Consistent patterns of vehicle collision risk for six mammal species

Casey Visintin1, Rodney van der Ree2, Michael A. McCarthy3

1Quantitative and Applied Ecology Group, School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia - Email: cvisintin@student.unimelb.edu.au

2Australian Research Centre for Urban Ecology, Royal Botanic Gardens Victoria and School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia - Email: rvdr@unimelb.edu.au

3Quantitative 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: WVC model with multiple species

Word Count: XXXX  
 Summary: XXXX  
 Main Text: XXXX  
 Acknowledgements: XXXX  
 References: XXXX  
 Tables: XXXX  
 Figure Legends: XXXX

Number of Tables: 5

Number of Figures: 6

Number of References: XXXX

* 1. Summary

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 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 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 26.2 +/-0.6% (common brushtail possum) and 56.6 +/-1.1% (koala). The collision models reduced the unexplained variation in collision event data between 21.3% (koala) and 41.6% (common ringtail possum), 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.

* 1. Keywords

animal, framework, multiple species, risk, species distribution model, speed limit, traffic volume, vehicle, wildlife-vehicle collision

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, Smith & Grilo, 2015). The mortality of wildlife from collisions with vehicles is a significant problem throughout most of the developed world, and will likely become so in the next few decades in the developing world (van der Ree, Smith & Grilo, 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, Kassar & Cook 2008; Huijser, et al., 2009; Rowden, Steinhardt & Sheehan, 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 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. 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 three 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 papers, 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 (Grilo, Bissonette & Santos-Reis, 2009). Although these recommendations are useful in the specific contexts of the studies, they may not generalise to varying spatial scales - specifically large areas. Many analyses are specific to road sections and not easily transferable to other areas, limiting their usefulness as planning tools for environmental management authorities.

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 with predictors that are manageable or suggest areas for more specific examination.

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 that relates risk to exposure (species occurrence) and hazard (traffic volume and speed) (Visintin, van der Ree & McCarthy, 2016). As this model framework is able to quantify risk over large spatial scales (covering a large region of south-east Australia), 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 both the existence and strength of common factors contributing to wildlife-vehicle collisions and 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 & Wierzchowski, 2001), 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.

Materials and Methods

* + 1. STUDY AREA

We used the 227,819 square kilometres 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. Seventy-nine municipal districts control the remaining sealed roads (approx. 125,000 kilometres). Our study combined all sealed roads within the state to analyse collision risk. To organise our spatial data and modelling, we overlaid a spatial grid of one square kilometre 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 655,348 segments as our modelling units for the collision model.

* + 1. 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 vehicle collisions (Croft, 2004; Coulson & Eldridge, 2010) with ungulates found in North America and Europe. They occur in groups and have home range sizes between 500 and 800 hectares (ref). Black swamp 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 (ref). Common wombats (Vombatus ursinus, Shaw) are medium-sized (22 to 39 kilograms), burrowing mammals and one of three wombat species in Australia (ref). 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.

* + 1. CONCEPTUAL MODEL FRAMEWORK

We employed a quantitative risk model (Visintin, van der Ree & McCarthy, 2016) to examine how collision risk of each species relates to occurrence of the species and traffic volume and speed:

|  |  |
| --- | --- |
|  | (1) |

where is the probability of a collision occurring, is species occurrence, is traffic volume, is traffic speed, in a given place .

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, van der Ree & McCarthy, (2016).

* + 1. SPECIES OCCURRENCE MODELS

We obtained observation records of the six study species from the Victorian Biodiversity Atlas (VBA) that satisfied the following criteria: survey date between 1 January 2000 and 31 December 2013 and spatial coordinate certainty of ≤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 spatially unique records of occurrence to represent presences across the study area. As we did not have access to recorded absence data, we generated 10,000 randomly sampled points across the study area as background data and combined them with presence data (Table 1). For each species, we fit SDM to the data using tools and methods by Elith & Leathwick (2009). We chose Boosted Regression Trees (BRT) (Friedman, 2002) as the statistical method. BRT fit complex non-linear relationships and automatically incorporate interaction effects between predictors (Elith, Leathwick & Hastie, 2008). 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).

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. The Pearson correlation coefficients were less than 0.70 for all pairs of explanatory variables with the exception of elevation and slope (0.71), which is expected to exhibit some collinearity, and tree density and precipitation of the driest month (0.72). 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).

Using the model fits, we predicted relative likelihood of occurrence from the species distribution models to the 1-km spatial grids for six species (Figure 2) and 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 plotted against distance in one kilometre bins (Figure S2).

* + 1. 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, including records with a spatial accuracy <300 metres. 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'.

For each target species, we selected the road segments that intersected with the reported location of each collision and re-coded them with ones (Table 1). We randomly selected twice the number of 'collision' road segments in the study area, coded them with zeros to represent background data, and combined them with the 'collision' records. We sampled the relative species occurrence predictions at mid-points of the road segments in the collision dataset by intersecting the road features with the spatial grid. Traffic modelling in previous work had already predicted traffic speed and volumes for each road segment (Visintin, van der Ree & McCarthy, 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 (complementary log-log link) collisions on the predictor variables and added a quadratic term for traffic volume to test for road avoidance effects as some species are repelled by roads when traffic volume increases beyond a threshold value (Seiler, 2005; Gagnon et al., 2007). Our final model is expressed as:

|  |  |
| --- | --- |
|  | (2) |

where is the relative likelihood of a collision occurring, is species occurrence, is traffic volume, and is traffic speed in a given grid cell .

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) curve scores (Metz, 1978) to assess the ability of the models to correctly predict collision events.

