Applying a conceptual framework to determine patterns of wildlife-vehicle collisions on two continents: mule deer in North America and grey kangaroos in Australia.

Casey Visintin1, Fraser Shilling2, Rodney van der Ree3, Michael A. McCarthy4

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

2Department of Environmental Science and Policy, University of California, Davis, CA 95616-8576 , U.S.A - Email: fmshilling@ucdavis.edu

3Ecology and Infrastructure International and School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia - Email: rvdr@unimelb.edu.au

4Quantitative 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 modelling on two continents

Word Count: 5964

Summary: 294

Main Text: 3955

Author's Contributions: 68

Acknowledgements: 65

References: 1264

Table Legends: 125

Figure Legends: 193  
Number of Tables: 3

Number of Figures: 7 (3 supplemental)

Number of References: 49

* 1. Summary

1. Wildlife-vehicle collisions (WVC) are an unfortunate consequence of human development and activity. The high costs and safety issues mean WVC are well studied in developed nations (e.g. United States, European countries, Australia), with mitigation strategies aiming to reduce WVC by influencing human behaviour or animal behaviour. We use a conceptual model framework to determine collision risk in USA and Australia. Despite being different species occurring in different contintents, we examine whether collision risk correlates similarly with animal behaviour (occurrence of the species) and collisions risk and human behaviour (traffic volume and speed).

2. We test our framework in two geographic locations: the State of Victoria in south-east Australia and a section of central California in the USA. We trained our collision models with spatially unique, citizen-science collected, collision/carcass record coordinates for each species. We used a very different set of data, police records of road accidents, to validate our model predictions. Relative collision risk is predicted to all road segments in each study area.

3. The kangaroo and deer occurrence models fit the data as expected and produced ecologically-reasonable predictions about the relative likelihood of species presence across each study area. The predictions for relative collision risk on all road segments were plausible throughout most of the two study areas.

4. Marginal effects of each predictor on relative collision risk were similar for both species. All variables were highly significant in the collision models. Speed had the highest relative importance followed by species occurrence.

5. *Synthesis and applications* Our analysis suggests appropriate mitigation for both species may include reduced speeds or fencing on road segments where species occurrence is predicted to be high. Our methods enable managers to conduct sensitivity analyses and calculate overall reductions in expected collisions based on applying mitigation strategies on different road segments.

* 1. Keywords

animal, framework, multiple species, risk, species distribution model, speed limit, traffic volume, vehicle, WVC, deer, kangaroo

Introduction

Wildlife-vehicle collisions (WVC) are an unfortunate outcome of human development and activity. Approximately one-million vertebrates are killed per day in the United States (Forman & Alexander, 1998) and up to 27-millions birds are killed annually in select European countries (Erritzoe J., Mazgajski T. D., Rejt L., 2003). Costs of WVC, including vehicle repair, human medical costs and the value of a human life, are estimated at over eight-billion dollars annually in the United States (Huijser et al., 2007) with deer collisions comprising the largest share (see Bissonette, Kassar & Cook, 2008). These costs and impacts will continue to worsen globally as new roads are constructed and existing roads are upgraded in developing countries; increasing five-fold in developing countries and doubling of worldwide traffic volume by 2050 (van der Ree, Smith & Grilo, 2015).

Mitigation strategies aim to reduce the rate and severity of WVC by influencing either human or animal behaviour (see Huijser & McGowen, 2010). Traffic planning (e.g. the magnitude and routing of vehicles), speed control (e.g. signage, traffic calming mechanisms and penalty enforcement), and education/training can influence the operation of vehicles on roads. Exclusion (e.g. fencing) and re-routing, either over or under the road, can influence the presence and movement of wildlife on the roads. Other strategies involve visual or auditory discouragement such as flashing lights and ultra-sonic whistles, however, these are largely ineffective (Reeve & Anderson, 1993; Bender, 2003; Scheifele, Browning & Collins-Scheifele, 2003; Ramp & Croft, 2006). Choices to implement mitigation strategies are influenced by an understanding of the factors contributing to WVC in a particular area.

Decisions about where and when to mitigate are important because both employing a strategy and damages resulting from employing an ineffective strategy entail costs. To address this issue, many studies examine patterns of collisions and related variables to inform management, often involving spatial modelling and prediction (see Gunson, Mountrakis & Quackenbush, 2011). Most research is specific to an area (i.e. section of road) or problem and not easily transferable due to variation in scale, geographic location and/or species traits, and confounding effects of explanatory variables. Although these issues cannot be totally eliminated - partly due to the uncertainty present in ecological systems - it is possible to develop analytical frameworks to assist managers to deal with uncertainty and make inferences regarding wildlife-vehicle collision risk (see Visintin et al., 2016).

We tested the general applicability of a previously developed conceptual model framework for predicting collision risk in California and Victoria. We had two aims; first we were interested in the applicability of our methods to planning for WVC mitigation irrespective of geographical location or scale (see van der Ree et al., 2011). We propose that our analytical methods will supplement management efforts by allowing clearer inferences about the factors related to collision risk, identifying potential biases and uncertainties in the analysis (thereby suggesting further research or data collection), and suggesting appropriate mitigation for WVC. Second, collisions of vehicles occur at a similar frequency for both kangaroos and deer, have similar impacts (Langley & Mathison, 2008), and are subject to similar management practices (McShea & Underwood, 1997; Croft, 2004). So while models of collision risk should differ between species because of the different contexts, we expect that our previous framework should reveal similar patterns of collision risk for the two species as a function of traffic attributes and species occurrence. To our knowledge, this work represents the first study to test a broadly applicable model for two analogous species on different continents for the purpose of satisfying contrasting needs of management.

