Estimation of Turning Flows from Automatic Counts

E. HAUER, E. PAGITSAS, AND B.T. SHIN

Traffic flows on intersection approaches can be obtained by using automatic counting machines. A method for the estimation of turning movements from approach counts is developed and tested. The problem is solved by identifying the most likely traffic-flow matrix that agrees with the given approach counts by using characteristic turning proportions. Traffic flows from 145 intersections in metropolitan Toronto have been coded. From this information, characteristic right- and left-turning proportions are estimated for five types of intersection approaches. Estimates of vehicle flows are derived by using these characteristic turning proportions. These estimates are then compared with the observed flows, and the accuracy of the estimation is explored. It appears that the method may in some cases serve as a useful tool.

Knowledge of traffic flows at intersections is needed for a variety of planning and management purposes. Obtaining an estimate of flows into and out of intersection approaches is relatively cheap. It can be done by automatic counting machines. However, for many purposes, the split of each traffic stream into left-turning, right-turning, and straight-through flows is of interest. The most common way of obtaining estimates of turning movements is from manual counts by observers, and such counts tend to be expensive.

Recently, alternative methods for the estimation of turning movements at intersections have been suggested (1-4). These methods use approach counts and other relevant information to obtain estimates of turning flows.

In this paper, we follow closely the method proposed by van Zuylen (3). The theoretical part in the following section contains no innovation and is included for completeness. The only difference is in the motivation of the method. Whereas van Zuylen relies on minimization of "information", our argument is based on the maximization of "likelihood".

The main purpose of the study is to explore the accuracy of the estimation obtained when the method is applied to a large number of urban intersections.

PROBLEM FORMULATION AND SOLUTION

Let i take on values 1,2,...,m and represent the label assigned to an intersection approach carrying traffic into the intersection. Similarly, let j take on values 1,2,...,n and represent the label assigned to an approach carrying traffic out of the intersection. With i and j as subscripts, let T_{ij} be the flow from approach i into approach j. When the T_{ij} are written in a table with m rows and n columns, a "flow matrix" is obtained. The task is to find estimates T_{ij}^* of T_{ij}--that is, to find an estimate of the flow matrix.

Two pieces of information are brought to bear on the estimation task:

1. It is assumed that the total flow into the intersection from approach i (O_i) and the total flow out of the intersection by approach j (D_j) are known for all i and j. This information is obtainable from, say, automatic traffic counters. By using this notation, $\sum_{j=1}^{n} T_{ij} = O_i$ and $\sum_{i=1}^{m} T_{ij} = D_j$. Barring errors of counting and negligible "end of count period" effects, the sum of entering flows (S) equals the sum of leaving flows ($\sum_{i=1}^{m} O_i = \sum_{j=1}^{n} D_j = S$).

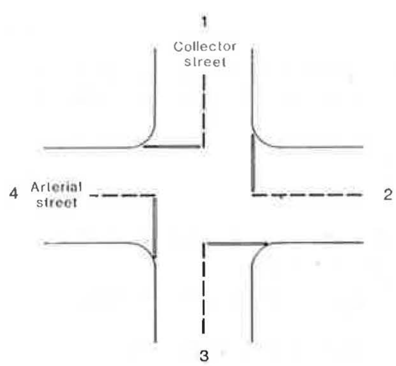
2. There exists some prior knowledge about the proportion of left and right turns at similar intersection approaches. Then let p_{ij} denote the proportion of the traffic emanating from approach i that at similar intersections turns in the direction of approach j. Naturally, $\sum_{j=1}^{n} p_{ij} = 1$.

To illustrate, consider the intersection shown in Figure 1 and the corresponding flow matrix. The cells on the diagonal are shaded to indicate that vehicles entering the intersection do not turn back by the same approach, and therefore these cells will have no entries. Any turn restrictions can be represented similarly. For example, if left turns were not permitted from the arterial street into the collector, the two cells (row 2, column 3 and row 4, column 1) would also be shaded. Assuming that the proportion of leftand right-turning traffic from an arterial street into a collector street is normally 0.02 each, and that the corresponding value for flow from the collector into the arterial is 0.30, the p_{ij} values can be entered into the matrix as shown. Finally, imagine that two automatic traffic counters were placed on each leg of the intersection, one counting entering traffic and the other counting traffic leaving the intersection. These eight counts are entered as row and column sums in the flow matrix. For example, the flow entering the intersection from approach 1 has been counted as 100 vehicles for the count period, whereas the flow leaving the intersection by approach 1 during the same period has been counted as 50 vehicles.

