Updating Outdated Predictive Accident Models

A. G. Wood[1], L. J. Mountain[1], R. D. Connors[2], M. J. Maher[2] and K. Ropkins[2]

[1] School of Engineering, University of Liverpool, UK

[2] Institute for Transport Studies, University of Leeds, UK

Corresponding author: Alan Wood, School of Engineering, University of Liverpool, Brownlow Hill, Liverpool, UK. L69 3GH. alanwood@liverpool.ac.uk

Abstract

Reliable predictive accident models (PAMs) are essential to design and maintain safe road networks however, ongoing changes in road and vehicle design coupled with road safety initiatives, mean that these models can quickly become dated. Unfortunately, because the fitting of sophisticated PAMs including a wide range of explanatory variables is not a trivial task, available models tend to be based on data collected many years ago and seem unlikely to give reliable estimates of current accidents. Large, expensive studies to produce new models are likely to be, at best, only a temporary solution. This paper thus seeks to develop a practical and efficient methodology to allow currently available PAMs to be updated to give unbiased estimates of accident frequencies at any point in time. Two principal issues are examined: the extent to which the temporal transferability of predictive accident models varies with model complexity; and the practicality and efficiency of two alternative updating strategies. The models used to illustrate these issues are the suites of models developed for rural dual and single carriageway roads in the UK. These are widely used in several software packages in spite of being based on data collected during the 1980s. It was found that increased model complexity by no means ensures better temporal transferability and that calibration of the models using a scale factor can be a practical alternative to fitting new models.

Keywords: Predictive Accident Model.

1. Introduction

Reliable predictive accident models (PAMs) are essential to provide and maintain safe road networks. Designers can, for example, use PAMs in the appraisal of the safety impacts of alternative design decisions, with PAMs essential to forecast accidents with and without possible interventions. Genuine high risk locations can be identified by comparing observed accidents with those predicted by PAMs given the type of site and level of traffic flow. Whereas scheme appraisal takes place prior to implementation using predicted outcomes, evaluation takes place after the event, normally using observed data. In the evaluation of safety impacts, however, simple comparisons of observed before and after accidents are known to exaggerate the effectiveness of treatments because of the regression-to-the-mean effect. This problem can be overcome using an empirical Bayes (EB) approach but its use relies on the availability of suitable PAMs (Mountain et al., 2005; Persaud and Lyon, 2007; Elvik, 2008). While the importance of PAMs is clear the quality of available models is rather less certain.

PAMs are derived by fitting regression models to data obtained from a large number of road sections or junctions. In their simplest form, such models relate expected accident frequencies to some measure of exposure (traffic flow). In more sophisticated models, additional variables describing the design features or geometry of the sites are also included. However, model fitting is by no means straightforward. High quality data are required for a large enough number of locations and accidents. The relevant data may not always be readily available and typically requires the interlinking of separately maintained databases for accidents, traffic flows and design features. There is no accepted theory to indicate how accident frequency should increase with traffic flow or, indeed, with other characteristics such as hilliness or bendiness. There is now a general recognition of the need to allow for overdispersion in accident modelling and the assumption of a negative binomial error structure is commonly used (see, for example, Maher and Summersgill, 1996). However, this is primarily for mathematical convenience, with recent research suggesting that alternative forms of error structure are now not only feasible but may also be more appropriate (Maher and Mountain, 2009, Lord and Mannering, 2010, Connors et al., 2012). Perhaps the most serious difficulty arises, however, because, over time, there will inevitably be changes in road, vehicle and driver characteristics such that the relationship between the dependent and independent variables may also tend to change. While some models include a term to allow for long-term trends in accident risk (see, for example, Walmsley and Summersgill, 1998; Walmsley et al., 1998a; Walmsley et al., 1998b) it is by no means clear that the pattern of change will remain stable over time (Elvik, 2010). The temporal transferability of PAMs is thus questionable, particularly when the elapsed time is large.

Unfortunately, because the fitting of sophisticated PAMs including a wide range of explanatory variables is not a trivial task, available models tend to be based on data collected many years ago. In the case of UK roads, for example, the Transport Research Laboratory (TRL) carried out a comprehensive series of accident studies during the 1980s and 1990s. TRL developed models for various junction and link types at various levels of detail and were, indeed, amongst the first to recognise the need to model overdispersion and to propose the use of a negative binomial error structure. However, the earliest of these studies used accident data for 1974-79 to fit models for 4-arm roundabouts (Maycock and Hall, 1984) while the most recent used accident data for 1979-92 to fit models for rural dual carriageways (Walmsley et al. 1998a). These models remain the industry standard. They are incorporated into several standard software packages and widely used in design and scheme appraisal in the UK in spite of the age of the modelled data. However, given that, in the UK, annual personal injury accidents fell by 30% between 1985 and 2009 while annual total traffic increased by 61% (DfT 2010a, DfT 2010b) it seems unlikely that PAMs derived using data from some 20 to 30 years ago could provide accurate estimates of current accidents. Fitting new models for the whole range of link and junction types would be both time-consuming and expensive because of the size and level of detail of the database required. In any case, new models would only provide a temporary solution since the new models themselves would soon become outdated. A more sensible long-term solution would appear to be to develop an updating strategy so that updated versions of the existing models can be used, not only now but also in the future. That is the objective of the present research study, of which this paper is a part.

This research study has two principal objectives. Firstly to establish the extent to which the temporal transferability of predictive accident models varies with model complexity and, in particular, the extent to which the inclusion of design variables in more complex models increases temporal stability. Secondly to develop an approach that will allow currently available predictive accident models to be readily and reliably updated to any point in time. Although new predictive accident models have been developed in the last few years (for example Hashim and Bird, 2005) the updating strategy is applied to the TRL models as they remain the industry standard. To allow the objectives of the study to be achieved a database has been compiled containing accident data, flow data and geometric design parameters for six site categories; modern rural single carriageway A-roads, modern rural dual carriageway A-roads, urban single carriageways, urban 3-arm signalised junctions, urban 4-arm signalised junctions and 4-arm roundabouts. The aim was to include a range of link and junction types, and a range of model ages.

1.1 Aims of this paper

In an earlier paper, the authors addressed the principal methodological issues that arise in seeking practical and efficient ways to update PAMs (Connors et al., 2012). These issues were illustrated by application to a basic model for rural single carriageway roads, and include: the choice of distributional assumption for overdispersion; the choice of goodness of fit measures; questions of independence between observations at the same site in different years, and between links on the same scheme; the estimation of trends in the models; the uncertainty of predictions; the most efficient and convenient ways to fit the required models, given the considerable advances that have been seen in statistical computing software in recent years. The focus of this paper is to apply this methodology to establish the extent to which the temporal transferability of currently available PAMs varies with model complexity and to assess how best to update existing models.

