RESOURCES AGENCY
California Department of Fish and Wildlife

Development and implementation of DNA-based survey methods for population monitoring of tule elk (*Cervus canadensis nannodes*) in Colusa and Lake Counties, California

---

A bachelor group of bull tule elk near Cowboy Camp, Colusa County, Calif. Image captured via remote trail camera June 27, 2019.

Thomas J. Batter[a,b], Joshua P. Bush[c], and Benjamin N. Sacks[a,d]

[a]*Mammalian Ecology and Conservation Unit, Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis*
[b]*California Department of Fish and Wildlife, Inland Deserts Region, Ontario, CA*
[c]*California Department of Fish and Wildlife, North-Central Region, Rancho Cordova, CA*
[d]*Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis*

December 10, 2020

Final report to the California Department of Fish and Wildlife, Agreement No. P1680034

Development and implementation of DNA-based survey methods for population monitoring of tule elk (*Cervus canadensis nannodes*) in Colusa and Lake Counties, California

Suggested Reference
Batter, T.J., J.P. Bush, and B.N. Sacks. 2020. Development and implementation of DNA-based survey methods for population monitoring of tule elk (*Cervus canadensis nannodes*) in Colusa and Lake Counties, California. Draft final report to the California Department of Fish and Wildlife. December 10, 2020.

Acknowledgements
This research was funded by the California Department of Fish and Wildlife (CDFW), Big Game Management Account through cooperative agreement with the University of California, Davis, supplemented through the Mammalian Ecology and Conservation Unit of the Veterinary Genetics Laboratory at UC Davis. Camilo Sanchez (UC Davis, CDFW), Kathryn Barnitz (Bureau of Land Management [BLM]), Alexia Hemphill (CDFW), Carly White (CDFW) and Julia Owen (UC Davis) assisted with field work. Stevi Vanderzwan (UC Davis) provided laboratory oversight, training, and assistance with laboratory analyses. Dr. Cate Quinn (UC Davis) and Dr. Jennifer Brazeal (UC Davis) provided advice and assistance with species distribution and population modelling, respectively. Sophie Preckler-Quisquater (UC Davis) also provided computational support for species distribution modeling. Russ Landers (CDFW) provided valuable statistical analysis advice. Thanks to Dr. Kristin Denryter (CDFW), Kathryn Barnitz (BLM), and Mario Klipp (CDFW) for offering helpful comments on earlier versions of this document. Finally, thanks to the United States Bureau of Land Management, the United States Forest Service, the United States Bureau of Reclamation, and the many private land owners, all of whom provided access to property to conduct research activities.

Executive Summary

Section I
We sought to develop a broadly applicable methodology for noninvasive DNA-based population monitoring of free-ranging elk (*Cervus canadensis*). Because elk tend to be more heterogeneously distributed across the landscape than other species, such as deer (*Odocoileus* spp), use of a simple random sampling design was expected to result in low survey efficiency. We therefore developed and evaluated a stratification approach involving a species distribution model (SDM) that we intended to be objective yet flexible enough to accommodate variable data sources. We constructed a preliminary SDM for three tule elk (*C. c. nannodes*) populations based on a variety of non-systematic sources of location data; used the model to guide a stratified random fecal pellet field sampling survey; estimated the improvement in detection efficiency obtained from the stratified surveys relative to a simple random sampling design; produced a refined model using a more rigorous procedure based on the systematically

collected detection data; and tested the preliminary and refined models against independent telemetry data and one another.

To develop the preliminary SDM in a ~9,000 km^2 region of Colusa and Lake Counties, we used 1,207 location data points gathered from historical sight surveys, anecdotal reports, and hunter-harvest locations, along with bioclimatic and land cover variables, to produce a presence-based Maxent model. Internal performance measures suggested reasonable discriminatory power of the preliminary model (AUC = 0.889). We then used this model to stratify the study area into 2-km^2 grid cells coded for elk presence as high probability (HP) or low probability (LP), and randomly sampled from each category at a ratio of 2:1, respectively. We surveyed 54 cells (37 HP, 17 LP) 98 times across three populations during Jun–Aug 2017–2019. We recorded 3,629 elk detections. Cross-validation of the preliminary model using these independent elk detections resulted in good model performance (AUC = 0.838). We also observed a strong correlation between model predictions and elk detections per km surveyed (r^2 = 0.87), with an expected range of 2.8–26.0 detections per km corresponding to the model-predicted relative probability of occurrence (RPO) ranging 0–1. Overall, the stratified random surveys resulted in 9.4 detections per km, compared to a predicted yield of 6.9 detections per km had we sampled randomly without stratification, resulting in a 36.7% increase in efficiency.

To assess whether a more rigorous modeling approach would significantly improve performance, we employed a stepwise approach using the systematically collected data to construct and select among refined SDMs. We then tested both the preliminary and top-refined models against one another and with another independent data set, 47,445 telemetry locations from 78 GPS-collared elk (39 male, 39 female). The preliminary and top-performing refined models both exhibited robust discriminatory power as determined by the AUC (AUC_{PRE} = 0.789, AUC_{REF} = 0.780), true skill statistic (TSS) (TSS_{PRE} = 0.445, TSS_{REF} = 0.472), and percent correctly classified (PCC) (PCC_{PRE} = 72%, PCC_{REF} = 78%) evaluation metrics as indicated by telemetered individuals. Spatial comparison of these two models revealed a small (6%) reduction in predicted presence habitat from the preliminary to the refined model. The reasonably high and similar performance between these two models suggests little gain was achieved by the more rigorous modeling approach. Together, these findings suggest that a model-guided approach such as ours utilizing available data, even if non-systematically collected, can qualitatively improve sampling efficiency in surveys designed to collect noninvasive DNA samples for capture-recapture density estimation, and includes additional benefits, such as tangible spatial hypotheses involving environmental relationships and useful baseline information for further inquiries.

Section II

Monitoring trends in abundance of big game species in California has traditionally relied upon air- or ground-based minimum count surveys, which can be affected by visibility biases and unknown precision. In principle, noninvasive fecal DNA (fDNA)-based spatially explicit capture-

recapture (SCR) approaches can provide a statistically robust means of estimating abundance, which has been demonstrated for deer (*Odocoileus* spp). However, fDNA SCR has not yet been widely used for more gregarious species, such as elk (*Cervus canadensis*). Because of their heterogeneous use of the landscape and grouping behavior, elk present novel challenges to sampling efficiency and possibly statistical validity of fDNA SCR. We employed fDNA SCR to estimate abundance in 3 northern California tule elk (*C. c. nannodes*) populations concurrent with but independent of GPS telemetry of 66 elk (32 male (M), 34 female (F)) in Colusa and Lake Counties, California, USA during Jun–Aug 2017–19. We used a species distribution model (SDM) to stratify the landscape for weighted sampling in higher-probability habitat. We collected 1,616 fecal pellet groups from the 3 populations, resulting in 1,002 fDNA genotypes (≥19 microsatellite loci, 1 sex marker) of 425 unique individuals. Based on SCR estimates from a model incorporating both sexes, elk density ranged from 0.31 (95% CI = 0.17–0.55) elk/km^2 to 1.7 (95% CI = 1.3–2.2) elk/km^2, translating approximately to 650 individuals (evenly split between M and F) among the three populations. The even sex ratio agreed with that directly observed from genotyped pellet groups (510 F, 494 M), and within each population as well. Spatial analyses of telemetry data indicated that activity centers of females, but not males, were clustered on the landscape, a violation of SCR assumptions. One population (Lake Pillsbury) exhibited extreme clustering of females, effectively sharing a single activity center. Comparison of combined-sex models to single-sex models indicated that SCR was robust to spatial clustering of females except in the most extreme case, Lake Pillsbury, and only when females were modeled without inclusion of males. In that case, the estimate of female abundance was considerably higher than other estimates and deemed an overestimate. Thus, the inclusion of both sexes was apparently sufficient to offset biases potentially stemming from aggregation of females. Altogether, our findings suggest SCR methods can be gainfully applied to socially gregarious species such as elk.

Table of Contents

I. **Evaluation of a species distribution model approach to guide noninvasive fecal-DNA surveys for tule elk (*Cervus canadensis nannodes*) in Northern California**

Introduction

Big game population monitoring in much of the western USA has customarily relied upon proxy-based measures of abundance to inform harvest quotas and other management strategies (Falcy et al. 2016). These traditional methods, which include aerial and ground-based minimum count indices, as well as opportunistic visual counts, provide no means of assessing uncertainty (Caughley 1974; McCullough et al. 1994; Bleich et al. 2001; CDFW 2018). Consequently, wildlife agencies have increasingly strived to improve monitoring practices to incorporate statistically robust survey methods (Mason et al. 2006; Clare et al. 2017; Bush et al. 2020). One such approach to estimating abundance is noninvasive genetic capture-recapture (Lukacs and Burnham 2005). This approach, which uses DNA left in the environment (typically hair or fecal material), has been recently integrated into monitoring programs for a variety of species, most notably mule deer (*Odocoileus hemionus*) and black-tailed deer (*O.h. columbianus*), for which it has proven especially useful in environments where direct observation through traditional means is impractical (Brinkman et al. 2011; Lounsberry et al. 2015; Furnas et al. 2018; Brazeal et al. 2017; Furnas et al. 2018, 2020). Thus far, however, similar methods have not been widely applied to more gregarious species, such as elk (*Cervus canadensis*).

Relative to deer, elk are more heterogeneously distributed across the landscape and are typically more gregarious, yet generally occur in lower frequencies, presenting unique challenges to sampling strategies (Mackie 1970; Jessup et al. 2014). For instance, most of the landscape may be unoccupied by elk, which impedes application of standard random sampling methods, which could result in high costs and countless hours of survey effort in areas unoccupied by elk that might otherwise have been excluded *a priori* (Stroud et al. 2014). On the other hand, using subjective criteria to exclude certain areas from sampling without independent verification risks biasing estimates (Lancia et al. 2005). Thus, the goal of the present study was to employ a species distribution model (SDM) approach to predict elk distribution for the purpose of stratifying the landscape to guide unbiased and efficient surveys (Guisan et al. 2013; Fois et al. 2015; Fois et al. 2018).

Traditionally, SDMs are created with presence-absence data. However, an increasingly popular alternative, Maxent, which uses presence-only data, performs well compared to other SDM approaches (Phillips et al. 2006; Baldwin 2009; Renner and Warton 2013; Law et al. 2017). The Maxent software applies maximum-entropy principles towards modeling species distributions based on a set of predictive environmental variables, and random background points as a surrogate for true absence data (Phillips et al. 2006).

To broadly apply an SDM approach, it is important that modeling be as standard as possible. However, sources of available data and particular environmental predictors differ among study populations and sites (Virgili et al. 2017; Bucklin et al. 2015), which potentially limits the degree to which modeling procedures can be standardized. By keeping models relatively simple, for example, using no more than 3–6 predictor variables, it may be possible to achieve the model's

purpose – enhancing the efficiency of surveys – despite sacrificing some realism that might be achieved with more sophisticated variable sets and procedures (Halvorsen et al. 2016).

In this study we explored the robustness of readily available data and a simple modeling technique and compared and evaluated models built simply from opportunistic data versus based on systematically collected data and a more rigorous modeling process. We conducted this study on three populations of tule elk (*C. canadensis nannodes*) in Colusa and Lake Counties, CA (Fig. 1). Our approach with Maxent was as follows: (1) use a combination of available data to generate a preliminary SDM for elk, (2) dichotomize the preliminary model and use it to guide a systematic, stratified random survey, (3) test the predictions of the preliminary SDM against the presence data collected during sampling and estimate the efficiency gained by stratification, (4) generate a refined SDM with the presence data collected during field surveys, (5) compare the predictive capabilities between the preliminary and refined SDMs against each other and an independent data set composed of telemetry locations from 78 individual elk.

Study Area

The study area consists of the Cache Creek, Bear Valley, Lake Pillsbury, and East Park Reservoir management units (MUs) as well as the entirety of Lake County, which occurs within approximately 9,000 km^2 of California's Coast and Interior Coast mountain ranges (CDFW 2015; CDFW 2018) (Fig. 1). The climate is described as Mediterranean with hot, dry summers and mild, wet winters (Kauffman 2003). Year-round temperatures generally range from below 0 ^{o}C in the winter to summer daytime temperatures exceeding 38 ^{o}C (Phillips 1976; CDFW 2018). Average annual precipitation is ~76 cm, most of which occurs from October through May (Ferrier and Roberts 1973; BLM 1986). Major year-round water sources include Cache Creek and Bear Creek, Lake Pillsbury and the Eel River, and East Park Reservoir and Stony Creek in the Cache Creek/Bear Valley, Lake Pillsbury, and East Park Reservoir MUs, respectively (Phillips 1976; CDFW 2018). Most of the topography involves rugged, broken terrain, with rolling foothills and flats interspersed throughout jagged peaks and valleys. Elevation ranges from 30 m in the plains of the Sacramento Valley to 2,176 m at Black Butte Mountain. The dominant vegetation communities are typical of the California Coastal range which include blue oak (*Quercus douglassii*) woodland, perennial grassland, chamise (*Adenostoma fasciculatum*)-chaparral, mixed conifer and hardwood forests, annual grassland, and agricultural pastures (McCullough 1969; O'Connor and Guitierrez 1986; O'Connor 1987; BLM 1982; CDFW 2018).

Methods

Population Range Boundary Estimates

We defined range boundaries for each population via reference of historical survey data from the United States Bureau of Land Management (BLM) (BLM 1972, 1977, 1978, 1979, 1980, 1981, 1982, 1983, 1985, 1986, 1989, 1992; Ferrier 1972a-d), the CDFW (Bower 1956; Brandvold 1969; Booth et al. 1988; California Department of Fish and Game [CDFG] 1974, 1978, 1980, 1982, 1987, 1989, 1991, 1995, 1998, 2002, 2004; CDFW 2018; Conover 1972; Curtis 1982; Smith 1973; Bush et al. 2020), and formal studies (McCullough 1969; Ferrier and Roberts 1973; Phillips

1976; O'Connor 1987). To generate contemporary borders, we digitized historical estimated range boundaries in GoogleEarth Pro (v. 7.3.2; Google Inc., 2005), converted the keyhole markup language (.kml) to polygon in ArcGIS (v. 10.4; ESRI 2011), and combined historical estimated range boundaries to form the greatest sized polygon for each population independent of one another. Digitization and consolidation of historical range boundary estimates and more recent observation data resulted in the production of range boundary estimates of 706 km^2 at Cache Creek, 189 km^2 at Lake Pillsbury, and 200 km^2 at East Park Reservoir (Fig. 1).

Selection of Environmental Variables for the Preliminary and Refined Models

We selected covariates from a pool of 27 environmental variables including abiotic, bioclimatic, ecological, topographic, and vegetation data (Supplementary Table 1). We referenced all raster layers using the WGS84 geographic coordinate system and resampled each layer to a resolution of 15 arc seconds (~0.50 km^2). We used the 19 bioclimatic variables from the WorldClim database (e.g., annual mean temperature, isothermality, annual precipitation, etc.) (Hijmans et al. 2005). We derived three topographic layers (elevation, slope, and aspect), as well as two vegetation layers (existing vegetation cover, existing vegetation height) from the Landscape Fire and Resource Management dataset to characterize terrain and vertical vegetation structure (LANDFIRE 2013a-b). We created a distance to water layer using the Model Builder tool in ArcGIS (v. 10.4) derived from the USA Surface Water Dataset (ESRI 2013). We used an available data set to characterize soil taxa (Web Soil Survey 2017), and a habitat layer derived from the California Wildlife Habitat Relationships (CWHR) database compiled by the California Department of Forestry and Fire Protection (CalFIRE 2015).

We used two separate approaches to select the variables for the preliminary and refined models, both of which used an "unsupervised" fitting procedure (e.g., Seoane et al. 2005). For the preliminary model, we began with the full dataset of 27 variables and allowed the internal regularization function within the Maxent algorithm to determine the most informative variables. We included or excluded variables depending on training gain, which indicates relative predictive power (Kumar et al. 2014). We ran multiple iterations of Maxent to explore variable contributions to achieve a reduced subset of environmental covariates. For example, we excluded variables that contributed <10% to the model in the previous run to achieve the most parsimonious output (Baldwin 2009; Kumar et al. 2014).

For the refined model, we again started with the full dataset of 27 variables. However, rather than allow Maxent's internal regularization function to determine the most informative variables, we first reduced the number of variables to the least inter-correlated using the vif.step function in Program R package *usdm* (Naimi 2015) to calculate the variation inflation factor (VIF) for all variables. The VIF is an index of collinearity derived from regressing the predictor variable against all other covariates. We used the vif.step function to exclude all variables with a VIF >4 (O'Brien 2007; Henseker et al. 2015).

Occurrence Records for the Preliminary Model

We compiled known detections of elk presence within the study area through four primary means to provide presence data for production of the Maxent model: historical survey data

(<2005), contemporary survey data (2005–2015), hunter harvest reports (2011–2015), and non-standardized road and ground surveys (2016). We considered elk presence to include direct observations of elk, elk harvests, and detections of elk sign including tracks, pellets, bedding areas, hair, antler casts, carcass remains/roadkill, rubs, and scrapes. Because these data were sourced over a cumulatively long period and using different approaches, we considered the data independent. Although we acknowledge these data did not constitute a random sample, they were the best data we could obtain and provided a good test case for the types of data likely to be more broadly available to other studies (Virgili et al. 2017).

Preliminary Model and Stratification of the Study Area

We used Maxent (v. 3.4.1; http://www.cs.princeton.edu/~schapire/maxent/) to produce the preliminary model (Phillips et al. 2006). While Maxent is capable of fitting highly complex models, simpler models are more interpretable, less vulnerable to model overfitting, and less sensitive to sampling bias (Yackulic et al. 2013). Keeping models simple was also important to our goal of keeping the procedures broadly applicable. Feature classes in Maxent determine the types of constraints and complexity allowed by a model, and can either be used singly or in various combinations (Baldwin 2009; Brown et al. 2017). We applied the least complex feature types, linear and quadratic, to produce the model (Syfert et al. 2013). We applied 10,000 random background ("pseudo-absence") points within the entirety of the study area (e.g., Fig. 1) and left all other parameters to the default Maxent setting (Pearson et al. 2007; Barbet-Massin et al. 2012). We used the cross-validate function to partition testing and training data, and also used a jackknife procedure to assess model performance for each covariate (Pearson et al. 2007).

We tentatively measured model performance using the receiver operating characteristic (ROC) curve, specifically, the area under the ROC curve (AUC) (Phillips et al. 2006), with the *evaluate* function in the R package *dismo* (Hijmans et al. 2017). The AUC is a threshold-independent measure interpreted as the probability that a randomly chosen presence location is ranked higher than a randomly chosen background location (Merow et al. 2013). The AUC ranges from 0.5 to 1, where 0.5 shows no discrimination between presence and background points (i.e., the model result is no better than a random selection), and 1 shows the highest level of discrimination (i.e., the model predictions are perfect); generally, an AUC ≥ 0.7 indicates good discriminatory power (Hosmer and Lemeshow 1989). Overfitting of models is indicated when the AUC estimated from the same data used to train the model (AUC_{TRAIN}) is substantially higher than the AUC estimated from data left out of the training set used in a cross-validation test (AUC_{TEST}). Therefore, we calculated both AUC_{TRAIN} and AUC_{TEST} and used the difference as an indication of overfitting.

We expressed model predictions using the logistic transformation to produce a relative probability of occurrence (RPO), estimates analogous (but technically not identical; Yackulic et al. 2013) to probability of occurrence (Baldwin 2009). We also converted the continuous SDM predictions into a binary surface of predicted presence versus predicted absence pixels (Hirzel et al. 2006). We chose a threshold designed to optimize between maximizing sensitivity (the proportion of presences correctly predicted) and minimizing the area of predicted presence (analogous to false positive rate or 1 – specificity [the proportion of absences correctly

predicted]; Phillips et al. 2006; Quinn et al. 2018). Specifically, we selected the maximum difference of the proportion of occurrence localities versus unconstrained (i.e., across the whole study area) background locations (Engler et al. 2004).

Using the Preliminary Model to Design Validation and Refinement Survey

To stratify the study area for sampling into high probability (HP) and low probability (LP) 2 km^2 sampling cells, we used a procedure analogous to that used to dichotomize the model. Specifically, we classified each cell according to its cumulative within-cell average RPO into one of two strata: HP (average RPO ≥ binary threshold) or LP (average RPO < binary threshold), where the binary threshold was the same one described above. To select a pool of 90 candidate cells for sampling, we randomly selected 60 HP cells and 30 LP cells for a 2:1 (HP:LP) ratio to favor cells expected to have more elk. Ultimately, the numbers and ratios varied somewhat due to constraints posed by accessibility, e.g., permission granted by private landowners and cooperative agreements with state and federal agencies.

Field Sampling

Once sampling cells were selected, we conducted foot surveys during Jun–Sep of three years, 2017–2019. During surveys, we collected fecal pellets, which were used for estimating abundance via mark-recapture analysis in a companion study (Batter et al., Ch 2), and, in this study, as indicators of elk presence. We additionally recorded direct observations of elk, hunter-harvested elk, and other elk sign, including tracks, bedding areas, hair, antler casts, carcass remains/roadkill, rubs, and scrapes. We surveyed cells using 6-km triangle transects whenever possible, but used 4.5-km out-over-and-back parallel transects when terrain or land-use permission did not permit triangular routes (Appendix A). During sampling, we used a hand-held Global Positioning System (GPS) device to record survey tracks and the location of each elk detection.

Model Performance and Efficiency

To compare the preliminary SDM model against survey detections, we measured model performance using the AUC as described above, except that we substituted the independent detection data set for the input data. Second, we directly quantified the correlation between model predictions (RPO) and the number of detections per km surveyed. We used detections per km surveyed to quantify the overall efficiency of the survey. To estimate the gain in efficiency obtained through our stratified random sampling versus a simple random design, we first had to estimate the efficiency that would have been obtained through a simple random design. To do so, we extracted the RPO value for each of the 10,000 preliminary background points and multiplied the proportions of locations falling in each of 5 RPO classes (RPO <0.1, 0.1–0.3, 0.3–0.5, 0.5–0.7, 0.7–1.0) by the corresponding number of detections per km surveyed in that range and by the total number of km surveyed; we then summed these to obtain the estimated number of detections that would have been obtained had we employed the same sampling effort randomly. The increase in efficiency was determined as the ratio of the number of detections recorded divided by this estimate minus 1.

Refined Model

Because the preliminary model was necessarily based on opportunistic data sources, we also wished to assess whether a more rigorous approach to modeling based on the systematically collected data might have qualitatively increased the value of the model for guiding surveys. We modeled the presence data collected during our 3-year survey using Maxent implemented in the R package *ENMeval* (Muscarella et al. 2014). This modeling process was similar to the preliminary modeling except that we made several refinements with the goal of optimizing model accuracy. First, to avoid biases from use of unrepresentative pseudo-absence data (Lobo et al. 2008; Elith et al. 2011; Royle et al. 2012), we constrained the background extent to limit model training closer to areas we surveyed (i.e., regions where we could have reasonably detected elk if they were present). We created a 10 km buffer layer around occurrence points from the full dataset in ArcGIS (v. 10.7.1; ESRI 2019) and populated the region with 10,000 random points within the aggregate buffered area for use as background points in model production and analysis (Barbet-Massin et al. 2012). Otherwise, our approach involved the following steps: (1) construct 4 datasets with varying degrees of thinning to ensure independence, (2) partition each of these datasets into a user-defined and a checkerboard scheme, totaling 8 competing datasets, (3) construct models using varying feature class and regularization multiplier (*β*) combinations using each dataset, and (4) select the best model for each of the 8 datasets.

Thinning.—In contrast to our first dataset for which independence of data points could be reasonably assumed in most cases, our survey data used for the refined model were not spatially independent and therefore required us to thin data to avoid pseudo-replication, which can lead to model overfitting and inaccurate predictions (Phillips et al. 2009). We used the R package *spThin* (Aiello-Lammens et al. 2015) to reduce the full dataset into subsets spatially filtered at values of 100 m, 500 m, and 1 km each (see below).

Partitions.—Next, to divide data for training and testing, we binned each of the 4 thinned datasets into two *a priori* masked geographically structured partitions (Muscarella et al. 2014). We first subdivided the study area into a two-scale checkerboard grid using the '*checkerboard2*' partitioning method (Radosavljevic and Anderson 2014). Occurrence and background points were assigned to 4 bins (Supplementary Fig. 1A). For the user-defined partition, we *a-priori* clustered 4 groups based on MUs with the '*user*' partition method. We assigned occurrence and background points into one of four groups based on the internally identified mean centroid of each cluster (Supplementary Fig. 1B).

Feature classes and regularization.—We considered 3 feature class combinations with the optimum regularization multiplier (*β*) value: linear (L), linear-quadratic (LQ), and linear-quadratic-product (LQP). The *β* multiplier is a coefficient designed to help reduce model overfitting by imposing a penalty for each term included in the model (Tibshirani 2011; Radosavljevic and Anderson 2014). Here we considered a *β* multiplier range from 0.5-7 with steps of 0.5.

Model selection.—To select the best model within each of the 8 datasets resulting from the thinning and partitioning, we used Akaike's Information Criterion, adjusted for small sample size (AICc; Hurvich and Tsai 1989), and created a pool of eight competing models from which to measure and compare overall model performance.

Model Evaluation and Comparison

Although internal validation measures (i.e., using some portion of the input dataset; Merow et al. 2013) are useful for assessing fits of alternative models during the model-construction process, they often exhibit the same biases as the data used to train the model and are therefore insufficient on their own to assess model accuracy (El-Gabbas and Dormann 2018). For this reason, use of independent data to test model performance is ideal (Halvorsen et al. 2016). To test our suite of models, we obtained an independent dataset of telemetry locations from 78 GPS-collared tule elk (39 male, 39 female) collected within the study region via satellite during December 2016–October 2019. Capture procedures involved ground-based free-range darting/chemical immobilization and air-based net-gunning/manual restraint of adult male and female elk. All capture activities were performed by CDFW personnel and followed guidance and approval from the CDFW Wildlife Investigations Laboratory (CDFW 2018). Elk were fixed with GPS Collars (Model: LifeCycle 800 GlobalStar, Lotek Wireless, Newmarket, Ontario, Canada) programmed to collect latitude and longitude of the elk's location every 13 hours and stored via the Lotek WEB Service (http://www.webservice.lotek.com). Collar data were downloaded and aggregated into one dataset. We also used the kernel density spatial analyst tool in ArcGIS (v. 10.7.1; ESRI 2019) to create a smoothly tapered surface to each point across all three populations (i.e., the "realized distribution") for visual comparison against detections recorded during field sampling.

In addition to testing with independent data, simultaneous use of different evaluation metrics is also recommended, particularly when true absence data are unavailable (Jiménez-Valverde 2012). While the AUC is useful in measuring how well presence locations can be discriminated from absences based on predictor variables, it provides little information on how well the model predictions fit the species distribution (Lobo et al. 2008). Therefore, in addition to the AUC, we evaluated models against the independent telemetry data set using the true skill statistic (TSS) (Allouche et al. 2006), and percent correctly classified (PCC) evaluation metrics (West et al. 2016; Luo et al. 2017).

In contrast to the AUC, the TSS is a threshold-dependent measure of the accuracy of a pre-determined binary SDM which places equal weight on model sensitivity and specificity (Allouche et al. 2006). Furthermore, under certain assumptions and with large sample sizes, this measure is independent of prevalence (the proportion of sites in which the species was observed present) (Somodi et al. 2017). Values can range from -1 to 1, with scores closer to 1 indicating greater model performance (West et al. 2016). Models with TSS scores ≥0.40 are considered to perform well (Landis and Koch 1977; Soultan and Safi 2017).

The PCC is a simple proportion of test observations correctly classified based on a predetermined threshold of a binary SDM (VanDerWal et al. 2012). Similar to AUC, PCC is a proportional index; resultant values closer to 100% indicate better model performance (York et al. 2011). We calculated the AUC as described above for the preliminary and refined models, and calculated the TSS and PCC using the *accuracy* function in the R package *SDMtools* (VanDerWal et al. 2012).

We thinned the telemetry data with the R package *spThin* to reduce the full independent datasets into three subsets matching the resolution of the calibration dataset for validation

(100 m, 500 m, and 1 km). Lastly, we used the ArcGIS toolkit *SDMtoolbox* (Brown et al. 2017) to visualize and quantify changes in binary distribution predictions between the preliminary and refined model (Khosravi et al. 2016).

Results

Preliminary SDM

We gathered 1,207 tule elk presence data points for the preliminary model input (Table 1). We identified a combination of 3 covariates through the internal regularization algorithm that best predicted tule elk distribution: habitat type (CalFIRE 2015), vegetation cover (LANDFIRE 2008), and mean diurnal temperature range ("bio2" in the WorldClim dataset). The resultant preliminary model yielded an AUC_{TRAIN} value of 0.889 and an AUC_{TEST} value of 0.885 supporting good model performance (Phillips and Dudik 2008). The small difference between these values (0.004) suggested the model was not overfit. The contribution of preliminary covariates in order of permutation importance was bio2, habitat type, and vegetation cover (Supplementary Table 2). The predicted RPO increased positively with bio2 (Supplementary Fig. 2A), and tended to be higher in perennial grassland, hardwood-conifer, lacustrine, annual grassland, and blue oak-valley oak habitat types (Supplementary Fig. 2B), and with increasing herb cover and reduced shrub and tree cover (Supplementary Fig. 2C). Visual inspection of the predictive surface revealed several concentrated areas of high RPO, connected by corridors corresponding approximately to waterways (Figs. 2A, Supplementary Fig. 3).

We next identified the optimal binary classification threshold at RPO = 0.15, which resulted in a dichotomized binary map (Fig. 2B) that we used to classify 2-km^2 cells into high probability (HP) or low probability (LP) cells (Fig. 2C). In total 487 (25%) of 1,944 cells were classified as HP.

We sampled 54 cells 98 times across all three populations over three survey years: 33 cells (24 HP; 9 LP) at Cache Creek (in 2017 and 2019; 66 surveys), 11 cells (7 HP; 4 LP) at Lake Pillsbury (2018, 2019; 22 surveys), and 10 cells (6 HP; 4 LP) at East Park Reservoir (2018; 10 surveys). Our field validation efforts produced 3,629 elk location points (Table 2). Elk were detected at Cache Creek in 22 (92%) of 24 HP cells and 5 (55%) of 9 LP cells across both survey years; in all cells at Lake Pillsbury across both survey years; and in all cells at East Park Reservoir in 2018.

Cross-validation of the independent elk detections with the preliminary logistic surface supported good model performance (AUC = 0.838), only slightly lower than that estimated using the cross-validated input data set (i.e., AUC = 0.885). Additionally, the number of detections per km surveyed were positively correlated with the RPO ($r = 0.933$, $r^2 = 0.87$, $F_{1,3} = 20.2$, $P = 0.021$; Fig. 3). The expected number of detections per km surveyed ($\hat{y}$) increased significantly as a function of RPO (x), resulting in a range of expected efficiencies of 2.8–26.0 detections per km corresponding to the observed range of model RPO on the study area.

In total, we obtained the 3,629 detections over 386.5 km of transect surveyed, for an estimated efficiency of 9.39 detections per km surveyed. Using the random background points to indicate composition of the study area with respect to the RPO ranges and the observed numbers of detections per km surveyed in each of the RPO ranges, we estimated that a simple random survey of 386.5 km of transects would have yielded 2,655 elk detections, resulting in 6.87

detections per km surveyed. Thus, a 36.7% increase in efficiency was attributable to the stratified design we employed (i.e., 9.39/6.87 – 1).

Comparison of the elk detections in our survey to the independent kernel density of all telemetry locations (n = 47,445) support the representativeness of our survey (Fig. 4). Elk detections across all three populations were in relatively high agreement with areas of elk use, and increased comparably within concentrated high-use areas (Fig. 4A–C). The 4 sampled cells where we did not detect elk also did not have any telemetry occurrences within them (Fig. 4C). Finally, all elk detections superimposed on the binary predictive surface occurred ≤500 m from predicted presence classified pixels (Fig. 5).

Refined SDM

To produce the refined model, we first eliminated all variables with a VIF > 4, which resulted in the following 6 least inter-correlated variables: mean diurnal temperature range (bio2), distance to water, habitat type, slope, soil, and vegetation cover. We thinned the 3,629-elk occurrence dataset to 1,169, 397, and 208 occurrence points for 100 m, 500 m, and 1 km subsets, respectively. As we increased spatial thinning distance, model performance declined for both partition methods (i.e., the full model outperformed the thinned models). The 'checkerboard2' models yielded an AUC_{TRAIN} range of 0.814–0.858, an AUC_{TEST} range of 0.764–0.845, and an AUC_{DIFF} range of 0.033–0.050, indicating good model performance across this class of partition type (Table 3). On the other hand, the 'user' models performed poorly in terms of model overfitting. While the range of the 'user' AUC_{TRAIN} (0.816–0.885) values indicates good model performance for each dataset, the ranges of AUC_{TEST} (0.500–0.615) and AUC_{DIFF} (0.242–0.333) values indicates a severe vulnerability to model overfitting, and renders these models uninformative.

To evaluate the models using the independent telemetry dataset, we first thinned the data to the same resolution as the occurrence dataset used to construct the Maxent models, resulting in 7,580, 924, and 323 location data points for 100 m, 500 m, and 1 km resolutions, respectively. We identified the optimal binary classification threshold for each of the eight competing models to measure model performance (Table 4). The preliminary model performed well as indicated by AUC_{FULL} (0.789), AUC_{BINARY} (0.722), and TSS (0.445) values, and was able to correctly predict 73% of the data over the binary surface (Table 4). The 'checkerboard2' models performed moderate-to-well, as indicated by the range of AUC_{FULL} (0.751–0.786), AUC_{BINARY} (0.685–736), and TSS (0.371–0.472) values, and correctly predicted between 64–78% of the data. Despite the vulnerability to model overfitting, the 'user' models performed similarly to the 'checkerboard2' models when tested with the independent data. The 'user' models yielded ranges of AUC_{FULL} (0.757–0.784), AUC_{BINARY} (0.689–724), and TSS (0.379–0.447) values, which support moderate-to-good model performance, and was able to correctly predict between 65–75% of the data over the binary surface. Overall, the top performing refined model was the 'checkerboard2'-500 m model according to the maximum TSS score (0.472), and its performance was supported by the AUC_{TRAIN} (0.832) and AUC_{TEST} (0.795) values. This model retained the desired qualities of a good model –low overfitting (AUC_{DIFF} = 0.033) and high discriminatory ability (78% proportion correctly classified).

