Consistent patterns of vehicle collision risk for six mammal species

Casey Visintin\textsuperscript{a}, Rodney van der Ree\textsuperscript{b}, Michael A. McCarthy\textsuperscript{c}

\textsuperscript{a}Quantitative and Applied Ecology Group, School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia - Email: cvisintin@student.unimelb.edu.au

\textsuperscript{b}Ecology and Infrastructure International and School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia - Email: rvdr@unimelb.edu.au

\textsuperscript{c}Quantitative and Applied Ecology Group, School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia - Email: mamcca@unimelb.edu.au

Corresponding Author: Casey Visintin, School of BioSciences, Bldg 122 - Rm 106A, University of Melbourne, Parkville, VIC 3010, Australia - Email: cvisintin@student.unimelb.edu.au, Phone: +61 4 34424084

Running Title: Collision risk model - multiple species

Word Count: 4871

Number of Tables: 5

Number of Figures: 9 (2 colour figures, 5 supplemental figures)

Number of References: 57
Abstract

The occurrence and rate of wildlife-vehicle collisions is related to both anthropocentric and environmental variables, however, few studies compare collision risks for multiple species within a model framework that is adaptable and transferable. Our research compares collision risk for multiple species across a large geographic area using a conceptually simple risk framework.

We used six species of native terrestrial mammal often involved with wildlife-vehicle collisions in south-east Australia. We related collisions reported to a wildlife organisation to the co-occurrence of each species and a threatening process (presence and movement of road vehicles).

For each species, we constructed statistical models from wildlife atlas data to predict occurrence across geographic space. Traffic volume and speed on road segments (also modelled) characterised the magnitude of threatening processes.

The species occurrence models made plausible spatial predictions. Each model reduced the unexplained variation in patterns and distributions of species between 29.5% (black wallaby) and 34.3% (koala). The collision models reduced the unexplained variation in collision event data between 9.1% (black wallaby) and 19.4% (common ringtail possum) with predictor variables correlating similarly with collision risk across species.

Road authorities and environmental managers need simple and flexible tools to inform projects. Our model framework is useful for directing mitigation efforts (e.g. on road effects or species presence), predicting risk across differing spatial and temporal scales and target species, inferring patterns of threat, and identifying areas warranting additional data collection, analysis, and study.
22 **Keywords**

23 animal, framework, species distribution model, speed limit, traffic volume, WVC
1. Introduction

Roads have considerable negative ecological impacts, prompting numerous scientific studies and mitigation projects over the past five decades (Forman & Alexander, 1998; Spellerberg, 1998; van der Ree et al., 2015). The mortality of wildlife from collisions with vehicles is a significant problem throughout most of the developed world, and will likely become more so in the next few decades in the developing world (van der Ree et al., 2015). Wildlife-vehicle collisions (WVC) are estimated to kill billions of fauna annually on transportation networks (Seiler & Helldin, 2006), fostering an interdisciplinary management problem. To address this problem, studies seek to understand the frequency, magnitude and distribution of WVC. Through this understanding, managers can identify and apply appropriate mitigation strategies.

WVC are financially costly (Bissonette et al., 2008; Huijser et al., 2009; Rowden et al., 2008) and thus knowing where and how to mitigate is important. If mitigation measures are not appropriately specified or located in the landscape, costs arise from both wasted installation labour, materials and time, and on-going collisions resulting from the omission. Moreover, some forms of mitigation, such as fencing, also create "barrier-effects" - a direct impediment to the movement of many species which affects movement and gene flow (Epps et al., 2005). Thus, strategic, informed use of mitigation is important for both conservation and road safety objectives.

The literature contains many studies on WVC with several papers utilising statistical modelling methods. Quantitative models are a useful way to support decision-making for managers by helping to clearly organise problems, test inputs, and make inferences (Anderson et al., 2015).
By determining the factors that contribute to collisions with quantitative analysis, managers can begin to understand relationships and optimise mitigation strategies.

We view two challenges in current WVC modelling and management practices. First, methods that generalise to multiple species and across taxonomic groups are under-represented in scientific studies (see Farmer & Brooks, 2012). This is likely due to the complexity in resource requirements for different taxa. Another challenge is identifying relationships between predictors and collisions that are able to be mitigated. Although management implications are highlighted in many studies, it is not always clear how the results of the analyses could be easily applied due to variable interactions and confounding effects in the model predictors (Gunson et al., 2011). Some studies explicitly suggest mitigation that relates directly to environmental variables based on model results (see Grilo et al., 2009). Although these recommendations are useful in the specific contexts of the studies, they may not generalise to varying spatial scales - specifically large areas. We argue that generality in methods is useful and it is important to create analytical tools that help managers identify risk to both individual species and taxonomic groups, across large areas, and with predictors that are within the scope of management.

In this study we have two objectives: testing a conceptual risk model framework and analysing patterns of WVC for multiple species. We first model wildlife-vehicle collisions by expanding upon a formerly developed conceptual analytical framework by Visintin et al. (2016) that relates risk to exposure (species occurrence) and hazard (traffic volume and speed). As this model framework is able to quantify risk over large spatial scales (demonstrated on a large region of south-east Australia as a case study), we apply it to six Australian terrestrial mammal species commonly involved in WVC to test its flexibility and potential to support management decision-
making. We analyse model outputs to determine the relationships of common factors contributing to wildlife-vehicle collisions and make predictions of risk to identify/prioritise areas requiring mitigation. As the need for and location of WVC mitigation is often derived using species characteristics and movements (e.g. Clevenger et al., 2002), we use the model framework to determine the importance of species occurrence in collision risks on road segments. Ignoring variables that influence habitat preferences of species may lead to less robust inference by managers (Roger & Ramp, 2009). Where occurrence is influential for multiple species, managers may decide to control animal presence on or near the roads and this may involve a mixture of mitigation strategies (see Beckmann et al., 2010). Likewise, if traffic speed or volume is a significant driver of collisions for many species, managers may consider control mechanisms such as speed enforcement or alternative transportation planning.
2. Materials and Methods

