

From: [Doherty, Kevin](#)
To: [Guinotte, John](#); [Stephen Torbit](#)
Subject: Re: Effective Population Size FYI
Date: Tuesday, May 16, 2017 11:47:22 AM
Attachments: [Fedy.etal.Cons.Gen.2016.pdf](#)

FYI,

Here is an attached paper from Fedy (who I am working with on Sage-grouse genetics). I asked them about sample sizes and estimating effective population size and he thought it was possible, however, he cautioned that these kind of results are easily over interpreted. In short, they were able to calculate effective population sizes with relatively small numbers of samples for sage-grouse in Wyoming (See table 1 in the attached paper).

Cheers
Kevin

On Mon, May 15, 2017 at 12:25 PM, Guinotte, John <john_guinotte@fws.gov> wrote:
Need to talk to you about this today.

John Guinotte
Fish and Wildlife Biologist
Ecological Services
U.S. Fish and Wildlife Service
Mountain Prairie Region 6
134 Union Blvd., Lakewood, CO 80228
303-236-4264
john_guinotte@fws.gov

----- Forwarded message -----

From: **Grizzle, Betty** <betty_grizzle@fws.gov>
Date: Mon, May 15, 2017 at 11:49 AM
Subject: Fwd: Effective Population Size FYI
To: Stephen Torbit <stephen_torbit@fws.gov>, John Guinotte <john_guinotte@fws.gov>

Here was what I sent in March. I attached Koskela's Dissertation.

----- Forwarded message -----

From: **Grizzle, Betty** <betty_grizzle@fws.gov>
Date: Wed, Mar 8, 2017 at 11:23 AM
Subject: Effective Population Size FYI
To: "Shoemaker, Justin" <justin_shoemaker@fws.gov>, John Guinotte <john_guinotte@fws.gov>, Stephen Torbit <stephen_torbit@fws.gov>
Cc: Caitlin Snyder <caitlin_snyder@fws.gov>

In preparation for the call on Friday re genetics, I read through Anni Koskela's 2013 dissertation (Finland), who developed estimates of effective population size for two wolverine subpopulations in that country.

Background/context: In the April 2016 court ruling, the judge specifically ruled in favor of the Plaintiffs that the Service's "unlawfully ignored best available science" regarding threats posed to wolverine by small population size and lack of genetic diversity, and remanded this to the Service (for further consideration).

I am sharing with you the following excerpts so you can see the level of effort that this type of analysis can entail, and also to contrast Koskela's methods with the methods that were used in the reference we cited in the previous proposed rule to conclude that wolverine effective population sizes are very low (Feb 4, 2013; 78 FR 7884). Highlights/notes added.

From Schwartz et al. (2009):

(from Methods section)

Effective population size

Using only those samples from the main part of the Rocky Mountains (i.e., eliminating samples from the Crazy and Belt Mountains, as Cegelski et al. 2003 suggested they form a different subpopulation), we estimated the effective population size and 95% credible limits of the estimate via summary statistics and approximate Bayesian computation methods using ONeSAMP (Tallmon et al. 2008). To determine the temporal stability of the wolverine population, we subdivided the samples into three time frames and used ONeSAMP to estimate effective population size in each frame. [I'm not sure how many samples they used for this analysis; somewhere between 142 and 210]

(from Results section)

Effective population size, calculated using ONeSAMP....was 35 (credible limits = 28–52) and was found not to change over time (N_e 1989–1994 = 33, credible limits 27–43; N_e 1995–2000 = 35, credible limits 28–57; N_e 2001–2006 = 38, credible limits 33–59). [Note credible limits of N_e exceed 50 in two of their subdivided samples. They separated these samples into 3 time periods in order to determine "temporal stability" - would be interesting to see the N_e estimate if they were kept together]

From Koskela Dissertation (2013):

To study the population genetics of wolverines in Finland, 406 scat samples, ten muscle tissue samples and two oral mucosa swabs as sources of DNA were collected during 2004–2011, mainly in northern and eastern Finland.

From Methods (pp. 32–34): The program DIYABC 1.0.4.46 (Cornuet *et al.* 2010) was used in order to infer the time at which population clusters had split from an ancestral population, as well as to infer the effective population sizes (N_e) of the assumed populations. The program makes inferences about population divergences, admixtures and changes in population size using approximate Bayesian computation. The program BOTTLENECK 1.2.02 (Cornuet & Luikart 1996) was used to detect recent reductions in effective population size. The program detects an excess of heterozygotes, as would be expected under a scenario of recent population bottlenecks, by using different tests: a sign test, a standardized differences test and a Wilcoxon rank test. The effective population size (N_e) was estimated using LD N_e 1.31 (Waples & Do 2008). The program estimates contemporary effective population sizes using a biased correction for the Hill method (Hill 1981, Waples 2006)

based on linkage disequilibrium data. Three estimates of N_e are given based on the lowest allele frequency used (0.01, 0.02 and 0.05).

From Results: $N_e = 14-17$ in this study.

--

Betty J. Grizzle, D.Env.
Fish and Wildlife Biologist
U.S. Fish and Wildlife Service
Carlsbad Fish and Wildlife Office
2177 Salk Ave, Suite 250
Carlsbad, CA 92008
760-431-9440, ext. 215
760-431-5901 fax

--

Betty J. Grizzle, D.Env.
Fish and Wildlife Biologist
U.S. Fish and Wildlife Service
Carlsbad Fish and Wildlife Office
2177 Salk Ave, Suite 250
Carlsbad, CA 92008
760-431-9440, ext. 215
760-431-5901 fax

--

Kevin Doherty, PhD
Spatial Ecologist
USFWS Region 6 --Science Applications
134 Union Blvd, Lakewood, CO 80228
Phone: (303) 921-0524
Email: kevin_doherty@fws.gov

Integration of genetic and demographic data to assess population risk in a continuously distributed species

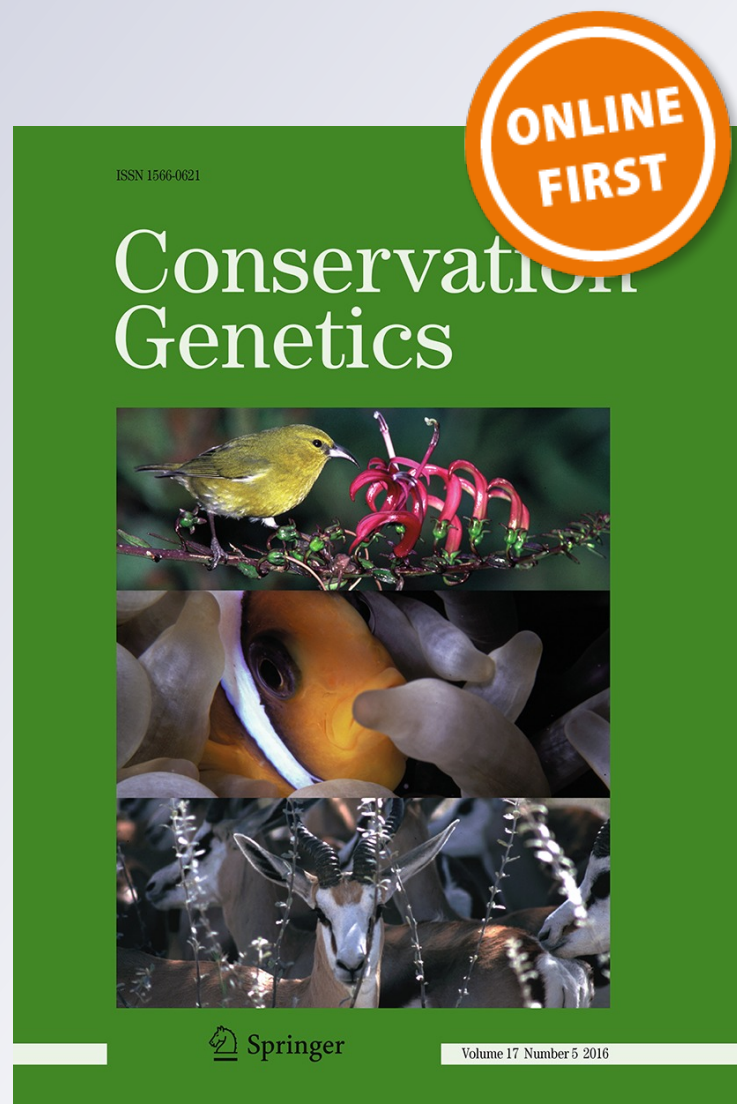
Bradley C. Fedy, Jeffrey R. Row & Sara J. Oyler-McCance

Conservation Genetics

ISSN 1566-0621

Conserv Genet

DOI 10.1007/s10592-016-0885-7



Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

RESEARCH ARTICLE

Integration of genetic and demographic data to assess population risk in a continuously distributed species

Bradley C. Fedy¹ · Jeffrey R. Row¹ · Sara J. Oyler-McCance²

Received: 31 March 2016 / Accepted: 8 September 2016
© Springer Science+Business Media Dordrecht 2016

Abstract The identification and demographic assessment of biologically meaningful populations is fundamental to species' ecology and management. Although genetic tools are used frequently to identify populations, studies often do not incorporate demographic data to understand their respective population trends. We used genetic data to define subpopulations in a continuously distributed species. We assessed demographic independence and variation in population trends across the distribution. Additionally, we identified potential barriers to gene flow among subpopulations. We sampled greater sage-grouse (*Centrocercus urophasianus*) leks from across their range ($\approx 175,000$ Km²) in Wyoming and amplified DNA at 14 microsatellite loci for 1761 samples. Subsequently, we assessed population structure in unrelated individuals ($n = 872$) by integrating results from multiple Bayesian clustering approaches and used the boundaries to inform our assessment of long-term population trends and lek activity over the period of 1995–2013. We identified four genetic clusters of which two northern ones showed demographic independence from the others. Trends in population size for the northwest subpopulation were statistically different from the other three genetic clusters and the northeast and southwest subpopulations demonstrated a general trend of

increasing proportion of inactive leks over time. Population change from 1996 to 2012 suggested population growth in the southern subpopulations and decline, or neutral, change in the northern subpopulations. We suggest that sage-grouse subpopulations in northern Wyoming are at greater risk of extirpation than the southern subpopulations due to smaller census and effective population sizes and higher variability within subpopulations. Our research is an example of incorporating genetic and demographic data and provides guidance on the identification of subpopulations of conservation concern.

Keywords Microsatellite · Management units · Population trends · Sage-grouse · *Centrocercus urophasianus*

Introduction

Understanding the current status of animal populations and how they change through time is key to adaptive management and can trigger important management actions, including legal protection (Walters 1986; Runge 2011). Local and regional population dynamics are influenced by the distribution of individuals on the landscape and thus, the extent of a study area and the associated population of interest can influence our conclusions and have profound management implications. Several authors have recommended that individuals should be grouped into biologically meaningful populations to avoid these potential biases related to population boundaries and for more effective species management (Waples and Gaggiotti 2006; Palsbøll et al. 2007; Lowe and Allendorf 2010). Yet defining demographically independent units within natural systems is often difficult. This is particularly challenging when individuals are distributed across continuous habitat

Electronic supplementary material The online version of this article (doi:10.1007/s10592-016-0885-7) contains supplementary material, which is available to authorized users.