We joined the collision risk predictions for each species with all road segments resulting in a dataset of 655,348 road segments, each with a unique identifier and relative collision likelihood value for each of the six species. We summed the total number of species with collision likelihoods equal to or greater than .75 for each segment, resulting in all road segments classified as zero through six. These results were mapped to highlight road segments with high risks for multiple species (Figure 5).

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 there were any major deviations from expected patterns based on season (Figure S1).

* 1. 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 26.2% and 56.6% and the ROC scores varied between 0.90 and 0.96 (Table 4). The best occurrence models were for the koala and common wombat, each reducing deviance on the null models by 56.6% and 45.2%, respectively.

Only four variables were not selected in every species model and the least selected variable was distance to major waterway (Table 3). The four most influential predictors were annual range of temperature, artificial light, precipitation in the driest month, and seasonal change in vegetation greenness (Figure 3). 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 twenty kilometres and correlation values exceeded 0.4 at distances of two kilometres or less for all species except the ringtail possum (Figure S2). The highest overall spatial autocorrelation was in the wombat model residuals; however, both the wombat and koala models demonstrated high levels of spatial autocorrelation at small distances.

All of the collision models fit the data as expected. Reduction on the null deviance was between 21.3% and 41.6% for all models, with ROC scores of 0.83-0.91 when comparing observations with fitted values (Table 4). At least three predictor variables (total of four) were highly significant (p < .001) for each species in the collision models.

All of the predictor variables demonstrated plausible relationships to collision likelihood in the partial dependency plots (Figure 4) and the signs of the coefficients were as expected. For all species, collision risk increased with predicted occurrence of the species, and with traffic speed. The smallest increases were for the two possum species; ringtail possums had very small increases in collision risk at high values of predicted occurrence and traffic speed. Traffic speed was not a highly significant variable for all three arboreal mammal species (koala and two possums) (Table 4).

Occurrence of species had a weaker effect than traffic speed in all of the models. The strongest effect of traffic speed was for the wombat model, which had a regression coefficient for speed of 3.036; meaning each one kilometre/hour increase in traffic speed increased the probability of a collision event by nearly 2000% (exp(3.036) = 20.82; this is the multiplication factor for traffic speed values so (20.82-1)x100=1982%) with all other variables held constant. In contrast, the regression coefficient for occurrence was 0.62; meaning an increase in the probability of a collision event of approximately 8.6% (exp(.62)=1.86 so (.186-.1)x100=8.6%) for each unit (one-tenth of a probability scale between 0 and 1) increase in species occurrence.

All of the models had highly significant coefficients for both linear and quadratic terms of traffic volume. All of the species model fits demonstrated curvilinear responses to AADT with a consistent peak around 3000-4000 vehicles day−1 for all species except the two possums. Collision risk peaked at higher values of AADT for the ringtail possum and did not increase or decrease as rapidly across all values of AADT for the brushtail possum. Collision risk increased rapidly with traffic volumes up to approximately 2000 vehicles day-1 and then slowly increased or was relatively constant for all species except the brushtail and ringtail possums where collision risk increased more gradual with traffic volume (up to 10000 vehicles day−1, or more, in the case of brushtail possums).

All of the collision models attained high ROC scores when predictions were compared to independent validation data (recorded in the year 2014). The lowest score was 0.81 and the highest was 0.90 which fall within the range of the original ROC values from the models fit on the training data.

* 1. Discussion

Our study successfully applied a collision risk model framework to six different mammal species in south-east Australia and identified individual and cumulative species 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 & Wierzchowski, 2001; Cserkesz et al., 2013; Jaarsma et al., 2006), reptiles (Gunson et al. 2012; Langen et al., 2012) and mixed taxonomic groups (Clevenger et al., 2001; 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, van der Ree, McCarthy, 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 on the roads. 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) - that 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, Bino & Ramp, 2012), indicating higher predicted likelihood of species occurrence resulted in higher likelihood of collision.

Collision risks for two arboreal mammals (possums) increased more slowly with occurrence likelihood than for the other mammal species. This makes ecological sense considering the species’ spend 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 between 2010 and 2013 from the Wildlife Victoria database used in this study, brushtail and ringtail possums are the second and third most reported species, respectively (Table 1). Thus, the possums could be much more abundant than wombats or wallabies. This suggests a potential avenue for further research – determining how much of the difference in collision risk among species is attributable to differences in abundance, and how much is attributable to other traits of the species that might pre-dispose the species differently to a collision. Studies on reporting bias (Snow et al., 2015) and spatial error (Gunson, et al., 2009) in WVC 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. Moreover, all of the non-arboreal species had highly significant traffic speed variables and respective coefficient values that were much higher than all other variables. Several other studies have reached similar conclusions 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, A.,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 4) indicates a possible threshold effect. Roads have been shown to affect animal presence on or near roads at varying levels of traffic (Jaeger et al., 2005; Rhodes et al., 2014). 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 and re-routing of vehicles and traffic-calming infrastructure.

Our model framework is correlative and disregards temporal patterns of collision. A qualitative comparison of the effects plots revealed similar trends for all species between seasons suggesting at most small seasonal differences in collision risk (Figure S1). 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); however, this was beyond the scope of this work. A robust temporal analysis would require the expansion of all explanatory variables to a complementary temporal scale, significantly increasing the computational requirements. Moreover, temporal differences in species movements must 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 several covariates to represent both sampling bias and the spatial autocorrelation arising from such in our models. Both distance to town and distance to road were selected by all species models (with the exception of brushtail possums) and explained between two and nine percent of the deviance. Moreover, boosted regression trees are also equipped to partially address spatial autocorrelation by reducing bias and variance (Elith, Leathwick & Hastie, 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).

