EXPLORING THE USE OF CORPORATE DATA FOR ROAD SAFETY RISK ASSESSMENT

Comparison of corporate and traditional datasets for use in AUSRAP and ANRAM assessments

The WARRIP team explored the suitability and applicability of using data extracted from the Main Roads WA corporate database for use in producing AusRAP / ANRAM data sets.

Background

A proactive approach to road safety risk assessment can help identify locations across the road network that experience high risk for fatal and severe injury (FSI) crashes. This enables Main Roads WA to flag and prioritise areas that might require treatments to improve safety. AusRAP and ANRAM are two tools that can provide insights into identifying crash risk across the road network. They require specific data to inform safety assessments. Using existing corporate data would ultimately save time and money if this data could be adapted to this purpose.

Approach

- Preparation of AusRAP data sets.
  - 'traditional' (manually coded) data set using TSD data (obtained during the initial TSD trial).
  - 'corporate' data set extracted from Main Roads WA corporate asset database.
- Preparation of ANRAM data sets ('traditional' and 'corporate'), incorporating FSI crash data.
- Analysis of coding standards between the data sets, looking for similarities and differences in coding.
- AusRAP and ANRAM analysis of each data set, to gain insights into differences in risk assessment results between the two data sets.
- Collaboration between Main Roads WA and ARRB, particularly in preparation of the 'corporate' data set.
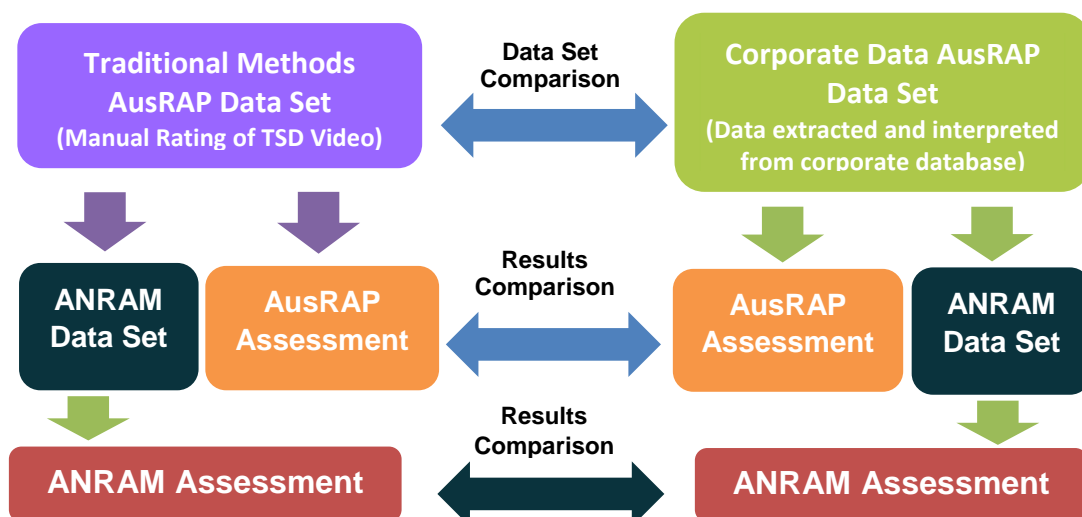
Differences between data sets

An analysis of the differences between the data sets found similarities in the coding of results of some attributes, but differences in many others. It was noted that coding differences for some attributes were likely due to the approach used for extraction and interpretation of attributes from corporate data base.

For the 'road condition' and 'skid resistance' attributes, traditional data sets rely on subjective visual rating of road sections. In comparison, the corporate data set used roughness and rutting information to classify the road sections. This difference in producing data resulted in differences in coding between the two data sets.

Differences between risk assessment results

A review of the AusRAP data results for each 100 m segment found that approximately only half of the results were the same.

A comparison of ANRAM results found that results were quite different for the two data methods. The 'traditional' method attributed approximately 50% of the FSI crash risk to run-off-road crash types, compared to approximately 30% for the 'corporate' data set.

Challenges and observations

It is challenging to align data from different sources. For this project, it was found that:

- Aligning road distances can be challenging, especially when methodology and original purpose of data gathered varies.
- Differences in terminology and/in interpretation of attributes can lead to confusion and differences in classification of attributes. Clear communication is crucial to collaborative projects. Interpretation and categorisation of attributes need to be consistently applied so that results aren't skewed by the different approaches, or where a different approach is to be employed, the impact on results needs to be recognised and understood.
- Data currency and accuracy can also be a factor. For some attributes, corporate data may provide more accurate information than the traditional (manual) rating approach, while for other attributes the traditional rating approach may be more up-to-date data than the corporate data. It should be noted that bias or subjectivity may be evident in any data set, particularly in relation to data that involves human interpretation or input. This may be present in corporate or traditionally coded data.
- Deterioration of assets needs to be considered for attributes like roughness and line marking which will change over time. Consideration may need to be given to issues such as currency of corporate data (for these attributes), and whether a rule may need to be applied as a proxy for condition.
- Roadside hazards are currently difficult to systematically or automatically ascertain; however, it is likely to become possible as technology in this space continues to evolve.

Next Steps

Corporate data could be used for some, although not currently all attributes for producing AusRAP and ANRAM data sets. With further refinement there is confidence additional attributes could be extracted from the corporate database.

New and evolving technology will continue to lead to improvements in accuracy of data sets. For example, automatic categorisation of roadside hazards and distance measurement.