Environmental Management (2009) 44:268–277
DOI 10.1007/s00267-009-9303-y

A Comparison of Data Sets Varying in Spatial Accuracy Used to Predict the Occurrence of Wildlife-Vehicle Collisions

Kari E. Gunson · Anthony P. Clevenger · Adam T. Ford · John A. Bissonette · Amanda Hardy

Received: 10 September 2008 / Accepted: 2 April 2009 / Published online: 19 May 2009

Abstract Wildlife-vehicle collisions (WVCs) pose a significant safety and conservation concern in areas where high-traffic roads are situated adjacent to wildlife habitat. Improving transportation safety, accurately planning highway mitigation, and identifying key habitat linkage areas may all depend on the quality of WVC data collection. Two common approaches to describe the location of WVCs are spatially accurate data derived from global positioning systems (GPS) or vehicle odometer measurements and less accurate road-marker data derived from reference points (e.g., mile-markers or landmarks) along the roadside. In addition, there are two common variable types used to predict WVC locations: (1) field-derived, site-specific measurements and (2) geographic information system (GIS)-derived information. It is unclear whether these different approaches produce similar results when attempting to identify and explain the location of WVCs. Our first objective was to determine and compare the spatial error found in road-marker data (in our case the closest mile-marker) and landmark-referenced data. Our second objective was to evaluate the performance of models explaining high- and low-probability WVC locations, using congruent, spatially accurate (<3-m) and road-marker (<800-m) response variables in combination with field- and GIS-derived explanatory variables. Our WVC data sets were comprised of ungulate collisions and were located along five major roads in the central Canadian Rocky Mountains. We found that spatial error (mean ± SD) was higher for WVC data referenced to nearby landmarks (516 ± 808 m) than for data referenced to the closest mile-marker data (401 ± 219 m). The top-performing model using the spatially accurate WVC locations contained all explanatory variable types, whereas GIS-derived variables were only influential in the best road-marker model and the spatially accurate reduced model. Our study showed that spatial error and sample size, using road-marker data for ungulate species, are important to consider for model output interpretation, which will impact the appropriate scale on which to apply modeling results. Using road-marker references <1.6 km or GPS-derived data locations may represent an optimal compromise between data acquisition costs and analytical performance.

Keywords Banff National Park · Data collection · Mortality · Road ecology · Spatial accuracy · Ungulate · Variable selection · Wildlife-vehicle collision

Wildlife-vehicle collisions (WVCs), particularly those involving ungulates, are increasingly recognized as a significant concern for traffic safety, socioeconomics, animal welfare, and wildlife management in the United States,

Electronic supplementary material The online version of this article (doi:10.1007/s00267-009-9303-y) contains supplementary material, which is available to authorized users.

K. E. Gunson
Eco-Kare International, P.O. Box 51522, Toronto, ON M4E 1C0, Canada
e-mail: kegunson@eco-kare.com

A. P. Clevenger · A. T. Ford · A. Hardy
Western Transportation Institute, Montana State University, P.O. Box 174250, Bozeman, MT 59717, USA

A. P. Clevenger (✉)
138 Birch Avenue, Harvie Heights, Alberta T1W 2W2, Canada
e-mail: apclevenger@gmail.com

J. A. Bissonette
U.S. Geological Survey, Utah Cooperative Fish and Wildlife Research Unit, Department of Wildland Resources, College of Natural Resources, Utah State University, Logan, UT 84322-5290, USA

Canada, and Europe (Groot-Bruinderink and Hazebroek 1996; Romin and Bissonette 1996; Joyce and Mahoney 2001). For instance, Conover and others (1995) estimated that >1 million deer (*Odocoileus* sp.)-vehicle collisions occurred annually in the United States, with an associated cost of more than U.S.$1 billion in vehicle damages and >29,000 human injuries.

Transportation agencies around the world have implemented highway mitigation measures such as fencing, animal detection systems, and variable message signs in an effort to reduce the frequency of WVCs (Spellerberg 1998; Clevenger and others 2001; Huijser and others 2007a). Often, WVC records are used to site these measures. For example, WVC data are being used to plan the location of wildlife crossing structures on highways (Marshik and others 2001; Evink 2002; Forman and others 2003). WVC data are also analyzed in conjunction with landscape and road variables to make predictions about where high WVC rates might occur to help inform the underlying mechanisms leading to future WVCs (Hubbard and others 2000; Malo and others 2004; Ramp and others 2005; Seiler 2005; Litvaitis and Tash 2008). For both purposes, it is imperative that transportation agencies acquire the necessary data to ensure that the analyses being performed are rigorous and reliable.

There are four common types of variable resolution used in most WVC-related studies, i.e., two response and two predictor variables. A review of published WVC data analyses (Table 1 and Supplementary Material) indicates that 48% of the studies used response variables derived from spatially accurate (SA) data (i.e., spatial error <10 m), while 68% used response variables derived from spatially inaccurate or road-marker (RM) data (i.e., spatial error >800 m). SA data are normally collected through regularly scheduled driving surveys using a global positioning system (GPS) to georeference WVC locations (e.g., Clevenger and others 2003; Ramp and others 2005; Langen and others 2007). Spatially inaccurate data may be referenced to a prominent geographical feature (e.g., nearest creek crossing or nearest township road [herein, landmark-referenced] or to the closest mile or one-tenth-mile highway marker [herein, road-marker] [Huijser and others 2007b]). In many jurisdictions, spatially inaccurate WVC data are collected by highway maintenance crews while performing other routine highway-related work, though some agencies are now implementing GPS data collection protocols for maintenance crews (Ament and others 2008).

