

RE-EVALUATION OF YELLOWSTONE GRIZZLY BEAR POPULATION DYNAMICS PROVIDES FLAWED INSIGHTS: RESPONSE TO DOAK & CUTLER

Frank T. van Manen^{1,*}, Michael R. Ebinger², Mark A. Haroldson¹, Richard B. Harris³,
Megan D. Higgs⁴, Steve Cherry⁴, Gary C. White⁵, and Charles C. Schwartz¹

¹U. S. Geological Survey, Northern Rocky Mountain Science Center, Interagency Grizzly Bear Study Team, 2327 University Way, Suite 2, Bozeman, MT 59715, USA

²University of Montana, College of Forestry and Conservation, University Hall, Room 309, Missoula, MT 59812, USA

³Washington Department of Fish and Wildlife, Game Division, Wildlife Program, Washington Department of Fish and Wildlife, 600 Capital Way North, Olympia, WA 98501, USA

⁴Montana State University, Department of Mathematical Sciences, Montana State University, P.O. Box 172400, Bozeman, MT 59717, USA

⁵Department of Fish, Wildlife, and Conservation Biology, 239 Wagar, Colorado State University, Fort Collins, CO 80523, USA

*Corresponding author email: franmanen@usgs.gov

SUPPLEMENTARY INFORMATION

Doak & Cutler (2013) critiqued methods used to estimate grizzly bear population size and trend in the Greater Yellowstone Ecosystem (GYE) by the Interagency Grizzly Bear Study Team (IGBST), a research partnership among federal and state agencies that was formed by the Secretary of the Interior in 1973. Peer review and critique are essential to scientific progress, and the IGBST has a long history of publishing study results, making reports publicly available, presenting background and justification for analysis decisions, and providing careful interpretation of population trend and demographic results. Doak & Cutler's critique rests primarily on two claims: 1) the increase in grizzly bear population estimates from 1983 to 2001 can be attributed to factors other than population size, primarily increased observation effort and sightability of grizzly bears; and 2) estimates of population growth were biased high because reproductive and survival senescence were not properly accounted for in demographic analyses. Here, we discuss both of these claims and demonstrate the main conclusions made by Doak &

Cutler are unfounded when placed within the context of a thorough understanding of the data, the study system, and previous research findings and publications.

Population size and trend

IGBST has used counts of females with cubs-of-the-year (F_{COY}) in combination with the Chao2 estimator to estimate the number of F_{COY} in the study area and monitor population trend (IGBST 2012). First, sightings of F_{COY} from aerial surveys and ground observations are clustered into those from unique bears (unduplicated F_{COY}) using a “rule set” primarily based on spatial, temporal, and litter size criteria (Knight *et al.* 1995). Given the size of the study area, observation flights are partitioned into bear observation units, each one typically involving approximately 2 hours of flight time to completely survey. Ground observations are obtained by convenience sightings and are a source of data completely separate from the survey flights (IGBST 2012). Using the frequencies of sightings of individual F_{COY} obtained from the rule set, an annual estimate of the total number of F_{COY} in the study area is calculated using the Chao2 estimator, which is robust to heterogeneity among observations (Chao 1989, Keating *et al.* 2002, Cherry *et al.* 2007).

Doak & Cutler claim they simulated the “observation process” as a function of number of flight hours and then argue, based on the simulation results, that increases in bear numbers during the 1983–2001 time period can be explained simply by increases in search effort and sightability. We discuss the role of search effort and sightability in the next section, and provide additional data to support our arguments.

Search Effort.—Doak & Cutler based much of their argument on search effort, which they define solely by number of flight hours, although survey flights represented only about a third of all F_{COY} sightings during 1986–2010. Ground observations and aerial observations by

other researchers also are important sources of data for counts of F_{COY} . We were unable to detect a change in the proportion of these other observations during 1986–2010 ($\hat{\beta} = -0.26$, $P = 0.516$, $R^2 = 0.02$). Although we have no measure of effort, these observations are reported almost exclusively by agency personnel and we have no indication the reporting rate has changed over time.