The contribution of refined covariates in order of permutation importance for our best-fit refined model was soil, bio2, vegetation cover, habitat type, distance to water, and slope (Supplementary Table 4). The importance of soil in the refined model contrasted with the preliminary model, which did not include this variable. Otherwise, however, the models were similar. As with the preliminary model, the RPO exhibited a positive relationship with bio2, vegetation cover, and habitat type, and the predictive values for the specific parameters showed similar response curves to the preliminary model (Supplementary Fig. 4). The predicted RPO increased positively with bio2 (Supplementary Fig. 4A), and was much greater in perennial grassland, followed by hardwood-conifer, montane riparian, and lacustrine habitat types (Supplementary Fig. 4B), and with increasing herb and shrub cover and decreasing tree cover (Supplementary Fig. 4C). Soil type, which was both the highest contributing variable and of greatest permutation importance, also had the greatest positive relationship with predicted RPO (38% of soil types had an RPO > 0.40), the highest of which was the Riverwash-Orland-Los Robles-Cortina soil type (RPO = 0.93) (Supplementary Fig. 4D). Distance to water (m) had a slightly negative relationship with predictive ability, and slope did not have an effect on predictive ability.

Visual inspection of the predictive surface revealed several core areas of high RPO, which, as with the preliminary models, corresponded to the MUs (Fig. 6A/Supplementary Fig. 5). Corridors of higher RPO again appear to connect core suitable habitat across the study region, although to a lesser degree compared to the preliminary model. We produced a raw binary surface using the classification threshold 0.35 (Fig. 6B), and transformed the map to a binary grid surface and assigned HP and LP classifications (Fig. 6C). In total, 279 cells were classified as HP (14% of all cells).

To visualize the difference between the preliminary and refined model surface predictions, we produced a map that identified binary distribution prediction changes across these two models (Fig. 7). Predicted presence described ~2,790 km^2 (31%) of the preliminary model and ~2,250 km^2 (25%) of the refined model, representing a 6% reduction (~540 km^2) in the fraction of the landscape in which elk were predicted to occur. Model predictions of elk presence differed the least in and around the three populations and the most in the intervening regions potentially connecting them. In both models, the majority of the study area was predicted to be absent of elk.

Discussion

Our goal was to investigate the use of a SDM approach to stratify the landscape for more efficient systematic sampling of a heterogeneously distributed species. In particular, we sought to evaluate the approach with respect to use of opportunistic versus systematic sources of data, predictive accuracy, and gain in sampling efficiency. Given the purpose of models, their most important attribute was to accurately identify portions of the landscape least likely to contain elk. Although we constructed models using two different data sets and procedures, models resulting from both efforts identified most (>69%) of the landscape as low probability, and for the most part, overlapped in these predictions. Because our study was part of a broader research effort (Bush et al. 2016), we were uniquely positioned to take advantage of a completely independent data set afforded by 78 telemetered elk to independently validate

each model's predictive ability, a step recommended but rarely practiced for SDM validation (Greaves et al. 2006; West et al. 2016; Law et al. 2017; Rhoden et al. 2017; Wauchope-Drumm et al. 2020). Comparison of these models to over 45,000 telemetry locations demonstrated them to be reasonably accurate, and sufficient to increase the efficiency of a stratified random design 36% relative to a simple random sample of the landscape. However, the similarity of the surfaces despite incorporation of different variables also underscores the need for caution when interpreting these predictive models as causal models. For example, soil was the most predictive variable in our refined model yet was absent from the preliminary model.

Comparison of the predicted RPO surfaces between the preliminary and refined models revealed differences in spatial predictions. The refined model included more constraints (i.e., soil, distance to water, and slope) on the data compared to the preliminary model, and, as expected, its output indicated more restrictive RPO in geographic space when compared to the preliminary model (Royle et al. 2012; West et al. 2016). Even with this reduction in predicted RPO, our results did not yield a significant increase in model performance from the preliminary to refined outputs either in terms of RPO or sampling efficiency, likely a result of the large initial sample size (n = 1,207) and the ecologically narrow-range of tule elk, both factors known to generate better model performance compared to the respective alternatives (Tessarolo et al. 2014). Because historical occurrence datasets are frequently accessible to wildlife management agencies (Virgili et al. 2017) this result is welcoming, as it indicates existing data can be readily applied towards reliable SDM production, assuming the existing dataset is sufficiently large for its specific geographic extent (Wisz et al. 2008; van Proosdij et al. 2015).

While incorporation of systematic sampling and model refinement into the overall framework is important to improve spatial predictions (Quinn et al. 2018; Wauchope-Drumm et al. 2020), survey effort is often limited by time, funding, and personnel allocation (CDFW 2013). The ability to confidently generate reliable SDMs with pre-existing datasets, as demonstrated with our preliminary model, affords wildlife managers an additional tool that can be applied towards species management with less time and resource investment. Moreover, once a model is applied for sampling purposes, continued refinement based on resulting survey data, as in this study, can lead to improved accuracy and, therefore, efficiency (West et al. 2016; Law et al. 2017; Rhoden et al. 2017).

Understanding assumptions and possible vulnerabilities of our modeling process is important towards future applications of model-guided surveys and the model refinement processes. For our preliminary model, we relied heavily on historical occurrence localities, data most commonly possessed and utilized by producers of presence-only SDMs (Virgili et al. 2017). Although many of the reported localities were obtained through resource management agencies, the accuracy of some locations was questionable, and in some cases, particularly the hunter-harvest locations, were a best guess based on limited information provided. Although Maxent is somewhat robust to spatial error of location data (Baldwin 2009), use of these historical data resulted in uneven sampling, yielding clusters of occurrences in areas more easily or frequently surveyed, areas known to be of high elk use, or a combination of the two (e.g., Bower 1956).

The results of our refined model, in which we tested a range of thinned datasets, similarly supported minimal impact from spatial clustering of data. While focal sampling within range boundaries estimated *a priori* could result in inflated evaluation metrics (e.g., Westwood et al. 2020), testing predictive ability with independent GPS collar data indicated good model performance, and visual assessment showed high agreeance between field detections and telemetry data for both the preliminary and refined models (Figs. 4, S5, S6). In comparison to the putative population ranges, only ~2% of the GPS collar locations (the "realized distribution") occurred an average of 0.78 km outside the estimated boundaries, thus supporting our decision to survey only within these areas. By concentrating sampling efforts within these estimated range boundaries we were able to reduce the entire survey area by 85%. Nevertheless, elk may have been present in these intermediary predicted presence areas (Williams et al. 2002). To more fully understand interactions across populations, and because connectivity of fragmented populations is essential towards long term viability (Williams et al. 2004; Hilty et al. 2006), future surveys could incorporate deliberate sampling of potential distribution outside presumed range boundaries with little cost to survey efficiency (Elith et al. 2006; de Oliveira et al. 2019).

Additional Benefits of SDMs

In addition to their utility for guiding more efficient density and abundance surveys, SDMs also produce spatial hypotheses useful for testing and evaluating additional demographic processes including, for example, habitat affinities and population genetics (McCullough et al. 1996; Hilty et al. 2006; Buchalski et al. 2015). These factors, among others, are critical to better understand and effectively manage harvested species, especially those persisting in fragmented populations (Williams et al. 2004; Frankham et al. 2017; CDFW 2018).

A model's ability to produce measures of a species' environmental associations is an integral component of SDM production (Franklin 2010), and allows for qualitative comparison against expert knowledge and known life-history traits of the focal species (Franklin 2013). For example, tule elk have consistently been associated with habitat types and vegetation cover that our predictor variables identified as important, including perennial grassland, riparian, lacustrine, and oak-dominant habitats (i.e., Bower 1956; Brandvold 1969; McCullough 1969; Ferrier and Roberts 1973; Smith 1973; Phillips 1976; O'Connor and Guitierrez 1986; McCullough et al. 1996; CDFW 2018), habitat types generally characterized as more "open," all of which provide required nutrients and promotion of gregarious space use and group vigilance (McCullough 1969; Raedeke et al. 2002). Furthermore, integration of SDMs with other monitoring tools, such as noninvasive genetic sampling, can lend further insight into improving understanding of elk populations. Quantification of genetic variation can illuminate the genetic patterns underlying population dynamics, uncover barriers and linkages to population connectivity, and allow for inference of dispersal patterns and potential range expansion (Hilty et al. 2006; Hicks et al. 2007; Onorato et al. 2007; Sun et al. 2017; Frankham et al. 2017; Quinn et al. 2018).

Management Implications

Our study demonstrated that development and use of a SDM can significantly increase efficiency of sampling for elk by using it to stratify the study area to focus sampling in high-

probability areas. Moreover, while it is ideal to construct the most rigorous model possible—and our study provides one possible set of guidelines for doing so—our findings also suggest that use of readily available information, even if not entirely verifiable, can substantially increase the efficiency of elk surveys.

Section I Literature Cited

Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B. and Anderson, R.P., 2019. spThin: Functions for spatial thinning of species occurrence records for use in ecological models. R package version 0.1. 0. 2014.

Allouche, O., A. Tsoar, and R. Kadmon. 2006. Assessing the accuracy of species distribution models: Prevalence, kappa, and the true skill statistic (TSS). Journal of Applied Ecology 43(6):1223-1232.

Baldwin, R.A. 2009. Use of maximum entropy modeling in wildlife research. Entropy 11:854-866.

Barbet-Massin, M., F. Jiguet, C.H. Albert, and W. Thuiller. Selecting pseudo-absences for species distribution models: how, where and how many? Methods in Ecology and Evolution 3:327-338.

Batter, T.J., J.P. Bush, and B.N. Sacks. In prep. Application of noninvasive fecal-DNA and spatial capture-recapture to estimate tule elk Cervus canadensis nannodes abundance in Northern California. Report to the California Department of Fish and Wildlife.

Bleich, V.C., C.S.Y. Chun, R.W. Anthes, T.E. Evans, and J.K. Fischer. 2001. Visibility bias and development of a sightability model for tule elk. Alces 37(2):315-327.

Booth, J., J. Swanson, and D. Koch. 1988. Cache Creek tule elk management unit management plan. March 31, 1988. California Department of Fish and Game, Sacramento, CA.

Bower, J. 1956. Survey of the Cache Creek elk herd Colusa and Lake Counties. Game Management Branch, California Department of Fish and Game, Region II. Unpublished report.

Brandvold, G. 1969. Cache Creek Ridge wildlife habitat area. Unpublished report. Bureau of Land Management, Ukiah, CA.

Brazeal, J.L., T. Weist, and B.N. Sacks. 2017. Noninvasive genetic spatial capture-recapture for estimating deer population abundance. The Journal of Wildlife Management 81(4):629-640.

Brown, J.L., J.R. Bennett, and C.M. French. 2017. SDMtoolbox 2.0: the next generation python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ 5, c4095.

Buchalski, M.R., A.Y. Navarro, W.M. Boyce, T.W. Vickers, M.W. Tobler, L.A. Nordstrom, J.A. Garcia, D.A. Gille, M.C.T. Penedo, O.A. Ryder, and H.B. Ernest. 2015. Genetic population structure of Peninsular bighorn sheep (Ovis canadensis nelsoni) indicates substantial gene flow across US-Mexico border. Biological Conservation 184: 218-228.

Bucklin, D.N., M. Basile, A.M. Benscoter, L.A. Brandt, F.J. Massotti, S.S. Romañach, C. Speroterra, and J.I. Watling. 2015. Comparing species distribution modelsconstructed with different subsets of environmental predictors. Diversity and Distribution, 21(1): 23-35.

Bureau of Land Management (BLM). 1972. Tule elk range inventory. Final monthly report, Oct-Nov. 1972. BLM, Ukiah, CA. Unpublished report.

BLM. 1977. A report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1978. Second annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1979. Third annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1980. Fourth annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1981. Fifth annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1982. Sixth annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1983. Seventh annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1985. Cache Creek tule elk wildlife habitat management plan. United States Department of Interior, Bureau of Land Management, Ukiah District.

BLM. 1986. Eighth annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1989. Ninth report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1992. Tenth report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

Bush, J.P., T.J. Batter, R. Landers, and K. Denryter. 2020. Report on aerial surveys of tule elk (Cervus canadensis nannodes) in 2018-2019 in Bear Valley, Cache Creek, East Park Reservoir, and Lake Pillsbury tule elk hunt zones. Internal report. California Department of Fish and Wildlife, Sacramento, CA.

California Department of Fish and Game (CDFG). 1974. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1978. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1980. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1982. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1987. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA

CDFG. 1989. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1991. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1995. Report to the legislature regarding tule elk. California Department of Fish and Game, Sacramento, CA.

CDFG. 1998. Report to the legislature regarding tule elk. California Department of Fish and Game, Sacramento, CA.

CDFG. 2002 Final Environmental Document regarding elk hunting. April 5, 2002. California Department of Fish and Game. Fig. 24.

CDFG. 2004. Final Environmental Document Regarding Elk Hunting. April 12, 2004. California Department of Fish and Game. 178 pp.

California Department of Fish and Wildlife (CDFW). 2013. Wildlife restraint handbook. Wildlife Investigations Laboratory. California Department of Fish and Wildlife, Rancho Cordova, CA.

CDFW. 2015. California statewide wildlife action plan, 2015 update: a conservation legacy for Californians. Eds: A.G. Gonzales and J. Hoshi. Prepared with assistance from Ascent Environmental, Inc., Sacramento, CA.

CDFW. 2018. Conservation and management plan for elk. State of California Department of Fish and Wildlife. Sacramento, CA. Retrieved from: https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=162912&inline

California Department of Forestry and Fire Protection (CalFIRE). (2015). Fire and Resource Assessment Program (FRAP) Vegetation (FVEG15_1). Sacramento, California. Retrieved from http://frap.fire.ca.gov/data/frapgisdata-sw-fveg_download

Caughley, G. 1974. Bias in aerial survey. Journal of Wildlife Management 38: 921 -933.

Clare, J., S.T. McKinney, J.E. DePie, and C.S. Loftin. 2017. Pairing field methods to improve inference in wildlife surveys while accommodating detection covariance. Ecological Applications 27 (7): 2031-2047.

Conover, M. 1972. Cache Creek elk range distribution. July 1972. California Department of Fish and Game. Unpublished report.

de Oliveira, M.L., H.T. Zarate do Couto, and J.M. Barbanti Duarte. 2019. Distribution of the elusive and threatened Brazilian dwarf brocket deer refined by non-invasive genetic sampling and distribution modelling. European Journal of Wildlife Research 65:21.

El-Gabbas, A., and C.F. Dormann. 2018. Improved species-occurrence predictions in data-poor regions: using large-scale data and bias correction with down-weighted Poisson regression and Maxent. Ecography 41:1161-1172.

Elith, J., C.H. Graham, R.P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A.. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Mortiz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. Townsend Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberon, S. Williams, M.S. Wisz, and N.E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129-151.

Elith, J., S.J. Phillips, T. Hastie, M. Dudik, Y.E. Chee, and C.J. Yates. 2011. A statistical explanation of Maxent for ecologists. Diversity and Distributions 17:43-57.

ESRI. 2011. ArcGIS desktop: release 10.4. Redlands, CA, USA: Environmental Systems Research Institute.

ESRI. 2013. USA Surface Water dataset. Redlands, CA, USA: Environmental Systems Research Institute. Retrieved from: https://landscape2.arcgis.com/arcgis/rest/services/USA_Surface_Water/ImageServer

ESRI. 2019. ArcGIS desktop: release 10.7. Redlands, CA, USA. Environmental Systems Research Institute

Falcy, M.R., J.L. McCormick, and S.A. Miller. 2016. Proxies in practice: Calibration and validation of multiple indices of animal abundance. Journal of Fish and Wildlife Management 7(1): 117-128.

Ferrier, G.J. 1972a. Tule elk range inventory monthly report for June, 1972. Bureau of Land Management, Ukiah District. Unpublished report.

Ferrier, G.J. 1972b. Tule elk range inventory monthly report for July, 1972. Bureau of Land Management, Ukiah District. Unpublished report.

Ferrier, G.J. 1972c. Tule elk range inventory monthly report for August, 1972. Bureau of Land Management, Ukiah District. Unpublished report.

Ferrier, G.J. 1972d. Tule elk range inventory monthly report for October/November, 1972. Bureau of Land Management, Ukiah District. Unpublished report.

Ferrier, G.J, and E.C. Roberts. 1973. The Cache Creek tule elk range. Cal-Neva Wildlife. 10 pp.

Fois, M., G. Fenu, A. Cuena-Lombraña, D. Cogoni, and G. Bacchetta. 2015. A practical method to speed up the discovery of unknown populations using species distribution models. Journal for Nature Conservation 24:42-48.

Fois, M., A. Cuena-Lombraña, G. Fenu, and G. Bacchetta. 2018. Using species distribution models at local scale to guide the search of poorly known species: Review, methodological issues and future directions. Ecological Modelling 385:124-132.

Frankham, R., J.D. Ballou, K. Ralls, M.D.B. Eldridge, M.R. Dudash, C.B. Fenster, R.C. Lacy, and P. Sunnucks. 2017. Genetic management of fragmented animal and plant populations. Oxford University Press, UK.

Franklin, J. 2010. Mapping species distributions: spatial inference and prediction. Cambridge: Cambridge University Press.

Franklin, J. 2013. Species distribution models in conservation biogeography: developments and challenges. Diversity and Distributions 19 (10): 1217-1223.

Greaves, G.J., R. Mathieu, and P.J. Seddon. 2006. Predictive modelling and ground validation of the spatial distribution of the New Zealand long-tailed bat (Chalinolobus tuberculatus). Biological Conservation 132(2): 211-221.

Guisan, A, R. Tingley, J.B. Baumgartner, et al. 2013. Predicting species distributions for conservation decisions. Ecology Letters 16: 1424-1435.

Halvorsen, R., S. Mazzoni, J.W. Dirksen, E. Næsset, T. Gobakken, and M. Ohlson. 2016. How important are choice of model selection method and spatial autocorrelation of presence data for distribution modelling by MaxEnt? Ecological Modelling 328: 108-118.

Hicks, J.F., J.L. Rachlow, O.E. Rhodes, Jr., C.L. Williams, and L.P. Waits. 2007. Reintroduction and genetic structure: Rocky Mountain elk in Yellowstone and the Western states. Journal of Mammalogy, 88(1): 129-138.

Hijmans, R. J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978.

Hijmans, R.J., S. Phillips, J. Leathwick, and J. Elith. 2017. dismo: species distribution modeling. R package version 1.1-4.

Hilty, J.A., W. Lidicker, and A. Merenlender. 2006. Corridor ecology: the science and practice of linking landscapes for biodiversity conservation. Island Press, Washington, D.C.

Hirzel, A.H., G. Le Lay, V. Helfer, C. Randin, and A. Guisan. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological Modeling 199:142-152.

Hosmer, D.W., and S. Lemeshow. 1989. A goodness-of-fit test for the multiple logistic regression model. Communications in Statistics 10: 1043-1069.

Hurvich, C.M., and C.I. Tsai. 1989. Regression and time-series model selection in small samples. Biometrika 76:297-307.

Jessup, D.A., S.R. DeJesus, W.E. Clark, and V.C. Bleich. 2014. Evolution of ungulate capture techniques in California. California Fish and Game 100(3): 491-526.

Jiménez-Valverde, A. 2012. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. Global Ecology and Biogeography 21: 498-507.

Kauffman, E. 2003. Atlas of the biodiversity of California. Climate and Topography. California Department of Fish and Game, Sacramento, CA. pp. 12-15.

Khosravi, R., M.R. Hemami, M. Malekian, A.L. Flint, and L.E. Flint. 2016. Maxent modeling for predicting potential distribution of goitered gazelle in central Iran: the effect of extent and grain size on performance of the model. Turkish Journal of Zoology 40:574-585.

Kumar, S., J. Graham, A.M. West, and P.H. Evangelista. 2014. Using district-level occurrences in Maxent for predicting the invasion potential of an exotic insect pest in India. Computers and Electronics in Agriculture 103: 55-62.

Lancia, R.A., W.L. Kendall, K.H. Pollock, and J.D. Nichols. 2005. Estimating the number of animals in wildlife populations. Pages 106-153 in C.E. Braun, editor. Techniques for wildlife investigations and management. Sixth edition. The Wildlife Society, Bethesda, Maryland, USA.

LANDFIRE. 2013a. Aspect, elevation, and slope layers. U.S. Department of Interior, Geological Survey. Retrieved from: https://landfire.cr.usgs.gov/topographic.php

LANDFIRE. 2013b. Existing Vegetation Cover and Height layers. U.S. Department of Interior, Geological Survey. Retrieved from: https://landfire.cr.usgs.gov/evc.php

Landis, J.R., and G.G. Koch. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159-174.

Law, B., G. Caccamo, P. Roe, A. Truskinger, T. Brassil, L. Gonsalves, A. McConville, M. Stanton. 2017. Development and field validation of a regional, management-scale habitat model: A koala Phasolarctos cinereus case study. Ecology and Evolution 7: 7475-7489.

Lobo, J.M., A. Jimenez-Valverde, and R. Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17:145-151.

Lukacs, P.M. and K.P. Burnham. 2005. Review of capture-recapture methods applicable to noninvasive genetic sampling. Molecular Ecology 14:3909-3919.

Luo, M., H. Wang, and Z. Lyu. 2017. Evaluating the performance of species distribution models Biomod2 and Maxent using the giant panda distribution data. The Chinese Journal of Applied Ecology 28(12): 4001-4006.

Mackie, R.J. 1970. Range ecology and relations of mule deer, elk, and cattle in the Missouri River Breaks, Montana. Wildlife Monographs 20:3-79.

Mason, R., L.H. Carpenter, M. Cox, J.C. Devos, J. Fairchild, D.J. Freddy, J.R. Heffelginer, R.H. Kahn, S.M. McCorquodale, D.F. Pac, D. Summers, G.C. White, and B.K. Williams. 2006. A case for standardized ungulate surveys and data management in the Western United States. Wildlife Society Bulletin 34(4):1238-1242.

McCullough, D.R. 1969. The tule elk: its history, behavior, and ecology. University of California

Publications in Zoology, Berkeley. 88: 209 pp.

McCullough, D. R., F. W. Weckerly, P. I. Garcia, and R. R. Evett. 1994. Sources of inaccuracy in black-tailed deer herd composition counts. Journal of Wildlife Management 58:319-329.

McCullough, D.R., J.D. Ballou and J.K. Fischer. 1996. From bottleneck to metapopulation: recovery of the tule elk in California in Metapopulations and wildlife conservation 375-410. Island Press, Washington, D.C., USA.

Merow, C., M.J. Smith, and J.A. Silander, Jr. 2013. A practical guide to Maxent for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36: 1058-1069.

Muscarella, R., P.J. Galante, M. Soley-Guardia, R.A. Boria, J.M. Kass, M. Uriarte, and R.P. Anderson. 2014. ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods in Ecology and Evolution 5:1198-1205.

Naimi, B. 2015. Usdm: Uncertainty analysis for species distribution models. R package version 1.1-18.

O'Brien, R.M. 2007. A caution regarding rules of thumb for variance inflation factors. Quality and Quantity 41:673-690.

O'Connor, P.M. 1987. Tule elk ecology and home range characteristics at Cache Creek, California. Humboldt State University, Arcata, CA.

O'Connor, P.M., and R.J. Gutierrez. 1986. Tule elk ecology at Cache Creek: report to the California Department of Fish and Game. Humbolt State University. 45pp.

Onorato, D.P., E.C. Hellgren, R.A. Van Den Bussche, D.L. Doan-Crider, and J.R. Skiles, Jr. 2007. Genetic structure of American black bears in the desert southwest of North America: conservation implications for recolonization. Conservation Genetics 8: 565-576.

Pearson, R.G., C.J. Raxworthy, M. Nakamura, and A.T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography 34:102-117.

Phillips, W.E. 1976. The conservation of the California tule elk: a socioeconomic study of a survival problem. Edmonton: The University of Alberta Press. 120 pp.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

Phillips, S.J., and M. Dudik. 2008. Modeling of species distributions with MAXENT: new extensions and a comprehensive evaluation. Ecography 31(2):161-175.

Phillips, S.J., M. Dudik, J. Elith, C.H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19(1):181-197.

Quinn, C.B., J.R. Akins, T.L. Hiller, and B.N. Sacks. 2018. Predicting the potential distribution of the Sierra Nevada red fox in the Oregon Cascades. Journal of Fish and Wildlife Management 9:351-365.

Radosavljevic, A., and R.P. Anderson. 2014. Making better Maxent models of species distributions: complexity, overfitting, and evaluation. Journal of Biogeography 41: 629-643.

Raedeke, K.J., J.J. Millspaugh, and P.E. Clark. 2002. North American elk: ecology and management. D.E. Toweill and J.W. Thomas, eds. Smithsonian Institution Press, Washington, D.C.

Renner, I.W., and D.I. Warton. 2013. Equivalence of Maxent and Poisson point process models for species distribution modeling in ecology. Biometrics 69(1): 274-281.

Rhoden, C.M., W.E. Peterman, and C.A. Taylor. 2017. Maxent-directed field surveys identify new populations of narrowly endemic habitat specialists. PeerJ 5: e3632. Doi: 10.7717/peerj.3632.

Royle, J.A., R.B. Chandler, C. Yackulic, and J.D. Nichols. 2012. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. Methods in Ecology and Evolution 3:545-554.

Seoane, J. J. Bustamante, and R. Diaz-Delgado. 2005. Effect of expert opinion on the predictive ability of environmental models of bird distribution. Conservation Biology 19(2): 512-522.

Smith, E.S. 1973. Effects of three possible reservoir development projects on the Cache Creek tule elk herd. Environmental Services Branch, California Department of Fish and Game. Administrative Report No. 73-2, January, 1973.

Soil Survey Staff. 2013. Natural Resources Conservation Service, United States Department of Agriculture. Retrieved from: https://websoilsurvey.sc.egov.usda.gov/

Soil Survey Staff. 2015. Illustrated guide to soil taxonomy, version 2. United States Department of Agriculture, Natural Resources Conservation Service, National Doil Survey Center, Lincoln, Nebraska.

Somodi, I., N. Lepesi, and Z. Botta-Dukát. 2017. Prevalence dependence in model goodness measures with special emphasis on true skill statistics. Ecology and Evolution 7(3): 863-872.

Soultan, A., and K. Safri. 2017. The interplay of various sources of noise on reliability of species distribution models hinges on ecological specialization. PLoS One 12(11): e0187906.

Stroud, J.T., E. Rehm, M. Ladd, P. Olivas, and K.J. Freeley. 2014. Is conservation research money being spent wisely? Changing trends in conservation research priorities. Journal for Nature Conservation 22: 471-473.

Sun, C.C., A.K. Fuller, M.P. Hare, and J.E. Hurst. 2017. Evaluating population expansion of black bears using spatial capture-recapture. Journal of Wildlife Management 81(5): 814-823.

Syfert, M.M., M.J. Smith, and D.A. Coomes. 2013. The effects of sampling bias and model complexity on the predictive performance of Maxent species distribution models. PLoS One 8(2).

Tessarolo, G., T.F. Rangel, M.B. Araújo, and J. Hortal. 2014. Uncertainty associated with survey design in species distribution models. Diversity and Distributions 20:1258-1269.

Tibshirani, R. 2011. Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73:273-282.

VanDerWal, J., L. Falconi, S. Januchowski, L. Shoo, and C. Storlie. 2012. SDMTools: Tools for processing data associated with species distribution modelling exercises. Retrieved from http://CRAN.R-project.org/package=SDMTools

van Proosdij, A.S., M.S.M. Sosef, J.J. Wieringa, and N. Raes. 2015. Minimum required number of specimen records to develop accurate species distribution models. Ecography 39:542-552.

Virgili, A., M. Racine, M. Authier, P. Moneaties, and V. Ridoux. 2017. Comparison of habitat models for scarcely detected species. Ecological Modeling 346: 88-98.

Wauchope-Drumm, M., J. Bentley, L.J. Beaumont, J.B. Baumgartner, and D.A. Nipperess. 2020. Using a species distribution model to guide NSW surveys of the long-footed potoroo (Potorous longipes). Austral Ecology 45:15-26.

West, A.M., S. Kumar, C.S. Brown, T.J. Stohlgren, and J. Bromberg. 2016. Field validation of an invasive species Maxent model. Ecological Informatics 36: 126-134.

Westwood, R., A.R. Westwood, M. Hooshmandi, K. Pearson, K. LaFrance, and C. Murray. 2020. A field-validated species distribution model to support management of the critically endangered Poweshiek skipperling (Oarisma poweshiek) butterfly in Canada. Conservation Science and Practice 2:e163. DOI: 10.1111/csp2.163.

Williams, B.K., J.D. Nichols, and M.J. Conroy. 2002. Analysis and management of animal populations. Academic Press. San Diego, CA, USA.

Williams, C.L., B. Lundrigan, and O.E. Rhodes, Jr. 2004. Microsatellite DNA variation in tule elk. Journal of Wildlife Management 68:109-119.

Wisz, M.S., R.J. Hijmans, J. Li, A.T. Peterson, C.H. Graham, A. Guisan, and the NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. Diversity and Distributions 14(5): 763-773.

Yackulic, C.B., R. Chandler, E.F. Zipkin, J.A. Royle, J.D. Nichols, E.H.C. Grant, and S. Veran. 2013. Presence-only modelling using Maxent: when can we trust the inferences? Methods in Ecology and Evolution 4:236-243.

York, P., P. Evangelista, S. Kumar, J. Graham, C. Flather, and T. Stohlgren. 2011. A habitat overlap analysis derived from Maxent for Tamarisk and the South-western Willow Flycatcher. Frontiers of Earth Science 5(2): 120-129.

Section I Figures

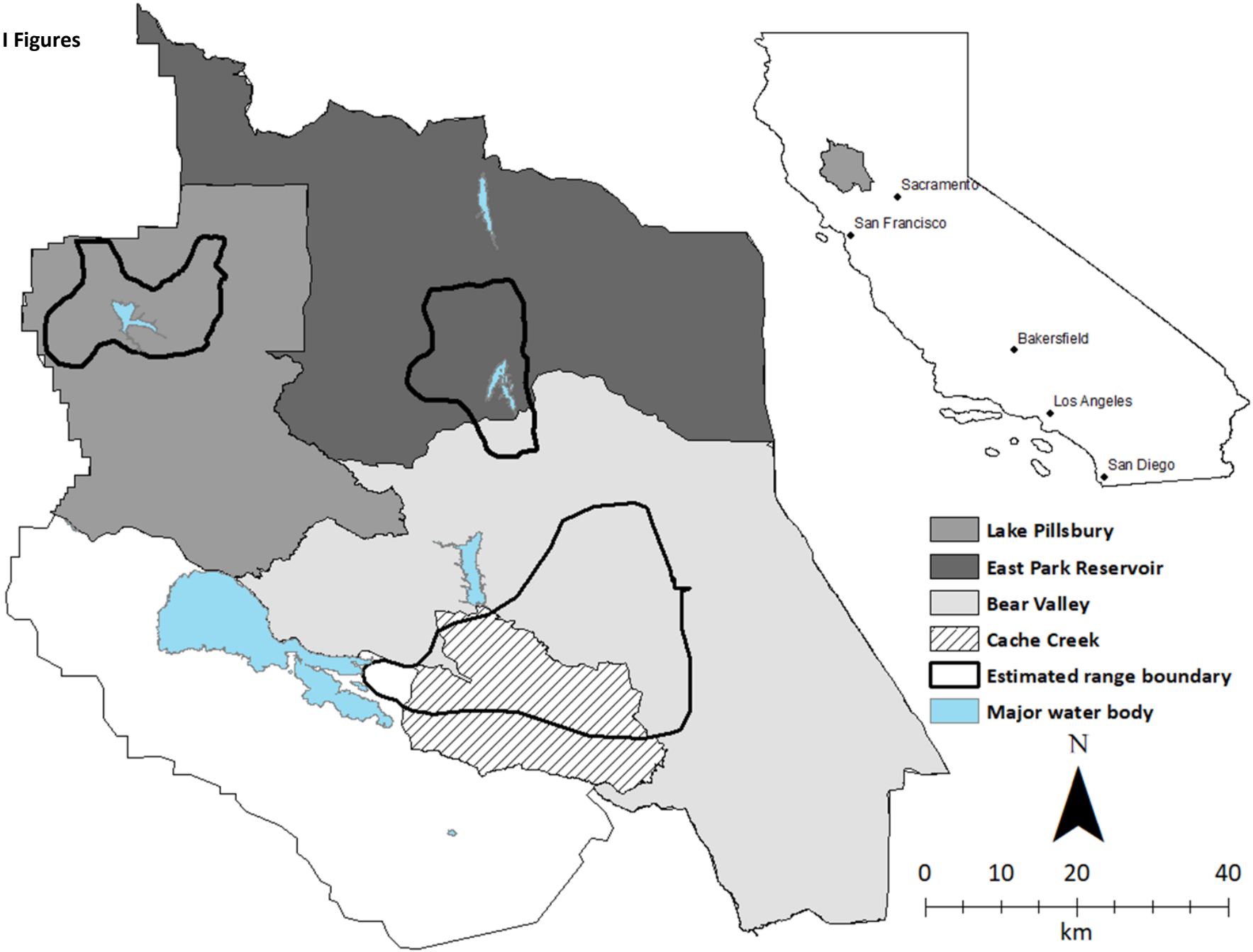
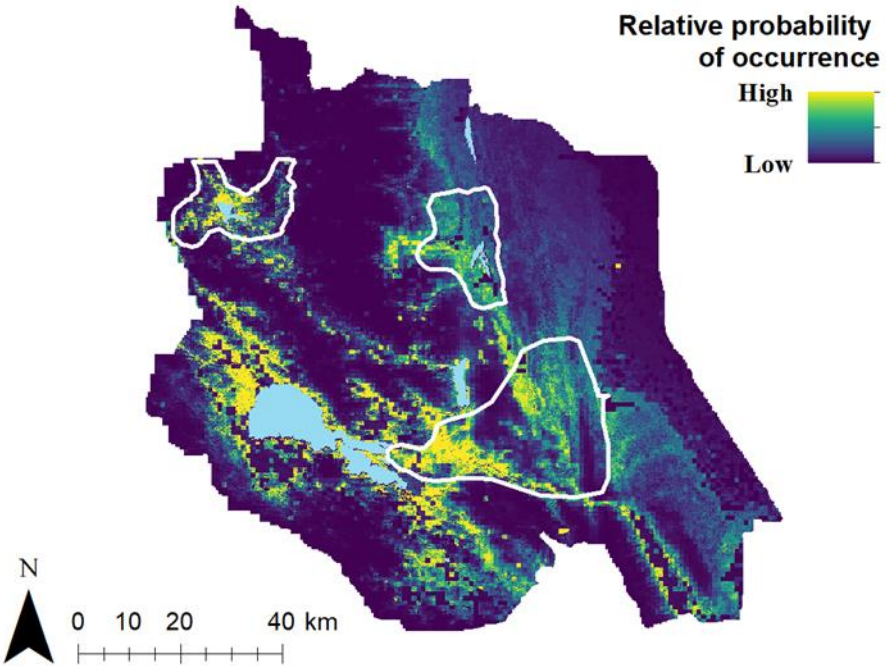
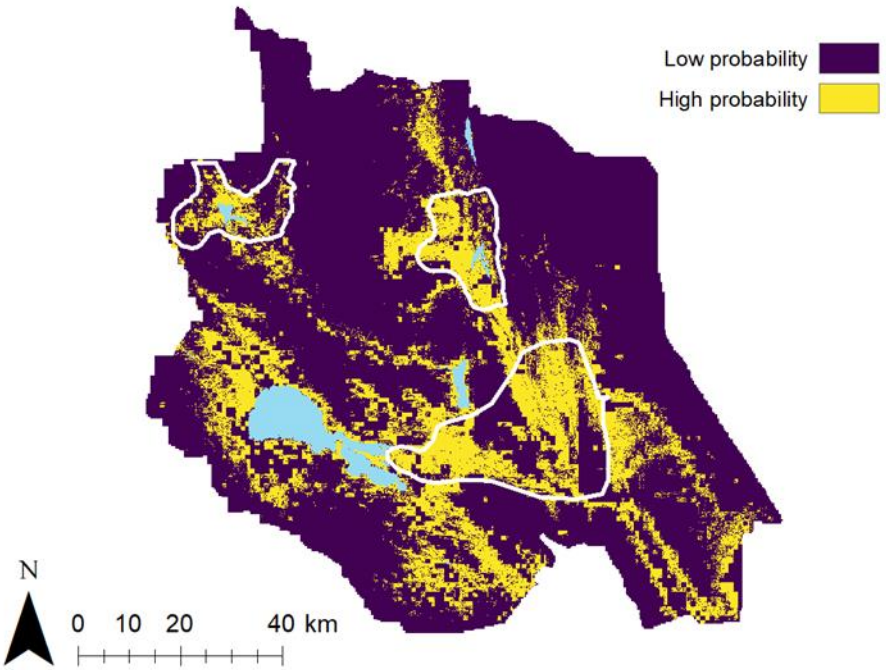


Figure 1. Study area in the California Coast and Interior Coast mountain ranges encompassing four management units (MUs) from CDFW (2018) plus the entirety of Lake County. Estimated range boundaries (clockwise from left) for the Lake Pillsbury, East Park Reservoir, and Cache Creek populations are outlined in black and overlaid atop the MUs. Major water bodies are indicated by blue polygons.