The two predictor variable types used to explain WVC locations involve field-derived site-specific variables (e.g., habitat, and motorist visibility) and geographic information system (GIS)-derived contextual variables (e.g., distance to nearest river, proportional land cover type). A review of published WVC data analyses (Table 1 and Supplementary Material) shows that 54% of the studies used field-measured, site-specific predictor variables and 77% used GIS-derived predictor variables. Field variables are generally more difficult and expensive to obtain, as each sample site must be visited and measurements recorded. On the other hand, GIS data are readily available to researchers as well as the public (e.g., the Canadian Federal Government's GEOGRATIS Internet service); however, GIS-derived variables may lack the resolution needed to detect local-scale processes related to WVC locations. Thus, like the SA and RM response variables, environmental predictor variables types vary in their resolution, accuracy, and cost of acquisition. It is unclear whether the trade-off between data acquisition costs can offset the need for high-resolution data when analyzing the effects of environmental variables on WVC locations.

The purpose of this study was first to determine the amount of spatial error present between a SA and two

Table 1 Trade-off summary and literature review of studies collecting response- and predictor-type variables for wildlife-vehicle collision data analyses

Data source	Variable type	Data resolution	Spatial error	Data acquisition cost[a]	Data availability[b]	Percent of studies using data type ($n = 27$)[c]
Spatially accurate WVC	Response	High	Low	High	Low	48[d]
Road-marker WVC	Response	Low	High	Low	High	61[d]
Field-measured, site-specific	Predictor	High	Low	High	Low	54[e]
GIS-derived	Predictor	Low	Low or high	Low	High	77[e]

High and low categories are relative to one another for within-variable type comparisons

[a] Refers to personnel training, equipment, and data collection protocols

[b] Refers to ubiquity of data across jurisdictions/study areas

[c] Based on a Web of Science search for the terms (roads OR highway) and (mortality OR collision) and (wildlife OR fauna OR mammal OR ungulate OR deer OR elk OR moose OR bird OR amphibian) conducted on October 22, 2007

[d] Four of the 27 studies did not report the spatial error of the WVC data or how these data were collected in the field, thus $n = 23$ for these rows

[e] Two of the 27 studies predated GIS technology and 3 were not designed to incorporate spatial predictors of wildlife vehicle collision locations, thus $n = 22$ for these rows

spatially inaccurate data sets that are (1) referenced to the closest geographical feature and (2) referenced to the closest mile-marker. We then wanted to compare the performance of models derived from SA and RM data and determine how predictor variable resolution interacts with the spatial accuracy of response variables to affect model performance. We expected that models derived from RM data would be best explained by variables less sensitive to spatial error (i.e., GIS-buffer variables), whereas models created from the SA data would likely include explanatory variables derived from both field and GIS measures.

Materials and Methods

Study Area

This study was conducted in the central Canadian Rocky Mountains, approximately 150 km west of Calgary, Alberta, straddling the continental divide in southwestern Alberta and southeastern British Columbia (Fig. 1). The study area encompassed 11,400 km^2 of mountain and subalpine landscapes in Banff, Kootenay, and Yoho National Parks and adjacent Alberta provincial lands. The region has a continental climate characterized by long winters and short summers (Holland and Coen 1983). Vegetation consists of open forests dominated by lodgepole pine (*Pinus contorta*), Douglas fir (*Pseudotsuga menziesii*), white spruce (*Picea glauca*), Englemann spruce (*Picea englemannii*), quaking aspen (*Populus tremuloides*), and natural grasslands.

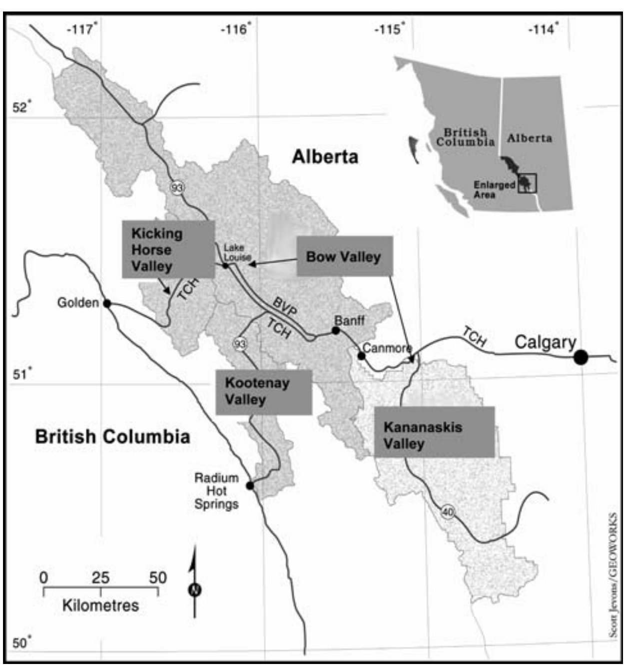

Fig. 1 The highways and their respective watersheds in the central Canadian Rocky Mountains where wildlife-vehicle collision data were collected

We selected five unfenced roads associated with the major watersheds in the study area (Fig. 1): (1) the Trans-Canada Highway (TCH) within the Kicking Horse Valley in Yoho National Park, (2) the TCH in the Bow Valley of Banff National Park, (3) the TCH in the Bow Valley in the province of Alberta, (4) Highway 93S in Kootenay National Park, and (5) Highway 40 in Kananaskis Country. The most prominent drainages, the Bow and Kicking Horse valleys, accommodate the TCH, which sustains the highest traffic volumes; e.g., the 2005 annual average daily traffic volume (AADT) ranged from 4600 to 16,960 (Parks Canada, unpublished data). The 2005 AADT for Highway 93S was 2000 (Parks Canada, unpublished data), while the 1999 AADT for Highway 40 was 3075 (Alberta Transportation, unpublished data). All AADT values were obtained from data taken from traffic counters placed on major highways by Alberta Transportation and Parks Canada.