Two site categories are assessed, namely, modern dual and single carriageway A-roads in rural areas which are amongst the most recent of the TRL models (Walmsley and Summersgill, 1998; Walmsley et al., 1998a; Walmsley et al., 1998b). The accident data used to fit these models were for the period 1979-90 for single carriageways and for 1979-92 for dual carriageways, although accident data were not available in every year for all schemes. A-roads are principal roads, designed to carry large volumes of long-distance traffic and here ‘modern’ refers to roads that meet post-1960 design standards. A rural road is defined here as a road which is not in a built up area. The highest quality dual carriageways are near motorway standard, while the lowest quality roads of both types can have sharper corners, steeper gradients and larger numbers of intersections without deceleration lanes. Initially, the goodness-of-fit of the existing suite of TRL models to current data was determined. The suite of TRL models for these roads covers a range of levels of detail: in the basic model, total accidents are simply related to a measure of traffic flow; in the most detailed models, accidents are disaggregated by type and for each type the accident frequency is related to traffic flow and a range of geometric parameters. Since the more detailed models can take account of the effects of improvements in highway design on the frequency of particular types of accidents, it seemed likely that they could offer better temporal transferability than the basic models, albeit at the expense of more input data. Two main updating strategies were trialled. The first re-estimated the TRL parameter values (while keeping the model format unchanged), thus both adjusting for trend and allowing the relationship between the dependent and independent variables to vary. The second used the TRL models but with a scaling factor to adjust for the trend between the time of the TRL study and the current study.

2. Database

The database compiled for this analysis contained 561 minor links distributed amongst 54 schemes for dual carriageways, and 341 minor links distributed amongst 73 schemes for single carriageways. A *scheme* refers to the largest feature studied, and is a section of road with similar flow characteristics, normally between two *major junctions* (defined as any junction where traffic on the scheme has to give way). Typically any given road number (e.g.

A14) appeared only once in the database. Within a scheme there are typically a number of *minor junctions* (defined as any other junction properly marked with a give way or stop line and a centre line on at least one junction arm) and *minor links* (the section of road between any two junctions commencing 20 m from the extended kerbline). These definitions are illustrated in Figure 1. On average, a dual carriageway scheme included 10.6 minor links (range: 2 to 45) and 8.0 minor junctions (range: 0 to 43); single carriageway schemes included an average of 4.7 minor links (range: 1 to 18) and 3.3 minor junctions (range: 0 to 17).

(Insert Figure 1 about here – caption is at the end of the manuscript)

Most of the schemes were analysed across a five year period (2005-2009): the exceptions were two single carriageway schemes which had undergone major changes in 2008 and 2009 and for these only data prior to the changes were used. In our study, accident data were obtained from the UK national STATS19 database or local authorities, and annual flow data were obtained from the UK Department for Transport (DfT) or local authorities. Table 1 compares the key features of our database with that used by TRL.

(Insert Table 1 about here)

It can be seen in Table 1 that the current study, included some 86 % of the road length used by TRL in the case of dual carriageways, and 60 % in the case of single carriageways. The total number of accidents included in the current study were some 44% of those used in the TRL study in the case of dual carriageways, and 71% in the case of single carriageways. The mean annual average daily traffic (AADT) on dual carriageways has more than doubled between the two studies, whilst on single carriageways it increased by a third (Table 1). It is, however, worth noting that the accident rates (annual accidents per 100 million veh-km) were not significantly different in either of the studies with an annual rate of some 12 accidents per 100 million veh-km on dual carriageways and some 19 accidents per 100 million veh-km on single carriageways. The breakdown of accidents between minor links and minor junctions on dual carriageways was broadly similar in both studies but, on single carriageways, a higher proportion of accidents took place at minor junctions in the TRL study than in the present study. A possible explanation was the higher minor junction density in the TRL study: with 2.7 (0.1) minor junctions per km in the TRL study compared with 1.0 (0.1) minor junctions per km in the present study. (The errors quoted in brackets here and elsewhere in this paper are the standard error of the sample mean or estimate.) The lower minor junction density in the present study may reflect design improvements over the intervening years.

In order to test the full range of TRL models it was necessary to disaggregate the accidents which occurred on minor links by type. Tables 2 and 3 compare the number and proportion of each type of accident in the TRL study and our study on dual and single carriageways respectively. The percentage of total accidents column does not add up to 100% in the case of single carriageways as not all link accidents could be assigned to a category. It will be noted that the proportion of some types of accident has changed between the two studies, for example on dual carriageways (Table 2), the proportions of pedestrian accidents (type 1) and the proportion of accidents involving two or more vehicle of types 11 and 12 more than halved between the two studies. On single carriageways (Table 3) the proportion of most types of accident were similar. However, it is notable that the proportions of two types of accident have shown a statistically significant increase (P=0.05), both more than doubling between the two studies. These are accidents at accesses (type 2) and shunts (type 9).

(Insert Table 2 and 3 about here)

TRL obtained details of a range of design features and geometry for each minor link. In the present study only those parameters which TRL found to be significant in the models at the 5 % level were recorded in the database. For dual carriageways these parameters were:

- Presence or absence of hardstrip on the minor link or scheme.
- Presence or absence of kerb on the minor link.
- Proportion of the minor link with a kerb.
- Presence or absence of a safety fence on the minor link.
- Proportion of the scheme with a safety fence.
- Proportion of the minor link with a continuous obstruction on the offside other than a safety fence.
- Quality of scheme, classified as a two level factor: *higher* quality for schemes of near motorway standard with mostly grade separated junctions and few roundabouts and *lower* quality for schemes with smaller and more frequent roundabouts and more minor junctions.
- Quality of link, classified as a four level factor: *best* quality for roads of near motorway standard with mostly grade separated junctions and no roundabouts; *good* quality for roads of near motorway standard with some large roundabouts; *lower* quality for roads with T-junctions or crossroads with short deceleration lanes; *urban bypass* for modern roads in semi-urban areas.
- Bendiness (sum of the angles turned, divided by the length (degrees km $^{-1}$)).
- Hilliness (sum of the height gain and loss, divided by the length (m km^{-1})).
- Mean density of offside accesses (accesses km^{-1}).

For single carriageways these parameters were:

- Width of scheme, classified as a two level factor according to whether the road was closer to 7.3 m wide (i.e. a normal 2-lane single carriageway) or 10 m wide (i.e. a wide 2-lane single carriageway).
- Presence or absence of hardstrip on the minor link or scheme.
- Bendiness (sum of the angles turned, divided by the length (degrees km^{-1})).
- Hilliness (sum of the height gain & loss, divided by the length (m km^{-1})).
- Net gradient on the minor link (single carriageways only, height difference between the end points of a link, divided by the link length. (%)).
- Mean density of accesses on the scheme (accesses km^{-1}).