2.1 Study Area

We used the 227,819 square kilometre state of Victoria in south-east Australia as a study area (Figure 1). The Victorian road authority (VicRoads) manages 25,256 kilometres of major roadway within the state. The remaining sealed roads (approx. 125,000 kilometres) are controlled by seventy-nine municipal districts. Our study analysed collision risk across 147,970 kilometres of sealed roads. To organise our spatial data and modelling, we overlaid a spatial grid of 1km$^2$ resolution (extents: -58000,5661000 x 764000,6224000, projection: GDA94 MGA zone 55, number of cells: 462,786) on the study area. Each grid cell was the modelling unit for species occurrence. All roads in the study area were bisected by the grid resulting in road segments that were approximately one kilometre or less in length. We used these 612,791 segments as our modelling units for the collision model.
2.2 Study Species

We obtained records of WVC from the Wildlife Victoria database (see 'Collision Model' section below), and selected the species of mammal most frequently recorded as roadkill (Table 1). Only mammals with at least 80 reported collision events were selected for use in the study resulting in six study species. Eastern grey kangaroos (*Macropus giganteus*, Shaw) are the second largest mammal (up to 85 kilograms for males) in Australia and share many management issues regarding abundance with ungulates found in North America and Europe (Croft, 2004; Coulson & Eldridge, 2010). They occur in groups and have home range sizes between 25 and 125 hectares (Dawson, 2012). Black wallabies (*Wallabia bicolor*, Desmarest) are medium-sized (13 to 17 kilograms), solitary mammals more often found at higher elevations and in areas of denser...
foliage (Van Dyck & Strahan, 2008). Common wombats (Vombatus ursinus, Shaw) are medium-sized (22 to 39 kilograms), burrowing mammals and one of three wombat species in Australia (Van Dyck & Strahan, 2008). Although wombats occasionally share burrows, they are typically solitary animals and occupy home range sizes of about 20 hectares. Common possums, brushtail (Trichosurus vulpecula, Kerr) and ringtail (Pseudocheirus peregrinus, Boddaert), are small arboreal mammals (one to four kilograms) that are often more abundant throughout Victorian urban and suburban environments than other arboreal mammals. Koalas (Phascolarctos cinereus, Goldfuss) are medium-sized (eight to 15 kilograms) arboreal mammals. They are mainly sedentary due to the exclusive diet of Eucalyptus leaves which have low caloric content and nutritional value and are toxic to many other species (Van Dyck & Strahan, 2008).

2.3 Conceptual Model Framework

We employed a quantitative risk model framework (Visintin et al., 2016) to examine how collision risk of each species relates to occurrence of the species, traffic volume, and traffic speed on all road segments. Using the open-source software package ‘R’ version 3.3.0 (R Core Team, 2016) to perform all statistical analyses, we developed species distribution models (SDM) to predict occurrence across the study area for each of the six species, and linear regression models to predict traffic volume and speed. Predicted traffic volume and speed values for all road segments were modelled by regressing annual average daily traffic (AADT) counts and posted speed limit data on anthropogenic variables using random forests; detailed methods are provided in Visintin et al. (2016).

2.4 Species Occurrence Models

We obtained observation records of the six study species from the Victorian Biodiversity Atlas
that satisfied the following criteria: survey date between 1 January 2000 and 31 December 2013 and spatial coordinate certainty of $\leq 500$ metres (DELWP, 2015). As our occurrence models are correlative, we grouped the records based on identical spatial coordinates regardless of observation dates; thus multiple observations were aggregated to single presence observations in space. For each of the six study species, we selected all 1km$^2$ grid cells that contained at least one occurrence record to represent presences across the study area. This ensured that we maintained at least 1000 meters between all observations. As we did not have access to recorded absence data, we randomly selected 10,000 1km$^2$ grid cells (without replacement) within the study area as background data and combined them with presence data (Table 1). Note, any spatially duplicated presence and background grid cells were reduced to single presence observations; therefore, some datasets may incorporate less than 10,000 background observations.

We selected predictor variables which influenced the biology, behaviour, and characteristics of all species based on literature review, consultation with species experts, and ecological principles. We considered a range of variables that describe three strata that our six species occupy; the subsurface zone for burrowing animals, the ground plane for ranging species, and the canopy for tree-dwelling wildlife. Each predictor was represented as a 1km$^2$ raster grid and values were sampled using the respective grid centroid coordinates in the presence/background datasets for each species. To reduce effects of collinearity, we used a predictor set that exhibited Pearson correlation coefficients less than 0.70 for all pairs of explanatory variables. There were two exceptions; elevation and slope (0.71), which were expected to exhibit some collinearity, and tree density and precipitation of the driest month (0.72).

Table 1: Top six species most frequently reported in wildlife-vehicle collisions in the Wildlife Victoria database (WVD) between the years of 2010 and 2014. Species observations from the Victorian Biodiversity Atlas were used
for the species occurrence models and recorded collisions from the WVD for the collision risk models. The datasets used in the models are comprised of both presence (P) and background (B) observations. Note, some species occurrence modelling datasets incorporate less than 10,000 background observations due to reductions in spatial duplicates - i.e., presence and background grid cells sharing the same coordinates were reduced to single presence observations.