The stage is now set for analyzing the problem of finding an estimate of the flow matrix T.

There are many different combinations of numbers that, when listed in the flow matrix, will meet the

Figure 1. Intersection, flow matrix, and p_{ij} values.



From \ To	1	2	3	4	
1	(0)	(0.30)	(0.40)	(0.30)	100
2	(0.02)	(0)	(0.02)	(0.96)	600
3	(0.40)	(0.30)	(0)	(0.30)	200
4	(0.02)	(0.96)	(0.02)	(0)	700
	50	800	100	650	1600

Figure 2. Computations for example intersection: use of the algorithm.

From \ To	p_{ij} 1	2	3	4	O	A^0	Iteration 1 A^1	2 A^2	3 A^3	4 A^4	5 A^5
1	0.00	0.30	0.40	0.30	100	2.5	2.06	1.94	1.91	1.90	1.90
2	0.02	0.00	0.02	0.96	600	15.0	15.37	15.86	16.29	16.65	16.95
3	0.40	0.30	0.00	0.30	200	5.0	6.28	6.54	6.59	6.60	6.62
4	0.02	0.96	0.02	0.00	700	17.5	16.70	16.24	15.89	15.71	15.46

		1	2	3	4	
	D	50	800	100	650	1600
Iteration 1	B^1	18.9	42.0	60.6	39.0	
Iteration 2	B^2	15.86	43.16	68.24	37.66	
Iteration 3	B^3	15.35	44.11	70.52	36.58	
Iteration 4	B^4	15.24	44.93	71.04	35.74	
Iteration 5	B^5	15.21	45.37	71.06	35.07	

given row and column sums. The question is which of these many "feasible solutions" is most likely to have occurred during the period in which the automatic counts were taken.

Each of these feasible solutions could have arisen in a large number of ways. The number of ways in which flows $T_{i1}, T_{i2}, \ldots, T_{in}$ can be selected from a total of O_i is known to be

$$O_i!/\prod_{1}^{n} T_{ij}! \tag{1}$$

To keep the notation simple, all sum (Σ) and product (π) operators will be understood to skip over the "shaded" cells of the flow matrix.

In view of the specified proportions p_{ij}, each of the events included in the total given by Equation 1 arises with a probability of

$$\prod_{1}^{n} (p_{ij})^{T_{ij}} \tag{2}$$

It follows that the relative frequency w(T) with which some specific flow matrix T should be expected to arise is given by

$$w(T) = \prod_{i=1}^{m}\left[O_i! \prod_{1}^{n} (p_{ij})^{T_{ij}} / \prod_{1}^{n} T_{ij}!\right] \tag{3}$$

The task is to identify the flow matrix T* for which the value of w(T) is the largest.

Since the product $\prod_{1}^{n} O_i!$ is fixed, the task can be transformed into finding T that maximizes

$$\ln \prod_{1}^{m} \prod_{1}^{n} (p_{ij}^{T_{ij}}/T_{ij}!) = \sum_{1}^{m} \sum_{1}^{n} (T_{ij} \ln p_{ij} - \ln T_{ij}!) \tag{4}$$

According to Stirling's formula,

$$\ln T_{ij}! = T_{ij} \ln T_{ij} - T_{ij} + \tfrac{1}{2}(\ln T_{ij} + \ln 2\pi) + [r(T)_{ij}/12T_{ij}] \tag{5}$$