✉ Bradley C. Fedy
bfedy@uwaterloo.ca

¹ School of Environment, Resources and Sustainability, University of Waterloo, Waterloo, ON N2L 3G1, Canada

² U.S. Geological Survey, Fort Collins Science Center, Fort Collins, CO 80526, USA

where gradients of change, and not hard boundaries, are likely the norm.

Population boundaries are often determined inductively, based on species biology and expert opinion (Pearce et al. 2001; Martin et al. 2012; Krueger et al. 2012). In other cases, populations are defined using demographic or genetic data (Moritz 1994; Allendorf and Luikart 2007; Palsbøll et al. 2007; Funk et al. 2012). Often the detailed demographic data required to accurately assess demographic independence are not available over large landscapes. Additionally, the demographics of some species (e.g., population trends, reproduction, survival) are highly correlated across space and time (e.g., Fedy and Doherty 2011) making the identification of boundaries based solely on demographic independence untenable. Yet, different regions across a species distribution may be demographically independent due to isolation-by-distance (IBD), movement barriers, or variation in habitat quality across the distribution.

When species are continuously distributed, genetic data can help define clusters or the thresholds beyond which IBD or other factors result in demographically independent groups of individuals. The identification of these cores can help guide demographic assessment. Often, in animal conservation, it is our goal to assess or monitor population responses to particular threats or pressures. These threats are seldom distributed continuously across a landscape. In populations that demonstrate IBD, it is useful to understand the patterns and deviations from IBD in order to assess the potential impacts of threats that are not continuously distributed. This requires the explicit consideration of the spatial autocorrelation that exists in allele frequencies in the presence of IBD (Meirmans 2012).

Specifically, boundaries or gradients of change can be defined through the use individually-based genetic clustering approaches (e.g., STRUCTURE, GENELAND, TESS) to assess whether a priori divisions are reflected in the genetic data (Höglund et al. 2011; Kolomyjec et al. 2012; Olsen et al. 2014) or to delineate populations without any a priori expectations (Oyler-McCance et al. 2005; Fedy et al. 2008; Row et al. 2010). These can be used to define the central core of subpopulations under patterns of IBD and guide the assessment of spatial demographic variation. However, three main challenges arise in certain situations. First, IBD must be addressed explicitly to determine its role is spatial clustering (Frantz et al. 2009). Second, the uncertainty in the posterior probabilities of population membership often is not addressed in a spatially-explicit manner. For example, it is unclear how to group and treat those individuals with uncertain population membership or transition zones which should be recognized, estimated, and predicted spatially. Third, the Bayesian clustering approaches often provide incongruent estimates of the

number of clusters (K) and the individual posterior probabilities of membership for each individual. Multiple studies have addressed the relative performance of these approaches under specific scenarios (Chen et al. 2007; Safner et al. 2011; Blair et al. 2012). However, when there is no clear ‘best approach’ that can be determined a priori we generally look for convergent answers. If outputs from different clustering approaches are incongruent, and in our experience they seldom agree perfectly, it is necessary to develop an approach to formally combine the output from multiple approaches into a single combined result rather than speak generally of agreement among approaches. The alternatives are to a priori choose a single preferred method and present only those results or—potentially more concerning—running all three and selecting the answer that best fits a researcher’s preferred outcome and succumbing to confirmation bias.

Greater sage-grouse (*Centrocercus urophasianus*; hereafter sage-grouse) are a gallinaceous species that occur in sagebrush (*Artemisia* spp) habitats throughout the North American West. Sage-grouse have a lek mating system in which males gather at central breeding locations in the spring and have been the subject of long-term monitoring. Until recently, sage-grouse were considered “warranted but precluded” from listing under the U.S. Endangered Species Act (ESA). However, a status review by the U.S. Fish and Wildlife Service in 2015 determined that protection for sage-grouse under the ESA was no longer warranted. The species seems to have recovered somewhat from historic lows in the mid-1990s; however, the species remains one of conservation concern throughout their range. Wyoming contains an estimated 37 % of the remaining range wide population (Doherty et al. 2010) and is expected to remain a stronghold for the species persistence (Knick and Hanser 2011). Previous range-wide genetic research on sage-grouse suggested three clusters within the state based on five sample locations with an average of 21 individuals per location (Oyler-McCance et al. 2005).

Here, we combine demographic and genetic data to estimate population trends, effective population sizes, IBD, demographic independence and population risk relative to other populations for spatially-explicit population clusters identified using genetic data. Our objectives were to define genetically distinct subpopulations and then identify which of these populations are more at risk due to lek extirpations, decreasing population trends, greater genetic isolation, or lower effective population sizes. Specifically, in order to identify genetically distinct subpopulations we integrated results from multiple Bayesian clustering algorithms and developed spatially-explicit predictions of population membership. Next, we used a spatially explicit individually-based approach to assess patterns of IBD and

identify potential barriers and corridors. Third, we identified regional subpopulation risk using the spatially-predicted clusters to group genetic and demographic data. We integrated these data to assess the effective subpopulation sizes and demographic independence of each cluster, the differences in subpopulation size and variability, and the differences in the percentage of active leks through time. Our approach provides an example of how passively collected genetic and demographic data can be combined to identify subpopulation clusters, their population trends and identify at-risk subpopulations while identifying potential barriers and corridors across large spatial scales in continuously distributed wildlife populations.

Methods

Study system

Wyoming is 253,300 km² and contains some of the largest remaining intact sagebrush steppe habitats in North America. Primary land uses include cattle grazing, tillage agriculture, and oil and gas energy production. The current sage-grouse distribution in Wyoming covers approximately 70 % of the state (174,580 km²).

Counts of males attending leks in the spring are one of the primary methods of monitoring sage-grouse populations. Despite some concerns regarding the accuracy of male lek counts (Walsh et al. 2004), the lek count data appear useful for assessing population change, particularly at broad scales (Connelly et al. 2004; Walker et al. 2007; Fedy and Aldridge 2011; Fedy and Doherty 2011). Wyoming has one of the longest running and extensive lek monitoring programs (Connelly et al. 2004).

Genetic data collection

Genetic data were collected using DNA isolated from either feather or blood samples. Most samples were extracted from feathers collected non-invasively at 88 leks during the spring counts (mean samples per lek = 9.3, standard error = 0.5). Feathers were collected and placed in a plastic bag (one feather per bag) and stored in a -20 °C freezer until processing. A few samples were blood collected from captured birds stored at room temperature on paper cards (Whatman FTA). As such, most samples were associated with a lek collected during the breeding season. A few (6 %) were from birds captured outside of the breeding season and thus not associated with any lek.

Our feather sampling approach was to gather many samples over a short time period and a large spatial extent. We collected feathers from leks across most of the

distribution of sage-grouse in Wyoming and selected sites from which to amplify DNA to maximize spatial coverage and evenness across the state. We did not collect samples from a small, isolated population of sage-grouse in the Jackson area because they are known to be genetically isolated from surrounding populations (Schulwitz et al. 2014). A quantitative sampling study was conducted after our analysis to identify priority sample locations across the entire sage-grouse distribution (Hanks et al. 2016). There was good congruence between the Wyoming sites identified through the Hanks et al. (2016) analysis and our approach.

DNA isolation, amplification, and post processing

DNA from 1761 samples was extracted using a DNeasy Blood & Tissue Kit (Qiagen) following the user-developed protocol for feather samples. Portions of the extraction process were automated using a QiaCube (Qiagen). Samples were amplified across 14 microsatellite loci developed for species of the subfamily Tetraoninae (MSP11, MSP18, SGMS06.6, SGMS06.8; Oyler-McCance and St John 2010, SGCA5, SGCA11, SGCTAT1; Taylor et al. 2003, TUT3; Segelbacher et al. 2000, TTT3; Caizergues et al. 2003, SG21, SG28, SG36, SG29, SG39; Fike et al. 2015). The primers for 2 loci were redesigned for better performance (SGCA5F: CGGACAGGTACATCCTGGAA, SGCA5R: GGGAAAAGATGTCAGAATCTACAAA, SGCA11F: G CAGTAAAGAAAATTTGGAAGCA, SGCA11R: TCTTG AACTGATGTTGGATTG). Amplifications were performed in 10 µL PCRs consisting of approximately 10 ng of template DNA, 0.2 mM of each dNTP, 0.25 µM dye-labeled forward primer, 0.25 µM reverse primer, 0.625 U GoTaq Flexi DNA polymerase (Promega), 2.25 mM MgCl₂ and 1X GoTaq Flexi Buffer (Promega). The amplification conditions for all loci were as follows: 94 °C for 2 min, then 94 °C for 1 min, annealing temperature (52 °C: MSP11, SGMS06.8; 55 °C: MSP18, SGCA5; 58 °C: TTT3; 60 °C: SGCA11, SGCTAT1, TUT3, SGMS06.6, SG21, SG28, SG36, SG29, SG39) for 1 min, and 72 °C for 1 min for 40 cycles, then 60 °C for 45 min and a final extension at 72 °C for 5 min. PCR products were multi-loaded based on product size and primer label, combined with GeneScan LIZ 600 internal lane size standard (Applied Biosystems), and electrophoresed through a capillary gel matrix using an AB3500 Automated DNA Sequencer (Applied Biosystems). Allele sizes were determined for each locus using GeneMapper v4.1 software (Applied Biosystems). After the first round of genotyping, poor quality samples (that amplified at <7 loci) and exact genotypic matches were removed prior to the identification of unique individuals (see below). The remaining 1056 samples were then genotyped at least one more time to

verify the genotype and identify instances of genotyping error and allelic dropout.

Most of our DNA samples came from noninvasive, passively collected tissues (e.g., feathers). Therefore, we did not know the number of unique individuals represented in the dataset prior to our analyses. Additionally, missing data occurred in some samples; however, since all samples were run twice, genotyping error was likely low. Thus, our first step, prior to assessing spatial genetic structure, was to identify all unique genotypes (i.e., individuals) in our dataset. We used ALLELEMATCH (Galpern et al. 2012) and DROPOUT (McKelvey and Schwartz 2005) to identify unique multilocus genotypes. Briefly, DROPOUT indicated 117 pairs of samples matching at 3 or fewer loci and these were flagged as potential duplicates. ALLELEMATCH, which accounts for missing data, suggested 27 of these pairs were not matches and these were further examined. Of the 27, nine were retained as different individuals because they had at least one allele with different data (not simply a heterozygote/homozygote mismatch) and 18 were classed as the same individuals because they were heterozygote/homozygote mismatches. The process is fully described in Row et al. (2015).

Spatial genetic structure

The number of clusters estimated using genetic clustering software programs (e.g., GENELAND, STRUCTURE) can be strongly influenced by the presence of family groups (Rodríguez-Ramilo and Wang 2012). This concern has been addressed in three ways. Some studies have argued that the presence of family groups is an important biological signature of interest to the research questions addressed and therefore, the researchers included family groups in their analyses (Woltmann et al. 2012). Other studies using passively sampled individuals (e.g. hair snares) have randomly sampled a portion of the samples to exclude prior to running the clustering algorithms (Schregel et al. 2012). Finally, as recommended by Rodríguez-Ramilo and Wang (2012), family groups can be identified and removed prior to clustering analysis. We were most interested in identifying coarse-scale patterns in genetic variability and clustering for our research, not individual family groups. Thus, to avoid inflation of the number of genetic clusters within Wyoming we followed the third approach presented above and identified family groups using COLONY (Jones and Wang 2010).