<table>
<thead>
<tr>
<th>Species Name</th>
<th>Species Code</th>
<th>Species Occurrence Model Data</th>
<th>Collision Model Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern Grey Kangaroo</td>
<td>EGK</td>
<td>10508 : 703(P) / 9805(B)</td>
<td>612791 : 1344(P) / 611447(B)</td>
</tr>
<tr>
<td>Common Brushtail Possum</td>
<td>BTP</td>
<td>10808 : 1021(P) / 9787(B)</td>
<td>612791 : 257(P) / 612534(B)</td>
</tr>
<tr>
<td>Common Ringtail Possum</td>
<td>RTP</td>
<td>10590 : 791(P) / 9799(B)</td>
<td>612791 : 212(P) / 612579(B)</td>
</tr>
<tr>
<td>Black Wallaby</td>
<td>BSW</td>
<td>10548 : 755(P) / 9793(B)</td>
<td>612791 : 108(P) / 612683(B)</td>
</tr>
<tr>
<td>Common Wombat</td>
<td>WOM</td>
<td>10330 : 515(P) / 9815(B)</td>
<td>612791 : 156(P) / 612635(B)</td>
</tr>
<tr>
<td>Koala</td>
<td>KOA</td>
<td>10304 : 489(P) / 9815(B)</td>
<td>612791 : 81(P) / 612710(B)</td>
</tr>
</tbody>
</table>

For each species, we fit Boosted Regression Tree (BRT) models (Friedman, 2002) to the occurrence data using tools and methods by Elith & Leathwick (2008). These methods included the use of internal cross-validation to reduce over-fitting. BRT fit complex non-linear relationships and automatically incorporate interaction effects between predictors (Elith et al., 2009). We selected a tree complexity of five (limit on number of terminal nodes per tree used to regulate interactions), a learning rate of .005 (contribution of each tree to the model), and bag fraction of 0.5 (decimal percent of data values used to cross validate the model predictions). For each model, we employed a selection mechanism - based on Elith et al. (2008) - to determine the most useful and parsimonious combination of predictors (Table 3) from a full set of 18 variables (Table 2).

**Table 2:** Predictor variables used in species occurrence (SO) models and collision risk (CR) models. The spatial
coordinates of centroids for grids with species presences and 10,000 randomly selected background grids were used
to sample from 1km² resolution predictor variable grids for occurrence models. The mid-points of road segments
were used to sample from 1km² resolution occurrence model predictions. Note, reported means and ranges are for
entire study area.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Units</th>
<th>Range : Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELEV (SO)</td>
<td>Elevation of terrain above sea level</td>
<td>m</td>
<td>-0.39-1890.66 : 252.83</td>
</tr>
<tr>
<td>GRASS (SO)</td>
<td>Mean annual grass growth (Carter et al., 2003)</td>
<td>kg ha⁻¹</td>
<td>0-15.74 : 5.45</td>
</tr>
<tr>
<td>GREEN (SO)</td>
<td>Mean seasonal change in vegetation greenness</td>
<td>–</td>
<td>-0.62-0.68 : 0.21</td>
</tr>
<tr>
<td>HYDRODIST (SO)</td>
<td>Distance from major waterway</td>
<td>km</td>
<td>0.95 : 10.09</td>
</tr>
<tr>
<td>HYDROFLOW (SO)</td>
<td>Hydrological flow accumulation</td>
<td></td>
<td>1-32767 : 197.95</td>
</tr>
<tr>
<td>ISOTHERM (SO)</td>
<td>Isothermality</td>
<td>°C*10</td>
<td>0-53 : 48.40</td>
</tr>
<tr>
<td>LIGHT (SO)</td>
<td>Remote-sensed artificial light intensity</td>
<td>–</td>
<td>0-63 : 1.23</td>
</tr>
<tr>
<td>MNTTEMPWQ (SO)</td>
<td>Mean temperature of wettest quarter</td>
<td>°C*10</td>
<td>-8-162 : 102.18</td>
</tr>
<tr>
<td>PRECDM (SO)</td>
<td>Precipitation of driest month</td>
<td>mm</td>
<td>0-81 : 34.70</td>
</tr>
<tr>
<td>PRECSEAS (SO)</td>
<td>Precipitation seasonality</td>
<td>mm</td>
<td>0-43 : 23.99</td>
</tr>
<tr>
<td>ROADDIST (SO)</td>
<td>Distance from roadway</td>
<td>km</td>
<td>0-36.55 : 2.02</td>
</tr>
<tr>
<td>SLOPE (SO)</td>
<td>Slope of terrain</td>
<td>%</td>
<td>0-41.61 : 2.69</td>
</tr>
<tr>
<td>SOILBULK (SO)</td>
<td>Mean soil bulk density at 2.5cm depth</td>
<td>kg m⁻³</td>
<td>0-1365 : 1129.13</td>
</tr>
<tr>
<td>SOILEXCH (SO)</td>
<td>Mean soil cation exchange capacity at 2.5cm depth</td>
<td>cmolc kg⁻¹</td>
<td>10-102 : 24.84</td>
</tr>
<tr>
<td>SOILSAND (SO)</td>
<td>Mean soil texture fraction sand at 2.5cm depth</td>
<td>%</td>
<td>23-100 : 55.11</td>
</tr>
<tr>
<td>TEMPANRANGE (SO)</td>
<td>Temperature annual range</td>
<td>°C*10</td>
<td>0.289 : 245.10</td>
</tr>
<tr>
<td>TOWNDIST (SO)</td>
<td>Distance from town center</td>
<td>km</td>
<td>0-46.10 : 5.76</td>
</tr>
<tr>
<td>Predictor</td>
<td>EGK</td>
<td>BTP</td>
<td>RTP</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>ELEV</td>
<td>9.73</td>
<td>7.67</td>
<td>4.72</td>
</tr>
<tr>
<td>GRASS</td>
<td>8.33</td>
<td>6.64</td>
<td>5.44</td>
</tr>
<tr>
<td>GREEN</td>
<td>6.71</td>
<td>7.27</td>
<td>5.68</td>
</tr>
<tr>
<td>HYDRODIST</td>
<td>2.31</td>
<td>2.88</td>
<td>2.07</td>
</tr>
<tr>
<td>HYDROFLOW</td>
<td>---</td>
<td>2.56</td>
<td>---</td>
</tr>
<tr>
<td>ISOTHERM</td>
<td>2.49</td>
<td>2.82</td>
<td>3.09</td>
</tr>
<tr>
<td>LIGHT</td>
<td><strong>14.85</strong></td>
<td>10.19</td>
<td><strong>13.44</strong></td>
</tr>
<tr>
<td>MNTEMPWQ</td>
<td>3.57</td>
<td>5.44</td>
<td>6.32</td>
</tr>
<tr>
<td>PRECDM</td>
<td>7.13</td>
<td>7.49</td>
<td>7.98</td>
</tr>
<tr>
<td>PRECSEAS</td>
<td>3.41</td>
<td>5.27</td>
<td>4.23</td>
</tr>
</tbody>
</table>

