METHODOLOGICAL ISSUES IN THE IMPLEMENTATION OF ACCIDENT REMEDIAL SCHEMES

CJ Mollett Pr Eng
Chief Engineer : Traffic Engineering
Department of Economic Affairs , Agriculture and Tourism : Branch - Transport
25 Alfred Street, Cape Town, 8001

Dr CJ Bester
Department of Civil Engineering
University of Stellenbosch
Private Bag X1, Matieland, 7602

1.   Introduction

A programme of action for the remedial treatment of accident prone locations basically involves a four-phase process :

a)   the identification of accident prone locations (APL's),
b)   the prioritisation of APL's,
c)   the identification and implementation of appropriate remedial measures, and
d)   the evaluation of the effectiveness and efficiency of remedial measures.

An accident prone location is a location that experience an significantly above average number of accidents when compared to similar sites.

The identification phase involve the initial identification of 'unsafe' sites based on measures of accident risk, such as an accident rate.

The prioritisation phase involves ranking the sites according to the likely cost-effectiveness of remedial measures. It is important that scarce resources are allocated to those schemes that provide the most cost-effective return on investment.

The identification of appropriate remedial measures requires a detailed investigation into the factors contributing to accidents, with a view of devising strategies to reduce the negative impact of these factors.

The objective of the evaluation phase is to determine the effectiveness of remedial treatments in order to determine the extent to which targets are met and for the quantified effects of various treatments to be fed back into future studies. Once the effectiveness is known, the efficiency i.e. actual cost effectiveness can be determined.

Before one any of these phases can be embarked upon it is important to first analyse accident data in a systematic manner to obtain a reliable and valid measure of the true level of safety at a location.

For the purposes of this paper the level of safety at a location is defined as the 'true' underlying accident rate. An *accident rate* represents the risk to a road user for using the transportation system.

A measure of safety is the end product of a process of measurement, (the principles of which will be described), when raw accident data is partitioned and validated (by means of classifications and categorisations) in a manner that ensure these indicators accurately represent the true level of safety at a location. These measures of safety form the foundation for all methods and procedures relating to accident remedial schemes.

It is important that the results obtained from applying *identification*, *prioritisation* and *evaluation* methods and procedures to reflect reality as closely as possible. The extent to which 'reality' is achieved will depend firstly on the accuracy of the safety measures, and secondly on the reliability and validity of the analytical method chosen to do the analysis. The mathematical and statistical procedures can be fairly straight forward but there are always special considerations and pitfalls that the analyst must be aware of.

The objectives of this paper are to a) define reliable estimates of safety at a location in order to *identify*, *prioritise* an *evaluate* accident prone locations, b) to provide guidelines on choosing appropriate methodologies and c) to highlight where appropriate the special considerations and pitfalls the analyst should be aware of.

2. THE MEASUREMENT OF SAFETY

2.1 INTRODUCTION

This section will first explain the concepts of reliability and validity and what steps can be taken to improve the reliability and validity of a safety measure.

The statistical properties of accident data will then be investigated. Knowledge of these properties is important to decide on the most appropriate methodology to measure safety. Thereafter a methodology based on an emperical Bayesian approach will be presented.

2.2 MEASUREMENT THEORY

2.2.1 Validity and Reliability

With any type of measurement *validity* and *reliability* are two very important considerations.

Validity is concerned with the soundness and the effectiveness of the indicator. It raises the question whether an indicator measures what it is suppose to measure as well as how comprehensively and accurately.

Reliability deals with how dependable and accurate an indicator is. An indicator must be reliable across time (stability reliability) as well as across sub-populations (representative reliability).

2.2.2 Measurement Error

Any measurement (X) or observed value consists of 3 different components: a)True perfect value (T), b) Systematic Error (S) and c) Random Error (R)

Where $X = T + S + R$

Under conditions of perfect validity $X = T$ i.e. there are no systematic and random errors. The validity and reliability of a measurement can be improved by focussing on eliminating S and R.

