Re-Evaluation of Yellowstone Grizzly Bear Population Dynamics not Supported by Empirical Data: Response to Doak & Cutler

Frank T. van Manen1, Michael R. Ebinger2,3, Mark A. Haroldson1, Richard B. Harris2,4, Megan D. Higgs5, Steve Cherry3, Gary C. White6, & Charles C. Schwartz1

1 U. S. Geological Survey, Northern Rocky Mountain Science Center, Interagency Grizzly Bear Study Team, 2327 University Way, Suite 2, Bozeman, MT 59715, USA
2 College of Forestry and Conservation, University of Montana, University Hall, Room 309, Missoula, MT 59812, USA
3 Department of Ecology, Montana State University, P.O. Box 173460, Bozeman, MT 59717, USA
4 Washington Department of Fish and Wildlife, Game Division, Wildlife Program, 600 Capital Way North, Olympia, WA 98501, USA
5 Department of Mathematical Sciences, Montana State University, P.O. Box 172400, Bozeman, MT 59717, USA
6 Department of Fish, Wildlife, and Conservation Biology, Colorado State University, 239 Wagar, Fort Collins, CO 80523, USA

Keywords
Chao estimator; demographic projection; Doak & Cutler; grizzly bears; population trend; rebuttal; scientific review; Ursus arctos; Yellowstone.

Correspondence
Frank T. van Manen, U. S. Geological Survey, Northern Rocky Mountain Science Center, Interagency Grizzly Bear Study Team, 2327 University Way, Suite 2, Bozeman, MT 59715, USA.
Tel: 406-994-5043, fax: 406-994-6556. E-mail: fvanmanen@usgs.gov

Received
4 October 2013
Accepted
10 February 2014

doi: 10.1111/j.1469-8131.2014.06555.x

Abstract
Doak and Cutler critiqued methods used by the Interagency Grizzly Bear Study Team (IGBST) to estimate grizzly bear population size and trend in the Greater Yellowstone Ecosystem. Here, we focus on the premise, implementation, and interpretation of simulations they used to support their arguments. They argued that population increases documented by IGBST based on females with cubs-of-the-year were an artifact of increased search effort. However, we demonstrate their simulations were neither reflective of the true observation process nor did their results provide statistical support for their conclusion. They further argued that survival and reproductive senescence should be incorporated into population projections, but we demonstrate their choice of extreme mortality risk beyond age 20 and incompatible baseline fecundity led to erroneous conclusions. The conclusions of Doak and Cutler are unsubstantiated when placed within the context of a thorough understanding of the data, study system, and previous research findings and publications.

Introduction
The Yellowstone grizzly bear (Ursus arctos) was listed as a threatened species under the U.S. Endangered Species Act in 1975 (Federal Register 40 FR:31734-31736). Since listing, recovery efforts have focused on increasing population size, improving habitat security, managing bear mortalities, and reducing human–bear conflicts. The Interagency Grizzly Bear Study Team (IGBST), a research partnership among federal, state, and tribal agencies was formed by the Secretary of the Interior in 1973. Since its inception, IGBST has played an important supportive role in the management of Greater Yellowstone Ecosystem (GYE) grizzly bears by producing science that informs management and policy decisions (Lynch et al. 2008:826). The grizzly bear population began to recover in the mid 1980s and increased in numbers from approximately 200–350 bears (Eberhardt & Knight 1996) to >600 in 2012 (Haroldson et al. 2013). As the population increased, range expansion occurred that has continued to date (Bjornlie et al. 2013).

We recognize the concerns expressed in the critique by Doak & Cutler (2013; hereafter D&C) of methods used by IGBST to estimate grizzly bear population size and trend in the GYE. Here, we respond to the work of D&C because of the vital role that scientific debate plays in gaining reliable knowledge (Banobi et al. 2011). D&C focused mainly on two claims: (1) increases in grizzly bear population estimates from 1983 to 2001 can be attributed to factors other than actual increases in population size,
primarily increased observation effort, and sightability of female grizzly bears with cubs-of-the-year (FCOV); and (2) estimates of population growth were biased high because survival senescence and reproductive senescence were not properly accounted for in demographic analyses. We examine both claims and demonstrate the conclusions of D&C are unfounded. We present our main arguments and supporting evidence here and provide technical details in Supporting Information.