COLONY implements a maximum likelihood method to assign parentage and sibship among individuals using their multi-locus genotypes (Jones and Wang 2010). The software considers all individuals within a population in assigning relationships. COLONY requires a defined 'population' in order to identify highly related individuals.

We defined populations in COLONY as spatially proximate sample groups across Wyoming. We selected an 18 km radius for our definition of sample groups as this distance represents the 90th percentile for the maximum pairwise distance that female sage-grouse in Wyoming move between nesting and summer (19.0 km) and the mean distance moved between summer and winter locations (17.27 km; Fedy et al. 2012). Additionally, an 18 km buffer distance value parallels the connectivity assessment of Knick and Hanser (2011). We buffered all sample locations with an 18 km radius in ArcGIS Desktop v.10.0, SP3 (<http://www.esri.com>). We merged overlapping buffers, and the resulting boundaries formed our sample groups (Fig. 1). We ran COLONY on each sample group to identify full siblings and parent-offspring relationships and set the COLONY project parameters to match the biology of sage-grouse. In particular, we set the mating system parameters to polygamy, without inbreeding, and used the full-likelihood analysis method, with the complexity sibship prior, assigning no known relationships. When full-sib or parent-offspring relationships were detected we randomly selected one member of the family group to retain. This data set was used for all subsequent analyses.

To account for variation in the clustering algorithm we assessed population genetic structure using three genetic clustering software programs. First, we used STRUCTURE (Pritchard et al. 2000; Falush et al. 2003) and specified the admixture model, a separate α -inferred for each population

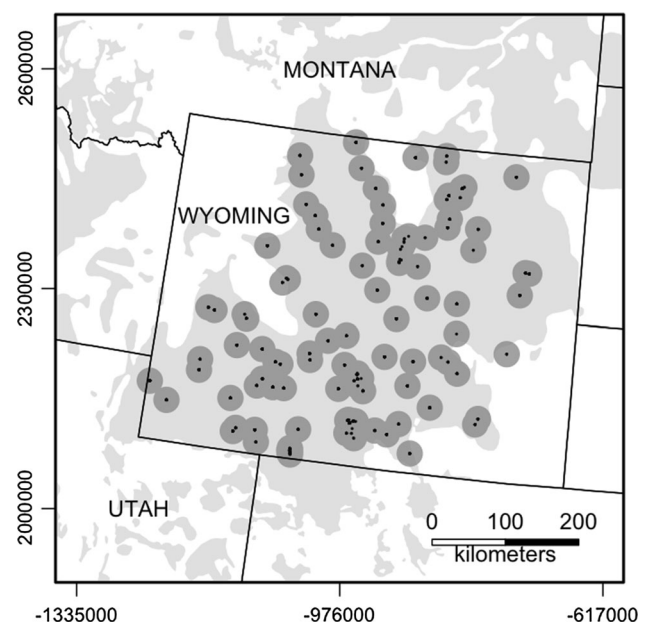


Fig. 1 Sample locations (black dots) across Wyoming, USA. Light gray area shows the current distribution of sage-grouse in Wyoming. Darker gray represent the sample groups created by placing an 18 km buffer around each sample location. The sample groups were used in COLONY to identify related individuals

and correlated allele frequencies (λ) estimated for each population. We generated 20 replicates for each number of genetic clusters (K) from one to six without prior population information. We used 500,000 Markov Chain Monte Carlo repetitions and used 250,000 initial iterations as the burnin period. We inferred K with the aid of STRUCTURE HARVESTER, which produces visualizations following the methods of Evanno et al. (2005) and Pritchard et al. (2000). We obtained averaged individual admixture coefficients for selected values of K by running 100 replicates for the selected values of K and averaging the top 10 % in CLUMMP (Jakobsson and Rosenberg 2007).

In addition to the STRUCTURE analysis, we also estimated the number of genetic clusters using the spatial clustering algorithm implemented in the program TESS 2.1 (Chen et al. 2007). As in STRUCTURE and the initial stages of the GENELAND analysis described below, we determined the optimal number of K between two (TESS does not allow $K = 1$) and six. We conducted 20 replicate runs for each K with 200,000 total sweeps (100,000 burnin) for each run. As suggested by the TESS manual (Durand et al. 2009a), we used non-admixture analysis for the selection of K with the optimal number being selected when the mean Deviance Information Criteria (DIC) reached a minimum asymptote or the Q-matrix of the posterior probabilities was stabilized. Sage-grouse can travel substantial distances for a non-migratory avian species (Fedy et al. 2012) and we suspected some level of admixture among clusters. Thus, using the value of K selected with the non-admixture analysis we conducted an admixture analysis to determine admixture coefficients for each individual. In the admixture analysis we used the linear conditional auto-regressive (CAR) Gaussian model (Besag et al. 1991; Durand et al. 2009b) for the spatial prior on admixture coefficients and the default spatial parameter (0.6). Similar to STRUCTURE, we ran an additional 100 replicate simulations for the selected K and averaged the top 10 % of this simulation in CLUMMP.

Lastly, we examined genetic structure using a Bayesian model executed in a Markov Chain Monte Carlo (MCMC) scheme and implemented in GENELAND version 4.0.3 package (Guillot et al. 2005) in R version 3.0. GENELAND seems to perform well at detecting barriers to gene flow when compared to other, similar, programs (Safner et al. 2011; Blair et al. 2012). We did not have strong a priori expectations for the number of genetic clusters in Wyoming, but previous non-spatial analysis across Wyoming suggested a minimum of three genetic clusters (Oyler-McCance et al. 2005). It was likely that we would have additional power to detect the spatially distributed genetic structure, given that we were using additional individuals and including spatial information in the

clustering analysis (Guillot 2009). Thus, to determine the optimal number of genetic clusters (K), we initially conducted ten independent GENELAND runs allowing K to vary from 1 to 6 to determine the most probable number of clusters. However, GENELAND consistently estimated K at the highest possible value. We increased K in the GENELAND analysis until the posterior estimates peaked and declined prior to our max K . Ultimately, we conducted 10 independent GENELAND runs allowing K to vary from 1 to 40. Each run was executed with the following parameters: 250,000 MCMC iterations, 25 % of iteration discarded as burnin, iterations thinned at 100, correlated allele frequencies, and no uncertainty in spatial coordinates. We determined the modal K from these ten runs. The GENELAND user manual (Guillot et al. 2012) suggests the user may want to treat spatial coordinates as uncertain if samples share the same coordinates. Under these conditions, treating coordinates as uncertain will allow assignment of samples from the same location to different populations. To ensure the spatial uncertainty parameter was not affecting our estimated K we ran the MCMC 5 times using the same parameters except with spatial uncertainty set to 18 km. The modal posterior estimate for the number of clusters was similar to model runs with no spatial uncertainty. However, the GENELAND-inferred model number of clusters was substantially larger than the number estimated using STRUCTURE and TESS (GENELAND: $K = 28$, STRUCTURE and TESS: $K = 4$, see Results below). In order to compare and combine the results from the three clustering approaches we fixed K in the GENELAND analysis at the number estimated by both STRUCTURE and TESS and we then ran the MCMC ten times with the fixed K and all other parameters as above. We calculated the mean logarithm of the posterior probability for each of the ten runs. Cluster membership for each individual was determined by the most probable assigned cluster based on the run with K fixed at four that resulted in the highest posterior probability.

We spatially interpolated the individual assignments from the STRUCTURE (Murphy et al. 2008), TESS and GENELAND analyses using the Krig function in the R package *fields* (Nychka et al. 2015). This function fits a surface to irregularly spaced locational data and we used the default parameters, which fit a 1st order polynomial (i.e., linear) and assumes that correlation between locations is an exponential function of the distance between them. The range parameter was set to the mean distance between locations and the resolution of the fitted surface was set to approximately 3.5 km². We accounted for differences in the admixture coefficients between STRUCTURE, GENELAND and TESS and derived a consensus surface by taking the mean values of the three surfaces at each cell (Row et al. 2010).

The consensus surface was used to define population clusters. We overlaid the predicted cluster boundaries with Wyoming lek locations. We then assigned each lek and genetic sample to the associated population cluster. All leks or samples that fell within the 0.70 predicted isocline for a cluster (i.e., individuals assignment >0.7 to respective cluster) were assigned a value for that cluster. Leks or samples that fell between clusters were not assigned to a cluster and were excluded from the following genetic and population trend analyses. In addition to the demographic analysis (below), for each defined cluster we calculated standard population genetic metrics of genetic diversity (e.g., mean observed and expected heterozygosity across loci) and differentiation (G_{ST} and D_{est}) using the R packages *adegenet* (Jombart 2008) and *mmod* (Winter 2012). We also estimated contemporary effective population size using the linkage-disequilibrium method (assuming a random mating model and a critical value of 0.01; (Hill 1981) for each genetic cluster as implemented in NeEstimator 2.0.1 (Do et al. 2014). Although our populations likely deviate from random mating, this approach has been shown to be relatively robust to violations from non-ideal populations, including migration (Waples and England 2011; Gilbert and Whitlock 2015) and non-equal sex ratios (Waples and England 2011), which would be expected under the lek-based mating system (e.g., few males mating with more females).

We estimated recent migration rates among the identified clusters to assess the levels of demographic independence for each of the clusters. BAYESASS 3.0 assigns individuals and their ancestors (three generations in the past) to their population of origin (Wilson and Rannala 2003). We conducted four independent runs of 5×10^8 total iterations (thinned every 2000th iteration) with unique random seeds and mixing parameters to ensure convergence of the MCMC chains. Trending was visualized with trace plots resulting in 1.0×10^5 iterations being discarded as burnin. Convergence was further assessed by comparing posterior estimates of the four runs. Finally, we combined the chains and estimated the mean and standard deviation of the sampling distribution for the dispersal parameters. Visualization and posterior estimates of the combined chains were conducted using the *coda* package in R (Plummer et al. 2006).

It is possible for these types of clustering approaches to estimate multiple clusters when there is an underlying pattern of isolation by distance (IBD; Frantz et al. 2009; Guillot et al. 2009). We examined this possibility for each of our clusters. If the genetic differentiation between clusters was greater than expected under isolation by distance, we expected the intercept of the relationship between genetic distance and geographic distance (i.e., IBD) to be higher for between-cluster comparisons than

within-cluster comparisons (McRae et al. 2005; Rosenberg et al. 2005). We assessed this expectation by first assigning the 37 lek-based groupings defined in Row et al. (2015) to one of the four averaged clusters and quantifying pairwise F_{ST} between groups. Seven groupings fell outside cluster boundaries and were removed prior to calculating F_{ST} . Subsequently, we compared the IBD relationship for pairwise comparisons between each adjacent cluster (i.e., northwest-to-southwest; northwest-to-northeast; northeast-to-southeast; southeast-to-southwest) to all within cluster IBD relationships using scatter plots and trend lines.