* 1. Materials and Methods
     1. *Study Areas*

We use two unique geographic locations to test our framework; the 227,819 km2 State of Victoria in south-east Australia and a 146,478 km2 section of central California in North America (Figure 1). Each locale has a single agency responsible for the major roads (e.g. high-speed, high-capacity, or both) while all other roads are managed by municipal districts. We predict collision risk for all sealed roads within the study areas. To organise our spatial data and modelling, we overlay spatial grids of one km2 resolution on the study area. Each grid cell is used as the modelling unit for species occurrence. All roads in the study area are bisected by the grid resulting in road segments that are approximately one kilometre or less in length. We use 655,348 road segments for Victoria and 646,705 road segments for California as our modelling units for each respective collision model.

* + 1. *Data Preparation*

We selected one species from each study area that are frequently involved in WVC . Eastern grey kangaroos (Macropus giganteus, Shaw) are the second largest mammal in Australia - up to 85 kilograms for males (Van Dyck & Strahan, 2008) - and share management issues with ungulates found in North America and Europe (Croft, 2004; Coulson & Eldridge, 2010). Mule deer (*Odocoileus hemionus*) are common across western North America and adults range in size up to 120 kilograms (Kays & Wilson, 2009).

To train our collision models, we obtained spatially unique collision/carcass records for each species from two citizen-science databases; the Wildlife Victoria Database, or WVD, (Wildlife Victoria, 2016) for kangaroos, and the California Roadkill Observation System, or CROS, (see Shilling & Waetjen, 2015) for deer. The collision/carcass records spanned a six year period between 1 January, 2010 and 1 January, 2016 for kangaroos and a ten year period between 1 June, 2006 and 1 June, 2016 for deer. We limited our datasets to records with associated global positioning system (GPS) coordinates. Citizen-science sources of data can be reliable and robust (see Paul, et al., 2014). Kangaroo records were initially inspected for spatial accuracy and we only retained observations with less than three-hundred metre error. Deer records from CROS have greater than ninety-nine percent species identification accuracy and less than one-hundred metre spatial error (Shilling and Waetjen, unpublished observations).

For each target species, we selected road segments that intersected with reported species’ collision records and coded them with ones. We coded all other road segments in each study area with zeros, to represent background data, and combined them with the collision record segments. After removing spatial duplicates, there were 4,245 presence and 640,470 background points in the kangaroo collision modelling dataset and 933 presence and 644,296 background points in the deer collision modelling dataset.

We used police records of road accidents to evaluate our model fits. For each species, we extracted spatially-unique records where the incidents were collisions with our target species. For kangaroos, we used data reported between 1 January, 2010 and 1 October, 2016, for deer, the reporting period was 1 February 2015 to 1 December 2016. Again, we used all road segments as background and coded the segments with collisions as ones and the remainder with zeros. The final validations datasets had 481 presence and 644,234 background points for kangaroo collision model validation and 1,795 presence and 643,434 background points for deer.

* + 1. *Collision Model*

We used a single-species quantitative risk model (Visintin, van der Ree & McCarthy, 2016) to fit and compare the relationship of species presence and road threat to collision likelihood

We used the open-source software package ’R’ version 3.3.0 (R Development Core Team, 2004) to perform all statistical analyses. For each species, we developed species distribution models (SDMs) to predict occurrence across their respective geographic areas. Traffic volume and speed values for all road segments in both Victoria and California were predicted following methods of Visintin, van der Ree & McCarthy (2016). We regressed annual average daily traffic (AADT) and speed on explanatory variables of distance to anthropogenic development (derived from remotely-sensed land use), distance to highway/freeway, road class, road density within 1 km of each road segment, and population density (from the Australian Bureau of Statistics and the United States Census Bureau, respectively) in random forest models (Brieman, 2001).

Each road segment with an observed collision was coded as one, and all other segments were coded with zeros, with each set of spatial duplicates reduced to a single observation. We used the mid-points of the road segments to sample each species occurrence prediction grids. This resulted in two modelling datasets, 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 with an added quadratic term for traffic volume. Our model is expressed as:

(1)

where is the relative likelihood of a collision occurring, is species occurrence, is traffic volume, is traffic speed, on a road segment .

To test for spatial autocorrelation, we calculated randomised quantile residuals (Dunn & Smyth, 1996) using 5000 simulations from each model fit. We projected the residuals to a standard normal distribution and calculated Moran's I at 20 spatial lags of both one kilometre and 250 metres. For each spatial resolution, we repeated the Moran's I calculation twenty times for 5000 randomly sampled road segments (due to computational limitations) and plotted the twenty trend lines for visual inspection at both spatial scales (Figure 3).