APPLIED MANAGEMENT IMPLICATIONS

Our study analyses 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 speed limit; here, the model framework functions as a decision planning tool. For example, the relationship between vehicle speed and collision risk was similar for the six species in our study. Moreover, the model predictions may also be spatially aggregated to suggest areas on the road network that require mitigation or further study. In the case of Victoria, high-speed freeways and highways north of Melbourne had high risks for multiple species. 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.

* 1. Data Accessibility

Model Dataset - Archived on Github

R Code - Archived on Github

* 1. Acknowledgements

James Johnson, on behalf of Wildlife Victoria, queried and provided the collision statistics used for the study. 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

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2015). An introduction to management science: quantitative approaches to decision making. Cengage learning.

Augustin, N.; Mugglestone, M. & Buckland, S. (1996) An autologistic model for the spatial distribution of wildlife. Journal of Applied Ecology. 33: 339-347

Beckmann, J. P.; Clevenger, A. P.; Huijser, M.; Hilty, J. & Ament, R. (2010) Safe passages: highways, wildlife, and habitat connectivity. Island Press Washington, DC,

Bissonette, J. A., Kassar, C. A., & Cook, L. J. (2008). Assessment of costs associated with deer-vehicle collisions: human death and injury, vehicle damage, and deer loss. *Human-Wildlife Interactions*, 61.

Clevenger, A. P. & Wierzchowski, J. (2001) GIS-based modeling approaches to identify mitigation placement along roads. ICOET 2001 Proceedings: A Time for Action

Coulson, G. & Eldridge, M. D. B., (2010). *Macropods: The Biology of Kangaroos, Wallabies, and Rat-kangaroos*. CSIRO Publishing, Collingwood, VIC, Australia.

Croft, D. B. (2004). Kangaroo management: individuals and communities.. *Australian Mammalogy* 26 : 101-108.

Cserkesz, T.; Ottlecz, B.; Cserkesz-Nagy, A. & Farkas, J. (2013) Interchange as the main factor determining wildlife-vehicle collision hotspots on the fenced highways: spatial analysis and applications. European Journal of Wildlife Research *59:* 587-597

DELWP (2015) Department of Environment, Land, Water & Planning Victorian Biodiversity Atlas fauna records (unrestricted) for sites with high spatial accuracy. Accessed December, 2015, URL https://www.data.vic.gov.au/data/dataset/victorian-biodiversity-atlas-fauna-records-unrestricted-for-sites-with-high-spatial-accuracy.

Dormann, C. F.; McPherson, J. M.; Araújo, M. B.; Bivand, R.; Bolliger, J.; Carl, G.; Davies, R. G.; Hirzel, A.; Jetz, W.; Kissling, D. W. & others (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* 30 : 609-628

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., et al. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36(1) : 27–46

Elith, J. & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40 : 677-697.

Elith, J.; Leathwick, J. R. & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology* 77 : 802-813.

Farmer, R. G. & Brooks, R. J. (2012). Integrated risk factors for vertebrate roadkill in southern Ontario. The Journal of Wildlife Management 76: 1215 - 1224

Firth D (1993). Bias reduction of maximum likelihood estimates. *Biometrika* 80, 27–38.

Forman, R. T. T. & Alexander, L. E. (1998). Roads and their major ecological effects. *Annual Review of Ecology and Systematics* 29 : 207-231.

Forman, R. T. T., et al. (2003). Road ecology: Science and solutions. Island Press, Washington, D.C., USA.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis* 38 : 367 - 378.

Gagnon, J. W.; Theimer, T. C.; Dodd, N. L.; Boe, S. & Schweinsburg, R. E. (2007). Traffic volume alters elk distribution and highway crossings in Arizona. *The Journal of Wildlife Management* 71 : 2318-2323.

Garrah, E.; Danby, R. K.; Eberhardt, E.; Cunnington, G. M. & Mitchell, S. (2015). Hot spots and hot times: wildlife road mortality in a regional conservation corridor. Environmental management. 56: 874-889

Gkritza, K.; Souleyrette, R. R.; Baird, M. J. & Danielson, B. J. (2013). Empirical Bayes approach for estimating urban deer-vehicle crashes using police and maintenance records. Journal of Transportation Engineering. 140

Grilo, C.; Bissonette, J. A. & Santos-Reis, M. (2009). Spatial--temporal patterns in Mediterranean carnivore road casualties: Consequences for mitigation  
Biological Conservation *142:* 301-313

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

Gunson, K. E.; Mountrakis, G. & Quackenbush, L. J. (2011). Spatial wildlife-vehicle collision models: A review of current work and its application to transportation mitigation projects  
Journal of Environmental Management 92: 1074 - 1082

Gunson, K. E.; Clevenger, A. P.; Ford, A. T.; Bissonette, J. A. & Hardy, A. (2009). A comparison of data sets varying in spatial accuracy used to predict the occurrence of wildlife-vehicle collisions. *Environmental Management* 44 : 268-277.