To explicitly test the influence of IBD in our cluster analysis we used a DResD analysis, which is a spatially explicit individually-based approach that corrects genetic distances for IBD and interpolates the resulting residual values (i.e., genetic distances higher or lower than would be expected under isolation by distance) throughout the study area (Hindrikson et al. 2013; Keis et al. 2013). Residual values are interpolated using an inverse distance weighting based on the mid-point locations between pairs of samples. Here, we used one minus the proportion of shared alleles as our individual based genetic distance, because it has a short temporal lag time (i.e., should be reflective of recent genetic structure; Landguth et al. 2010) and does not rely on an underlying genetic assumptions (i.e., Hardy–Weinberg Equilibrium), which may be violated in parts of the study area. We excluded pairwise locations that were greater than 100 km, which is beyond maximum seasonal movement distances for sage-grouse (Fedy et al. 2012) and interpolated the residual variance on a 20 km grid. We ran 1000 random iterations (i.e., residual values randomly distributed among grid locations) to determine where residual values were significantly above (migration barrier) or below (migration corridor) average IBD. Changing the maximum distance for pairwise locations within a biologically meaningful range (50–200 km), had little effect on the results.

Demographic analysis

We used sage-grouse lek data to develop population trend estimates using generalized additive models (GAMs; Fewster et al. 2000) for each of the genetically defined clusters. The trend modeling approach was developed by Fewster et al. (2000) and allows for the description of non-linearity in the data and estimation of confidence intervals. Models were fitted in the GAM R-Package procedure (R-package *mgcv*, Wood 2006, R version 3.0.2, <http://www.r-project.org>). The model estimates were used to generate the abundance index curve, which represented the relative abundance and is a ratio of exponentially transformed year effects (Fewster et al. 2000). This approach also allows for the investigation of population change points with a

numerical analysis of the second derivatives of the smoothed indices. Essentially, change points indicate significant changes in the rate of growth or decline in the population trends from the previous year. Details on the calculation and determination of significance of change points are provided in Fewster et al. (2000). To facilitate comparison among clusters, all index values were scaled so that the index for the first year in the data series was equal to 1. Details of the trend modeling approach applied to Wyoming sage-grouse can be found in Fedy and Aldridge (2011) and Fedy and Doherty (2011). In addition to using GAMs to characterize population trends for each population, we also developed a single integrated model that included population identification as a factor variable in the model development to determine whether the inclusion of a population term in the trend model plays a statistically significant role in the overall trend model. For this model the smoothing parameter for year was estimated using an un-biased risk estimator criterion within the *gam* function. Model fit diagnostics were inspected to ensure proper model specification.

Sage-grouse lek count data were obtained from the Wyoming Game and Fish Department (WGFD). We gathered count data from all active leks within each population cluster. When leks were counted more than once within a year, we followed convention and chose the maximum count. Using GAMs to describe population trends over time can be affected by >50 % missing data at each site (Atkinson et al. 2006). Therefore, we removed all lek sites with >50 % missing data across the time series. Sage-grouse populations in Wyoming were at historic lows in 1995 (Fedy and Aldridge 2011; Fedy and Doherty 2011; Garton et al. 2011) and the number of leks surveyed in Wyoming was also substantially lower in the early 1990s than in the mid 1990s to 2000s (Fedy et al. 2015). The ubiquitous low population numbers in 1995 and the increasing number of sites surveyed led us to select 1995 as the beginning year for our trend assessment. Thus, we developed trend models for all leks meeting the above criteria in each population cluster from 1995 to 2013. This approach ensured sufficient data for precise trend estimates and facilitated comparison among clusters.

We calculated the population change based on the estimated abundance index for two time periods for each population to assess the similarities in percent change among populations. The approach uses the bootstrap replicates to estimate the percent population change and associated confidence intervals between any 2 years (Fewster et al. 2000). We assessed two intervals for all populations and calculated the percent population change between these intervals. One interval represented the length of the time series (1996–2012) and the other interval represented the most recent population decline from a peak in

2007 to 2012. We did not consider the first or last year in the time series when calculating the percent pairwise change because GAMs produce less reliable estimates at the beginning and end of time series due to the smoothing process (Hewson and Noble 2009).

To assess whether there was a statistically significant difference among the trends in the different populations we developed an additional GAM model that was fit using data from all four populations and contained a population factor variable that identified each population. The smoothing parameter for the year-term was estimated using the unbiased risk estimator criterion within the *gam* function. Model fit diagnostics were inspected to ensure proper model specification. When estimating the test statistics and significance of factors the *gam* function compares all factors to the first factor entered into the model (e.g., is there a statistically significant difference between population 1 and all other populations). Therefore, we iteratively changed the base population and a Bonferroni correction was applied to the population pairwise comparisons.

Modeling trends at lek sites over time does not explicitly address the distribution of leks that have gone inactive. For example, it is possible for overall estimated trends in a population to demonstrate positive growth even if the proportion of active leks within the area declines. This is because of the GAM model structure defined above and is true of other trend estimation approaches that include site-level covariates (e.g., route regression). Additionally, we removed inactive sites from the trend analysis because they had reached zero consistently and therefore, could not inform population fluctuations. Therefore, we included an additional analysis that explicitly described the extirpation of leks within each population. The WGFD defines active leks as those that have been attended by male sage-grouse during the strutting season. A lek is only classified as inactive if sufficient data indicates that no strutting activity took place throughout the strutting season. The inactive designation requires documentation of no birds in attendance during at least two surveys separated by a week and conducted during ideal conditions. We calculated the percentage of active leks during each year from 1995 to 2013.

Results

DNA isolation, amplification, and post processing

We attempted to extract DNA from 1761 samples. Of these, 1195 amplified at 7 or more loci. The AlleleMatch and Dropout analysis identified 949 unique individuals (see Row et al. 2015 for details).

Spatial genetic structure

The 18 km buffer applied to each sample location produced 33 unique sample groups (Fig. 1). Sample sizes varied within each buffer ($n = 7\text{--}132$). After iteratively removing all potential full-sib or parent offspring relationship, we retained one individual from each closely related group and were left with 872 unique and unrelated individuals.

In the STRUCTURE analysis, the log-posterior probability of K appeared to level off at $K = 4$, but ΔK method favored a $K = 2$ (A. 1). As expected, when spatial information was included using TESS, the number of clusters was at the higher end of this range with the DIC values reaching a minimum asymptote at $K = 4$. Plotting the Q-matrices of individual admixture coefficients, it was clear that both programs were in general agreement for individual assignments with more uncertainty or admixture in the STRUCTURE and TESS outputs than in the GENELAND output (Fig. 2). The consensus value of $K = 4$ estimated by STRUCTURE and TESS was reasonable given the data availability and the goal of the study to estimate broad-scale population structure.

The modal K from the 10 GENELAND runs varying K from 1 to 40 was 28 clusters with good mixing around that value across runs. We re-ran the GENELAND MCMC 10 times with K fixed to 4, the inferred number of clusters from TESS and STRUCTURE to facilitate integration. Plotting the Q-matrices of individual admixture coefficients revealed general agreement for individual assignments among TESS, STRUCTURE, and GENELAND (A.

2). The overall trends were the same, but GENELAND had little uncertainty in individual assignments (Fig. 2).

Similar to the Q-matrices, the spatially interpolated individual assignments from STRUCTURE (A. 2), TESS (A. 3), and GENELAND (A. 4) for a $K = 4$ demonstrated general agreement in the location of the clusters between the three analytical approaches. The averaged interpolated surfaces from the three clustering approaches identified four discrete, spatially distributed clusters, with clear zones of transition. All individuals (including related individuals) that fell within the 0.70 predicted isocline for given cluster were retained for the calculation of genetic summary statistics (734 assigned individuals with a range of 96–280 individuals per cluster; Table 1).

Average allelic richness across loci and clusters was 12.37, with mean observed heterozygosity = 0.79, mean expected heterozygosity = 0.83, and mean $F_{IS} = 0.40$ (Table 1). As expected, Fisher exact tests suggested significant genetic structure between clusters ($P < 0.001$) and pairwise G_{ST} and D_{est} of the inferred populations showed weak differentiation, (mean $G_{ST} = 0.012$, standard deviation = 0.0038; mean $D_{est} = 0.11$, standard deviation = 0.035) and all pairwise G_{ST} and D_{est} and were highly correlated (0.99; Table 2). Effective population sizes were smallest in the northwest and largest in the southwest, with sizes in the south being approximately twice as large as those in the north (Table 1).

BAYESASS estimated that >90 % of individuals in both northern clusters originated from their own cluster in each generation (northeast 0.94 ± 0.02 ; northwest 0.94 ± 0.02) and thus had less than 10 % immigration

Fig. 2 Histograms of averaged assignment probabilities for three Bayesian clustering approaches (**a** STRUCTURE, **b** TESS, **c** GENELAND). Each vertical bar represents an individual ($n = 872$) and its assignment probability to one of four clusters. Individuals are arranged in order by Northwest, Northeast, Southwest and Southeast

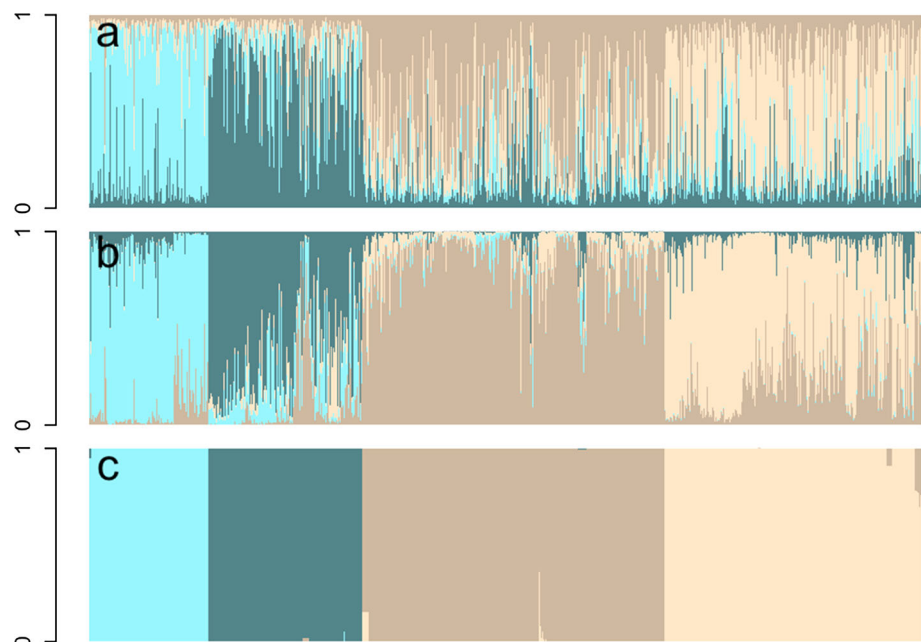


Fig. 3 Sage-grouse population delineation across Wyoming, USA based on $n = 872$ unique and unrelated individuals across the state. Population membership was estimated by averaging population membership from STRUCTURE, GENELAND, and TESS. Dotted black lines represent the population boundaries used to assign leks to each population cluster for the trend analysis. Blue dots represent all sample locations included in the analyses