Doak & Cutler’s simulation of the number of sightings per F_{COY} , for a known population of F_{COY} (i.e., 30, 50, or 70), was based on set of independent Bernoulli trials $B_i(n, p_i)$, where i represents individual bears ($i = 1, \dots, 70$) and p_i is the probability of observing bear i in any of the n flight hours (rounded to nearest integer) for the given year over the entire study area (~20,000–45,000 km² for 1986–2010). The outcome of the n trials represents the number of observations (“successes”) of an individual for a given year, with bear-specific probabilities drawn from a beta distribution based on data from Keating *et al.* (2002); these observations served as the input data for the Chao2 estimate. For example, in their simulated population of 70 bears and 35 flight hours (i.e., 1988, lowest number of flight hours), and using the mean sightability of 0.00868 in 1,000 simulations, the mean number of bears sighted at least once was obtained as the number of successes in 35 trials, or 18 out of 70 bears (26%). For the maximum of 199 flight hours (i.e., 2008), a mean of 58 out of 70 bears (83%) would be sighted at least once. We find three major flaws with this approach.

First, we disagree that Doak & Cutler’s method of simulation takes “a more mechanistic approach that explicitly simulates the observation process.” (p. 3). Their approach makes sweeping assumptions about the method of data collection and the study system that do not represent reality. In their simulations Doak & Cutler explicitly allow each F_{COY} to be sampled every flight hour, regardless of the spatial location of the flight. For this to be realistic, we

would need to envision observation flights covering the entire area of occupied range during each flight hour, which is clearly impossible. Even if it were possible, most F_{COY} would be too far away to be seen during most flight hours so their probability of being sighted should be zero for most flight hours. In recent work (Higgs *et al.* 2013), we used a mark-resight estimator based solely on counts of marked (radio collar) and unmarked F_{COY} from the standardized observation flights. For that estimator, we have assumed an individual F_{COY} can only be sighted one time per flight survey and there are only two flights over each observation area per year. This assumption is supported by data on sightings of radio-collared F_{COY} during survey flights without aid of telemetry (Higgs *et al.* 2013). Because Doak & Cutler assumed that all bears in the entire study area are available to be sighted each flight hour, in their simulations the number of F_{COY} sighted once ($f1$) or twice ($f2$) quickly stabilize as flight hours increase and are relatively low compared with m because many are sighted more than twice (Fig. S1). The consequence is that the essence of the Chao2 estimator disappears because the estimate for the simulated population of known N is determined essentially by the number of unique F_{COY} observed (m), which, unsurprisingly, shows a direct correspondence between flight hours and Chao2 estimates (Fig. S1). However, this simplistic relationship to flight hours is the foundation of arguments made by Doak & Cutler regarding Chao2 estimates and trends.

Second, Doak & Cutler base their conclusion that the increasing trend in Chao2 was driven by increases in search effort on a simple visual comparison of their simulation results with the empirical Chao2 data. IGBST has always carefully compared simulated data characteristics with those of empirical data (e.g., Cherry *et al.* 2007), which Doak & Cutler did not do. Accordingly, we applied linear regression to each of 1,000 realizations of data produced from Doak & Cutler's simulation method of the 25-year time series (not 20 years, as Doak & Cutler

mistakenly described) to obtain a distribution of slope (trend) estimates under their model. We then compared the slope of our empirical estimates of $\ln(\text{Chao2})$ for 1986–2010 ($\hat{\beta} = 0.0404$, $P < 0.001$, $R^2 = 0.61$,; Fig. S2A) with the distribution of slopes based on Doak & Cutler's simulations: 99.9% of Doak & Cutler's realizations resulted in slopes lower than observed from the empirical data (Fig. S2A). This finding suggests it is extremely unlikely that the increasing trend can be accounted for simply due to increasing flight hours, even under the model used by Doak & Cutler that assumes a much stronger relationship between flight hours and sightings than is realistic for the system.