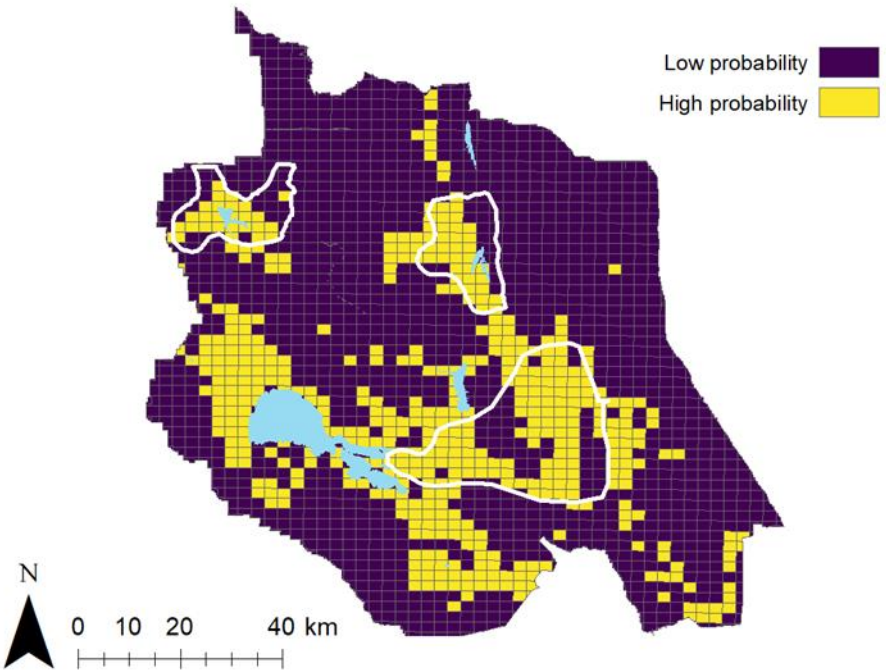


Figure 2. Preliminary model predictive surfaces for tule elk in Colusa and Lake Counties, CA. Predictive surfaces are shown for the (A) relative probability of occurrence (RPO) (B) the raw binary surface (classification threshold value = 0.15), and (C) its corresponding dichotomized 2 km^2 gridded surface. Putative range boundaries are outlined in white.

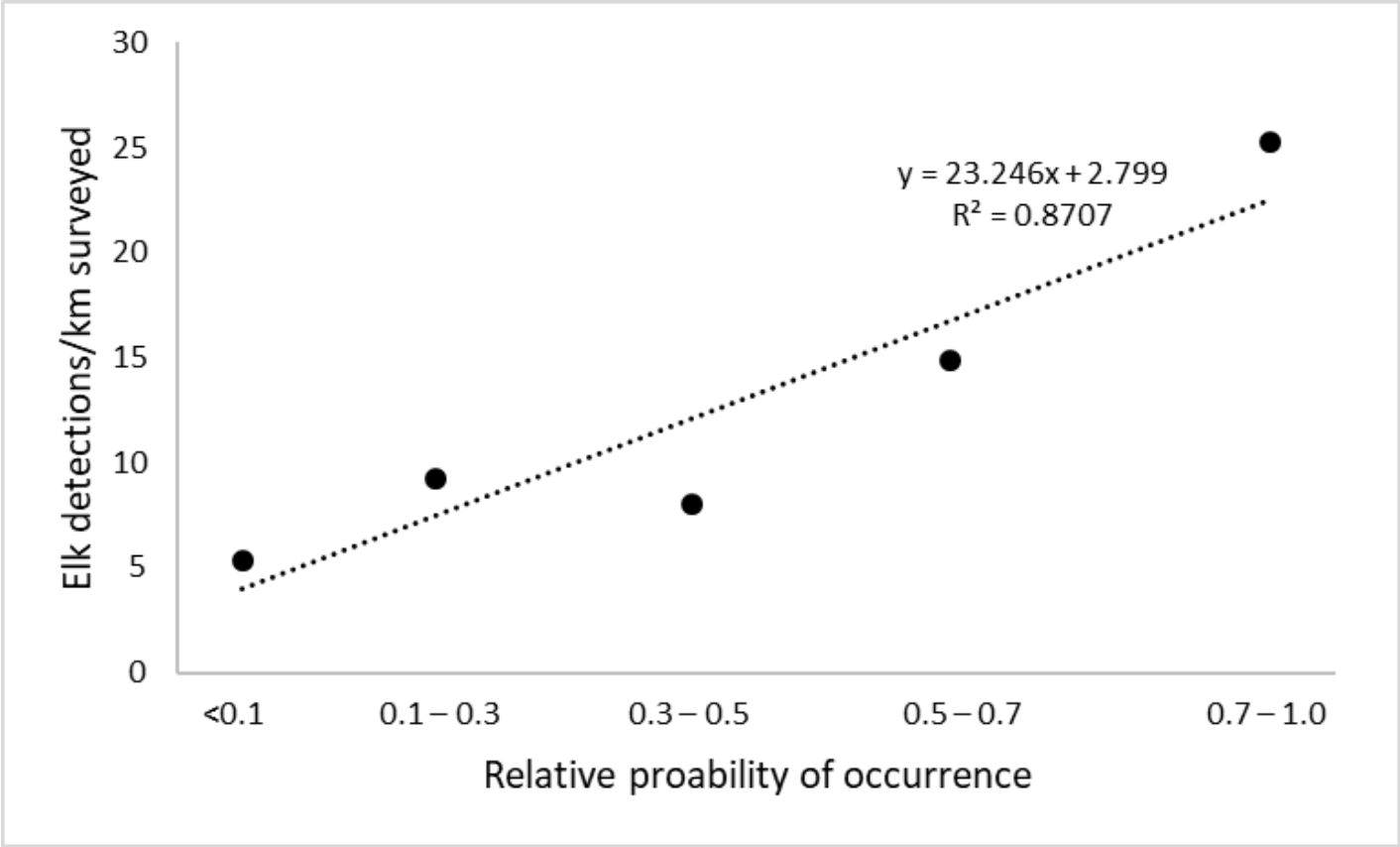


Figure 3. Linear least-squares regression of elk detections per kilometer surveyed as a function of 5 relative probability of occurrence (RPO) classes according to the preliminary Maxent model. Elk detections (n = 3,629) were gathered during field surveys from Jun–Aug 2017–19 in Colusa and Lake Counties, CA. The regression equation was $\hat{y}$ = 23.25x + 2.8, with mean square error = 10.8 detections per km.

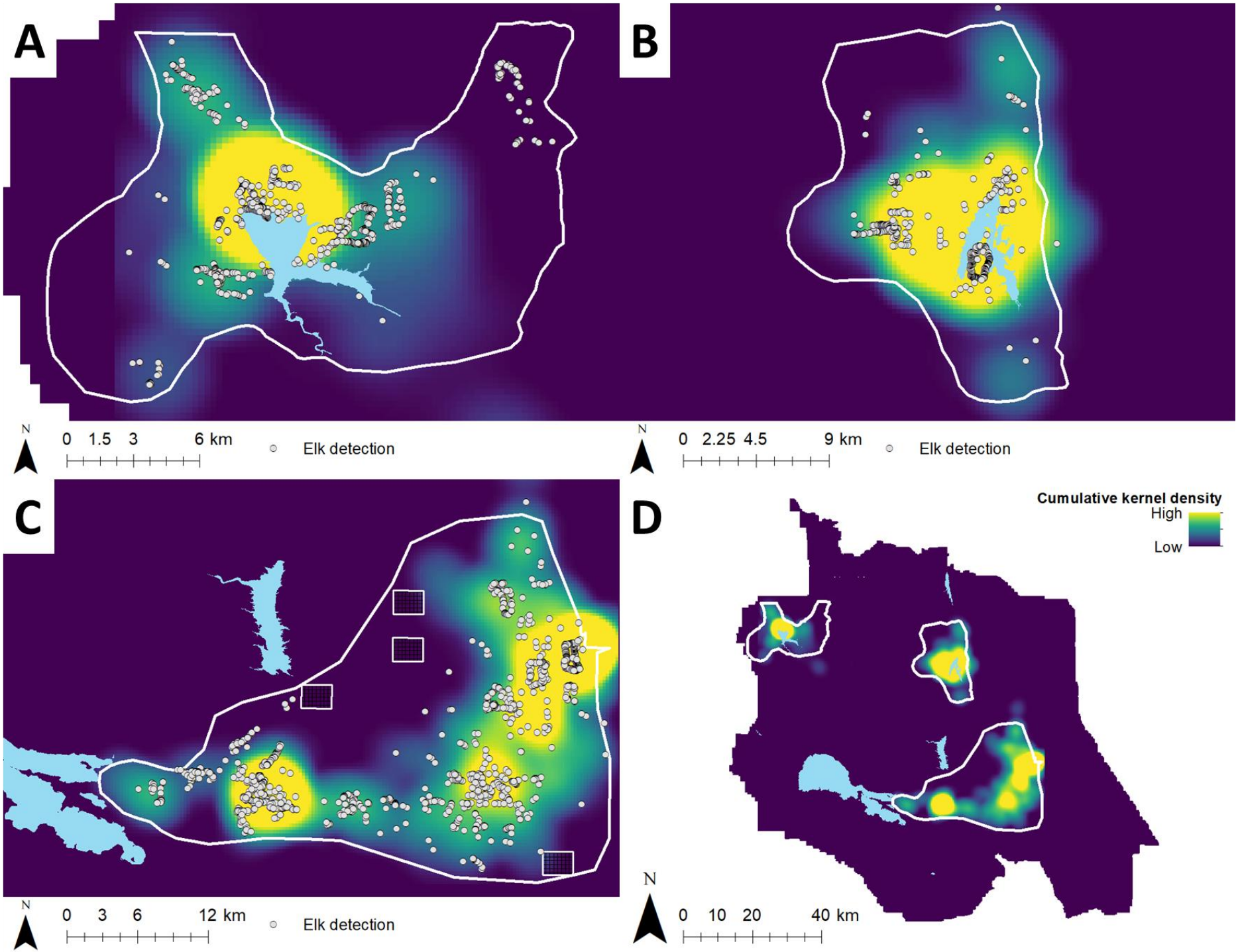


Figure 4. Cumulative kernel densities of GPS collar points from 78 tule elk (39 male, 39 female) accumulated during January 2017–October 2019 and elk detections (grey circles) gathered during field surveys from Jun–Aug 2017–19 in Colusa and Lake Counties, CA. Surfaces are shown for elk at (A) Lake Pillsbury, (B) East Park Reservoir, (C) Cache Creek, and within (D) the entire study area. Putative population range boundaries are outlined in white. The cells where elk were not detected at Cache Creek (n = 4) are outlined in white with black cross-hatch (C).

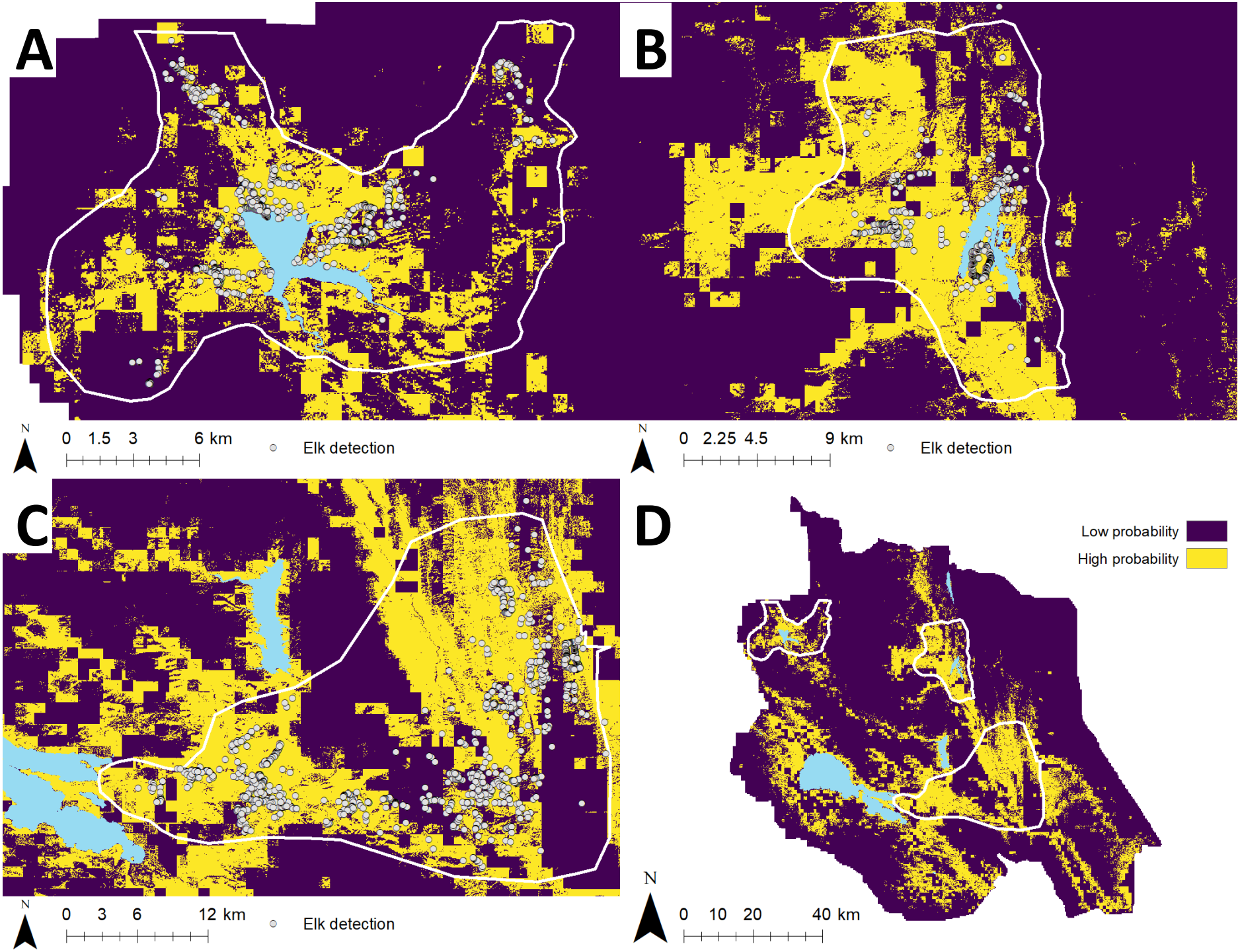
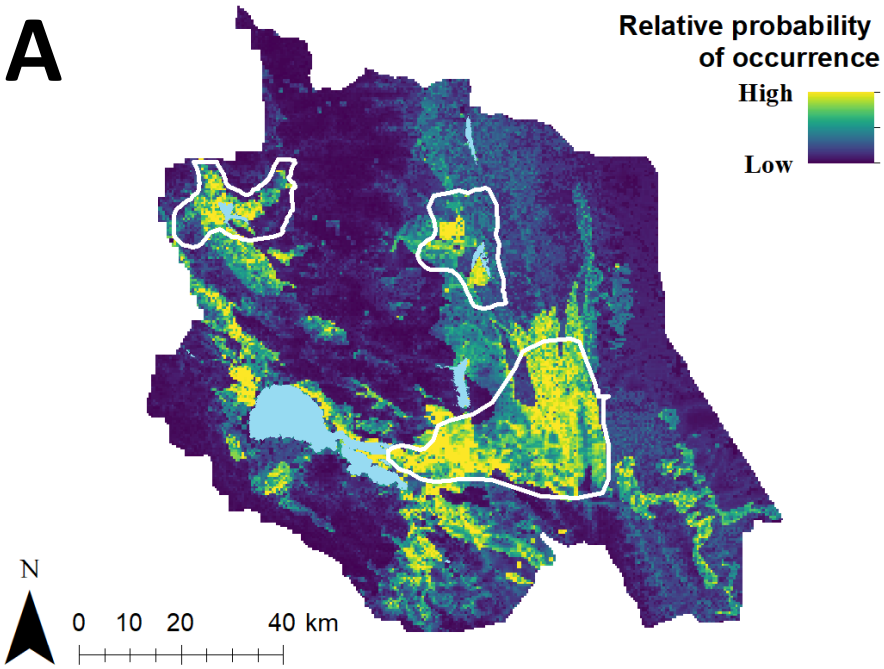
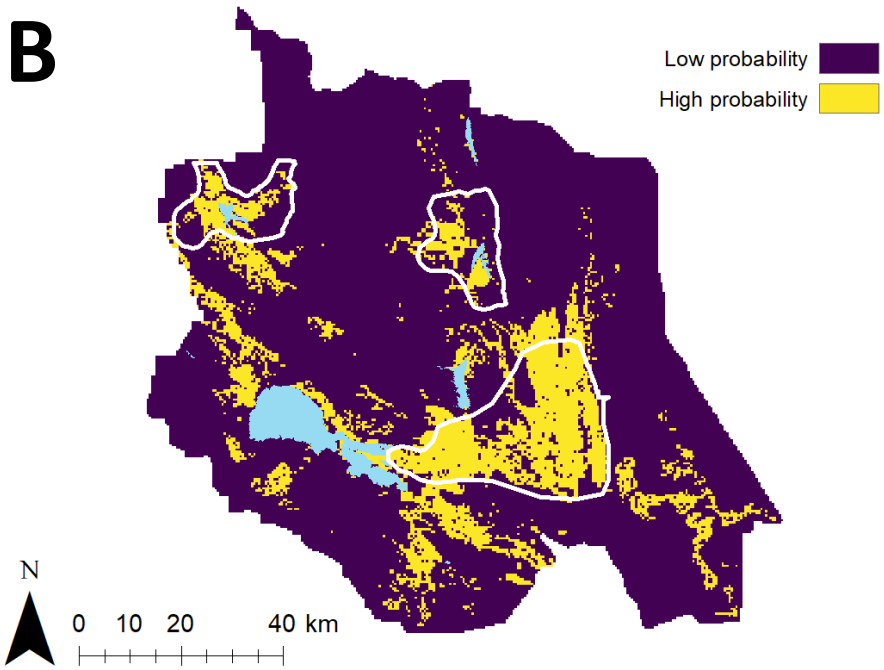


Figure 5. Elk detections recorded during the three field seasons overlaid upon the preliminary model raw binary surface. Surfaces are shown for (A) Lake Pillsbury, (B) East Park Reservoir, (C) Cache Creek within (D) the entire study area. Gold areas indicate areas of high probability of elk occurrence; purple areas indicate areas of low probability of elk occurrence. Putative range boundaries are outlined in white. Elk detections occurred in ≤500 m from areas of predicted presence.





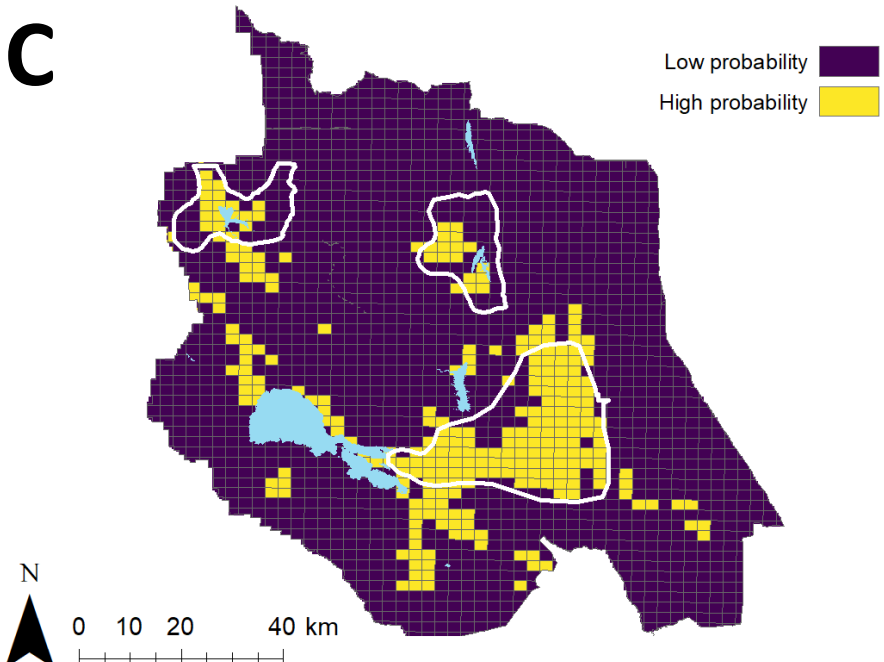


Figure 6. Refined model predictive surfaces for tule elk in Colusa and Lake Counties, CA. Predictive surfaces are shown for the (A) relative probability of occurrence (RPO), (B) the raw binary surface (classification threshold value = 0.35), and (C) its corresponding dichotomized 2-km^2 sampling grid surface. Putative range boundaries are outlined in white.

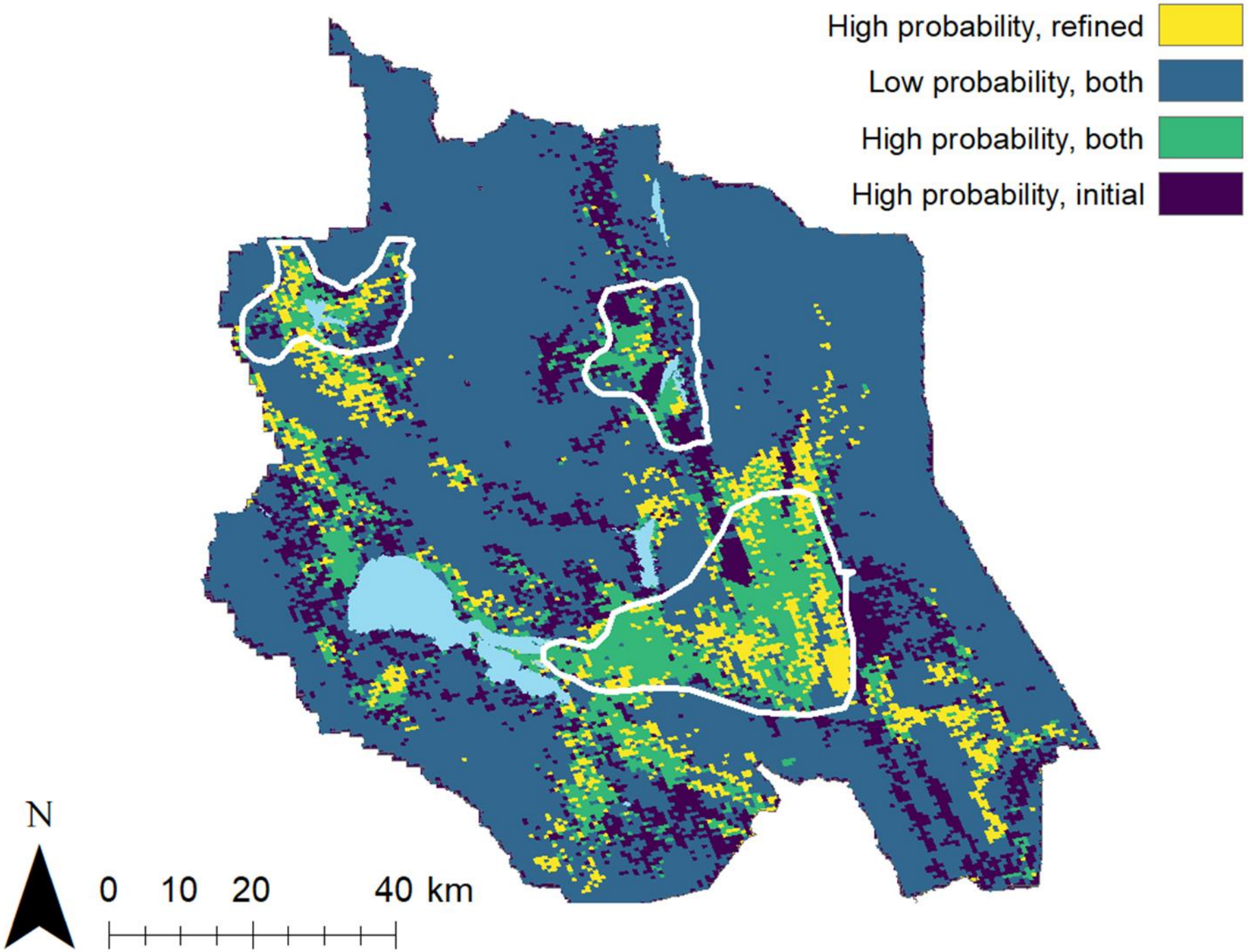


Figure 7. Changes in binary distribution predictions between the preliminary and refined models. Surface depicts areas of: high probability in the refined model only (gold), low probability in both models (blue), high probability in both models (green), and high probability in the preliminary model only (purple). The refined model reduced the high probability area by ~540 km^2 (6%).

Section I Tables

Table 1. Presence type and total count for tule elk detection data for the preliminary Maxent model input.

Presence Type	Count
Antler shed	1
Bedding area	6
Camera detection	3
Carcass remains	3
Direct observation	187
Hair	1
Hunter harvest	72
Pellet groups	825
Rub	2
Scrape	1
Tracks	106
Total	**1207**

Table 2. Presence type and total count for tule elk detection data for the refined Maxent model input.

Presence Type	Count
Antler shed	11
Bedding area	242
Camera detection	31
Carcass remains	13
Direct observation	277
Failed GPS collar	3
Hunter harvest	34
Pellet groups	2605
Rub	2
Scrape	7
Tracks	404
Total	**3629**

Table 3. Average area under the receiver operating characteristic curve (AUC) values for different model parameterizations estimated from cross-validation (preliminary model, gray fill) and geographically structured k-fold cross-validation (refined model) of Maxent predictions for locations of tule elk in Colusa and Lake Counties, CA. Results are shown for the best-fit model from each dataset (Occurrence Dataset), with each model's corresponding partition type (Partition), number of occurrence points used as model input (No. Occ. Points), interface used to produce the model (Interface), feature class or feature class combination (FC), and regularization multiplier (RM).

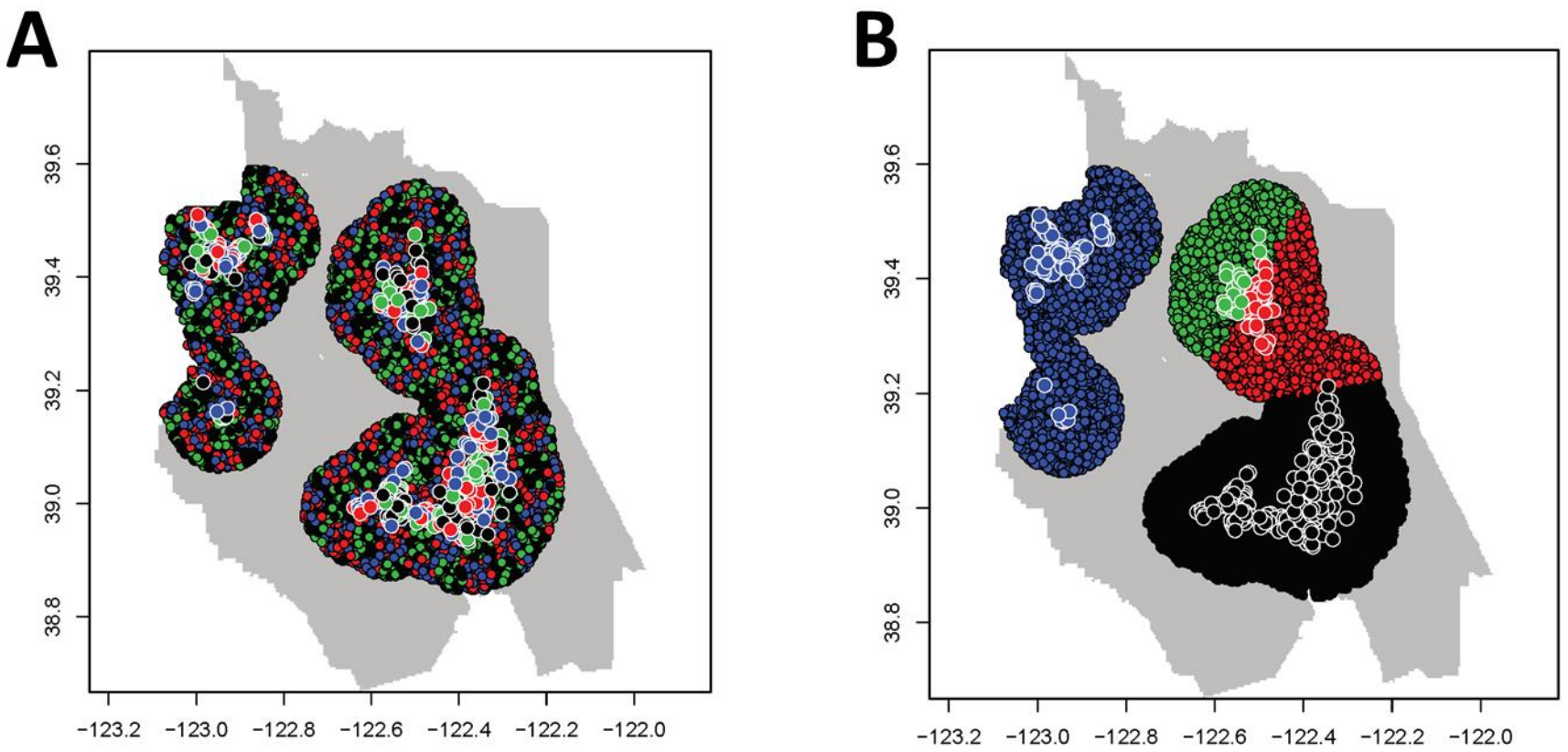
Occurrence Dataset	Partition	AUC_{TRAIN}	AUC_{TEST}	AUC_{DIFF}	No. Occ. Points	Interface	FC*	RM
Prelim	Cross-validate	0.889	0.885	0.004	1,207	Maxent GUI	LQ	1
Full	CheckerBoard2	0.885	0.845	0.040	3,629	ENMeval	L	2.5
100m	CheckerBoard2	0.858	0.825	0.033	924	ENMeval	LQP	0.5
500m	CheckerBoard2	0.832	0.795	0.037	323	ENMeval	LQ	0.5
1km	CheckerBoard2	0.814	0.764	0.050	208	ENMeval	LQ	1.5
Full	User	0.885	0.615	0.270	3,629	ENMeval	LQP	0.5
100m	User	0.858	0.561	0.297	924	ENMeval	LQ	0.5
500m	User	0.833	0.500	0.333	323	ENMeval	LQP	0.5
1km	User	0.816	0.574	0.242	208	ENMeval	LQ	1

*Feature classes considered singly and in combination include linear (L), quadratic (Q), and product (P)

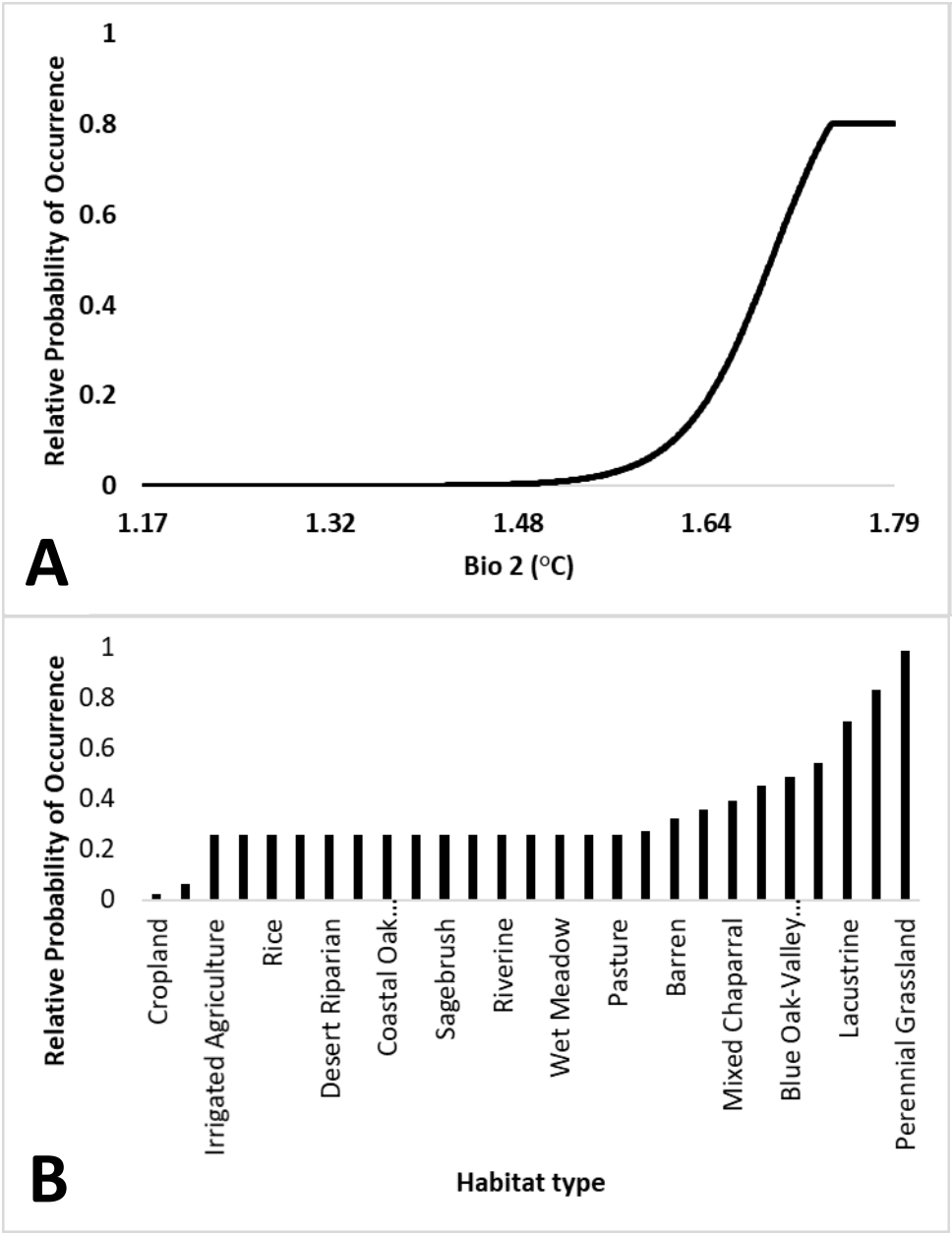
Table 4. Measurements of model performance for Maxent models using full and thinned independent GPS collar data from 78 elk (39 male, 39 female). Results are shown for the best-fit model from each dataset (Occurrence Datatset), with each model's corresponding partition type (Partition), and number of GPS points used as model input (No. GPS Points). Performance metrics include average area under the receiver operating characteristic curve (AUC) for the threshold independent logistic model (AUC_{FULL}), the binary classification threshold (Threshold), the AUC for the threshold-dependent binary (un-gridded) model (AUC_{BINARY}), true skill statistic (TSS) value, and percent correctly classified (PCC).

