Emerging Crash Trend Analysis

Mark Logan
Department of Main Roads, Queensland

Patrick McShane
Queensland Transport

ABSTRACT

The identification and treatment of "black spots" or hazardous road locations is a core business activity for road authorities. Evaluation of the resulting black spot treatment programs has revealed high returns on investment. Tools used for the identification of black spots include crash frequency (crashes per site or per km), crash rate (crashes per vehicle km travelled) or relative severity index (the aggregate of crash type costs). Australian road authorities, based upon data collected over a 5-year period, use a combination of these tools.

Experience has shown that clusters of crashes can develop over a relatively short time of one to two years. Under such circumstances, the use of the 5 year crash data analysis period would mask the seriousness of the problem and resulting in the late discovery of the site for treatment. This prompted the need to investigate and develop a means with which to provide early warning of emerging high crash locations. The Department of Main Roads and Queensland Transport have developed a methodology called Interpretive Modelling, which compares a suite of five crash indicators, over three different time periods (short, medium and long term), using a rolling segmentation along the road to ensure that no road location is overlooked.

Experience to date with the application of the model has confirmed that it is an effective means of monitoring emerging crash trends.

BACKGROUND

The Queensland Department of Main Roads generally uses road crash data over a 5-year period to analyse crash rates and to identify black spots. This analysis can be run at any time but it is usually run on a yearly basis. Once run, the crash locations are ranked according to their crash rates which are used to identify locations that need remedial treatments.

This method has a deficiency in that where there is a sharp short term change in the number of crashes on a road segment, the longer-term analysis of 5 years does not pick this up this change as it is masked in the 5-year data. The quick changes are not detected early enough for management to take appropriate actions.

Figure 1 – Plot showing emerging crash trend (y axis = time, x axis = chainage)
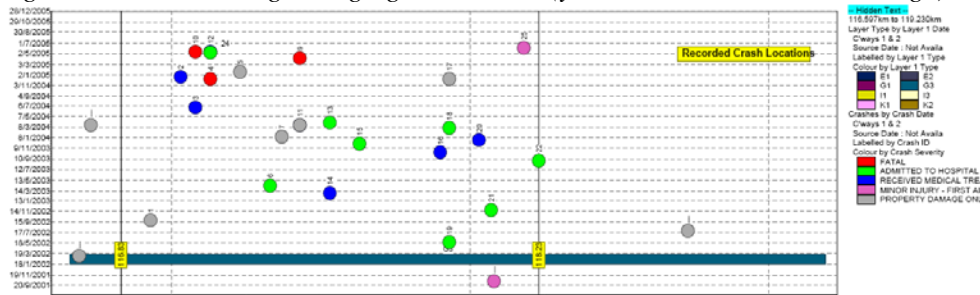
(Produced from the Department of Main Roads Chartview Program)

Figure 1 above shows a section of road over time with the crashes plotted on it. This shows a section of road that has had a significant change in crashes over a very short period of time. This was difficult to detect, as it was not fully evident in the longer-term analysis. In order to detect these short sharp changes, the Department of Main Roads and Queensland Transport have developed a methodology called "Interpretive Modelling".

WHAT IS INTERPRETIVE MODELLING?

Interpretive modelling is a methodology where the road network is segmented so that no location can be missed (1km segments rolling 100m). There are 3 periods of time used, 1 year versus 1 year, 3 years versus 3 years and 1 year versus 5 years, this represents a short, medium and long-term view of the crash history. There are 5 key indicators for which rankings have been created for the comparison between the time periods above. It is the accumulation of these ranking values that give the final ranking and provide the means for identifying those segments that have a significant sharp short-term change.

HOW DOES INTERPRETIVE MODELLING WORK?

Initially, the following 11 potential indicators were chosen for the crash analysis:
- Cost of crashes (based upon average cost by DCA code group)
- Number of crashes
- Number of killed and serious injury (KSI) crashes
- Number of fatalities
- Number of units
- Number of fatal crashes
- Casualty cost (based upon social cost)
- Fatalities per fatal crash
- KSI injuries per KSI crash
- Percentage of fatal crashes per segment
- Percentage of KSI crashes per segment

To determine the most suitable indicators for the model, each indicator was presented in a chart of sequential analyses (rolling in 3 monthly intervals) for a known section of road. Each chart was then analysed to identify the time at which significant increases in the indicator occurred at known crash hot spots.