The two models were used to predict relative collision risk of all road segments in each respective study area (Figure 4). We tested the ability of our model to predict independent validation data using a receiver-operator characteristic (ROC) score - a score of one indicating perfect discrimination ability while 0.5 suggesting a performance no better than random (see Metz, 1978).

* 1. Results
  2. All of the predictor variables demonstrated plausible relationships to collision likelihood in the partial dependency plots. The probability of collision for both species changed similarly with occurrence of the species (Figure 2a & b), traffic volume (Figure 5c & d), and traffic speed (Figure 5e & f).

Reduction on the null deviance was 10.5% for kangaroos and 18.6% for deer in the collision model fits (Table 2).

A comparison of the validation observations to the fitted values produced ROC scores of 0.79 for kangaroos and 0.88 for deer (Table 2).

Spatial correlation in the random quantile residuals was low and showed no patterns across distances at two spatial scales (Figure 3).

Speed had the highest relative importance for both kangaroos and deer based on ROC analysis of each predictor. Increasing speed from 80 to 100 km hr-1 approximately doubled the relative collision risk for both species (Figure 2). Collision risk with deer was less than kangaroos at lower speeds and increased at a faster rate at higher speeds. At 110 km hr-1 relative collision risk for kangaroos was more than 150% that of deer.

Species occurrence was the next most influential variable influencing risk of collision for both species. The response shape of collision risk against likelihood of occurrence was similar for both species, however, collision risk was 30% higher for deer at predicted relative occurrences of one (Figure 2).

Relative collision risk peaked near 5000 vehicles day-1 for both species (Figure 2). Collision risk at the peaks relative to baseline (0) were different between species by more than ten orders of magnitude. Kangaroos showed a more drastic decrease in collision likelihood at higher traffic volumes, however, deer had larger confidence intervals around the marginal response curve.

The predictions for relative collision risk across all road segments were plausible in most areas of the two study areas. Victoria has more consistently plausible predictions as the highest likelihoods of collisions were in the northern suburbs of Melbourne which has relatively high rates of collision. In contrast, predictions in California were highest on road segments in both the San Francisco Bay Area (SFBA), which has high relative rates of deer collisions, and along the Interstate 5 and 99 corridors through the Central Valley region, where deer collisions are less observed/reported.

* 1. Discussion

Our results suggest that the proposed conceptual framework has utility irrespective of locality, spatial scale, or species. Species occurrence, traffic volume and speed, and predictor variable data is publicly available for many jurisdictions (see list of online sources in supplemental information), therefore, an analyst only requires data on vehicle collisions with wildlife species to train the models in the conceptual framework, make inferences and predict risk. Moreover, the collection of data on collisions with vehicles by road authorities (via carcass collection operations) or citizen scientists such as wildlife groups is increasing due to human safety or animal welfare concerns and technological innovations (see Olson 2014; Shilling, 2015). Analysis from our framework may guide mitigation actions in several problematic areas to reduce collisions.

Forman (2003) identifies both animal and human behaviour as major drivers for WVC. Our framework operates on this hypothesised relationship and indicates maximum relative risk where animal presence and vehicles moving at speed co-occur in space. Road authorities mitigateby managing human activity (traffic volume and speed) or animal behaviour (occurrence on or movement across the road) and these have different costs and implications. Our results demonstrated that speed is highly significant for both kangaroos and deer and therefore mitigation that either addresses animal presence near high speed road segments (e.g. fencing) or traffic speeds (e.g. controls and enforcement) in collision hotspot areas should be considered. This pattern is consistent with other studies on deer collisions (Gkritza, 2013; Meisingset, 2014; Sudharsan, 2009), kangaroo collisions (Rowden, 2008), and other taxa (Gunson, 2011). Moreover, the traffic models demonstrate plausible fits to the speed data (Victoria: 58.7% variance explained, n=42439; California: 60.2% variance explained, n=7292) therefore suggesting greater reliability in their predictions. However, when assessing where to place mitigation for speed, the analyst and managers should also consider the uncertainty around predictions at each road segment for both speed and animal presence.

The conceptual framework with sub-models allows us to more clearly identify bias and uncertainty in the analysis. For example, the occurrence models are both produced using presence-only data that explicitly assumes perfect detection and includes potential sampling bias. Considering these shortcomings, an analyst may choose to weight their inference, or augment the data and statistical methods (Dorazio, 2014). The deer occurrence predictions may also be less reliable than the roo predictions? due to a smaller number of training observations. Likewise, the lower statistical significance and wide confidence intervals in the marginal effects of traffic volume on collision likelihood for deer may warrant further analysis and, thus, reducing vehicles on problematic road segments may not have the desired effect. Moreover, reported or observed collision data are also subject to the same limitations as species occurrence data although predictions from models with reliable statistical estimates and accuracy are robust to under-reporting (Snow, 2015).