Huijser, M. P., J. W. Duffield, A. P. Clevenger, R. J. Ament, and P. T. McGowen. (2009). Cost–benefit analyses of mitigation measures aimed at reducing collisions with large ungulates in the United States andCanada; a decision support tool. *Ecology and Society* 14(2): 15

Jaarsma, C. F.; van Langevelde, F. & Botma, H. (2006). Flattened fauna and mitigation: Traffic victims related to road, traffic, vehicle, and species characteristics. Transportation Research Part D: Transport and Environment. 11: 264 - 276

Jaeger, J. A.; Bowman, J.; Brennan, J.; Fahrig, L.; Bert, D.; Bouchard, J.; Charbonneau, N.; Frank, K.; Gruber, B. & von Toschanowitz, K. T. (2005). Predicting when animal populations are at risk from roads: an interactive model of road avoidance behavior. Ecological Modelling. 185: 329-348

Langen, T. A.; Gunson, K. E.; Scheiner, C. A. & Boulerice, J. T. (2012). Road mortality in freshwater turtles: identifying causes of spatial patterns to optimize road planning and mitigation. Biodiversity and conservation. 21: 3017-3034

Langen, T. A.; Ogden, K. M. & Schwarting, L. L. (2009). Predicting hot spots of herpetofauna road mortality along highway networks. The Journal of Wildlife Management. 73: 104-114

Lao, Y.; Wu, Y.-J.; Corey, J. & Wang, Y. (2011). Modeling animal-vehicle collisions using diagonal inflated bivariate Poisson regression. Accident Analysis & Prevention. 43: 220 - 227

Litvaitis, J. A. & Tash, J. P. (2008). An approach toward understanding wildlife-vehicle collisions. *Environmental Management* 42 : 688-697.

Metz C.E. (1978) Basic principles of ROC analysis. *Seminars in Nuclear Medicine*. 8(4) : 283-98.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramp, D.; Wilson, V. K. & Croft, D. B. (2006). Assessing the impacts of roads in peri-urban reserves: Road-based fatalities and road usage by wildlife in the Royal National Park, New South Wales, Australia Biological Conservation, Elsevier. 129: 348-359

Rhodes, J.; Lunney, D.; Callaghan, J. & McAlpine, C. (2014). A Few Large Roads or Many Small Ones? How to Accommodate Growth in Vehicle Numbers to Minimise Impacts on Wildlife. PLoS One: 9, e91093

Roger, E.; Bino, G. & Ramp, D. (2012). Linking habitat suitability and road mortalities across geographic ranges. *Landscape Ecology* 27 : 1167-1181.

Roger, E. & Ramp, D. (2009). Incorporating habitat use in models of fauna fatalities on roads. *Diversity and Distributions* 15 : 222-231.

Rowden, P., Steinhardt, D., & Sheehan, M. (2008). Road crashes involving animals in Australia. *Accident Analysis & Prevention*, 40(6), 1865-1871.

Seiler, A. (2005). Predicting locations of moose–vehicle collisions in Sweden. *Journal of Applied Ecology* 42 : 371-382.

Seiler, A. & Helldin, J. O. (2006). *Mortality in wildlife due to transportation*. In: Davenport, J. & Davenport, J. L. (Ed.), *The Ecology of Transportation: Managing Mobility for the Environment*, Springer.

Snow, N. P.; Porter, W. F. & Williams, D. M. (2015). Underreporting of wildlife-vehicle collisions does not hinder predictive models for large ungulates. Biological Conservation. 181: 44 - 53

Spellerberg, I. A. N. (1998). Ecological effects of roads and traffic: a literature review. *Global Ecology and Biogeography* 7 : 317-333.

Sudharsan, K.; Riley, S. J. & Campa Iii, H. (2009). Relative risks of deer--vehicle collisions along road types in southeast Michigan. Human Dimensions of Wildlife. 14: 341-352

van der Ree, R.; Smith, D. J. & Grilo, C., (2015). *Handbook of Road Ecology*. John Wiley & Sons.

van Langevelde, F.; van Dooremalen, C. & Jaarsma, C. F. (2009). Traffic mortality and the role of minor roads. Journal of Environmental Management. 90: 660-667

Wintle, B. & Bardos, D. (2006). Modeling species-habitat relationships with spatially autocorrelated observation data. Ecological Applications. 16: 1945-1958

Zhang, L.; Gove, J. H. & Heath, L. S. (2005). Spatial residual analysis of six modeling techniques. Ecological Modelling 186: 154-177

* 1. Tables

Table 1: Top six species most frequently reported in wildlife-vehicle collisions in the Wildlife Victoria database (WVD). Species observations from the Victorian Biodiversity Atlas are 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 randomly-generated background (B) observations.

|  |  |  |  |
| --- | --- | --- | --- |
| Species name | Species code | n species observations (2000-2014) | n recorded collisions (2010-2014) |
| Eastern Grey Kangaroo | EGK | 10619 : 858P / 9761B | 3977 : 1344P / 2633B |
| Common Brushtail Possum | BTP | 11060 : 1299P / 9761B | 794 : 257P / 537B |
| Common Ringtail Possum | RTP | 10762 : 1001P / 9761B | 685 : 212P / 473B |
| Black Swamp Wallaby | BSW | 10739 : 978P / 9761B | 352 : 108P / 244B |
| Common Wombat | WOM | 10447 : 686P / 9761B | 461 : 156P / 305B |
| Koala | KOA | 10622 : 861P / 9761B | 277 : 81P / 196B |

Table 2: Predictor variables used in species occurrence (SO) models and collision risk (CR) models. Spatial coordinates for species presences and psuedo-absences were used to sample from predictor variable grids at one square kilometre resolution for occurrence models. The centroids of road segment with reported collisions were used to sample from one square kilometre occurrence model predictions. Note, reported means and ranges are for entire study area.

