	4	554	81	17	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
360	2	327	384	23	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
240	22	958	23	11	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
120	1356	24	19	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2883	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 6. Flow-Occupancy Scatter Plot for NB1-1, 890908, at Critical Speed 57 km/h

Total Number of Data Points: 20159
Number of Good Points: 8680

Critical Points: 7552

Flow	Occupancy 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
3480	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3360	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3240	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3120	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2760	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2640	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2520	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2400	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2280	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2160	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2040	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1920	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1800	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1680	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1560	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
1440	0	0	0	0	0	0	0	0	8	2	1	0	0	0	0	0	0	0	0	0
1320	0	0	0	0	0	0	0	21	5	3	0	0	0	0	0	0	0	0	0	0
1200	0	0	0	0	0	0	28	28	3	0	0	0	0	0	0	0	0	0	0	0
1080	0	0	0	0	0	44	76	14	1	0	0	0	0	0	0	0	0	0	0	0
960	0	0	0	0	20	214	25	2	1	0	0	0	0	0	0	0	0	0	0	0
840	0	0	0	4	279	30	11	0	1	0	0	0	0	0	0	0	0	0	0	0
720	0	0	2	198	286	10	2	0	0	0	0	0	0	0	0	0	0	0	0	0
600	0	0	27	540	29	7	0	1	0	0	0	0	0	0	0	0	0	0	0	0
480	0	3	554	80	17	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
360	2	327	382	23	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
240	22	957	23	11	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
120	1356	24	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2821	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 7. Flow-Occupancy Scatter Plot for NB1-1, 891201, at Critical Speed 80 km/h

Total Number of Data Points: [illegible]
Number of Good Points: 1958[illegible]

Critical Points: 15214

Flow \ Occupancy	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
3480	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3360	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3240	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3120	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2760	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2640	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
2520	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
2400	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	5	1	0	0	0
2280	0	0	0	0	0	0	0	0	0	0	0	0	3	4	10	5	2	1	0	0
2160	0	0	0	0	0	0	0	0	0	0	0	2	9	17	8	1	0	1	0	0
2040	0	0	0	0	0	0	0	0	0	0	0	10	31	17	3	2	1	0	0	0
1920	0	0	0	0	0	0	0	0	0	1	14	44	27	9	4	0	2	0	0	0
1800	0	0	0	0	0	0	0	0	0	11	60	57	12	0	1	0	0	0	0	0
1680	0	0	0	0	0	0	0	0	3	65	65	21	10	1	0	0	1	0	0	0
1560	0	0	0	0	0	0	0	2	64	135	50	12	6	2	0	0	0	0	0	0
1440	0	0	0	0	0	0	1	69	105	72	34	5	0	1	0	0	0	0	0	0
1320	0	0	0	0	0	0	14	271	60	43	12	6	2	0	0	0	0	0	0	0
1200	0	0	0	0	0	5	301	234	50	14	10	2	0	0	0	0	0	0	0	0
1080	0	0	0	0	1	195	411	65	25	8	1	2	0	0	0	0	0	0	0	0
960	0	0	0	0	58	632	62	37	6	3	0	0	0	0	0	0	0	0	0	0
840	0	0	0	13	732	214	75	15	5	0	0	0	0	0	0	0	0	0	0	0
720	0	0	2	447	665	80	20	7	1	0	0	0	0	0	0	0	0	0	0	0
600	0	0	70	1124	71	43	10	1	0	0	0	0	0	0	0	0	0	0	0	0
480	0	3	1328	153	61	2	3	1	0	0	0	0	0	0	0	0	0	0	0	0
360	0	641	907	62	7	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
240	36	1614	40	20	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
120	2244	30	36	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	4956	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 8. Regression Equations for 890908 and 891201

890908

Station-lane	Constant		Equation		Coefficients
NB1-1	0.462	+	1.67x	-	0.00449x↑2 †
NB1-2	-0.0553	+	1.56x	-	0.0457x↑2 *
NB2-1	negative critical speed--equation not available				
NB2-2	0.161	+	1.52x	-	0.0407x↑2 *
NB3-1	high percent bad data--equation not available				
NB3-2	high percent bad data--equation not available				
NB4-1	0.608	+	1.47x	-	0.0156x↑2 *
NB4-2	-0.105	+	1.56x	-	0.0425x↑2 *
NB5-1	0.836	+	1.16x	-	0.00963x↑2 *
NB5-2	-0.257	+	1.24x	-	0.0233x↑2
NB6-1	high percent bad data--equation not available				
NB6-2	high percent bad data--equation not available				

891201

Station-lane	Constant		Equation		Coefficients
NB1-1	0.810	+	1.47x	-	0.0183x↑2 *
NB1-2	1.03	+	1.37x	-	0.0419x↑2
NB2-1	0.712	+	1.25x	-	0.00928x↑2 *
NB2-2	0.717	+	1.30x	-	0.0307x↑2
NB3-1	0.532	+	1.68x	-	0.0239x↑2 *
NB3-2	high percent bad data--equation not available				
NB4-1	0.579	+	1.45x	-	0.0144x↑2
NB4-2	0.441	+	1.37x	-	0.0344x↑2
NB5-1	0.814	+	1.16x	-	0.00923x↑2
NB5-2	0.0295	+	1.15x	-	0.0215x↑2
NB6-1	0.572	+	1.44x	-	0.0133x↑2
NB6-2	0.0972	+	1.30x	-	0.0294x↑2

† ↑2 is used to express the x value to the exponent 2.

* These are the lanes chosen to be recommended to the Burlington FTMS according to the critical speeds chosen in Table 4.

Figure 1 compares the two regression lines for NB1-1 for the two weeks. The difference in the number of data points and how they are scattered is reflected in the graph. Both weeks have a lot of points in the lower percent occupancies but differ more in the higher percent range. As observed previously, Table 6 has no data points past 10 percent occupancy while Table 7, NB1-1 for 891201, has a scatter of data points up to 19 percent occupancy. The regression lines in Figure 1 reflect this as they would diverge from each other if plotted beyond the range of data. This does not occur with the majority of the stations (see Appendix A).