Occurrence Dataset	Partition	No. GPS Points	AUC_{FULL}	Threshold	AUC_{BINARY}	TSS	PCC
Prelim	Cross-validate	47,445	0.789	0.1506	0.722	0.445	73%
Full	CheckerBoard2	47,445	0.759	0.3008	0.704	0.407	64%
100m	CheckerBoard2	7,580	0.786	0.3229	0.721	0.443	73%
500m	CheckerBoard2	924	0.780	0.3577	0.736	0.472	78%
1km	CheckerBoard2	323	0.751	0.3292	0.685	0.371	65%
Full	User	47,445	0.782	0.2689	0.705	0.410	65%
100m	User	7,580	0.784	0.4178	0.716	0.433	73%
500m	User	924	0.780	0.4171	0.724	0.447	75%
1km	User	323	0.757	0.4099	0.689	0.379	74%

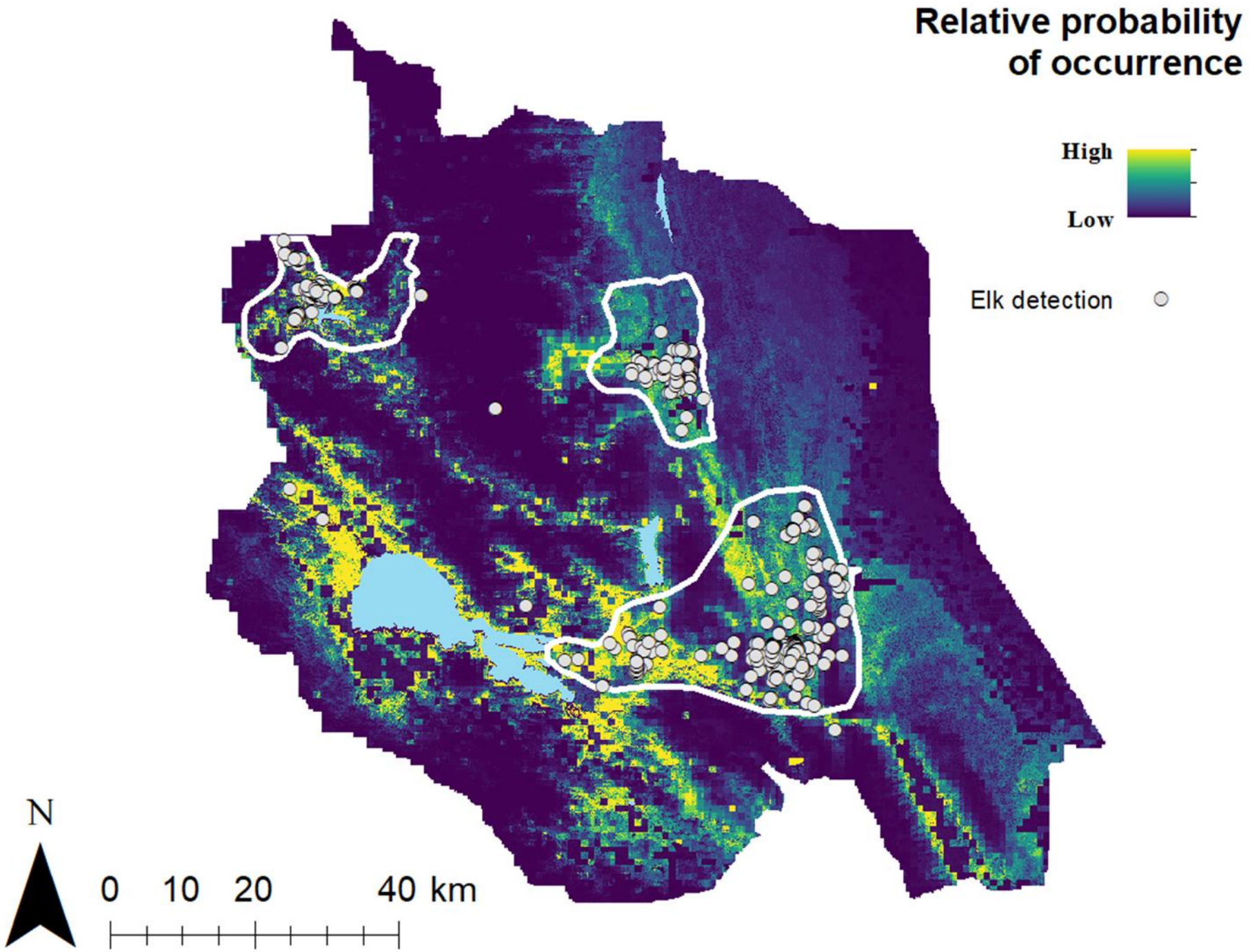
Section I Supplementary Figures



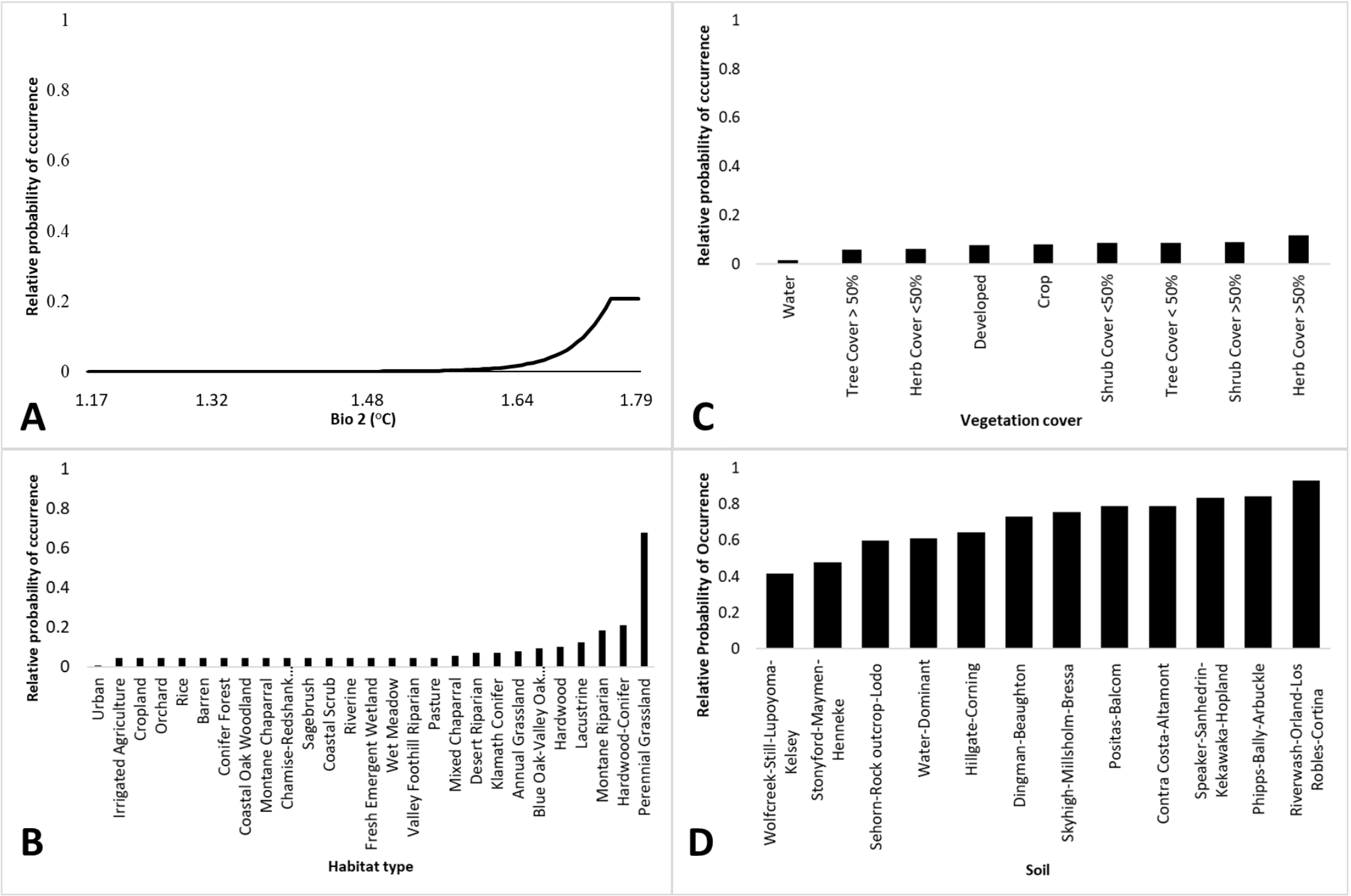
Supplementary Fig. 1. *Checkerboard2* (A) and *user* (B) partition methods to divide data for training and testing in the *R* program *ENMeval* (Muscarella et al. 2014).



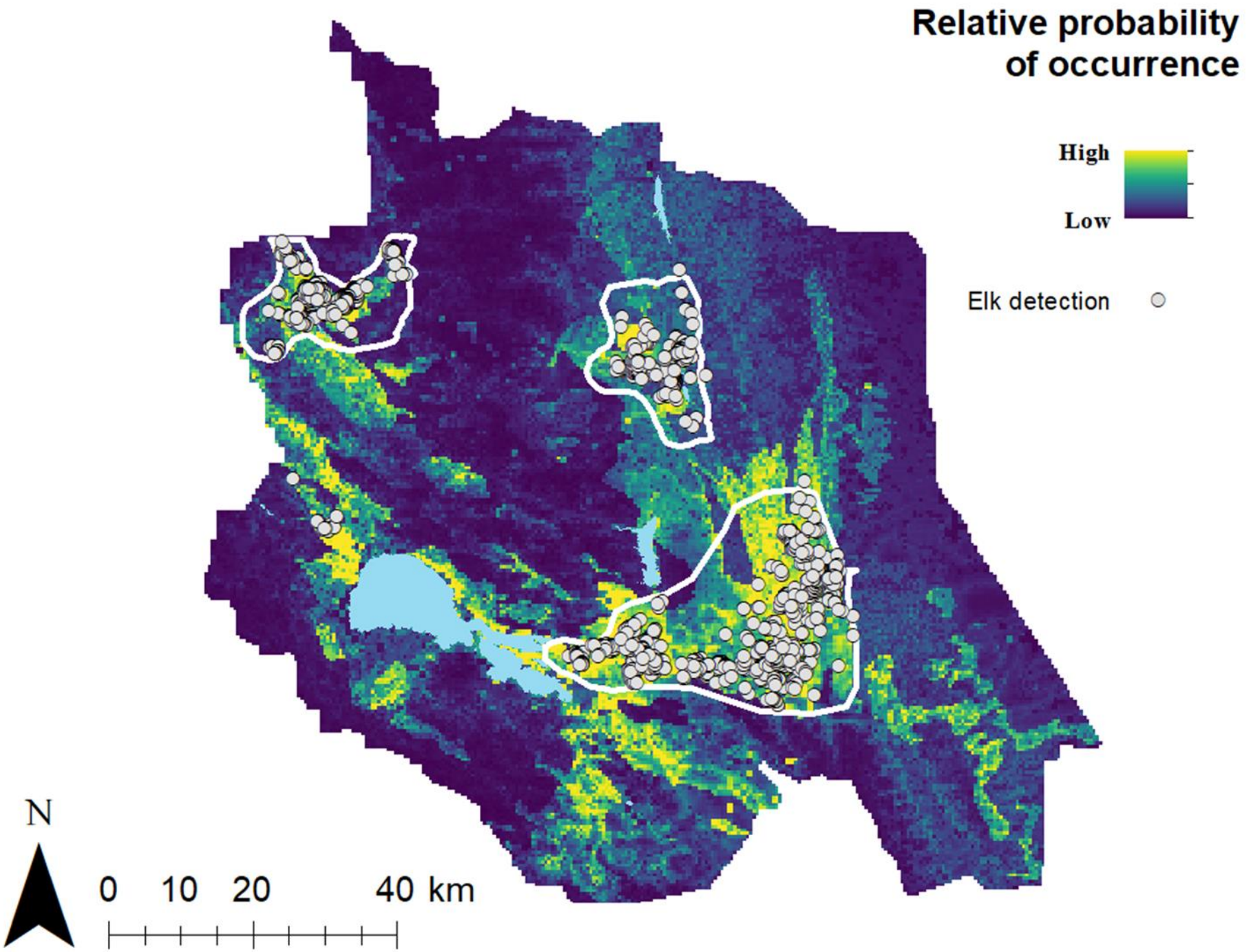
Supplementary Fig. 2. Preliminary model response curves for bio2 (mean diurnal temperature range) (A), habitat type (B), and vegetation cover (C). Each graph shows how the predicted relative probability of elk occurrence changes as each environmental variable is varied while keeping all other variables at their average sample value.



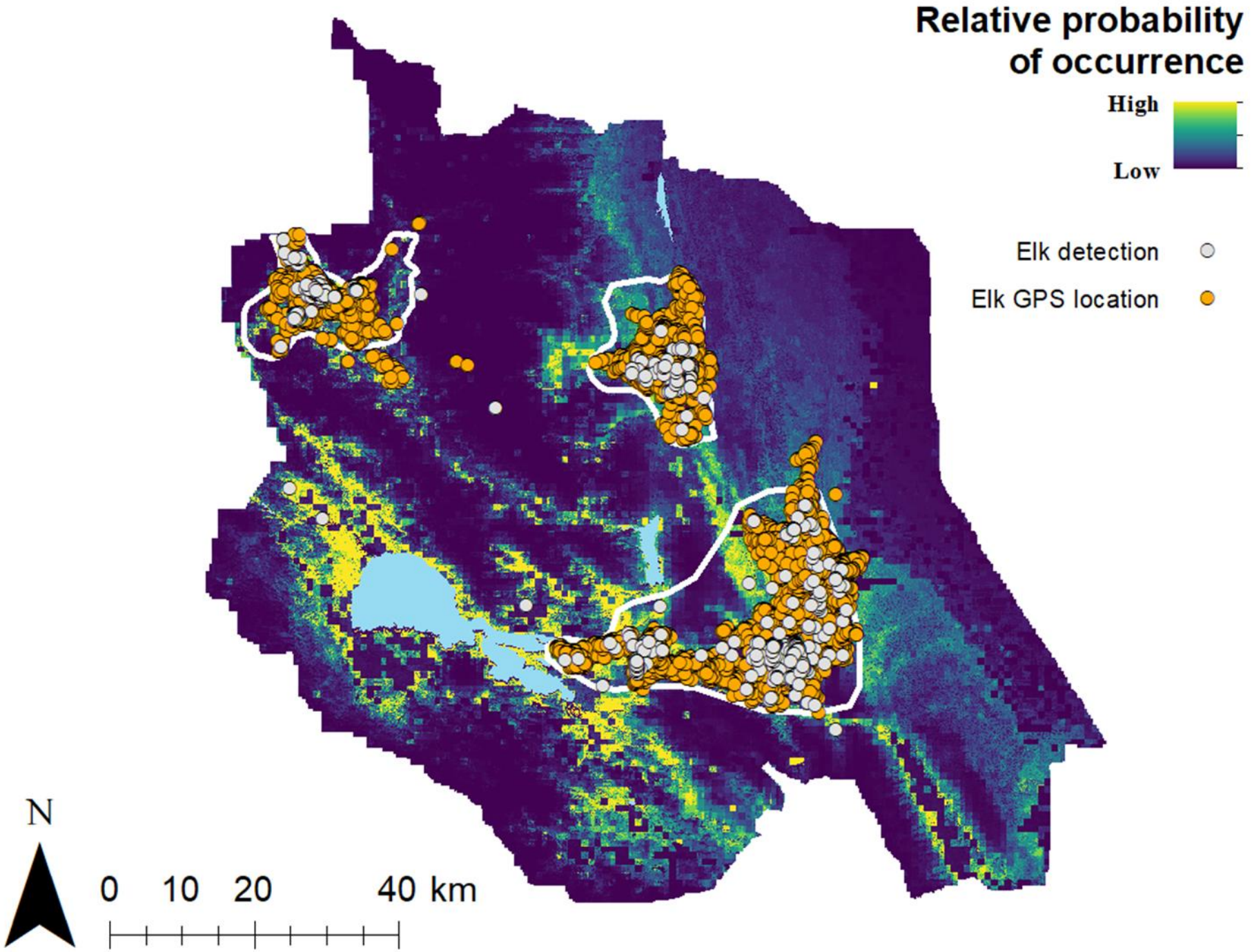
Supplementary Figure 3. Preliminary model logistic predictive surface showing the relative probability of occurrence of tule elk in Colusa and Lake Counties, CA. Elk detections used to produce the model are shown in gray (n = 1,207). Warmer surface colors indicate areas predicted to have more suitable environmental conditions. Putative range boundaries are outlined in white.



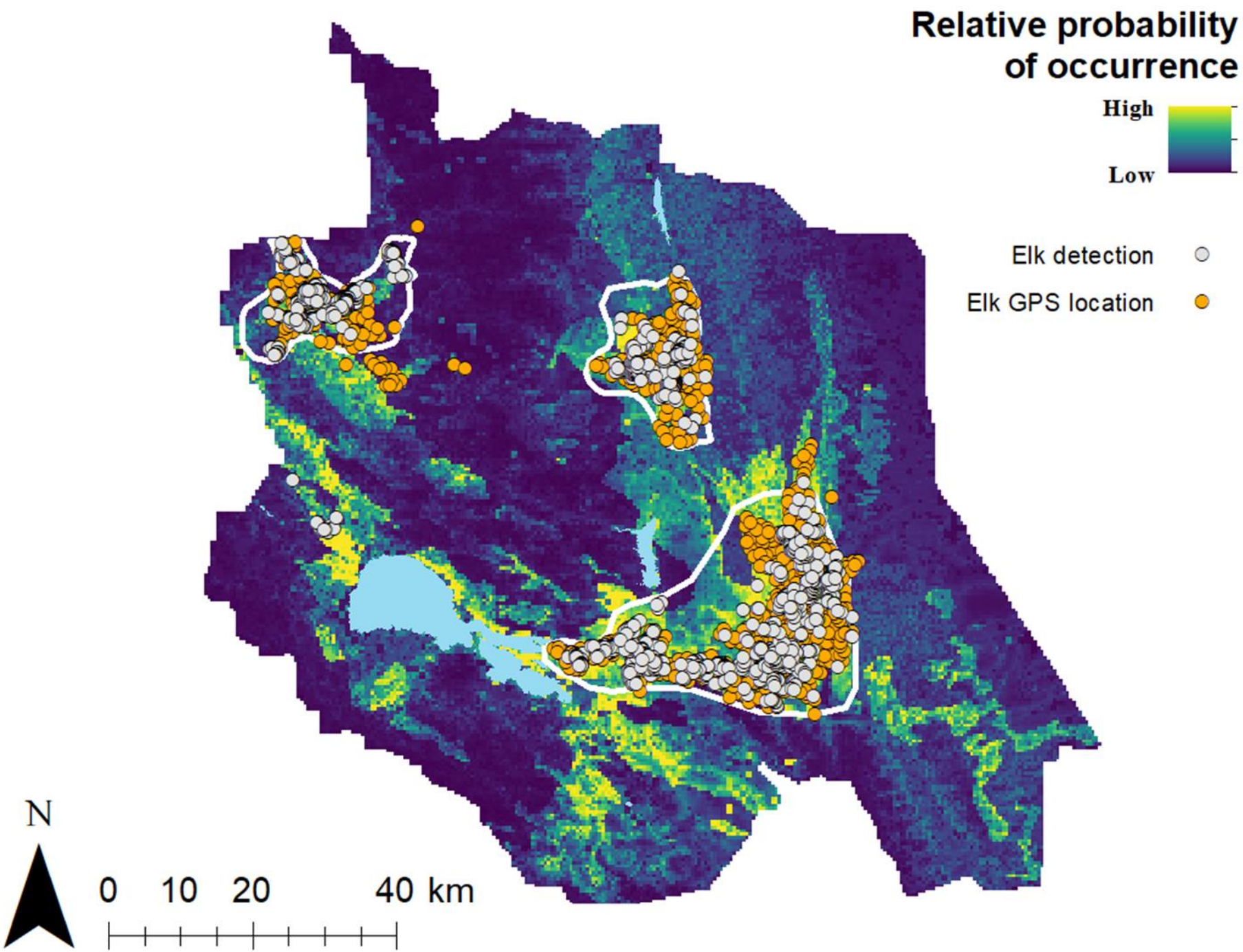
Supplementary Figure 4. Refined model response curves for bio2 (mean diurnal temperature range) (A), habitat type (B), vegetation cover (C) and soil (only soil types with relative probability of occurrence values > 0.40 are shown) (D). Each graph shows how the predicted relative probability of elk occurrence changes as each environmental variable is varied while keeping all other variables at their average sample value.



Supplementary Figure 5. Refined model logistic predictive surface showing the relative probability of occurrence of tule elk in Colusa and Lake Counties, CA. Spatially thinned (500 m) elk detections used to produce the model (n = 323) are shown in gray. Warmer surface colors indicate areas predicted to have more suitable environmental conditions. Putative range boundaries are outlined in white.



Supplementary Fig. 6. Preliminary model logistic predictive surface showing the relative probability of occurrence of tule elk in Colusa and Lake Counties, CA. Elk detections used to produce the model (n = 1,207) are shown in gray; independent GPS collar location data (n = 47,445) from 78 elk (39 male, 39 female) used to evaluate model performance are shown in orange. Warmer surface colors indicate areas predicted to have more suitable environmental conditions. Putative range boundaries are outlined in white.



Supplementary Fig. 7. Refined model logistic predictive surface showing the relative probability of occurrence of tule elk in Colusa and Lake Counties, CA. Spatially thinned (500 m) elk detections used to produce the model (n = 323) are shown in gray; spatially thinned (500 m) independent GPS collar location data (n = 924) from 78 elk (39 male, 39 female) used to evaluate model performance are shown in orange. Warmer surface colors indicate areas predicted to have more suitable environmental conditions. Putative range boundaries are outlined in

Section I Supplementary Tables

Supplementary Table 1. Covariates used for Maxent modelling of suitable elk habitat in Colusa and Lake Counties, California, USA.

Variable	Variable Subcategories	Notes	Source
Abiotic			
Soil			WSS 2013
	Cibo-Ayar-Altamont		
	Contra Costa-Altamont		
	Dingman-Beaughton		
	Glenview-Bottlerock-Arrowhead		
	Goldridge		
	Hillgate-Corning		
	Maxwell-Leesville		
	Maymen-Etsel		
	Millsholm-Los Osos-Dibble-Balcom		
	Millsholm-Maymen-Los Gatos-Dibble		
	Parrish-Los Gatos-Hulls-Goulding		
	Parrish-Maymen-Los Gatos-Etsel		
	Phipps-Bally-Arbuckle		
	Positas-Balcom		
	Riverwash-Orland-Los Robles-Cortina		
	Sehorn-Millsholm-Lodo		
	Sehorn-Rock outcrop-Lodo		
	Sheetiron-Millich-Goulding		
	Sheetiron-Rubble land-Neuns		
	Skyhigh-Millsholm-Bressa		
	Sobrante-Hambright		
	Sodabay-Konocti-Benridge		
	Speaker-Sanhedrin-Kekawaka-Hopland		
	Stonyford-Maymen-Henneke		
	Tehama-Hillgate-Arbuckle		
	Wappo-Manzanita-Forbesville		
	Water-Dominant		
	Wolfcreek-Still-Lupoyoma-Kelsey		
	Yollabolly-Rock outcrop		
	Yolo-Sycamore-Brentwood-Artois		

	Yorkville-Yorktree-Witherell-Squawrock-Shortyork Zamora-Willows-Marvin-Capay		
Distance to water	Continuous, meters	Manually derived from USA Surface Water layer	ESRI 2013
Bioclimatic			
Worldclim			Hijmans et al. 2005
	Bio1	Annual mean temperature	
	Bio2	Mean diurnal temperature range	
	Bio3	Isothermality	
	Bio4	Temperature of seasonality	
	Bio5	Maximum temperature of warmest month	
	Bio6	Maximum temperature of coldest month	
	Bio7	Temperature of annual range	
	Bio8	Mean temperature of wettest quarter	
	Bio9	Mean temperature of driest quarter	
	Bio10	Mean temperature of warmest quarter	
	Bio11	Mean temperature of coldest quarter	
	Bio12	Annual precipitation	
	Bio13	Precipitation of wettest month	
	Bio14	Precipitation of driest month	
	Bio15	Precipitation seasonality	
	Bio16	Precipitation of wettest quarter	
	Bio17	Precipitation of driest quarter	
	Bio18	Precipitation of warmest quarter	
	Bio19	Precipitation of coldest quarter	
Ecological			
Habitat type			CalFIRE 2015
	Annual Grassland		
	Barren		
	Blue Oak-Valley Oak Dominant		
	Chamise-Redshank Chaparral		
	Coastal Oak Woodland		
	Coastal Scrub		
	Conifer Forest		

	Cropland	
	Desert Riparian	
	Fresh Emergent Wetland	
	Hardwood	
	Hardwood-Conifer	
	Irrigated Agriculture	
	Klamath Conifer	
	Lacustrine	
	Mixed Chaparral	
	Montane Chaparral	
	Montane Riparian	
	Orchard	
	Pasture	
	Perennial Grassland	
	Rice	
	Riverine	
	Sagebrush	
	Urban	
	Valley Foothill Riparian	
	Wet Meadow	
Topographic		
Elevation	Continuous, meters	LANDFIRE 2013a
Slope	Continuous, percent rise	LANDFIRE 2013a
Aspect	Continuous, degrees	LANDFIRE 2013a
Vegetation		
Vegetation cover		LANDFIRE 2013b
	Crop	
	Developed	
	Herb Cover <50%	
	Herb Cover >50%	
	Shrub Cover <50%	
	Shrub Cover >50%	
	Tree Cover < 50%	
	Tree Cover > 50%	
	Water	
Vegetation height		LANDFIRE 2013b
	Agriculture - General	

Barren
Cultivated Crops
Developed - General
Developed - High Intensity
Developed - Low Intensity
Developed - Medium Intensity
Developed - Open Space
Developed-Herbaceous Wetland Vegetation
Developed-Roads
Developed-Upland Deciduous Forest
Developed-Upland Evergreen Forest
Developed-Upland Herbaceous
Developed-Upland Mixed Forest
Developed-Upland Shrubland
Developed-Woody Wetland Vegetation
Fallow
Forest Height > 50 meters
Forest Height 0 to 5 meters
Forest Height 10 to 25 meters
Forest Height 25 to 50 meters
Forest Height 5 to 10 meters
Herb Height > 1.0 meter
Herb Height >= 0.5m
Herb Height >0 and < 0.5m
Herb Height 0 to 0.5 meters
Herb Height 0.5 to 1.0 meters
Herbaceous Semi-dry
Herbaceous Semi-wet
Herbaceous Wetlands
NASS-Aquaculture
NASS-Bush fruit and berries
NASS-Close Grown Crop
NASS-Fallow/Idle Cropland
NASS-Orchard
NASS-Pasture and Hayland
NASS-Row Crop
NASS-Row Crop-Close Grown Crop

NASS-Vineyard
NASS-Wheat
Open Water
Pasture/Hay
Quarries-Strip Mines-Gravel Pits
Recently Disturbed Forest
Shrub Height > 3.0 meters
Shrub Height >= 0.5 and < 1.5m
Shrub Height >= 1.5m
Shrub Height >0 and < 0.5m
Shrub Height 0 to 0.5 meters
Shrub Height 0.5 to 1.0 meter
Shrub Height 1.0 to 3.0 meters
Small Grains
Snow/Ice
Sparse Vegetation Height
Sparse Vegetation Height
Tree Height > 0 and < 10m
Tree Height >= 10m
Urban-Recreational Grasses

Supplementary Table 2. Percent contribution and permutation importance of three variables to the preliminary model

Variable	Percent Contribution	Permutation Importance
Bio2	53.8	65.1
Habitat type	25.4	15.2
Vegetation cover	20.8	19.6

Supplementary Table 3. Field survey detections of tule elk classified by the preliminary model predictions in all populations (all), Cache Creek (CC), Lake Pillsbury (LPB), East Park Reservoir (EPR), and Bachelor Valley (BV) in Colusa and Lake Counties, CA, from 2017-2019. Occurrence locations are classified by low probability (LP) cells, high probability (HP) cells, total detections (Total), percent low probability (%LP), and percent high probability (%HP).

Location	LP	HP	Total	%LP	%HP
All	581	3048	3629	16%	84%
Cache Creek	226	1853	2079	11%	89%
Lake Pillsbury	265	782	1047	25%	75%
East Park Reservoir	90	394	484	19%	81%
Bachelor Valley	0	19	19	0%	100%

Supplementary Table 4. Percent contribution and permutation importance of six variables to the refined model

Variable	Percent Contribution	Permutation Importance
Soil	64.8	61.1
Habitat type	13.0	9.1
Vegetation cover	12.0	10.0
Bio2	8.1	17.8
Distance to water	2.1	2.0
Slope	0.0	0.0

II. Noninvasive spatial capture-recapture estimation of tule elk (*Cervus canadensis nannodes*) abundance in Northern California using fecal DNA

Introduction

Effective conservation and management of wildlife populations requires periodic estimation of density and abundance (Pollock 1991; Lancia et al. 2005). Aerial and ground-based minimum count data are often used to infer population sizes of big game species (Williams et al. 2002; California Department of Fish and Wildlife [CDFW] 2018). These traditional methods, however, typically yield underestimates of density and abundance with unknown precision and potentially variable relationships to true abundance; they can be affected by visually obstructive habitat and terrain, unknown detection probabilities, and lack of bounded statistical estimates (Caughley 1974; McCullough et al. 1994; Bleich et al. 2001; Williams et al. 2004; Stephens et al. 2015; Schoenecker and Lubow 2016).

Alternatively, statistical methods, such as line-transect, distance-sampling, or capture-recapture (CR) methods enable point estimates in addition to confidence intervals that quantify the uncertainty associated with them (Pollock et al. 1990; Williams et al. 2002; Lancia et al. 2005). In recent years, noninvasive fecal DNA (fDNA)-based spatial capture-recapture (SCR) techniques have proven highly versatile for local and regional efforts to estimate and monitor abundance of wildlife populations for species that conform to SCR assumptions such as deer (*Odocoileus* spp.) in a variety of habitats (Brinkman et al. 2011; Goode et al. 2014; Lounsberry et al. 2015; Brazeal et al. 2017; Furnas et al. 2018, 2020).

However, these methods have not been widely applied to more gregarious species such as elk (*Cervus canadensis*) that use the landscape more heterogeneously (Mackie 1970) and exhibit grouping behavior that changes seasonally (Murie 1951; Raedeke et al. 2002). Although SCR methods can account for heterogeneous landscape use, randomly sampling landscapes for which only a small fraction is occupied by target species may be unpractical and inefficient. Second, space use and grouping behavior of elk, which differ between the sexes, can affect assumptions of independence. Females especially tend to aggregate (Geist 2002), potentially introducing biases that can lead to overestimated precision and, in extreme cases, abundance (Bischof et al. 2020).