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Description | Units | Range : Mean |
| ELEV (SO) | Elevation of terrain above sea level | m | -0.39-1890.66 : 252.83 |
| GRASS (SO) | Mean annual grass growth | kg/ha | 0-15.74 : 5.45 |
| GREEN (SO) | Mean seasonal change in vegetation greenness | – | -0.62-0.68 : 0.21 |
| HYDRODIST (SO) | Distance from major waterway | km | 0-95 : 10.09 |
| HYDROFLOW (SO) | Hydrological flow accumulation | – | 1-32767 : 197.95 |
| ISOTHERM (SO) | isothermality | °C\*10 | 0-53 : 48.40 |
| LIGHT (SO) | Remote-sensed artificial light intensity | – | 0-63 : 1.23 |
| MNTEMPWQ (SO) | Mean temperature of wettest quarter | °C\*10 | -8-162 : 102.18 |
| PRECDM (SO) | Precipitation of driest month | mm | 0-81 : 34.70 |
| PRECSEAS (SO) | Precipitation seasonality | mm | 0-43 : 23.99 |
| ROADDIST (SO) | Distance from roadway | km | 0-36.55 : 2.02 |
| SLOPE (SO) | Slope of terrain | % | 0-41.61 : 2.69 |
| SOILBULK (SO) | Mean soil bulk density at 2.5cm depth | kg/m3 | 0-1365 : 1129.13 |
| SOILEXCH (SO) | Mean soil cation exchange capacity at 2.5cm depth | cmolc/kg | 10-102 : 24.84 |
| SOILSAND (SO) | Mean soil texture fraction sand at 2.5cm depth | % | 23-100 : 55.11 |
| TEMPANRANGE (SO) | Temperature annual range | °C\*10 | 0-289 : 245.10 |
| TOWNDIST (SO) | Distance from town center | km | 0-46.10 : 5.76 |
| TREEDENS (SO) | Tree cover | decimal % | 0.00-1.00 : 0.28 |
|  |  |  |  |
| BSW (CR) | Relative probability of wallaby occurrence. per km2 | – | 0.01-0.98 : 0.18 |
| BTP (CR) | Relative probability of brushtail possum occurrence per km2 | – | 0.01-0.98 : 0.20 |
| EGK (CR) | Relative probability of kangaroo occurrence per km2 | – | 0.01-0.97 : 0.13 |
| KOA (CR) | Relative probability of koala occurrence per km2 | – | 0.01-0.98 : 0.12 |
| RTP (CR) | Relative probability of ringtail possum occurrence per km2 | – | 0.01-0.98 : 0.16 |
| TSPD (CR) | Predicted traffic speed per road segment | km/h | 20.36-107.55 : 61.41 |
| TVOL (CR) | Predicted traffic volume per road segment | vehicles/day | 0.29-47639.76 : 4309.27 |
| WOM (CR) | Relative probability of wombat occurrence per km2 | – | 0.01-0.96 : 0.12 |

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.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Predictor | EGK | BTP | RTP | BSW | WOM | KOA |
| ELEV | 10.44 | 7.96 | 5.19 | 7.8 | 8.06 | 5.69 |
| GRASS | 8.68 | 7.49 | 5.32 | 7.54 | 7.31 | 7.53 |
| GREEN | 7.17 | 9.13 | 8.07 | 8.02 | 6.9 | 9.8 |
| HYDRODIST | --- | 2.62 | --- | --- | --- | 2.03 |
| HYDROFLOW | 2.38 | 3.02 | --- | 2.62 | 3.24 | --- |
| ISOTHERM | 2.86 | 2.17 | 3.14 | 2.56 | 2.65 | --- |
| LIGHT | **14.11** | 9.11 | 11.57 | 7.17 | 7.49 | 3.57 |
| MNTEMPWQ | 4.08 | 4.54 | 6.76 | 4.05 | 4.02 | 3.72 |
| PRECDM | 7.52 | 8.06 | 6.6 | 4.68 | **15.1** | 8.6 |
| PRECSEAS | 3.06 | 4.54 | 4.6 | 4.7 | 4.58 | 12.69 |
| ROADDIST | 4.69 | 6.53 | 7.64 | 8.98 | 5.8 | 5.56 |
| SLOPE | 6.55 | 5.41 | 6.09 | 6.65 | 6.46 | 4.78 |
| SOILBULK | 7.52 | 5.52 | 4.6 | 4.19 | 5.4 | 4.62 |
| SOILEXCH | 2.58 | 2.93 | 3.71 | 3.61 | 3.93 | 2.15 |
| SOILSAND | 4.46 | 2.65 | 3.48 | 3.42 | 3.28 | 2.95 |
| TEMPANRANGE | 4.98 | **12.3** | **12.12** | 8.76 | 7.13 | **13.51** |
| TOWNDIST | 2.26 | --- | 3.19 | 3.52 | 3.5 | 2.55 |
| TREEDENS | 6.67 | 6.02 | 7.91 | **11.75** | 5.14 | 10.26 |

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 SO models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species Name | % Deviance (SO) | ROC (SO) | % Deviance (CR) | ROC (CR) |
| Eastern Grey Kangaroo | 36.0 +/-1.1 | 0.90 +/-0.005 | 24.0 | 0.83 |
| Common Brushtail Possum | 26.2 +/-0.6 | 0.91 +/-0.005 | 29.6 | 0.84 |
| Common Ringtail Possum | 41.2 +/-1.2 | 0.92 +/-0.003 | 41.6 | 0.90 |
| Black Swamp Wallaby | 37.7 +/-1.0 | 0.90 +/-0.004 | 24.9 | 0.83 |
| Common Wombat | 45.2 +/-1.5 | 0.94 +/-0.004 | 41.4 | 0.91 |
| Koala | 56.6 +/-1.1 | 0.96 +/-0.002 | 21.3 | 0.84 |