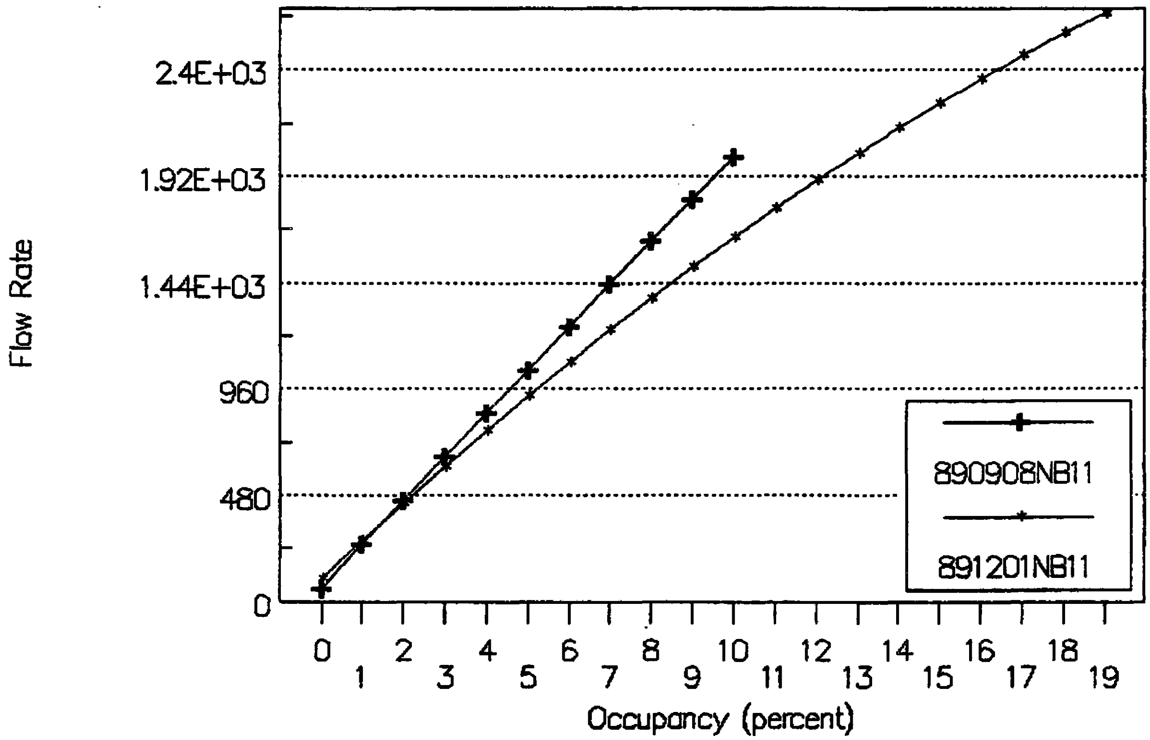

Figure 1. Regression for NB1-1, 890908 and 891201

Figure 2 illustrates the regression lines for NB4-2 in which the critical speeds are both 73 km/h, and the data plots are in general similar in scatter of data. This results in very similar curves in the regression lines.

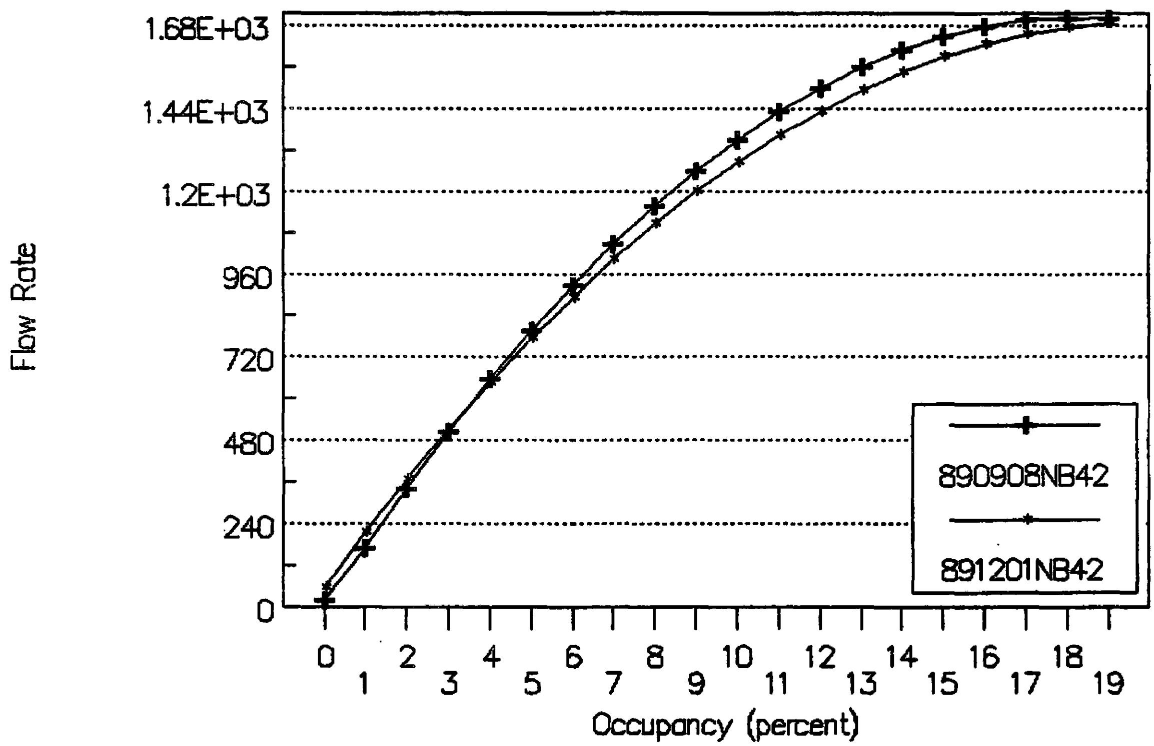

Figure 2. Regression for NB4-2, 890908 and 891201

The Lower Bound Calculation

The lower bound is the line that separates uncongested flow-occupancy data points from the transitional and congested points on a flow-occupancy scatter plot. According to Hall and Hall (1989) this is obtained manually by finding the constant difference value for each lane and subtracting it from the corresponding regression function.

The constant difference is half of the range of the flow rate data at each percent occupancy in vehicles per hour. It is assumed that the rage of flow data is the same for all the occupancy percentages. In this study, it is the average number of rows for each of the 17 columns, 3 to 19 percent occupancy, divided by two to get half the range, and multiplied by 120 to convert the data from vehicles per second to vehicles per hour. the 30-second flow data. Table 9 illustrates the average range of values for 890908 and 891201:

Table 9. Average Range of Values

	Summary Files	
	890908	891201
Station-lanes		
NB1-1	2.824	7.235 *
NB1-2	11.941 *	10.529
NB2-1	6.294	8.471 *
NB2-2	12.176 *	11.647
NB3-1	0.353	6.118 *
NB3-2	0.118	2.353
NB4-1	6.294 *	5.706
NB4-2	11.647 *	10.941
NB5-1	8.176 *	7.529
NB5-2	10.412	10.059 *
NB6-1	6.824	6.059 *
NB6-2	10.059	10.529 *

* These are the lanes chosen to be recommended to the Burlington FTMS according to the critical speeds of Table 4.

These values give a clue to the type of data in the flow-occupancy scatterplots. The small average range of rate of flow for NB1-1 in 890908 and Station 3's high percent of bad data are further illustrated by this table. The values marked with a star (*) are chosen to be recommended to the Burlington FTMS on the basis of the corresponding critical speed recommendations. The chosen range values average flow range of 9 rows per column indicating a fairly even spread of data points across the occupancy columns.