2.2.2.1 R - Random error

Accidents are random events. It is therefor unlikely that an observed measurement of accident data will be equal to the true value T. For more details please refer to Section 2.3 on the statistical properties of accident data.

The random error associated with a safety measurement can be reduced by using an appropriate methodology for estimating safety. For more details please refer to Section 2.4 on methods for the estimation of safety.

2.2.2.2 S - Systematic Error

Whereas random error is a statistical phenomenon, systematic errors are 'man-made' and can be avoided. It is the result of external influences, which distorts a measurement in a systematic way. The most common causes of systematic errors relating to accident data is a) the underreporting of accidents, and the b) inaccurate, incorrect and incomplete completion of accident report forms.

In order to minimise the systematic error the following remedial measures could be taken:

a) Exclude from the analysis data that is associated with significant systematic errors such as for example 'damage only' accident data.
b) Systematic errors as a result of inaccurate or missing data can really only be reduced by having a good, well managed, accident data collection system.

2.3   STATISTICAL PROPERTIES OF ACCIDENT DATA

The number of accidents (x) at a location is statistically described as a variable with a Poisson distribution :

$$P(x/m, E) = \frac{(E.m)^x e^{-m.E}}{x!} \qquad ...[1]$$

m -   the true level of safety – e.g. accidents/mvkm (denoted by T in Section 2.2.2).
E -   the level of exposure - e.g. million vehicle kilometres (mvkm)

The MLE (Maximum Likelihood Estimator) of *m* is given by $R = \dfrac{x}{E}$.

According to Nembhard(15) *R* is a reasonable estimate of *m* provided *E* is large. For small values of *E* the estimate becomes unstable. The MLE estimator can be improved on by using a Bayesian estimator.

The difference between *x* and *m*E* is related to the random error mentioned in Section 2.2.2.1.

The value of *m* varies between different sites according to the Gamma distribution:

$$f(m) = \frac{\alpha^\beta m^{\beta-1} e^{-\alpha m}}{\Gamma(\beta)} \qquad ...[2]$$

Where

α - The shape parameter and β - The scale parameter.

$$E(m) = \frac{\alpha}{\beta} \qquad ...[3] \qquad\qquad VAR(m) = \frac{\alpha}{\beta^2} = \frac{E(m)^2}{\beta} \qquad ...[4]$$

The distribution of *x* between different sites can be described by the Negative Binomial distribution:

$$P(x_i) = \frac{\Gamma(x_i + k)}{x_i! \Gamma(k)} \left(\frac{k^{-1}\overline{x}}{1 + k^{-1}\overline{x}} \right)^{x_i} \left(\frac{1}{1 + k^{-1}\overline{x}} \right)^k \qquad ...[5]$$

Where K = 1/k is the dispersion parameter and the mean and variance are:

$$E(x_i) = \bar{x} \qquad \text{...[6]} \qquad\qquad VAR(x_i) = \bar{x} + \frac{(\bar{x})^2}{k} \qquad \text{...[7]}$$

The smaller the value of K the less the dispersion. If K = 0 then Equation 5 reverts to the Poisson distribution. The value of K is a measure of the degree of similarity between a sample of sites. The smaller the value of K the more similar the sites are, and vice versa.

2.4   SAFETY ESTIMATION

A common method of defining the safety at a location is by means of a risk estimate. Risk estimates are used to describe the safety level of transportation systems in a manner that is invariable to their exposure level. The most widely used measure of transportation-system risk is the *accident rate*.(8)

An accident rate is determined by dividing the observed accident frequency by a measure of exposure.

Exposure can be defined as the number of opportunities for an accident to occur. Exposure is thus related to the 'intensity' of use. The determination of an accident rate in this manner serves to compensate for differences in 'intensity of use' thereby making comparisons between different entities and time periods more meaningful.