TRL used a specially adapted van to make measurements of the geometric parameters. In the present study hilliness was obtained from the Digital Elevation Maps (DEM) in Google Earth and the bendiness was obtained from Ordnance Survey maps. Other geometric parameters were established using Google Street View. If a factor (such as presence or absence of hardstrip) changed over a minor link or scheme then it was allocated to whichever category was appropriate to the greatest length of road.

3. Methods

3.1 TRL models

To assess how well, or otherwise, the TRL suite of models fit current data, the models were used to predict the number of accidents in our database. TRL developed models for rural roads at three units of *size* and at four *levels of detail* and these are summarised in Table 4. Models were not fitted for all combinations of size and level of detail. So, for example, Level 1A models (models of total accidents without geometric parameters) were available for all units of size. If the addition of geometric parameters to the model improved the fit then TRL also presented a Level 1B model (models of total accidents with geometric parameters). Level 2 and 3 models (models for particular types of accident, without and with geometric parameters respectively) were only available for accidents on individual minor links (size 1). Size 1 models give estimates of accidents on *individual* minor links while size 2 link models give estimates of the total accidents on *all* minor links within a scheme. Total scheme accidents can be estimated directly using a size 3 model or by summing separate estimates of link accidents (size 1 or size 2 link) and junction accidents (size 2 junction).

(Insert Table 4 about here)

The first step was to establish to what extent the existing TRL models represent current conditions. The models for links and schemes fitted by TRL are of the general form;

$$\mu_{it} = k \cdot L_i^{\lambda} \cdot Q_{it}^{\alpha} \cdot \exp(Junction_i + Geom_i) \cdot \exp(Trend) \qquad [1]$$

where μ_{it} is the expected number of accidents at site i in year t, L_i is length (in km), Q_{it} is the two-way AADT (in thousand veh/day), $Junction_i$ describes the details of the minor junctions, $Geom_i$ are various relevant geometric design features, *Trend* describes the year-on-year trend in accident risk and k, α and λ are constants. In many of the models λ is unity (i.e. predicted accidents are proportional to link length). Although accidents at major junctions are not included in these models, the numbers of both major and minor junctions were found to be significant in many of the models, effectively accounting for any *spillover* effects from the junctions. (The *spillover* effect arises because, in the UK, junction accidents are defined as accidents occurring within 20m of the extended kerb line of a junction (see Figure 1). In practice accidents may be assigned to links even though they arise because of the presence of a junction more than 20m from the accident.)

TRL presented a total of 70 models relating to rural A-roads. These are fully documented in Walmsley and Summersgill (1998), Walmsley et al. (1998a) and Walmsley et al. (1998b). Space prohibits giving the detail of each of these models however, as an example, the dual carriageway level 1B, size 2 model (the model for total accidents on all of the minor links of a dual carriageway scheme with geometric features) is:

$$\mu_{it} = 0.0393 \cdot Q_{it}^{0.913} \cdot L_{L_i} \cdot \exp(Junction_i + Geom_i) \cdot \exp(Trend) \qquad [2]$$

In this model $Junction_i$ and $Geom_i$ are defined as:

$$Junction_i = 0.120 \cdot \frac{N_{m_i}}{L_{S_i}} + 0.210 \cdot \frac{N_{n_i}}{L_{S_i}} \quad [3]$$

$$Geom_i = -0.255 \cdot g_{HS2_i} + 0.082 \cdot g_{HS1_i} - 0.231 \cdot g_{Q_i} + 0.025 \cdot g_{B_i}$$

$$+ 0.089 \cdot g_{H_i} + 0.071 \cdot g_{NX_i} - 0.030 \cdot g_{SF_i} \quad [4]$$

where N_m is the number of major junctions on the scheme,
N_n is the number of minor junctions on the scheme,
L_L is the total length of the links on the scheme (excluding the length of the minor junctions),
L_S is the total length of the scheme (including the length of the minor junctions),
g_{HS2} is a factor set to unity if hardstrips are present on both the nearside and the offside and set to zero otherwise;
g_{HS1} is a factor set to unity if hardstrips are present on either the nearside or the offside (but not both) and set to zero otherwise;
g_Q is a two level factor describing the quality of the scheme set to unity for lower quality roads and zero for higher quality roads (as defined in Section 2);
g_B is the bendiness (degrees km^{-1});
g_h is the hilliness (m km^{-1});
g_{NX} is the offside access density (accesses km^{-1}); and
g_{SF} is a factor set to unity if a safety fence is present and set to zero otherwise.

In this example $Junction_i$ did not separate minor junctions by type (i.e. N_n is simply a count of the total number of minor junctions within the scheme). In the various TRL models up to four categories of minor junction were included, with separate counts of minor junctions in each category. The categories were:

- Dual carriageways with four types of minor junction: grade-separated, N_1; other junctions with no access across the central reservation to the opposite carriageway, N_2; staggered junctions with access across the central reservation to the opposite carriageway, N_3; other junctions with access across the central reservation to the opposite carriageway, N_4.

- Dual carriageways with two types of minor junction: junctions with no access across the central reservation to the opposite carriageway, N_{1+2}; junctions with access across the central reservation to the opposite carriageway, N_{3+4}.

- Single carriageways with two types of minor junction: 3-arm junctions, N_1; 4-arm junctions, N_2.

The term $\exp(Trend)$ allows for trend, where $Trend = \beta \cdot t$. In these models β is the annual rate of change in accident risk and t is the number of years since 1990 (the base year for the TRL models). TRL estimated three values of β depending on carriageway type and the road element modelled: -2.25x10^{-2} for size 1 models of dual carriageways, 0 for size 1 models of single carriageways and -2.0x10^{-2} for all other models. (This suggests a reduction in accident risk over the 17 years from 1990 to the midpoint of our study period (2007) of 100(1-$e^{17\beta}$)%

or 32% for size 1 models of dual carriageways, no change for size 1 models of single carriageways and 29% for all other models.) The data used to estimate these trend values were, as previously noted, for the years 1979-90 and 1979-92 for single and dual carriageways respectively.

The goodness of fit of each of these models (with the exception of some level 3 models which relied on geometric factors that could not be readily measured such as maximum visibility on a link) was established. In each case the model was used to predict accidents in each year (2005 to 2009, with t = 15 to 19). These predictions were summed and compared to the total number of accidents observed during this five year period. In the case of the level 2 and 3 models (Table 4) which predict accidents by type, predictions were summed and compared to the total number of observed accidents. If a level 3 model was not available for a particular type of accident then a level 2 model was substituted.