As expected, collision risk for kangaroos and deer were similar in response to species occurrence, traffic volume and traffic speed. Collision risk increased monotonically with increasing species occurrence for both kangaroos and deer, however, there was an order of magnitude greater change for kangaroos, despite both datasets having a similar number of road segments and study period. Assuming species occurrence is of equal importance as the other collision model predictors, the smaller increase in collision risk for deer may be attributed to lower predictions arising from a mis-specified species distribution model (Guillera-Arroita, 2015). Collision risk response to traffic volume was unimodal for both species as expected, however, the change in collision risk was an order of magnitude greater per unit of traffic volume for kangaroos. This response shape can arise from different mechanisms. First, there may be a potential observer bias meaning individuals are less likely to stop or report collisions on high-volume roads out of concern for their own safety or simply not seeing the dead animal. Second, there may be road avoidance effects as some species are repelled by roads when traffic volume increases beyond a threshold value due to high noise loads or small crossing intervals (see van Langevelde, 2004; Seiler, 2005; Gagnon et al., 2007). Our results are consistent with other studies on behavioural responses to roads that include kangaroos and deer (see Jacobson 2016). With the exception of slope variation along the curve, collision risk increased monotonically with increasing speed with a similar magnitude for both species. Kangaroos exhibited higher collision risk at lower traffic speeds than deer. This may be due to Melbourne suburbs expanding into natural kangaroo habitat and therefore human-kangaroo conflicts are prevalent.

We did not explore temporal effects in this study, however, this may be easily incorporated into our framework. One such method is to add a function-based term to the model that allows collision risk to vary as a function of time and season (Visintin et al., manuscript in prep). The functional form may be expressed in relation to the known activity periods of the target species. Kangaroos arecrepuscular, with peaks of activity at dawn and dusk, and deer are either crepuscular or nocturnal depending on the season. Hour of day and time of year are useful predictors of collisions with ungulates (Meisingset, 2014 ; Mountrakis, 2009), however, not yet fully tested for kangaroos.

The inferences made from the conceptual framework will be more robust as more observations are added to train the models. We currently use a mixture of professional and citizen-science collected data - each with unique advantages, disadvantages, and implications for analysis. Citizen-science data is used to supplement data-deficient ecological studies and, more recently, been shown to be effective in road ecology (Paul, 2014). Moreover, citizen-science data is relatively low-cost and can cover large spatial scales. For example, the collision records in both the Wildlife Victoria database and the California Roadkill Observation System are systematically collected and stored in databases and employ mechanisms to elicit data from unrestrained spatial distances. As innovative and standardised data collection techniques are implemented (see Aanensen, 2009; Donaldson, 2010; van der Ree et al?, 2015), inferences from the model framework will improve.

Our analysis suggests appropriate mitigation for both species may include reduced speeds or fencing on road segments where species occurrence is predicted to be high. Although omitted for this work, our methods enable managers to conduct sensitivity analyses and calculate overall reductions in expected collisions based on mixed applications of mitigation on different road segments (assuming the mitigation is 100% effective, but see Huijser, 2009). As costs are often known for specific mitigation measures, simulations can determine optimal uses of resources to maximise reductions of collisions and minimise costs.

* 1. Authors' Contributions

All authors made equally valuable contributions to the manuscript and gave final approval for publication. Casey Visintin analysed the data and led the writing; Casey Visintin and Michael McCarthy conceived the ideas and developed the methodology; Rodney van der Ree, Fraser Shilling and Michael McCarthy provided critical review of and contributions to the manuscript content; Fraser Shilling provided access to the California Roadkill Observation System and California crash data.

* 1. Acknowledgements

Wildlife Victoria provided the collision statistics for Victoria in the study. Data for California were obtained from the Road Ecology Center at UC Davis. Nick Golding provided statistical and computational advice. This project was supported by a University of Melbourne International Research Scholarship and the Australian Research Council Centre of Excellence for Environmental Decisions.

* 1. Supplemental Materials

Model Datasets - Archived on GitHub

R Code - Archived on GitHub

List of online data sources

References

Aanensen, D. M.; Huntley, D. M.; Feil, E. J.; Spratt, B. G. & others (2009) EpiCollect: linking smartphones to web applications for epidemiology, ecology and community data collection *PLoS One*, 4, e6968

Bender, H. (2003) Deterrence of kangaroos from agricultural areas using ultrasonic frequencies: efficacy of a commercial device. *Wildlife Society Bulletin*, 1037-1046

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 , 17-21

Breiman, L. (2001). Random forests. Machine learning. 45: 5-32

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.

Department of Environment, Land, Water & Planning (DELWP) (2016) *Victorian Biodiversity Atlas fauna records (unrestricted) for sites with high spatial accuracy.*

Donaldson, B. & Lafon, N. (2010) Personal digital assistants to collect data on animal carcass removal from roadways *Transportation Research Record: Journal of the Transportation Research Board*, 18-24

Dorazio, R. M. (2014) Accounting for imperfect detection and survey bias in statistical analysis of presence-only data *Global Ecology and Biogeography*, 23, 1472-1484

Dunn, P., & Smyth, G. (1996) Randomized Quantile Residuals. *Journal of Computational and Graphical Statistics*, 5(3), 236-244

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.

Erritzoe, J.; Mazgajski, T. D. & Rejt, L. (2003) Bird casualties on European roads-a review *Acta Ornithologica*, 38, 77-93

Ferguson, K. I. (2005). *Mule deer (Odocoileus hemionus).* Washington, D.C. : Natural Resources Conservation Service.

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.; Sperling, D.; Bissonette, J. A.; Clevenger, A. P.; Cutshall, C. D.; Dale, V. H.; Fahrig, L.; France, R.; Goldman, C. R.; Heanue, K. & others (2003) Road ecology: science and solutions. *Island Press*, Washington, D.C.