Figure 2 shows an example of the visual analysis performed on a section of road.   The circled section shows an emerging crash location appearing in the 6th chart. This type of analysis was performed on all of the 11 indicators.

Figure 2 – Plot of crashes per VKT (using a 1 year data period) rolling by 3 monthly intervals.
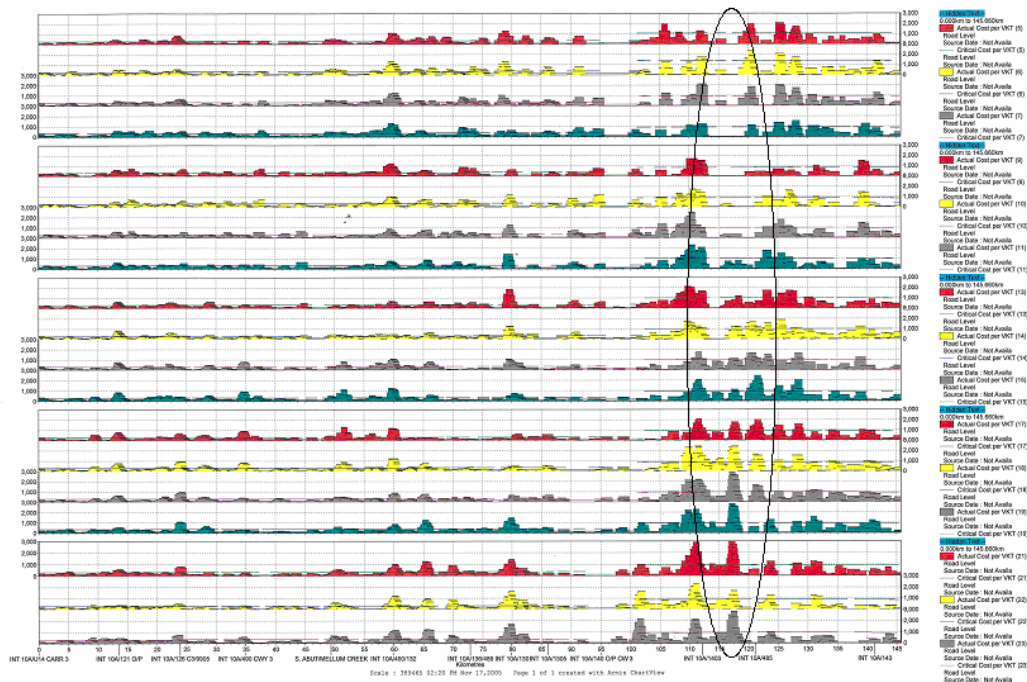
Department of Main Roads and Queensland Transport, 2006

From this analysis, the following 5 indicators that showed the increase at the earliest times:
- Cost of crashes (based upon average cost by DCA code group)
- Number of crashes
- Number of KSI crashes
- Number of fatalities
- Percentage of KSI crashes on a segment

While the selection of indicators utilised exposure based data (VKT), it was not used in the final time series based comparisons, as relatively small changes in VKT did not have a significant effect. This also allows the analysis tool to be applied on roads where traffic information is not available.

A number of considerations were taken into account to establish the segmentation method used in the model. Firstly, a 1 km length segment was adopted, as this was long enough to enable a problem to be identified but not so short as to mask a problem area.

The next consideration was to 'move' the segment along the road. Either fixed sections or rolling sections can be used to identify hazardous links. Fixed sections involve subdividing the road into successive sections of equal length. However, this can prevent the detection of crash concentrations located at the boundary of two adjacent sections. A rolling section approach that shifts a constant section length (1 km) along the road by means of a short increment (100 m) was adopted as it eliminates the fixed length problem. The adoption of this approach resulted in the generation of approximately 360 000 segments on the Queensland state-controlled road network. This allows for the comparison of multiple time periods as the segments locations are the same for each period.