We applied fDNA SCR methods to elk in northern California during summer months, after calving and before the rut, when elk were expected to exhibit relatively stable space use (McCullough 1969). To enhance the efficiency of fDNA samples, we used a species distribution model (SDM) to stratify the landscape into high- and low-probability strata, which allowed us to sample each stratum randomly but with effort disproportionately expended in the high-probability stratum (Batter et al. Ch. 1). This procedure resulted in a ~35% increase in efficiency relative to a simple random design (Batter et al. Ch. 1). Pellet surveys were conducted concurrently with but independently of GPS telemetry and aerial counts in 3 tule elk (*C. c. nannodes*) populations. We compared fDNA SCR estimates both for males (expected to conform better to assumptions of independence) and females (expected to violate assumptions) modeled separately and for combined-sex models and used telemetry data to explore the magnitude of spatial clustering of each sex.

Methods

Study Area

The study area encompassed ~9,000 km^2 of the California Coast and Interior Coast mountain ranges, including four management units (MUs): the Cache Creek, Bear Valley, Lake Pillsbury (LPB), and East Park Reservoir (EPR) MUs (CDFW 2015; CDFW 2018) (Fig. 1). Based on historical records, range boundaries were defined for three populations corresponding to LPB, EPR, and the combined Cache Creek and Bear Valley MUs, hereafter referred to as Cache Creek (CC) (Batter et al. Ch. 1). The region generally experiences hot, dry summers and mild, rainy winters, described as a Mediterranean climate (Kauffman 2003). Year-round temperatures range from 0 ^{o}C in the winter to summer daytime temperatures exceeding 38 ^{o}C (Phillips 1976; CDFW 2018). Average annual precipitation is ~76 cm, a majority of which occurs from October through May (Ferrier and Roberts 1973; Bureau of Land Management [BLM] 1986). Perennial water sources include Cache Creek and Bear Creek, Lake Pillsbury and the Eel River, and East Park Reservoir and Stony Creek in the Cache Creek/Bear Valley, Lake Pillsbury, and East Park Reservoir MUs, respectively (Phillips 1976; CDFW 2018). This region is characterized by harsh, rugged terrain; rolling foothills and flats permeate jagged peaks and valleys. Elevation ranges from 30 m at its lowest point in the Sacramento Valley to 2,176 m at its highest point at Black Butte Mountain. Dominant vegetation communities include blue oak (*Quercus douglasii*) woodland, perennial grasslands, chamise (*Adeonostoma fasciculatum*) –redshank (*A. sparsifolium*) chaparral, blue oak-foothill pine (*Pinus sabiniana*) woodlands, mixed conifer and hardwood forests, annual grasslands, lacustrine, and agricultural pastures. Land ownership is a mix of public and privately held parcels, with cattle ranching and outdoor recreation (i.e. camping, hunting, hiking, horseback riding, etc.) as primary uses (Brandvold 1969; Booth et al. 1988; Burns 2004; CDFW 2018).

Field Methods

We conducted pellet surveys during Jun–Aug of 2017–2019. Summer sampling was conducted to avoid changes in group dynamics during the calving and rutting seasons and to ensure a closed population. We sampled CC in 2017 and 2019, and LPB in 2018 and 2019 for replication. After the initial season sampling EPR (2018), which was severely limited by access, we opted not to repeat that survey for a second year.

To select survey units, we stratified the landscape according to a SDM to ensure that we focused most pellet-sampling effort in portions of the landscape more likely to be used by elk but also sampled other portions of the landscape sufficiently to avoid biasing estimates (Batter et al. Ch. 1). For each of the three populations, we overlaid grids comprised of 2-km^2 cells, which we expected to approximate average home range radius of tule elk (Efford and Fewster 2013; Brazeal and Sacks 2018). Cells were dichotomized into high-probability (HP) and low-probability (LP) based on the SDM (Batter et al. Ch. 1). We assigned each grid cell a random cell identification number and used a random number generator to determine cells eligible for sampling. When a cell was selected for sampling, all direct-bordering cells were then disqualified for sampling unless a border cell did not have the same prediction of occurrence as the cell sampled, or the bordering cell was diagonal from the sampled cell. With some exceptions due to constraints associated with land ownership and access, sampled grid cells

were ≥2 km apart. We ultimately selected 21% of the high-probability cells and 15% of the low-probability cells to sample within the three estimated range boundaries (21% HP, 16% LP at Cache Creek; 25% HP, 12% LP at Lake Pillsbury; 16% HP, 18% LP at East Park Reservoir) (Batter et al. Ch. 1).

We surveyed each cell using 6-km triangle transects (Helle et al. 2016) when practical and 4.5-km out-over-back transects otherwise. For triangle transects, we nested an equilateral triangle within the 2-km grid cell, which began and ended at the same point. For out-over-back transects, we assigned starting locations randomly on the grid cell boundary from which the starting direction was typically dictated by land ownership/access. Transects extended 2 km out to the opposite boundary, 500 m over along that boundary, and 2 km back to the beginning boundary parallel to the first segment. Orientation of both transect methods was determined by topography, landownership, parking/access points, potential for elk use, safety and feasibility, or a combination of these factors.

We applied a flexible sampling approach along the transect allowing the surveyor to deviate ≤150 m from the transect line on either side. This approach was used to maximize efficiency to collect fecal pellets (e.g., if an obvious bedding area was observed adjacent to the transect, or if the surveyor was given a choice between suitable and unsuitable/inaccessible vegetation, terrain, etc.). Using a hand-held GPS device we recorded locations of each pellet group sampled and saved our tracks to measure effort. We collected 5–8 fecal pellets from each relatively fresh pellet group we encountered and placed them in 50 mL specimen jars. On the same day as collection, we filled specimen jars with 95–100% ethanol for preservation, and stored them at room temperature for 1–4 months until DNA extraction. We visited each transect once, with spatial re-detections of individuals acting as replicates during sampling occasions (Efford 2011; Sollmann et al. 2012).

Telemetry

To explore aggregation of female and male elk, we used telemetry data and aerial survey counts collected as part of a broader study (Bush et al. 2020; Bush *in prep.*). Briefly, we collected GPS telemetry locations from 66 elk, including 34 F and 32 M, which were monitored simultaneous with the surveys using only locations from the same 3-month periods pellet surveys were conducted. Elk were captured December 2016–October 2019 using ground-based free-range darting with a combination of tiletamine/zolazepam (1,000 mg) and xylazine (400 mg) and air-based net-gunning/manual restraint of adult male and female elk. All capture activities were performed by CDFW personnel and followed guidance and approval from the CDFW Wildlife Investigations Laboratory (CDFW 2018b). Elk were fixed with GPS Collars (Model: LifeCycle 800 GlobalStar, Lotek Wireless, Newmarket, Ontario, Canada) programmed to collect latitude and longitude of the elk's location every 13 hours.

Genetic Analyses

We completed all laboratory analyses at the Mammalian Ecology and Conservation Unit of the University of California, Davis Veterinary Genetics Laboratory. We evaporated ethanol from 1 to 2 pellets at 21^{o}C overnight; we then agitated the outside of the pellets with ≥2 mL of buffer ATL (Qiagen, Valencia, CA, USA) for 1 hour to remove epithelial cells from the outer surface of the

pellet(s) into the buffer. We extracted DNA from the cells in the buffer using the Qiagen Blood and Tissue kit according to the manufacturer's protocol, except we eluted DNA in 50 µL of buffer AE (Qiagen) to acquire sufficiently concentrated DNA samples. We genotyped the DNA samples using a multiplex assay developed specifically for tule elk to address low polymorphism in previous elk markers (Sacks et al. 2016). We included 19 microsatellite markers: TE179, TE85, TE132, TE84, TE185, TE45, TE182, TE68, TE83, T501, TE169, TE105, TE88 (Sacks et al. 2016), T26, T193, T501, T172, T108 (Jones et al. 2002), and a sex-typing marker from the Y chromosome (SRY; Wilson and White 1998). We used an ABI 3730 (Applied Biosystems, Grand Island, NY, USA) and internal size standards (500-LIZ; Applied Biosystems) for electrophoresis, with alleles scored manually using electropherograms visualized in Program STRand (version 2.4.89; Toonan and Hughes 2001). We attempted to genotype each DNA sample in two independent Polymerase Chain Reaction (PCR) replicates. All PCR sets were conducted with two negative PCR controls to confirm that samples were not contaminated.

Sample Genotypes and Individual Identification

We constructed sample genotypes from each pair of replicate genotypes. To maximize accuracy and resolution, we excluded sample genotypes with <19 loci. To assign individuals based on the sample genotypes, we needed to allow for some number of allele mismatches due to genotyping error, while minimizing the risk of erroneously assigning genotypes from two closely related but distinct individuals (i.e. siblings, parents-offspring, etc.) to the same individual. We used the R package *allelematch* (v. 2.5, http://ecologics.ucalgary.ca/lab/software/), which uses a dissimilarity matrix from pairwise comparisons among multilocus genotypes to identify clusters of genotypes identified as unique individuals (Galpern et al. 2012). Using *allelematch*, we determined that the optimal threshold for number of allele mismatches allowed between sample genotypes considered the same individual was ≤2. For pairwise genotypes with exactly 2 allelic mismatches, we considered the number of samples sharing the genotype, whether mismatching loci were heterozygous, and the quality of the genotypes based on peak heights in the electropherogram. After assigning individuals, we derived consensus genotypes for each individual across all such sample genotypes.

Using the consensus genotypes for each individual, we estimated genetic summary statistics for both the total study area (i.e., elk across all three populations) and separately for each population in the R package *diveRsity* (Keenan et al. 2013). We tested for departures from Hardy-Weinberg equilibrium (HWE) at each locus in each population and estimated the probability of any two unique genotypes matching using GENALEX 6.5 (Peakall and Smouse 2012). We estimated both the probability of identity (P_{ID}; the probability of two randomly selected individuals sharing the same genotype) and the probability of identity of siblings (P_{SIB}; the probability that two siblings share the same genotype) (Waits et al. 2001). We calculated sex ratio (% male: % female) and 95% confidence limits directly from the sample genotypes for each population (Lounsberry et al. 2015).

Spatial Capture-Recapture Modeling

We estimated elk density using the R package *secr* (v. 4.2; Efford 2004, 2020) which applies maximum likelihood estimation to SCR analyses (Efford et al. 2009a). SCR models directly depend on adequate number of unique individuals captured and recaptured at multiple spatial locations to derive density estimates (Efford and Boulanger 2019). The approach combines a "state model", which describes the density ($\hat{D}$) of animal activity centers, and an "observation model", which describes the probability of detection relative to the distance from the activity center (Efford 2004). The observation model takes a functional form, such as half-normal or negative exponential, that describes probability of detection in relation to distance from the activity center according to two parameters: probability of detection of an individual at its activity center (*g0*) and the individual's scale of movement (σ), which is proportional to the home range radius (Efford 2020). Similar to conventional CR, SCR assumes a closed population, however, in contrast to traditional CR, SCR is associated with an explicit spatial area and accounts for spatial heterogeneity of individual animals and detectors, which gives rise to variation in detection probability (Borchers and Efford 2008; Royle et al. 2014).

An additional advantage of SCR models is to allow use of covariates in both the "state" and "observation" models: the "state model" can incorporate covariates that can be used to predict explicitly how density is distributed across the landscape in response to environmental factors, such as vegetation communities; the "observation model" can incorporate covariates, such as sex, to allow for distinct home range sizes of males and females, or, using hybrid mixture models (*hcov*), latent variables, such as when unobserved demographic classes exhibit different space use (Williams et al. 2002; Efford and Fewster 2013). As *secr* models are computationally intensive, we used a step-wise approach (Brazeal et al. 2017; Loosen et al. 2018). For each population, we first identified the optimal detection function (half-normal vs. negative exponential) to use for the entirety of the modeling process (step 1). We then tested for an effect of sex using both the hybrid mixture model and the null model (step 2). Lastly, we applied the best observation model as a base to fit state models with both single and multiple variables (step 3; Supplementary Table 1). Because we sampled the same populations during the same seasons across the same space, we expected neither the probability of detection nor the scale of movement to vary temporally or spatially and therefore restricted parameters in the detection function to sex. For the null detection model, we assumed a homogeneous observation model ($g0$~1 σ~1) for each population across sample years; we similarly assumed a homogeneous observation model for the *hcov* detection model, except we included the *pmix* parameter as a constant effect ($g0$~h2 σ~h2).

Once we settled on the optimal observation model (detection function) for each population, we modeled density relative to spatial heterogeneity. We created 7 two-year (Cache Creek and Lake Pillsbury) and a one-year (East Park Reservoir) state models that differed in sources of variation affecting density. We modeled spatial heterogeneity using environmental covariates considered important predictor variables for tule elk (Batter et al. Ch. 1): average daily temperature change ("bio2" in the WorldClim dataset; Hijmans et al. 2005), habitat quality (Supplementary Table 2; CalFIRE 2015), and soil taxonomy (Supplementary Table 3; Web Soil Survey 2017). For habitat quality, we grouped habitat into "prime", "moderate", and "poor" habitat classes based on predictive values from our habitat suitability model (Batter et al. Ch. 1), while soil was classified by taxonomic subgroup.

Unique individuals identified via genetic analyses were used to construct capture histories for density modeling. We first discretized sample transects at 75-m intervals to act as proximity detectors in ArcGIS (v. 10.7.1) (ESRI 2019). Proximity detectors allow an individual or multiple individuals to only be detected a single time at a single detector, but the individual(s) may be detected at multiple detectors during one sampling occasion (i.e. DNA hair snares, fecal collection; Efford et al. 2009b; Borchers 2012; Sollmann et al. 2012). We therefore snapped pellet samples to the nearest detector, then thinned data so that individuals were only detected once at a given proximity detector (≤75 m). Detections of the same individual at different detectors (>75 m) serve as recaptures, which can occur within the same sampling occasion (Efford et al. 2009b; Borchers 2012; Sollmann et al. 2012; Efford 2020). Thus, proximity detectors enable density estimates from a single field sampling occasion within a buffered extent (Efford et al. 2009a, b; Royle et al. 2014). The buffered extent of the state space has no effect on density estimates so long as the mask is sufficiently large to include the activity centers of all animals potentially exposed to sampling (Borchers and Efford 2008; Royle and Young 2008). We specified a buffer width of 4 times the initial σ for each population ($4\sigma_{CC}$ = 4.75 km, $4\sigma_{LPB}$ = 2.95 km, $4\sigma_{EPR}$ = 5.02 km) (Sollmann et al. 2012; Efford 2020). Each model's density estimates reached a plateau well before its respective buffer width, suggesting buffer width was sufficiently large enough to include all potential animals with activity centers in the sample area for each population under consideration (Efford 2020). We evaluated support for competing models by comparing Akaike's Information Criterion (AIC) (Burnham and Anderson 2002).

We applied the *derived* function in *secr* to obtain density estimates for each population and for each of its sample years. This function also provided 95% confidence intervals and estimates of relative standard error of density (RSE($\hat{D}$)) for each density estimate. We projected abundance ($\hat{N}$) estimates according to the best-supported model for each population across the corresponding range boundary for each population using the *predictDsurface* function. As a heuristic check we calculated the proportion of female and male pellet samples to independently validate the *pmix* parameter for each population. Assuming male and female elk defecate at comparable rates and deposit a similar amount of pellet groups in the environment, and any observation or detection biases being equal, the ratio of male pellets sampled relative to female pellets provides an unbiased estimate of sex ratio (Lounsberry et al. 2015).

Home Range Size and Independence of Activity Centers

We used the R package *adehabitatHR* (Calenge 2006) to estimate the 95% kernel utilization distribution (KUD) for each individual during the sample seasons in which they were monitored. We averaged these KUDs across the two seasons of monitoring for Cache Creek and Lake Pillsbury elk. We also examined spatial clustering of home ranges and activity centers (centroids). The SCR approach assumes a Poisson point observation process for animal activity centers, implying that home ranges are independent and randomly distributed over the landscape. In practice, these assumptions are often unrealistic and violated, particularly in more social species, but estimates are relatively robust except when violations are extreme (Bischof et al. 2018). In the most extreme cases of correlated space use (e.g., all individuals

share a single activity center), the primary effect is to overestimate precision. To explore the magnitude of group association for males and females, we used GPS collar locations collected during our fecal sampling seasons (Jun–Aug in 2017 and 2019 at Cache Creek, in 2018 and 2019 at Lake Pillsbury, and 2018 at East Park Reservoir) from 66 elk (32 male, 34 female). To estimate each elk's activity center we used the mean center spatial statistics tool, then quantified both observed and expected (if randomly distributed and, therefore, independent of one another) nearest-neighbor distances among activity centers using the average nearest neighbor spatial statistics tool in ArcGIS (v. 10.7.1) (ESRI 2019). We then tested the null hypothesis for each sex that the slope of a regression line describing observed nearest-neighbor distances versus expected nearest-neighbor distances was equivalent to 1 (Zar 1999).

Results

Genotyping and Individual Identification

We collected a total of 1,616 pellet groups across three field seasons (Table 1). After we eliminated samples with <19 out of 20 loci amplified, 1,002 samples remained for analyses. We identified 425 unique individual elk across the study region (265 at Cache Creek, 128 at Lake Pillsbury, and 32 at East Park Reservoir) from the 1,002 successfully genotyped samples (Table 1).

Based on the 425 individual genotypes, microsatellite loci exhibited an average of 3.3 alleles (range: 2–7), H_O averaged (± SD) 0.31 ± 0.10, and H_E averaged 0.39 ± 0.12 (Table 2). When adjusted for multiple comparisons, no loci were significantly out of HWE in >1 population (Supplementary Tables 4-6). Therefore, deviations from HWE were likely due to population substructure rather than allelic dropout or null alleles (see also Batter et al. Ch. 3). All loci were retained in analyses. The cumulative polymorphism of these markers was sufficient to achieve P_{ID} of 3.6×10^{-8} and P_{SIB} of 2.5×10^{-4} in the pooled dataset (Table 2) and ranged from 4.3×10^{-7} to 2.2×10^{-6} (P_{ID}) and 6.3×10^{-4} to 1.7×10^{-3} (P_{SIB}) among the three population data subsets (Supplementary Tables 4–6).

Spatial Capture-Recapture Modeling

After eliminating 130 sample genotypes that were assigned to the same individual as other samples at the same detector, we retained 872 genotyped samples for SCR analysis (Table 1). The exponential detection function was at least as well or better supported than the half normal function for all 3 populations in both combined- and single-sex models (Table 3; Supplementary Table 4). Therefore, we adopted the exponential function for all subsequent observation models.

Among the combined-sex models, the spatially heterogeneous hybrid mixture models were similarly or better supported than the null model in all populations (Table 3) and, therefore, adopted for modeling habitat covariates. The most highly supported models differed across populations: the Cache Creek top model included bio2, habitat quality, and soil (Fig. 2A); the Lake Pillsbury top model included two variables habitat quality and soil (Fig. 2B); and East Park Reservoir's two top models received similar support, so we model-averaged these: a single-variable model, habitat quality, and a two-variable model, bio2 and habitat quality (Fig.2C;

Table 3). The null model was either the least well-supported or second least well-supported model in all three populations. Among the single-sex models, the null model was generally less well-supported than models including covariates (Supplementary Table 7). Top single-sex models were very similar to those of the combined-sex models for both sexes.

Density and Abundance Estimates

Density estimates were similar between combined-sex models with and without (null) covariates and single-sex models, although confidence intervals were wider for the latter two categories (Fig. 3; Table 4; Supplementary Table 8). Thus, estimates of density were not sensitive to the particular model. Overall, the Lake Pillsbury population had higher estimated density than both Cache Creek and East Park Reservoir populations. Due to its larger spatial extent, however, the Cache Creek population had the highest estimated abundance. Based on the top combined-sex models, for example, Cache Creek numbers were estimated at $\hat{N}$ = 325 (95% CI = 261–395) in 2017 and $\hat{N}$ = 318 (95% CI = 261–395) in 2019), followed by the Lake Pillsbury ($\hat{N}$ = 265 [95% CI = 208-340] in 2018 and $\hat{N}$ = 284 [95% CI = 227–359] in 2019), and East Park Reservoir ($\hat{N}$ = 62 [95% CI = 34–110] in 2018) populations (Table 4).

Sex Composition

Based on the two better-sampled populations, the proportional composition of males in the pellet groups was similar to those estimated from the hybrid mixture models, both of which were higher than male composition estimated from the single-sex models. In particular, at Cache Creek, 47% (95% CI: 43–51%) of pellet groups were male, compared to 43% (38–47%) males based on *pmix* of the combined-sex model, and 34.7% (28–39%) males based on the two single sex models. Similarly, at Lake Pillsbury, 50% (44–56%) of pellet groups were male, compared to 49% (41–56%) males based on *pmix* of the combined-sex model, and 36.7% (30–41%) males based on the two single sex models. These patterns suggest the presence of bias in at least one of the single sex models that was potentially compensated for by the use of both sexes in the combined sex model.

Home Range Size and Clustering of Activity Centers

Home range sizes (95% KUDs) varied across populations for both sexes during sample seasons, as illustrated at Cache Creek (Fig. 4A, B) and Lake Pillsbury (Fig. 5A, B). Females averaged 40 km^2, 5 km^2, and 22 km^2, while males averaged 71 km^2, 75 km^2, and 9 km^2, at Cache Creek, Lake Pillsbury, and East Park Reservoir, respectively. Activity centers and 95% KUDs were much more clustered for females compared to males within each population and across sample seasons. The slope of the regression line of observed nearest neighbor distances on expected nearest neighbor distances for females (β = 0.475, SE = 0.058, $P < 0.001$) but not males (β = 0.932, SE = 0.235, $P > 0.10$) was significantly lower than 1, as would be expected if they were distributed according to a Poisson distribution (Table 5).

Discussion

We undertook this study to address the need for a method of estimating the abundance of elk in a heterogeneous landscape. Because of its success in other ungulates, we chose to apply a fecal DNA-based SCR approach. However, elk populations come with the added challenge of a

group-living social system and, as a consequence, both more heterogeneous distribution on the landscape and the potential for highly correlated space use among individuals (McCullough 1969; Mackie 1970; Boyce et al. 2003; Proffitt et al. 2015). These qualities pose challenges both in terms of efficiency, in that strictly random or systematic sampling approaches would result in considerable time wasted searching in locations where no elk are present, and a lack of independence among elk activity centers, which violate basic model assumptions.

Our use of a species-distribution model to stratify the study area, allowing us to place more effort in locations likely to have more elk, was largely successful in allowing us to collect sufficient numbers of samples to obtain acceptably precise estimates in two of the three populations surveyed. In Cache Creek and Lake Pillsbury populations, our RSE estimates for the hybrid mixture models (10–14%) and single-sex models (14–21%) compared well with those generated for most other ungulates using an SCR approach, including deer (20–28%, Brazeal et al. 2017), giant eland (*Taurotragus derbianus*) (25–36%, Jůnek et al. 2015), as well as a study that used a captive population of tule elk of known size (12–29%, Brazeal and Sacks 2018). The precision of our estimates at East Park Reservoir, on the other hand, was poor, as expected from our small sample size. Our inadequate sampling of this population was due to a lack of access to private land. In the future, either more access to private lands needs to be secured or alternative approaches to abundance estimation, such as aerial line-transect approaches, will be needed to provide reliable abundance estimates for this population.

Although our estimates were sufficiently precise as measured by model variances, for estimates to be accurate, they and the associated estimates of precision must also have low bias. Our concern over the application of SCR to elk was based on expected violations to basic assumptions, independence of individual movements, in particular, that could potentially bias parameter and especially precision estimates (Bischof et al. 2020). To minimize violations in the assumption of independence, we sampled during a time of year when group sizes tended to be smaller and social interactions more fluid (McCullough 1969). Nevertheless, using the telemetry data, we found relatively strong clustering among females. This concern was greatest in the Lake Pillsbury population, where females were clustered in a single small location, effectively sharing a single activity center. Within the Cache Creek estimated range, we observed at least 4 distinct cow groups, each of which had highly correlated activity centers. Bischof et al. (2020) found that SCR models were relatively robust to clustered activity centers when there were multiple aggregations, such as at Cache Creek in our study, but potentially more severely biased when many individuals essentially shared a single activity center, which characterizes the Lake Pillsbury cow group.

Because of concerns over clustering, an auxiliary study was conducted in October 2019 (soon after our study of the same year) to estimate abundance of the Lake Pillsbury population using an alternative sampling and analysis approach (Bush *in prep*.; Sacks et al. 2020). Pellets were sampled late in the rut during a one-day intensive search focused on the Basin where the cow group was centered, and abundance was estimated using Capwire (Miller et al. 2005), which does not consider spatial distribution. The auxiliary study produced an estimate of female abundance (159, 95% CI: 106–189) that agreed with that of our combined-sex SCR model (147 females, 95% CI: 108–187), but which was considerably lower than our female-only

SCR model (283, 95%CI: 189–397). Thus, it appears that under the most extreme level of spatial clustering, SCR estimates can be highly biased. However, males did not exhibit significant clustering of activity centers, suggesting no reason to expect SCR estimates to be biased. Correspondingly, the inclusion of males in the modeling process was apparently sufficient to counter the effect of female clustering at Lake Pillsbury. In the other population, where females occurred in multiple cow groups, we found little difference between the estimates from single-sex models and combined-sex models, except that the latter were more precise. Altogether, these findings suggest that the SCR method is robust to the level of gregariousness we observed in female elk during summer in situations where multiple distinct cow groups occur or even in a case of extreme female clustering when males are included in the modeling (e.g., Fig. 5A,B).

Indeed, the general agreement among our different approaches to modeling elk density (i.e., combined sex, with and without covariates, single-sex) bodes well for the accuracy of parameter estimates, despite violations in model assumptions by the female component of the populations. Previous SCR estimation of density in captive ungulate populations of known size (including tule elk) also have proven to be relatively unbiased (Jůnek et al. 2015; Brazeal and Sacks 2018). A more likely issue in need of further investigation is the potential for overestimating precision which can lead to false inference (Bischof et al. 2020). Incorporating simulations specific to the study system is one possible solution to this problem (Kristensen and Kovach 2018; Royle et al. 2014). In our case, simulations of known numbers of aggregating individuals could help to improve understanding of SCR accuracy and precision for social species in general (López-Bao et al. 2018; Bischof et al. 2020; McFarlane et al. 2020). Beyond simulations, future research opportunities exist to develop models that address group association either within an SCR framework or those used in other abundance estimation methods (e.g., Hickey and Sollmann 2018). Accounting for group association within models is desirable not only to address problems related to biases, but to also allow for other useful measures including strength of cohesion, group size, and even group membership as a latent variable (Bischof et al. 2020).

Beyond estimates—habitat modeling and other added benefits. An added benefit of SCR methods is the inclusion of landscape covariates to determine factors associated with local density of small, isolated tule elk populations. Elk density was highest in areas of fertile, loam-dominant soil types, prime habitat classification (i.e. perennial grassland; Supplementary Table 2), and within an average daily temperature-change range of 1.62-1.70° C. These variables have been shown to be important predictors in tule elk habitat suitability both in past field observations as well as previous research (Van Dyke 1902; Grinnell 1933; Brandvold 1969; McCullough 1969; Ferrier and Roberts 1973; Phillips 1976; Booth et al. 1988; McCullough et al. 1994). These results also align with variables that most influenced the predicted relative probability of occurrence (RPO) of tule elk in a companion study to guide field sampling efforts (Batter et al. Ch. 1). Our model predicted elk presence to be positively associated with an average daily temperature-change range between 1.64–1.79° C, prime habitat types (RPO > 0.10; including perennial grassland, hardwood montane, and blue oak-valley oak woodland, among others), and loam-dominant, highly fertile soils.

It is important to note that our habitat model predicts areas of suitable conditions across the landscape relative to conditions of sites where elk were observed, and these predictions are unassociated with projections of density or abundance. However, visual comparison of areas of agreement (higher RPO and relatively greater density) across surfaces can lend insight into the utility of each method for describing elk spatial assortment. Our results here indicate areas of relatively greater surface density extrapolated from our *secr* predictions generally paralleled areas of higher RPO predictions at Cache Creek and Lake Pillsbury (Supplementary Figs. 1-2), while East Park Reservoir predictions were less complimentary (again, likely influenced by small sample size and lower precision of *secr* estimates; Supplementary Fig. 3). Furthermore, as we expected, predicted surface density across populations was spatially heterogeneous. In contrast, fDNA-SCR studies performed on California mule deer (*O.h. californicus*) and Columbian blacktail deer (*O.h. columbianus*) were projected to be distributed more homogeneously across the landscape ($\hat{D}$ = 4.0–6.0 deer/km^2) (Brazeal et al. 2017).

In addition to its value in estimating density, abundance, and associated landscape covariates, noninvasive fecal pellet sampling provides genetic identification of individuals which has applied conservation and management value, and can offer insight into the genetic patterns underlying population dynamics (Buchalski et al. 2015; Sun et al. 2017). For example, with these genetic data, wildlife managers can quantify levels of genetic diversity and differentiation within and among populations, and reveal population structure allowing for inference of landscape connectivity, patterns of dispersal, and possible range expansion (Hicks et al. 2007; Onorato et al. 2007). Evaluation of genetic diversity as well as gene flow and relatedness of elk populations is not only desirable but essential to maintain long-term population viability for those that exist in small, isolated populations occupying patchy habitat (McCullough et al. 1996; Buchalski et al. 2015). These features describe most, if not all, tule elk populations and, given their life history and past genetic bottlenecks, only intensifies the need for periodic genetic monitoring of tule elk (Williams et al. 2004; CDFW 2018).

Management Implications

Previous studies have successfully applied noninvasive fecal DNA-based SCR to more uniformly distributed, and less gregarious species, such as mule and black-tailed deer (Brazeal et al. 2017; Furnas et al. 2020). This study demonstrated the applicability of SCR to a highly social species in which individuals use the landscape less independently of one another and more heterogeneously, and represents a step forward in CDFW's goals for more robust statistical population monitoring for elk (CDFW 2018). A fDNA-SCR framework can provide improved estimates of elk density and abundance in comparison to minimum count indices facilitating more informed harvest quotas and information fundamental to conservation and restoration actions. In addition, the utility of fDNA sampling can provide data on additional population parameters important for management of ungulates, such as sex ratio (Brazeal et al. 2017; Furnas et al. 2018), population genetics (e.g., genetic diversity, population structure) (Buchalski et al. 2015; Sun et al. 2017), diet (e.g. DNA metabarcoding) (Kartzinel et al. 2015), as well as spatial and temporal distribution (Azad et al. 2019; Furnas et al. 2020).

Section II Literature Cited

Azad, S., K. McFadden, J.D. Clark, T. Wactor, and D.S. Jachowski. 2019. Applying spatially explicit capture-recapture models to estimate black bear density in South Carolina. Wildlife Society Bulletin 43(3):500-507.

Batter, T.J., J.P. Bush, and B.N. Sacks. In prep. Evaluation of a species distribution model approach to guide noninvasive fecal-DNA surveys for tule elk (Cervus canadensis nannodes) in Northern California. Part I. Report to the California Department of Fish and Wildlife.

Batter, T.J., J.P. Bush, and B.N. Sacks. In prep. Assessing genetic diversity and connectivity among tule elk (Cervus canadensis nannodes) populations in Northern California. Part III. Dissertation, University of Califronia, Davis.

Bischof, R., C. Bonenfant, I. M. Rivrud, A. Zedrosser, A. Friebe, T. Coulson, A. Mysterud, and J.E. Swenson. 2018. Regulated hunting re-shapes the life history of brown bears. Nature Ecology and Evolution 2:116-123.

Bischof, R., P. Dupont, C. Milleret, J. Chipperfield, and J.A. Royle. 2020. Consequences of ignoring group association in spatial capture-recapture analysis. Wildlife Biology (1):1-10.

Booth, J., J. Swanson, and D. Koch. 1988. Cache Creek tule elk management unit management plan. March 31, 1988. California Department of Fish and Game, Sacramento, CA.

Borchers, D. 2012. A non-technical overview of spatially explicit capture-recapture models. Journal of Ornithology 152:435-444.

Borchers, D., and M.G. Efford. 2008. Spatially explicit maximum-likelihood methods for capture-recapture studies. Biometrics 64:377-385.

Boyce, M.S., J.S. Mao, E.H. Merrill, D. Fortin, M.G. Turner, J. Fryxell, and P. Turchin. 2003. Scale and heterogeneity in habitat selection by elk in Yellowstone National Park. Écoscience 10(4): 421-431.

Brandvold, G. 1969. Cache Creek Ridge wildlife habitat area. Unpublished report. Bureau of Land Management, Ukiah, CA.

Brazeal, J.L., and B.N. Sacks. 2018. Use of an enclosed elk population to assess two noninvasive methods for estimating population size. Report to the California Department of Fish and Wildlife, November 6, 2018, 34 pp.

Brazeal, J.L., T. Weist, and B.N. Sacks. 2017. Noninvasive genetic spatial capture-recapture for estimating deer population abundance. The Journal of Wildlife Management 81(4):629-640.

Brinkman, T.J., M.K. Schwartz, D.K. Person, K.L. Pilgrim, and K.J. Hundertmark. 2010. Effects of time and rainfall on PCR success using DNA extracted from deer fecal pellets. Conservation Genetics 11: 1547-1552.

Brinkman T.J., D.K. Person, F.S. Chapin, W. Smith, and K.J. Hundertmark. 2011. Estimating abundance of Sitka black-tailed deer using DNA from fecal pellets. Journal of Wildlife Management 75:232-242.

Buchalski, M.R., A.Y. Navarro, W.M. Boyce, T.W. Vickers, M.W. Tobler, L.A. Nordstrom, J.A. Garcia, D.A. Gille, M.C.T. Penedo, O.A. Ryder, and H.B. Ernest. 2015. Genetic population structure of Peninsular bighorn sheep (Ovis canadensis nelsoni) indicates substantial gene flow across US-Mexico border. Biological Conservation 184: 218-228.

Bureau of Land Management (BLM). 1972. Tule elk range inventory. Final monthly report, Oct-Nov. 1972. BLM, Ukiah, CA. Unpublished report.

BLM. 1977. A report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1978. Second annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1979. Third annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1980. Fourth annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1981. Fifth annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1982. Sixth annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1983. Seventh annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1985. Cache Creek tule elk wildlife habitat management plan. United States Department of Interior, Bureau of Land Management, Ukiah District.

BLM. 1986. Eighth annual report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1989. Ninth report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

BLM. 1992. Tenth report to congress: the tule elk in California. Bureau of Land Management, Sacramento, CA.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer, New York, New York, USA.