The challenge facing the analyst is to find a reliable and valid measure of safety (accident rate) that will satisfy the research objective. This requires that decisions be made on an appropriate *accident measure* as well as an appropriate *measure of exposure.*

2.4.1   Accident Measure

The chosen accident measure must support the objective of the research. If, for example, the objective of the research is to make inferences about pedestrian safety only pedestrian related accidents should be considered, or if the objective is to assess the safety impact of surface treatment on intersection approaches only head-rear type collisions on the treated approaches could be considered.

A decision will also be required as to which severity classes to include in the accident measure. As shown before the reliability of a safety estimate could be enhanced by excluding 'damage-only' type accidents. However excluding 'damage-only' type accidents will not always be appropriate depending on the objective of the research.

If one of the objectives is also to make inferences about the severity of accidents at a location then the accident measure could be expressed in terms of an EAN (Equivalent Accident Number). The EAN is determined by applying a weight factor to the total number of accidents in each severity class (fatal, serious, slight and damage only) and then summing up across all the classes. The magnitude of the weighting factor is related to the relative cost of an accident where 'damage-only' type accident has a weight equal to 1 (one).

2.4.2   Exposure measure

The main reason for the inclusion of an exposure measure is to compensate for differences in intensity of use so as to make comparisons meaningful. In order for exposure to do the job of compensating the safety performance function (the relationship between the accident measure and exposure) must be a straight line i.e. linear.

When the Safety Performance Function is not a straight line, the accident rate will change as the amount of traffic change even if there was no intervention and the road remained the same. It is possible for the accident rate to decrease even as the facility becomes less safe. It is also possible when two facilities are compared with each other for the safer facility to have a higher accident rate than the other facility.

For an exposure measure to be completely valid the relationship between it and the accident measure must be completely linear. This can be achieved by identifying an exposure measure that is highly correlated with the total number of opportunities for an accident, as determined by the accident measure, to occur.

The most common measure of exposure for estimating the accident rate on road segments is as follows:

$$E = AADT.*L*n*365*10^{-6} \qquad\qquad ...[8]$$

and for intersections :

$$E = \sum AADT_{in} *n*365*10^{-6} \qquad\qquad ...[9]$$

AADT	- Annual Average Daily Traffic
L	- Length of segment in km.
n	- Number of years over which data was collected
$\Sigma AADT_{in}$	- Sum of all flows entering an intersection

These measures of exposure do not provide reliable estimates of safety for the following reasons:

a) Various researchers have found (5) that the relationship between E and most accident measures (e.g. all accidents, injury accident etc.) to be non-linear. For instance the number of single-vehicle accidents tend to increase initially with E, then reach a maximum and, as traffic grows further, the number of single vehicle accidents declines.

Hauer (5) claims that the use of the exposure measure as shown in Equation 9, especially for particular intersection accident types (e.g. head-rear accidents), to be logically unsatisfactory. The number of rear-end collisions on approach A of an intersection will strongly depend on the flow on approach A and depend only weakly on the flow on approaches B, C and D. Similarly it is expected that collisions between vehicles from streams A and B moving at right angles to each other might be related to the product of flows A and B.

In order to evaluate the safety effect of converting angle parking to parallel parking along a street McCoy et al. (9) used as a measure of exposure, not the total vehicle miles over the study section, but the total number of vehicle mile-hours per parking bay. This measure of exposure was determined from the product of travel and parking use per parking bay. It is considered a more reliable measure of exposure than the traditional total vehicle miles measure because it takes into account the level of parking activity. In a similar fashion the level of pedestrian safety at a location cannot be determined reliably by using a measure of exposure that does not account for the level of pedestrian activity.

b) In the exposure function a time-average value of traffic flow, the AADT is used. There exist a causal relationship between flow and accidents, The effect (accidents) is observed over a long period of time during which the cause (traffic flows) has assumed widely different values, but for which only the average value of the cause is known. The causal link between accidents and AADT is therefor indirect. This lead to the issue of *argument averaging* which according to Mensah et al. (11) could introduce, depending on the form of the SPF, a large and significant bias.

