ISSUES RELATED TO THE APPLICATION OF ACCIDENT PREDICTION MODELS FOR THE 
COMPUTATION OF ACCIDENT RISK ON TRANSPORTATION NETWORKS

By
Dominique Lord

March 25th 2001

Center for Transportation Safety*
Texas Transportation Institute
Texas A&M University System
3135 TAMU
College Station, TX
77843-3135

Tel. (979) 458-1218
Fax. (979) 845-4872
e-mail: d-lord@tamu.edu

Paper presented to the 81st Annual Meeting of the Transportation Research Board

*This paper was written while the author was a post-doctoral researcher at Ryerson University, Ontario.
ABSTRACT

Accident risk has been applied extensively in transportation safety analysis. Risk is often used to describe the level of safety in transportation systems by incorporating a measure of exposure, such as traffic flow or kilometers driven. The most commonly applied definition of accident risk states that risk is a linear function of accidents and traffic flow. This definition, however, creates problems for transportation systems that are characterized by a non-linear relationship between these variables. This original research on which this paper is based aimed, as the primary objective, to illustrate the application of accident prediction models (APMs) to estimate accident risk on transportation networks (APMs are useful tools for establishing the proper relationship between accidents and traffic flow). The secondary objective was to describe important issues and limitations surrounding the application of APMs for this purpose. To accomplish these objectives, APMs were applied to a computerized transportation network with the help of EMME/2. The accident risk was computed with the traffic flow output of the computer program. The results were dramatic and unexpected: In essence, the individual risk of being involved in a collision decreases as traffic flow increases. The current and most common model form of APMs explains this outcome. The author anticipates that the application of these results may have significant effects on transportation policy and ITS strategies.
INTRODUCTION

Accident risk has been extensively applied in transportation safety analysis. Risk is often used to describe the level of safety of transportation systems by incorporating a measure of exposure. Accident risk has been applied in studies aiming at comparing different highways, at evaluating the effects of treated highway facilities, at estimating the probability of an accident for a driver and at finding the safest routes for the transportation of hazardous materials.

The most traditional definition of accident risk states that the relationship between the number of collisions (e.g., by type, severity, etc.) and the amount of exposure on a transportation facility (e.g., kilometers driven, traffic flows, etc.) is linear with a zero intercept (1, 2, 3, 4, 5). The assumption herein is that the individual probability of being involved in a traffic accident increases linearly as the exposure increases. This definition, however, creates problems for transportation systems that are characterized by a non-linear relationship between accidents and exposure. The relationship between accidents and traffic flow, the most common measure of exposure, has been shown to actually follow a non-linear relationship, in which, accident counts usually increase at a decreasing rate as traffic flow increases (6, 7). Consequently, the methods used in the literature to estimate accident risk to individual vehicles may be not be appropriate for transportation network safety analyses.

The primary objective of this paper is to illustrate the application of accident prediction models (APMs) to estimate accident risk on transportation networks. The main focus of this research is on finding the safest routes on a macroscopic representation of a transportation network. Although the risk of being involved in a collision is a function of various characteristics of the driver, the vehicle and the environment, the definition of risk as applied to this paper refers only to a function of traffic accidents and traffic flow. Our aim, therefore, is to determine the probability of a driver being involved in an accident when, on average, X number of accidents occurs on a segment of a transportation network and Y number of vehicles travels through this segment. In other words, the individual risk of being involved in a collision is based on the (prior) safety level of various components of the transportation network.

The paper is divided into four sections. The first presents a brief review of previous work on accident risk. The second describes how the models were applied to estimate accident risk. The third shows how the safest routes were computed. The last section presents the issues surrounding the application of APMs to estimate individual accident risk on transportation networks.