Three different time period comparisons were adopted:
- Current 1 year verus previous 1 year (short term)
- Average of current 3 years versus average of previous 3 years (medium term).
- Current 1 year versus average of the previous 5 years (long term)

In order to rank each segment a severity scoring table was created for each indicator. These severity scoring indicators were generated by banding each indicators result into ranges and plotting these results for each indicator. Each band was given a number from 0 to n signifying the level of severity. These are then added together for each indicator and time period to give a final rating.

The example plot on figure 3 shows the number of occurrences for the difference in the total number of crashes. This chart was used to determine the ranges to use based on the spread of differences.

Figure 3 – Example plot of results for average number of crashes differences
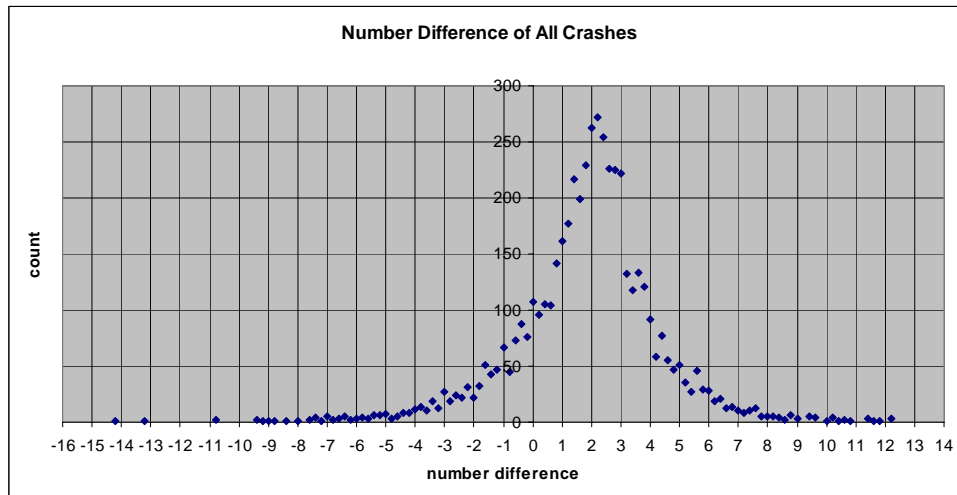
Department of Main Roads and Queensland Transport, 2006

Initially, all segments were used to rank the crashes. However, it was found that a cluster of overlapping segments appeared in the ranked list. Therefore, two overall rankings were generated; the first to rank on every segment and the second to identify the highest scoring segment in a series of overlapping segments and rank only on that segment.

Table 1 shows the results of the analysis above. The ranking disregards any difference less than or equal to 0 as this is a decrease in the number of crashes and is good. The ranking then has bands increasing in 2's going to 8, this would be different for each indicator as they have different ranges of values. Each band is also broken down to >= %50 change and < %50 change. The >=%50 change is given a higher ranking as this shows there a significant change between periods.

Table 1 – Example Indicator 2 ratings for Average Number of Crashes

Number Difference		% Change Difference			SCORE
>	<=	NULL	<	>=	
-	0				0
0	2		50		1
0	2			50	2
0	2	NULL			2
2	4		50		3
2	4			50	4
2	4	NULL			4
4	6		50		5
4	6			50	6
4	6	NULL			6
6	8		50		7
6	8			50	8
6	8	NULL			8
8	-		50		9
8	-			50	10
8	-	NULL			10

Department of Main Roads and Queensland Transport, 2006

Initial runs of the model indicated that there were quite a few segments appearing that only had 1 or 2 crashes in one period and none in the previous period. These locations scored high on the comparison due to one not having a score. A decision was made to restrict the segments to greater than or equal to 2 crashes in the current 1 year period.

It was decided that the model be run every 3 months in order to provide an appropriate balance between early identification of emerging black spots and the desktop data analysis requirements imposed upon districts. In addition, as the process has been implemented as a standard report within Queensland Transport's WebCrash 3 data query tool, it can be initiated at any time using other constraints (for example, by constraining units to motor cycles).

The final results were verified against crash time vs. chainage plots (such as the example in Figure 1). Several Main Roads, Queensland Transport and Queensland University of Technology staff also applied a "reasonableness test".