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.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Species** | **Variable** | **Coefficient** | **Std. Error** | **Z-value** | **Pr(Z)** | **ANOVA** |
| Eastern Grey Kangaroo | Intercept | -36.61 | 2.149 | -17.03 | 2E-16 | --- |
|  | EGK | 0.4764 | 0.02391 | 19.92 | 2E-16\* | 57.49 |
|  | TVOL | 5.846 | 0.4727 | 12.37 | 2E-16\* | 0.038 |
|  | TVOL2 | -0.3488 | 0.02945 | -11.84 | 2E-16\* | 10.52 |
|  | TSPD | 3.087 | 0.1718 | 17.97 | 2E-16\* | 31.95 |
|  |  |  |  |  |  |  |
| Common Brushtail Possum | Intercept | -40.72 | 9.789 | -4.159 | 0.00003 | --- |
|  | BTP | 0.6266 | 0.08664 | 7.232 | 4.7E-13\* | 77.16 |
|  | TVOL | 8.08 | 2.05 | 3.942 | 0.00008\* | 16.36 |
|  | TVOL2 | -0.4363 | 0.1186 | -3.679 | 0.0002\* | 4.892 |
|  | TSPD | 0.9248 | 0.433 | 2.136 | 0.03 | 1.584 |
|  |  |  |  |  |  |  |
| Common Ringtail Possum | Intercept | -123.8 | 27.94 | -4.432 | 0.000009 | --- |
|  | RTP | 0.7908 | 0.09552 | 8.28 | 2E-16\* | 76.07 |
|  | TVOL | 26.73 | 6.115 | 4.372 | 0.00001\* | 11.32 |
|  | TVOL2 | -1.49 | 0.3457 | -4.311 | 0.00002\* | 11.15 |
|  | TSPD | 1.154 | 0.515 | 2.241 | 0.03 | 1.462 |
|  |  |  |  |  |  |  |
| Black Swamp Wallaby | Intercept | -42.27 | 8.025 | -5.268 | 0.0000001 | --- |
|  | BSW | 0.5768 | 0.09598 | 6.009 | 1.9E-09\* | 57.89 |
|  | TVOL | 7.702 | 1.867 | 4.125 | 0.00004\* | 0.031 |
|  | TVOL2 | -0.4686 | 0.1182 | -3.965 | 0.00007\* | 20.14 |
|  | TSPD | 2.802 | 0.6277 | 4.464 | 0.000008\* | 21.95 |
|  |  |  |  |  |  |  |
| Common Wombat | Intercept | -41.18 | 7.809 | -5.273 | 0.0000001 | --- |
|  | WOM | 0.6207 | 0.06801 | 9.126 | 2E-16\* | 79.64 |
|  | TVOL | 7.079 | 1.712 | 4.135 | 0.00004\* | 0.483 |
|  | TVOL2 | -0.418 | 0.1064 | -3.93 | 0.00009\* | 6.217 |
|  | TSPD | 3.036 | 0.5894 | 5.151 | 0.0000003\* | 13.66 |
|  |  |  |  |  |  |  |
| Koala | Intercept | -42.84 | 9.32 | -4.597 | 0.000004 | --- |
|  | KOA | 0.3342 | 0.07314 | 4.569 | 0.000005\* | 55.9 |
|  | TVOL | 8.808 | 2.26 | 3.898 | 0.0001\* | 0.344 |
|  | TVOL2 | -0.557 | 0.146 | -3.815 | 0.0001\* | 27.42 |
|  | TSPD | 2.082 | 0.7081 | 2.941 | 0.003 | 16.34 |
|  |  |  |  |  |  |  |