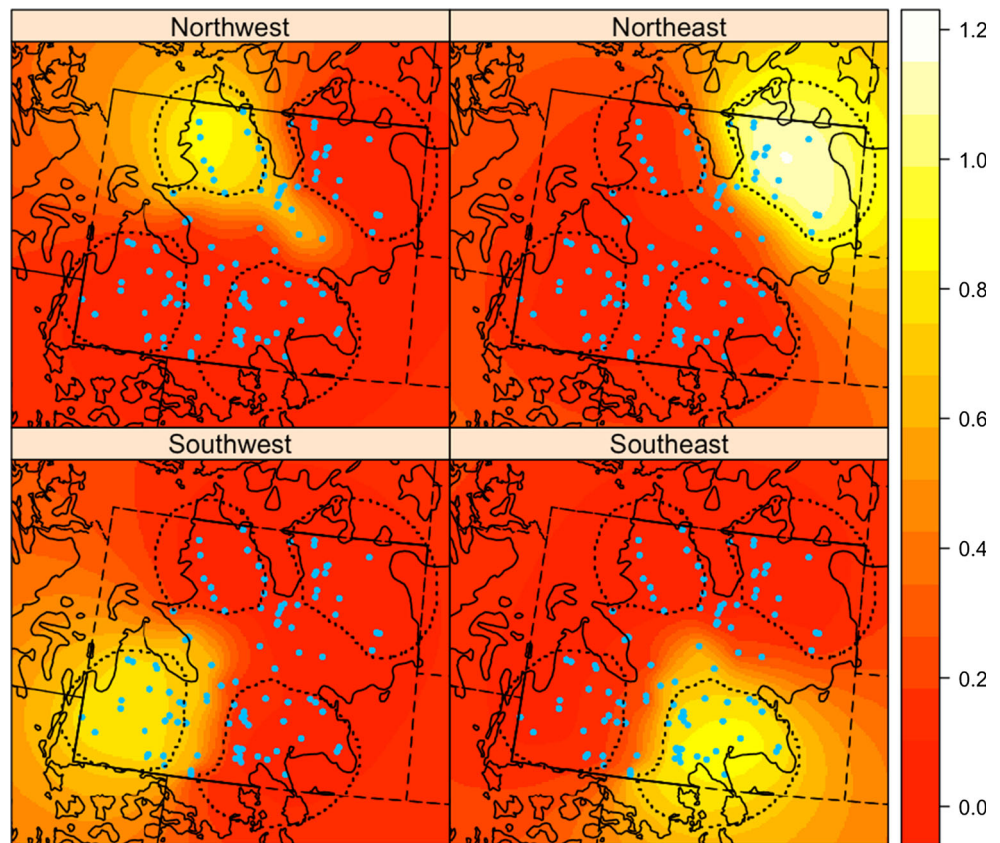


Table 1 Summary of genetic parameters by cluster for the greater sage-grouse assigned to the four clusters ($n = 734$) in Wyoming

Cluster	n	N_a	A_R	H_O	H_E	F_{IS}	N_e
Northwest	137	12.43 (5.50)	11.94 (5.20)	0.76 (0.11)	0.80 (0.11)	0.053 (0.05)	197 (158–257)
Northeast	96	12.43 (5.05)	12.43 (5.05)	0.78 (0.07)	0.81 (0.06)	0.034 (0.03)	247 (195–332)
Southwest	280	13.79 (5.29)	12.22 (4.44)	0.79 (0.07)	0.83 (0.06)	0.046 (0.05)	569 (423–845)
Southeast	221	13.93 (5.62)	12.91 (4.95)	0.81 (0.08)	0.83 (0.07)	0.046 (0.05)	527 (424–685)

Variables presented are n sample size, N_a number of alleles, A_R , allelic richness, H_O observed heterozygosity, H_E expected heterozygosity, and F_{IS} , calculated as Robertson and Hill (1984). Means and standard deviations (in brackets) across loci are shown. Effective population sizes (N_e) and their 95 % confidence intervals (in brackets) calculated using the linkage-disequilibrium method are also shown

Table 2 Pairwise G_{ST} (lower matrix) and Jost D_{est} (upper matrix) between four genetically defined clusters of sage-grouse in Wyoming

	Northwest	Northeast	Southwest	Southeast
Northwest		0.095	0.131	0.144
Northeast	0.011		0.123	0.132
Southwest	0.014	0.014		0.047
Southeast	0.015	0.015	0.005	

from all other clusters combined (Table 3). These estimates were consistent across each of the four independent runs. The southeastern and southwestern clusters had greater

than 10 % immigration and the majority of this migration was between the southern clusters (Table 3). Although the migration between the southern clusters was greater than 10 % in all runs the direction of migration was not consistent among runs.

Our assessment of the potential for an underlying pattern of IBD indicated different patterns for the northern and southern populations. For all comparisons using either of the two northern clusters, the intercept for the IBD relationship was greater for between than within cluster relationships (A. 5). This suggests that differentiation from the northern populations was greater than expected under IBD. This was not the case for the southern populations (A. 5) and therefore, IBD may be an underlying force influencing

Table 3 Pairwise asymmetric migration rates among sage-grouse genetic clusters

	Northeast	Northwest	Southeast	Southwest
Northeast	0.94 (0.02)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
Northwest	0.01 (0.01)	0.94 (0.02)	0.01 (0.01)	0.03 (0.03)
Southeast	0.04 (0.02)	0.04 (0.02)	0.86 (0.07)	0.25 (0.03)
Southwest	0.01 (0.01)	0.01 (0.01)	0.13 (0.07)	0.71 (0.02)

Standard deviations are presented in brackets. Each column represents the proportion of individuals that originate from the cluster identified by row name

the patterns of genetic variation in the southern populations.

The DResD analysis provided evidence of spatial genetic differentiation that was greater than expected under the null hypothesis of IBD (Fig. 4). Several areas across the state were identified in which the residuals were significantly higher or lower than expected by IBD. In particular, the northeastern region suggested blending of individuals and higher relatedness than expected by IBD. Central regions of the state also demonstrated significant deviations from IBD indicating less movement (i.e., barriers) to movement through the central portion of the state and separating the northern and eastern portion of their range in the state (Fig. 3). It is of note that these

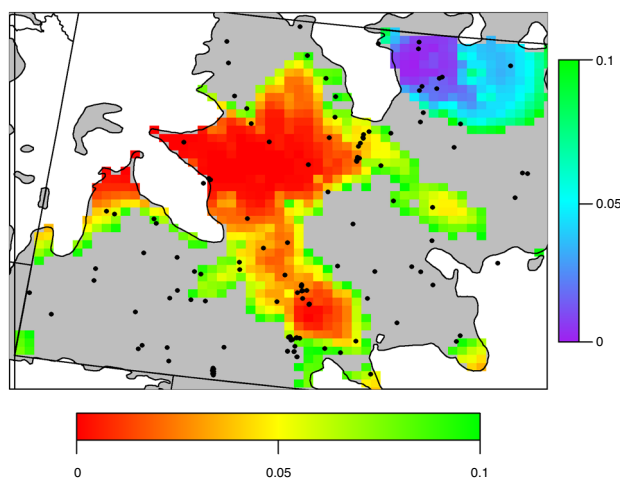


Fig. 4 Putative areas of high movement (e.g., corridors) and high resistance (i.e., barriers) for sage-grouse in Wyoming, USA. Colors represent significant deviations from IBD and the test value. The vertical color bar shades from green to purple and represents the P value associated with the lower tail of the test from $P = 0.10$ to $P < 0.001$ and areas of high movement. The horizontal color bar shades from green to red and represents the P value associated with the upper tail of the test from $P = 0.10$ to $P < 0.001$ and areas of high resistance. Gray areas represent no significant deviation and support of the null hypothesis of IBD. White areas are outside of the sage-grouse range in Wyoming. Black dots represent sample locations

barriers roughly corresponds with the division between southwestern and southeastern subpopulations and the north–south divide estimated through the Bayesian cluster analyses.

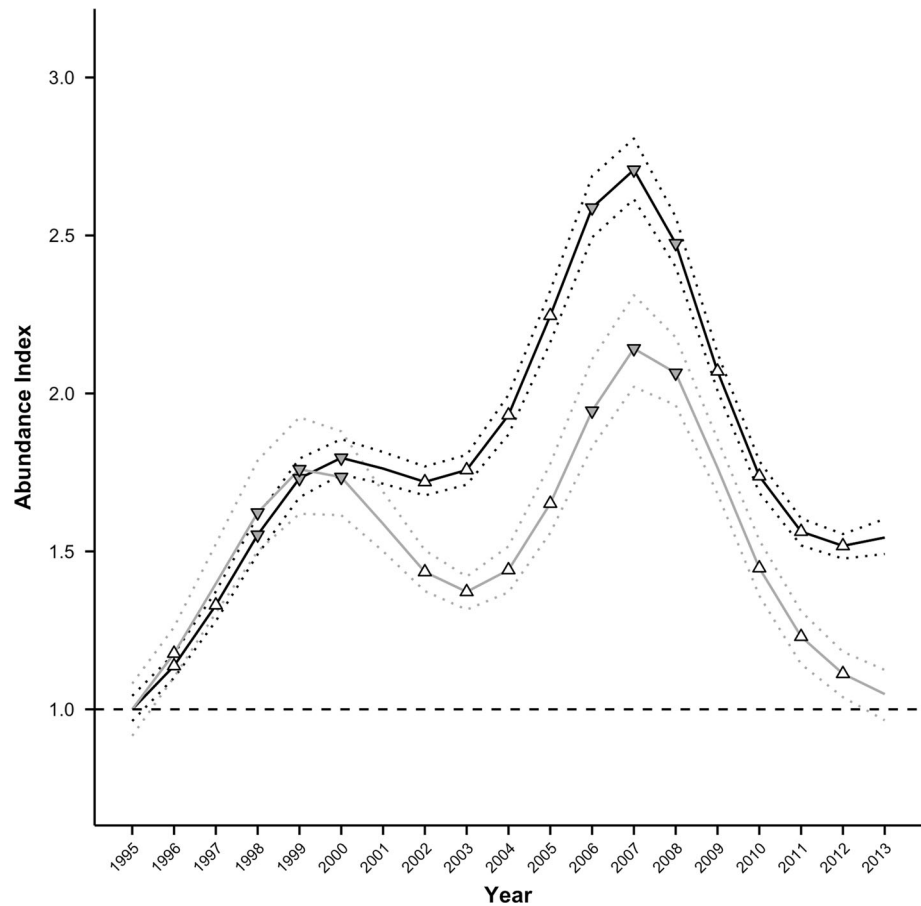
Demographic analysis

Across all four clusters there were 1053 active lek sites monitored that met the criteria for inclusion in our trend analyses from 1995 to 2013. Each cluster contained a sufficient number of leks to describe population trends using the GAMs (northwest $n = 164$; northeast $n = 248$; southeast $n = 298$; southwest $n = 344$). Overall, fluctuations in the number of males attending leks were similar across the four clusters including similar change points (A. 6). The model including a factor for cluster membership suggested that the cluster term played a statistically significant role in model specification ($P < 0.001$). The iterative assessment of each cluster indicated that the northwest was significantly different from each of the other clusters (northeast $\beta = -4.65 \pm 1.08$, $P < 0.001$; southeast $\beta = -3.67 \pm 0.84$, $P < 0.001$; southwest $\beta = -4.43 \pm 1.10$, $P < 0.001$) and there were no other significant differences among clusters ($P > 0.05$; range of P values 0.44–0.78). Examination of the model predicted values (Fig. 5) indicated trends in the northwest were generally lower, but parallel to the combined trends from the other three clusters with close or overlapping confidence intervals from 1995 until the population peak in 2000. After 2000 the northwest values seemed to diverge substantially, though generally parallel trends with confidence intervals that no longer overlapped with the other three clusters.