Spatially Accurate Data Set

We created a spatially accurate (SA) data set by providing national park staff, Alberta highway maintenance contractors, and natural resource agencies in our study area with colored pin-flags to mark site-specific locations in the right-of-way where they observed road-killed large mammals (coyote [*Canis latrans*] sized and larger). Each collaborator then reported the location of the kill site to our field staff by referencing it to a nearby geographical landmark (e.g., 0.3 km west of Banff National Park entrance). We then relocated the kill site (marked with a pin-flag) within 48 h to record the actual location in two formats: (1) using a differentially correctable GPS unit (Trimble Navigation, Sunnyvale, CA, USA) with a high spatial accuracy (<3 m) and (2) using the odometer distance from the reported landmark to the pin-flagged WVC site. In the case where the pin-flag was no longer at the reported location, we took a GPS reading only if there was some evidence of the collision at the site, e.g., broken glass, animal fur, or blood stain.

To minimize variability in the analyses caused by differences in guilds and body size (Ford and Fahrig 2007), we focused our analysis on WVCs involving ungulate species. We only used WVC data that had a spatially accurate location (GPS coordinates; $n = 499$) from the complete data set collected from January 1998 to November 2003. This was approximately 85% of all the WVCs that occurred in our study area for that time period. White-tailed deer (*Odocoileus virginianus*) and mule deer (*O. hemionus*) were most frequently involved in vehicle collisions, comprising

58% of the total, followed by elk (*Cervus elaphus*; 27%), moose (*Alces alces*; 7%), bighorn sheep (*Ovis canadensis*; 3%), and all others (5%).

Road-Marker Data Set

We created a road-marker (RM) data set with a spatial accuracy similar to that commonly used in other studies (e.g., ±0.5 mi or ±800 m). We first divided each of the five highways in the study area into 1.6-km (1.0-mi) segments using ArcView 3.3 (Environmental Systems Research Institute, Redlands, CA, USA). We then 'moved' each of the 499 SA WVC locations to its nearest mile-marker reference point ($n = 120$) and recorded the total number of WVCs within each mile-marker segment. The only difference between the SA and the RM data is the spatial distribution, with each of the SA locations being moved a maximum of 800 m to create the RM data.

Reduced, Spatially Accurate Data Set

A further issue in comparing SA to RM data is related to differences in statistical power; i.e., for the same length of highway, the SA data have one data point for each WVC location (e.g., $n = 499$), whereas the RM data have one data point for each associated mile-marker (e.g., $n = 120$). We addressed this issue by reducing the density of WVC locations in the SA data in order to create a reduced SA data set (RSA). This data set was created by selecting the first occurrence of a WVC from the complete SA data and subsequently selecting the next nearest WVC along the same road that was at least 1.6 km away. The RSA data thus have a spatial distribution and sample size ($n = 120$) similar to those of the RM data set.

Definitions of High- and Low-WVC Frequency

In all data sets, we delineated 1.0-mi highway segments, defined as 800 m (0.5 mi) in each direction of a given mile-marker. We categorized each 1-mi segment as having either a high WVC probability (high kill) or a low WVC probability (low kill) by comparing the total number of WVCs associated with a given segment to the mean number of WVCs per mile for each of the five highways. If the summed number of ungulate collisions within the 1.6-km segment was greater than or equal to the rounded mean for that highway, the RM segment was considered a high-kill zone; if it was less, the segment was considered a low-kill zone. We selected this cutoff point because it was a natural break in the data and because it allowed for sufficient sample size for each high-low category on each road (see Hubbard and others 2000).

Predictor Variables

We used expert opinion and the literature to determine potential field-derived variables that may have influenced the occurrence of a collision (Table 2). To determine the GIS-derived variables we selected the most suitable variables that could be derived from the available digital data for the entire study region. We grouped the variables into three measurement types: (1) field-derived variables, defined as site-specific measurements at each WVC and RM location; (2) GIS-derived proximity variables, defined as distance measurements from each collision/marker location to the nearest landscape feature; and (3) GIS-derived buffer variables, defined as measurements within a given distance of each collision/marker location.

Site-Specific Predictor Variables

From April 2003 to July 2005 we used a GPS unit to locate the 499 WVC locations and the 120 mile-marker locations on the highways and we measured 14 site-specific variables at each location (Table 2). We used a digital rangefinder (Yardage Pro 1000; Bushnell Inc., Denver, CO, USA) to measure the distance to the nearest vegetative cover and visibility. Vegetative cover, habitat, and topography within an approximately 100-m radius were all estimated visually at each WVC site (defined in Table 2). We measured the slope for the roadside, verge, and adjacent land at each site along a 30-m perpendicular transect from the highway edge. We obtained elevation on-site using a GPS unit.

Visibility variables estimated the extent to which it was assumed that a motorist could see wildlife on the highway right-of-way. Standing at the pavement edge (in-line visibility), and at 5 and 10 m from the pavement edge (angular visibility), we measured the distance where a passing vehicle could no longer be observed using a digital rangefinder. We took 12 visibility measurements at each site: facing each direction of traffic (at, and away from, oncoming traffic), on both sides of the highway (eastbound and westbound lanes), and at each distance from the pavement within each site (0, 5, and 10 m). Since we could not always determine from what side or from which direction a vehicle struck an animal, we averaged the four measurements taken at each distance from the pavement to provide three visibility values at each location: in-line visibility (0 m), 5-m angular visibility, and 10-m angular visibility.