3.2 New parameter estimates

It is not unreasonable to expect that the value of both parameter estimates and the trend may have changed over time. Clearly the ageing of PAMs can arise from a variety of sources, including: improvements in highway design and traffic control; changes in driver training, culture and attitudes; changes in legislation and enforcement strategies; improvements in vehicle design; changes in vehicle fleet mix and so on. Whereas improvements in highway design features might be expected to primarily impact on risk (k), many recent road safety interventions have targeted driver behaviour: such measures might reasonably be expected to impact on the way in which flow affects accidents (the power of flow term (α)). The form of the relationship between accidents and other explanatory variables may thus change over time.

The next step was therefore to obtain new parameter estimates by re-fitting the TRL models using our database. These models had the same independent variables, factors and form as those given by TRL. Connors et al. (2012) tested assumptions about the distributional form of the model and recommended a method in which the model was considered in two linked parts. In the first part the aggregated data across all years (2005-2009) was used to estimate the model parameters using a Poisson distribution for the accident distribution at each site and a Variable Shaped Gamma (VS-G) distribution for the overdispersion (i.e. the between site variation). The second part was to disaggregate the data by year and model the distribution of the accidents at each site across the five years 2005-2009 using a multinomial distribution to estimate the trend within the data. This will be referred to as the *recent trend* and denoted by β^l to distinguish it from the *long term trend* (i.e. from the base year of the TRL models (1990) to the mid-point of our study period (2007)). Estimating the parameters involved an iterative approach in which the new parameter estimates were used to update the recent trend and then this trend was used to update the new parameter estimates and so on. Iterations continued until none of the new parameter estimates varied significantly (at the 5 % level) from one iteration to the next. This gave a model of the form:

$$\mu_{it} = k \cdot L_i^{\lambda} \cdot Q_{it}^{\alpha} \cdot \exp(Junction_i + Geom_i) \cdot \exp(\beta^l \cdot t^l) \qquad [5]$$

As the recent trend was estimated using data from 2005 to 2009, t^l is the number of years after 2007 (for 2005 to 2009 $t^l = -2 \ldots +2$). The revised estimate of the long term trend (1990 to 2007) is effectively incorporated in the estimate of k in the re-fitted model.

3.3 Calibration of TRL models

An alternative approach to model updating is calibration of the TRL models using a *scale factor*. In this approach it is assumed that the relationship between accidents and the explanatory variables has not changed over time and that the discrepancies between the model predictions and the data are simply due to changes in the trend in risk over time. The TRL model was used to predict accidents in each year of the present study (2005 to 2009) with the *Trend* term set to zero. These predictions were summed (over all years and all sites) and a scale factor, *sf*, calculated as the ratio of total observed to predicted accidents during this five year period i.e. $sf = \sum y_j \big/ \sum \mu_j$. The reduction in accident risk over the 17 years from 1990 (the base year of the TRL models) to the midpoint of our study period (2007) can thus be estimated 100(1-*sf*)%.

There were several alternative possibilities for the optimal scaling factor depending on the choice of measure of goodness-of-fit. Possibilities include: minimising the absolute mean error; minimising root mean square error (RMSE); minimising the mean absolute deviation (MAD); or maximising the log likelihood. There is no unique, best way of determining the scaling factor as all of the goodness-of-fit criteria are sensible and desirable. These issues are discussed in more detail in Connors et al. (2012). We chose to estimate the scaling factor as $sf = \sum y_j \big/ \sum \mu_j$ which minimises the absolute value of the mean error as this is simple, intuitive and minimises bias. In this study the choice of scaling factor was shown not to affect the choice of best-fitting model.

3.4 Goodness-of-fit

The same considerations apply when we wish to compare the performance of alternative models: which is the best fitting model will be influenced by the criterion used to measure goodness of fit. Three goodness-of-fit statistics were selected. These were: the Akaike information criterion (AIC) which allows an assessment of whether the model fit is improved by the inclusion of additional variables; the root mean square error (RMSE) which measures the precision of the model; the mean error (ME) which measures bias. As the purpose of the study was to establish how well the models represented current conditions, AIC was chosen as the principle criterion to select the best fitting model. However, as users of these models wish to know the uncertainty in the predictions, the RMSE and the ME were also stated.

3.5 Comparison with simple rate models

The TRL models are relatively sophisticated models in that, even in their simplest form, they do not assume that accidents are proportional to flow. The simplest PAMs are, however, *accident rate models* in which accidents are assumed to be proportional to flow and length. It was decided to compare the fit of the TRL models with simple accident rate models. In the UK, while the TRL models are used in standard software packages such as ARCADY, PICADY and OSCADY, simple accident rate models are used in Cost Benefit Analysis (COBA) software (DfT, 2006). The COBA models are designed to be used to value the accident savings associated with alternative scheme proposals. Although junction and link accidents can be modelled separately in COBA, the junction models require traffic counts on all arms of the junctions, and thus could not be applied to the minor junctions in our data since only the major road flows were available. Sufficient data were available, however, to allow the COBA models to be used to predict accidents on all links of a scheme and accidents on all links and junctions of a scheme. The COBA models have a base year of 2000 but incorporate a correction for trend in accident risk in the form of an accident rate change coefficient raised to a power equal to the number of years after 2000. For rural single and dual carriageways the annual change coefficient is 0.973 suggesting a reduction in accident risk over the 7 years from 2000 to the midpoint of our study period (2007) of 17% (i.e. 100(1-0.973^7)%) on all schemes.

4. Results and discussion

4.1 Rural dual carriageways

Initially the 5–year accidents on each of the 54 dual carriageway schemes were estimated using each size 2 and 3 model with the methods outlined in section 3 and the results are shown in Table 5. In Table 5, the column headed *TRL* contains the results using the unadjusted TRL models, the column headed *New Parameter* contains the results for the re-fitted models and the column headed *Scale Factor* contains the results for the calibrated models. The models have been numbered (D1 to D15) for ease of reference.

(Insert Table 5 about here.)

4.1.1 Trend in accident risk

Table 5 shows the various estimates of trend in accident risk: the original TRL estimate of trend, the recent trend (for 2005 to 2009) estimated by re-fitting the models, and the long term trend (1990 to 2007) estimated by the scale factor. The TRL estimate of long term trend for all of these models was exp(-0.02t) giving an estimated total reduction in accident risk over the 17 years 1990-2007 of 29% (i.e. for t=17, exp(-0.02t) = 0.71). From Table 5 it is clear that, for junctions, the TRL estimate of the reduction in accident risk over these 17 years is reasonably close to that estimated by the scale factor, particularly for model D8 (the best fitting junction model) which has a scale factor of 0.70. However, for the link models, the scale factors are larger suggesting that the TRL models overestimate the actual long term trend in risk on links. Comparison of the estimates of recent trend (2005-2009) obtained by re-fitting the models with the TRL estimates suggest larger trends in risk in recent years and the differences are statistically significant in the case of the size 2 link models and the size 3 models. Clearly the estimates of scale factor represent the average change in risk over the 17-year period. While the data suggest that, on average, the trends in risk at junctions were similar to those reported by TRL, there may have been periods of higher or lower trends in the intervening years.. For links the estimates of trend are much more variable. Given the potential for such variability it seems preferable to avoid applying a trend estimated over a particular time period to a different time period.