Burns, R. 2004. Cache Creek coordinated resource management plan/environmental assessment final report. Ukiah Field Office, Ukiah, CA.

Bush, J.P. In prep. Colusa/Lake tule elk project: final report on objectives. Internal report. California Department of Fish and Wildlife, Sacramento, CA.

Bush, J.P., T.J. Batter, R. Landers, and K. Denryter. 2020. Report on aerial surveys of tule elk (Cervus canadensis nannodes) in 2018-2019 in Bear Valley, Cache Creek, East Park Reservoir, and Lake Pillsbury tule elk hunt zones. Internal report. California Department of Fish and Wildlife, Sacramento, CA.

California Department of Fish and Game (CDFG). 1974. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1978. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1980. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1982. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1987. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA

CDFG. 1989. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1991. Tule elk in California: a report to the legislature. California Department of Fish and Game, Sacramento, CA.

CDFG. 1995. Report to the legislature regarding tule elk. California Department of Fish and Game, Sacramento, CA.

CDFG. 1998. Report to the legislature regarding tule elk. California Department of Fish and Game, Sacramento, CA.

CDFG. 2002 Final Environmental Document regarding elk hunting. April 5, 2002. California Department of Fish and Game. Fig. 24.

CDFG. 2004. Final Environmental Document Regarding Elk Hunting. April 12, 2004. California Department of Fish and Game. 178 pp.

California Department of Fish and Wildlife (CDFW). 2013. Wildlife restraint handbook. Wildlife Investigations Laboratory. California Department of Fish and Wildlife, Rancho Cordova, CA.

CDFW. 2015. California statewide wildlife action plan, 2015 update: a conservation legacy for Californians. Eds: A.G. Gonzales and J. Hoshi. Prepared with assistance from Ascent Environmental, Inc., Sacramento, CA.

CDFW. 2018. Conservation and management plan for elk. State of California Department of Fish and Wildlife. Sacramento, CA. Retrieved from: https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=162912&inline

California Department of Forestry and Fire Protection (CalFIRE). (2015). Fire and Resource Assessment Program (FRAP) Vegetation (FVEG15_1). Sacramento, California. Retrieved from http://frap.fire.ca.gov/data/frapgisdata-sw-fveg_download

Calenge, C. 2006. The package "adehabitat" for the R software: A tool for the analysis of space and habitat use by animals. Ecological Modelling 197: 516–519.

Caughley, G. 1974. Bias in aerial survey. Journal of Wildlife Management 38: 921 -933.

Efford, M.G. 2004. Density estimation in live-trapping studies. Oikos 106:598-610.

Efford, M.G. 2011. Estimation of population density by spatially explicit capture-recapture analysis of data from area searches. Ecology 92(12):2202-2207.

Efford, M.G. 2020. secr 4.2 – spatially explicit capture-recapture in R. https://www.otago.ac.nz/density/pdfs/secr-overview.pdf.

Efford, M.G., D.L. Borchers, and A.E. Byrom. 2009a. Density estimation by spatially explicit capture-recapture: likelihood-based methods. Pages 255-269 in D.L. Thomson, E.G. Cooch, and M.J. Conroy eds. Modeling demographic processes in marked populations. Springer, NY, USA.

Efford, M., D.K. Dawson, and D.L. Borchers. 2009b. Population density estimated from locations of individuals on a passive detector array. Ecolog 90:2676-2682.

Efford, M.G., and R.M. Fewster. 2013. Estimating population size by spatially explicit capture-recapture. Oikos 122:918-928.

Efford, M.G., and J. Boulanger. 2019. Fast evaluation of study designs for spatially explicit capture-recapture. Methods in Ecology and Evolution 10(9): 1529-1535.

ESRI. 2011. ArcGIS desktop: release 10. Redlands, CA, USA: Environmental Systems Research Institute.

ESRI. 2019. ArcGIS desktop: release 10.7. Redlands, CA, USA: Environmental Systems Research Institute.

Ferrier, G.J., and E.C. Roberts. 1973. The Cache Creek tule elk range. Cal-Neva Wildlife. 10 pp.

Furnas, B.J., R.H. Landers, S. Hill, S.S. Itoga, and B.N. Sacks. 2018. Integrated modeling to estimate population size and composition of mule deer. Journal of Wildlife Management 82(7):1429–1441.

Furnas, B.J., R.H. Landers, R.G. Paiste, and B.N. Sacks. 2020. Overabundance of black-tailed deer in urbanized coastal California. The Journal of Wildlife Management Accessed: https://wildlife.onlinelibrary.wiley.com/doi/pdf/10.1002/jwmg.21849

Galpern, P., M. Manseau, P. Hettinga, K. Smith, and P. Wilson. Allelematch: an R package for identifying unique multilocus genotypes where genotyping error and missing data may be present. Molecular Ecology Resources 12: 771-778.

Geist, V. 2002. Adaptive behavioral strategies. In D.E. Toweill, and J.W. Thomas, eds. North American elk: ecology and management. Smithsonian Institution Press, Washington, D.C., USA. Pages 389-433.

Goode, M.J., J.T. Beaver, L.I. Muller, J.D. Clark, F.T. van Manen, C.A. Harper, and P.S. Basinger. 2014. Capture-recapture of white-tailed deer using DNA from fecal pellet groups. Wildlife Biology (20): 270-278.

Google Inc. 2005. Google Earth. Mountain View, CA.

Grinnell, J. 1933. Review of the recent mammal fauna of California. University of California Publications in Zoology. 40 (2): 71-234.

Helle, P., K. Ikonen, and A. Kantola. 2016. Wildlife monitoring in Finland: online information for game administration, hunters, and the wider public. Canadian Journal for Research 46:1491-1496.

Hickey, J.R., and R. Sollmann. 2018. A new mark-recapture approach for abundance estimation of social species. PLoS One 13:e0208726.

Hicks, J.F., J.L. Rachlow, O.E. Rhodes, Jr., C.L. Williams, and L.P. Waits. 2007. Reintroduction and genetic structure: Rocky Mountain elk in Yellowstone and the Western states. Journal of Mammalogy, 88(1): 129-138.

Hijmans, R. J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978.

Jones, K.C., K.F. Levine, and J.D. Banks. 2002. Characterization of 11 polymorphic tetranucleotide microsatellites for forensic applications in California elk (Cervus elaphus canadensis). Molecular Ecology Notes 2:425-427.

Jůnek, T., P.J. Vymyslická, K. Hozdecka, and P. Hejcmanová. 2015. Application of spatial and close capture-recapture models on known population of the Western Derby Eland (Taurotragus derbianus derbianus) in Senegal. PLoS ONE DOI:10.1371/journal.pone.013652

Kartzinel, T.R., P.A. Chen, T.C. Coverdale, D.L. Erickson, W.J. Kress, M.L. Kuzmina, D.L. Rubenstein, W. Wang, and R.M. Pringle. 2015. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. PNAS 112(26): 8019-8024.

Kauffman, E. 2003. Atlas of the biodiversity of California. Climate and Topography. California Department of Fish and Game, Sacramento, CA. pp. 12-15.

Keenan, K., P. McGinnity, T.F. Cross, W.W. Crozier, P.A. Prodöhl. 2013. diveRsity: An R package for the estimation of population genetics parameters and their associated errors. Methods in Ecology and Evolution 4(8):782-788.

Kristensen, T.V., and A. I. Kovach. 2018. Spatially explicit abundance estimation of a rare habitat specialist: implications for SECR study design. Ecosphere (5):e02217. 10.1002/ecs2.2217

Lancia, R.A., W.L. Kendall, K.H. Pollock, and J.D. Nichols. 2005. Estimating the number of animals in wildlife populations. Pages 106-153 in C.E. Braun, editor. Techniques for wildlife investigations and management. Sixth edition. The Wildlife Society, Bethesda, Maryland, USA.

Loosen, A.E., A.T. Morehouse, and M.S. Boyce. 2018. Land tenure shapes black bear density and abundance on multi-use landscape. Ecology and Evolution 9:73-89.

López-Bao, J.V., R. Godinho, C. Pacheco, F.J. Lema, E. Garcia, L. Llaneza, V. Palacios, and J. Jimenez. 2018. Toward reliable population estimates of wolves by combining spatial capture-recapture models and non-invasive DNA monitoring. Nature 8:2177. DOI:10.1038/s41598-018-20675-9

Lounsberry Z.T., T.D. Forrester, M.T. Olegario, J.L. Brazeal, H.U. Wittmer, and B.N. Sacks. 2015. Estimating sex-specific abundance in fawning areas of a high-density Columbian black-tailed deer population using fecal DNA. Journal of Wildlife Management 79(1): 39-49.

Lukacs, P.M. and K.P. Burnham. 2005. Review of capture-recapture methods applicable to noninvasive genetic sampling. Molecular Ecology 14:3909-3919.

Mackie, R.J. 1970. Range ecology and relations of mule deer, elk, and cattle in the Missouri River Breaks, Montana. Wildlife Monographs 20:3-79.

McCullough, D.R. 1969. The tule elk: its history, behavior, and ecology. University of California

Publications in Zoology, Berkeley. 88: 209 pp.

McCullough, D.R., F.W. Weckerly, P.I. Garcia, and R.R. Evett. 1994. Sources of inaccuracy in black-tailed deer herd composition counts. Journal of Wildlife Management 58(2):319-329.

McCullough, D.R., J.D. Ballou and J.K. Fischer. 1996. From bottleneck to metapopulation: recovery of the tule elk in California in Metapopulations and wildlife conservation 375-410. Island Press, Washington, D.C., USA.

McFarlane, S., M. Maneau, R. Steenweg, D. Hervieuz, T. Hegel, S. Slater, and P.J. Wilson. 2020. An assessment of sampling designs using SCR analyses to estimate abundance of boreal caribou. Ecology and Evolution 10(20): 11631-11642.

Murie, O.J. 1951. The elk of North America. Wildlife Management Institute, Washington D.C. 376 pp.

O'Connor, P.M. 1987. Tule elk ecology and home range characteristics at Cache Creek, California. Master's Thesis. Humboldt State University, Arcata, CA.

Onorato, D.P., E.C. Hellgren, R.A. Van Den Bussche, D.L. Doan-Crider, and J.R. Skiles, Jr. 2007. Genetic structure of American black bears in the desert southwest of North America: conservation implications for recolonization. Conservation Genetics 8: 565-576.

Peakall, R., and P.E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537-2539.

Phillips, W.E. 1976. The conservation of the California tule elk: a socioeconomic study of a survival problem. Edmonton: The University of Alberta Press. 120 pp.

Pollock, K.H. 1991. Modeling capture, recapture, and removal statistics for estimation of demographic parameters for fish and wildlife populations: past, present, and future. Journal of the American Statistical Association 86: 225-238.

Pollock, K.D., J.D. Nichols, C. Brownie, and J.E. Hines. 1990. Statistical inference for capture-recapture experiments. Wildlife Monographs 107.

Proffitt, K.M., N. Anderson, P. Lukacs, M.M. Riordan, J.A. Gude, and J. Shamhart. 2015. Effects of elk density on elk aggregation patterns and exposure to brucellosis. Journal of Wildlife Management 79(3): 373-383.

Raedeke, K.J., J.J. Millspaugh, and P.E. Clark. 2002. North American elk: ecology and management. D.E. Toweill and J.W. Thomas, eds. Smithsonian Institution Press, Washington, D.C.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. Available from: https://www.R-project.org/.

Royle, J.A., R.B. Chandler, R. Sollmann, and B. Gardner. 2014. Spatial capture-recapture. Elsevier Press. 577 pp.

Royle, J.A., and K.V. Young. 2008. A hierarchical model for spatial capture-recapture data. Ecology 89(8):2281-2289.

Sacks, B.N., Z.T. Lounsberry, T. Kalani, E.P. Meredith, and C. Langner. 2016. Development and characterization of 15 polymorphic dinucleotide microsatellite markers for tule elk using HiSeq3000. Journal of Heredity 107(7): 666-669.

Sacks, B.N., J.P. Bush, and T.J. Batter. 2020. Lake Pillsbury basin auxiliary elk survey. Report to the California Department of Fish and Wildlife. October 8, 2020.

Schoenecker, K.A., and B.C. Lubow. 2016. Application of a hybrid model to reduce bias and improve precision in population estimates for elk (Cervus elaphus) inhabiting a cold desert ecosystem. Journal of King Saud University – Science 28 (3): 205-215.

Sollmann, R., B. Gardner, and J.L. Belant. 2012. How does spatial study design influence density estimates from spatial capture-recapture models? PLoSone 7(4):1-8.

Stephens, P.A., N. Pettorelli, J. Barlow, M.J. Whittingham, and M.W. Cadotte. 2015. Management by proxy? The use of indices in applied ecology. Journal of Applied Ecology 52:1-6. doi: 10.1111/1365-2664.12383

Sun, C.C., A.K. Fuller, M.P. Hare, and J.E. Hurst. 2017. Evaluating population expansion of black bears using spatial capture-recapture. Journal of Wildlife Management 81(5): 814-823.

Toonan, R.J., and S. Hughes. 2001. Increased throughpout for fragment analysis on ABI prism 377 automated sequencer using a membrane comb and STRand software. Biotechniques 31:1320-1324.

Van Dyke, T.S. 1902. The deer and elk of the Pacific coast. Pp. 167-256 in The Deer Family, by T.S. Roosevelt, T.S. Van Dyke, D.G. Elliott, and A.J. Stone. New York: MacMillan. 334 pp.

Waits, L.P., G. Luikart, and P. Taberlet. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. Molecular Ecology 10(1):249-256.

Web Soil Survey. 2013. Natural Resources Conservation Service, United States Department of Agriculture. Retrieved from: https://websoilsurvey.sc.egov.usda.gov/

Williams, B.K., J.D. Nichols, and M.J. Conroy. 2002. Analysis and management of animal populations. Academic Press. San Diego, CA, USA.

Williams, C.L., B. Lundrigan, and O.E. Rhodes, Jr. 2004. Microsatellite DNA variation in tule elk. Journal of Wildlife Management 68:109-119.

Wilson, P.J., and B.N. White. 1998. Sex identification of elk (Cervus elaphus canadensis), moose (Alces alces), and white-tailed deer (Odocoileus virginianus) using the polymerase chain reaction. Journal of Foresnic Science 43:477-482.

Section II Figures

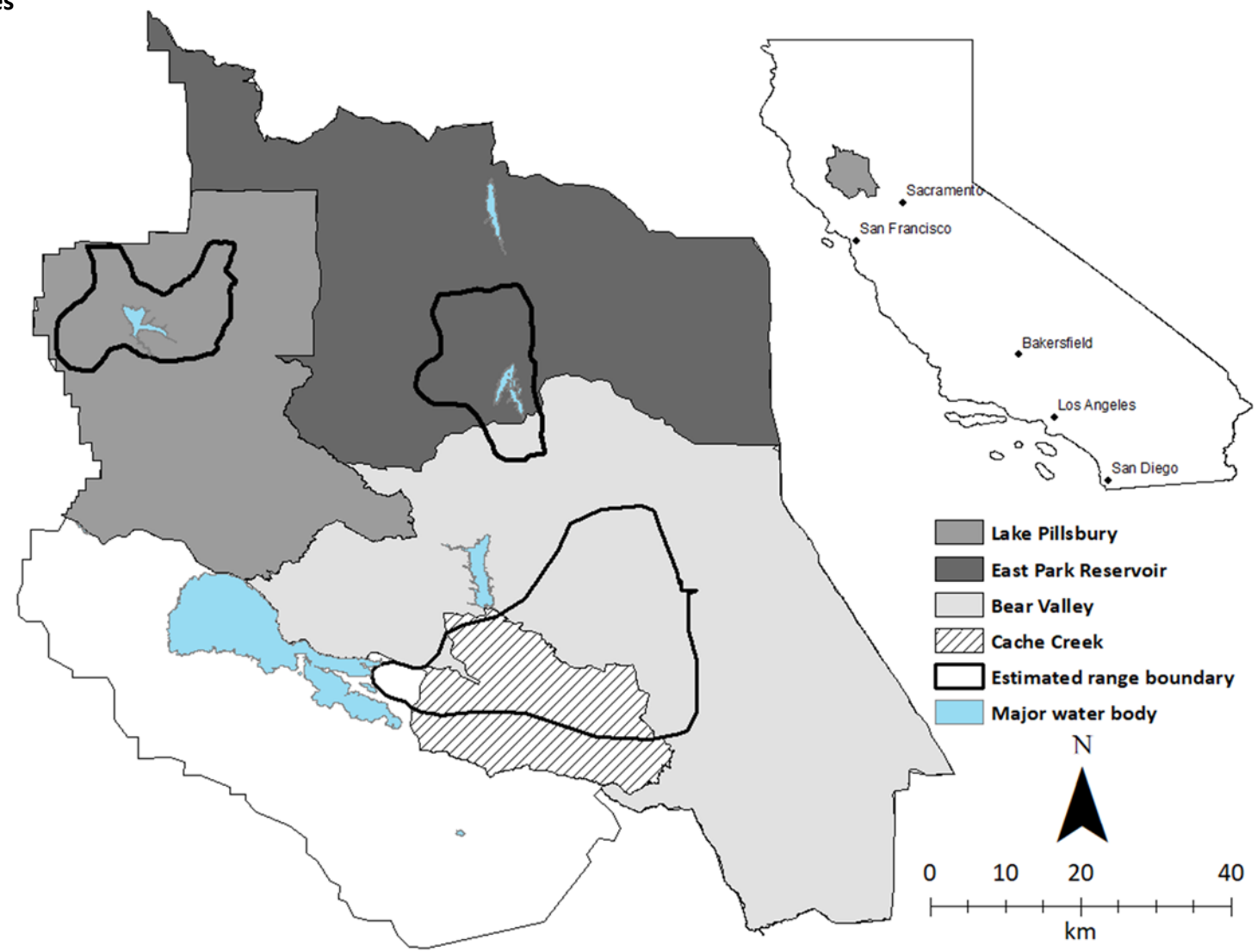


Fig. 2. Study area in the California Coast and Interior Coast mountain ranges encompassing four management units (MUs) from CDFW (2018) plus the entirety of Lake County. Estimated range boundaries (clockwise from left) for the Lake Pillsbury, East Park Reservoir, and Cache Creek populations are outlined in black and overlaid atop the MUs. Major water bodies are indicated by blue polygons.

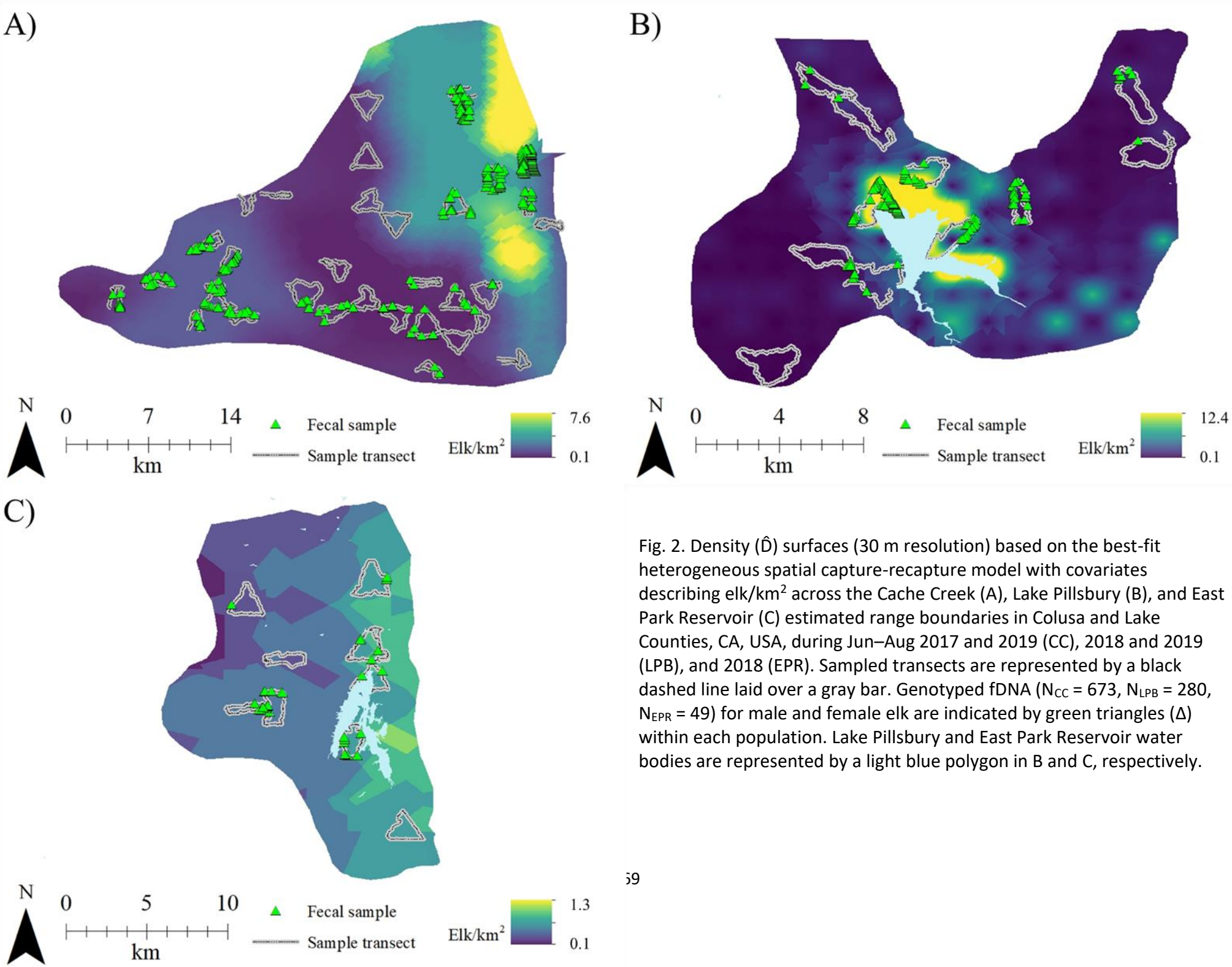


Fig. 2. Density ($\hat{D}$) surfaces (30 m resolution) based on the best-fit heterogeneous spatial capture-recapture model with covariates describing elk/km^2 across the Cache Creek (A), Lake Pillsbury (B), and East Park Reservoir (C) estimated range boundaries in Colusa and Lake Counties, CA, USA, during Jun–Aug 2017 and 2019 (CC), 2018 and 2019 (LPB), and 2018 (EPR). Sampled transects are represented by a black dashed line laid over a gray bar. Genotyped fDNA (N_{CC} = 673, N_{LPB} = 280, N_{EPR} = 49) for male and female elk are indicated by green triangles (Δ) within each population. Lake Pillsbury and East Park Reservoir water bodies are represented by a light blue polygon in B and C, respectively.

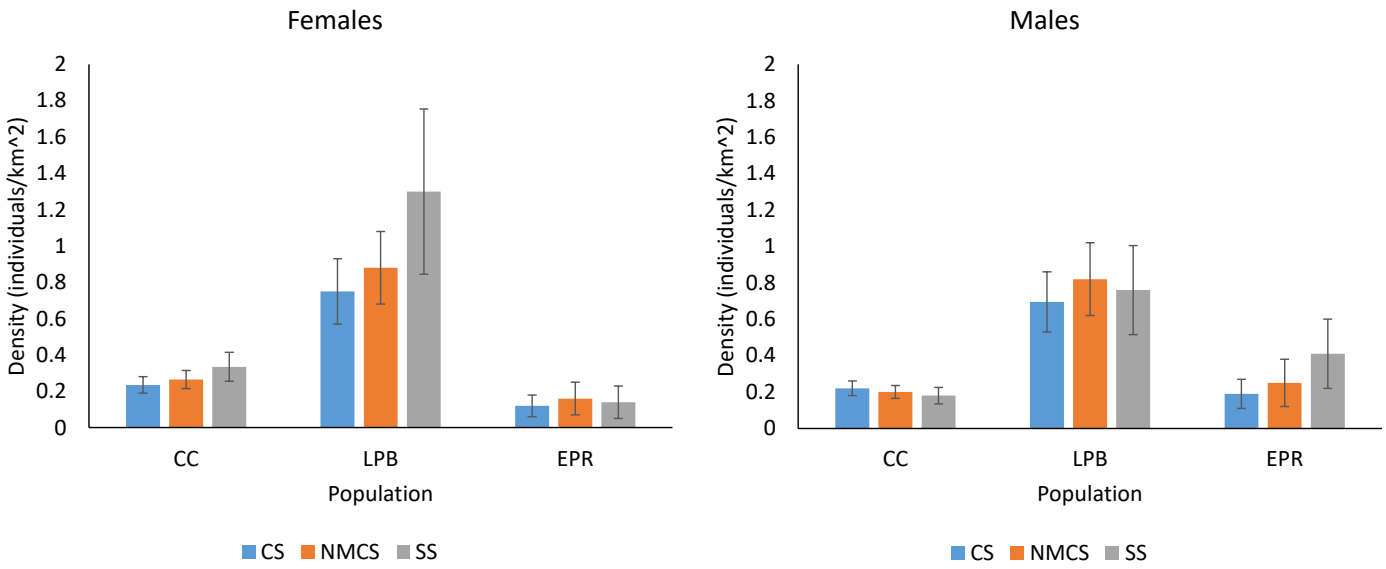


Fig. 3. Comparison of spatially explicit capture-recapture (SCR) density estimates (+/- 95% CI) based on 1,002 fecal DNA genotypes for female and male tule elk derived from combined-sex top models (CS), combined sex null models (NMCS), and single-sex best models (SS) for three populations in Lake and Colusa Counties, CA, 2017–2019: Cache Creek (CC), Lake Pillsbury (LPB), and East Park Reservoir (EPR), averaged between both years' estimates (CC, LPB). EPR was only sampled in 2018.

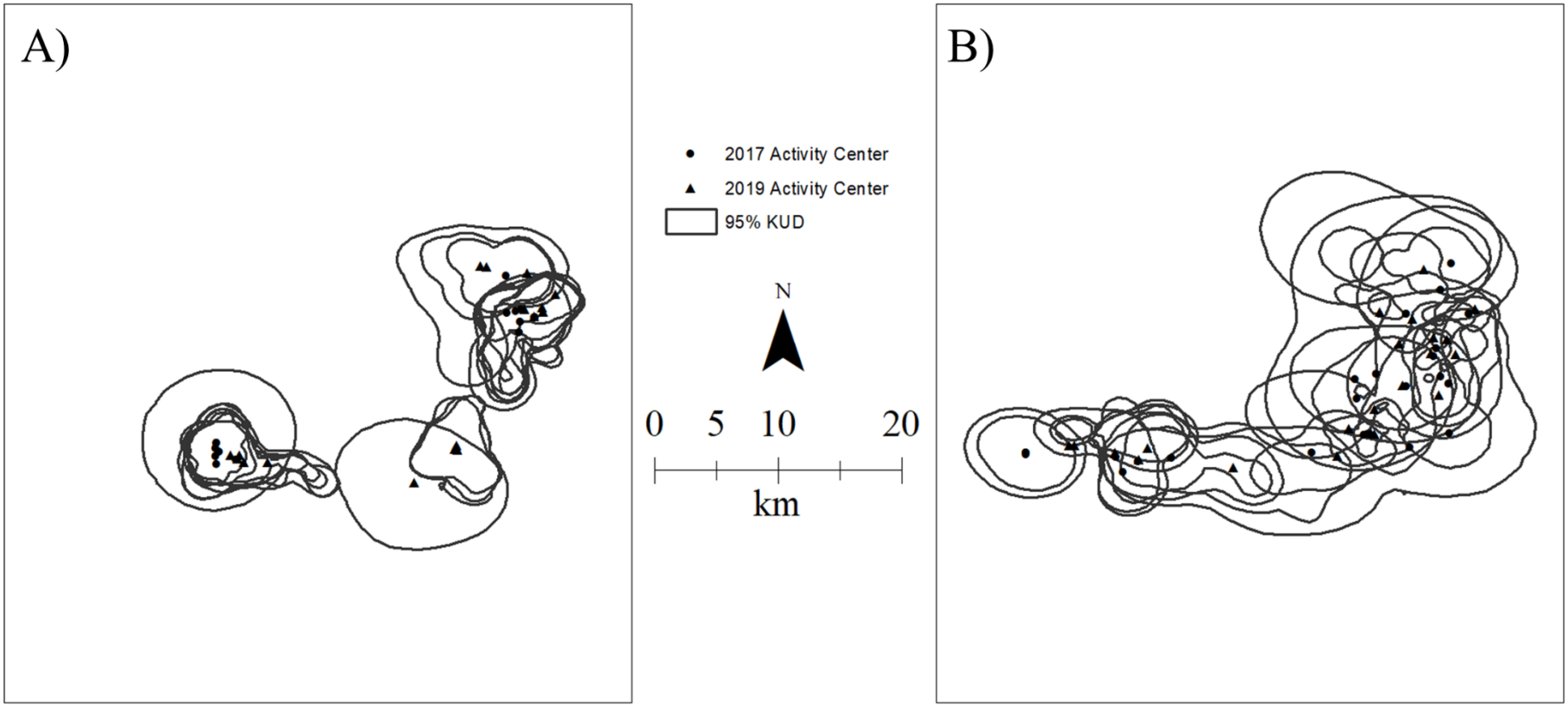


Fig. 4. 95% Kernel Utilization Density (KUD) and activity centers (AC) and of 22 female (A) and 23 male (B) tule elk at Cache Creek, Colusa and Lake Counties, CA. ACs are displayed for each sample season while 95% KUDs are averaged across sample seasons (Jun–Aug , 2017 and 2019).

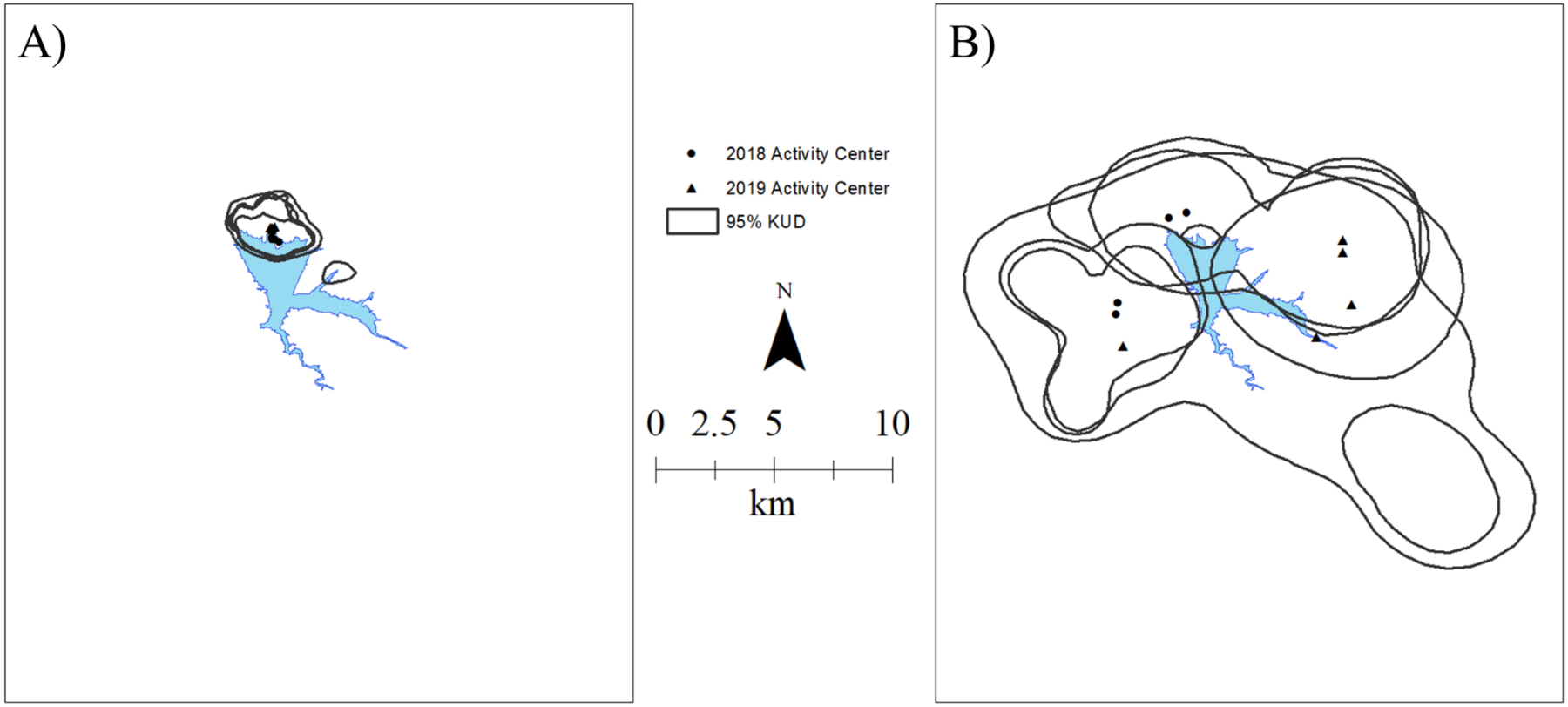


Fig. 5. 95% Kernel Utilization Density (KUD) and activity centers (AC) and of X female (A) and 5 male (B) tule elk at Lake Pillsbury, Lake County, CA. ACs are displayed for each sample season while 95% KUDs are averaged across sample seasons (Jun–Aug , 2018 and 2019). Lake Pillsbury is indicated by a blue polygon.

Section II Tables

Table 1. Numbers of elk pellet samples collected, samples successfully genotyped at ≥ 19 loci, capture history, individuals identified, individual males, individual females, and total recaptures used in spatially explicit capture-recapture (SCR) analyses during summers 2017-19 in Colusa and Lake Counties, California, USA.

	Year	Samples	Genotypes	Capture History[a]	Total Individuals	M	F	Recaptures[b]
Cache Creek	2017	513	358	300	161	72	89	139
	2019	497	315	273	157[c]	65[c]	92[c]	116
	Total	1,010	673	573	265	118	147	255
Lake Pillsbury	2018	191	125	120	80	46	34	40
	2019	291	155	132	79[d]	31[d]	48[d]	53
	Total	482	280	252	128	62	66	93
East Park Res.	2018	124	49	46	32	20	12	14
	Total	124	49	46	32	20	12	14
Colusa/Lake Counties	Total	1,616	1,002	871	425[e]	200[e]	225[e]	363

[a]Capture history is the total detections used for secr input data. Proximity detectors only allow one detection of an individual per detector.