The calculation of percent population change demonstrated substantial variation in the abundance from 1996 to 2012. In particular, the northeast population had an estimated decline from 1996 of -43% (CI -60 , -21). The northwest cluster estimate overlapped 0 with a point estimate of -9% (CI -25 , 9). These contrast significantly with the positive estimates for the southeast (20% , CI 1 , 39) and the southwest (46% , CI 29 , 69) clusters. All clusters showed similar declines throughout the most recent population declines of about 60% from 2007 to 2012 (Fig. 6).

Our assessment of active and inactive leks indicated different population patterns than the trend analysis. Overall, the percent of active leks was relatively stable for the northwest (84 – 100%) and the southeast (67 – 83%). However, the northeast and the southwest demonstrated overall declines in the percent of active leks over the time series. The northeast, in particular, decreased from a high of 84% in 2006 to a low of 48% in 2013 (Fig. 7).

Fig. 5 Estimated sage-grouse abundance index trends for lek sites from the Northwest population cluster and from all other population clusters combined. Trends are based on male lek counts from 1995 to 2013 and were modeled using generalized additive models. The *black line* represents the estimate trend in abundance index for the leks combined across the Southwest, Southeast, and Northeast populations. The *gray line* represents the estimate trend in abundance index for the Northwest population. *Triangles* represent significant change points, which indicate a significant change in the rate of increase or decrease of the abundance index. *Gray triangles* represent significant downturns and *white triangles* represent significant upturns. *Dotted lines* represent the bootstrap-estimated 95 % confidence intervals



Discussion

Many studies have examined population trends and dynamics in wildlife populations, but it is less common that demographic and genetic data are combined into a single analysis and assessment of population risk (Gaggiotti et al. 2002; Mayer et al. 2009; Olsen et al. 2014). The combination of genetics and demographics, as demonstrated here, led to greater insight into population fluctuations. For example, previous work demonstrated highly correlated population trends for sage-grouse across Wyoming (Fedy and Doherty 2011), but in this present study, we were able to identify four genetic clusters and assess trends and lek activity for each separately. Through this process, we detected significant regional differences not previously recognized and identified regional subpopulations that are likely more at risk due to lek extirpations, decreasing population trends, greater genetic isolation, and lower effective population sizes.

Population genetic analyses typically apply to evolutionary time scales and may not inform contemporary processes of relevance to management (Pearse and Crandall 2004; Palsbøll et al. 2007). Hence, demographics, rather than genetics, may lead to more relevant definition

of populations and management units (Waples and Gaggiotti 2006; Palsbøll et al. 2007; Lowe and Allendorf 2010). The concern with applying a strictly genetic approach to defining populations is that genetic differences may take a long time to accumulate, and recently fragmented, yet demographically independent populations, may go undetected (Höglund et al. 2011).

Our application of genetic approaches to population definition differed from those critiqued by other authors (Waples and Gaggiotti 2006; Palsbøll et al. 2007; Lowe and Allendorf 2010). We first used clustering to identify our groups, assigned membership based on spatial predictions of membership probability, and then assessed measures of demographic independence that are relevant to species management and upon which managers make decisions. Thus, as others have suggested we delineated intra-specific units through an assessment of genetics and the geographic, and demographic factors that influence gene flow (Serrouya et al. 2012). Our combined genetic and demographic analyses (e.g., overall trends and lek activity) suggested the northern subpopulations are likely demographically separate populations from each other and from the southern clusters. Therefore, it may be useful to generate specific management plans tailored to the needs of

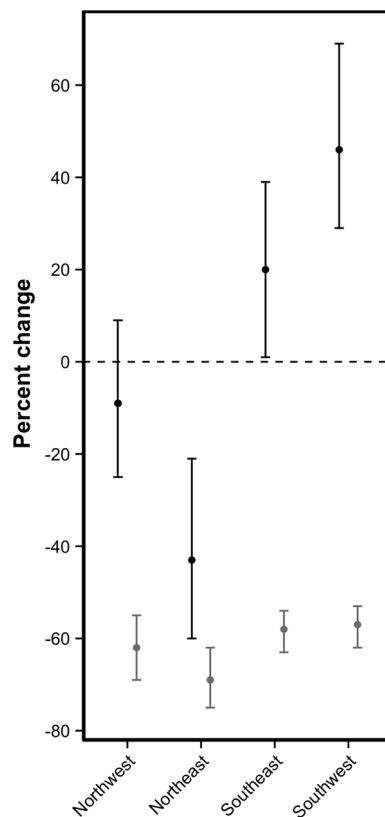


Fig. 6 Estimated percent population change based on the trend models for all four populations. *Black dots* and associated error bars represent the percent change and 95 % confidence intervals from 1996 to 2012. *Grey dots* and associated error bars represent the percent change and 95 % confidence intervals from 2007 to 2012. Populations are indicated along the x-axis NW Northwest, NE Northeast, SE Southeast, SW Southwest

each subpopulation. Regional management plans will be particularly relevant if the threats and impacts to subpopulations also vary regionally. However, if the threats are similar across all, or some, of the subpopulations then perhaps management targeted across subpopulations is most appropriate. Similar to Olsen et al. (2014) our clustering approach does not unequivocally demonstrate demographic independence among subpopulations. However, in a simulation analysis migration rates of <10 % between populations led to demographic independence (Hastings 1993) leading some to suggest this should be used as a cutoff value for defining populations (Waples and Gaggiotti 2006). The two northern subpopulations both had immigration rates around 6 % suggesting immigration from other regions would not have a large effect on population trends.

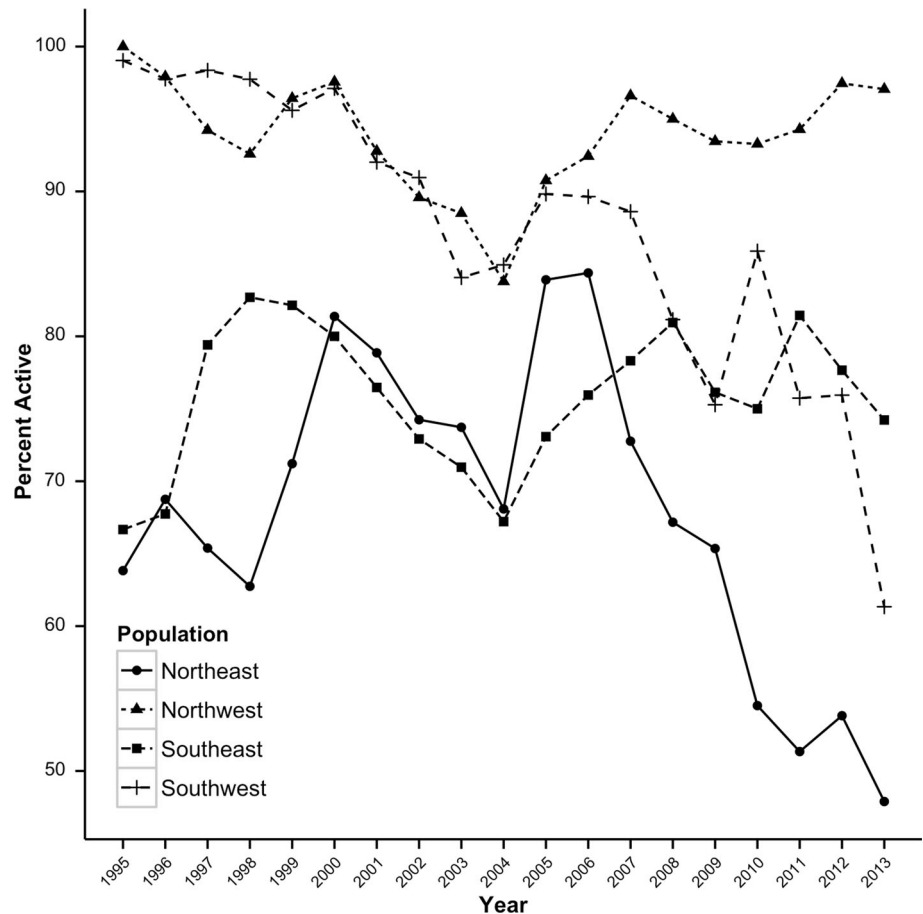
Collectively, the results suggest these subpopulations could be managed as three demographically independent units with two northern subpopulations and one southern subpopulation, which demonstrated isolation-by-distance based on our assessment of the changes in intercepts.

However, our DResD analysis indicated the potential for greater than expected resistance to movement (i.e., a barrier) between some portions of the two southern subpopulations. Additionally, this putative barrier roughly coincided with the divisions suggested through our multiple Bayesian clustering analyses. The weight of evidence seems to suggest the variation in the southern portion of the state may be due to more than simply isolation-by-distance, but we recommend further research to address this issue conclusively. Additionally, the general location of the predicted east–west division detected in the southern portion of the state coincides with previous estimates of subpopulation boundaries (Connelly et al. 2004) and connectivity (Knick and Hanser 2011).

Our research demonstrates one of the first steps to integrating genetic and demographic data. The population clusters we identified could also be used to focus additional studies into the demography of populations such as population viability analyses. Additionally, with the rapid development of potentially adaptive markers in sage-grouse (e.g., single-nucleotide polymorphisms), these clusters would provide an excellent place to begin testing for adaptive differentiation using genomic data (Funk et al. 2012). Further analysis of demographics could consider multiple parameters in addition to lek counts and activity, where data are available, such as age-specific demographic parameters.

Many ecological factors could have influenced the observed pattern of genetic clustering in Wyoming. Multiple studies have demonstrated the positive association between structural connectivity (i.e., habitat) and functional connectivity (i.e., gene flow; Laiolo and Tella 2006; Wang et al. 2008; Row et al. 2010). However, structural habitat connectivity does not always result in functional connectivity (Ribe et al. 1998; Spear et al. 2010). The structural connectivity of seasonal habitats (Fedy et al. 2014; Row et al. 2015) likely influences functional gene flow across the state; however, these habitat may not be as influential as large-scale landscape features that represent substantial areas of unsuitable habitat such as mountain ranges. In Wyoming, the Wind River Mountains in the southwest and the Big Horn Mountains in the north likely contribute to the observed patterns of genetic differentiation. The definition of our study extent (i.e., the sage-grouse range within the political boundaries of the state of Wyoming) likely influenced the patterns observed. Varying spatial extents to larger (e.g., multiple states) and smaller (e.g., National Oceanic and Atmospheric Administration climate divisions; Fedy and Aldridge 2011) would allow formal assessment of the influence of varying spatial extents on the estimation of genetic (e.g., cluster identification, migration rates, effective population size) and demographic (e.g., lek counts) parameters.