Table 3: Final predictors selected for each species in species occurrence models and relative importance in model expressed as percentage contribution to reduction in unexplained variation in the data. “—” indicates variable was not selected for use in the model and bold text signifies most influential variable per species.
We refit the models using the best combination of predictors and predicted relative likelihood of occurrence to 1km$^2$ spatial grids for six species (Figure 2). We examined the rasters for unrealistic predictions of distributions by plotting atlas records of occurrences on prediction grids and also cross-referencing with expert-derived species range maps. Spatial autocorrelation was also assessed for each species by calculating Moran’s $I$ from model residuals and spatial coordinates and plotting against distance in one kilometre bins.
2.5 Collision Model

We obtained spatially unique collision record coordinates from the Wildlife Victoria database for each species spanning a four year period between 1 January, 2010 and 31 December, 2013.
Wildlife Victoria is a not-for profit organisation that maintains a professional database of state-wide reported wildlife injuries and includes spatial and temporal data for more than 80,000 records. For this study, we extracted records with cause of injury reported as 'hit by vehicle' and only included records with a spatial accuracy <300 metres.

For each target species, we selected the road segments that intersected with the reported location of each collision and re-coded them with ones. We selected the remaining road segments in the study area, coded them with zeros to represent background data, and combined them with the 'collision' records (Table 1). We sampled the relative species occurrence predictions at midpoints of the road segments in the collision dataset by intersecting the road features with the respective occurrence spatial grids. Traffic modelling in previous work had already predicted traffic speed and volumes for each road segment (Visintin et al., 2016). This resulted in six modelling datasets (one for each species), each with a binary dependent variable of collision (1) or background (0) and three continuous predictors of species occurrence, traffic volume, and traffic speed.

We regressed collisions on the predictor variables using a complementary log-log (cloglog) link. This link considers the collision data as resulting from a truncated count process which agrees with both our underlying theory about how collisions occur (as a rate following a Poisson distribution) and our data construction methodology. Further, this link can be used to specify a model that is invariant to the spatial scale of the observations allowing greater flexibility as a modelling tool. The cloglog link contrasts with the popular logit link, commonly used for binomial data, which does not have the same interpretation. In this study, we included two variations on the original model presented in Visintin et al., (2016). First, we added an offset
term to account for variation in road lengths and the temporal span of the data. We intend our
model outputs to be suitable for management and the offset allows us to represent risk across the
network using annual collisions per kilometre. Second, we included a quadratic term for traffic
volume to allow a unimodal response shape. Former studies indicate that some species are
repelled by roads when traffic volume increases beyond a threshold value (Seiler, 2005; Gagnon
et al., 2007). Other possible determinants of this response shape include changes in local wildlife
abundance resulting from previous wildlife-vehicle collisions, absence of wildlife in highly
urbanised areas (i.e., heavy traffic areas), and reporting bias due to observers having difficulty
stopping to report collisions on roads with high traffic volumes. We tested quadratic terms for the
other predictors, however, they were not informative in this work. Our final model is expressed
as:

\[
cloglog(p_i) = \beta_0 + \beta_1 \ln(O_i) + \beta_2 \ln(V_i) + \beta_3 (\ln(V_i))^2 + \beta_4 \ln(S_i) + \ln(L_i \times T)
\]

where \(p_i = \Pr(Y_i = 1)\) is the relative likelihood of a collision occurring, \(O_i\) is species occurrence, \(V_i\)
is traffic volume, \(S_i\) is traffic speed, and \(L_i\) is the length of a given road segment \(i\). \(T\) is a constant
representing the total time span of the data in years and is used in the road length offset to scale
the model to an annual baseline.

Using the model fits, we predicted relative collision risk to all roads segments for each of the six
study species. To validate the predictions of the model, we created independent datasets for each
species using the same methods as the modelling data but for collisions recorded in the year
2014. We calculated receiver operator characteristic (ROC) scores (Metz, 1978), also known as
area under the curve (AUC), to assess the ability of the models to correctly predict collision
events. Although other metrics are used to measure model discrimination (e.g. true skill statistic
(TSS), see Allouche et al., 2006), they are often only suited to presence-absence data where prevalence is known. ROC uses non-specific threshold values and is an appropriate measure for assessing the discriminative ability of presence-only models (Lawson et al., 2014).