REVIEW OF PREVIOUS WORK

The risk of collision for drivers is usually a function of the current safety level of a facility, either estimated by recorded accident counts or in some instances traffic conflict counts, and a measure of exposure such as traffic flow or kilometers driven. According to Cameron and Milne (8), accident risk can be defined as “a measure of the probability of a potential accident event resulting in an accident.” The relationship between risk, exposure and safety can basically be given by the following equation:

\[
\text{Risk} = \frac{\text{Safety}}{\text{Exposure}}
\]

This relationship has been used extensively in various studies, such as comparing various geographic areas, drivers or highways, or the prediction of accidents through traffic conflict counts. Examples of such applications can be found in (4), (9), (10), (11) and (12). The merit of the validity of such comparisons is however left for others to debate and will not be addressed here. The reader is referred to (1), (3), (13), (14), (15) and (16) for additional information on the relationship between safety, exposure and risk.
An important use of accident risk for the individual vehicle has been for estimating the safest route on a network where hazardous materials are being transported. Since hazardous materials can cause serious harm both to the population and to the environment if an accident occurs, private carriers of such materials are always interested in the application of models that minimize the risk of such accidents. These models usually involve five components: (1) accident likelihood, (2) containment failure, given an accident, (3) volume of rate of material released, (4) the hazard area associated with each potential threat for different releases and materials, and (5) the population affected by the hazard. In short, the models attempt to estimate both the risk of an accident on a given route and the potential damage to the environment and population given this accident and its location. Obviously, the risk is also dependent on the type of product carried and the characteristics of the vehicle carrying the product.

A wide variety of algorithms are available for determining the minimum risk route on such networks and all of them use traffic flow as the measure of exposure. The proposed algorithms can be simple, such as the one put forward by Ashtakala and Eno (18), based on the Moore algorithm. Alternately, they can be highly complex such as the algorithm proposed by Miller-Hooks and Mahmassani (19), which predicts the optimal routing on a stochastic, time-dependent transportation network. Many proposed models also include a constraint imposed by the operating cost, which, in effect, makes the safest routes only partially optimized for safety. The models are not described here since they fall outside the scope of this research.

Accident risk has also been recently employed for intelligent transportation systems (ITS) applications. Similar to the above, the goal is to find the safest routes for an individual vehicle within a given network aimed at re-routing traffic to routes with lower risk with the help of dynamic route guidance (DRG) systems. Al-Deek et al. (20), Burrow and Taylor (21), Lord et al. (22), and Look and Abdulhai (23) have applied accident risk on urban networks incorporating ITS technology.

The estimation of accident risk for transportation network safety analyses is unfortunately plagued by an important drawback: most minimizing risk models currently employ accident counts per unit of time and segment (nodes or links) as a measure of safety. By using accident counts, one assumes a linear relationship between accidents and traffic flow. However, as explained above, the relationship between accidents and traffic flow has been shown to follow a non-linear relationship. In addition, accident counts have been found to be an inaccurate estimation of safety since they are usually random and independent events (24). Thus, one should use the long-term average or the expected number of accidents instead of actual accident counts.

APMs are extremely useful tools for the safety analysis of transportation facilities, as they have a wide range of applications. They can be used to: 1) find the proper relationship between accidents and traffic flow (and other variables); 2) determine the long-term average in accidents for transportation entities; 3) compute the expected number of accidents for sites that have yet to be built or upgraded.

The issue of using accident counts as a measure of accident risk has been investigated by Mahalel (16). Mahalel proposed a method for estimating accident risk as a function of exposure, accidents and probability modelling. He argued, correctly so, that risk varies in a non-linear fashion with exposure since traffic flow is a function of vehicle density and speed. Hence, similar flows could be attributed to different combinations of density and speed, leading to different accident risk values. In essence, his method shows that accident risk should be estimated solely from APMs, which describe the proper relationship between accidents and traffic flow.
ESTIMATION OF ACCIDENT RISK

The APMs used in this work were specifically created for computerized transportation network applications. These models are used to predict accidents on links and nodes. Three different model types were applied. The first type was used for predicting accidents at nodes. The second and third types were utilized for predicting accidents on links. As suggested by Jackett (25), the prediction of collisions on links could be separated into two components: mid-block and intersection. The mid-block component models predict the number of accidents between minor intersections located on the physical network but not encoded as nodes on the computerized or digital network. The intersection component models estimate the number of accidents at these minor intersections. They only use the link flow as input. A detailed description on how the models were developed and should be applied on networks can be found in (26), (27) and (28).