4.1.2 The goodness-of-fit of the models

The best fitting models are highlighted in Table 5. Based on the AIC values, the best fitting unadjusted TRL models are D3, D8 and D15 for size 2 link, size 2 junction and size 3 respectively. The best fitting re-fitted models were D4, D8 and D13 while the best fitting calibrated models were D6, D8 and D15. With the exception of the size 2 junction models (for which there are no models including geometric design features) and the unadjusted size 2 link model (D3), the best fitting models include geometric design features suggesting that increased model complexity improves the fit.

Comparison of the AIC values in Table 5 shows that the best fit is achieved using the re-fitted models. Not unexpectedly, the least good fits are generally for the unadjusted TRL models although, for the junction models, the calibrated models give similar AIC values to the unadjusted models and indeed for models D7 and D8 the AIC is actually lower for the unadjusted TRL model. This is perhaps not surprising given that the TRL estimates of long term trend in accident risk at these junctions was similar to the scale factor estimates whereas, for the links and whole schemes, the TRL models tend to overestimate the reductions due to trend and thus to underestimate accidents. This bias is reflected in the overall mean errors (ME): for the link and whole scheme models the unadjusted TRL models result in large negative values of ME. The values of ME for the calibrated models are zero: recall that the scale factor used here minimises the mean error. The values of ME for the re-fitted models for links and junctions are generally close to zero suggesting little bias in the estimates although for the size 3 models the mean errors are rather larger. This is probably because the size 3 models effectively use an average value of trend for both links and junctions whereas the size 2 models suggest that this may not be appropriate since the trends differ. In terms of the RMSE, the re-fitted models tend to perform slightly better than the calibrated models but the differences are generally small. Both the re-fitted models and the calibrated models generally give smaller RMSE than the unadjusted TRL models: the exception is for the calibrated size 3 models where the RMSEs are slightly larger using the scale factor than the unadjusted model.

4.1.3 Parameter estimates in the re-fitted models

The re-fitted models take exactly the same form as the equivalent unadjusted model although some variables/factors were not significant in the re-fitted model, possibly as a consequence of improved design standards. For example, bendiness was not significant and the mean value of bendiness was smaller in the present study (25.0 (2.0) degrees km^{-1}) than in the TRL study (35.1 degrees km^{-1}) suggesting improvements in horizontal alignment since the 1980s. Nevertheless non-significant parameters were retained as the purpose of the present study was to update existing models rather than to develop new models.

Comparison of the parameter estimates for the re-fitted models with the unadjusted TRL models suggests that relatively few of these have changed significantly. Most notably, there was no significant difference in the estimates of the power of flow term for the link and scheme models. In the best fitting junction model (D8), the power of flow was 1.20 (standard error not given) in the unadjusted TRL model as compared with 0.82(0.18) in the re-fitted model suggesting that the influence of increasing traffic flows on accident frequencies at junctions may be less now than in the 1980s: doubling the flow would increase expected junction accidents by a factor of 2.29 using the unadjusted TRL model and by 1.77 using the re-fitted model.

4.1.4 Modelling accidents on individual minor links

The relatively poor performance of the calibrated size 3 models suggests that it may be preferable to estimate accidents on a scheme by separately estimating accidents on links and junctions and summing the resulting estimates. This has the advantage of allowing separate estimates of trends in accident risk to be made for links and junctions. Recall that accidents on links can be estimated in two ways: modelling accidents on all links on a scheme (size 2 link models) or modelling accidents on individual minor links (size 1 models). These size 1 models are available at all four levels of detail (Table 4). All link accidents on a scheme can thus either be estimated directly using a size 2 model or by summing the estimates for individual minor links obtained using a size 1 model. Before accidents on schemes are estimated it was necessary to establish which of these approaches gives the best estimates.

To establish the best fitting size 1 model, estimates of accidents on individual links in the current data set were made using the unadjusted TRL models, re-fitted models and calibrated models and the goodness-of-fit of these models are summarised in Table 6. (There were insufficient data to fit new parameter estimates for the accident type models (levels 2 and 3).)

Consideration of the goodness-of-fit of the unadjusted TRL models suggests that increasing model complexity does give some improvement in the temporal transferability of the models: the best fitting unadjusted model (D17) includes geometric design features although modelling accidents by type (D18 and D19) does not improve the model fit. Comparison of the AIC values suggest that refitting the models gives a better fit than calibration and, not unexpectedly, the least good fits are for the unadjusted TRL models. The best fitting model for accidents on individual links is the refitted level 1B model (D17).

(Insert Table 6 about here.)

The next step was to compare the estimates of accidents on all links on the schemes obtained by summing the estimates obtained using the size 1 models with the estimates obtained using the size 2 models. In the case of the scale factor approach, the best fitting calibrated size 1 model for individual minor links (D16, Table 6) was applied to each minor link and the predictions summed and compared to the observed value. This gave a RMSE of 17.0 and a ME of 0.0. The best fitting calibrated size 2 model (D6, Table 5) gave a better fit to the data with a RMSE of 14.0 and a ME of 0.0. A similar procedure was tried for the re-fitted model. The best fitting re-fitted size 1 model for individual minor links (D17, Table 6) gave a RMSE of 13.9 and a ME of +0.1. The best fitting re-fitted size 2 model (D4, Table 5) gave a better fit to the data with a RMSE of 12.8 and a ME of +0.4.

Having determined that the use of size 2 links models are preferable to size 1 models, estimates of accidents on whole schemes were obtained by summing the estimates obtained using the best fitting size 2 models for links and junctions and these were compared with the estimates obtained using size 3 models. In the case of the calibrated models, the best fit was achieved using size 2 models: the best fitting size 2 models (D6 and D8, Table 5) gave estimates of scheme accidents with a RMSE of 21.7 and a ME of 0.0 which compares favourably with the best fitting calibrated size 3 model (D15, Table 5) which had a RMSE of 30.7 and a ME of 0.0. A similar procedure was used with the re-fitted models but for these a marginally better fit was achieved using the size 3 model. The best fitting re-fitted size 2 models for links and junctions (D4 and D8, Table 5) gave estimates of scheme accidents with a RMSE of 19.8 and the ME was +2.0 while the best fitting re-fitted size 3 model (D13, Table 5) had a RMSE of 18.5 and a ME of +1.9.