To determine collective risk on each road segment, we summed the total relative collision rates for all species on each segment. These results were mapped to highlight road segments with generally high risk regardless of target species. For each species, we also identified road segments that ranked in the top 0.1% for predicted collision rates within each respective range; collision rates varied by species due to differences in reported collisions. We recorded the frequency that individual road segments appeared in this category, the values ranging from zero (no species) to six (all species).

All of the modelling in our study was correlative, however, we explored the influence of seasonal phenomena by partitioning the collision data into four seasons (summer, autumn, winter, spring) and running the models for each subset. Note, due to the small sample bias through partitioning, we fit the data using penalised likelihood methods (Firth, 1993). We plotted the marginal effects of the predictors on collisions to determine if any season deviated from expected patterns.
3. Results

The species occurrence models produced ecologically-plausible predictions about the relative likelihood of species presence across Victoria. The deviance (unexplained variation in the data) reduced by the models varied between 29.5% and 34.3% and the mean cross-validated ROC scores varied between 0.87 and 0.92 (Table 4). The model that best fit the data was for the koala.

**Table 4:** Statistical model performance as percent reduction of deviance on the null model and receiver operator characteristic (ROC) scores for species occurrence (SO) and collision risk (CR) models by species. Note, due to the nature of gradient boosting, errors on both deviance and ROC scores are reported as +/- values for species occurrence models.

<table>
<thead>
<tr>
<th>Species Name</th>
<th>% Deviance (SO)</th>
<th>ROC (SO)</th>
<th>% Deviance (CR)</th>
<th>ROC (CR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern Grey Kangaroo</td>
<td>30.7 +/- 0.7</td>
<td>0.89 +/- 0.004</td>
<td>11.8</td>
<td>0.80</td>
</tr>
<tr>
<td>Common Brushtail Possum</td>
<td>32.6 +/- 1.2</td>
<td>0.89 +/- 0.005</td>
<td>16.1</td>
<td>0.83</td>
</tr>
<tr>
<td>Common Ringtail Possum</td>
<td>33.7 +/- 1.1</td>
<td>0.90 +/- 0.005</td>
<td>19.4</td>
<td>0.87</td>
</tr>
<tr>
<td>Black Wallaby</td>
<td>29.5 +/- 1.5</td>
<td>0.87 +/- 0.007</td>
<td>9.1</td>
<td>0.80</td>
</tr>
<tr>
<td>Common Wombat</td>
<td>33.7 +/- 1.9</td>
<td>0.91 +/- 0.006</td>
<td>15.7</td>
<td>0.89</td>
</tr>
<tr>
<td>Koala</td>
<td>34.3 +/- 1.9</td>
<td>0.92 +/- 0.007</td>
<td>7.4</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Five variables were not selected in every species model and the least selected variables were distance to major waterway and hydrological flow accumulation (Table 3). The three most influential predictors were annual range of temperature, artificial light, and precipitation in the driest month (Figure S1). The lowest variation in relative likelihood of occurrence for all species was for annual range of temperature (0.00 to 0.13). Brushtail possums and black wallabies both had the largest spikes in relative occurrence at values of approximately 18 degrees Celsius. Relative likelihood of occurrence increased at higher values of artificial light for all species except koalas and wombats. At high values of precipitation in the driest month, occurrence of the
two possums deviated considerably from the trends observed in all other species.

Spatial autocorrelation patterning was evident in the species occurrence model residuals between two and approximately twelve kilometres (Figure S2). Correlation values (Moran's I) exceeded 0.4 at distances of two kilometres or less for all species except the ringtail possum. The highest overall spatial autocorrelation was in the koala model residuals, however, the wombat, koala, and black wallaby models all demonstrated high levels of spatial autocorrelation at one kilometre distances.

All of the collision models fit the data as expected. Reduction on the null deviance was between 7.4% and 19.4% and all of the collision models attained high ROC scores when predictions were compared to the independent validation data (Table 4). The black wallaby model had the lowest ability to discriminate between false positives and false negatives (0.80) and the koala model had the highest (0.89). All of the predictors were highly significant (p < .001) in all of the models except traffic speed which was not a highly significant variable for the three arboreal mammal species (Table 5).

Table 5: Summary of collision models fits. Coefficients and significance of variables are shown with relative percent contribution (ANOVA) to model fits for each species. Highly significant variables (p<.001) are marked with asterisks.