Since $T_{ij} >> 1$ and $r(T_{ij}) < 1$, all but the first two terms can be neglected. In preparation for the forthcoming minimization, note that, by using this approximation,

$$\ln T_{ij}! \cong T_{ij} \ln T_{ij} - T_{ij} \tag{6}$$

and

$$d \ln T_{ij}!/d T_{ij} \cong \ln T_{ij} \tag{7}$$

Making use of the approximation in Equation 6, we wish to find T that maximizes

$$\sum_{1}^{m} \sum_{1}^{n} T_{ij} \ln p_{ij} - T_{ij} \ln T_{ij} + T_{ij} \tag{8}$$

subject to

$$\sum_{j=1}^{n} T_{ij} = O_i \quad \text{for } i = 1, \ldots, m \tag{9}$$

and

$$\sum_{i=1}^{m} T_{ij} = D_j \quad \text{for } j = 1, \ldots, n \tag{10}$$

Let α_i be the Lagrange multiplier for the m-row constraints and β_i the Lagrange multiplier for the n-column constraints. Taking derivatives of the Lagrange expression, we find that

$$T_{ij}^* = p_{ij} e^{-\alpha_i} e^{-\beta_i} = p_{ij} A_i B_j \tag{11}$$

Equation 11 constitutes the solution of the problem as formulated. It remains to determine the value of the unknown multipliers A_i and B_j so as to satisfy the row and column sums as specified in Equation 10. The algorithm for doing so is described below. The underlying rationale for this estimation procedure and its weaknesses are discussed more fully by Hauer and Shin (5).

COMPUTATIONS

The problem formulated above and its solution belong to the general class of "biproportional" models. A brief description of the biproportional problem and its history, applications, and properties is given by Murchland (6).

Possibly the simplest solution algorithm consists of repeated balancing of the vectors A and B and is named after Kruithof (7). It consists of the following steps:

1. Set A_i (current) $= O_i/S^{1/2}$.
2. Find B_j from $B_j = D_j/\sum_{i=1}^{m} p_{ij}A_i$ (current).
3. Find A_i (new) from A_i (new) $= O_i/\sum_{j=1}^{m} p_{ij}B_j$.
4. Compare A_i (new) and A_i (current). If the largest difference is sufficiently small, use last A_i and B_j to find $T_{ij}* = p_{ij}A_iB_j$. Otherwise, set A_i (current) = A_i (new) and return to step 2.

To illustrate, consider the intersection described in Figure 1. The proportions (p_{ij}), inflows (O_i), and outflows (D_j) are reproduced in Figure 2. Step 1 of the algorithm produces the column under A^0. Thus, for example, $A_1 = O_1/S^{1/2} = 100/1600^{1/2} = 2.5$. In step 2, by using A^0, the vector B^1 is calculated. For instance, $B_4{}^1 = 650/(2.5 \times 0.30 + 15.0 \times 0.96 + 5.0 \times 0.30 + 17.5 \times 0.00) = 39.0$. Going to step 3, vector A^1 is found by using B^1. For instance, $A_1{}^1 = 100/(18.9 \times 0.00 + 42.0 \times 0.30 + 60.6 \times 0.40 + 39.0 \times 0.30) = 2.06$. The first round of computations ends by comparing the new vector A with the previous one. Unless the desired closure is attained, a new round of computations is carried out. In this example, results of five iterations are listed. A gradual convergence of the multiplier vectors A and B is evident. Were one to calculate estimates T* on the basis of A^3, B^3, then Figure 3a would apply. Estimates of T* using A^5, B^5 are shown in Figure 3b.

PROPORTION OF TURNING MOVEMENTS AT INTERSECTIONS IN METROPOLITAN TORONTO

The procedure described and illustrated above produces estimates of vehicle flows by using data from automatic counters and prior knowledge about typical proportions of turning movements at intersections. The better the estimates of these typical proportions (p_{ij}), the more accurate the resulting estimate of the flow matrix will be. If no information is available about the intersection at hand or about the area in which it is situated, one may be inclined to rely on some gross average proportions. Thus, for example, the "average conditions" for signalized intersections on which the charts in the Highway Capacity Manual (8, p. 133) are based correspond to 10 percent left turns and 10 percent right turns. In an attempt to improve the quality of turning-flow estimation, the turning proportions at 145 intersections in metropolitan Toronto have been examined.

For each intersection, complete traffic-flow counts were available for four periods:

Period	Hours
Morning peak	7:00-9:00 a.m.
Evening peak	4:00-6:00 p.m.
Off-peak	11:00 a.m.-12:00 noon
Daytime	7:00 a.m.-6:00 p.m.

Figure 4 shows a typical traffic-flow diagram that served as a source of data. The information has been coded and keypunched as shown in Figure 5.