**Observation process for population size and trend**

D&C “mechanistically” modeled the annual observation frequency of individual females with cubs-of-the-year (FCOV), providing data with which to estimate FCOV population size using the Chao2 estimator (Supporting Information). Using a constant population size of 70 animals during 1983–2001, but varying observation process (e.g., number of flight observation hours), D&C argue their simulations show that a positive trend (i.e., positive slope) similar to the empirical trend in Chao2 reported by IGBST can be observed simply as an artifact of increased survey effort and sightability, even when population size remains constant. Whereas their simulation approach is potentially useful for illustration, within the context of the critique their conclusions hinge on several unrealistic and unjustified assumptions. Additionally, although D&C mechanistically generated the “counting process” of bears, they did not evaluate the estimated trends using the simulated Chao2 estimates in the same way the empirical Chao2 trend is estimated by the IGBST. Although we believe this to be unintentional, it provided a misleading interpretation of the simulation results.

**Search effort**

We identified three major issues with D&C’s simulations of the observation process. First, they explicitly gave each FCOV a probability of being counted during every flight hour, regardless of the spatial location of the flight (Supporting Information). For this to be realistic, observation flights would have to cover all occupied range in the GYE during each flight hour, which is not possible as evidenced by the established protocol of using 48 separate flight observation areas. The direct consequence of this in D&C’s simulations is that the numbers of simulated FCOV sighted once (f1) and twice (f2) quickly stabilize as flight hours increase and are relatively low compared with the number of unique FCOV observed (m) because many are sighted more than twice (Figure 1). Accordingly, a Chao2 estimate for a simulated population of known N is determined essentially by m, rather than the frequency of sightings (f1 and f2), which, unsurprisingly, results in a direct correspondence between flight hours and Chao2 estimates (Figure 1).

Second, the conclusion of D&C that the increasing trend in Chao2 was driven by increased search effort was based on visual interpretation of the trend in the annual expected values (i.e., the annual mean FCOV from all simulations) over time. Whereas mathematically correct, a meaningful comparison should focus on the estimated trend for each realization of simulated data, calculated using methods actually deployed by the IGBST for the empirical data. Accordingly, we applied linear regression to the natural log-transformed estimates produced from 1,000 realizations of 25-year time series under D&C’s simulation method. We thus constructed a distribution of trend estimates under the assumptions of the simulation model with which we could compare the estimate obtained in the same way from empirical data. Of D&C’s realizations, 99.9% resulted in estimated slopes less than the empirically estimated slope (Figure 2A). Therefore, given the known, constant population of 70 individuals that D&C simulated to support their argument regarding search effort (figure 5 in D&C), we found little evidence that increased search effort alone would result in detection of the population trend documented by IGBST based on empirical data, even if we accept their unrealistic assumption that all bears are potentially observable during each flight hour, which we do not.
Finally, occupied grizzly bear range in the GYE approximately doubled from the late 1970s to 2010 (Blanchard et al. 1992; Schwartz et al. 2002, 2006a; Bjorndal et al. 2013) and observation areas outside the recovery zone were added in 1989, 1998, and 2007. Of course, a larger survey area required additional flight hours (Figure S1) but we argue that it is search intensity (effort/area surveyed) that matters because the probability of seeing a $F_{COV}$ in these expansion areas was very small, if not zero, prior to expanding the survey area. To demonstrate the importance of accounting for increasing survey area, we adjusted annual flight hours based on total area surveyed and repeated the simulations with which D&C generated their figure 5 (Supporting Information). Of 1,000 realizations, all had estimated slopes from linear models lower than that from the empirical data (Figure 2B). Our findings do not support D&C’s conclusion that “much of the apparent increasing trend in bear numbers during this time period can be parsimoniously explained as a result of increasing search effort.” (p. 4).