Thus, while the comparison of size 1 and 2 models indicate that it is preferable to estimate accidents using a model based on the largest spatial scale (size 2), the comparison of the size 2 and 3 models is less clear cut. With the re-fitted models, accidents on a scheme are marginally better estimated using a model based on the largest spatial scale (size 3 model) but with the calibrated models, the best estimates are obtained using size 2 models. This is a likely consequence of the different long-term trends for accident risk on links and junctions. When a scale factor is applied to a size 3 model the total number of accidents predicted by the model matches the observations, however the proportions of accidents predicted to occur on links or at junctions does not match the observations. With a re-fitted model a new estimate is made of the proportion of accidents occurring on links or at junctions. Overall these results indicate that accidents are best estimated using a model for the largest spatial size provided that the trend in accident risk is the same in all of the component parts.

Predictions and observations for the best fitting models are shown in Figure 2 (for the calibrated size 2 models D6 and D8) and Figure 3 (for refitted size 3 model D13). In each of these figures the top left panel shows the observed 5-year accident frequency plotted against the predicted accident frequency. The residuals (the difference between the predicted and observed accident frequencies) versus the predicted values are shown in the top right panel. These panels show a strong relationship between predicted and observed accidents, albeit with a certain amount of scatter. The large number of sites with small numbers of accidents meant that numerous points were plotted in a small area. To enhance the clarity of these plots the data were "binned": the predicted values were sorted in ascending order of the predicted value and grouped into bins of nine sites (chosen to give 6 bins of equal size for the 54 schemes). The lower plots in Figures 2 and 3 are similar to the upper plots but for the binned data. In the lower left-hand plot, the mean of the observed values in each bin is plotted against the mean of the predictions, and in the lower right-hand plot the mean of the residuals is plotted against the mean of the predictions. By aggregating the data, a lot of the variability is removed and this permits a more informative view of the data and any systematic variations (Connors et al., 2012). These lower panels show that there are no systematic variations across the data.

(Insert Figures 2 and 3 about here – captions are at the end of the manuscript.)

4.1.5 Data requirements for different methods

In seeking a practical updating strategy, it is important to consider the data input requirements of alternative strategies. At the extremes are the use of an unadjusted outdated model and the fitting of a new model. When an outdated model is used to predict expected accidents at a site, the only input data required are the traffic flows and relevant design features for the site of interest. Fitting a reliable new model will require an extensive database and specialist statistical expertise. Updating outdated models, both by refitting the model or calibration using a scale factor, requires traffic flows and relevant design features for a group of suitable sites. Clearly if these data are needed for a large number of sites then the benefits of updating existing models rather than simply fitting a new model could be lost. In theory it is only necessary to have more observations than parameters to be estimated. For calibration of the model using a scale factor there is only one parameter to be estimated whereas with the re-fitted models there may be several, depending on the model used: in this study there are typically around a dozen parameters. In re-fitting the models we also require sites with variation across the data set (for example, factors taking values of both 0/1 showing the presence/absence of particular design features). This is a particular problem when, as here, the models contain several categorical variables since, for each significant design feature, we need sites with the full range of combinations of other features. For example, for roads with and without hardstrips, it is necessary to include both lower and higher quality roads and, for each of these combinations, roads with and without safety fences. In principal then, calibrating models can be achieved with less data than is needed for re-fitting, although better estimates are likely to be achieved with larger data sets: uncertainty measures like the standard error of the parameter estimates and goodness-of-fit statistics such as RMSE or ME can be expected to be inversely proportional to the square root of the sample size. In order to illustrate how the accuracy of the estimates varies with the number of sites used, model D6 was re-fitted and calibrated using a range of sample sizes of randomly selected schemes. The key parameters and their standard errors, together with goodness-of-fit statistics are shown in Table 7. These data confirm that the models can only be updated using a scale factor when the number of available sites with appropriate data becomes small (in this case less than about 40 sites). In the case of the calibrated model, the estimates of the scale factor did not vary significantly as progressively smaller samples were used although, as expected, the standard errors increased. In the case of the re-fitted model, when less than 40 sites were used, new parameter estimates could not be made because there was an inadequate range of combinations of the categorical variables.

(Insert Table 7 about here.)

4.1.6 Summary

The fit of the outdated models can be improved either by re-fitting the models or by using a scale factor. Although the re-fitted models suggest that the value of the parameter estimates have changed somewhat over time, the main issue affecting the model fit appears to be the variation in the trend in accident risk over time. Outdated models which correctly estimate long term trend (e.g. size 2 junction model D8, Table 5) can give good estimates of current accidents but, when the trend is incorrectly estimated (e.g. size 2 link model D3, Table 5), the bias can be large. Because changes in accident risk from year to year can vary unpredictably, estimates of trend based on data for one time period can be no means guaranteed to give good estimates for another time period. Thus PAMs that incorporate estimates of trend do not necessarily give good estimates of current accidents. The use of current data to either re-fit the model or estimate a scale factor give better estimates (lower RMSE) and less biased estimates (lower ME) than the unadjusted models essentially because the trend in accident risk between the modelled period and the current period is estimated from the data rather than predicted using data from an earlier time period. The tendency for the re-fitted models to give a somewhat better fit than the calibrated models suggests that the form of the relationship between accidents and other explanatory variables has changed somewhat over time but this is much less important than the variations in the trend in accident risk over time.

The best fitting re-fitted models (D4, D8 and D13) are shown in the appendix. Since the re-fitting process allows for changes in the proportions of accidents on links and at junctions, the best fitting re-fitted model is a whole scheme (size 3) model. With the calibrated models these proportions are fixed in the size 3 model and, because trends on links and at junctions differ, better estimates are obtained by summing the estimates from separate calibrated models for links and junctions (size 2). For dual carriageways, increased model complexity improves the temporal transferability the models somewhat, with the inclusion of design features improving the fit of the link and whole scheme models for both the re-fitted and calibrated models. However, modelling accidents by type does not improve the fit of the models. Although the re-fitted model gives slightly better estimates (lower RMSE) than the calibrated model, the calibrated model minimises bias (ME=0), can be applied using data for fewer sites and does not rely on specialist statistical knowledge. On balance the calibration of models using a scale factor appears to be the most practical approach to model updating.

4.2 Rural single carriageways

The models for rural single carriageways were tested using the same procedures as outlined in Section 4.1 for the rural dual carriageways. The trends for each model are shown in Table 8 along with the goodness-of-fit estimated the same three ways as for dual carriageways. As in the case of dual carriageways, the fit of the outdated models was improved by both re-fitting and by calibration: the re-fitted models provided a better fit to the data than the calibrated models, but did not offer a substantial improvement in the model predictions compared to the calibrated models. The best fitting re-fitted models were S4, S5 and S10 and the best fitting calibrated models were S2, S5 and S7 for links, junctions and schemes respectively.