For coding, five types of intersections have been defined:

1. Central business district (CBD) (within rectangle bounded by Spadina, Dundas, Jarvis, and Front Streets),
2. Arterial with arterial,
3. North-south arterial with east-west collector,
4. East-west arterial with north-south collector, and
5. Collector with collector.

In this study, an arterial is a main thoroughfare with two or more lanes in one direction. In metro-

Figure 3. Computations for example intersection: (a) T*³ and (b) T*⁵.

a

From \ To	1	2	3	4	O
1		26	54	20	100
2	5		24	571	600
3	40	90		70	200
4	5	673	22		700
D	50	789	100	661	

b

From \ To	1	2	3	4	O
1		25	54	21	100
2			23	572	600
3	40	87		72	199
4	5	673	22		700
D	50	785	99	665	

Figure 4. Typical traffic-flow diagram.

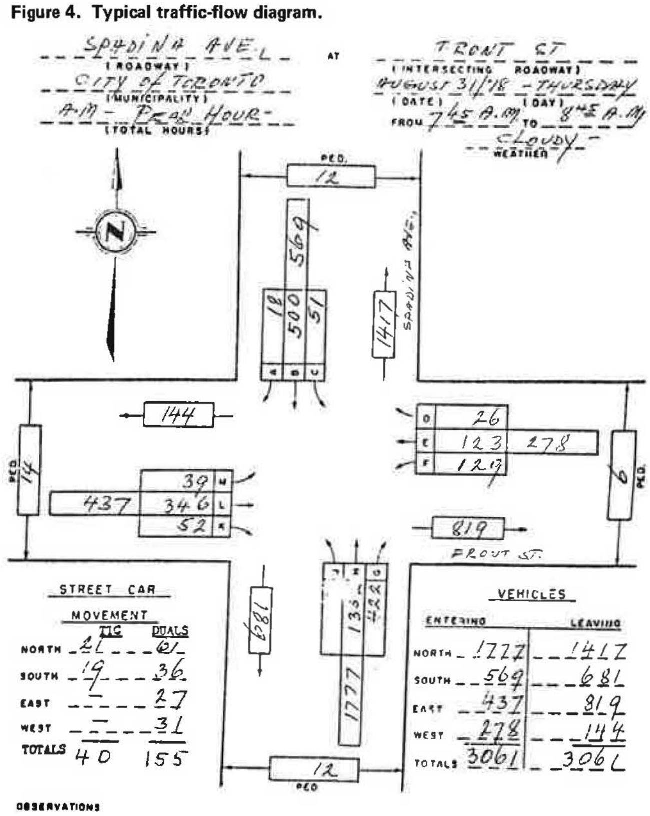
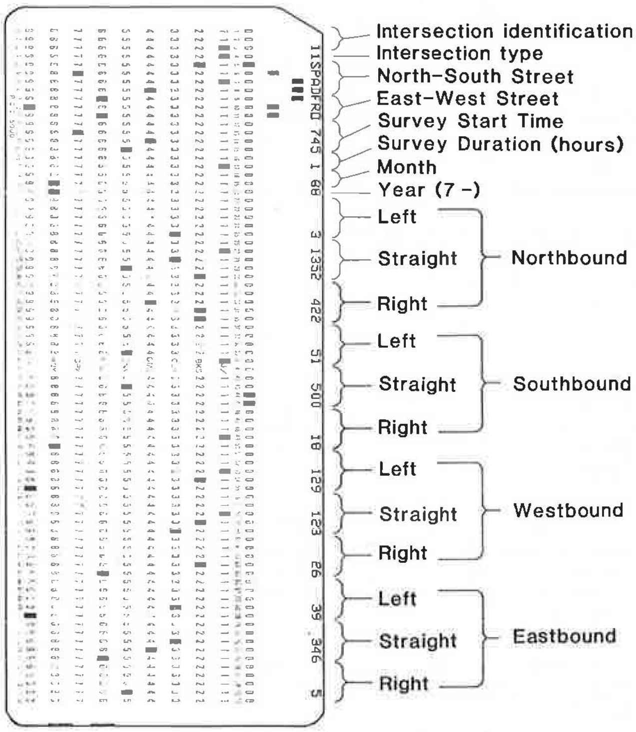


Figure 5. Data card for traffic-flow diagram in Figure 4.