Proximity and Buffer Predictor Variables

We obtained GPS coordinates for all concrete Jersey barriers, guardrails, and streams near each highway. We digitized landscape features (lakes, rivers, and human use facilities) within an 800-m buffer surrounding the road

Table 2 Definition and description of variables used in the analysis of factors explaining ungulate-vehicle collision occurrence

Variable	Definition
Habitat class	Dominant habitat within a 100-m radius on both sides of the highway measured as open (meadows, and barren ground; O), open deciduous forest (some deciduous forest interspersed with open grass patches; ODF), and coniferous forest (CF)
Topography[a]	Landscape-scale terrain measured as flat (1), completely raised or buried (2), or partially buried or raised (3)
Forest cover	Mean percentage of continuous forest cover (trees >1 m tall) in a 100-m transect line perpendicular to the highway, taken from both sides of the road
Shrub cover	Mean percentage of shrub cover (trees and shrubs <1 m tall) in a 100-m transect line perpendicular to the highway, taken from both sides of the road
Openness	Mean percentage of area devoid of shrub or forest (e.g., rock, gravel, grass, pavement) in a 100-m transect line perpendicular to the highway, taken from both sides of the road
Vegetative cover	Mean distance (m) to vegetative cover (trees and shrubs >1 m high) taken from both sides of the road
Roadside slope	Mean slope (deg) of the land 0–5 m perpendicular to the pavement edge taken from both sides of the road
Verge slope	Mean slope (deg) of the land 5–10 m perpendicular to the pavement edge taken from both sides of the road
Adjacent land slope	Mean slope (deg) of the land 10–30 m perpendicular to the pavement edge taken from both sides of the road
Elevation	GPS height (m)
Road width	Distance (m) from one side of the highway pavement to the other
In-line visibility	Mean distance (m) at which an observer standing at the pavement edge could no longer see passing vehicles taken from each direction on both sides of the highway
5-m angular visibility	Mean distance (m) at which an observer standing 5 m from the pavement edge could no longer see passing vehicles taken from each direction on both sides of the highway
10-m angular visibility	Mean distance (m) at which an observer standing 10 m from the pavement edge could no longer see passing vehicles taken from each direction on both sides of the highway
Drainage distance	Distance (m) to the nearest waterway (river, stream, or creek) which crossed the road
Human use distance	Distance (m) to the nearest human use feature (e.g., rest stop, campground or lodging) along the highway
Barrier distance	Distance (m) to the nearest Jersey barrier[b] or guardrail[c]
Road length	Length (m) of each highway segment within each buffer
Open water area	Area (km^2) of open water within each buffer
River length	Length (m) of all rivers within each buffer
Barrier length	Length (m) of all Jersey barriers and guardrails in each buffer
Slope mean	Mean slope (deg) within each buffer
Slope maximum	Maximum slope (deg) within each buffer

[a] (1) Flat; (2) raised; (2) buried-raised; (2) buried; (3) part-buried; (3) part-raised

[b] A Jersey barrier stands ≈1.0 m and is made of poured concrete. It is designed to separate lanes of traffic (often opposing lanes of traffic), with the goal of minimizing vehicle crossover in the case of accidents

[c] Guardrail stands ≈1.0 m and is designed to provide a barrier between traffic and hazardous features in the landscape, such as an open body of water

from 1:50,000-scale topographic maps. We used a digital elevation model (30-m resolution) to obtain slope measurements.

We used ArcView 3.3 GIS to aid in measuring the distance from each WVC location and mile-marker sampling site to a landscape feature (i.e., watercourse, human use area, Jersey barrier; Table 2). We also used 800-m radius buffers around each sampling site to calculate either the area (i.e., open water, human use features) or the length (i.e., rivers, roads, road safety barriers) of specific landscape features. We chose 800 m, as this would encompass the entire length of each mile-marker section.

Addressing Spatial Error

We calculated the mean, standard error, and range for two types of spatial error associated with WVCs. The spatial error for the landmark-referenced WVC data was calculated by finding the distance between the actual and the reported WVC location using the vehicle odometer measures. The spatial error in the RM data was calculated by finding the distance between the actual location of the WVC record and the distance the WVC was 'moved' to its closest mile-marker using a GIS.

Model Performance

We used a binomial maximum likelihood logistic regression (Hosmer and Lemeshow 1989) with high- or low-kill location as the dependent variable to create models from the SA, RM, and RSA data for all analyses. To minimize overfitting and to meet our objectives of determining the effects of spatial error on model performance for different variable types, we entered all predictor variables into the model by variable type, rather than performing a stepwise procedure (Burnham and Anderson 2002; Shtatland and others 2003). Thus, for each data set we created seven models that included all combinations of each of the three predictor variable types: field-derived, GIS proximity, and GIS buffer. To determine the best-fitting model for each variable type, we used Akaike's Information Criterion with a correction for small sample sizes (AIC_c) and Akaike weights (w_i) for each of the three data sets and seven model types (21 models in total) (Burnham and Anderson 2002). We reported beta coefficients for models where the ΔAIC_c for a given set of competing models was <10. We calculated model-averaged parameter estimates and 95% confidence intervals (Burnham and Anderson 2002) highlighting those estimates whose intervals do not overlap with 0. To compare the explanatory power of models derived from different data sets, we used Nagelkerke R^2 values. We used SAS version 9.1 for all logistic regression analyses (SAS Institute 2003).