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.

Global Biodiversity Information Facility (GBIF). (2016). GBIF data portal

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, 04013002

Guillera-Arroita, G.; Lahoz-Monfort, J. J.; Elith, J.; Gordon, A.; Kujala, H.; Lentini, P. E.; McCarthy, M. A.; Tingley, R. & Wintle, B. A. (2015) Is my species distribution model fit for purpose? Matching data and models to applications *Global Ecology and Biogeography*

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

Huijser, M. P. & McGowen, P. T. (2010) Reducing wildlife-vehicle collisions in  
Safe Passages: Highways, Wildlife, and Habitat Connectivity, Island Press: Washington, DC, USA, 51-74

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 and Canada; a decision support tool. *Ecology and Society* 14(2): 15.

Huijser M.P., McGowen P.T., Fuller J., Hardy A. & Kociolek A. (2007) Wildlife-vehicle collision reduction study: report to Congress. Bozeman: Western Transportation Institute.

Kays, R. W. & Wilson, D. E. (2009). Mammals of North America. Princeton University Press

Jacobson, S. L.; Bliss-Ketchum, L. L.; Rivera, C. E. & Smith, W. P. (2016) A behavior-based framework for assessing barrier effects to wildlife from vehicle traffic volume *Ecosphere*, 7(4), e01345

Langley, R. L. & Mathison, J. (2008) Worldwide characteristics and mitigation strategies for motor vehicle-animal collisions in Traffic accidents: causes and outcomes. Nova Science Publishers, New York , 75-96

McShea, W. J. & Underwood, H. B. (1997) Deer management and the concept of overabundance in The Science of Overabundance: Deer Ecology and Population Management, Smithsonian Institution Press, 1-7

Meisingset, E. L.; Loe, L. E.; Brekkum, Ø. & Mysterud, A. (2014) Targeting mitigation efforts: The role of speed limit and road edge clearance for deer–vehicle collisions The Journal of Wildlife Management, 78, 679-688

Metz C.E. (1978) Basic principles of ROC analysis. Seminars in Nuclear Medicine 8(4): 283-298

Mountrakis, G. & Gunson, K. (2009) Multi-scale spatiotemporal analyses of moose--vehicle collisions: a case study in northern Vermont *International Journal of Geographical Information Science*, 23, 1389-1412

Ng, J. W.; Nielson, C. & St Clair, C. C. (2008) Landscape and traffic factors influencing deer-vehicle collisions in an urban environment. Human-Wildlife Interactions, 79

Olson, D. D.; Bissonette, J. A.; Cramer, P. C.; Green, A. D.; Davis, S. T.; Jackson, P. J. & Coster, D. C. (2014) Monitoring Wildlife-Vehicle Collisions in the Information Age: How Smartphones Can Improve Data Collection PLoS One, 9, e98613

Paul, K.; Quinn, M. S.; Huijser, M. P.; Graham, J. & Broberg, L. (2014) An evaluation of a citizen science data collection program for recording wildlife observations along a highway  
Journal of Environmental Management, 139, 180 - 187

R Core Development Team. (2004). *R: A language and environment for statistical computing*

Ramp, D. & Croft, D. B. (2006) Do wildlife warning reflectors elicit aversion in captive macropods? *Wildlife Research*, 33, 583-590

Reeve, A. F. & Anderson, S. H. (1993) Ineffectiveness of Swareflex reflectors at reducing deer-vehicle collisions *Wildlife Society Bulletin* 21, 127-132

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

Scheifele, P. M.; Browning, D. G. & Collins-Scheifele, L. M. (2003) Analysis and effectiveness of deer whistles for motor vehicles: frequencies, levels, and animal threshold responses. *Acoustics Research Letters Online*, 4, 71-76

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

Shilling, F., Perkins, P. & Collinson, W. (2015) Wildlife/roadkill observation and reporting systems. In: van der Ree R., Smith D., Grilo C. (eds) *Handbook of Road Ecology*

Shilling, F. M. & Waetjen, D. P. (2015) Wildlife-vehicle collision hotspots at US highway extents: scale and data source effects. *N*ature Conservation 11: 41–60

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.; Jaeger, J. A.; van der Grift, E. A. & Clevenger, A. P. (2011) Effects of roads and traffic on wildlife populations and landscape function: road ecology is moving toward larger scales *Ecology and society*, 16, 48-48

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

van Langevelde F, Jaarsma CF (2004) Using traffic flow theory to model traffic mortality in mammals. *Landscape Ecology* 19:895–907

Van Dyck, S. & Strahan, R. (2008). *The mammals of Australia.* New Holland Pub Pty Limited

Wildlife Victoria. (2016). Australian Wildlife Emergency Response Organisation

Visintin, C.; van der Ree, R. & McCarthy, M. A. (2016). A simple framework for a complex problem? Predicting wildlife–vehicle collisions. Ecology and Evolution.