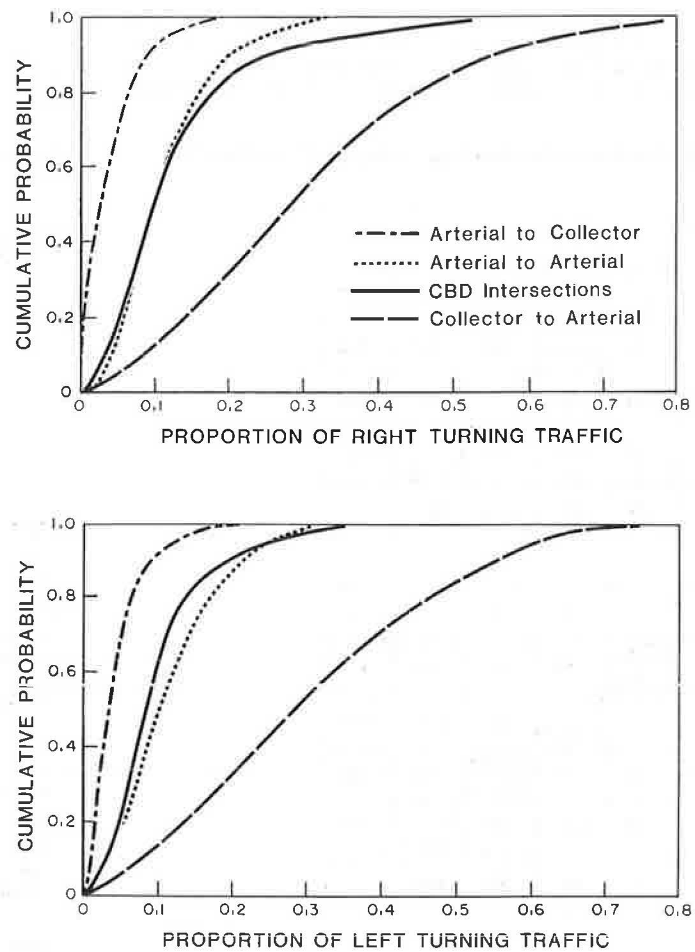
politan Toronto, arterials are approximately 1.25 miles apart. All lesser streets are included in the category "collector".

Many factors can serve to explain the differences between the turning proportions; these factors include the type of intersection or approach, the time of day, the direction of movement, the location in the urban area, nearby turn restrictions, and local land use. After detailed exploration of the various factors, it has been found that a large part of the difference can be attributed to the function of the road from which the vehicles enter the intersection and the function of the road by which they leave the intersection. The location in the urban area is another important factor. Accordingly, the estimates of average turning proportions given in Table 1 have been obtained. The empirical probability distribution of p_{ij} for approach types 1-4 is shown in Figure 6.

Table 1. Average turning proportions.

		Proportion		
No.	Type of Approach	Turning Left	Turning Right	No. of Approaches in Sample
1	CBD	0.10	0.12	92
2	Arterial to arterial	0.12	0.12	83
3	Arterial to collector	0.04	0.05	52
4	Collector to arterial	0.30	0.32	53
5	Collector to collector	0.10	0.20	3

Figure 6. Probability distribution of p_{ij}: left- and right-turning traffic.



COMPARISON OF OBSERVED FLOWS AND THEIR ESTIMATES

By using the average turning proportions in Table 1 and the observed approach flows, estimates of the entire flow matrix were obtained for all 145 intersections. This allows a comparison of observed flows and the estimates that could be obtained if turning flows were not counted.

To illustrate, the correspondence of flows and estimates for the intersection shown in Figure 4 is as follows:

Observed Flows	Estimated Flows	Observed Flows	Estimated Flows
18	12	3	0
500	506	1352	1353
51	50	422	420
26	43	39	19
123	128	346	348
129	105	52	68

The plotting of observed versus estimated flows for each movement of each type of approach defined in Table 1 for all 145 intersections results in Figures 7-11.

ACCURACY OF ESTIMATION

One of the main purposes of this study has been to explore the accuracy of estimates obtainable by this method. A first impression of such accuracy can be obtained by scanning Figures 7-11. Several observations can be made:

1. Points seem to be distributed approximately symmetrically around the bisector (except, of course, near the origin).
2. The band of points surrounding the bisector appears to be of constant width. This is attributable to the minimization process that provides the rationale to the estimation algorithm.
3. The accuracy of estimation varies by approach type. One reason for this variation is the distribution of p_{ij} (Figure 6). The wider the distribution of p_{ij}, the greater is the chance that the average value used in a specific case differs sub-

Figure 7. Correspondence of observed and estimated flows: CBD intersections.