Results

Calculated Spatial Error

The mean spatial error for WVCs referenced to the closest landmark ($n = 233$) was 516 m (SD, 808 m; range, 0–6500 m). The spatial error of the road marker data set, i.e., the distance each observed location was 'moved' to the nearest mile-marker location, was a mean distance of 401 m (SD, 219 m; range, 7–794 m).

Model Performance

The rounded WVC mean and cutoff for high- vs. low-WVC differentiation per 1.6 km for each highway segment was 2 (TCH within the Kicking Horse Valley in Yoho National Park), 3 (TCH within the Bow Valley of Banff National Park), 13 (TCH within the Bow Valley east of Banff National Park), 2 (Highway 93S within Kootenay National Park), 2 (Highway 40 within Kananaskis Country). The spatially accurate (SA) data set had 385 high- and 114 low-kill zones for all roads combined, while the road-marker (RM) data set had 63 high- and 57 low-kill zones and the reduced-SA (RSA) data set had 73 high- and 47 low-kill zones.

The model using all the variables, developed from the SA data, was unequivocally the best model supporting the prediction of high- and low-kill WVC locations (Table 3).

Table 3 Model performance using three variable types as a group for three spatial data sets (with sample size for high- and low-kill zones), where R^2 is the Nagelkerke value, K is the number of parameters, and ω_i is the Akaike weight

Model	R^2	K	AIC_c	ΔAIC	ω_i
Spatially accurate (high = 385, low = 114)					
All	0.39	26	438.78	0.00	1.00
Distance-buffer	0.28	10	455.51	16.74	0.00
Field-buffer	0.30	23	472.75	33.97	0.00
Buffer	0.18	7	486.96	48.18	0.00
Field-distance	0.23	20	494.26	55.47	0.00
Distance	0.07	4	520.04	81.26	0.00
Field	0.13	17	524.28	85.50	0.00
Road-marker (high = 63, low = 57)					
Distance-buffer	0.23	10	165.30	0.00	0.92
Buffer	0.11	7	171.20	5.90	0.05
Distance	0.02	4	172.40	7.10	0.03
Field	0.24	17	182.55	17.25	0.00
Field-buffer	0.37	23	184.79	19.49	0.00
All	0.42	26	187.96	22.66	0.00
Field-distance	0.25	20	189.90	24.59	0.00
Reduced, spatially accurate (high = 73, low = 47)					
Distance-buffer	0.23	10	156.70	0.00	0.72
Buffer	0.14	7	158.64	1.94	0.27
Distance	0.01	4	168.03	11.34	0.00
Field-buffer	0.36	23	177.94	21.24	0.00
All	0.41	26	181.88	25.18	0.00
Field	0.16	17	184.33	27.64	0.00
Field-distance	0.17	20	192.10	35.40	0.00

For the RM data set, the distance-buffer model had substantial support,with lower support for the buffer-only and distance-only models. For the RSA data set, the distance-buffer model had some support, followed by the buffer model. Among the top-performing models from each data set, the Nagelkerke R^2 was highest for the SA data ($R^2 = 0.39$), followed by a tie between the RM and the RSA data (at $R^2 = 0.23$).

Table 4 shows that several factors contributed to high-frequency WVC locations using the SA data. WVCs were more likely to be found in coniferous forest than in open habitat. In addition, WVCs were found closer to drainages intersecting the road and when the maximum surrounding slope was lower. In the SA and RM data sets, WVCs were more likely to be found closer to barriers and when there were fewer barriers along road segments at WVC locations. Except for slope maximum, all of the 95% confidence intervals for the model-averaged estimates derived from RSA data overlapped with zero.

Table 4 Model-averaged coefficient estimates and 95% confidence intervals (CI) for the top models in the three data sets

Variable	Coefficient estimate	Upper CI	Lower CI
Spatially accurate (high = 385, low = 114)			
HC: open[a,]*	−1.4676	−0.5462	−2.9575
HC: open deciduous forest[a]	0.6999	1.2463	−0.9813
T: completely raised or buried[b]	−0.3694	0.1902	−2.0439
T: partially raised or buried[b]	0.4858	0.9150	−1.2552
Forest cover	0.0047	0.0921	−1.9107
Shrub cover	−0.0015	0.0915	−1.9141
Openess	0.0213	0.1089	−1.8940
Vegetative cover	−0.0065	0.0037	−1.9613
Roadside slope	0.0314	0.0733	−1.9072
Verge slope	0.0254	0.0579	−1.9180
Adjacent land slope	−0.0122	0.0103	−1.9607
Elevation	−0.0009	0.0012	−1.9598
Road width	0.0403	0.0670	−1.9061
In-line visibility	−0.0017	0.0006	−1.9605
5-m angular visibility	0.0003	0.0027	−1.9585
10-m angular visibility	0.0005	0.0020	−1.9587
Drainage distance*	−0.0003	−0.0002	−1.9602
Human use distance	−0.0004	<0.0001	−1.9602
Barrier distance*	−0.0004	−0.0002	−1.9603
Road length	−0.0014	0.0071	−1.9571
Open water area	<0.0001	<0.0001	−1.9598
River length	0.0001	0.0002	−1.9598
Barrier length*	−0.0022	−0.0012	−1.9617
Slope mean	0.0117	0.0928	−1.9069
Slope maximum*	−0.0631	−0.0282	−2.0053
Road-marker (high = 63, low = 57)			
Drainage distance	−0.0002	0.0001	−0.0004
Human use distance	<−0.0001	0.0006	−0.0006
Barrier distance*	−0.0005	−0.0001	−0.0009
Road length	−0.0021	0.0057	−0.0100
Open water area	<0.0001	<0.0001	<−0.0001
River length	0.0001	0.0004	−0.0002
Barrier length*	−0.0020	−0.0005	−0.0035
Slope mean	0.0137	0.1239	−0.0965
Slope maximum	−0.0541	0.0001	−0.1083
Reduced, spatially accurate (high = 73, low = 47)			
Drainage distance	<0.0001	−0.0001	<0.0001
Human use distance	<0.0001	−0.0002	0.0001
Barrier distance	−0.0004	−0.0012	0.0004
Road length	−0.0029	−0.0086	0.0028
Open water area	<0.0001	−0.0007	0.0007
River length	0.0002	0.0005	−0.0002
Barrier length	−0.0015	−0.0046	0.0015
Slope mean	0.0546	0.1616	−0.0524
Slope maximum*	−0.0717	−0.2123	−0.0688