2.4.3 Estimation Methods

Even after choosing a reliable and valid accident measure and a measure of exposure to calculate an accident rate the reliability of such an accident rate can still be in question because of the *regression-to-mean effect*.

Because of the random nature of accidents it is unlikely that the accident rate (R) at a location is equal to the true underlying accident rate (m). According to Nicholson (6) five years is the most suitable time period upon which to base an estimate of the true level of safety. According to Berlanger (2) using such long periods does not solve the problem because many factors can influence the value of m over time, such as changes in speed limits, traffic patterns etc. Therefor R is not necessarily a reliable measure of the safety even if it was estimated from data collected over extended periods of time. Using R instead of m to make inferences could introduce the *regression- to-mean* bias. If R was used to choose sites for treatment because of a poor accident record then the value of m will be over estimated and a potentially large positive regression-to-mean effect will result from this.

The first step towards determining and eliminating the regression-to-mean effect is to identify a sample of sites, called a reference group. The requirement is that the expected number of accidents of a treated site with a set of given characteristics (geometry, traffic flow) be the same as the expected accident frequency of a reference group site with identical characteristics.

An initial estimation of safety for a particular location can be obtained by determining the average accident rate for all reference group sites.

$$E(m) = \frac{1}{n} \sum_{i=1}^{n} R_i \qquad \qquad \text{...[10]}$$

n – Number of sites in reference group

In spite of being similar each site could have its own regional character, driver population, etc. giving it a unique level of safety *m*. Therefor *m* varies from site to site. The distribution of *m's* between sites in the reference group can be described by a Gamma density function with mean *E(m)* and variance *VAR(m).*

A more accurate Bayesian estimate of the level of safety at a location can be obtained by combining the initial estimate E(m) and the observed accident rate at a site R using the following equations . This procedure will ensure that the true level of safety reflects also the characteristics that are unique to that site.

$$E(m_i \mid x_i, E_i) = \left(\frac{E_i}{E_i + \alpha} \right) R_i + \left(\frac{\alpha}{E_i + \alpha} \right) E(m) \qquad \text{...[11]}$$

where

$$\alpha = \frac{E'\overline{R}}{E'S_r^2 - \overline{\overline{R}}} \qquad \qquad \text{...[12]}$$

where *E' – the harmonic mean of the exposures (E) of all sites in the reference group.*

The true level of safety at a site is a convex combination of the observed accident rate (*R*) and the underlying Gamma mean, *E(m).* As mentioned before at low levels of exposure the accident rate *R,* the maximum likelihood estimator of *m*, becomes unstable. With the Bayesian estimator (Equation 11) the significance of *R* reduce with decreasing values of *E* while the significance of *E(m)* increases. Thus when a site has a low level of exposure its true level of safety will be closely related to the average level of safety for similar locations while the level of safety at a site with a high level of exposure will be closely related to the observed accident rate.

To allow for, amongst others, the non-linear nature of safety performance functions Hauer et al.(4) has supplemented the empirical Bayesian approach with multivariate modelling techniques. Instead of using data from the reference group to calculate E(m) and VAR(m) directly, the data is used to develop a prediction model by regressing the observed accident frequency at a site to its traffic flows (AADT's).

E(m) and *VAR(m)* are determined from the calibrated accident prediction model. These values are then combined with th observed accident frequency at a particular site in order to determine E*(m|x)* and *VAR(m|x)* using the following equations :

$$E(m \mid x) = aE(m) + (1-a)x \qquad \qquad \text{...[13]}$$

$$VAR(m \mid x) = a(1-a)E(m) + (1-a)^2 x \qquad \qquad \text{...[14]}$$

where

$$a = \frac{E(m)}{E(m) + VAR(m)} \qquad \ldots[15]$$

Mountain et al. (12) has indicated how *E(m)* and *VAR(m)* can be estimated using an existing prediction model without having to go through the multivariate regression exercise.

The methods proposed by Hauer et al.(4), Mountain et al.(12) are all based on an empirical Bayesian approach. Since the Bayesian approach estimates and work with true levels of safety the regression to mean effect is automatically accounted for.