Sightability

D&C claim that sightability of $F_{COV}$ has potentially increased over time, which would inflate Chao2 estimates. Indeed, under these assumptions, their simulations show estimates of Chao2 increase with greater mean sightability of individuals. However, details of these relationships are known and were previously reported in Keating et al. (2002) and Cherry et al. (2007). D&C provide no quantitative, empirical evidence that mean sightability of individuals has increased over time. Instead, they speculate that changes known to be occurring in the ecosystem that may affect grizzly bears (e.g., use of army cutworm moth [Euxoa auxiliaris] aggregation sites and increased researcher knowledge of those sites; decline of cutthroat trout [Oncorhynchus clarkii] and whitebark pine [Pinus albicaulis]; reintroduction of wolves [Canis lupus]) must have led to a progressive increase in mean sightability of grizzly bears. We do not disagree that sightability may increase or decrease due to these factors. However, during telemetry flights, attempts to observe each radio-collared $F_{COV}$ (that are well distributed throughout the GYE; Schwartz et al. 2006b:12) are made and such visual observations provide no empirical justification to conclude sightability increased over time (Figure S2).

Finally, D&C cite Link (2003) and argue that population size is nonidentifiable for estimators using observation frequencies when frequencies are small and variance in sightability is high. Link (2003) was making the point that mark-recapture data cannot identify the true underlying distribution that generated the data. However, Link (2003) did not show that estimators could not approximate true population size. Cherry et al. (2007:198) referenced this issue and clearly demonstrated that Chao2 estimates asymptotically converge on truth (based on simulated datasets with known properties) with increasing effort. Unlike three other nonparametric estimators, however, Chao2 consistently underestimated population size at all levels of effort (Cherry et al. 2007:206).
**Demographic analyses**

The authors argue that past IGBST estimates of population growth were biased high because survival senescence and reproductive senescence were not properly accounted for in demographic analyses. Although modeling senescence in long-lived animals remains a substantial challenge, primarily because of individual heterogeneity (Reiske et al. 2010), we do not disagree with D&C that these parameters deserve consideration. However, the underlying assumptions of survival and reproductive senescence that D&C used resulted in misinterpretation of effect size. We again address our main concerns here and provide additional technical details in Supporting Information.

**Incorporating survival senescence in population projections**

The assumption of survival senescence in D&C’s simulations fits the “sudden” senescence pattern (Nussey et al. 2011), however, our empirical data indicate their senescence function is too abrupt and too early (age 20) in the potential lifespan of GYE grizzly bears. To examine the influence of survival senescence on population projections, we compared models with survival as parameterized by D&C with models using empirical estimates of survival. We used data for 1983–2011 to estimate age-specific survival (data provided in Supporting Information). Our procedures followed Haroldson et al. (2006) but instead of generating a single mean survival rate for the entire period, we estimated age-specific survival of females using age and age² covariates (Figure 3, solid black line), similar to the analyses of Johnson et al. (2004). We used these age-specific, empirical parameter estimates to repeat the survival senescence population projections of D&C. In contrast to D&C’s reduction of mean annual population growth compared with the reference model of no survival senescence (Figure 3, dashed line) for high (1.12%) and low (1.03%) survival scenarios (Tables 1A and 1B), we observed a 0.07% reduction (high survival, Table 1A) and a 0.41% increase of population growth (low survival, Table 1B). These findings suggest little contribution of age-specific survival on population projections and confirm those of Haroldson et al. (2006), who found age class did not explain much variation beyond other covariates in the model. We do not dismiss D&C’s point that survival senescence may become relevant as this population ages; however, information gained from incorporating age-specific survival should be weighed against increased model complexity and, so far, we have found no support for incorporating age-specific survival.