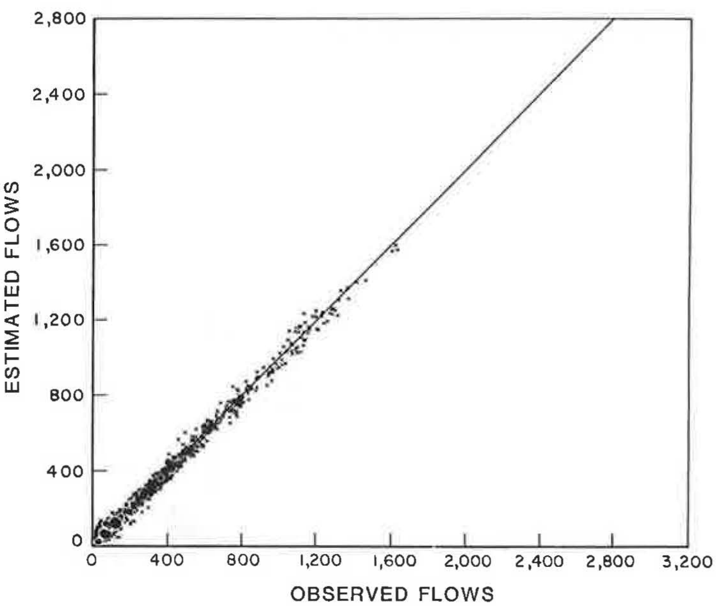


Figure 8. Correspondence of observed and estimated flows: arterial to arterial.

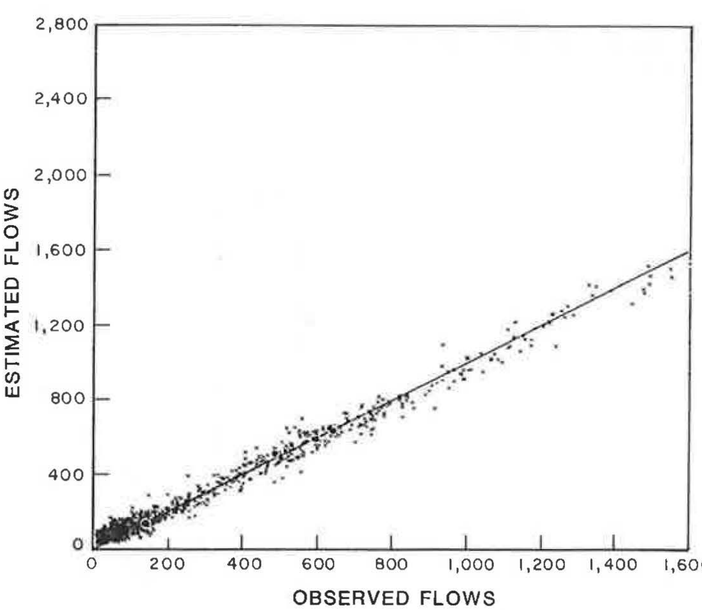


Figure 9. Correspondence of observed and estimated flows: arterial to collector.