The APMs were applied on a hypothetical network that was built with the help of EMME/2 (29). This software is a transportation planning computer program that is often used to assess traffic flows at a regional level. It is based on the traditional 4-stage transportation modelling process (trip generation, trip distribution, modal split, and traffic assignment). EMME/2 employs a static optimization algorithm (i.e., not time-dependent) and provides flows in vehicles per hour. Since the mid-70s, it has been used extensively throughout the world by various local agencies.

The network has 15 nodes, 6 centroids, 18 links, and 6 centroid connectors (Figure 1). Nine nodes are used to represent intersections, while the other six nodes are employed as gateway nodes located between the centroids at the outskirts of the network and the street system. An origin-destination matrix consisting of 9800 vehicle-trips was assigned to this network. The product of this assignment was expanded to AADT flows with an average expansion factor computed from the recorded traffic in Toronto, Canada. Under the assumption that the estimated flows were predicted accurately by EMME/2, the expected number of accidents was computed on this network with the models detailed in (26), (27) and (28). The models were developed with AADT as the input variable since their original application was related to the annual prediction of accidents for a given computerized transportation network.

![Sample network](image-url)
The risk was estimated with equation (2) for each link and node on the sample network. The definition based on this equation is a widely-used measure of individual risk. It has been applied in various traffic safety studies (4, 30, 31). Equation (2) can be defined as the driver’s risk of being involved in a collision given a certain level of exposure. That is, the probability that a driver would be involved in an accident on a link or node that has, on average, x accidents per day (or year) and y vehicles per day. The equation is as follows:

\[
\text{Accident Risk} = \frac{\text{Expected Number of Accidents}}{\text{Exposure}}
\]  

(2)

The accident risk in equation (2) is estimated by dividing the expected number of accidents per unit of time with the traffic flow in vehicles per unit of time for any given link or node, as applied by (31). By computing the expected number of accidents with an APM, equation (2) becomes:

\[
\text{AR} = \frac{\text{E}\{K\}}{\text{Flow} \times 365}
\]  

(3)

where,

- \(\text{AR}\) = the accident risk defined as the number of accidents per vehicle per link or per passing intersection;
- \(\text{E}\{K\}\) = the expected number of accidents per year (estimated with an APM);
- \(\text{Flow}\) = the traffic flow (in AADT) on the link or for the manoeuvre at a node (right, through, left) in vehicles per day;
- 365 = the number of days in one year.

COMPUTATION OF SAFEST PATHS FOR INDIVIDUAL VEHICLES

The safest path in this analysis was found by locating the route between each pair of centroids for which a driver has the lowest probability of being involved in an accident. Equation (3) was used to compute the accident risk for the links (those found in the original network). In the analysis, the same risk was assigned to drivers travelling in opposite directions on each link. The disaggregation of accident risk by direction was not possible with the models developed for this research.

The computation of accident risk for the nodes was somewhat more complicated and required further manipulation. Many commercially available transportation planning software programs, including EMME/2, cannot assign external cost factors for different turning manoeuvres at nodes. Therefore, it was necessary to modify each node (13 to 17, as illustrated in the previous Figure 1) in view of the fact that the risk of a collision depends on the type of manoeuvre at a node (e.g., turning right, left or travelling through). Each node was split into eight “sub-nodes”, as depicted in Figure 2. These each required connection to three separate links—one for left-turns, one for right turns and one for vehicles going through straight through the intersection.
A few steps were necessary to separate the predicted accidents by type. The models developed for the nodes in this research were not separated by collision type, due to the lack of available data. It should be mentioned that the best approach, if such typing is necessary, would have been to develop specific APMs for each accident type. These types of models are more appropriate for the computation of accident risk, as detailed below. In the first step, the proportion of accidents by type was examined for signalized 3- and 4-legged intersections in the Toronto accident database. Based on this exploratory analysis, it was possible to compute the proportion of accidents for drivers who turn left, right, or travel straight through the intersection. In the second step, the total expected number of accidents computed for each intersection was multiplied by the proportions calculated in the first step. This resulted in finding the predicted number of accidents for each turning movement at the intersection. In the third step, the risk was computed with equation (3) for each manoeuvre at the node. In the end, a different accident risk value was used depending on whether a driver turns left, right, or travels through a node. The limitations regarding for approach is discussed further below.