**Incorporating reproductive senescence in population projections**

Whereas D&C clearly state they used “an overall lower reproductive rate” (D&C Supporting Information, p. 8) for their population modeled under reproductive senescence, they do not provide convincing evidence that this
Table 1  Population size, mean annual population growth (λ), and percent decline in λ for female grizzly bears in the Greater Yellowstone Ecosystem (GYE) based on simulations of Doak & Cutler (2013). We compared influence of survival senescence as implemented by Doak & Cutler with Interagency Grizzly Bear Study Team (IGBST) estimates of age-specific survival. Doak & Cutler used fecundity (m) for 1983–2001 from Harris et al. (2007) as their reference. We created a second reference model based on Schwartz et al. (2003: model D) to demonstrate the effect of having a compatible reference dataset to assess influence of age-specific fecundity; this model is not representative of the GYE (see text and Supporting Information for further explanation). Population projections are based on survival estimates in which bears with unknown fates were censored (A) or assumed dead (B; Haroldson et al. 2006), which Doak & Cutler refer to as “high” and “low” survival scenarios, respectively. We use Doak & Cutler’s terminology of survival and reproductive senescence but note they are based on age-specific survival and age-specific fecundity. Data for all simulation scenarios are provided in Supporting Information.

<table>
<thead>
<tr>
<th>Population projection scenarios</th>
<th>m in reference model based on Harris et al. (2007) (N_{2001} = 0.322)^a</th>
<th>m in reference model based on Schwartz et al. (2003) (N_{2001} = 0.2635)^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. High survival scenario</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference model</td>
<td>219 1.073 0</td>
<td>148 1.051 0</td>
</tr>
<tr>
<td>Survival senescence^c</td>
<td>176 1.061 1.12</td>
<td>110 1.035 1.55</td>
</tr>
<tr>
<td>Survival senescence (IGBST data)</td>
<td>216 1.073 0.07</td>
<td>143 1.050 0.18</td>
</tr>
<tr>
<td>Reproductive senescence^d,h</td>
<td>136 1.047 2.46</td>
<td>136 1.047 0.43</td>
</tr>
<tr>
<td>Survival senescence^c and reproductive senescence</td>
<td>110 1.035 3.56</td>
<td>110 1.035 1.55</td>
</tr>
<tr>
<td>Survival senescence (IGBST data) and reproductive senescence</td>
<td>138 1.048 2.39</td>
<td>138 1.048 0.36</td>
</tr>
<tr>
<td>B. Low survival scenario</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference model</td>
<td>131 1.045 0</td>
<td>88 1.023 0</td>
</tr>
<tr>
<td>Survival senescence^c</td>
<td>108 1.034 1.03</td>
<td>67 1.008 1.44</td>
</tr>
<tr>
<td>Survival senescence (IGBST data)</td>
<td>142 1.049 –0.41^i</td>
<td>93 1.026 –0.32^i</td>
</tr>
<tr>
<td>Reproductive senescence^d,h</td>
<td>81 1.019 2.51</td>
<td>81 1.019 0.44</td>
</tr>
<tr>
<td>Survival senescence^c and reproductive senescence</td>
<td>66 1.008 3.55</td>
<td>66 1.008 1.49</td>
</tr>
<tr>
<td>Survival senescence (IGBST data) and reproductive senescence</td>
<td>90 1.024 1.96</td>
<td>90 1.024 –0.12^i</td>
</tr>
</tbody>
</table>

^a As implemented by Doak & Cutler (2013) in their simulations.

^b As implemented by authors for additional simulations with different m, for reference model; based on R code in Doak & Cutler (2013) Supporting Information.


^d Geometric mean of annual population growth (N_{t+1}/N_t).

^e Senescence as modeled by Doak & Cutler (2013) based on parameter estimates for age and age^2 from Johnson et al.’s (2004) proportional hazards model (Figure 3, solid gray line).

^f Age-specific survival based on IGBST (IGBST data) for high survival scenario.