3.   IDENTIFICATION OF ACCIDENT PRONE LOCATIONS

A location can be considered accident prone it experiences an abnormally high accident rate or accident potential relative to locations of similar characteristics or relative to a chosen criterion.

The Institution of Highways and Transportation's 1986 Guidelines refer to four types of accident reduction plans ; a) single site (blackspot) plans , b) route action plans (blackroutes), c) area action plans and d) mass action plans.

According to Nicholson (15, 16) experience with accident reduction programmes indicates that the expected percentage reduction and economic return is higher for site plans than route plans, and higher for route plans than area plans. The effectiveness and economic return of an accident reduction plan, however, depends on the extent to which accidents are clustered at particular sites or along particular routes or areas. The implementation of a single site reduction plan while the level of clustering is low is unlikely to yield better benefits then route or area plans.

The higher the level of clustering the more suitable single site plans are, and the less the level of clustering the more suitable route and area plans become. Nicholson (15, 16) has developed a number of indices that can be used to assess the level of accident clustering.

In choosing an appropriate identification methodology the following issues require consideration. (Deacon)

a) The availability of financial and personnel resources.
b) The type and availability of non-accident data such as traffic volumes, road geometric etc.
c) The efficiency of the identification method. An efficient method maximises the probability of identifying a truly hazardous location as hazardous, and minimises the probability of identifying a non-hazardous site as hazardous.
d) The method should identify sites that are most correctable using engineering countermeasures and which will yield maximum benefits per Rand invested.
e) The method should be compatible with existing administrative and information systems.

The known methods for identifying hazardous locations can be classified onto the following three broad categories: a) Classical Methods, b) Bayesian methods and c) Neural network and fuzzy logic methods.

a)   Classical Methods

The most common classical methods to identify accident prone locations are :

* Accident Number
* Rate method

* Rate Number method
* Rate Control method

The Accident Rate method tends to favour locations with low exposures. This could lead to sites with very low levels of exposure to be incorrectly identified as accident prone locations. In order to prevent this the Accident Rate method is combined with the Accident Number method which tend to favour high volume locations. This method is called the Rate Number method.

The Rate Control method involve calculating for each site its critical accident rate. If the observed accident rate exceeds the critical rate the location is then tagged as an accident prone location.

$$R_{cr} = R_a + K\sqrt{\frac{R}{E_i}} + \frac{1}{2E_i}$$ ...[16]

R_{cr} - Critical accident rate
R_a - Average critical rate of reference group
E_I - Exposure at site i.
K - Statistical constant = 1.645 for a 95 % degree of confidence.

This equation is based on the assumption that accidents between the sites in the reference group is Poisson distributed. It is also assumed that all sites have the same level of safety and that is estimated by R_a . This assumption is only valid if the dispersion factor k (see Equation 5) is equal to zero, which would be the case if there is a large degree of similarity between the sites. If k > 0 then the accidents are distributed according to the Negative Binomial distribution and equation becomes invalid.

b)   Bayesian methods

Bayesian methods basically involves determining the *pdf's* (probability density functions) of E(*m)* and E(*m|x, E)* and then determining *P{E(m|x, E) > E(m)}.* If this probability exceed say 95 % then the site is considered hazardous.

Bayesian methods have been shown to be more efficient (). I.e. it produces less false negatives (the no. of truly hazardous sites not selected), more true positives (the no. of truly hazardous selected) and less false positives (the number of non-hazardous sites selected).

c)   Neural network and fuzzy classification techniques

This method utilise safety experts' knowledge to classify accidents according to their contributing factors and causes into the three road system components (vehicles, driver and road environment. The membership of each accident in the road environment component was used as a measure of accident correctability by road improvement.

The advantage of this method is its ability to identify accident prone locations that are most promising to be treated by road improvements, thus increasing the potential effectiveness of safety improvements.