After the accident risk was computed, it was then assigned to the modified network. The risk computed from the original link had to be allocated to all the modified links (between the receiving and outgoing sub-nodes for each leg) located between any adjacent nodes. For every node in the network, the risk was assigned to all additional links that represented the various turning movements at the intersection.

The description can be better understood with the help of an example. With the network described in Figure 2, let us assume that the risk between node 5 and each of the other adjacent nodes is equal to 0.0500 acc/veh; that a left-turn is equal to 0.0045 acc/veh; that travelling through an intersection is equal to 0.0035 acc/veh and that turning right is equal to 0.0020 acc/veh. The assignment of the various risk factors is presented in Figure 3.
Different accident risk values, as computed by equation (3), were applied on the computerized network presented in Figure 1. Each node on this network had to be modified according to the procedure described above. Hence, the number of nodes increased from 15 to 66 in the new modified network.

Figure 4 shows the accident risk on the sample network. The number in parentheses in this figure is the expected number of accidents per year. The different values of accident risk at nodes are presented in Table 1. Note that the high-risk values in this table are due to extremely low flows at the nodes. The risk applied in this situation is disproportionately large, owing to the low denominator and the use of aggregated accident prediction models. Since the predicted accidents for the turning flows were estimated with average proportions, the individual accident risk can, in certain occasions, be very high when the average predicted accidents are divided by low flows, as detailed in equation (3). One way to avoid this problem is to set the nodes with a minimum flow. For instance, Chatterjee and McDonald (31) propose setting the minimum flow as equal to the lowest flow in the sample of sites producing the APMs. The blank cells indicate that this particular turning manoeuvre is not possible for that node.
FIGURE 4 Assignment of risk on the sample network
<table>
<thead>
<tr>
<th>Node</th>
<th>Acc/yr</th>
<th>NB_L</th>
<th>NB_T</th>
<th>NB_R</th>
<th>SB_L</th>
<th>SB_T</th>
<th>SB_R</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>12.7</td>
<td>1.2E-6</td>
<td>6.3E-8</td>
<td>1.0E-6</td>
<td>1.3E-6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8.0</td>
<td>1.9E-6</td>
<td>8.6E-8</td>
<td>6.5E-7</td>
<td>1.9E-6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>12.3</td>
<td>1.0E-7</td>
<td>2.7E-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>25.2</td>
<td>1.0E-6</td>
<td>8.3E-7</td>
<td>7.1E-7</td>
<td>1.9E-5</td>
<td>7.5E-7</td>
<td>2.0E-7</td>
</tr>
<tr>
<td>15</td>
<td>11.6</td>
<td>2.2E-7</td>
<td></td>
<td></td>
<td>7.9E-8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>8.5</td>
<td>6.7E-7</td>
<td>1.1E-6</td>
<td></td>
<td>8.9E-7</td>
<td>2.5E-7</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>37.1</td>
<td>8.3E-7</td>
<td>1.8E-6</td>
<td>4.3E-7</td>
<td>9.8E-7</td>
<td>2.2E-6</td>
<td>1.6E-7</td>
</tr>
<tr>
<td>18</td>
<td>15.2</td>
<td></td>
<td></td>
<td></td>
<td>1.0E-7</td>
<td></td>
<td>3.3E-7</td>
</tr>
<tr>
<td>19</td>
<td>20.6</td>
<td>1.2E-6</td>
<td>7.2E-7</td>
<td>1.2E-6</td>
<td>1.1E-6</td>
<td>7.4E-7</td>
<td>1.2E-7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acc/yr</th>
<th>EB_L</th>
<th>EB_T</th>
<th>EB_R</th>
<th>WB_L</th>
<th>WB_T</th>
<th>WB_R</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>12.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.8E-8</td>
</tr>
<tr>
<td>13</td>
<td>12.3</td>
<td>1.2E-6</td>
<td>7.9E-8</td>
<td></td>
<td>1.9E-5</td>
<td>8.6E-7</td>
</tr>
<tr>
<td>14</td>
<td>25.2</td>
<td>1.2E-6</td>
<td>1.5E-6</td>
<td>5.0E-7</td>
<td>2.2E-6</td>
<td>3.6E-6</td>
</tr>
<tr>
<td>15</td>
<td>11.6</td>
<td>9.7E-7</td>
<td>2.2E-7</td>
<td>5.3E-7</td>
<td>1.0E-6</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>8.5</td>
<td>2.1E-7</td>
<td></td>
<td>1.1E-7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>37.1</td>
<td>8.2E-7</td>
<td>1.5E-6</td>
<td>1.4E-7</td>
<td>1.9E-6</td>
<td>1.2E-6</td>
</tr>
<tr>
<td>18</td>
<td>15.2</td>
<td>2.2E-6</td>
<td>1.5E-6</td>
<td></td>
<td>1.5E-6</td>
<td>8.1E-8</td>
</tr>
<tr>
<td>19</td>
<td>20.6</td>
<td>1.2E-6</td>
<td>1.8E-6</td>
<td>2.3E-7</td>
<td>6.5E-6</td>
<td>2.0E-6</td>
</tr>
</tbody>
</table>