A standard "crash data interpretive model" report has been implemented in WebCrash 3. The report provides a summary cover page (a single line for each segment, ranked from highest to lowest). Detailed report forms, that include charts and tables showing the number of crashes and injuries, are also generated.

An additional useful reporting format is to present sequential analyses in a spreadsheet. The use of conditional formatting that assigns different colours to the top 50, 50-200 and 200-500, provides a visual tool to highlight areas where a change in trend has emerged. The numbers in Figure 4 relate to the ranking of the road sections in the total network analysis. For example, if it is number 3, it is the third highest rank segment in the analysis. The colour red relates to the top 50 ranked segments, yellow relates to ranks from 51 to 200, and pale yellow is for ranks from 201 to 500.

Figure 4 – multiple run with colour coded rankings showing an emerging section (red)

DISTRICT_ID	ROAD_SECTION_ID	TDIST_START	TDIST_END	MARCH_01	JUNE_01	SEPT_01	DEC_01	MARCH_02	JUNE_02	SEPT_02	DEC_02	MARCH_03	JUNE_03	SEPT_03	DEC_03	MARCH_04	JUNE_04	SEPT_04	DEC_04	MARCH_05	JUNE_05	SEPT_05	DEC_05
99	XXX	10.1	11.1																				
99	XXX	10.2	11.2							553	1002												
99	XXX	10.3	11.3							713	1003								854				
99	XXX	10.4	11.4							1100	1004								158	1701	95	121	1345
99	XXX	10.5	11.5																176	1102	23	32	1072
99	XXX	10.6	11.6													719	498	497	106	1617	103	128	637
99	XXX	10.7	11.7													187	123	139	30	1396	102	84	241
99	XXX	10.8	11.8										597			150	57	76	13	1159	341	270	331
99	XXX	10.9	11.9										279		582	20	25	21	7	936	600	487	328
99	XXX	11	12										280		583	55	44	33	14	944	461	378	329
99	XXX	11.1	12.1										281		104	21	26	22	73	937	601	488	330
99	XXX	11.2	12.2												165	2	5	1	4	726	704	593	333
99	XXX	11.3	12.3									179	471	466	149	3	17	5	339	719	754	637	386
99	XXX	11.4	12.4									94	369	330	57	4	18	6	671	785	777	668	662
99	XXX	11.5	12.5									95	232	189	45	5	8	7	672		778	669	663
99	XXX	11.6	12.6									96			264	195	6	3	669		966	670	656
99	XXX	11.7	12.7									97			265	1057	62	90	938		1594	993	1219
99	XXX	11.8	12.8												1165	656	207	279			2256	1354	988
99	XXX	11.9	12.9													1403	1002	1097			2601	1355	1220
99	XXX	12	13							618	88	98				1404	1003	904			2602	1335	1493
99	XXX	12.1	13.1							619	89	99				1405	1004	905			2603	1336	1494
99	XXX	12.2	13.2																			1415	1294
99	XXX	12.3	13.3																				
99	XXX	12.4	13.4																				

WHAT HAVE WE DONE SO FAR?

The process has been finalised and checked and has been implemented for mid block and intersection crashes for the Queensland's state-controlled road network in WebCrash 3.

The reports are now run every 3 months and the results are sent to our Districts for investigation.

WHERE TO FROM HERE?

The process will be implemented in the Main Roads data centre, where it will be set to run automatically every 3 months. This will then be available for packages like Chartview and Mapview (MR products) so that sequential analyses can be displayed.

It might be necessary to have a closer look at the rating tables to check that the values given for each indicator are still valid and to check if changing these will give better results.

Queensland Transport is investigating means of implementing the model for local government controlled roads.

CONCLUSION

Experience with the application of the model has confirmed that it is an effective means of identifying emerging crash trends. While the results often include a number of "false positives" (locations which exhibit an increase in crashes, but which upon closer examination cannot be attributed to road environment conditions), the time required to undertake a desktop review of the ranked sites and strip out these sites is short (the time investment is minor relative to the significant benefits).

The model will enable the Department of Main Roads to closely monitor emerging high crash road locations, and to take action before they become black spots.

REFERENCES

Department of Main Roads and Queensland Transport, *Interpretative Modelling,* 2006.