<table>
<thead>
<tr>
<th>Species</th>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Z-value</th>
<th>Pr(Z)</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern Grey Kangaroo</td>
<td>Intercept</td>
<td>-46.55</td>
<td>1.96</td>
<td>-23.74</td>
<td>&lt;.001*</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>EGK</td>
<td>0.67</td>
<td>0.02</td>
<td>27.6</td>
<td>&lt;.001*</td>
<td>69.53</td>
</tr>
<tr>
<td></td>
<td>TVOL</td>
<td>7.98</td>
<td>0.46</td>
<td>17.23</td>
<td>&lt;.001*</td>
<td>7.18</td>
</tr>
<tr>
<td></td>
<td>TVOL2</td>
<td>-0.47</td>
<td>0.03</td>
<td>-16.17</td>
<td>&lt;.001*</td>
<td>11.63</td>
</tr>
<tr>
<td></td>
<td>TSPD</td>
<td>2.19</td>
<td>0.13</td>
<td>17.35</td>
<td>&lt;.001*</td>
<td>11.66</td>
</tr>
<tr>
<td>Common Brushtail Possum</td>
<td>Intercept</td>
<td>-57.19</td>
<td>9.55</td>
<td>-5.99</td>
<td>&lt;.001*</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>BTP</td>
<td>0.86</td>
<td>0.09</td>
<td>9.78</td>
<td>&lt;.001*</td>
<td>73.29</td>
</tr>
<tr>
<td></td>
<td>TVOL</td>
<td>11</td>
<td>2.01</td>
<td>5.48</td>
<td>&lt;.001*</td>
<td>19.16</td>
</tr>
<tr>
<td></td>
<td>TVOL2</td>
<td>-0.59</td>
<td>0.12</td>
<td>-5.12</td>
<td>&lt;.001*</td>
<td>7.52</td>
</tr>
<tr>
<td></td>
<td>TSPD</td>
<td>0.18</td>
<td>0.36</td>
<td>0.49</td>
<td>0.62</td>
<td>0.03</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>TSPD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common Ringtail Possum</td>
<td>Intercept</td>
<td>-146.7</td>
<td>25.95</td>
<td>-5.65</td>
<td>&lt;.001*</td>
<td>---</td>
</tr>
<tr>
<td>RTP</td>
<td>1.08</td>
<td>0.1</td>
<td>10.51</td>
<td>&lt;.001*</td>
<td>74.44</td>
<td></td>
</tr>
<tr>
<td>TVOL</td>
<td>30.47</td>
<td>5.68</td>
<td>5.37</td>
<td>&lt;.001*</td>
<td>13.93</td>
<td></td>
</tr>
<tr>
<td>TVOL2</td>
<td>-1.68</td>
<td>0.32</td>
<td>-5.27</td>
<td>&lt;.001*</td>
<td>11.04</td>
<td></td>
</tr>
<tr>
<td>TSPD</td>
<td>0.92</td>
<td>0.41</td>
<td>2.24</td>
<td>0.03</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Black Wallaby</td>
<td>Intercept</td>
<td>-64.59</td>
<td>8.02</td>
<td>-8.05</td>
<td>&lt;.001*</td>
<td>---</td>
</tr>
<tr>
<td>BSW</td>
<td>0.73</td>
<td>0.09</td>
<td>8.38</td>
<td>&lt;.001*</td>
<td>60.03</td>
<td></td>
</tr>
<tr>
<td>TVOL</td>
<td>12.46</td>
<td>2.01</td>
<td>6.19</td>
<td>&lt;.001*</td>
<td>6.23</td>
<td></td>
</tr>
<tr>
<td>TVOL2</td>
<td>-0.76</td>
<td>0.13</td>
<td>-5.94</td>
<td>&lt;.001*</td>
<td>27.77</td>
<td></td>
</tr>
<tr>
<td>TSPD</td>
<td>1.88</td>
<td>0.52</td>
<td>3.61</td>
<td>&lt;.001*</td>
<td>5.97</td>
<td></td>
</tr>
<tr>
<td>Common Wombat</td>
<td>Intercept</td>
<td>-64.38</td>
<td>7.18</td>
<td>-8.97</td>
<td>&lt;.001*</td>
<td>---</td>
</tr>
<tr>
<td>WOM</td>
<td>0.85</td>
<td>0.06</td>
<td>13.81</td>
<td>&lt;.001*</td>
<td>75.19</td>
<td></td>
</tr>
<tr>
<td>TVOL</td>
<td>12.01</td>
<td>1.76</td>
<td>6.82</td>
<td>&lt;.001*</td>
<td>5.41</td>
<td></td>
</tr>
<tr>
<td>TVOL2</td>
<td>-0.72</td>
<td>0.11</td>
<td>-6.49</td>
<td>&lt;.001*</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>TSPD</td>
<td>2.3</td>
<td>0.41</td>
<td>5.57</td>
<td>&lt;.001*</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>Koala</td>
<td>Intercept</td>
<td>-60.98</td>
<td>9.01</td>
<td>-6.77</td>
<td>&lt;.001*</td>
<td>---</td>
</tr>
<tr>
<td>KOA</td>
<td>0.51</td>
<td>0.08</td>
<td>6.58</td>
<td>&lt;.001*</td>
<td>53.63</td>
<td></td>
</tr>
<tr>
<td>TVOL</td>
<td>11.75</td>
<td>2.31</td>
<td>5.09</td>
<td>&lt;.001*</td>
<td>10.28</td>
<td></td>
</tr>
<tr>
<td>TVOL2</td>
<td>-0.72</td>
<td>0.15</td>
<td>-4.83</td>
<td>&lt;.001*</td>
<td>30.73</td>
<td></td>
</tr>
<tr>
<td>TSPD</td>
<td>1.66</td>
<td>0.63</td>
<td>2.66</td>
<td>0.01</td>
<td>5.37</td>
<td></td>
</tr>
</tbody>
</table>

All of the predictor variables demonstrated plausible relationships to collision likelihood in the partial dependency plots (Figure 3) and the signs of the coefficients were as expected (Table 5). Further, these response shapes were similar across all seasons for each species except the two possums (Figures S3-S5). Based on analysis of variation (ANOVA), species occurrence had the largest contribution to reduction in deviance for all of the models.
Figure 3: Marginal effects of predictor variables on relative likelihood of collision per species. Each predictor is represented by an 1) overall plot showing relative collision rates for all species and 2) six sub-panels depicting
response shapes with rescaled collision risk (minimum to maximum) for each species. Species symbols are:  

`Eastern Grey Kangaroo';  
`Common Brushtail Possum';  
`Common Ringtail Possum';  
`Black Wallaby';  
`Wombat';  
`Koala'.

The response of collision risk to occurrence likelihood was most similar in shape across species (Figure 3a) compared to traffic volume (Figure 3b) and traffic speed (Figure 3c). Koalas had the highest increases in collision risk at lower values of occurrence likelihood and gradually reduced risk at higher values, while the responses of the other species were closer to linear.