* 1. Tables

Table 1: 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.

|  |  |  |
| --- | --- | --- |
| Variable | Description | Units |
| ELEV (SO) | Elevation of terrain above sea level | m |
| GREEN (SO) | Mean seasonal change in vegetation greenness (relative) | – |
| LIGHT (SO) | Remote-sensed artificial light intensity (relative) | – |
| MNTEMPWQ (SO) | Mean temperature of wettest quarter | °C\*10 |
| PRECDM (SO) | Precipitation of driest month | mm |
| SLOPE (SO) | Slope of terrain | % |
| TREEDENS (SO) | Tree cover | decimal % |
| X (SO) | X spatial coordinate of intersecting 1km2 grid centroid | m |
| Y (SO) | Y spatial coordinate of intersecting 1km2 grid centroid | m |
|  |  |  |
| DEER (CR) | Relative probability of deer occupancy per km2 | – |
| EGK (CR) | Relative probability of kangaroo occupancy per km2 | – |
| TSPD (CR) | Predicted vehicle speed per road segment | km/h |
| TVOL (CR) | Predicted traffic volume per road segment | vehicles/day |

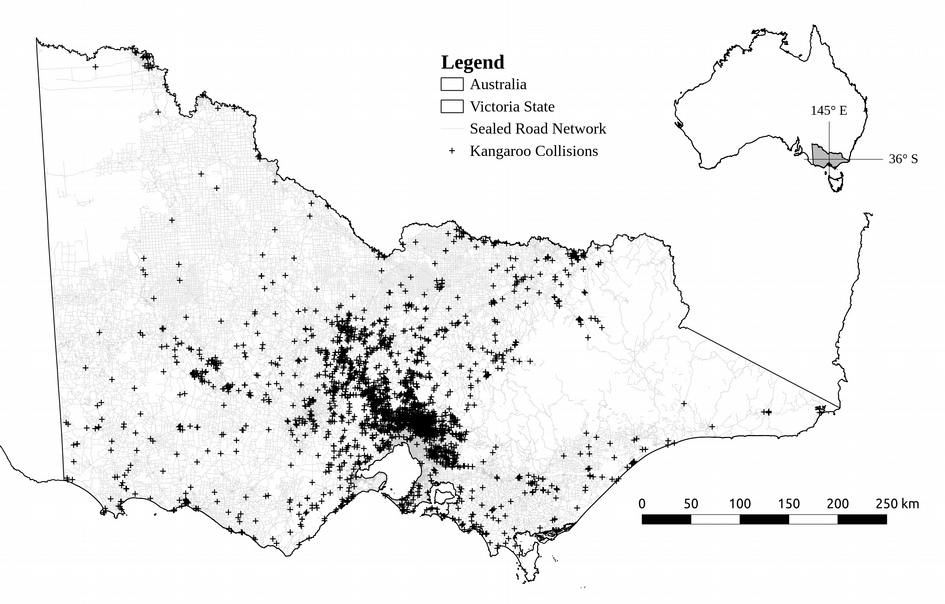
Table 2: Statistical model performance metrics shown as percent reduction of deviance on the null model using all training data and receiver operating characteristic (ROC) scores and regression coefficients using the independent validation data. All regression coefficients were highly significant (p<.001).

|  |  |  |
| --- | --- | --- |
|  | Eastern Grey Kangaroo | Mule Deer |
| % Reduction in Deviance | 11.8 | 18.6 |
| ROC Score | 0.79 | 0.88 |
| Regression Intercept | -2.22 (0.17 standard error) | 0.35 (0.08 standard error) |
| Regression Coefficient | 0.99 (0.04 standard error) | 0.94 (0.01 standard error) |

Table 3: Summary of collision models fits. Highly significant variables (p<.001) are marked with asterisks.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Species | Variable | Coefficient | Std. Error | Z-value | Pr(Z) |
| Eastern Grey Kangaroo | Intercept | -42.30 | 1.07 | -39.66 | <1.0x10-14\* |
|  | EGK | 0.70 | 0.02 | 47.77 | <1.0x10-14\* |
|  | TVOL | 5.55 | 0.25 | 22.68 | <1.0x10-14\* |
|  | TVOL2 | -0.33 | 0.02 | -21.76 | <1.0x10-14\* |
|  | TSPD | 3.93 | 0.07 | 54.49 | <1.0x10-14\* |
|  |  |  |  |  |  |
| Mule Deer | Intercept | -39.61 | 3.05 | -12.98 | <1.0x10-14\* |
|  | DEER | 0.61 | 0.02 | 29.12 | <1.0x10-14\* |
|  | TVOL | 2.49 | 0.66 | 3.80 | .0001\* |
|  | TVOL2 | -0.14 | 0.04 | -3.44 | .0006\* |
|  | TSPD | 6.29 | 0.21 | 30.67 | <1.0x10-14\* |