* 1. Figures

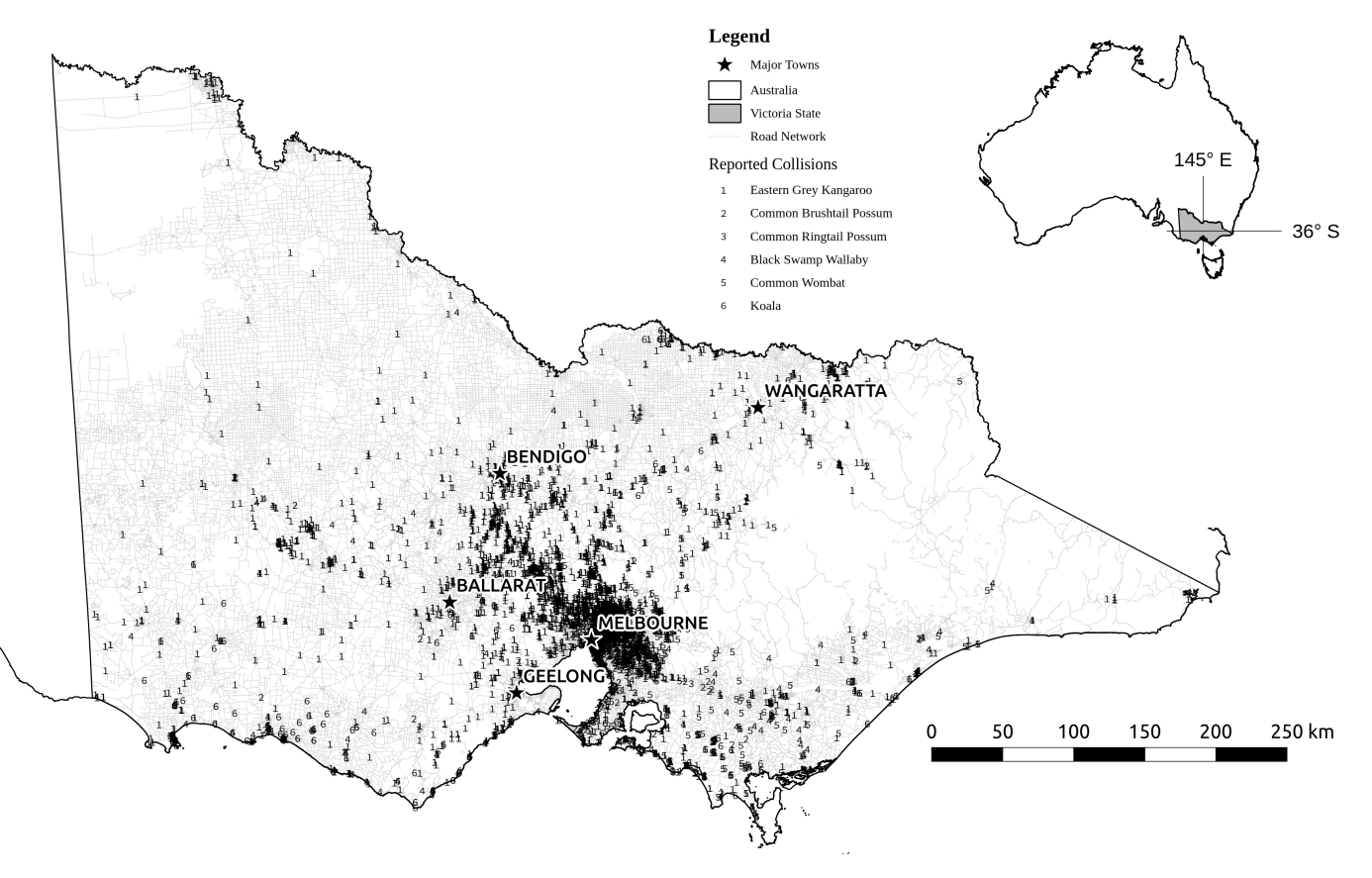
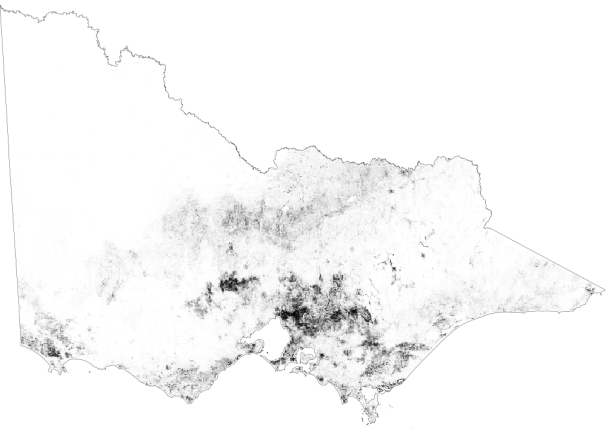


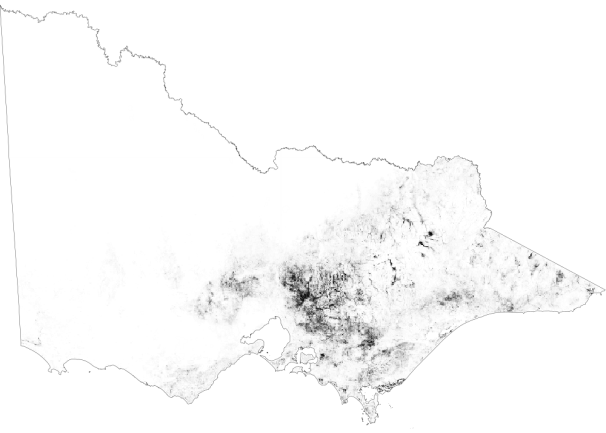
Figure 1: Study area (State of Victoria, in south-east Australia) showing location of wildlife-vehicle collisions. Each number denotes one of six species (see legend).

Eastern Grey Kangaroo Common Brushtail Possum

Common Ringtail Possum Black Swamp Wallaby

Common Wombat Koala

Figure 2: Predicted relative likelihood of occurrence of each species across the State of Victoria. Darker shading indicates higher relative likelihood of occurrence.





Figure 3: Marginal effects plots of four most influential predictor variables on relative likelihood of occurrence per species. Species abbreviations are given in Table 1.







Figure 4: Effects of predictor variables on relative likelihood of collision per species. Species abbreviations are given in Table 1.