The relationship between the regression line and the lower bound is illustrated using NB4-2 for 890908 as an example because of its tight cluster of data points. Table 10 is a section of the NB4-2 flow-occupancy scatter plot. The number column range, from 7 to 13, is marked as well as the regression line and the lower bound. The regression points are found using a spreadsheet and the lower bound points found by taking the constant difference value, 11.647 * 120 / 2, and subtracting it from the regression points for each percent occupancy. The regression line is found locating the number at each column according to the numbers in the flow (volume) as illustrated in Table 10. The lower bound points are also located using according to the values of the flow column. In the case of NB4-2, one more data point is cut out of the plot at the lowest point of the scatter at 7 percent occupancy, and one at 10 percent occupancy.

CONCLUSIONS AND RECOMMENDATIONS

As illustrated in the tables and figures, the data sets are not constant between the two weeks of 890908 and 891201. Table 11 illustrates the critical speed, regression function, and the average of the range of values recommended to the Burlington FTMS for testing the MacAlg. These values are chosen from the better of the two weeks for

Table 10. Flow-Occupancy Scatter Plot for NB4-2, 890908, Illustrating the Regression Line and the Lower Bound

Total Number of Data Points: 20159
Number of Good Points: 19928

Critical Points: 19001

Flow	Occupancy 0	1	2	3	4	5	6	7	8	9	10
3480	0	0	0	0	0	0	0	0	0	0	0
3360	0	0	0	0	0	0	0	0	0	0	0
3240	0	0	0	0	0	0	0	0	0	0	0
3120	0	0	0	0	0	0	0	0	0	0	0
3000	0	0	0	0	0	0	0	0	0	0	0
2880	0	0	0	0	0	0	0	0	0	0	0
2760	0	0	0	0	0	0	0	0	0	0	0
2640	0	0	0	0	0	0	0	0	0	0	0
2520	0	0	0	0	0	0	0	0	0	0	0
2400	0	0	0	0	0	0	0	0	0	0	0
2280	0	0	0	0	0	0	0	0	0	0	0
2160	0	0	0	0	0	0	0	0	0	0	0
2040	0	0	0	0	0	0	0	0	0	0	0
1920	0	0	0	0	0	0	0	0	0	0	3
1800	0	0	0	0	0	0	0	0	1	2	23
1680	0	0	0	0	0	0	0	0	4	26	115
1560	0	0	0	0	0	0	0	2	18	167	227
1440	0	0	0	0	0	0	0	13	161	346	219 X
1320	0	0	0	0	0	0	5	197	436	304 X	177
1200	0	0	0	0	0	2	114	528	328 X	245	136
1080	0	0	0	0	1	45	579	452 X	241	153	128
960	0	0	0	0	8	418	613 X	317	144	145	67
840	0	0	0	2	199	665 X	272	219	104	62	37
720	0	0	0	55	749 X	221	209	79	48	20	18
600	0	0	7	625 X	273	185	61	47	17	11	4 ◆
480	0	2	344	453	170	71	46	15	8	1 ◆	1
360	0	94	810	171	95	43	13	6 ◆	0 ◆	0	0
240	19	916	170	214	51	13	0 ◆	1	0	0	0
120	987	138	178	4	1	6 ◆	0	0	0	0	0
0	849	2	2	0 ◆	0 ◆	0	0	0	0	0	0

* Range of Flow at 10 percent Occupancy

X Approximate location of the regression line

◆ Approximate location of the lower bound each lane as a result of data testing evaluation using the Resky and Cal Fortran programs. Although there are enough data points to calibrate functions for NB3-2 for 891201, testing indicates that these functions are unacceptable for recommendation and are therefore discarded.

Table 11. Values Recommended to the Burlington FTMS for Northbound Sations 1 Through 6

Station -Lane	Critical Speed	Regression Function	Average of the Range of Values
NB1-1	80	0.81 + 1.47x - 0.0183x↑2	7.235
NB1-2	78	-0.0553 + 1.56x - 0.0457x↑2	11.941
NB2-1	80	0.712 + 1.25x - 0.00928x↑2	8.471
NB2-2	77	0.161 + 1.52x - 0.0407x↑2	12.176
NB3-1	83	0.532 + 1.68x - 0.0239x↑2	6.118
NB3-2	high percent of bad data--values not available		
NB4-1	79	0.608 + 1.47x - 0.0156x↑2	6.294
NB4-2	73	-0.105 + 1.56x - 0.0425x↑2	11.647
NB5-1	80	0.836 + 1.16x - 0.00963x↑2	8.176
NB5-2	79	0.0295 + 1.15x - 0.0215x↑2	10.059
NB6-1	80	0.572 + 1.44x - 0.0133x↑2	6.059
NB6-2	78	0.0972 + 1.30x - 0.0294x↑2	10.529

A second objective that was not possible to accomplish within the research time frame questions how well the algorithm wirks using the functions in Table 11. To evaluate the algorithm's ability to detect incidents, it is compared with the on-duty operator's log at the Burlington FTMS for incidents detected, incidents missed, or false alarms.