(Insert Table 8 about here.)

As with dual carriageways, in addition to the models for accidents on all minor links or all minor junctions on a scheme (size 2 models), TRL also fitted models for individual minor links (size 1 models). As in the dual carriageways case, the individual minor links models were tried but again the best estimates were obtained using the best fitting model for all minor links (size 2 models) rather than summing estimates for individual minor links. As for dual carriageways, estimates of accidents on whole schemes were obtained by summing the estimates from the best fitting size 2 link and junction models and these were compared with those obtained using the best fitting size 3 model. The results were similar to those for the dual carriageways, with the slightly better estimates for the calibrated models obtained by summing the estimates from the size 2 models; for the re-fitted models there was nothing to choose between the size 2 or size 3 models. The estimates from the calibrated size 2 models (Table 8, S2 and S5) gave a RMSE of 8.6 and a ME of -0.2 while the best fitting calibrated size 3 model (Table 8, S7) gave a RMSE of 9.2 and a ME of -0.1 . With the re-fitted models, the summed predictions from the size 2 models (Table 8, S4 and S5) gave a RMSE of 7.8 and a ME of -0.2 while the best fitting size 3 model (Table 8, S10) gave a RMSE of 8.0 and ME of +0.1 . The best fitting re-fitted models (S4, S5 and S10) are given in the appendix.

It is perhaps surprising that the best fitting models for single carriageways did not always include geometric features. This result does not mean that the geometric parameters have no effect upon the number of accidents, but it may be that improvements in design since the 1980s mean that the most dangerous geometric features no longer arise.

4.3 Comparison with COBA models

The simple rate models for rural dual and single carriageway links and schemes used in COBA were tested using the same procedures as outlined in Section 4.1. The trends and the goodness-of-fit for each model are shown in Table 9. As in the case of the TRL models, the re-fitted models gave the best fit to the data (lowest AIC). The calibrated models gave a better fit than the unadjusted models, although for links the improvement was marginal. This suggests that the trend in accident risk currently used in COBA is appropriate for links. Comparing the re-fitted COBA links models with the TRL size 2 link models indicates that, for both single and dual carriageway links, the best fitting re-fitted TRL models (S4, Table 8 and D4, Table 5) gave only marginally better estimates than the re-fitted COBA models. For schemes, the unadjusted COBA models over-estimate accidents, particularly for dual carriageway schemes, suggesting that the trend in accident risk is under-estimated in the models. However, comparing the re-fitted COBA scheme models with the TRL size 3 models indicates that, for both single and dual carriageway schemes, the best fitting re-fitted TRL models (S10, Table 8 and D13, Table 5) again gave only marginally better estimates than the re-fitted COBA models.

These results confirm that increased model complexity by no means ensures better temporal transferability and that the main issue affecting the model fit is the variation in the trend in accident risk over time. Outdated models which correctly estimate the trend between the modelled period and the current period can give good estimates of current accidents but, because trends in risk can vary over time, it is preferable to estimate trend from the data (by re-fitting or calibration) rather using estimates based on data from an earlier time period.

(Insert Table 9 about here.)

5. Conclusions

Reliable predictive accident models (PAMs) are essential to design and maintain safe road networks but on going changes in road and vehicle design, coupled with a range of road safety initiatives, mean that these models can quickly become dated. While it is quite possible to fit new models, reliable new models would require an extensive (and thus expensive) database which includes high quality data concerning accidents, traffic flows and design characteristics for a large sample of links and/or junctions. In addition, and perhaps more importantly, the fitting of new models is likely to provide only a temporary solution since they will in turn also become outdated. Certainly the difficulties in developing PAMs mean that many widely used models are based on data collected many years ago. While some available models do include a term to adjust for trends in accidents it is by no means certain that accident trends remain constant over time.

This paper has examined two principal issues in predictive accident modelling. Firstly, it has examined the extent to which the temporal transferability of predictive accident models varies with model complexity and, in particular, the extent to which the inclusion of design variables, and modelling by accident type and different road elements in more complex models, increases temporal stability. Secondly, it has investigated the reliability and practicality of updating strategies based both on re-fitting the outdated model and calibrating the model using a scale factor that could allow currently available PAMs to be readily updated to any point in time. The models used to illustrate these issues are the suites of models developed by TRL for rural dual and single carriageway roads in the UK and which are widely used in several software packages in spite of being fitted using data from the 1980s. Some simple accident rate models were also used. The database used to test the models includes data for the 5 years 2005-2009.

On the issue of model complexity, it was found that increased model complexity by no means ensured better temporal transferability. Modelling accidents by type did not improve the goodness-of-fit and, although the inclusion of design features improved the goodness-of-fit of some models (particularly re-fitted models), this was by no means universal. Indeed simple accident rate models could be updated to give estimates that were almost as good as those from models which included a range of design features. Although the re-fitted models suggest that the value of the parameter estimates have changed somewhat over time, the main issue affecting the model fit is the variation in the trend in accident risk over time. While some more complex models include a term to allow for long-term trends in accident risk, our data show that it is by no means clear that the pattern of change will remain stable over time and this can lead to significant bias in the estimates. Simpler models, with trend estimated from the data (by re-fitting or calibration), give better estimates than more complex models with trend predicted using data from an earlier time period.

The fit of the outdated models can be improved either by re-fitting the models or by calibrating the models using a scale factor. Both methods require less data than fitting new models and, because they can be applied at any point in time, avoid the possibility of becoming outdated. Slightly better estimates (lower RMSE) were obtained using re-fitted models but the calibrated models minimise bias (ME=0), can be applied using data for fewer sites and do not rely on specialist statistical knowledge. On balance the calibration of models using a scale factor appears to be the most practical, cost-effective alternative to developing new models.

Acknowledgements

Financial support for this paper was provided by the UK Engineering and Physical Sciences Research Council under grant EP/H046739/1 and the project partners were Lancashire County Council (LCC). The authors would like to thank Peter Andrews, David Cronshaw, Nigel Cleave, Richard Donkin, Ashley Newnham and their colleagues in the Environment Directorate, LCC for their extensive assistance in compiling the data for this paper.

References:

Connors, R., Maher, M.J., Wood, A.G., Mountain, L.J., Ropkins, K., 2012. Methodology for fitting and updating predictive accident models with trend. *Submitted at Accident Analysis and Prevention*.

DfT, 2006. The COBA manual. Department for Transport, London.

DfT, 2010a. Road Casualties Great Britain: 2009. Department for Transport, London.

DfT, 2010b. Transport Trends 2009. Department for Transport, London.

Elvik, R., 2008. The predictive validity of empirical Bayes estimates of road safety. Accident Analysis & Prevention, 40, 1964-1969.