* 1. Figures



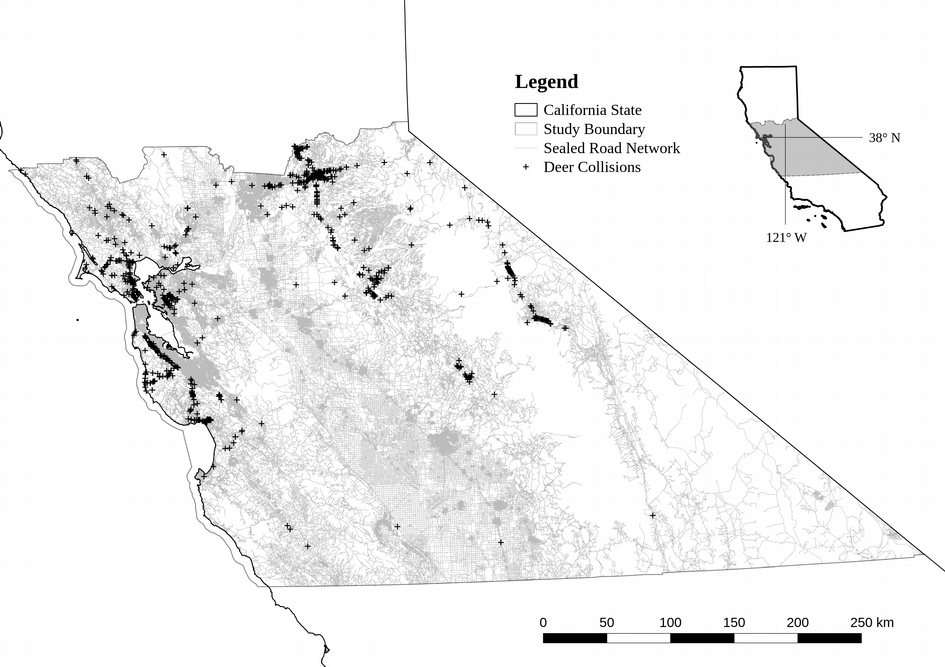


Figure 1: Location of study areas for analysis of factors influencing rate of WVC of eastern grey kangaroos in Victoria (above) and mule deer in central California (below). Insets show state of Victoria in Australia and portion of state of California referenced with latitude and longitude coordinates. Sealed roads are shown as light gray lines and locations of reported collisions/carcasses are shown as black crosses.

.

a) b)

c) d)

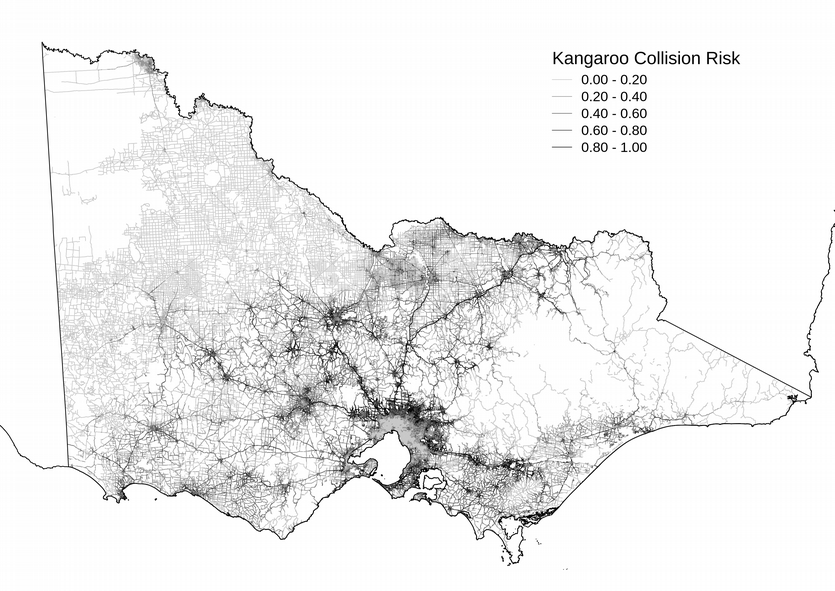
e) f)

Figure 2: Effects of predictor variables on relative likelihood of collision per species. Each variable is expressed with all other variables set at mean values. Likelihood of collision is expressed as a rate across all road segments for the total period of the observation data. To convert relative collision rate to expected annual number of collisions, multiply rate by total road segments divided by years of data (107,452.5 for kangaroos and 64,522.9 for deer).





Figure 3: Spatial autocorrelation in randomised quantile model residuals for each species at two spatial lags (1km and 250m). Each trend line is for a randomly selected subset of the data (20 total, each species).



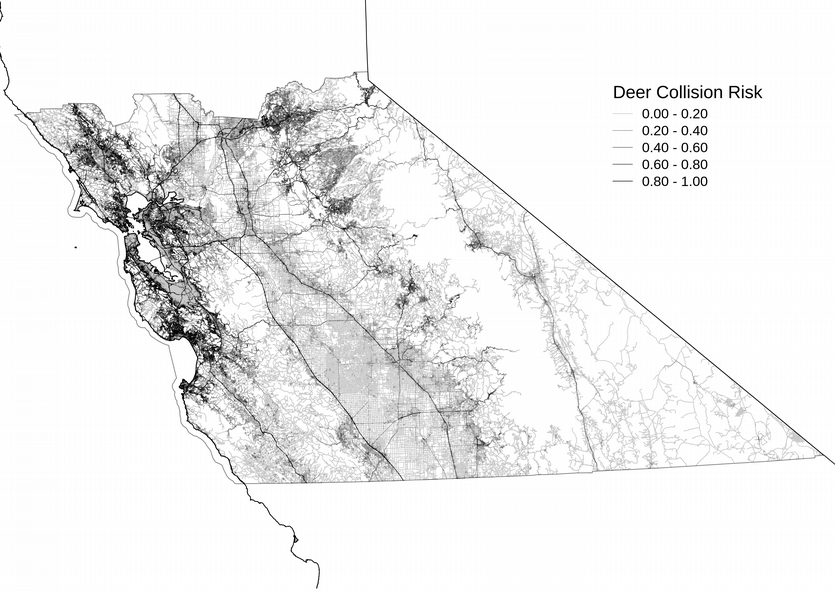


Figure 4: Predicted relative collision risk for all road segments in each study area. Darker, heavier lines indicate higher predicted values.