APPENDIX A

Figures A1 to A9

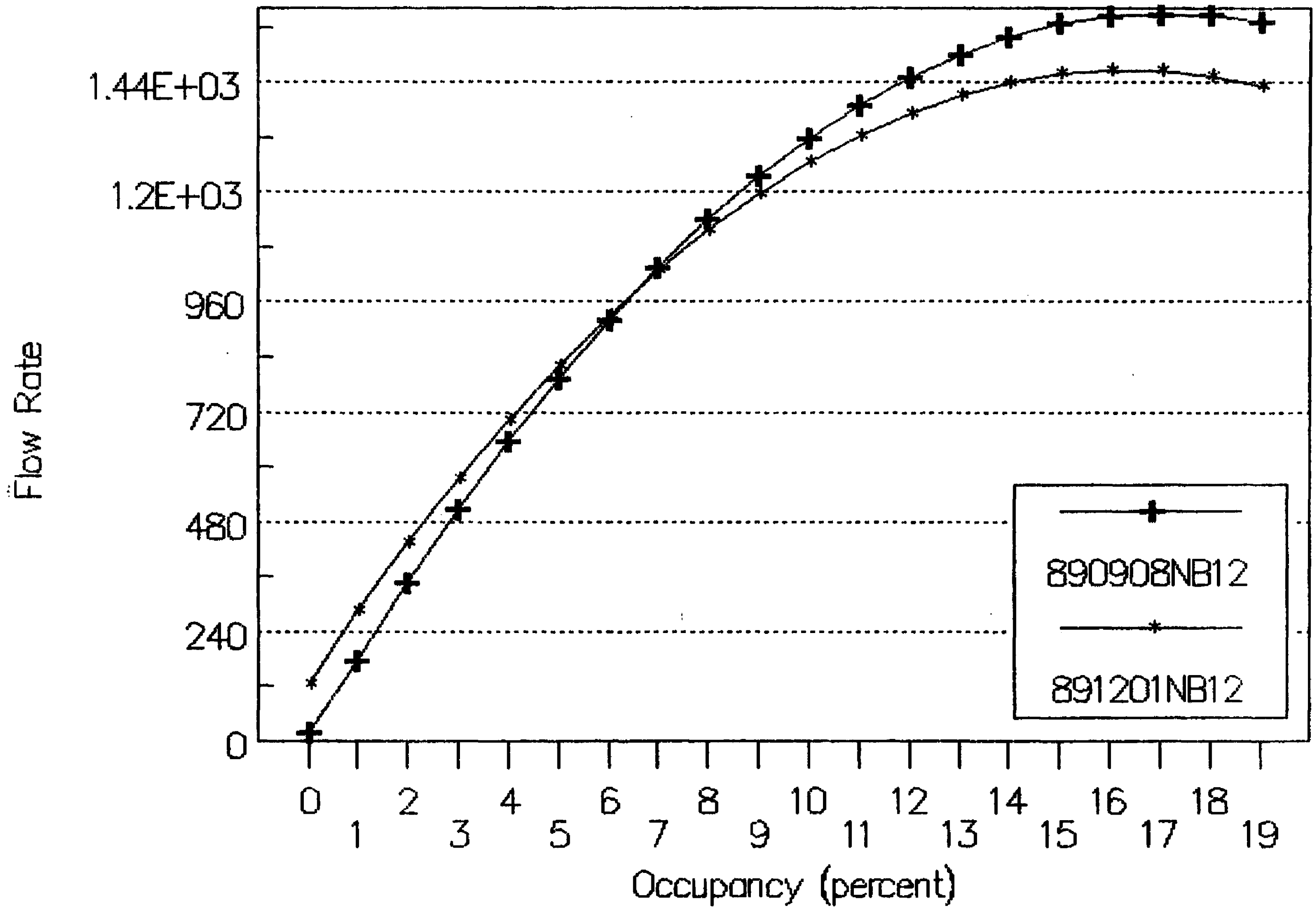

Figure A1. Regression for Skyway NB1-2

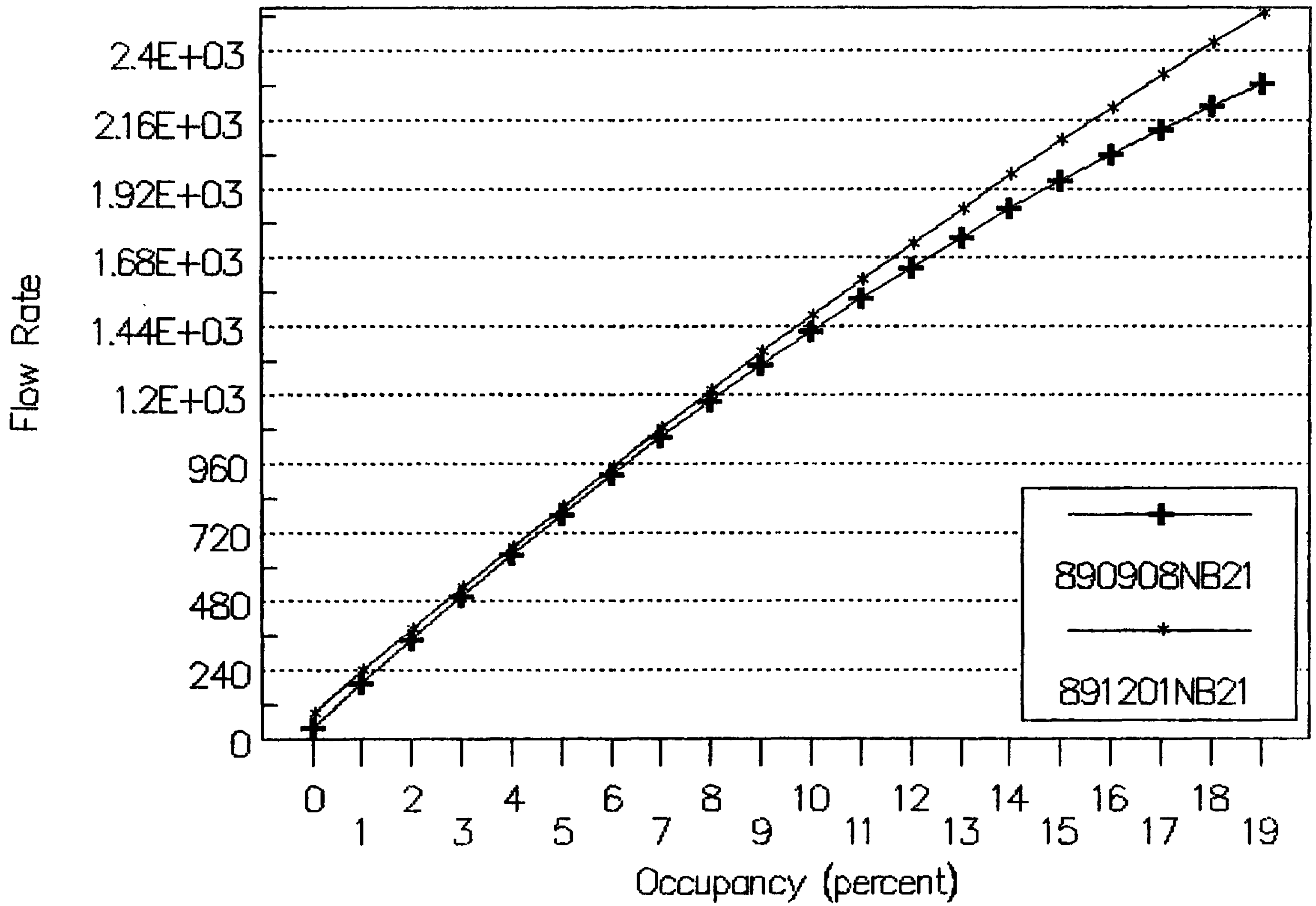

Figure A2. Regression for Skyway NB2-1

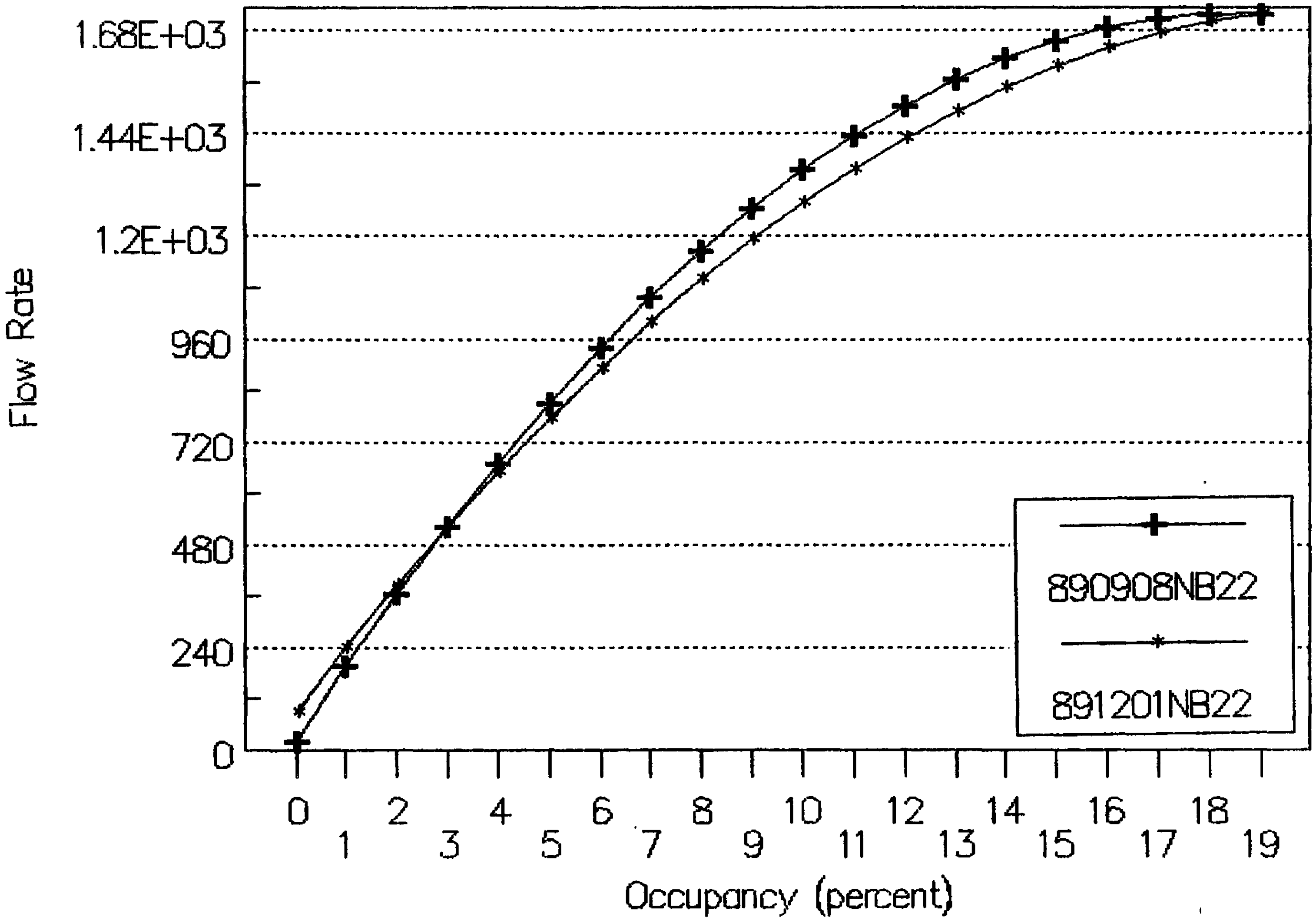

Figure A3. Regression for Skyway NB2-2

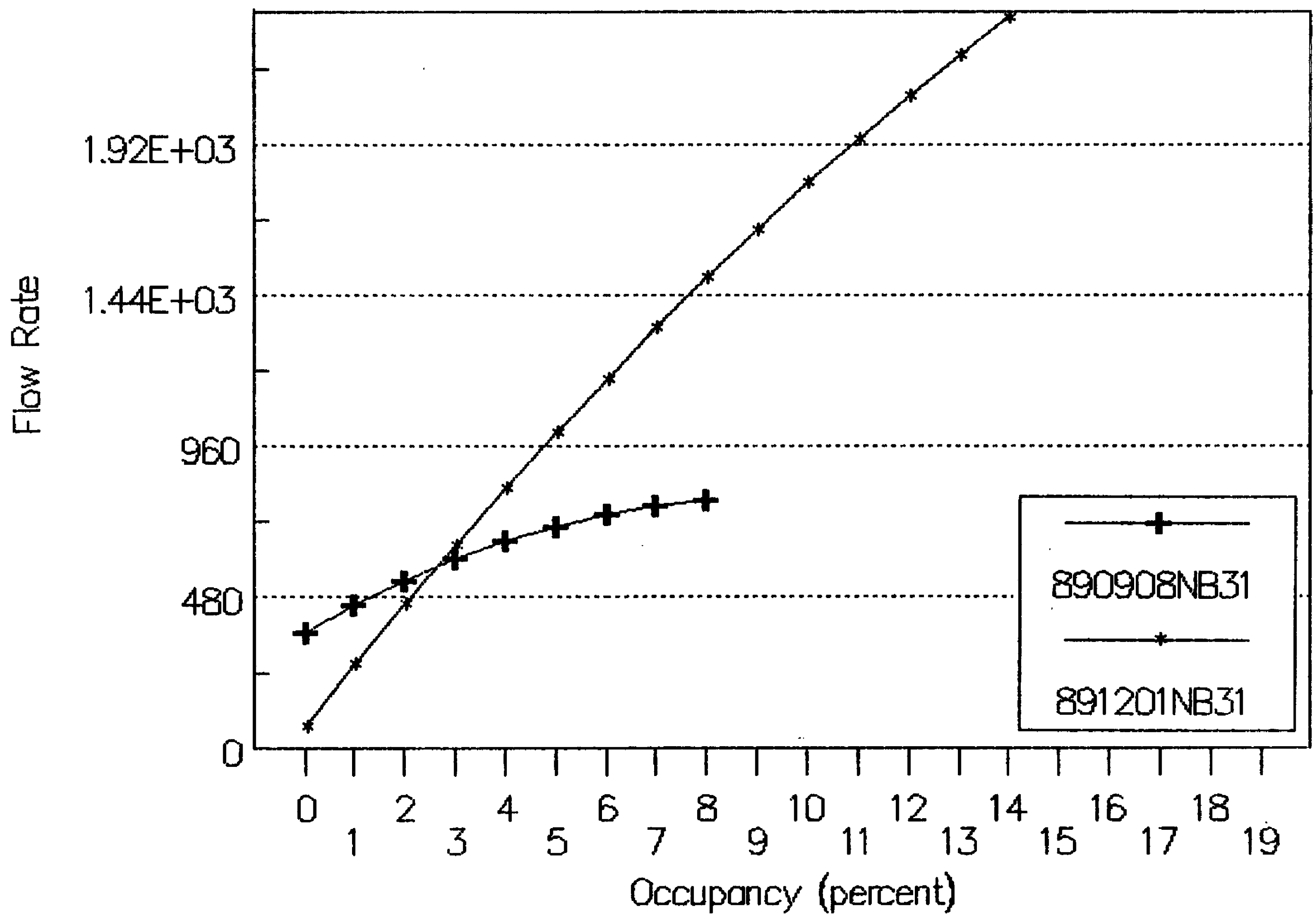

Figure A4. Regression for Skyway NB3-1

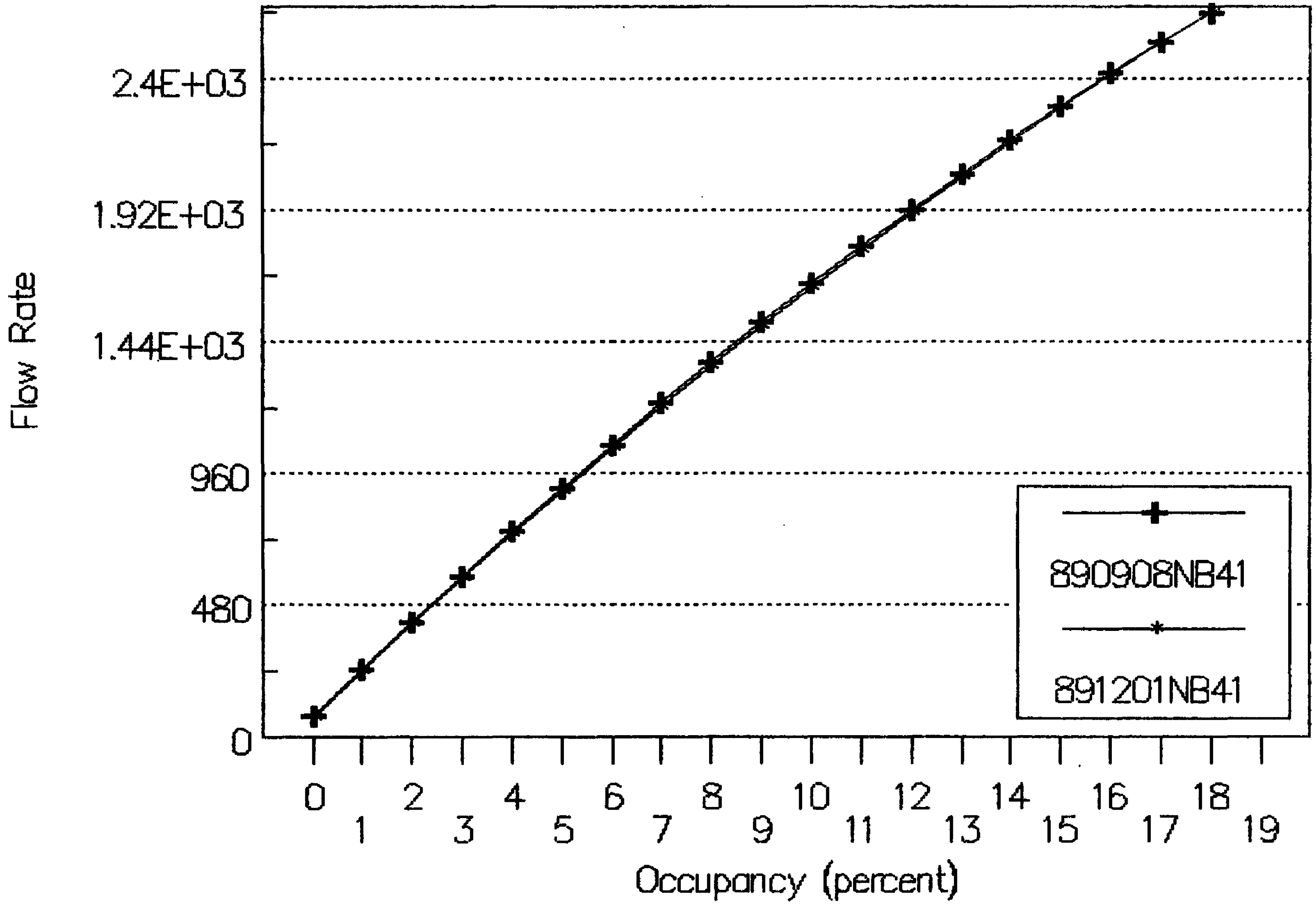

Figure A5. Regression for Skyway NB4-1

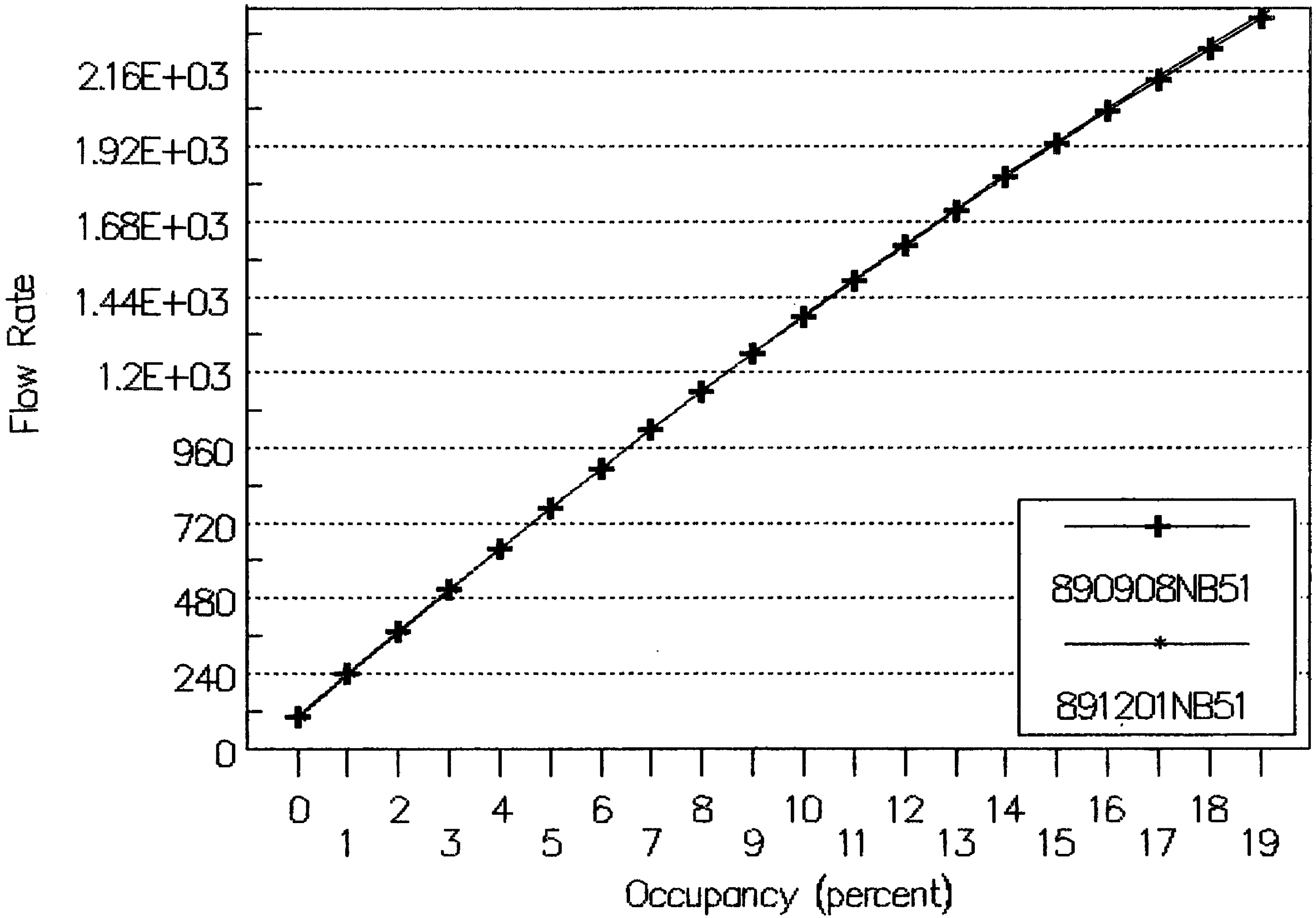

Figure A6. Regression for Skyway NB5-1

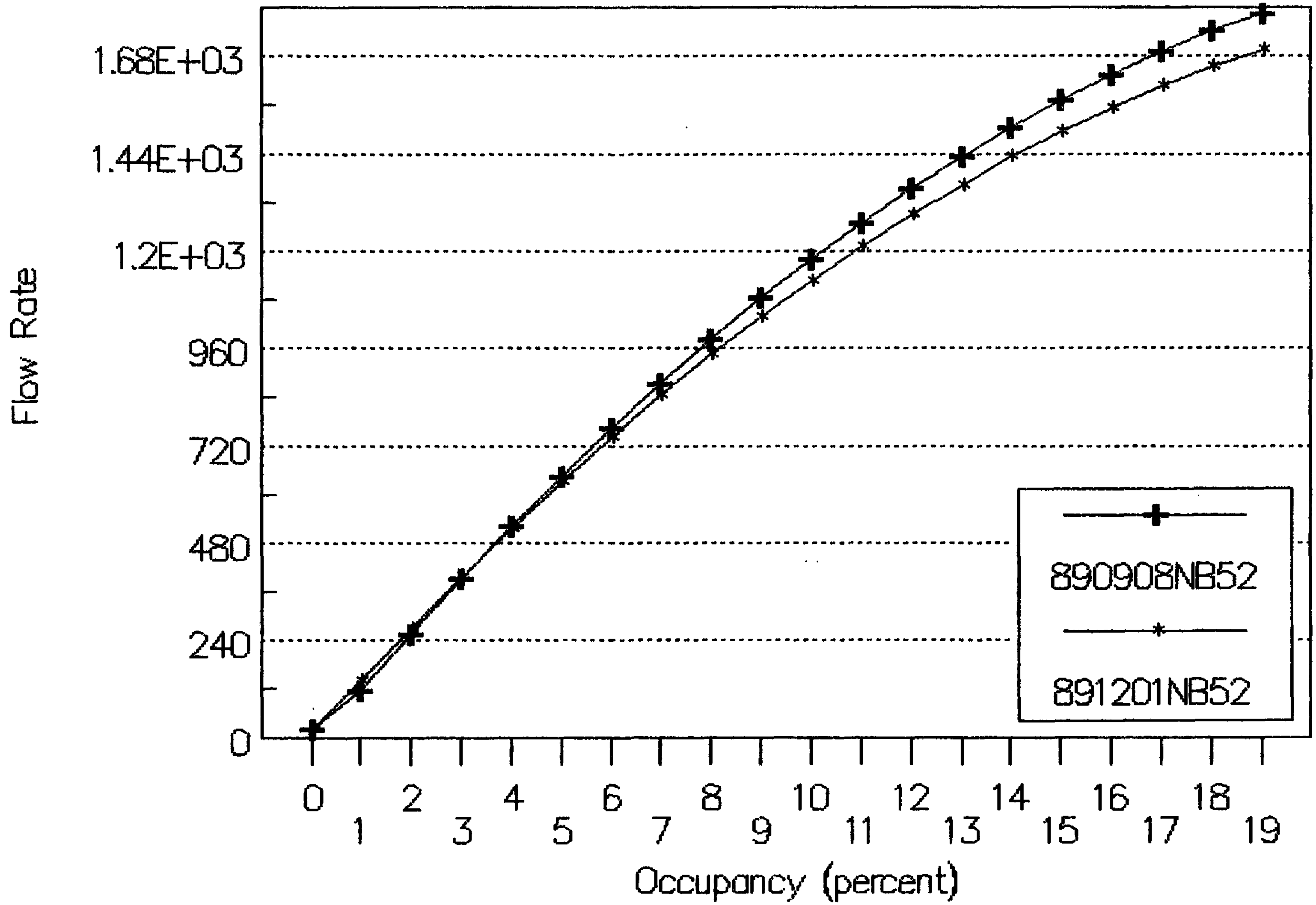