All of the models had highly significant coefficients for both linear and quadratic terms of traffic volume. All of the model fits demonstrated curvilinear responses to AADT with consistent peaks around 3,000-5,000 vehicles day$^{-1}$ for all species except the two possums which peaked between 9,000-12,000 vehicles day$^{-1}$. Between the two possums, collision risk peaked at higher values of AADT for the brushtail possum and did not increase or decrease as rapidly across all values of AADT. Koalas and black wallabies had very similar response shapes and also showed the quickest rise and fall of collision risk between 0-10,000 vehicles day$^{-1}$ compared to all other species.

Collision risk increased faster than linearly with traffic speed for all species except the two possums. For brushtail possums, collision risk increased rapidly for low values of traffic speed and increased slowly for values above 30 km hr$^{-1}$. It should be noted, however, that the predictor of traffic speed was not significant for brushtail possums and had less than 0.1% relative influence on the model fit. As with species occurrence, ringtail possums demonstrated a nearly linear relationship between collision risk and traffic speed. Wombats demonstrated the highest rate of
increase in collision risk for increasing collision speed compared to all other species.

The aggregated predictions of collision likelihood were highest on several sections of the Hume Highway and portions of the Calder Freeway just north of the urban centre of Melbourne (Figure 4). There were 60 road segments that were identified with the highest 0.1% of collision risk for all six species and 3,678 road segments for at least one species.

**Figure 4:** Road segments with summed collision rates for all six species. The entire road network is shown as faint gray lines for context.
4. Discussion

Our study successfully applied a collision risk model framework to six different mammal species in south-east Australia and identified individual species' and cumulative collision risk for all road segments. We acknowledge that our study is not the first to model collision risk for multiple taxa; others have studied mammals (Clevenger et al., 2002; Cserkesz et al., 2013; Jaarsma et al., 2006), reptiles (Gunson et al. 2012; Langen et al., 2012) and mixed taxonomic groups (Clevenger et al., 2002; Garrah et al., 2015; Langen et al., 2009; Litvaitis & Tash, 2008). And like ours, all these studies used mixtures of specific environmental and anthropogenic variables in multivariate models. What our research provides is a useful analytical framework for managers that generalises well and allows ongoing scrutiny and tuning of the working parts (Visintin et al., 2016). For example, if species occurrence is a strong predictor of collisions but the performance of the SDMs are not satisfactory, managers may choose to further improve the sub-models or re-consider how much emphasis is placed on installing mitigation measures that influence animal occurrence. Similarly, spatial autocorrelation, detectability, bias, and other factors affecting model calibration and robustness may be assessed accordingly. Moreover, our methodology can be applied to single road segments or across an entire linear network, and on a single species, taxonomic class or mixed group.

The conceptual framework utilised predictors that connect to clear management objectives - species occurrence on road, traffic volume, and driver behaviour (i.e. traffic speed) - and are also consistent with theories of road ecology (see Forman et al., 2003). The models identified common trends in the relationships between the predictors and collision risk. Collision risk increased consistently with higher likelihoods of species occurrence suggesting the usefulness of
separating out the exposure parameter via SDMs. Roger & Ramp (2009) demonstrated the importance of habitat variables (a surrogate for species occurrence) for modelling collisions. The work was extended to use an SDM for wombats (Roger et al., 2012), indicating higher predicted likelihoods of species occurrence resulted in higher likelihoods of collisions.

Collision risks for ringtail possums increased more slowly with occurrence likelihood than for the other mammal species. This makes ecological sense considering the species spends most of the time elevated (mostly in highly connected urban development via power lines and street trees) and would come into contact with moving vehicles less frequently than ground-dwelling species. However, it is difficult to make direct comparisons between species due to uncertainties in the data. Potential reporting biases (e.g. not reporting smaller species) and errors (e.g. possum was assumed hit by car due to discovery on or near road verge) may result in artificially increased or decreased sample sizes. Of the 6546 collision records from the Wildlife Victoria database used in this study, ringtail possums were the third most reported species (Table 1). Thus, possums could be much more abundant than wombats or wallabies. Grilo et al. (2014) explore collision risk among species based on differences in abundance and other traits that might pre-dispose species to collisions. Studies on reporting bias (Snow et al., 2015) and spatial error (Gunson et al., 2009) in WVC also examine some of these issues in more detail.

Collision likelihoods also increased with higher traffic speeds for all species. With the exception of the brushtail and ringtail possums, the magnitudes of increase were significant, suggesting high vehicle speed is a hazard for wildlife. One plausible explanation is that higher traffic speeds reduce the window of time available to avoid collisions for both wildlife and drivers. Previous work has used simulation analysis to model the relationship of speed and collisions by using
species traits to determine crossing success (Jaarsma et al., 2006). Several other studies have shown similar results on the effects of traffic speed on road mortality of wildlife (e.g. Farmer & Brooks, 2012; Gkritza et al., 2013; Lao et al., 2011; Ramp et al., 2006; Seiler, 2005; Seiler & Helldin, 2006; Sudharsan et al., 2009; van Langevelde et al., 2009).

The consistent quadratic shape observed in the responses to traffic volume for all species (Figure 4b) indicates a possible threshold effect. Varying levels of traffic volume has been shown to affect animal presence on or near roads (Jaeger et al., 2005; Rhodes et al., 2014). Aside from species behavioural responses, traffic volumes above a certain level may affect species abundance due to mortality caused by wildlife-vehicle collisions. Lower abundances result in less animals on the road and, therefore, less collisions. Under this hypothesis, traffic volume is especially problematic for species that are attracted to roads (Forman et al., 2003). Nonetheless, identifying traffic volumes that have high relative collision risks for multiple species may help reduce roadkill. For example, high likelihood of collisions at particular traffic speed and volume levels may warrant further examination to determine potential mitigation such as redistribution or re-routing of vehicles and traffic-calming infrastructure.