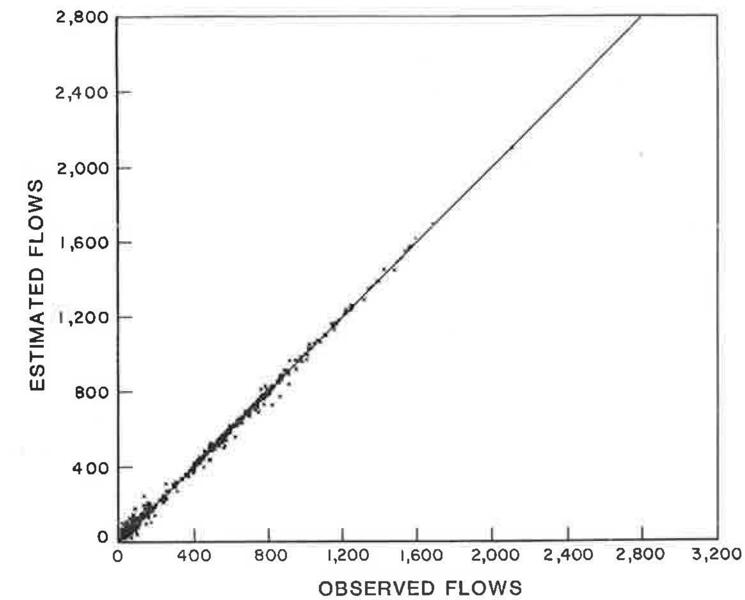


Figure 10. Correspondence of observed and estimated flows: collector to arterial.

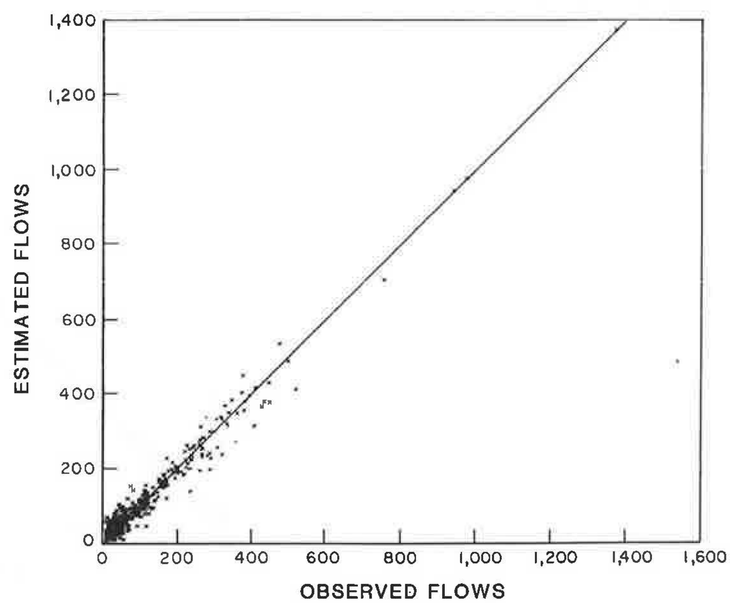


Figure 11. Correspondence of observed and estimated flows: collector to collector.

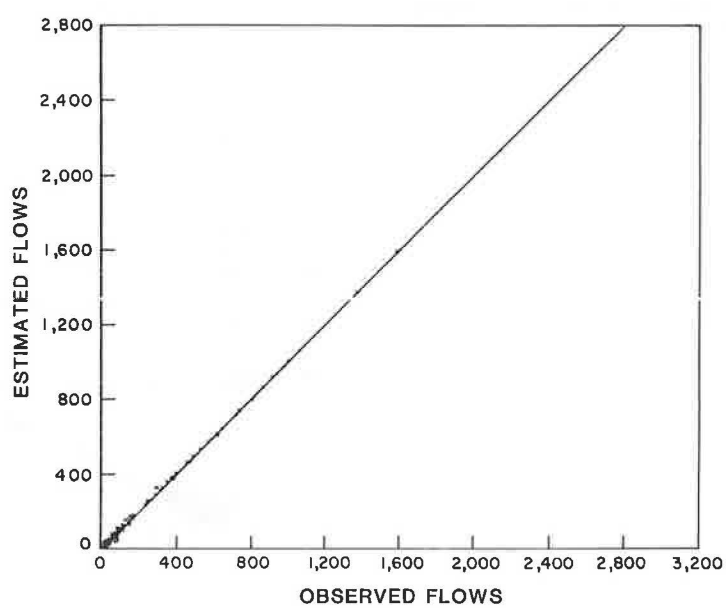
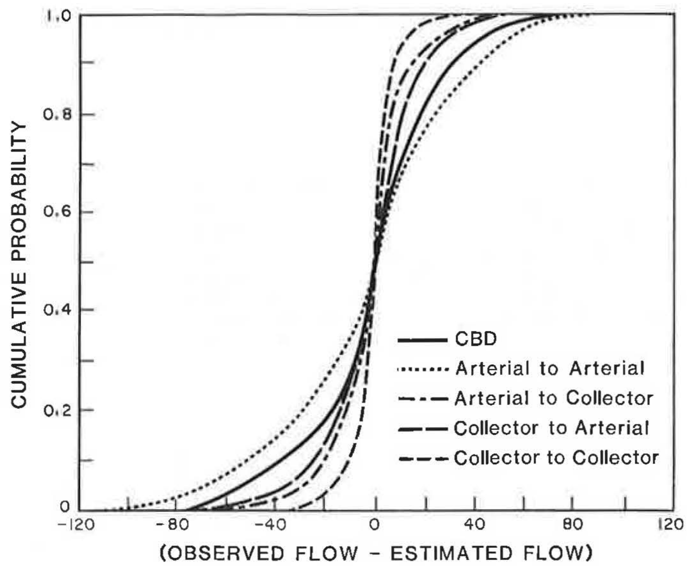