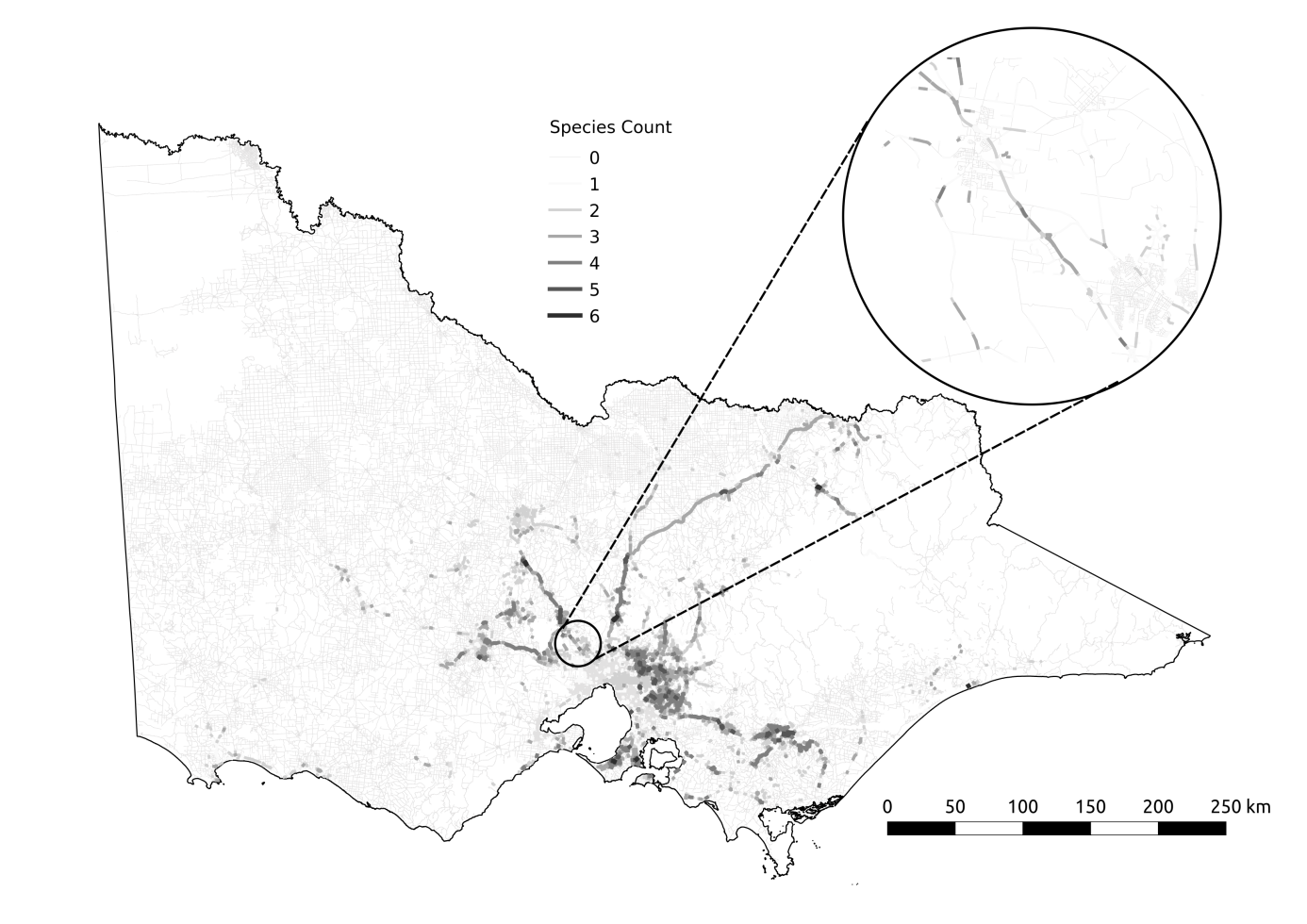


Figure 5: Road segments in which predictions of collision likelihood over 0.75 for at least one of the six species. Lighter, thinner lines represent a few species with high collision likelihood (minimum one) and darker, thicker lines represent more species (maximum six). The entire road network is shown as faint gray lines for context.







Figure S1: Effects of predictor variables on relative likelihood of collision per species by season. Species abbreviations are given in Table 1.

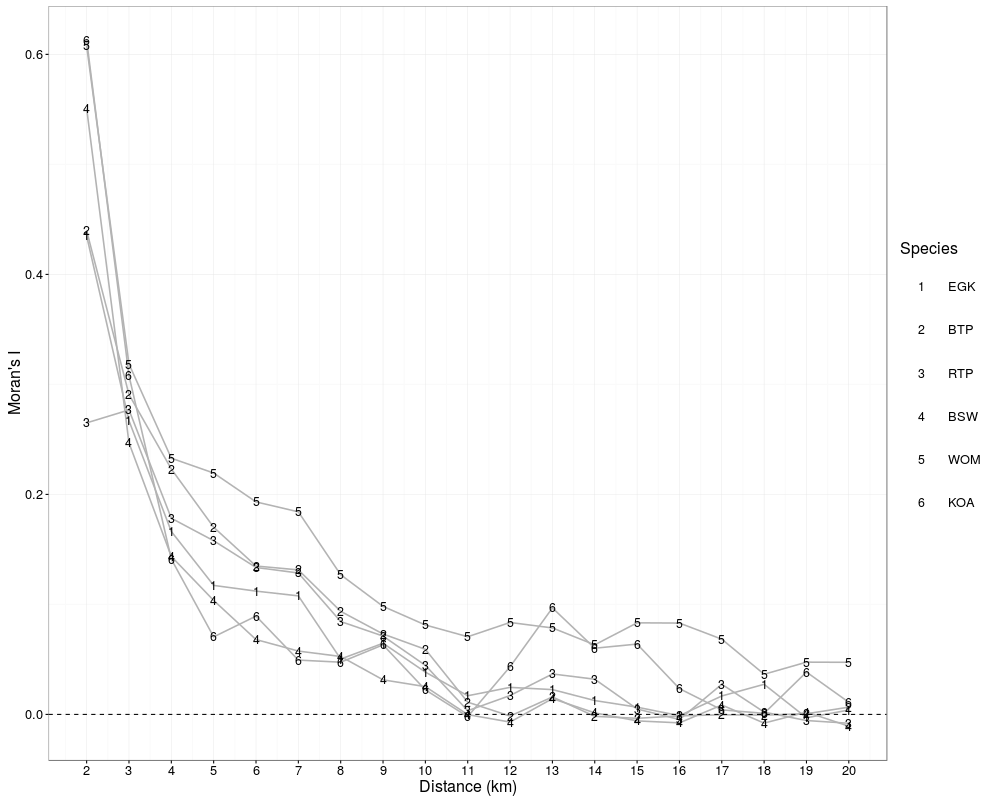


Figure S2: Spatial autocorrelation in occupancy models residuals for each species grouped by distance between observations. Trend lines use numbers to indicate species (see legend); EGK is Eastern grey kangaroo, BTP is Brushtail Possum, RTP is Ringtail Possum, BSW is Black Swamp Wallaby, WOM is Wombat, and KOA is Koala.