As mentioned previously, our model framework is correlative and disregards temporal patterns of collisions. A qualitative comparison of the effects plots revealed similar trends between seasons for all species except ringtail and brushtail possums, however, the small sample sizes used in modelling warrant a cautionary interpretation. None of our species demonstrate migratory patterns and have relatively stable patterns of activity throughout the year. For species that exhibit seasonal patterns of movement, it would be useful to incorporate this information into the collision model (see Neumann et al., 2012), potential by including temporal variation in
occurrence in the species distribution model. When activity of the species varies diurnally, partitioning the temporal component at the hourly scale may be more useful as systematic patterns of collisions exist throughout the day (e.g. Litvaitis & Tash, 2008 or Rhodes et al., 2014). This was beyond the scope of this work as a robust temporal analysis would require time of collision data (we only had data on time of reporting). The expansion of all explanatory variables to a complementary temporal scale would significantly increase the computational requirements. Temporal differences in species movements and lags in collision reporting must also be considered.

The spatial autocorrelation in the occurrence models' residuals was high at small distances and similar for all species. Any spatial correlation or patterning in the residuals violates statistical independence assumptions and several methods exist in the SDM literature to deal with this issue (Augustin et al., 1996; Dormann et al., 2007 & 2013). Patterning and higher values of spatial autocorrelation are plausible in certain ecological data and for some species. For example, correlations in model residuals based on the locations of koalas may be expected as they are selective in their choice of tree species, which are often found in homogeneous patches. However, high values of spatial autocorrelation or patterning should be carefully reviewed and appropriately addressed when used in statistical models (Wintle & Bardos, 2006). We included covariates to represent both sampling bias and the spatial autocorrelation arising from such in our models. Distance to road was selected by all species models and distance to town was selected by all models except brushtail possum and koala. Moreover, boosted regression trees are equipped to partially address spatial autocorrelation by reducing bias and variance (Elith et al., 2008).
We did not find significant spatial autocorrelation or patterning in the collision model residuals and, thus, we did not choose to address it as part of the scope of this project. As mentioned previously, our model framework easily allows for this correction by the inclusion of auto- covariates in the model specification (e.g. Dwyer et al., 2016; Farmer & Brooks, 2012; Gomes et al., 2008), or by selecting alternative statistical modelling methods (Zhang et al., 2005).

4.1 Conclusions & implications for management

Our study analysed reported vehicle collisions for six terrestrial mammal species in Australia to test the generality of a model framework and to help infer cause and effect. Species occurrence modelling is useful to determine wildlife exposure to risk (on a relative scale). Species occurrence, traffic volume and traffic speed, as single modelled predictors, allow managers to improve models as new data becomes available. Managers may also predict effects by varying predictor values that reside in their authoritative control such as species occurrence or speed limit; here, the model framework functions as a decision planning tool. For example, the relationship between species occurrence and collision risk was similar for the six species in our study. Thus, managers may simulate changes in expected collision risk resulting from fencing parts of the road network by reducing values of occurrence on road segments in the model training data and predicting collision rates with the model. Further, the model predictions may also be spatially aggregated to suggest areas on the road network that require general mitigation or further study. In the case of Victoria, high-speed freeways and highways north of Melbourne had the highest overall collision risks. Our framework may also be extended to explore additional threats and predict risk to other taxonomic groups and vulnerable species (i.e., focus of conservation efforts) arising from the use of linear infrastructure.
Supporting Information

Data Accessibility

Model Datasets - Archived on Github

R Code - Archived on Github
Acknowledgements

James Johnson, on behalf of Wildlife Victoria, provided the collision data used for the study. Nick Golding advised on statistical methods and provided useful feedback on manuscript revisions. We also received insightful comments from two anonymous reviewers. This project was supported by a University of Melbourne International Research Scholarship and the Australian Research Council Centre of Excellence for Environmental Decisions. Rodney van der Ree was supported by the Baker Foundation.
References


Conservation Biology 16: 503-514


Gunson, K. E., Ireland, D. & Schueler, F. (2012) A tool to prioritize high-risk road mortality locations for wetland-forest herpetofauna in southern Ontario, Canada. North-Western Journal of Zoology 8: 409-413


victims related to road, traffic, vehicle, and species characteristics. Transportation Research Part D: Transport and Environment 11: 264-276


collisions. Biological Conservation, 145: 70-78


Figure S1: Marginal effects plots of three most influential predictor variables on relative likelihood of occurrence per species.
**Figure S2:** Spatial autocorrelation in occupancy models residuals for each species grouped by distance between observations. Moran’s I measures the amount of similarity between values at each spatial distance grouping - zero
indicating no correlation and one indicating perfect correlation.
**Figure S3:** Marginal effects of species occurrence on relative likelihood of collision per species by season.
Figure S4: Marginal effects of traffic volume on relative likelihood of collision per species by season.
Figure S5: Marginal effects of traffic speed on relative likelihood of collision per species by season.
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Visintin, C; van der Ree, R; McCarthy, MA

Title:
Consistent patterns of vehicle collision risk for six mammal species

Date:
2017-10-01

Citation:
Visintin, C; van der Ree, R; McCarthy, MA, Consistent patterns of vehicle collision risk for six mammal species, JOURNAL OF ENVIRONMENTAL MANAGEMENT, 2017, 201 pp. 397 - 406

Persistent Link:
http://hdl.handle.net/11343/216849

File Description:
Accepted version