4.   **PRIORITISATION**

Once a list of APL's have been identified the next step is to investigate and identify appropriate and cost effective solutions to the safety problems at each site on the list. The investigation of and the implementation of remedial measures obviously consume resources. To maximise the benefit from using resources it is important to focus on those locations with the highest potential for accident reduction.

A common approach is to list locations according their accident rates. According to McGuigan(10) using the accident rate produce a list of blackspots that are biased towards low accident totals and low traffic usage (exposure) at which the scope for accident reduction is severely restricted.

An important principle to remember is that accident rates, or other measures of safety, can only be compared if there exist a reasonable *expectancy of equality* between sites. The fact that a 4-legged intersection has a higher accident rate than a 3-legged intersection does not mean that the 4-legged intersection is a better candidate for remedial action. The accident rate for a 4 legged intersection is logically expected to be higher than that of a 3-legged intersection because of the higher number of conflicts at a 4-legged intersection.

If the 'blackspot' list consist of locations where the '*expectancy of equality'* principle is not valid the accident rate <u>cannot</u> be used to prioritise the sites.

What makes a site more suitable than another for remedial action is the extent to which its level of safety exceeds the average of all similar locations.

There are two ways in which this could be measured :

$$S_1 = E(m \mid x_i, E_i) - E(m) \qquad \text{...[17]}$$

$$S_2 = \frac{E(m \mid x_i, E_i)}{E(m)} \qquad \text{...[18]}$$

Another useful ranking criterion is the PAR (Potential Accident Reduction) :

$$PAR_i = S_1 * E_i \qquad \text{..[19]}$$

The *PAR* is not a direct measurement of cost effectiveness since it does not include for the cost of remedial measures. It does however provide a strong indication of the potential benefit that can be obtained from applying remedial measures.

5. EVALUATION OF REMEDIAL MEASURES

The evaluation of safety improvements is considered a vital component of safety management. The objectives are to establish the effectiveness of improvements and develop accident warrant guidelines, both of which are required to make safety decisions and to set priorities.

The evaluation of a remedial measure has two main objectives :
- Estimating the effectiveness of a treatment i.e. the degree to which safety improved.
- Estimating the efficiency of a treatment i.e. the degree to which the treatment was economically feasible.

An effective treatment is not always an efficient treatment.

Determining the effectiveness of an treatment requires first of all an estimation of what the level of safety would have been had no remedial measures been undertaken. This level of safety is then compared to the observed level of safety after treatment. The general approach is to estimate the '*what would have been*' level of safety by determining the level of safety before treatment, and then to assume that this level of safety would have continued to prevail had no treatment taken place.

The Accident Reduction Factor (ARF) is determined by the difference between the expected 'before' level of safety and the expected 'after' level of safety divided by the expected 'before' level of safety.

$$ARF = \frac{E(m_b \mid x_b, E_b) - E(m_a \mid x_a, E_a)}{E(m_b \mid x_b, E_b)} \qquad \text{...[20]}$$

5.1 METHODOLOGICAL ISSUES

i) After the implementation of an engineering measure there is often a 'settling in' period in which the traffic adjust to the 'new' environment. To eliminate the bias that this could introduce the treatment year is normally excluded from the analysis.

ii) Because locations are selected for treatment on the basis of an abnormally poor level of safety during the 'before' period, the regression-to-mean effect could cause safety effects to be significantly overestimated.

iii) To eliminate the regression-to-mean effect and to account for changes in external factors between the 'before' and 'after' periods most evaluation methods make use of a reference group of untreated sites similar to the site under investigation.

iv) The successful treatment of a site could cause the accidents at adjoining sites to increase. This is called accident migration. One explanation put forward to explain this phenomenon is that drivers compensate for the reduction in risk produced by an treatment by behaving in a more risky manner e.g. higher speed, reduced caution etc. This effect could be global or local. Global compensation

would redistribute the accidents over the whole network in a manner that is virtually undetectable, whereas local compensation simply reduce the saving at the treated site itself.