[b]Recaptures are the total number of samples representing redetections at > 1 detector (i.e. not number of individuals recaptured), this number excludes eliminated samples that violate this rule.

[c]Total individuals include 38 (14 M, 24 F) of the individuals initially sampled in 2017.

[d]Total individuals include 24 (10 M, 14 F) of the individuals initially sampled in 2018.

[e]Total individuals does not include recaptured individuals. In other words, this number does not include double counted individuals, but rather reflects the realized total number of individuals detected through genotyping.

Table 2. Genetic summary statistics for tule elk from fecal pellet samples gathered across three herd ranges in Colusa and Lake Counties, California, USA, from June to August 2017-19. For each microsatellite locus, and across all loci, the table displays the average number of alleles observed per locus (N_A), observed heterozygosity (H_O), expected heterozygosity (H_E), the inbreeding coefficient (F_{IS}), probability of identity (P_{ID}), and probability of sibship (P_{SIB}).

	N_A	H_O	H_E	F_{IS}	P_{ID}	P_{SIB}
T193	3	0.388	0.451	-0.062	0.387	0.622
T26	5	0.402	0.581	0.063	0.258	0.524
TE132	2	0.138	0.329	0.045	0.504	0.712
TE145	3	0.341	0.476	-0.034	0.340	0.597
TE159	2	0.242	0.277	0.070	0.561	0.752
TE167	3	0.153	0.184	0.085	0.679	0.828
TE179	2	0.251	0.292	0.066	0.543	0.740
TE182	3	0.465	0.596	-0.046	0.246	0.514
TE185	7	0.317	0.376	0.003	0.408	0.664
TE45	5	0.263	0.284	-0.066	0.531	0.741
TE84	4	0.348	0.481	0.139	0.380	0.605
TE85	3	0.360	0.463	0.196	0.395	0.617
T108	5	0.560	0.561	-0.276	0.233	0.528
T172	3	0.410	0.525	0.032	0.328	0.570
T501	3	0.277	0.387	-0.002	0.410	0.659
TE105	2	0.229	0.278	0.038	0.560	0.751
TE169	2	0.297	0.332	0.037	0.501	0.709
TE68	3	0.177	0.193	-0.040	0.661	0.819
TE83	3	0.334	0.390	-0.010	0.432	0.663
TE88	2	0.313	0.328	0.010	0.506	0.713
All loci[a]	3.3	0.313	0.389	0.012	3.59E-08	2.45E-04

[a]We averaged N_A, H_O, H_E, and F_{IS} across loci. We multiplied the P_{ID} and P_{SIB} values across loci.

Table 3. Stepwise model rankings for spatially explicit capture-capture (SCR) analysis for density and abundance estimation of tule elk in Colusa and Lake Counties, California, USA. We used Akaike's Information Criterion (AIC) to rank the models in each step. We used the optimal detection function (detectfn) identified in step 1 for each population for the entirety of each population's respective modelling process. In step 2 we compared null model with the effect of the mixing proportion (pmix) on the observation model. In step 3 we modelled heterogeneous density with variation influenced by three covariates: mean diurnal temperature range ("bio2" in the WorldClim dataset; Hijmans et al. 2005), habitat classification, and soil classification.

Pop'n		Model	detectfn	K^a	logLik	AIC	ΔAIC	$AIC_w{}^b$
	Step 1							
		D~1 g0~1 sigma~1	exponential	3	-2786.79	5579.576	0	1
		D~1 g0~1 sigma~1	halfnormal	3	-2849.5	5705.009	125.432	0
	Step 2							
		D~1 g0~1 sigma~1 pmix~h2	exponential	4	-3004.16	6016.31	0	0.6306
		D~1 g0~h2 sigma~h2 pmix~h2	exponential	6	-3002.69	6017.38	1.07	0.3694
	Step 3							
Cache Creek		D~bio2 + habitat + soil g0~h2 sigma~h2 pmix~h2	exponential	24	-2843.23	5734.469	0	0.9869
		D~habitat + soil g0~h2 sigma~h2 pmix~h2	exponential	23	-2848.56	5743.119	8.65	0.0131
		D~soil g0~h2 sigma~h2 pmix~h2	exponential	21	-2861.15	5764.305	29.836	0
		D~bio2 + soil g0~h2 sigma~h2 pmix~h2	exponential	22	-2873.87	5791.74	57.271	0
		D~bio2 + habitat g0~h2 sigma~h2 pmix~h2	exponential	9	-2971.88	5961.767	227.298	0
		D~habitat g0~h2 sigma~h2 pmix~h2	exponential	8	-2977.85	5971.69	237.221	0
		D~1 g0~h2 sigma~h2 pmix~h2	exponential	6	-3002.69	6017.38	282.911	0
		D~bio2 g0~h2 sigma~h2 pmix~h2	exponential	7	-3002.81	6019.627	285.158	0
	Step 1							
		D~1 g0~1 sigma~1	exponential	3	-1185.59	2377.179	0	1
		D~1 g0~1 sigma~1	halfnormal	3	-1192.5	2390.997	13.818	0
	Step 2							
Lake Pillsbury		D~1 g0~1 sigma~1 pmix~h2	exponential	4	-1295.72	2599.443	0	0.6909
		D~1 g0~h2 sigma~h2 pmix~h2	exponential	6	-1294.53	2601.052	1.609	0.3091
	Step 3							
		D~habitat + soil g0~h2 sigma~h2 pmix~h2	exponential	13	-1156.6	2339.2	0	0.9721
		D~bio2 + habitat + soil g0~h2 sigma~h2 pmix~h2	exponential	14	-1159.15	2346.302	7.102	0.0279

		D~bio2 + soil g0~h2 sigma~h2 pmix~h2	exponential	13	-1178.49	2382.987	43.787	0
		D~soil g0~h2 sigma~h2 pmix~h2	exponential	12	-1186.15	2396.304	57.104	0
		D~bio2 + habitat g0~h2 sigma~h2 pmix~h2	exponential	8	-1225.12	2466.244	127.044	0
		D~habitat g0~h2 sigma~h2 pmix~h2	exponential	7	-1240.28	2494.567	155.367	0
		D~bio2 g0~h2 sigma~h2 pmix~h2	exponential	7	-1280.73	2575.467	236.267	0
		D~1 g0~h2 sigma~h2 pmix~h2	exponential	6	-1294.53	2601.052	261.852	0
	Step 1							
		D~1 g0~1 sigma~1	exponential	3	-239.893	485.786	0	0.7785
		D~1 g0~1 sigma~1	halfnormal	3	-241.15	488.3	2.514	0.2215
	Step 2							
		D~1 g0~h2 sigma~h2 pmix~h2	exponential	6	-258.915	529.831	0	0.5368
		D~1 g0~1 sigma~1 pmix~h2	exponential	4	-261.063	530.126	0.295	0.4632
	Step 3							
East Park Reservoir		D~habitat g0~h2 sigma~h2 pmix~h2	exponential	8	-246.773	509.547	0	0.7211
		D~bio2 + habitat g0~h2 sigma~h2 pmix~h2	exponential	9	-246.777	511.555	2.008	0.2642
		D~habitat + soil g0~h2 sigma~h2 pmix~h2	exponential	19	-240.076	518.152	8.605	0.0098
		D~bio2 + habitat + soil g0~h2 sigma~h2 pmix~h2	exponential	20	-239.768	519.535	9.988	0.0049
		D~soil g0~h2 sigma~h2 pmix~h2	exponential	17	-245.199	524.399	14.852	0
		D~bio2 + soil g0~h2 sigma~h2 pmix~h2	exponential	18	-245.271	526.542	16.995	0
		D~bio2 g0~h2 sigma~h2 pmix~h2	exponential	7	-256.702	527.403	17.856	0
		D~1 g0~h2 sigma~h2 pmix~h2	exponential	6	-258.909	529.817	20.27	0

[a]K is the number of parameters the model estimated.

[b]We used the AIC weights (AICw) to model average results for models with ΔAIC ≤2 in step 3.

Table 4. Parameter estimates from homogeneous density and best-fit heterogeneous density secr hybrid mixture (hcov) models for tule elk populations (pop'n) in Colusa and Lake Counties, CA. Results are displayed for both sexes combined, with estimates for females and males derived from the mixture proportion (pmix) of the 2 sexes, for each population. Parameters include density (D̂), abundance (N̂), probability of detection (g0), scale of movement (σ), as well as pmix, and relative standard error of density estimates (RSE(D̂)), a measure of the precision of a fitted secr model. The standard error (SE) and 95% confidence intervals are shown. Elk density is elk/km2 and sigma is in meters.

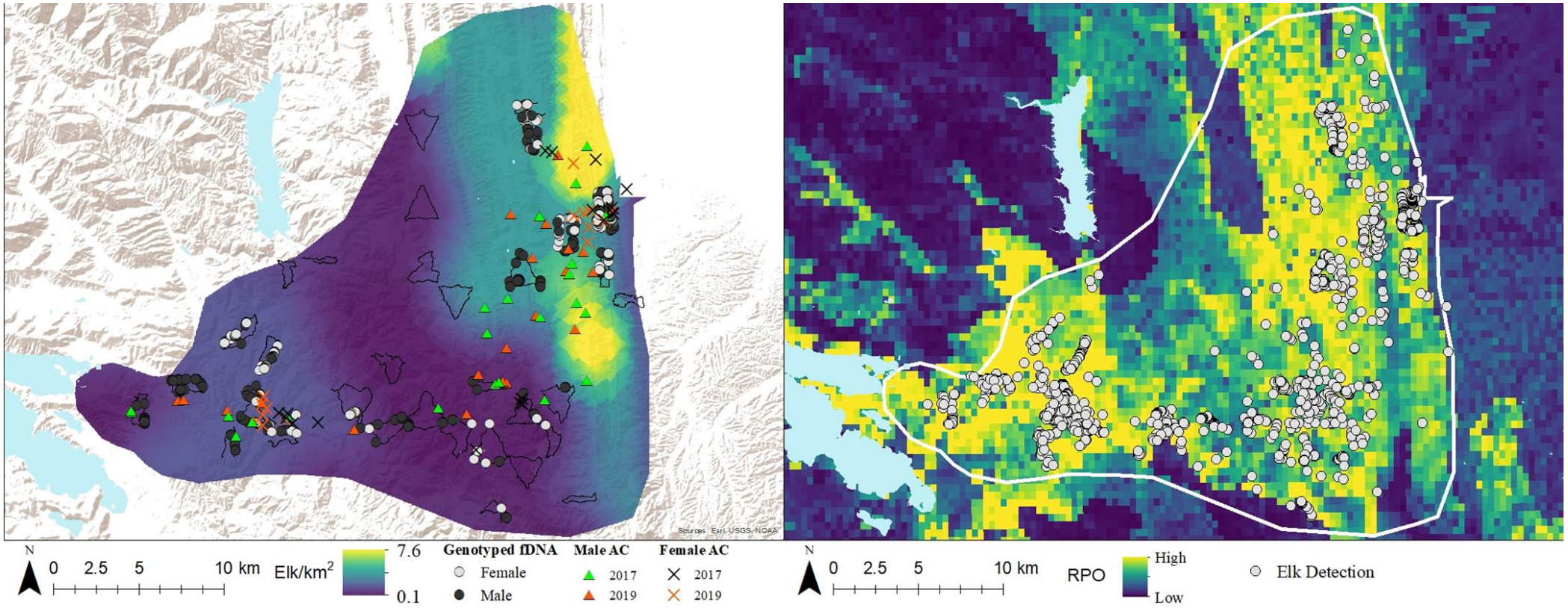
Pop'n	Year	Model	D̂	N̂	*g0*	σ	pmix	RSE(D̂)
Cache Creek 706 km^2		D~1						
	2017	Combined	0.47 ± 0.05 (0.39-0.57)	332 ± 35 (275-402)	-	-	-	0.10
		Female	0.27 ± 0.03 (0.22-0.33)	190 ± 20 (155-233)	0.06 ± 0.01 (0.04-0.07)	728 ± 51 (635-836)	0.57	-
		Male	0.20 ± 0.02 (0.17-0.25)	142 ± 15 (120-177)	0.05 ± 0.01 (0.04-0.07)	857 ± 62 (744-987)	0.43	-
	2019	Combined	0.46 ± 0.04 (0.38-0.56)	325 ± 28 (268-395)	-	-	-	0.10
		Female	0.26 ± 0.02 (0.21-0.32)	184 ± 14 (148-226)	0.06 ± 0.01 (0.04-0.07)	728 ± 51 (635-836)	0.57	-
		Male	0.20 ± 0.02 (0.16-0.24)	141 ± 14 (113-169)	0.05 ± 0.01 (0.04-0.07)	857 ± 62 (744-987)	0.43	-
		D~bio2+habitat +soil						
	2017	Combined	0.46 ± 0.05 (0.37-0.56)	325 ± 35 (261-395)	-	-	-	0.10
		Female	0.24 ± 0.03 (0.19-0.29)	169 ± 20 (134-205)	0.07 ± 0.01 (0.06-0.09)	688 ± 58 (584-811)	0.51	-
		Male	0.22 ± 0.02 (0.18-0.27)	155 ± 15 (127-191)	0.06 ± 0.01 (0.05-0.09)	861 ± 59 (753-985)	0.49	-
	2019	Combined	0.45 ± 0.05 (0.37-0.56)	318 ± 35 (261-395)	-	-	-	0.11
		Female	0.23 ± 0.03 (0.19-0.29)	162 ± 20 (134-205)	0.07 ± 0.01 (0.06-0.09)	688 ± 58 (584-811)	0.51	-
		Male	0.22 ± 0.02 (0.18-0.27)	155 ± 15 (127-191)	0.06 ± 0.01 (0.05-0.09)	861 ± 59 (753-985)	0.49	-
Lake Pillsbury 189 km^2		D~1						
	2018	Combined	1.7 ± 0.23 (1.3-2.2)	321 ± 43 (246-416)	-	-	-	0.14
		Female	0.88 ± 0.12 (0.68-1.1)	166 ± 23 (129-208)	0.06 ± 0.01 (0.04-0.09)	389 ± 42 (315-480)	0.52	-
		Male	0.82 ± 0.11 (0.62-1.1)	155 ± 21 (117-208)	0.05 ± 0.01 (0.03-0.08)	484 ± 60 (380-616)	0.48	-
	2019	Combined	1.7 ± 0.24 (1.3-2.2)	321 ± 45 (246-416)	-	-	-	0.14
		Female	0.88 ± 0.13 (0.68-1.1)	166 ± 25 (129-208)	0.06 ± 0.01 (0.04-0.09)	389 ± 42 (315-480)	0.52	-
		Male	0.82 ± 0.12 (0.62-1.1)	155 ± 23 (117-208)	0.05 ± 0.01 (0.03-0.08)	484 ± 60 (380-616)	0.48	-
		D~habitat + soil						

		Combined	1.4 ± 0.18 (1.1-1.8)	265 ± 34 (208-340)	-	-	-	0.13
	2018	Female	0.73 ± 0.09 (0.57-0.94)	138 ± 17 (108-178)	0.06 ± 0.02 (0.04-0.10)	417 ± 34 (355-490)	0.52	-
		Male	0.67 ± 0.09 (0.53-0.86)	127 ± 17 (100-163)	0.03 ± 0.01 (0.02-0.05)	669 ± 72 (543-825)	0.48	-
		Combined	1.5 ± 0.19 (1.1-1.9)	284 ± 36 (208-359)	-	-	-	0.13
	2019	Female	0.78 ± 0.10 (0.57-0.99)	147 ± 19 (108-187)	0.06 ± 0.02 (0.04-0.10)	417 ± 34 (355-490)	0.52	-
		Male	0.72 ± 0.09 (0.53-0.91)	136 ± 17 (100-172)	0.03 ± 0.01 (0.02-0.05)	669 ± 72 (543-825)	0.48	-
East Park Res. 200 km^2	D~1							
		Combined	0.41 ± 0.16 (0.19-0.86)	82 ± 32 (38-172)	-	-	-	0.4
	2018	Female	0.16 ± 0.06 (0.07-0.33)	32 ± 12 (14-66)	0.005 ± 0.006 (0.001-0.05)	2940 ± 4209 (371-23290)	0.38	-
		Male	0.25 ± 0.09 (0.12-0.53)	50 ± 18 (24-106)	0.02 ± 0.01 (0.01-0.05)	759 ± 211 (445-1297)	0.62	-
	D~habitat							
		Combined	0.31 ± 0.09 (0.17-0.55)	62 ± 18 (34-110)	-	-	-	0.3
	2018	Female	0.12 ± 0.03 (0.06-0.21)	24 ± 6 (12-42)	0.002 ± 0.001 (0.0003-0.006)	8970 ± 7852 (2045-39339)	0.38	-
		Male	0.19 ± 0.06 (0.11-.34)	38 ± 12 (22-68)	0.02 ± 0.01 (0.01-0.06)	876 ± 201 (561-1366)	0.62	-

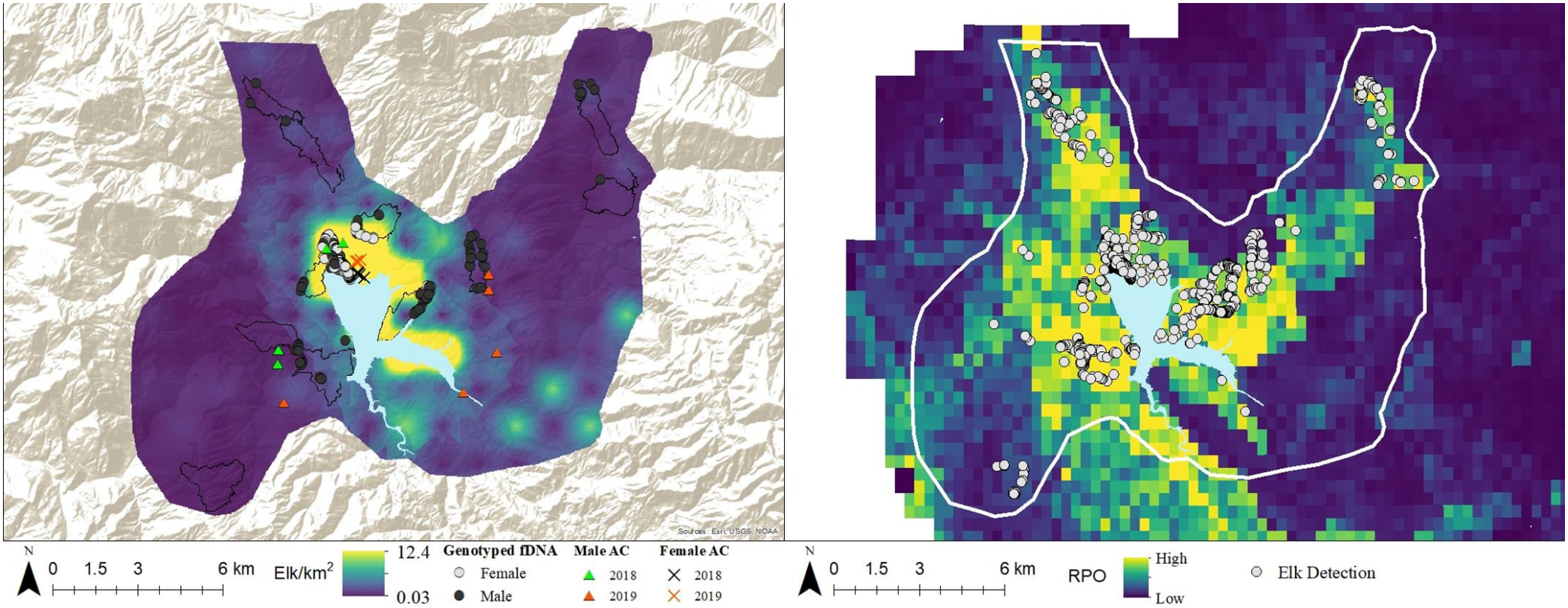
Table 5. Spatial clustering data of activity centers of GPS collared tule elk in Colusa and Lake Counties, CA, including for males (M) and females (F) number of elk (n), average nearest-neighbor distance (in meters) expected if activity centers are randomly (Poisson) distributed on the landscape (Exp) and observed average nearest-neighbor distances (Obs), for 3 populations sampled in one or two years in Lake and Colusa County, California, during 2017–2019.

Population	Year	n-M	Exp-M (m)	Obs-M (m)	n-F	Exp-F (m)	Obs-F (m)
Cache Creek	2017	23	1877	1134	22	1287	607
Cache Creek	2019	15	1704	1486	18	1827	778
Lake Pillsbury	2018	6	309	337	5	25	77
Lake Pillsbury	2019	5	1139	2218	4	26	65
East Park Reservoir	2018	3	87	52	7	690	568

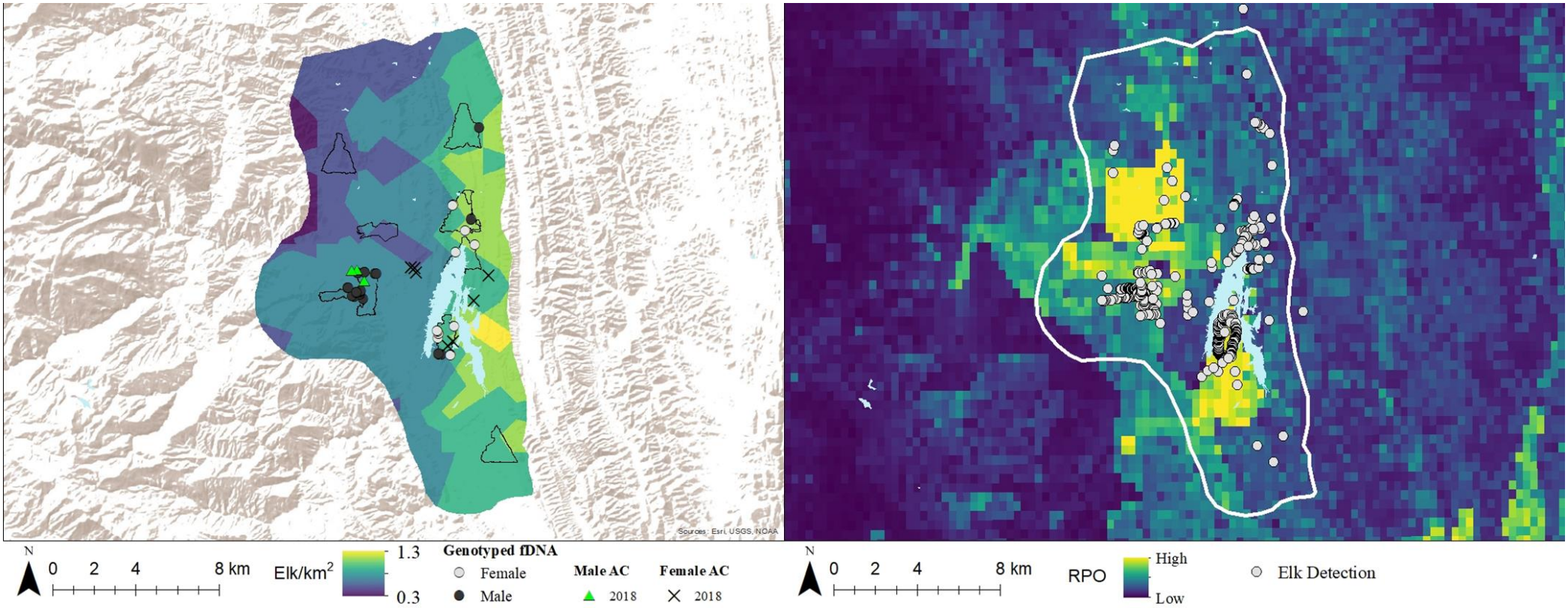
Section II Supplementary Figures



Supplementary Figure 1. Density ($\hat{D}$) surface (30 m resolution) (left) and predictive habitat model (right) across the Cache Creek estimated range boundary in Colusa County, CA, USA. The $\hat{D}$ surface is based on the best-fit spatial capture-recapture model describing elk/km^2 across during summers of 2017 and 2019. Warmer colors indicated higher levels of predicted density. Sampled transects are outlined in black. Genotyped fDNA for male and female elk are indicated by gray and black circles, respectively. The activity centers (AC) for males (Δ) and females (X) are color coded by year. The predictive habitat model surface shows the relative probability of occurrence (RPO) of tule elk in Colusa and Lake Counties, CA. Elk detections used to produce the model are displayed as gray circles. Warmer surface colors indicate areas predicted to have more suitable environmental conditions (higher RPO). Putative range boundary is outlined in white.



Supplementary Figure 2. Density ($\hat{D}$) surface (30 m resolution) (left) and predictive habitat model (right) across the Lake Pillsbury estimated range boundary in Lake County, CA, USA. The $\hat{D}$ surface is based on the best-fit spatial capture-recapture model describing elk/km^2 during summers of 2018 and 2019. Warmer colors indicated higher levels of predicted density. Sampled transects are outlined in black. Genotyped fDNA for male and female elk are indicated by gray and black circles, respectively. The activity centers (AC) for males (Δ) and females (X) are color coded by year. The predictive habitat model surface shows the relative probability of occurrence (RPO) of tule elk in Colusa and Lake Counties, CA. Elk detections used to produce the model are displayed as gray circles. Warmer surface colors indicate areas predicted to have more suitable environmental conditions (higher RPO). Putative range boundary is outlined in white.



Supplementary Figure 3. Density ($\hat{D}$) surface (30 m resolution) (left) and predictive habitat model (right) across the East Park Reservoir estimated range boundary in Colusa County, CA, USA. The $\hat{D}$ surface is based on averaged spatial capture-recapture models <2 ΔAIC describing elk/km^2 across, during summer 2018. Warmer colors indicated higher levels of predicted density. Sampled transects are outlined in black. Genotyped fDNA for male and female elk are indicated by gray and black circles, respectively. The activity centers (AC) for males (Δ) and females (X) are color coded by year. The predictive habitat model surface shows the relative probability of occurrence (RPO) of tule elk in Colusa and Lake Counties, CA. Elk detections used to produce the model are displayed as gray circles. Warmer surface colors indicate areas predicted to have more suitable environmental conditions (higher RPO). Putative range boundary is outlined in white.

Section II Supplementary Tables

Supplementary Table 1. Stepwise selection process for secr models for estimating density and abundance of tule elk in Colusa and Lake Counties, CA.

Step Number	Model Number	Model Description	Detection Function	Additional Parameters
1	1	D~1 g0~1 σ~1	halfnormal	-
	2	D~1 g0~1 σ~1	exponential	-
2	3	D~1 g0~1 σ~1 pmix~h2	step 1 df	pmix
	4	D~1 g0~h2 σ~h2 pmix~h2	step 1 df	pmix
3	5	D~bio2 g0~h2 σ~h2 pmix~h2	step 1 df	pmix
	6	D~habitat g0~h2 σ~h2 pmix~h2	step 1 df	pmix
	7	D~soil g0~h2 σ~h2 pmix~h2	step 1 df	pmix
	8	D~bio2+habitat g0~h2 σ~h2 pmix~h2	step 1 df	pmix
	9	D~bio2+soil g0~h2 σ~h2 pmix~h2	step 1 df	pmix
	10	D~habitat+soil g0~h2 σ~h2 pmix~h2	step 1 df	pmix
	11	D~bio2+habitat+soil g0~h2 σ~h2 pmix~h2	step 1 df	pmix

Supplementary Table 2. Classification of habitat types based on the relative probability of occurrence (RPO) as determined by a predictive habitat suitability model (Section I) for *secr* modeling of tule elk density and abundance in Colusa and Lake Counties, CA. We divided habitat types into prime, moderate, and poor classifications for *secr* modeling.

Habitat Classification	Habitat Type
Prime	Relative probability of occurrence (RPO) > 0.095
	Perennial Grassland
	Hardwood-Conifer
	Montane Riparian
	Lacustrine
	Hardwood
	Blue Oak-Valley Oak Dominant
Moderate	RPO: 0.046-0.079
	Annual Grassland
	Klamath Conifer
	Desert Riparian
	Mixed Chaparral
	Irrigated Agriculture
	Cropland
	Orchard
	Rice
	Barren
	Conifer Forest
	Coastal Oak Woodland
	Montane Chaparral
	Chamise-Redshank Chaparral
	Sagebrush
	Coastal Scrub
	Riverine
	Fresh Emergent Wetland
	Wet Meadow
	Valley Foothill Riparian
	Pasture
Poor	RPO < 0.046
	Urban

Supplementary Table 3. Taxonomic subgroup classification of soils for secr modeling of tule elk density and abundance in Colusa and Lake Counties, CA.

Soil Taxonomic Subgroup
Aeric Fluvaquents
Aridic Haploxererts
Cumulic Haploxerolls
Lithic Argixerolls
Lithic Dystroxerepts
Lithic Haploxerepts
Lithic Haploxerolls
Lithic Xerorthents
Mollic Haploxeralfs
Mollic Palexeralfs
Mollic Xerofluvents
Pachic Argixerolls
Sodic Endoaquerts
Typic Argixerolls
Typic Dystroxerepts
Typic Haplohemists
Typic Haploxeralfs
Typic Haploxerepts
Typic Haploxererts
Typic Palexeralfs
Typic Vitrixerands
Ultic Haploxeralfs
Ultic Haploxerolls
Ultic Palexeralfs
Xeric Endoaquerts

Supplementary Table 4. Genetic summary statistics for tule elk from fecal pellet samples gathered from Cache Creek in Colusa and Lake Counties, California, USA from June to August 2017-19. For each microsatellite locus, and across all loci, the table displays the number of individuals per population sample genotyped per locus (N), the total number of alleles observed per population sample per locus (A), the percentage of total alleles observed across population samples per population sample per locus (%), allelic richness per locus (A_r), observed heterozygosity (H_O), expected heterozygosity (H_E), deviation from Hardy-Weinberg Equilibrium (HWE), the inbreeding coefficient (F_{IS}), probability of identity (P_{ID}), and probability of sibship (P_{SIB}).

	Cache Creek									
	N	A	%	A_r	H_O	H_E	HWE	F_{IS}	P_{ID}	P_{SIB}
T193	260	3	100	2.45	0.496	0.492	0.818	-0.008	3.7E-01	6.0E-01
T26	251	5	100	4.08	0.530	0.593	0.001***	0.106	2.3E-01	5.1E-01
TE132	261	2	100	1.47	0.019	0.019	0.876	-0.010	9.6E-01	9.8E-01
TE145	261	3	100	2.98	0.238	0.232	0.723	-0.024	6.0E-01	7.8E-01
TE159	260	2	100	2	0.215	0.266	0.002**	0.189	5.7E-01	7.6E-01
TE167	259	3	100	2.75	0.193	0.197	0.757	0.019	6.6E-01	8.2E-01
TE179	261	2	100	2	0.257	0.330	0.000***	0.223	5.0E-01	7.1E-01
TE182	261	3	100	3	0.502	0.569	0.002**	0.117	2.7E-01	5.3E-01
TE185	253	6	85.71	4.73	0.257	0.333	0.000***	0.229	4.6E-01	7.0E-01
TE45	261	5	100	3.29	0.130	0.128	0.100	-0.018	7.6E-01	8.8E-01
TE84	261	4	100	2.23	0.360	0.435	0.107	0.172	4.1E-01	6.4E-01
TE85	261	2	66.67	2	0.418	0.486	0.023*	0.140	3.8E-01	6.0E-01
T108	261	4	80	3.4	0.483	0.481	0.585	-0.003	3.3E-01	5.9E-01
T172	258	3	100	2.96	0.496	0.538	0.286	0.078	3.1E-01	5.6E-01
T501	261	3	100	2.81	0.180	0.211	0.006**	0.146	6.4E-01	8.0E-01
TE105	261	2	100	1.99	0.138	0.128	0.231	-0.074	7.7E-01	8.8E-01
TE169	261	2	100	2	0.188	0.206	0.144	0.090	6.5E-01	8.1E-01
TE68	261	3	100	2.78	0.230	0.237	0.854	0.031	6.0E-01	7.8E-01
TE83	261	3	100	2.46	0.230	0.240	0.663	0.043	6.0E-01	7.8E-01
TE88	261	2	100	2	0.291	0.337	0.028*	0.136	5.0E-01	7.1E-01
All loci[a]	259.75	62	96.62	2.67	0.293	0.323	0.512	0.079	8.2E-07	1.1E-03

[a]We summed A, and averaged N, %, Ar, HO, HE, HWE, and FIS across loci; We multiplied the PID and PSIB values across all loci.
*p<0.05, **p<0.01, ***p<0.001

Supplementary Table 5. Genetic summary statistics for tule elk from fecal pellet samples gathered from Lake Pillsbury in Lake County, California, USA from June to August 2018-19. For each microsatellite locus, and across all loci, the table displays the number of individuals per population sample genotyped per locus (N), the total number of alleles observed per population sample per locus (A), the percentage of total alleles observed across population samples per population sample per locus (%), allelic richness per locus (A_r), observed heterozygosity (H_O), expected heterozygosity (H_E), deviation from Hardy-Weinberg Equilibrium (HWE), the inbreeding coefficient (F_{IS}), probability of identity (P_{ID}), and probability of sibship (P_{SIB}).