Figure A7. Regression for Skyway NB5-2

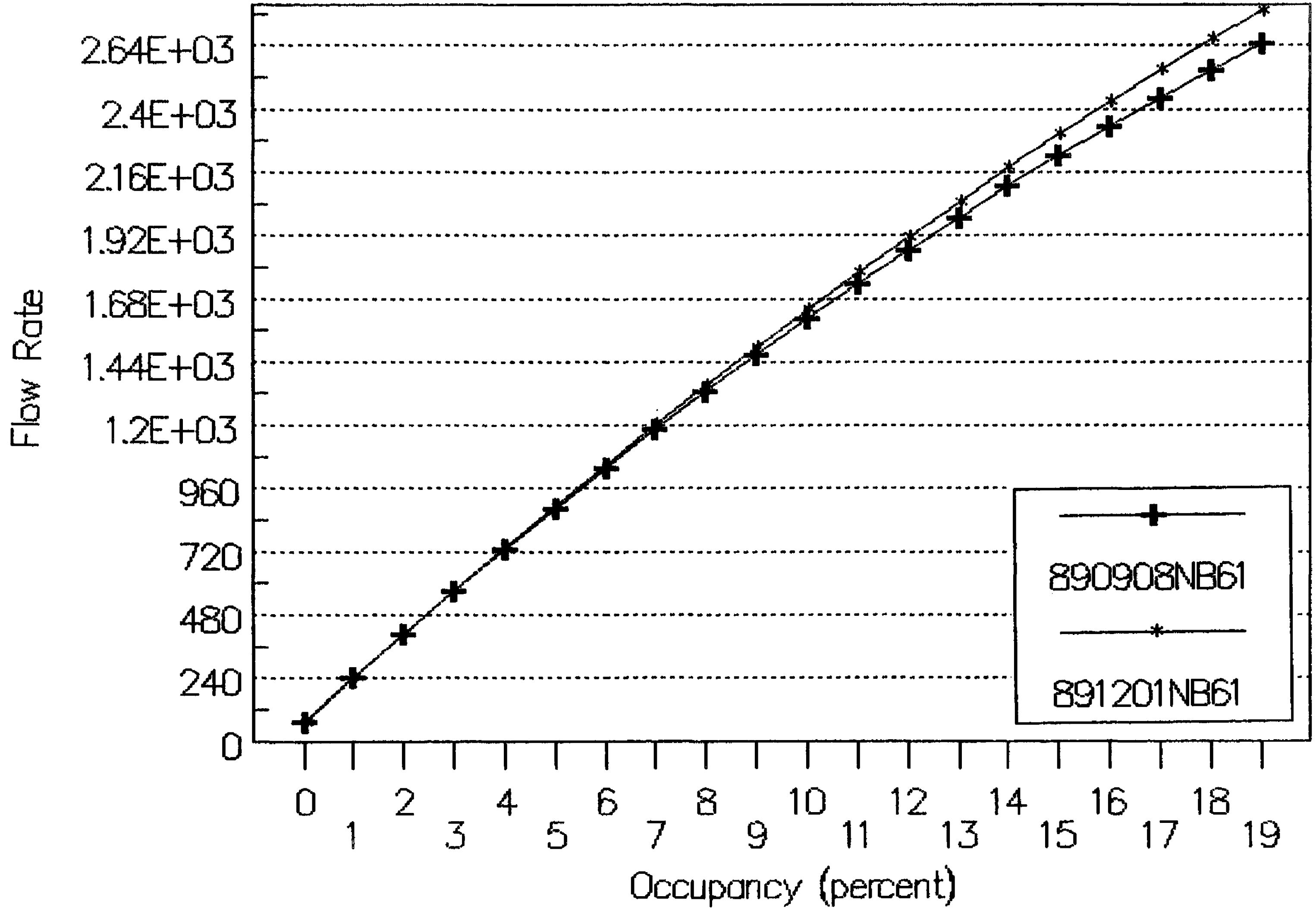

Figure A8. Regression for Skyway NB6-1

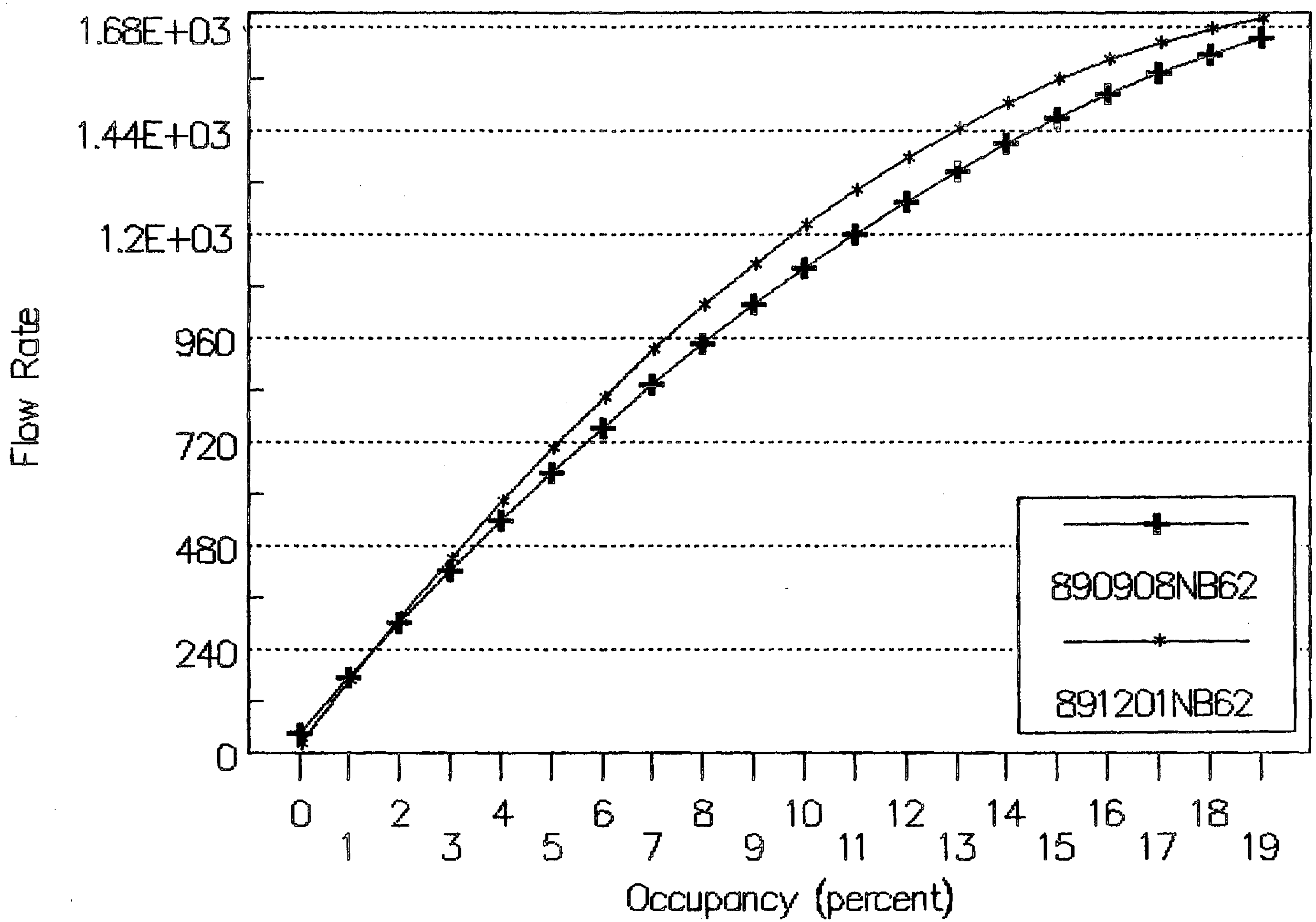

Figure A9. Regression for Skyway NB6-2

REFERENCES

1. Case, E.R. and K.M. Williams. 1978. "Queen Elzabeth Way freeway surveillance and control system demonstration project." **Transportation Research Record** 682, pp. 84-93.

2. Cook, A.R. and D.E. Cleveland. 1974. "Detection of freeway capacity-reducing incidents by traffic stream measurements." **Transportation Research Record** 495, pp. 1-11.