HC habitat class variables, *T* road topography variables. Variables were significant (designated with an asterisk) if the upper and lower CI did not include 0

[a] Compared to the reference category 'coniferous forest'

[b] Compared to the reference category 'flat'

Discussion

In many studies, the spatial error of WVC data is often estimated (e.g., Hubbard and others 2000; Seiler 2005) or not reported at all (e.g., Nielson and others 2003; Bashore and others 1985). Studies using collision location data without knowledge of the inherent spatial error may compromise the informed selection of model inputs and interpretation of model outputs. Our study is the first published attempt that we are aware of to explicitly show the spatial error of WVC locations. We found that the mean spatial error and standard deviation were higher for WVCs referenced to a nearby landmark than for WVCs referenced to the nearest mile marker. This result implies that when reporting WVC locations without the aid of a GPS unit, collisions locations are more accurate if recorded to a consistent feature along roads, such as road markers (e.g., Dussault and others 2006).

Comparing model performance among data sets we found that the top-performing model using the spatially accurate (SA) data explained 16% more of the variation in WVC locations than the top-performing models derived from both the road-marker (RM) and the reduced SA (RSA) data. The top-performing model using the SA data included field and GIS variables; however, field-derived predictor variables were not associated with the best-performing models in either the RM or the RSA data. Since there were no field variables in the top-performing RSA models, this suggests that greater statistical power in the SA data may be responsible for greater model performance. Data collection methods that maximize sample size in RM data, such as using 0.1- or 1.0-km RMs (e.g., Malo and others 2004; Saeki and Macdonald 2004) to record kill locations, may represent an optimal balance between data the acquisition costs associated with increased spatial resolution of the data (Clevenger and others 2003; Ramp and others 2005, 2006) and the inferential strength of modeling results.

This study does not unequivocally confirm our initial prediction that field-derived predictor variables are sensitive to spatial error, however, it is intuitive that field- and distance-measured variable types are best suited for spatially accurate WVC locations. This is especially true if the variable measurements vary significantly within the known spatial error of the WVC data. Careful consideration should be given to the selection of predictor variables that match the spatial resolution of the response variable and the habitat characteristics of the landscape.

Our study limited the scale of GIS predictors influencing collisions to 1.6 km to be consistent with common WVC collection protocols (Puglisi and others 1974; Hubbard and others 2000; Huijser and others 2007b). In other studies, it would be worthwhile to extend the analytical resolution of spatial processes affecting WVC locations to larger scales. For example, Clevenger and others (2003) found the peak clustering of WVCs at 20 km for small mammals in our study area, while Mountrakis and Gunson (in press) found that the peak clustering occurred at 10 km for moose in Vermont. Future WVC analyses designed to test the influence of resolution and spatial scale of both predictor and response variables on model inferential strength would assist in further developing the results obtained in this study.

In addition to characterizing the effects of spatial error in the analysis of WVC data, our results also provide insight into the mechanisms affecting WVC locations. In the SA model, high-kill sites were more closely associated with forested habitats (open-deciduous and coniferous) than with open habitat. Other studies have found similar results, where a combination of wooded areas for cover and open grasslands for grazing facilitates ungulate movement toward and across roads (Puglisi and others 1974; Bashore and others 1985; Malo and others 2004). The distance to drainage was inversely related to the probability of high-kill locations, which can be explained by the use of drainage systems as travel routes for wildlife (Bellis and Graves 1971; Feldhamer and others 1986; Finder and others 1999). WVCs also tended to occur near the ends of roadside barriers (i.e., Jersey barriers and guardrails), suggesting a 'drift-fence' effect of these features on wildlife movement (Malo and others 2004). Additionally, barrier length negatively affected WVC probability, suggesting that road safety barriers may deter animals from crossing roads at these locations (Barnum 2003; Malo and others 2004). Since ungulates are capable of crossing these barriers (<1.5 m in height), this result is most likely due to landscape factors associated with barrier locations, such as steep topography or the presence of wetlands. It is more likely that these landscape features deter animals from crossing the highway at these locations than the safety barriers per se.

In contrast to other studies reporting environmental variable associations with WVC locations, we did not find a large number of statistically significant variables. For example, Seiler (2005) showed that 14 of the 19 variables were included in their final best model and significantly ($p < 0.05$) contributed to the occurrence of moose-vehicle collisions in Sweden. A probable explanation is that our study area was in a relatively homogeneous landscape characteristic of a mountainous ecosystem contained within provincial and federally protected parks. Other studies have shown that where preferred habitat is extensive and common, deer accident sites were observed to be more randomly distributed (Bashore and others 1985; Allen and McCullough 1976; Feldhamer and others 1986). As land cover in our study area becomes more heterogeneous through proposed habitat modification programs, e.g., prescribed burning (White and others 1998), the spatial distribution of WVC locations may shift as animal distributions respond to habitat change.