Figure S1: Predicted relative likelihood of occurrence kangaroos in Victoria and deer in central California area. Darker shading indicates higher relative likelihood of occurrence. (Victoria: extents=(-58000,5661000) x (764000,6224000), projection=GDA94 MGA zone 55, ncells=462,786; California: extents=(445000,3962000) x (1165000,4329000), projection=NAD83 UTM zone 10N, ncells=264,240)

Figure S2: Spatial autocorrelation in occupancy model residuals for Mule deer in central California (dot) and Eastern Grey Kangaroo in Victoria (triangle) each species by distance grouping (spatial lag).

Figure S3: Effects of three most significant predictor variables on relative likelihood of occurrence per species.

Appendix A

To establish their occurrence probabilities in their respective native geographic areas , we sourced presence records from two on-line, publicly-available databases; the Victorian Biodiversity Atlas, (DELWP, 2016) and Global Biodiversity Information Facility, (GBIF, 2016). We obtained presence records that satisfied the following criteria: survey date between 1 January 2000 and 1 December 2015 and spatial coordinate certainty of 500 metres or less. 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 two study species, we selected all spatially unique records of occurrence. To reduce potential effects of spatial dependency, we thinned each species occurrence points to maintain a minimum separation distance of 1000 metres between observations. As we did not have access to recorded absence data, we generated 10,000 randomly sampled points across each study area to represent background data. After omitting null values from sampling predictor variable grids, there were 700 presence and 9957 background points in the kangaroo occurrence modelling dataset and 366 presence and 9986 background points in the deer occurrence modelling dataset.

We developed species distribution models 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 0.005 (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 that influenced the biology, behaviour, and characteristics of each species based on existing literature and ecological principles (e.g. Coulson & Eldridge, 2010; Ferguson, 2005). Although each species is unique in its respective biology, we considered similar behaviours (sheltering and movements) and traits (foraging requirements) as the basis for choosing broad environmental variables that covered both species (Table 1). As bioclimatic variables exhibit spatial gradients, we also included the spatial coordinates of grid cell centroids (X, Y) as predictor variables in the models. This reduces biases in the influence of variables with high spatial correlation. It should be noted that our framework allows any combinations of data deemed appropriate for species modelling, however, using the same environmental variables for both kangaroos and deer demonstrates the generality of our framework and allows a more direct comparison in the analysis.

The species distribution models predicted relative likelihood of occurrence across the spatial grids for each species (Figure S1).

Although the mechanism of gradient boosting (BRT in this case) does not assume independence in the dependent variable, we reviewed spatial trends in the model fit for each species by calculating Moran’s I from model residuals and spatial coordinates and plotting against distance in one kilometre bins (Figure S2).

Each occurrence model fit the data as expected. The portions of the uncertainty in the distribution of the recorded occurrences of each species explained by the predictor variables (deviance) in the models were 26.7% for kangaroos and 30.8% for deer. Both models had good discriminative ability; the in-sample cross-validated ROC scores were 0.88 for kangaroos and 0.91 for deer. All of the environmental variables were influential for both species.

The three most non-spatial influential non-spatial predictors were slope (14.9% relative influence), artificial light (13.5%) and vegetation greenness (9.5%) for deer and artificial light (16.5%), elevation (12.5%) and vegetation greenness (10.5%) for kangaroos (Figure S3). Kangaroo occurrence showed a potential bi-modal response to artificial light with large peaks at both 25 and 55 units of relative light intensity. Deer occurrence did not have the same response shape, rather, there was a general increasing trend in occurrence probability with light intensity and a sharp peak between 45 and 50 units. Kangaroo occurrence was higher at low levels of elevations (<500m) with a peak at approximately 200m. Deer occurrence increased with higher slope aspects.

Spatial patterning was evident in the species occurrence model residuals between one and nine kilometres for both deer and kangaroos (Figure S2) and both species showed similar trends. Kangaroos demonstrated higher values of spatial autocorrelation than deer at closer spatial distances. The spatial covariates consistently ranked high among the most influential variables in both occurrence models (ranked 2nd & 3rd for kangaroos and 1st and 4th for deer). The X and Y variables accounted for 27.8% of the influence in the kangaroo model and 33.3% in the deer model.

The species occurrence models produced ecologically-reasonable predictions about the relative likelihood of species presence across each study area (Figure S1). Eastern grey kangaroo occurrence was predicted to be lower in north eastern Victoria which is consistent with historical knowledge. Predictions of deer occurrence were higher in areas of topographically-varied natural (parks and undeveloped) space around the San Francisco Bay Area which is also consistent with historical knowledge. However, lower relative occurrence was predicted than expected in the eastern portion of California (Sierra Nevada foothills) where deer are known to be abundant