* shaded values indicate high risk movements
A computer program written in QuickBASIC (32) was created to facilitate the search of the safest paths in the sample network. The Dantzig algorithm was coded according to the method proposed by Yen (33). Another computer program was also written with the Floyd algorithm to validate the results of the first program’s output.

The total risk for the safest path between each pair of centroids in the sample network is presented in Table 2. An example of the safest path between centroids 1 and 5, and between centroids 3 and 4 is shown in Figures 5 and 6 respectively. The individual accident risk was calculated by adding the risk in acc/veh for each segment of the network (nodes and links) for all possible routes between each pair of centroids. The safest path is the route with the lowest value. The risk for each route was computed with the values illustrated in Table 1 and Figure 4.

### TABLE 2 Total risk (acc/veh) for the safest path between each pair of centroids

<table>
<thead>
<tr>
<th>Centroids</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>6.05E-6</td>
<td>1.12E-5</td>
<td>13.2E-6</td>
<td>13.2E-6</td>
<td>8.41E-6</td>
</tr>
<tr>
<td>2</td>
<td>4.74E-6</td>
<td>0</td>
<td>9.62E-6</td>
<td>10.0E-6</td>
<td>11.1E-6</td>
<td>11.3E-6</td>
</tr>
<tr>
<td>3</td>
<td>9.95E-6</td>
<td>5.21E-6</td>
<td>0</td>
<td>6.14E-6</td>
<td>7.32E-6</td>
<td>1.42E-6</td>
</tr>
<tr>
<td>4</td>
<td>12.7E-6</td>
<td>9.10E-6</td>
<td>5.23E-6</td>
<td>0</td>
<td>1.92E-6</td>
<td>10.2E-6</td>
</tr>
<tr>
<td>5</td>
<td>1.32E-5</td>
<td>9.53E-6</td>
<td>5.67E-6</td>
<td>0.44E-6</td>
<td>0</td>
<td>9.86E-6</td>
</tr>
<tr>
<td>6</td>
<td>8.27E-3</td>
<td>1.21E-6</td>
<td>13.3E-6</td>
<td>9.59E-6</td>
<td>9.15E-6</td>
<td>0</td>
</tr>
</tbody>
</table>

**FIGURE 5** Safest path between centroids 1 and 5
Figure 5 reveals that the safest route is laid out in a stepwise fashion, with three right turns and two left turns. This path initially may appear inappropriate, since the addition of the risk factors associated respectively with links 16-19, and 16-17 (without including the risk at the nodes) is less than those associated with links 18-19 and 17-18. Since the accident risk for a right-turn is 8 to 16 times safer than travelling through at nodes 19 and 17, the sum of the risk factors on route 19-18-17 (7.5E-6 acc/veh) becomes less than that on route 19-16-17 (8.0E-6 acc/veh).

Figure 6 shows that it is safer to perform a detour through nodes 19-16-15 than to make a left turn at node 14 and travel on link 14-15. In real traffic conditions, it would be unlikely that a driver would take such a path. In fact, the shortest path is often the path with the lowest risk (31). Since the left-turning traffic flow at node 14 is very low, the accident risk is very high. The risk for that particular left turn is about 20 to 30 times higher than the risk of turning left at other nodes in the network. Thus, it is safer to perform a detour than to use the most direct route, as unrealistic as this may seem. It is important to mention that this hypothetical digital transportation network did not represent an actual physical network and that all the variables were set arbitrarily. Furthermore, the expected number of accidents for turning manoeuvres was not estimated from APMs but rather from an average value that was extracted from accident data. Therefore, it is not surprising that unusual routes could be selected in this network.