	Lake Pillsbury									
	N	A	%	A_r	H_O	H_E	HWE	F_{IS}	P_{ID}	P_{SIB}
T193	126	2	66.67	1.96	0.095	0.091	0.997	-0.050	0.831	0.912
T26	125	4	80	2.63	0.216	0.220	0.591	0.018	0.628	0.797
TE132	126	2	100	2	0.421	0.442	0.402	0.048	0.409	0.631
TE145	126	2	66.67	2	0.524	0.487	0.820	-0.075	0.382	0.602
TE159	126	2	100	2	0.357	0.350	N/A	-0.020	0.484	0.696
TE167	126	1	33.33	1	0.000	0.000	0.147	N/A	1.000	1.000
TE179	126	2	100	2	0.302	0.267	0.601	-0.129	0.573	0.760
TE182	126	2	66.67	2	0.357	0.341	0.836	-0.047	0.492	0.702
TE185	125	4	57.14	3.02	0.424	0.421	0.070	-0.007	0.403	0.640
TE45	126	3	60	3	0.540	0.516	0.635	-0.045	0.300	0.567
TE84	126	3	75	2.4	0.413	0.462	0.002	0.107	0.389	0.616
TE85	126	2	66.67	2	0.325	0.447	0.409**	0.272	0.406	0.628
T108	125	4	80	3.56	0.648	0.579	0.890	-0.120	0.257	0.525
T172	125	2	66.67	2	0.304	0.300	0.059	-0.012	0.535	0.734
T501	126	2	66.67	2	0.365	0.439	0.343	0.168	0.411	0.633
TE105	126	2	100	2	0.389	0.359	0.331	-0.084	0.476	0.690
TE169	125	2	100	2	0.520	0.478	N/A	-0.087	0.387	0.607
TE68	126	1	33.33	1	0.000	0.000	0.574	N/A	1.000	1.000
TE83	126	2	66.67	2	0.468	0.493	0.081	0.050	0.379	0.598
TE88	126	2	100	2	0.373	0.323	N/A	-0.155	0.511	0.716
All loci[a]	125.75	46	74.27	2.13	0.352	0.351	0.461	-0.009	4.3E-07	6.3E-04

[a]We summed A, and averaged N, %, Ar, HO, HE, HWE, and FIS across loci; We multiplied the PID and PSIB values across all loci.
*p<0.05, **p<0.01, ***p<0.001

Supplementary Table 6. Genetic summary statistics for tule elk from fecal pellet samples gathered from East Park Reservoir in Colusa County, California, USA from June to August 2018. For each microsatellite locus, and across all loci, the table displays the number of individuals per population sample genotyped per locus (N), the total number of alleles observed per population sample per locus (A), the percentage of total alleles observed across population samples per population sample per locus (%), allelic richness per locus (A_r), observed heterozygosity (H_O), expected heterozygosity (H_E), deviation from Hardy-Weinberg Equilibrium (HWE), the inbreeding coefficient (F_{IS}), probability of identity (P_{ID}), and probability of sibship (P_{SIB}).

East Park Reservoir										
	N	A	%	A_r	H_O	H_E	HWE	F_{IS}	P_{ID}	P_{SIB}
T193	32	3	100	3	0.656	0.592	0.707	-0.108	0.236	0.513
T26	32	2	40	1.99	0.125	0.117	0.706	-0.067	0.786	0.888
TE132	32	1	50	1	N/A	N/A	N/A	N/A	1.000	1.000
TE145	32	3	100	2.88	0.469	0.471	0.952	0.004	0.365	0.606
TE159	32	1	50	1	N/A	N/A	N/A	N/A	1.000	1.000
TE167	32	2	66.67	2	0.438	0.492	0.530	0.111	0.379	0.599
TE179	32	1	50	1	N/A	N/A	N/A	N/A	1.000	1.000
TE182	32	3	100	2.95	0.594	0.479	0.477	-0.240	0.350	0.598
TE185	32	2	28.57	2	0.375	0.305	0.192	-0.231	0.530	0.730
TE45	32	2	40	2	0.250	0.219	0.419	-0.143	0.634	0.799
TE84	32	1	25	1	N/A	N/A	N/A	N/A	1.000	1.000
TE85	32	2	66.67	1.64	0.031	0.031	0.928	-0.016	0.940	0.970
T108	32	2	40	2	0.844	0.488	0.000***	-0.730	0.381	0.601
T172	32	2	66.67	1.99	0.125	0.117	0.706	-0.067	0.786	0.888
T501	32	3	100	3	0.719	0.611	0.375	-0.177	0.234	0.503
TE105	32	2	100	2	0.344	0.417	0.318	0.177	0.426	0.648
TE169	32	2	100	2	0.313	0.375	0.346	0.167	0.461	0.678
TE68	32	2	66.67	2	0.438	0.404	0.642	-0.082	0.437	0.657
TE83	32	3	100	3	0.656	0.608	0.734	-0.080	0.222	0.501
TE88	32	2	100	2	0.250	0.264	0.769	0.052	0.577	0.762
All loci[a]	32	41	69.51	2.02	0.414	0.374	0.587	-0.089	2.2E-06	1.7E-03

[a]We summed A, and averaged N, %, Ar, HO, HE, HWE, and FIS across loci; We multiplied the PID and PSIB values across all loci.
*p<0.05, **p<0.01, ***p<0.001

Supplementary Table 7. Stepwise model rankings for spatially explicit capture-capture (SCR) analysis for density and abundance estimation of tule elk in Colusa and Lake Counties, California, USA. We modeled each sex seperately for each population. We used Akaike's Information Criterion (AIC) to rank the models in each step. We used the optimal detection function (detectfn) identified in step 1 for each population for the entirety of each population's respective modelling process. In step 2 we compared null model with the effect of the mixing proportion (pmix) on the observation model. In step 3 we modelled heterogeneous density with variation influenced by three covariates: mean diurnal temperature range ("bio2" in the WorldClim dataset; Hijmans et al. 2005), habitat classification, and soil classification.

Pop'n		Female Only							Male Only						
		model	detectfn	K^a	logLik	AIC	ΔAIC	AIC_w^b	model	detectfn	K^a	logLik	AIC	ΔAIC	AIC_w^b
Cache Creek	Step 1														
		D~1 g0~1 sigma~1	exponential	3	-1643.97	3293.94	0	1	D~1 g0~1 sigma~1	exponential	3	-1349.1	2704.21	0	1
		D~1 g0~1 sigma~1	halfnormal	3	-1671.15	3348.297	54.356	0	D~1 g0~1 sigma~1	halfnormal	3	-1381.52	2769.042	64.832	0
	Step 2														
		D~bio2 + habitat + soil g0~1 sigma~1	exponential	21	-1519.82	3081.647	0	0.8189	D~bio2 + habitat + soil g0~1 sigma~1	exponential	21	-1295.18	2632.351	0	0.4315
		D~habitat + soil g0~1 sigma~1	exponential	20	-1522.33	3084.665	3.018	0.1811	D~bio2 + soil g0~1 sigma~1	exponential	19	-1297.32	2632.637	0.286	0.3744
		D~soil g0~1 sigma~1	exponential	18	-1531.51	3099.017	17.37	0	D~habitat + soil g0~1 sigma~1	exponential	20	-1296.97	2633.945	1.594	0.1945
		D~bio2 + soil g0~1 sigma~1	exponential	19	-1547.53	3133.058	51.411	0	D~soil g0~1 sigma~1	exponential	18	-1312.76	2661.513	29.162	0
		D~bio2 + habitat g0~1 sigma~1	exponential	6	-1603.03	3218.054	136.407	0	D~bio2 g0~1 sigma~1	exponential	4	-1337.18	2682.356	50.005	0

	D~habitat g0~1 sigma~1	exponential	5	-1615.81	3241.622	159.975	0	D~habitat g0~1 sigma~1	exponential	5	-1342.35	2694.693	62.342	0
	D~bio2 g0~1 sigma~1	exponential	4	-1637.19	3282.383	200.736	0	D~bio2 + habitat g0~1 sigma~1	exponential	6	-1343.36	2698.712	66.361	0
	D~1 g0~1 sigma~1	exponential	3	-1643.97	3293.94	212.293	0	D~1 g0~1 sigma~1	exponential	3	-1349.1	2704.21	71.859	0
Step 1															
	D~1 g0~1 sigma~1	exponential	3	-669.779	1345.558	0	1	D~1 g0~1 sigma~1	exponential	3	-617.094	1240.188	0	0.7823
	D~1 g0~1 sigma~1	halfnormal	3	-675.705	1357.41	11.851	0	D~1 g0~1 sigma~1	halfnormal	3	-618.372	1242.745	2.558	0.2177
Step 2															
	D~bio2 + habitat_class + soil g0~1 sigma~1	exponential	11	-493.552	1009.105	0	1	D~habitat + soil g0~1 sigma~1	exponential	10	-573.483	1166.967	0	0.8504
	D~soil g0~1 sigma~1	exponential	9	-514.481	1046.962	37.857	0	D~soil g0~1 sigma~1	exponential	9	-576.222	1170.443	3.476	0.1496
	D~bio2 + soil g0~1 sigma~1	exponential	10	-514.227	1048.455	39.35	0	D~bio2 + habitat g0~1 sigma~1	exponential	5	-585.595	1181.189	14.222	0
	D~bio2 + habitat_class g0~1 sigma~1	exponential	5	-519.76	1049.519	40.414	0	D~bio2 + soil g0~1 sigma~1	exponential	10	-580.771	1181.542	14.575	0
	D~habitat_class + soil g0~1 sigma~1	exponential	10	-516.621	1053.243	44.138	0	D~bio2 + habitat + soil g0~1 sigma~1	exponential	11	-580.234	1182.468	15.501	0
	D~bio2 g0~1 sigma~1	exponential	4	-542.888	1093.777	84.672	0	D~bio2 g0~1 sigma~1	exponential	4	-588.583	1185.167	18.2	0

Lake Pillsbury

East Park Reservoir

	D~habitat_class g0~1 sigma~1	exponential	4	-621.949	1251.899	242.794	0	D~habitat g0~1 sigma~1	exponential	4	-600.082	1208.163	41.196	0
	D~1 g0~1 sigma~1	exponential	3	-669.779	1345.558	336.453	0	D~1 g0~1 sigma~1	exponential	3	-617.094	1240.188	73.221	0
Step 1														
	D~1 g0~1 sigma~1	exponential	3	-97.777	201.555	0	0.5217	D~1 g0~1 sigma~1	halfnormal	3	-158.965	323.93	0	0.7917
	D~1 g0~1 sigma~1	halfnormal	3	-97.8643	201.729	0.174	0.4783	D~1 g0~1 sigma~1	exponential	3	-160.301	326.601	2.671	0.2083
Step 2														
	D~bio2 + soil g0~1 sigma~1	exponential	16	-82.6746	197.349	0	0.3395	D~habitat g0~1 sigma~1	halfnormal	5	-138.185	286.37	0	0.691
	D~bio2 + habitat g0~1 sigma~1	exponential	6	-92.7276	197.455	0.106	0.3219	D~bio2 + habitat g0~1 sigma~1	halfnormal	6	-138.129	288.258	1.888	0.2688
	D~habitat g0~1 sigma~1	exponential	5	-94.3598	198.72	1.371	0.171	D~habitat + soil g0~1 sigma~1	halfnormal	14	-132.205	292.41	6.04	0.0337
	D~habitat + soil g0~1 sigma~1	exponential	17	-83.1353	200.271	2.922	0.0788	D~bio2 + soil g0~1 sigma~1	halfnormal	13	-134.862	295.725	9.355	0.0064
	D~1 g0~1 sigma~1	exponential	3	-97.777	201.555	4.206	0.0414	D~bio2 + habitat + soil g0~1 sigma~1	halfnormal	15	-135.792	301.583	15.213	0
	D~bio2 g0~1 sigma~1	exponential	4	-97.2953	202.591	5.242	0.0247	D~bio2 g0~1 sigma~1	halfnormal	4	-154.275	316.551	30.181	0
	D~bio2 + habitat + soil g0~1 sigma~1	exponential	18	-83.3814	202.763	5.414	0.0227	D~1 g0~1 sigma~1	halfnormal	3	-158.965	323.93	37.56	0

D~soil g0~1 sigma~1	exponential	15	-97.777	225.555	28.206	0	D~soil g0~1 sigma~1	halfnormal	12	-158.965	341.93	55.56	0

[a]K is the number of parameters the model estimated.

[b]We used the AIC weights (AICw) to model average results for models with ΔAIC ≤2 in step 3

Supplementary Table 8. Parameter estimates from homogeneous density and best-fit heterogeneous density secr models for tule elk populations (pop'n) in Colusa and Lake Counties, CA. Results are displayed for both sexes modeled seperately for each population. Parameters include density (D̂), abundance (N̂), probability of detection (g0), scale of movement (σ), and relative standard error of density estimates (RSE(D̂)), a measure of the precision of a fitted secr model. The 95% confidence intervals are shown in parentheses. Elk density is elk/km2 and sigma is in meters.

Pop'n	Year	Model	D̂	N̂	*g0*	σ	RSE(D̂)
Cache Creek		D~bio2 + habitat + soil					-
	2017	Female	0.33 ± 0.05 (0.25-0.44)	233 ± 35 (177-311)	0.05 ± 0.01 (0.04-0.06)	702 ± 55 (602-819)	0.15
	2019		0.34 ± 0.05 (0.26-0.45)	240 ± 35 (184-317)			0.14
		Model avg: D~bio2+habitat+soil, D~bio2 + soil, D~habitat + soil					-
	2017	Male	0.19 ± 0.03 (0.14-0.24)	134 ± 21 (99-169)	0.06 ± 0.01 (0.04-0.08)	866 ± 65 (748-1,001)	0.14
	2019		0.17 ± 0.02 (0.13-0.22)	120 ± 14 (92-155)			0.14
Lake Pillsbury		D~bio2 + habitat + soil					-
	2018	Female	1.1 ± 0.22 (0.69-1.6)	208 ± 42 (130-302)	0.07 ± 0.01 (0.05-0.10)	291 ± 25 (245-344)	0.21
	2019		1.5 ± 0.30 (1.0-2.1)	283 ± 56 (189-397)			0.19
		D~habitat + soil					-
	2018	Male	0.90 ± 0.17 (0.63-1.3)	170 ± 32 (119-246)	0.04 ± 0.01 (0.03-0.06)	500 ± 54 (406-617)	0.19
	2019		0.61 ± 0.13 (0.40-0.92)	115 ± 25 (76-174)			0.21
East Park Res.		Model average: D~bio2 + soil, D~bio2 + habitat, D~habitat					-
	2018	Female	0.14 ± 0.07 (0.05-0.36)	28 ± 14 (10-72)	0.01 ± 0.008 (0.004-0.05)	1203 ± 300 (745-1948)	0.52
		Model average: D~habitat, D~bio2 + habitat					-
	2018	Male	0.41 ± 0.13 (0.22-0.76)	82 ± 26 (44-152)	0.02 ± 0.01 (0.007-0.03)	910 ± 102 (732-1133)	0.33

Appendix A – Fecal Pellet Sampling Guidelines for Elk

UC Davis – Mammalian Ecology and Conservation Unit – elk fecal pellet sampling guidelines

Tom Batter, Josh Bush, and Ben Sacks

Draft updated: 2018/02/21 by T. Batter

Overarching sample design scheme *(Fig. A1)*

- Grid cells overlaid within estimated herd range boundary
- Stratify grid cells according to Maxent model classification threshold into predicted presence and predicted absence
- Randomly select a portion of predicated presence grid cells and a portion of predicted absence grid cells to sample
- A grid cell selected for sampling will disqualify all bordering cells for sampling; the nearest the next sampled grid cell can be is 2km away (may have to make exceptions i.e. private land access)
- To qualify for sampling, a cell must contain half of the cell area within the estimated herd boundary (cells that are sliced by the approximate herd boundary and have >1/2 of the cell area outside the boundary line will be excluded)
- Once cells are selected for sampling, prepare a cell portfolio.
    - Include: cell, transect route, satellite image, topographic map, vegetation map
    - Electronically scout the area via Google Earth, ArcGIS, etc. to map out the expected route; determine which and to what extent travel routes must deviate from the planned route due to terrain, water bodies, vegetation type (specifically avoid dense chemise-type vegetation), land ownership, etc.

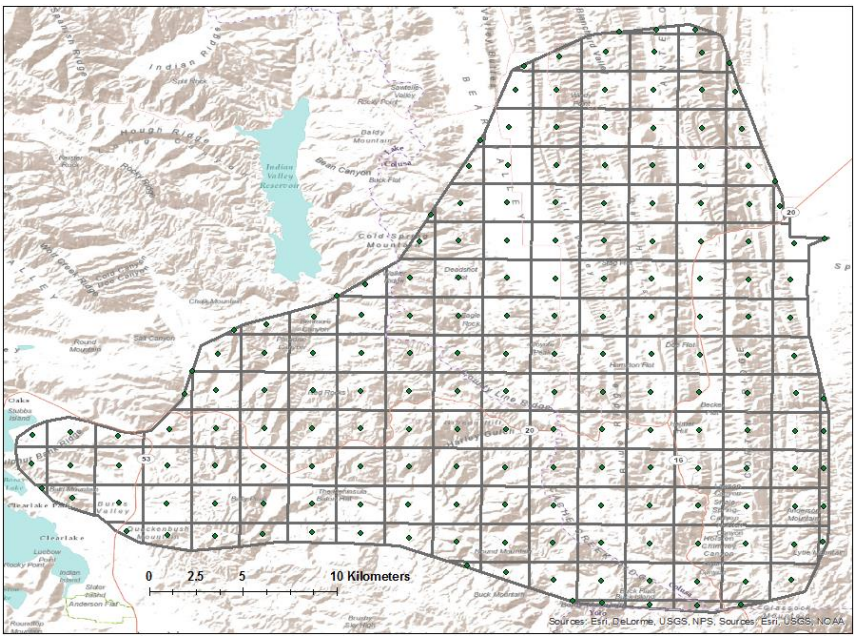


Figure A1 Cache Creek tule elk herd approximate boundary (cca. 2017) overlaid with 2km x 2km grid cells, each grid cell containing a centroid

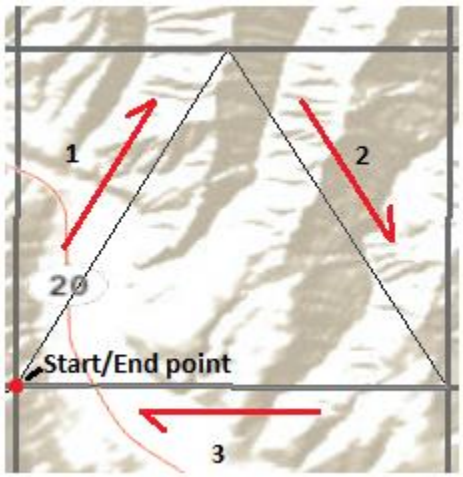


Figure A2 Close up of a triangle transect skeleton within a 2km x 2km grid cell; numbers indicate order of transect, arrows indicate direction of travel from the start point/towards the end point (indicated by the red point)

Sample scheme 1: Triangle transect – 2km segments (Fig. A2)

- Each cell contains a transect-based triangle survey skeleton
- The start/end point is to be located at one of the vertices, i.e. the southwest or southeast corner of the grid cell or the mid-point of the northern horizontal cell boundary line, depending on the point of access
- Each triangle skeleton is formed by three (3) segments, each ~2 km in length
    - Total distance covered: ~6km

Sampling scheme 2: Mirrored linear transect – 2km segments (Fig. A3)

- Each cell contains two linear transect survey skeletons
- The start point is to be located at the mid-point of the vertical cell boundary lines or the mid-point of the horizontal cell boundary lines, depending on the point of access
- Each mirrored skeleton is formed by two (2) segments, each ~2km in length, connected by a 500m segment
    - Total distance covered: ~4.25km
- A transect will run 2km in a North-South or East-West fashion; upon reaching 2km, the next 2km transect will be connected by traveling 500m to the east or west (for a N-S transect) or 500m to the north or south (for an E-W transect)
    - The direction to travel at the end of a segment will be determined by the sampler; the sampler will choose the direction they consider elk use to be more likely based on electronic scouting of the cell, and on-the-ground landscape observations and/or elk sign observed (see next section: "ideal search areas…")
- The end point should occur 500m from the start point along the same horizontal or vertical cell boundary

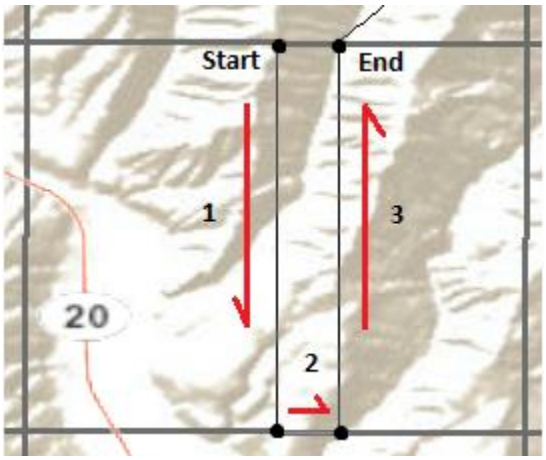


Figure A3 Close up of a mirrored linear transect skeleton within a 2km x 2km grid cell; numbers indicate order of transect, arrows indicate direction of travel; each black point indicates the start/end point of a segment with the transect start/end points labeled as such

<u>Sampling a transect</u>

- Use the GPS to navigate to the nearest point of the triangle depending on the point of access
    - GPS unit position format should be set to decimal degrees (hddd.ddddd°) and map datum to WGS 84
- Sampling a triangle skeleton
    - Name the starting point using the convention CellName_A
    - i.e. RockQuarryA; cells contain a Unique ID and will also be issued a name ahead of time based on some dominant feature contained within that cell
    - Triangle transect will have 3 points (1 point at each vertex); the second point and third point, when reached, can be saved using abbreviations. In this case, RQ_B… RQ_C
- Sampling a mirrored skeleton

    - Name the starting point and ending point of a segment using the convention CellName_Segment(A/B)#(1/2)
    - i.e. RockQuarry_SegmentA1; the second point, the end point of the first segment, will then be RQ_SA2
    - The first point of the second segment will be named RQ_SB1, and the end point RQ_SB2
    - Mirror transect will have 4 points (1 point at each start/end point for each segment)\
- From the starting point, travel towards the next point following the pre-determined triangle skeleton or pre-determined mirrored skeleton
- Search for pellets, elk tracks, or other elk sign along transect segment, as well as areas near, adjacent to, neighboring, close by, etc. that may attract elk use <u>without deviating >150 meters from the triangle skeleton on either side</u>
    - Other sign includes, but is not limited to, direct elk observation, hair (on barbed wire, power line poles, or other rubbing surface), antler scrapes, wallows, browse sign, antler sheds, bedding areas, etc. (see: Index I)
    - Document any elk sign observed on both the elk sign data sheet and log into the GPS
- Ideal search areas include the following:
    - Nearby watering holes, ponds, streams, creeks, etc.
    - Flat or gently sloping, grassy plains, especially flat areas contained within sloped terrain (i.e. the crest or peak of a slope)
    - Oak-savannah, especially later into the summer/fall when diet shifts to browse due to lack of palatable grasses and forbs
    - More obvious, relatively wide (compared to deer/hog) game trails
        - Elk don't typically travel single file such as cattle, so a "game trail" will occur over a wider area assuming a decent sized group travels the region; the exception are solitary bulls or small bands of elk which are less conspicuous
    - Bedding areas, usually in shady areas (i.e. under oak groves, high grass)
    - <u>Note:</u> thick, brushy areas (i.e. wooded areas with dense secondary canopy, or chaparral-dominant communities) are less likely to be used by elk; search these areas if/when obvious elk sign is observed within this type of habitat i.e. tracks, pellets, etc.
        - Collar data shows some bull use of rocky, chaparral-scrub dominant areas; the key difference is presumably lower vegetation density. Greater spacing allows for easy travel while still providing cover
- Save tracks once the transect has been completed using the convention: YYYYMMDD_CellName i.e. 20170503_RockQuarry
- Write total track distance traveled on the data sheet in the proper location

<u>Selecting and collecting pellets</u>

Here we define an <u>elk pellet group</u> ("pg") as:

> *An assemblage of 6 or more intact pellets judged by the collector to have been continuously voided, by the same individual animal, at the place where they are observed/to be collected*

<u>Freshest pellets in the given area will be collected</u> following classification guidelines adapted from the CDFW

NCR Deer abundance collection protocol based on the pellets' condition:

- ❖ *Eroded: Outer coating has been weathered, pellet is cracked. DO NOT COLLECT*
- ❖ *Dull: Whole but no longer shiny, may or may not be cracked. DO NOT COLLECT*
- ❖ *Malleable: Not wet or shiny, but contains enough moisture to be molded. COLLECT*
- ❖ *Shiny: Not wet, but still has a mucus coating or shiny dark color. COLLECT*
- ❖ *Wet: Some of the pellets (mostly inner) are wet, outer still dry. COLLECT*
- ❖ *Slimy: Every pellet is wet. COLLECT*
  - *Note: in very dry environments (i.e. Cache Creek, Colusa County, CA), a majority of the pellet samples encountered may be dull and dry. Collect the darkest pellets within the pellet group.*

- Elk pellets are significantly larger compared to deer pellets, in both size and in quantity, and may be clumped together or scattered in a concentrated or relatively wide area (see: Index II)
- If the pellets fall into a collectable category, scoop 4-6 pellets into a tube without touching them and seal the lid
  - Seek pellets toward the center of the group and towards the bottom or center of the pile (better protected from the elements)
  - Use a stick or sticks to separate clumped pellets (avoid breaking pellets, exposing vegetable material) and to move the pellets into the tube
  - For small pellets try to equal mass of 4-6 normal pellets, and denote this sample as likely to be "calf"
- If an exorbitant amount of pellet groups are encountered on the landscape, collect collector must use judgment to collect a subset of the freshest pellets available
  - Elk pellet groups generally occur in clusters; where there is one pellet group, more are likely to be near; therefore, upon detecting and collecting elk pellets, intensify the search effort over the immediate area
  - Collect samples from pellet groups that are more distant from one another, i.e. when encountering numerous pellet groups over an extended area, grid the area out to the best of your ability in your mind or on a piece of field notebook paper and attempt to collect pellets in a fashion that represents the whole area (i.e. not concentrated in one spot)
- If the only pellets that are encountered are older (i.e. eroded, dull) collect these pellets in the same fashion as described; document the quality of the sample on the data sheet and the sample tube
- Create a waypoint for each sample collected (waypoint ID can simply be the next autofilled number the GPS unit provides)
  - Write the corresponding waypoint on the data sheet, along with the GPS coordinates, the labelling convention, and the sample condition
- Label each sample tube as it becomes filled with a pellet group sample using the following convention:
  - GridCellNumber_PelletGroup#
  - i.e. If you are collecting within cell 2175 and collecting pellet group 17 you will write on the vial and lid: 2175 P17
- Upon completion of sampling, return samples to the lab, office, etc. and fill each vial containing pellets with 95-100% ethanol; be sure all pellets are COMPLETELY submerged in ethanol!
  - Pellets will absorb ethanol. Add a sufficient amount to account for absorption to still maintain complete submergence.
- Make sure vials are clearly labeled with pellet group number, plot number and date

- Transcribe pellets into master pellet spreadsheet
- Storage: Ethanol will readily remove the ink from the tubes. Place all vials from a particular transect in a single zip lock bag labeled with plot and date. Rubber band vials together to help keep them upright
- Put samples in a cooler out of direct sunlight, and out of contact with water or ice
- Store coolers inside vehicle while working transects or at night to minimize risk of bear damage

Data Transfer

- At the conclusion of each field week update Plot Completion Status spreadsheet on Region 2 office server to note statuses of all plots visited.

Data Type	Instructions
Transect Path	Create a folder labeled as the cell name/cell number (e.g., RockQuarry) and use DNRGPS or Basecamp software to download the transect path track as a txt file. Label this file as the concatenation of plot number and "track" (e.g., RockQuarrytrack).
Pellet Waypoints	Use DNRGPS or Basecamp software to download the transect points (observations and pellet locations) and save as a .csv file. Label this file as the concatenation of cell name, "pts," and date (e.g., RockQuarry_pts20170503). Place it in the same folder as above (e.g., RockQuarry)
Data Sheets	Create a folder labeled as the plot number (e.g., RockQuarry) and transfer scanned copies of data sheets to it. Label them as the concatenation of cell name and date (e.g., RockQuarry_20170503).

Appendix B – Elk Fecal Pellet Collection Datasheet

Cell Name/Number	Date	Sample Type (circle): Triangle / Mirror	Crew
Herd:	Start/End	Transect distance (km)	Page ______ of ______

Pellet ID & Waypoint ID	Condition	GPS Coordinates & Notes	Pellet ID & Waypoint ID	Condition	GPS Coordinates & Notes
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	
	[ ] slimy [ ] wet [ ] shiny [ ] other			[ ] slimy [ ] wet [ ] shiny [ ] other	

Appendix C – Photographic Examples of Elk Sign

Figure C1. Types of elk sign (clockwise left to right, top to bottom): direct observation, pellets, tracks (multiple), rub, track (single), antler scrape, hair, & carcass. Sign not pictured: hunter harvest, antler shed(s), elk bed(s), etc.

Appendix D – Protocol for Extracting Elk Pellets Using QIAGEN DNeasy 96 Blood & Tissue Kit

MATERIALS: P1200 and P200 multichannel pipettes, P1000 and P200 single channel pipettes and tips, 1000uL tips, Qiagen DNeasy 96 blood & tissue kit (proteinase K, DNeasy 96 (spin column) plate, two S-blocks, light blue collection plate and plastic caps, blue elution/permanent plate and rubber caps, Airpore tape sheets, and buffers ASL, AL, AW1, AW2, and AE), 100% EtOH, reagent reservoirs (troughs), plate map (printed out), and one autoclaved 50 ml centrifuge tube (and lid) for each sample.

SAFETY NOTES

Always be familiar with the locations of first aid kit, eye wash station, and chemical spill kit, and always wear proper PPE: latex gloves, lab coat, and eye protection.

DAY 1

The volumes listed below reflect the values for a single plate. If you are doing two plates, **use the number in parentheses/bold.**

1. Before starting:
    - **Make sure you have all of the reagents you need for both days.**
    - Turn on the large incubator to ~70°C.
    - In the **Intern Extraction Notebook**, write or print the lab/sample IDs for all samples to be extracted, giving each one a temporary extraction ID (i.e., 1-94 + 1 blank and 1 empty well per plate). Also include the date and initials of everyone working on the plate.
    - Label 95 autoclaved 50 ml centrifuge tubes with extraction IDs and use racks to keep the tubes upright.

2. Clean the lab bench (10% bleach and a Clorox wipe)

3. Place 1 pellet from each sample into its respective labeled centrifuge tube. Place the racks of 50 ml tubes (open, covered with a large Kimwipe) in the large incubator for 1.5+ hours to dry. The temperature is not super important; the only thing going on here is drying the EtOH. If you dry the pellets overnight, do so on the fecal extraction bench covered with a Kimwipe and proceed to day 2 the next morning.

DAY 2

4. Turn on the incubator to 56°C.

5. Add 1.5 ml Buffer ATL to each sample. Mix each one manually to ensure that the pellet is at least partially immersed in the buffer. If not, use a stick to push the pellet to the bottom of the tube or add more buffer. Also add more buffer to especially large and/or absorbent pellets.

6. Tape the racks on to the rocker and rock for ~30 minutes. Do not stack the racks. The shorter rocking time ensures that all pellets will have time on the rocker. After/before rocking, manually mix the tubes about every 10 minutes for an additional 30 minutes so that the pellets soak in buffer for at least 1 hour total.

7. Pipet 2 ml (**4 ml**) Proteinase K into a trough and multichannel 20 µl into each well of light blue collection plate. Also, label the blue elution/permanent plates you will need for later with plate ID, date, and extracting person(s) initials). Spin plates down if there is any proteinase K on the sides of the wells.

8. After rocking, add 350 µl of the lysate from each tube into its respective well on the collection plate using the P1000 single channel. If the pellets absorb too much ATL (as is occasionally the case), add 100-200 µl ATL and swirl around. Make sure you can pipette a *full* 350 µl aliquot because balance is crucial!

9. Using a large disposable pipette, add 34 ml (**68 ml**) Buffer AL (without EtOH) from the extraction kit to a 100 ml trough (NOT a 25 ml!) and multichannel out 350 µL into each well. Cover with plastic strip caps.

10. Shake/invert the plate vigorously for 10 seconds.

NOTE: DO NOT do this with just the plastic plate cover on, as there is some space between the tops of the caps and the lid and the terrible caps will pop up during this step. Instead, press your hand evenly and firmly on top of the caps to keep them from popping.

11. Spin plates down in the Sorvall (large plate spinner) at ~1000 rpm for 2 minutes.

12. Incubate at 56°C for ten minutes. **Caution: heat can cause the caps to pop off! Put something heavy on top or wrap an elastic band around the plastic plate cover.**

13. Remove plate(s) from incubator, **check to make sure the caps have not popped off (and re-cap the rows if they are)**, shake/invert as before, and centrifuge briefly (2 minutes) to get liquid off the cap. Remove and discard the caps.

14. Turn incubator to 70°C and put buffer AE in to warm it up.

15. Using a large disposable pipette, add 34 ml (**68 ml**) 100% EtOH from the extraction kit (or flammable cabinet) to a 100 ml trough (NOT a 25 ml) and multichannel out 350 µl into each well.

16. **Cover** with new caps and shake/invert for 10 seconds.

17. Centrifuge at 1000 rpm for 2 minutes to pellet any remaining debris and get liquid off caps.

18. Place a new DNeasy 96 plate onto an S-block. Using the P1200 multichannel, pipette 900 µl of the lysate mixture (avoiding any remaining solid particles) into the DNeasy 96 plate sitting on the S-block. To keep the S-block orientation straight, write an "x" or something on the side of the S-block underneath your first sample (so you know which well the A1 sample on the 96 plate goes into).

19. **Cover** with Airpore tape sheet and centrifuge the plate and S-block for **15 minutes at 5000 rpm**.

20. Using a large disposable pipette, add 48 ml (**96 ml**) Buffer AW1 to a trough and multichannel 500 µl to each well in the DNeasy 96 plate. Cover with an Airpore tape sheet, and centrifuge for **10 minutes at 5000 rpm**.

21. After centrifugation, put the DNeasy 96 (spin column) plate onto a clean S-block. Switching to a new block will allow us to spin the AW2 through without putting unnecessary stress on the machine due to heavy plates and also ensure that liquid won't come in contact with the bottom of the plate. During this step (or at the end of the extraction), dump the contents of the full S-block into the appropriate waste container.

22. Using a large disposable pipette, add 48 ml (**96 ml**) Buffer AW2 to a trough and multichannel 500 µl to each sample in the 96 plate and, **without covering it**, centrifuge for **24 minutes at 5000 rpm**. Put the orange lid on the rotor to avoid contamination or aerosolized stuff (since there is no Airpore sheet on). Lower the ramp up and down speeds to **4** to minimize the noise due to the lid.

23. Place the DNeasy 96 plate on a labeled blue elution/permanent plate. Double check that your wells line up (i.e., A1 on the DNeasy 96 plate is in A1 of the blue elution plate). Then add 5 ml (**10 ml**) Buffer AE (warmed to 70°C) to a trough and, using the P200 multichannel, multichannel 50 µl directly onto the filter in DNeasy 96 plate (switch tips between each row). Incubate at room temperature for **five minutes**.

24. After incubation, cover with an Airpore tape sheet and centrifuge the sample for **10 minutes at 5000 rpm**.

25. Cap the elution/permanent plate containing DNA sample and put into the refrigerator until ready for PCR.

26. Clean up the lab area around you. Check to make sure there are enough supplies for the next extraction – order another kit or supplies as necessary.