Figure 12. Distribution of difference between observed and estimated flows.



stantially from the value that actually prevails at that intersection.

4. The attainable accuracy is slightly exaggerated by the fact that the average turning proportions used (Table 1) have been derived from the same data that served for the plotting of Figures 7-11. However, limited sensitivity tests indicate that the accuracy of estimation is not affected significantly when p_{ij} are changed by a few percentage points.

The standard deviations of the difference between observed flows and their estimates are given below:

Type of Approach	Standard Deviation (no. of vehicles)
CBD	28.09
Arterial to arterial	40.83
Arterial to collector	17.67
Collector to arterial	20.07
Collector to collector	9.74

The probability distribution of this difference by approach type is shown in Figure 12. From Figure 12 one can read the probability of the difference between the observed and estimated flow to be in a certain range. To illustrate, for a CBD intersection the probability of an error less than ±60 vehicles is approximately 0.97 - 0.03 = 0.94.

SUMMARY

This paper presents a method for the estimation of vehicle turning movements from intersection approach flows. The estimation method identifies the most likely set of flows that agrees with the observed approach counts, taking also into account typical proportions of leftand right-turning flows. In principle, the method suggested by van Zuylen (3) is followed.

Flow estimates are obtained by iterative computations. The algorithm used in these computations is incorporated into a FORTRAN computer program.

To obtain realistic estimates of turning flows, prior information about characteristic leftand right-turning proportions is required. This has been obtained through the analysis of flows at 145 intersections in metropolitan Toronto. Results are summarized in the form of average turning proportions for five types of approaches.

By using the turning proportions so obtained and the observed approach flows, estimates of all flows on all 145 intersections for three periods of the day have been calculated. There appears to be a surprisingly close correspondence between the actual and the estimated flows. An empirical probability distribution curve for the difference between the actual and the observed flows is given for each of the five approach types. These can be used to anticipate the accuracy of estimation in similar circumstances.

When the obtainable accuracy is sufficient for the purpose at hand, the method described in this paper may be an attractive alternative to the conduct of a field survey by observers.

REFERENCES

1. M. Jeffreys and M. Norman. On Finding Realistic Turning Flows at Road Junctions. Traffic Engineering Control, Vol. 18, No. 1, Jan. 1977, pp. 19-21,25; Addendum, April 1977, p. 207.
2. A. Mekky. On Estimating Turning Flows at a Road Junction. Traffic Engineering Control, Vol. 20, No. 10, Oct. 1979, pp. 486-487.
3. H.J. van Zuylen. The Estimation of Turning Flows on a Junction. Traffic Engineering Control, Vol. 20, No. 11, Nov. 1979, pp. 539-541.
4. M. Norman and N. Hoffmann. Non-Iterative Methods for Generating a Realistic Turning Flow Matrix for a Junction. Traffic Engineering Control, Vol. 20, No. 12, Dec. 1979, pp. 587-589.
5. E. Hauer and B.T. Shin. Estimation of Origin-Destination Matrices from Observed Flows: Simple Systems. Toronto-York Joint Program in Transportation, Univ. of Toronto and York Univ., Toronto, Res. Rept. 65, Feb. 1980.
6. J.D. Murchland. Applications, History, and Properties of Biand Multi-Proportional Models. Transport Studies Group, University College, London, England, 1978.
7. J. Kruithof. Calculation of Telephone Traffic. De Ingenieur, Vol. 52, No. 8, 1937.
8. Highway Capacity Manual. HRB, Special Rept. 87, 1965.