DISCUSSION OF RESULTS

Although the definition of accident risk used in this work is a valid method of estimating risk, its application to transportation networks nonetheless leads to a very important result. Indeed, by looking more closely at the relationship of equation (3), one can rewrite the equation the following way:

\[
AR = \frac{E\{K\}}{\text{Flow} \times 365} = \frac{\alpha F^\beta}{\frac{F \times 365}{365}} = \frac{\alpha F^{(\beta - 1)}}{\text{Flow} \times 365} = \frac{\alpha F^{(\beta - 1)}}{365}
\]

(4)
Equation (4) exhibits a special characteristic for beta $\beta<1$. Most models found in the literature contain a coefficient $\beta$ below 1. When the power coefficient of the flow parameter $F$ is negative, equation (4) indicates that the individual risk decreases as the flow increases. In fact, the relationship implies that it is less dangerous for a driver to travel under heavy flow conditions than under light flow conditions. For example, let us assume that an APM has the following form $0.01x\text{Flow}^{-0.5}$ (single vehicle accident): when the model is applied on links with traffic flows equal to 100 veh/unit of time and 10 veh/unit of time, one gets 0.10 acc/unit of time and 0.03 acc/unit of time respectively. In principle, 10 vehicles out of the 100 vehicles and 3 vehicles out of the 100 vehicles should be involved in a collision respectively. Consequently, the probability for a driver to be involved in a collision becomes 0.001 and 0.003 (Figure 7). This result shows that a driver has a greater likelihood of being involved in an accident if he or she travels on the link with 10 veh/unit of time. The special characteristic of equation (4) merits a thorough analysis.

![Accident risk function](image)

**FIGURE 7 Accident risk function**

The characteristic exhibited in equation (4) appears to be counterintuitive. Common sense would dictate that the individual risk of being involved in an accident should increase as more road users share the same limited space on the road. Similarly, when one looks at human factors, the driving task should become more complicated as traffic increases. The driver should have greater difficulty absorbing information, thus, the probability for a driver to make errors should be larger leading to a greater likelihood of being involved in an accident. Recent work on traffic safety and capacity also showed that as the level of service (LOS) decreases, the risk per vehicle becomes higher. Persaud and Look (34), and Persaud and Nguyen (35) developed APMs in relation to the LOS for freeway sections and signalized intersections. The results showed that, for similar flows, more accidents were predicted when the LOS decreased.

Despite what intuition and the preliminary empirical work on this subject might indicate, the relationship of equation (4) shows that it is safer to travel on roads with higher flows than with lower flows when the coefficient $\beta<1$. The issue about the diminishing risk is directly related to the coefficient $\beta$. As explained in the first section, studies have shown that the number of accidents usually increases at a decreasing rate when the flow increases. Thus, the individual risk indeed becomes less as traffic flow increases. Even the models proposed by Persaud and Look (34), and Persaud and Nguyen (35) contain a $\beta<1$ for the flow parameter $F$. Consequently, for a given LOS, it is safer for an individual to travel on a road or through an intersection that has more vehicles than less vehicles, everything else being equal (i.e., same road characteristics).
A similar characteristic can be found when one looks at the societal risk. The goal, in this case, is to compute the accident likelihood for a given level of exposure. The accident likelihood is usually estimated from accident rates rather than from APMs. For this definition, the relationship of equation (3) becomes:

\[
\text{Accident likelihood} = E\{K\} = \alpha F^\beta
\]  

Equation (5) may be estimated for different categories of road users such as for heavy vehicles or trucks.

Equation (5) shares the same property as in equation (4) for beta \( \beta < 1 \) when used for the optimization of network safety. In minimizing societal risk, the flows on networks will tend to concentrate on few links rather than disperse on many links since risk decreases with flow. Maher \textit{et al.} (36), who attempted to simultaneously optimize accidents and vehicle delay on computerized transportation networks, acknowledged this unusual characteristic. Maher \textit{et al.} found that when the network is solely optimized for safety, flows tend to concentrate on a few links or road segments, leaving certain links empty.