Elvik, R., 2010. The stability of long-term trends in the number of traffic fatalities in a sample of highly motorised countries. Accident Analysis & Prevention, 42, 245-260.

Hashim I.H., Bird R.N., 2005. Analysis of Accident Rates and Geometric Consistency Measures on Sections of Rural Single Carriageway. *Road Safety on Four Continents (VTI/TRB Conference)*, Warsaw, Poland.

Lord, D., Mannering, F.L. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transportation Research Part A* 44 (5), 291-305.

Maher, M.J., Summersgill, I., 1996. A comprehensive methodology for the fitting of predictive accident models. *Accident Analysis and Prevention* 28(3), 1996, 281-296

Maher, M.J., Mountain, L.J., 2009. The sensitivity of estimates of regression to the mean. Accident Analysis and Prevention 41, 861-868.

Maycock G., Hall, R.D., 1984. Accidents at four-arm roundabouts. LR1120, TRL, Crowthorne, UK.

Mountain, L.J., Hirst, W.M., Maher, M.J., 2005. Are speed enforcement cameras more effective than other speed management measures? The impact of speed management schemes on 30mph roads. Accident Analysis and Prevention 37, 742-752.

Persaud, B., Lyon, C., 2007. Empirical Bayes before–after safety studies: Lessons learned from two decades of experience and future directions. Accident Analysis and Prevention 39, 546–555.

Walmsley, D.A., Summersgill, I., 1998. The relationship between road layout and accidents on modern rural trunk roads. Transport Research Laboratory Report 334.

Walmsley, D.A., Summersgill, I. and Payne, A., 1998a. Accidents on modern rural dual-carriageway trunk roads. Transport Research Laboratory Report 335.

Walmsley, D.A., Summersgill, I. Binch, C., 1998b. Accidents on modern rural single-carriageway trunk roads. Transport Research Laboratory Report 336.

Appendix:

The best fitting re-fitted models for dual carriageways (highlighted in Table 5) were:

- Size 2, link: Model D4

$$A = 0.048 \cdot Q^{0.99} \cdot L_L \cdot \exp\begin{bmatrix} -0.034 \cdot (N_m/L_S) + 0.52 \cdot (N_1/L_S) - 0.017 \cdot (N_2/L_S) \\ -0.78 \cdot (N_3/L_S) + 0.15 \cdot (N_4/L_S) - 0.24 \cdot g_{HS2} \\ +0.020 \cdot g_{HS1} - 0.17 \cdot g_Q + 7.0 \cdot 10^{-4} \cdot g_B \\ +0.014 \cdot g_H + 0.081 \cdot g_{NX} - 0.050 \cdot g_{SF} \end{bmatrix} \cdot \exp(-0.067 \cdot t)$$

- Size 2, junction: Model D8

$$A = 0.025 \cdot Q^{0.82} \cdot N_n \cdot \exp[-0.26 \cdot (N_n/L_S) + 0.37 \cdot (N_{3+4}/L_n) + 0.96 \cdot (L_n/N_n)] \cdot \exp(-0.038 \cdot t)$$

- Size 3: Model D13

$$A = 0.050 \cdot Q^{1.02} \cdot L_S \cdot \exp\begin{bmatrix} -0.21 \cdot (N_m/L_S) + 0.79 \cdot (N_1/L_S) + 0.17 \cdot (N_2/L_S) \\ -0.39 \cdot (N_3/L_S) + 0.56 \cdot (N_4/L_S) - 0.11 \cdot g_{HS2} + 011 \cdot g_{HS1} \\ -0.067 \cdot g_Q - 1.5x10^{-3} \cdot g_B + 0.012 \cdot g_H - 2.2x10^{-3} \cdot g_{NX} \\ -.0.44 \cdot g_{SF} \end{bmatrix} \cdot \exp(-0.060 \cdot t)$$

The best fitting re-fitted models for single carriageways (highlighted in Table 8) were:

- Size 2, link: Model S4

$$A = 0.052 \cdot Q^{0.93} \cdot L_L \cdot \exp\begin{bmatrix} -0.17 \cdot (N_m/L_S) - 0.43 \cdot (N_n/L_S) - 0.16 \cdot g_{CW} \\ -0.090 \cdot g_{HS2} + 9.1x10^{-4} \cdot g_B + 0.021 \cdot g_H + 0.031 \cdot g_{NX} \end{bmatrix} \cdot \exp(-0.059 \cdot t)$$

- Size 2, junction: Model S5

$$A = 0.25 \cdot Q^{0.67} \cdot N_n \cdot \exp[-0.60 \cdot (N_n/L_S) - 0.96 \cdot (N_1/L_n)] \cdot \exp(-0.065 \cdot t)$$

- Size 3: Model S10

$$A = 0.087 \cdot Q^{0.88} \cdot L_S \cdot \exp\begin{bmatrix} -0.26 \cdot (N_m/L_S) - 0.019 \cdot (N_n/L_S) - 0.16 \cdot g_{CW} \\ -0.11 \cdot g_{HS2} - 9.7x10^{-4} \cdot g_B + 0.017 \cdot g_H \\ +0.010 \cdot g_{NX} \end{bmatrix} \cdot \exp(-0.060 \cdot t)$$

Where, Q is the two way AADT (in thousand veh/day)

N_m is the number of major junctions on the scheme,

N_n is the number of minor junctions on the scheme,

N_1, N_2 ... are the numbers of minor junctions of types 1, 2 ...

L_L is the total length of the links on the scheme (excluding the length of the minor junctions),

L_S is the total length of the scheme (including the length of the minor junctions),

g_{HS2} is a factor set to unity if hardstrips are present on both the nearside and the offside and set to zero otherwise;

g_{HS1} is a factor set to unity if hardstrips are present on either the nearside or the offside (but not both) and set to zero otherwise;

g_Q is a factor describing the quality of the scheme set to unity for lower quality roads and zero for higher quality roads (as defined in Section 2);

g_B is the bendiness (degrees km^{-1});

g_h is the hilliness (m km^{-1});

g_{NX} is the offside access density (accesses km^{-1});

g_{NXT} is the access density on both sides of the road (accesses km^{-1});

g_{CW} is a factor describing the width of the scheme set to unity for wide roads (10 m) and zero for normal width roads (7.3 m); and

g_{SF} is a factor set to unity if a safety fence is present and set to zero otherwise; and

t is the number of years since 2007.

Captions for figures:

Captions for tables are in the word document containing the tables.

Figure 1 – Outline of scheme definitions.

Figure 2 – Plot of observed versus predicted and standardised residuals versus predicted for best fitting calibrated model for dual carriageways (for individual sites and in bins of 9 sites).

Figure 3 – Plot of observed versus predicted and standardised residuals versus predicted for best fitting re-fitted model for dual carriageways (for individual sites and in bins of 9 sites).