^g Not included in simulation models of Doak & Cutler (2013).

^h Reproductive senescence modeled as age-specific fecundity based on Schwartz et al. (2003: model D).

^i Negative value indicates population increase compared with reference data (due to greater survival among ages 6–18 years; Figure 3).

was justified. Furthermore, we demonstrate that the different fecundities for the baseline (i.e., no reproductive senescence) and reproductive senescence population projections overwhelmingly drive reported effect sizes in D&C’s reproductive senescence models (Supporting Information). To examine the implications of differences in baseline fecundity, we constrained D&C’s combined reproductive and survival senescence simulations to allow for a reproductive senescence-only scenario, which they did not include. We developed a population-level fecundity estimate based on age-specific fecundity and the stable age distribution used in D&C’s simulations, which resulted in a compatible population-level baseline of m_{y, pop} = 0.2635 (Supporting Information). We duplicated D&C’s simulations using this baseline value and observed only a 0.43% (high survival, Table 1A) and 0.44% (low survival, Table 1B) reduction in population growth during 1983–2001 upon inclusion of reproductive senescence, compared with a 2.46% and 2.51% reduction, respectively, using the higher, incompatible baseline of m_{y, pop} = 0.322 (Table 1). The lack of a strong reproductive senescence effect at the population level confirms findings from our earlier work (Supporting Information). Clearly, the choice of fecundity baseline changes interpretation regarding impacts of reproductive senescence.
Incorporating survival and reproductive senescence in population projections

Table 1 shows how D&C’s choice of assumptions for survival senescence and incompatible baseline for reproductive senescence influenced their simulation results. Using the compatible fecundity reference data for inference, reductions in population growth due to D&C’s combined reproductive and survival senescence are 1.55% (Table 1A) and 1.49% (Table 1B); these reductions are the same as those for survival senescence only (Tables 1A and 1B) because when both senescence functions are combined and an appropriate fecundity reference is used, the extreme survival senescence function of D&C (i.e., almost zero survival past age 20) overrides any reproductive senescence effects. We point this out because D&C implied a strong additive effect from reproductive senescence in their figure 6, but this is not possible because their simulations allow no females to survive long enough to reach reproductive senescence age. The additive effect from reproductive senescence was in fact a direct function of their fecundity reference, rather than senescence itself.

Finally, when using the compatible reference for fecundity, IGBST estimates of age-specific survival resulted in only a 0.18% reduction (Table 1A) and a 0.32% increase (Table 1B) in population growth with a slightly additive effect when combined with D&C’s reproductive senescence (0.36% [Table 1A] reduction and 0.12% [Table 1B] increase of population growth, respectively). These changes are not biologically relevant at the population level; thus these simulations do not support the importance of incorporating survival and reproductive senescence (however, see Rebke et al. [2010] regarding potential role at individual level). We conclude that D&C’s choice of an extreme survival senescence function, combined with inappropriate comparison among fecundity schedules to model reproductive senescence, substantially influenced their simulation results, leading to conclusions that are not supported by empirical evidence.

Discussion

D&C conclude that “we actually know very little about the past trends of this population” and that “with rapidly accelerating impacts, the flattening Chao2 estimates over the last decade, even as search effort has continued to increase, are consistent with a population that may now be, in fact, declining.” (p. 9). D&C further assert that the comprehensive population studies of IGBST (e.g., Keating et al. 2002; Schwartz et al. 2006b; Harris et al. 2006, 2007; IGBST 2006, 2012; Cherry et al. 2007) show a “lack of
attention to basic issues of wildlife data analysis (accounting for observation effort and realistic treatment of life history patterns)” and “are likely to have resulted in misunderstandings of the data collected, systematic bias in the inferences about the dynamics of this population, and overconfidence in apparent trends.” (p. 9). As we demonstrate here, these statements are not supported by their own analyses, are based on selective and inappropriate use of existing data, and fail to account for the broader context provided by available data and other published information.