Conclusions

This study is the first we are aware of that examines the role of spatial error in predicting the location of WVCs using environmental variables. Spatial error of WVC locations needs to be matched with appropriately scaled predictor variables. Confounding our ability to tease apart the effects of spatial error in WVC locations on model selection is the issue of statistical power. Power is generally greater in spatially accurate (SA) WVC data because the inferential strength of data are limited by the number of WVCs for a given stretch of highway, as opposed to the length of the highway as is the case for road-marker (RM) data. In hindsight, we would have gained additional information on WVC analysis practices if we had collected environmental data from the landmark-referenced WVC data ($n = 233$) we used to calculate spatial error. Future studies should assess the strengths and limitations of the landmark-referenced data relative to RM and SA WVC data to assess its contribution to WVC analyses.

Most transportation agencies have legal or policy mandates to minimize WVCs, especially as highways are modified and newly constructed (Huijser and others 2007a, b) and where mitigation measures need to be sited. Generally, RM WVC data, along with GIS-derived buffer variables, are useful for explaining large-scale interactions between WVCs and roads, as well as determining highway segments (e.g., $\geq$1.6 km in length) where collisions are most likely to occur (Puglisi and others 1974; Seiler 2005). These types of data are useful for large-scale mitigation planning, such as modified speed limits and fencing (Huijser and McGowen 2003; D'Angelo and others 2006; Krisp and Durot 2007). On the other hand, spatially accurate WVC data are appropriately used with field, proximity, and buffer variables to assist in identifying both large-scale and localized sections (<1.6 km) of highway that have high WVC rates. High-resolution information is essential for informing management actions at the local scale, such as the placement of wildlife crossing structures and grade-level crossings and the alteration of roadside habitat and road design (Putman 1997; Lenhert and Bissonette 1998; Clevenger and Waltho 2000; Forman and others 2003; Rea 2003).

Acknowledgments We thank A. Seiler, N. Dodd, and two anonymous reviewers for their thoughtful comments on an early draft of the manuscript. Support for the Banff research data collection came from Parks Canada and Public Works and Government Services Canada (contracts C8160-8-0010 and 5P421-010004/001). Preparation of this paper was funded by Parks Canada Ecological Integrity Innovation and Leadership funding to the Banff National Park Field Unit (Partnership for Research and Monitoring on the Impacts of Highways on Wildlife in Canada's Mountain Parks, 2005–2008) and Transportation Research Board, National Highway Cooperative Research Program funding to Utah State University. Finally, this work and continued long-term research on roads and wildlife populations in the Canadian Rocky Mountains would not have been possible without funding from the Woodcock, Henry P. Kendall, and Wilburforce foundations and their support in securing an international public-private partnership.

References

Allen RE, McCullough DR (1976) Deer car accidents in southern Michigan. Journal of Wildlife Management 40:317–325

Ament R, Clevenger AP, Yu O, Hardy A (2008) An assessment of road impacts on wildlife populations in U.S. national parks. Environmental Management 42:480–496

Barnum SA (2003) Identifying the best locations along highways to provide safe crossing opportunities for wildlife. Report CDOT-DTD-UCD-2003–9. Research Branch, Colorado Department of Transportation, Denver

Bashore TL, Tzilkowski WM, Bellis ED (1985) Analysis of deer-vehicle collision sites in Pennsylvania. Journal of Wildlife Management 49:69–774

Bellis ED, Graves HB (1971) Deer mortality on a Pennsylvania interstate highway. Journal of Wildlife Management 35:232–237

Burnham KP, Anderson DR (2002) Model selection and inference: a practical information-theoretic approach. Springer, New York

Clevenger AP, Waltho N (2000) Factors influencing the effectiveness of wildlife underpasses in Banff National Park, Alberta, Canada. Conservation Biology 14:47–56

Clevenger AP, Chruszcz B, Gunson KE (2001) Highway mitigation fencing reduces wildlife-vehicle collisions. Wildlife Society Bulletin 29:646–653

Clevenger AP, Chruszcz B, Gunson KE (2003) Spatial patterns and factors influencing small vertebrate fauna road-kill aggregations. Biological Conservation 109:15–26

Conover MR, Pitt WC, Kessler KK, DuBow TJ, Sanborn WA (1995) Review of human injuries, illnesses, and economic losses caused by wildlife in the US. Wildlife Society Bulletin 23:407–414

D'angelo GJ, D'angelo JG, Gallagher GR, Osborn DA, Miller KV, Warren RJ (2006) Evaluation of wildlife warning reflectors for altering white-tailed deer behavior along roadways. Wildlife Society Bulletin 34:1175–1183

Dussault CM, Roulin M, Courtois R, Ouellet JP (2006) Temporal and spatial distribution of moose-vehicle accidents in the Laurentides Wildlife Reserve, Quebec, Canada. Wildlife Biology 12: 415–426

Evink G (2002) Interaction between roadways and wildlife ecology: a synthesis of highway practice. National Cooperative Highway Research Program Synthesis 305. Transportation Research Board, Washington, DC

Feldhamer GA, Gates JE, Harman DM, Loranger AJ, Dixon KR (1986) Effects of interstate highway fencing on white-tailed deer activity. Journal of Wildlife Management 50:497–503