So far, there has been no explanation in the literature to interpret the relationship of beta \( \beta < 1 \). This relationship may be explained by various factors. It is possible that drivers may change or adapt their driving behaviour according to the level of traffic present on the road. They may become more alert as traffic increases, and/or undertake less risky manoeuvres knowing that space on the road is limited. Another explanation might be related to speed. As traffic flow increases, the vehicles travel at a lower speed. The speed usually slightly decreases until near capacity conditions where, at that point, it sharply goes down. The damage caused by an accident to both the vehicle and its occupants is less severe at lower speed. Thus, perhaps accidents that were once serious may become less serious; slight accidents may become property damage only (PDO) accidents (which are less likely to be reported), etc. The speed distribution may also play an important role. The variance in speed is probably higher for lower traffic flow than for higher traffic flow conditions. As explained by Garber and Gadiraju (37), a larger variance in the distribution of the speed leads to a decrease in safety. The speed may become more uniform when traffic density increases. Consequently, the likelihood for a conflicting situation to arise between vehicles is higher at a lower density.

Although the explanations put forward above may be valid, the issue of diminishing risk should nonetheless be explored in much greater detail since it may have important consequences on transportation policies and ITS strategies. For instance, based on the computation presented above, a network administrator should, in theory, send vehicles on a minimal number of links (through Dynamic Route Guidance (DRG) systems or otherwise). This traffic assignment scheme will lead to an enormous increase in vehicle delay in which no drivers will be willing to tolerate. Furthermore, the counterintuitive aspect of these findings would likely discourage commuter use of more-heavily used links. In principle, there is more risk involved travelling on a link by yourself than when other vehicles are present (see Figure 7). Obviously, commuters would prefer to travel alone rather than with other people on any given road.

Perhaps the current model forms found in the literature are not appropriate. Recent work on this subject has shown that the relationship between accidents and traffic flow follows a Gamma function for particular transportation facilities (38, 39). This relationship is characterized by the fact that accidents initially increase, then peak and then decrease as flow increases. The accident risk in this situation follows a similar pattern. In other instances, recent work on the study of the relationship between the volume-to-capacity ratio and accident rates have shown that it follows a U-shaped function (40, 41, 42). Finally, Garber and Subramanyan (43) found that the relationship between accidents and vehicle occupancy follows a polynomial function of the third order, very similar to the relationship established in (38) and (39).
It is therefore possible that traffic flow may not be a suitable measure of exposure. For a given flow, two possible traffic conditions can occur: high speed and low density or low speed and high density. Thus, it may be better to utilize density as a measure of exposure rather than flow for the computation of accident risk. Conversely, segregation of traffic flow according to whether traffic is moving under congested or uncongested conditions may be similarly useful. Unfortunately, information on density is not easily obtainable, as it is only measured infrequently. With new ITS technologies, it may become easy to record density information.

Finally, further work should be performed on the application of disaggregated prediction models to estimate the individual risk of a collision on transportation networks. These kinds of models should include variables such as the age and sex of the occupant, the use of alcohol by a driver, the longitudinal grade of the road and the presence of street lighting, among others.

CONCLUSION

This research has shown that accident risk computed with APMs leads to surprising - if not counterintuitive - results. In essence, the individual risk of being involved in a collision actually decreases as traffic flow increases. The current and most common model form of AMPs explains this outcome. Although this mathematical formulation of accident risk is indisputable, the application of such a definition could be difficult to implement, due to very the counterintuitive nature of the conclusion. It is therefore suggested that further research be pursued on the relationship between accidents and traffic flow and its use in the computation of accident risk. Additional work on the use of density for safety applications and as a measure of accident risk should be explored.

ACKNOWLEDGEMENT

The research was partly supported by an operating grant from the Natural Sciences and Engineering Research Council of Canada (NSERC). This support is greatly appreciated, as is the assistance of the city of Toronto Transportation Department who provided the data for calibrating the accident prediction models used in the research.

REFERENCES