Fig. 7 Percent of active leks within each population from 1995 to 2013. The value represents the percent of leks surveyed that were active for each year divided by the sum of all leks that were surveyed with sufficient intensity to classify as whether active or inactive. Leks of unknown status were not included in the calculation



Our demographic assessment suggested the two northern subpopulations in Wyoming are most at risk. The northwest subpopulation in particular demonstrated significantly different population trends over time compared to the other subpopulations. Our analyses suggest that the northwest has not performed as well as the other populations—particularly since 2009. Consistent with these results, the northern subpopulations had lower effective population sizes with the northwest subpopulation also having the lowest Allelic Richness and highest levels of inbreeding as estimated by F_{IS} . Calculation of pairwise population change from 1996 to 2012 indicated recovery from the statewide low numbers of the mid-1990s (Fedy and Aldridge 2011) in the southern, but not the northern populations. At best, the northern populations have similar population estimates based on lek counts as the mid-1990s, which represented a low across a 44-year data set from 1965 to 2008 (Fedy and Aldridge 2011). Percent pairwise change was positive for the southwest cluster from 1996 to 2012, but this estimate does not take into account the declining percent of active leks in the southwest. The pairwise estimate for the same period in the southeast cluster was positive, but the lower CI was close to zero (1 %) suggesting that the positive population change may

not be significantly different from 0 or no change from the lows of the mid 1990s. Our analysis of lek counts suggests greater recovery in the southwest than the southeast (southwest 46 %; southeast 20 %) over the study time-frame. However, it is important to note that the confidence intervals for these estimates overlap (Fig. 6). Comparison of percent change from the last population cycle peak to the most recent year available for comparison (2007–2012) indicated that all populations have declined at a similar rate since the last population peak.

If we had only considered population trends, we would not have identified the concerns with lek extirpation in the northeastern portion of the state. The assessment of trends in lek counts alone is insufficient to provide a complete picture of population risk. As noted in the methods, a lek is only classified as inactive if sufficient data indicates that no strutting activity took place throughout the strutting season. The inactive designation requires documentation of no birds in attendance during at least two surveys separated by a week and conducted during ideal conditions. Many of the inactive leks would have been excluded from trend analyses, leading to some differences in the interpretation of the trend and lek activity analyses. The additional assessment of lek-activity demonstrated clear declines in the

proportion of active leks within the northeast and potentially the southwest. However, we were not able to incorporate survey effort into our estimates of lek activity. Future studies to address lek activity should be designed to explicitly incorporate survey effort and spatial configuration into models of lek activity. Additionally, the cyclic nature of sage-grouse populations in Wyoming (Fedy and Doherty 2011) and potentially elsewhere in the range, remains largely ignored in the sage-grouse literature. The majority of studies on sage-grouse in Wyoming continue to use linear models to describe patterns in population trends that are clearly better described by non-linear models. We recommend future studies assess the nature of the population trends in their study extent of interest (i.e., linear vs. non-linear) and that future studies of sage-grouse in Wyoming attempt to explicitly address the cyclic patterns in sage-grouse trends. As noted by Fedy and Doherty (2011), the context of *where* you are in the cycle can fundamentally influence conclusions regarding species recovery or population stability. Critically, there has been no comprehensive assessment of the mechanisms underlying these large-scale population trends, which would help address many priority management actions that are driven by species population trends. As demonstrated here, population trends can vary substantially among subpopulations. We suggest a similar incorporation of genetic and demographic data across the range of sage-grouse.

Acknowledgments We thank Thomas Christiansen and Christopher Keefe for their support, coordination, and insights. Zachary Bowen was integral to the initiation of the project and we appreciate his contributions early on in the research. We also thank the many individuals who invested time and energy into collecting and submitting feathers and FTA cards for genetic analysis and the many individuals who have collected lek count data in Wyoming. The manuscript was improved through the comments from Christopher Kirol and three anonymous reviewers. Funding for the research was provided by the U.S. Bureau of Land Management, U.S. Geological Survey, and Wyoming Game and Fish Department. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

- Allendorf FW, Luikart G (2007) Conservation and the Genetics of Populations. Blackwell, Oxford
- Atkinson PW, Austin GE, Rehfish MM et al (2006) Identifying declines in waterbirds: the effects of missing data, population variability and count period on the interpretation of long-term survey data. *Biol Conserv* 130:549–559. doi:[10.1016/j.biocon.2006.01.018](https://doi.org/10.1016/j.biocon.2006.01.018)
- Besag J, York J, Mollie A (1991) Bayesian image-restoration, with 2 applications in spatial statistics. *Ann Inst Stat Math* 43:1–20. doi:[10.1007/BF00116466](https://doi.org/10.1007/BF00116466)
- Blair C, Weigel DE, Balazik M et al (2012) A simulation-based evaluation of methods for inferring linear barriers to gene flow.

- Mol Ecol Resour* 12:822–833. doi:[10.1111/j.1755-0998.2012.03151.x](https://doi.org/10.1111/j.1755-0998.2012.03151.x)
- Caizergues A, Ratti O, Helle P et al (2003) Population genetic structure of male black grouse (*Tetrao tetrix*) in fragmented vs. continuous landscapes. *Mol Ecol* 12:2297–2305
- Chen C, Durand E, Forbes F, Francois O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Mol Ecol Notes* 7:747–756
- Connelly JW, Knick ST, Schroeder MA, Stiver SJ (2004) Conservation assessment of greater sage-grouse and sagebrush habitats. Western Association of Fish and Wildlife Agencies, Cheyenne, Wyoming
- Do C, Waples RS, Peel D et al (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Mol Ecol Resour* 14:209–214
- Doherty K, Tack J, Evans J, Naugle D (2010) Mapping breeding densities of greater sage-grouse: a tool for range-wide conservation planning. BLM completion report: Interagency Agreement L10PG00911, p 30
- Durand E, Chen C, François O (2009a) TESS version 2.3—reference manual. Available at memberstmc. imag. fr/Olivier. Francois/tess. html
- Durand E, Jay F, Gaggiotti OE, François O (2009b) Spatial inference of admixture proportions and secondary contact zones. *Mol Biol Evol* 26:1963–1973. doi:[10.1093/molbev/msp106](https://doi.org/10.1093/molbev/msp106)
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587
- Fedy BC, Aldridge CL (2011) Long-term monitoring of greater sage-grouse populations: the importance of within-year repeated counts and the influence of scale. *J Wildl Manag* 75:1022–1033
- Fedy BC, Doherty KE (2011) Population cycles are highly correlated over long time series and large spatial scales in two unrelated species: greater sage-grouse and cottontail rabbits. *Oecologia* 165:915–924
- Fedy BC, Aldridge CL, Doherty KE et al (2012) Interseasonal movements of greater sage-grouse, migratory behavior, and an assessment of the core regions concept in Wyoming. *J Wildl Manag* 76:1062–1071
- Fedy BC, Doherty KE, Aldridge CL et al (2014) Habitat prioritization across large landscapes, multiple seasons, and novel areas: An example using greater sage-grouse in Wyoming. *Wildl Monogr* 190:1–39
- Fedy BC, Martin K, Ritland C et al (2008) Genetic and ecological data provide incongruent interpretations of population structure and dispersal in naturally subdivided populations of white-tailed ptarmigan (*Lagopus leucura*). *Mol Ecol* 17:1905–1917. doi:[10.1111/j.1365-294X.2008.03720.x](https://doi.org/10.1111/j.1365-294X.2008.03720.x)
- Fedy BC, O'donnell MS, Bowen ZH (2015) Large-scale control site selection for population monitoring: An example assessing sage-grouse trends. *Wildl Soc Bull* 39:700–712
- Fewster RM, Buckland ST, Siriwardena GM et al (2000) Analysis of population trends for farmland birds using generalized additive models. *Ecology* 81:1970–1984
- Fike JA, Oyler-McCance SJ, Zimmerman SJ, Castoe TA (2015) Development of 13 microsatellites for Gunnison Sage-grouse (*Centrocercus minimus*) using next-generation shotgun sequencing and their utility in Greater Sage-grouse (*Centrocercus urophasianus*). *Cons Gen Res* 7:211–214
- Frantz AC, Cellina S, Krier A et al (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously

- distributed population: clusters or isolation by distance? *J Appl Ecol* 46:493–505. doi:[10.1111/j.1365-2664.2008.01606.x](https://doi.org/10.1111/j.1365-2664.2008.01606.x)
- Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends Ecol Evol* 27:489–496. doi:[10.1016/j.tree.2012.05.012](https://doi.org/10.1016/j.tree.2012.05.012)
- Gaggiotti OE, Jones F, Lee WM et al (2002) Patterns of colonization in a metapopulation of grey seals. *Nature* 416:424–427. doi:[10.1038/416424a](https://doi.org/10.1038/416424a)
- Galpern P, Manseau M, Hettinga P et al (2012) ALLELEMATCH: an R package for identifying unique multilocus genotypes where genotyping error and missing data may be present. *Mol Ecol Resour* 12:771–778. doi:[10.1111/j.1755-0998.2012.03137.x](https://doi.org/10.1111/j.1755-0998.2012.03137.x)
- Garton EO, Connelly JW, Horne JS et al (2011) Greater sage-grouse population dynamics and probability of persistence. In: Knick ST, Connelly JW (eds) *Studies in avian biology*. University of California Press, Berkeley, California, USA, pp 293–381
- Gilbert KJ, Whitlock MC (2015) Evaluating methods for estimating local effective population size with and without migration. *Evolution (N Y)* 69:2154–2166. doi:[10.1111/evo.12713](https://doi.org/10.1111/evo.12713)
- Guillot G (2009) On the inference of spatial structure from population genetics data using the Tess program. *Bioinformatics* 25:1796–1801. doi:[10.1093/bioinformatics/btp267](https://doi.org/10.1093/bioinformatics/btp267)
- Guillot G, Estoup A, Mortier F, Cosson JF (2005) A spatial statistical model for landscape genetics. *Genetics* 170:1261–1280. doi:[10.1534/genetics.104.033803](https://doi.org/10.1534/genetics.104.033803)
- Guillot G, Leblois R, Coulon A, Frantz AC (2009) Statistical methods in spatial genetics. *Mol Ecol* 18:4734–4756. doi:[10.1111/j.1365-294X.2009.04410.x](https://doi.org/10.1111/j.1365-294X.2009.04410.x)
- Guillot G, Santos F, Estoup A (2012) Population genetics analysis using R and the Geneland program. <http://www2.imm.dtu.dk/~gigu/Geneland/Geneland-Doc.pdf>
- Hanks EM, Hooten MB, Knick ST et al (2016) Latent spatial models and sampling design for landscape genetics. *Ann Appl Stat* 10:1041–1062. doi:[10.1214/16-AOAS929](https://doi.org/10.1214/16-AOAS929)
- Hastings A (1993) Complex interactions between dispersal and dynamics: lessons from coupled logistic equations. *Ecology* 74:1362–1372
- Hewson CM, Noble DG (2009) Population trends of breeding birds in British woodlands over a 32-year period: relationships with food, habitat use and migratory behaviour. *Ibis (Lond 1859)* 151:464–486
- Hill G (1981) Estimation of effective population size from data on linkage disequilibrium. *Genet Res* 38:209–216
- Hindrikson M, Remm J, Männil P et al (2013) Spatial genetic analyses reveal cryptic population structure and migration patterns in a continuously harvested grey wolf (*Canis lupus*) population in North-Eastern Europe. *PLoS One* 8:1–12. doi:[10.1371/journal.pone.0075765](https://doi.org/10.1371/journal.pone.0075765)
- Höglund J, Larsson JK, Corrales C et al (2011) Genetic structure among black grouse in Britain: implications for designing conservation units. *Anim Conserv* 14:400–408. doi:[10.1111/j.1469-1795.2011.00436.x](https://doi.org/10.1111/j.1469-1795.2011.00436.x)
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806
- Jombart T (2008) ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. doi:[10.1093/bioinformatics/btn129](https://doi.org/10.1093/bioinformatics/btn129)
- Jones OR, Wang J (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour* 10:551–555. doi:[10.1111/j.1755-0998.2009.02787.x](https://doi.org/10.1111/j.1755-0998.2009.02787.x)
- Keis M, Remm J, Ho SYW et al (2013) Complete mitochondrial genomes and a novel spatial genetic method reveal cryptic phylogeographical structure and migration patterns among brown bears in north-western Eurasia. *J Biogeogr* 40:915–927. doi:[10.1111/jbi.12043](https://doi.org/10.1111/jbi.12043)
- Knick ST, Hanser SE (2011) Connecting pattern and process in greater sage-grouse populations and sagebrush landscapes. In: Knick ST, Connelly JW (eds) *Greater sage-grouse: ecology and conservation of a landscape species and its habitats*. University of California Press, Berkeley p, pp 383–405
- Kolomyjec SH, Grant TR, Johnson CN, Blair D (2012) Regional population structuring and conservation units in the platypus (*Ornithorhynchus anatinus*). *Aust J Zool*. doi:[10.1071/ZO13029](https://doi.org/10.1071/ZO13029)
- Krueger T, Page T, Hubacek K et al (2012) The role of expert opinion in environmental modelling. *Environ Model Softw* 36:4–18. doi:[10.1016/j.envsoft.2012.01.011](https://doi.org/10.1016/j.envsoft.2012.01.011)
- Laiolo P, Tella JL (2006) Landscape bioacoustics allow detection of the effects of habitat patchiness on population structure. *Ecology* 87:1203–1214
- Landguth EL, Cushman SA, Schwartz MK et al (2010) Quantifying the lag time to detect barriers in landscape genetics. *Mol Ecol* 19:4179–4191
- Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Mol Ecol* 19:3038–3051. doi:[10.1111/j.1365-294X.2010.04688.x](https://doi.org/10.1111/j.1365-294X.2010.04688.x)
- Martin TG, Burgman MA, Fidler F et al (2012) Eliciting expert knowledge in conservation science. *Conserv Biol* 26:29–38. doi:[10.1111/j.1523-1739.2011.01806.x](https://doi.org/10.1111/j.1523-1739.2011.01806.x)
- Mayer C, Schiegg K, Pasinelli G (2009) Patchy population structure in a short-distance migrant: evidence from genetic and demographic data. *Mol Ecol* 18:2353–2364. doi:[10.1111/j.1365-294X.2009.04200.x](https://doi.org/10.1111/j.1365-294X.2009.04200.x)
- McKelvey KS, Schwartz MK (2005) DROPOUT: a program to identify problem loci and samples for noninvasive genetic samples in a capture-mark-recapture framework. *Mol Ecol Notes* 5:716–718
- McRae BH, Beier P, Dewald LE et al (2005) Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American puma. *Mol Ecol* 14:1965–1977. doi:[10.1111/j.1365-294x.2005.02571.x](https://doi.org/10.1111/j.1365-294x.2005.02571.x)
- Meirmans PG (2012) The trouble with isolation by distance. *Mol Ecol* 21:2839–2846. doi:[10.1111/j.1365-294X.2012.05578.x](https://doi.org/10.1111/j.1365-294X.2012.05578.x)
- Moritz C (1994) Defining “evolutionarily significant units” for conservation. *Trends Ecol Evol* 9:373–375
- Murphy MA, Evans JS, Cushman SA, Storer A (2008) Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. *Ecography (Cop)* 31:685–697. doi:[10.1111/j.1600-0587.2008.05428.x](https://doi.org/10.1111/j.1600-0587.2008.05428.x)
- Nychka D, Furrer R, Paige J, Sain S (2015) Fields: tools for spatial data. National Center for Atmospheric Research, Boulder. <http://cran.r-project.org/package=fields>
- Olsen MT, Andersen LW, Dietz R et al (2014) Integrating genetic data and population viability analyses for the identification of harbour seal (*Phoca vitulina*) populations and management units. *Mol Ecol* 23:815–831. doi:[10.1111/mec.12644](https://doi.org/10.1111/mec.12644)
- Oyler-McCance SJ, St John J (2010) Characterization of small microsatellite loci for use in non invasive sampling studies of Gunnison Sage-grouse (*Centrocercus minimus*). *Conserv Genet Resour* 2:17–20. doi:[10.1007/s12686-009-9122-8](https://doi.org/10.1007/s12686-009-9122-8)
- Oyler-McCance SJ, Taylor SE, Quinn TW (2005) A multilocus population genetic survey of the greater sage-grouse across their range. *Mol Ecol* 14:1293–1310
- Palsbøll PJ, Bérubé M, Allendorf FW (2007) Identification of management units using population genetic data. *Trends Ecol Evol* 22:11–16. doi:[10.1016/j.tree.2006.09.003](https://doi.org/10.1016/j.tree.2006.09.003)
- Pearce JL, Cherry KMD et al (2001) Incorporating expert opinion and fine-scale vegetation mapping into statistical models of faunal

- distribution. J Appl Ecol 38:412–424. doi:[10.1046/j.1365-2664.2001.00608.x](https://doi.org/10.1046/j.1365-2664.2001.00608.x)
- Pearse DE, Crandall KA (2004) Beyond F-ST: analysis of population genetic data for conservation. Conserv Genet 5:585–602
- Plummer M, Best N, Cowles K, Vines K (2006) CODA: convergence diagnosis and output analysis for MCMC. R News 6:7–11
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
- Ribe R, Morganti R, Hulse D, Shull R (1998) A management driven investigation of landscape patterns of northern spotted owl nesting territories in the high cascades of Oregon. Landsc Ecol 13:1–13. doi:[10.1023/A:1007976931500](https://doi.org/10.1023/A:1007976931500)
- Robertson A, Hill WG (1984) Deviations from Hardy-Weinberg proportions: sampling variances and use in estimation of inbreeding coefficients. Genetics 107:713–718
- Rodriguez-Ramilo ST, Wang JL (2012) The effect of close relatives on unsupervised Bayesian clustering algorithms in population genetic structure analysis. Mol Ecol Resour 12:873–884. doi:[10.1111/j.1755-0998.2012.03156.x](https://doi.org/10.1111/j.1755-0998.2012.03156.x)
- Rosenberg NA, Mahajan S, Ramachandran S et al (2005) Clines, clusters, and the effect of study design on the inference of human population structure. PLoS Genet 1:e70. doi:[10.1371/journal.pgen.0010070](https://doi.org/10.1371/journal.pgen.0010070)
- Row JR, Blouin-Demers G, Loughheed SC (2010) Habitat distribution influences dispersal and fine-scale genetic population structure of eastern foxsnakes (*Mintonius gloydi*) across a fragmented landscape. Mol Ecol 19:5157–5171. doi:[10.1111/j.1365-294X.2010.04872.x](https://doi.org/10.1111/j.1365-294X.2010.04872.x)
- Row JR, Oyler-McCance SJ, Fike JA et al (2015) Landscape characteristics influencing the genetic structure of greater sage-grouse within the stronghold of their range: a holistic modeling approach. Ecol Evol. doi:[10.1002/ece3.1479](https://doi.org/10.1002/ece3.1479)
- Runge MC (2011) An introduction to adaptive management for threatened and endangered species. J Fish Wildl Manag 2:220–233. doi:[10.3996/082011-JFWM-045](https://doi.org/10.3996/082011-JFWM-045)
- Safner T, Miller MP, McRae BH et al (2011) Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. Int J Mol Sci 12:865–889. doi:[10.3390/ijms12020865](https://doi.org/10.3390/ijms12020865)
- Schregel J, Kopatz A, Hagen SB et al (2012) Limited gene flow among brown bear populations in far Northern Europe? Genetic analysis of the east-west border population in the Pasvik Valley. Mol Ecol 21:3474–3488. doi:[10.1111/j.1365-294X.2012.05631.x](https://doi.org/10.1111/j.1365-294X.2012.05631.x)
- Schulwitz S, Bedrosian B, Johnson JA (2014) Low neutral genetic diversity in isolated greater sage-grouse (*Centrocercus urophasianus*) populations in northwest Wyoming. Condor 116:560–573. doi:[10.1650/CONDOR-14-54.1](https://doi.org/10.1650/CONDOR-14-54.1)
- Segelbacher G, Paxton RJ, Steinbruck G et al (2000) Characterization of microsatellites in capercaillie *Tetrao urogallus* (AVES). Mol Ecol 9:1934–1935. doi:[10.1046/j.1365-294x.2000.0090111934.x](https://doi.org/10.1046/j.1365-294x.2000.0090111934.x)
- Serrouya R, Paetkau D, McLellan BN et al (2012) Population size and major valleys explain microsatellite variation better than taxonomic units for caribou in western Canada. Mol Ecol 21:2588–2601. doi:[10.1111/j.1365-294X.2012.05570.x](https://doi.org/10.1111/j.1365-294X.2012.05570.x)
- Spear SF, Balkenhol N, Fortin MJ et al (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Mol Ecol 19:3576–3591. doi:[10.1111/j.1365-294X.2010.04657.x](https://doi.org/10.1111/j.1365-294X.2010.04657.x)
- Taylor SE, Oyler-McCance SJ, Quinn TW (2003) Isolation and characterization of microsatellite loci in greater sage-grouse (*Centrocercus urophasianus*). Mol Ecol Notes 3:262–264
- Walker BL, Naugle DE, Doherty KE (2007) Greater sage-grouse population response to energy development and habitat loss. J Wildl Manag 71:2644–2654
- Walsh DP, White GC, Remington TE, Bowden DC (2004) Evaluation of the lek-count index for greater sage-grouse. Wildl Soc Bull 32:56–68
- Walters CJ (1986) Adaptive management of renewable resources. Macmillan, New York
- Wang YH, Yang KC, Bridgman CL, Lin LK (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. Landsc Ecol 23:989–1000. doi:[10.1007/s10980-008-9262-3](https://doi.org/10.1007/s10980-008-9262-3)
- Waples RS, England PR (2011) Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. Genetics 189:633–644. doi:[10.1534/genetics.111.132233](https://doi.org/10.1534/genetics.111.132233)
- Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol 15:1419–1439
- Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191
- Winter DJ (2012) mmod: an R library for the calculation of population differentiation statistics. Mol Ecol Resour 12:1158–1160
- Woltmann S, Kreiser BR, Sherry TW (2012) Fine-scale genetic population structure of an understory rainforest bird in Costa Rica. Conserv Genet 13:925–935. doi:[10.1007/s10592-012-0341-2](https://doi.org/10.1007/s10592-012-0341-2)
- Wood SN (2006) Generalized additive models: an introduction with R. Chapman and Hall, New York, USA