3. Dudek, C.L. and C.J. Messer. 1973. "Detecting stoppage waves for freeway control." **Highway Research Record** 469, pp. 1-15.

4. Dudek, C.L., C.J. Messer, and N.B. Nuckles. 1974. "Incident detection on urban freeways." **Transportation Research Record** 495, pp. 12-24.

5. Gall, Ana. Undated, 1988? "Automatic incident detection." Lecture Material for Civil Engineering 3K3.

6. Gilchrist and F.L. Hall, *Three-Dimensional Relationships Among Traffic Flow Theory Variables.*

7. Hall, L. 1989. "Using the Vax and Traffic Data." Lab Material for Geography 4C6 Students Under Dr. F.L. Hall (Advisor), September, 1989.

8. Hall, L. and F.L. Hall. 1989. "Calibration and Refinement of the McMaster Incident Detection Algorithm." Working Paper 3. Transportation Research Group, Departments of Civil Engineering and Engineering Mechanics and Geography, McMaster University, June 1989.

9. Levin, M. and G.M. Krause. 1979. "Incident detection algorithms, part 2: on-line evaluation." **Transportation Research Record** 722, pp. 58-64.

10. Payne, H.J. and S.C. Tignor. 1978. "Freeway incident-detection algorithms based on decision trees with states." **Transportation Research Record** 682, pp. 30-37.

11. Persaud, B.N. and F.L. Hall. 1989. "Catastrophe theory and patterns in 30-second freeway traffic data--implications for incident detection." **Transportation Research A**, Vol. 23A, No. 2, pp. 103-113.

12. Persaud, B.N., F.L. Hall and L.M. Hall. 1989. "Development and Testing of the McMaster incident detection algorithm." Submitted to the Transportation Research Board, August 1989.