v) The effect of an engineering countermeasure/s could decay over time, which if not accounted for, could lead to an underestimation of the true safety effect. It must be decided beforehand whether the decay effect or a portion thereof should form part of the treatment effect or whether it should be excluded.

vi) Improved sites could generate or attract traffic from other sites in the network. This increase in exposure could lead to an increase in accidents, which if not accounted for could cause the true effect to be underestimated. There is a school of thought that since the increase in exposure is a direct consequence of the treatment that the increase in exposure should not be taken into account.

vii) The magnitude of the safety effect at site could be a function of the level of safety at that site before improvements were undertaken. This is an important consideration in determining the safety effect of treatments that were applied to 'safe' and 'unsafe' sites alike e.g. large scale resurfacing programmes. The safety effect of a remedial measure is best determined by only analysing locations with an abnormally poor level of safety.

viii) The effect of external influences e.g. changes in legislation, law enforcement programmes, change in speedlimit, routine maintenance, etc. could compromise the validity and reliability of the results obtained from a before and after study.

5.2 EVALUATION METHODS

The known evaluation methods can be categorised as either a) conventional, or b) Bayesian in nature.

a) Conventional

The most common conventional methodologies are the a) simple before-and-after , and the b) before-and after with comparison group methods.

It has been shown that because of the regression-to-mean effect the simple before-and-after method consistently overestimates the true safety effect, especially if periods of less that 5 years are used for the before and after periods respectively. The method becomes more reliable when periods of 5 yeas and longer is used, but can be seriously compromised by temporal changes in the operational and environmental characteristic if the site and its surroundings.

According to Al-Masaeid (1) the theoretical basis of the before-and-after with comparison group method is theoretically sound (provided that the study site is similar to the comparison groups in all respects; accident history, geometric characteristic and operational characteristics. Using sites with similar geometric and operational characteristic but dissimilar levels of safety will cause that the RTM effect to remain unaccounted for. In such cases Bayesian methods are used to determine the RTM effect, which is then combined with the results of *the 'before-and after with comparison group'* study to get a better estimate of the true effect.

This method make use of an control ratio *C* to predict what the number of accidents would have been had no remedial measures been undertaken. The control ratio is determined by the ratio of the number of 'after' accidents to the number of 'before' accidents for the control group.

$$E(m_a \mid x_a, E_a) = x_a \qquad ...[21] \qquad\qquad E(m_b \mid x_b, E_b) = C.x_b \qquad ...[22]$$

This method assumes that changes in the level of exposure as a result of the remedial action equals the change in the exposure of the control group over the same period.

b) Bayesian Methods

The Bayesian procedure involves estimating the true 'before' accident rate , $E(m_b|x_b, E_b)$, and its *pdf* as well as the true 'after' accident rate $E(m_a|x_a, E_a)$ and its *pdf*. Having knowledge on the respective pdf's of *$E(m_b|x_b, E_b)$* and *$E(m_a|x_a, E_a)$* allow for the accurate determination of $P(m_a < m_b)$

Bayesian methods are considered superior to the conventional methods for the following reasons :

i)   The regression-to-mean effect is automatically accounted for.
ii)   Less accident information (at least 2 years) is required.
iii)   The reference group is used to determine the parameters of the prior distribution, and as these are updated with site information to obtain the posterior parameters, the significance of the prior parameters decrease. This has the implication that the criteria for selecting comparison groups can be relaxed compared to the requirements of the before-and after method with comparison groups.
iv)   The output from the Bayesian methods allow for statistical inferences to be made on the estimated value of the ARF.
v)   In the complete absence of a suitable reference group accident models can be used.

6.   CONCLUSION

To be effective an accident remedial scheme requires the systematic analysis of accident data during each phase of the process – *identification*, *prioritisation* and *evaluation*. This paper presented some general guidelines on how the analysis of accident data could be undertaken to first of all measure true levels of safety and how to use this information to identify, prioritise and evaluate accident prone locations.

7. LIST OF REFERENCES

1. Al-Masaeid HR ; Performance of safety evaluation methods ; Journal of Transportation Engineering ; September/October 1997 ; Pages 364 – 369.