We recognize that no two investigators are likely to make the exact same choices in dealing with complex data. It can be tempting, from afar, to imbue such choices with intent when it merely results from the inevitable differences in choice made by equally qualified and objective investigators. This is why IGBST has long favored a team approach. Current and past demographic analyses include about a dozen scientists and methods and results are critically evaluated, often involving intensive debates and exhaustive consideration of alternatives. Moreover, IGBST explicitly evaluates stochasticity in the modeling process whenever possible. When faced with the inevitable uncertainty associated with parameter estimation, the study team evaluated such relationships, chose to avoid overestimation of population size and mortality limits (Supporting Information), and provided complete transparency of the process (e.g., IGBST 2012:20).

Decisions by federal, state, and tribal land managers regarding Yellowstone grizzly bears are not based on any single data source or analysis but rather on the totality of available data, rigorous analyses, and knowledge of the system by a team of experienced scientists. To that regard, we point out that three independent datasets (Chao2 based on counts of unique $F_{\text{com}}$; population projections based on demographic analyses of intensive known-fate data; mark-resight analysis of $F_{\text{com}}$ from systematic observation flights only) have shown very similar trends of population growth (Harris et al. 2006, 2007; Haroldson 2011; Haroldson et al. 2013; Higgs et al. 2013). Moreover, our interpretation of a positive population trend since 1983 (with slowing of growth during the last decade; IGBST 2012) is supported by additional indicators. First, our systematic observation flights within the Recovery Zone (excluding army cutworm feeding sites where bears are more observable), where search effort has remained constant or even declined slightly, show observation rates of bears increased from <1/hour in the mid 1990s to ~3/hour in late 2000s (Figure 4). Second, the proportion of unmarked bears in the capture sample (i.e., first capture) remained constant and high (50–70%) while total individuals captured/year increased (Figure 5). Finally, there is substantial evidence that stable or positive population trends, or high density, are a prerequisite for range expansion among animal populations, as has been demonstrated for species ranging from butterflies in the United Kingdom (Mair et al. 2014) to brown bears in Europe (Swenson et al. 1998). During the 1970s and into the 1980s, grizzly bear range in the GYE was likely contained within the recovery zone (23,828 km$^2$) but expanded to 34,416 km$^2$ by the end of the 1990s (Schwartz et al. 2002) and 50,280 km$^2$ in 2010 (Bjornlie et al. 2013), the latter study documenting a 38% increase during 2004–2010 alone. The combined inference from our analyses presented here, our ongoing monitoring efforts, and complementary data sources provide substantial evidence that contradicts D&C’s assertion the GYE grizzly bear population is likely declining.

We conclude the inferences of D&C are unfounded when placed within the context of a thorough understanding of the data, the study system, and previous research findings and publications. The IGBST and collaborating scientists have a long history of publishing study results, making reports publicly available, presenting background and justification for analysis decisions, and providing careful interpretation of estimated population trends and demographic results. Careful reading of those papers and reports addresses most, if not all, of the claims presented by D&C. The task of IGBST is to use the best available science to ensure that federal and state agencies have objective and reliable data upon which to base their policy and management decisions (e.g., Schwartz et al. 2006b). Scientific debate is critical to the scientific process and can lead to important new insights. However, based on our examination, the simulations and analyses that formed the foundation of D&C’s
critique led to conclusions that are not supported by empirical data.

Acknowledgments

We thank J. D. Clark for review of this manuscript as part of the U.S. Geological Survey Fundamental Science Practices and J. L. Kershner for helpful comments on an earlier draft. We appreciate the insightful comments from those involved in the journal review process of this manuscript. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher’s web site:

Fig. S1.
Fig. S2.
Fig. S3.
Fig. S4.

References

Interagency Grizzly Bear Study Team. (2012) Updating and evaluating approaches to estimate population size and sustainable mortality limits for grizzly bears in the Greater Yellowstone Ecosystem. Interagency Grizzly Bear Study Team, U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT.