Finder RA, Roseberry LL, Woolf A (1999) Site and landscape conditions at white-tailed deer-vehicle collision locations in Illinois. Landscape and Urban Planning 44:77–85

Ford A, Fahrig L (2007) Diet and body size of North American road mortalities. Transportation Research Part D 12:498–505

Forman RTT, Sperling D, Bissonette JA, Clevenger AP, Cutshall CD, Dale VH, Fahrig L, France R, Goldman CR, Heanue K, Jones JA, Swanson FJ, Turrentine T, Winter TC (2003) Road ecology: science and solutions. Island Press, Washington DC

Groot-Bruinderink GWTA, Hazebroek E (1996) Ungulate traffic collisions in Europe. Conservation Biology 10:1059–1067

Holland WD, Coen GM (1983) Ecological land classification of Banff and Jasper national parks. The Wildlife Inventory, vol III. Canadian Wildlife Service, Edmonton Alberta

Hosmer DW, Lemeshow S (1989) Applied logistic regression. Wiley, New York

Hubbard MW, Danielson BJ, Schmitz RA (2000) Factors influencing the location of deer-vehicle accidents in Iowa. Journal of Wildlife Management 64:707–712

Huijser MP, McGowen PT (2003) Overview of animal detection and animal warning systems in North America and Europe. In: Irwin CL, Garrett P, McDermott KP (eds) Proceedings of the 2003 international conference on ecology and transportation. Center for Transportation and the Environment, North Carolina State University, Raleigh, pp 268–382

Huijser MP, McGowen P, Fuller J, Hardy A, Kociolek A, Clevenger AP, Smith D, Ament R (2007a) Wildlife-vehicle collision reduction study. Report to U.S. Congress. U.S. Department of Transportation, Federal Highway Administration, Washington, DC

Huijser MP, Fuller J, Wagner ME, Hardy A, Clevenger AP (2007b) Animal-vehicle collision data collection: a synthesis of highway practice. National Cooperative Highway Research Program (NCHRP) Synthesis 370. Transportation Research Board, Washington, DC

Joyce TL, Mahoney SP (2001) Spatial and temporal distributions of moose-vehicle collisions in Newfoundland. Wildlife Society Bulletin 29:281–291

Krisp JM, Durot S (2007) Segmentation of lines based on point densities-An optimisation of wildlife warning sign placement in southern Finland. Accident Analysis and Prevention 39:38–46

Langen TA, Machniak A, Crowe EK, Mangen C, Marker DF, Liddle N, Roden B (2007) Methodologies for surveying herpetofauna mortality on rural highways. Journal of Wildlife Management 71:1361–1368

Lenhert MA, Bissonette JA (1998) Effectiveness of highway crosswalk structures at reducing deer-vehicle collisions. Wildlife Society Bulletin 25:809–818

Litvaitis JA, Tash JP (2008) An approach to understanding wildlife-vehicle collisions. Environmental Management 42:688–697

Malo JE, Suarez F, Diez A (2004) Can we mitigate animal-vehicle accidents using predictive models? Journal of Applied Ecology 41:701–710

Marshik JL, Renz JL, Sipes D, Becker D, Paulson D (2001) Preserving spirit of place: U.S. Highway 93 on the Flathead Indian Reservation. In: Irwin CL, Garrett P, McDermott KP (eds) Proceedings of the 2001 international conference on ecology and transportation. Center for Transportation and the Environment, North Carolina State University, Raleigh, pp 244–256

Mountrakis G, Gunson KE (in press) Multi-scale spatiotemporal analyses of moose-vehicle collisions: a case study in northern Vermont. International Journal of GIS

Nielson CK, Anderson RG, Grund MD (2003) Landscape influences on deer-vehicle accident areas in an urban environment. Journal of Wildlife Management 67:46–51

Puglisi MJ, Lindzey JS, Bellis ED (1974) Factors associated with highway mortality of white-tailed deer. Journal of Wildlife Management 38:799–807

Putman RJ (1997) Deer and road traffic accidents: options for management. Journal of Environmental Management 51:43–57

Ramp DJ, Caldwell J, Edwards KA, Warton D, Croft DB (2005) Modelling of wildlife fatality hotspots along the Snowy Mountain Highway in New South Wales, Australia. Biological Conservation 126:474–490

Ramp D, Wilson VK, Croft DB (2006) Assessing the impacts of roads in periurban reserves: road-based fatalities and road usage by wildlife in the Royal National Park, New South Wales, Australia. Biological Conservation 129:348–359

Rea RV (2003) Modifying roadside vegetation management practices to reduce vehicular collisions with moose *Alces alces*. Wildlife Biology 9:81–91

Romin LA, Bissonette JA (1996) Deer-vehicle collisions: status of state monitoring activities and mitigation efforts. Wildlife Society Bulletin 24:276–283

Saeki M, Macdonald DW (2004) The effects on the raccoon dog (*Nyctereutes procyonoides viverrinus*) and other mammals in Japan. Biological Conservation 118:559–571

SAS, version 9.1 (2003) SAS Institute Inc. Cary, NC, USA

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

Shtatland ES, Kleinman K, Cain EM (2003) Stepwise methods in using SAS PROC LOGISTIC and SAS Enterprise minter for prediction. SUGI 28 Proceedings, paper 258-28, Cary, NC

Spellerberg IF (1998) Ecological effects of roads and traffic: a literature review. Global Ecology and Biogeography Letters 7:317–333

White C, Olmsted C, Kay C (1998) Aspen, elk, and fire in the Rocky Mountain national parks of North America. Wildlife Society Bulletin 26:449–462