2. Belanger C ; Estimation of safety of four-legged unsignalized intersections ; Transportation Research Record 1467 ; Pages 23 – 29.

3. Davis GA and Koutsoukos K ; Statistical method for identifying locations of high crash risk to older drivers ; Transportation Research Record 1375 ; Pages 61 – 67.

4. Hauer E, Ng JCN and Lovell J ; Estimation of safety at signalised intersections ; Transportation Research Record 1185 ; Pages 48 – 61.

5. Hauer E ; On exposure and accident rate ; Traffic Engineering and Control ; March 1995 ; Pages 134 - 157.

6. Higle JL and Hecht MB ; A comparison of techniques for the identification of hazardous locations ; Transportation Research Record ; Pages 10 – 19.

7. Higle JL and Witkowski JM ; Bayesian identification of hazardous locations ; Transportation Research Record 1185 ; Pages 24 – 31.

8. Mahalel A ; A note on accident risk ; Transportation Research Record 1068 ; Pages 85 – 89.

9. McCoy PT, Ramanujam M, Moussavi M and Ballard JL ; Safety comparison of types of parking on urban streets in Nebraska ; Transportation Research Record 1270; Pages 28 – 41.

10. McGuigan DRD ; Non-junction accident rates and their use in 'black-spot' identification ; Traffic Engineering and Control ; February 1982 ; Pages 60 – 65.

11. Mensah A and Hauer E ; Two problems of averaging arising in the estimation of the relationship between accidents an traffic flow ; Transportation Research Record 1635 ; Pages 37 – 43.

12. Mountain L and Fawaz B ; The accuracy of estimates of expected accident frequencies obtained using an Empirical Bayes approach ; Traffic Engineering and Control ; Volume 32 Number 5; May 1991 ; Pages 246 – 251.

13. Mountain L, Maher M and Fawaz B ; Improved estimates the safety effects of accident remedial schemes ; Traffic Engineering and Control ; October 1998 : Pages 554 – 558.

14. Nembhard DA and Young M ; Parametric empirical Bayes estimates of truck accident rates ; Journal of Transportation Engineering ; July/August 1995 ; Pages 359 – 375.

15. Nicholson AJ ; Accident clustering : some simple measures ; Traffic Engineering and Control ;May 1989 ; Pages 241 – 246.

16. Nicholson AJ ; Indices of accident clustering : a re-evaluation ; Traffic Engineering and Control ; May 1995 ; Pages 291 – 292.

17. Nicholson AJ ; The estimation of accident rates and coutermeasure effectiveness ; Traffic Engineering and Control ; October 1987 ; pages 518 – 523.

18. Sayed T, de Leur P and Abdelwahab W ; A comparison of fuzzy and neural classifiers for road accident analysis ; XIIIth International Road Federation World Meeting ; Toronto ; June 1997

METHODOLOGICAL ISSUES IN THE IMPLEMENTATION OF ACCIDENT REMEDIAL SCHEMES

CJ Mollett Pr Eng
Chief Engineer: Traffic Engineering
Department of Economic Affairs , Agriculture and Tourism : Branch - Transport
25 Alfred Street, Cape Town, 8001

Dr CJ Bester
Department of Civil Engineering
University of Stellenbosch
Private Bag X1, Matieland, 7602

CURRICULUM VITAE

CJ Mollett   Pr Eng

Mr CJ Mollett currently holds the position of Chief Engineer: Traffic Engineering in the Transport Branch of the Provincial Administration of the Western Cape. He has been in the employ of the Provincial Administration: Western Cape since 1990. He has been working in the Traffic Engineering division of the Branch since 1994.

Mr Mollett completed his B.Sc Civil Engineering degree at UCT in 1989 and is currently busy with his masters degree in Traffic and Transportation Engineering at the University of Stellenbosch. Mr Mollett has completed a number of courses towards a B.Com degree at UNISA, amongst others majoring in Statistics and Quantitative Management.