Finally, Doak & Cutler's definition of effort as the number of flight hours fails to account for the increase in area surveyed, with their methods assuming each hour is used to survey the same spatial area. IGBST survey efforts have evolved to account for documented range expansion of this population as reported by Blanchard *et al.* (1992), Schwartz *et al.* (2002, 2006c) and, most recently, by Bjornlie *et al.* (2013), who documented a 38% increase in occupied range between 2004 and 2010 alone. In response to this expansion, bear observation areas outside the recovery zone were added in 1989, 1999, and 2007 based on new records of bear presence (e.g., sightings, conflicts). Because dispersal among grizzly bears is substantially male biased (Proctor *et al.* 2004), such expansion initially involved subadult and young adult males. Thus, establishment of F_{COY} occurred subsequent to expansion of observation flight areas. Indeed, only 24 out of 297 initial sightings of F_{COY} were from observation flights outside the recovery zone during 1986–2010, or an average increase of 1 unique F_{COY} /yr. Of course, a larger survey area required additional flight hours (Fig. S3) but we argue that it is search intensity (effort/area surveyed) that matters because the probability of seeing an F_{COY} in these expansion areas was very small, if not zero, prior to expanding the survey area. In other words,

it is unrealistic to assume the population was actually stable or declining but appeared to be increasing because of search effort expanding to new areas.

To demonstrate this point, we adjusted annual flight hours based on total area surveyed and repeated the simulations with which Doak & Cutler generated their Fig. 5. To account for this spatial component of search effort in a way that is tractable under the binomial framework of the Doak & Cutler simulations, we rescaled the original search hours using the area flown as the adjustment parameter. We used the smallest search area (1988) as the reference category and adjusted flight hours for all other years accordingly. We note that the choice of reference year influences the absolute numbers for flight hours but does not affect relatively change in counts across years. Of all realizations generated by these simulations, 100% had slopes lower than the Chao2 trend of our empirical data (Fig. S2B). In other words, given the known, constant population of 70 individuals, we found no support for the notion that increase in search effort alone, adjusted for area surveyed, would result in detection of a trend as large as documented by IGBST based on empirical data.

We conclude that Doak & Cutler's assertion that IGBST's estimates of grizzly bear population trend were a direct consequence of search effort is based on a substantially flawed simulation of the F_{COY} observation process, and is not even consistent with results from their own simulations particularly after adjusting number of flight hours for increased survey area. Thus, we find no support for their 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.—Doak & Cutler claim that sightability of F_{COY} 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. They also show an increase of Chao2 with decreasing coefficient of variation (CV) of sightability (although they incorrectly state that “rising CV in sightabilities results in increased F_{COY} and Chao2 estimates” [Fig. 3 caption]; the relationship is actually opposite because CV *declined* along the horizontal axis). However, details of these relationships were previously reported and published in Keating *et al.* (2002) and Cherry *et al.* (2007). Thus, their argument hinges on the premise that mean sightability of individuals has increased over time, for which Doak & Cutler provide no quantitative evidence. Instead, they speculate changes known to be occurring in the ecosystem that affect grizzly bears (e.g., greater use of army cutworm moth [*Euxoa auxiliaris*] aggregation sites and increased researcher knowledge of those sites; decline of cutthroat trout [*Oncorhynchus clarki*] and whitebark pine [*Pinus albicaulis*]; reintroduction of wolves [*Canis lupus*]) 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, observations of radio-collared F_{COY} from telemetry flights, which are well distributed throughout the GYE (Schwartz *et al.* 2006d:12), provide no basis for assuming a major shift in sightability (Fig. S4).

Additional Notes.—We find it remarkable that Doak & Cutler failed to mention that Chao2 estimates of F_{COY} in the GYE are known to be biased low and that this bias increases with population size. This underestimation bias has 2 sources: (1) the Chao2 estimator itself and (2) the count of unduplicated F_{COY} obtained from the Knight *et al.* (1995) rule set to which the Chao2 estimator is applied (IGBST 2012). Keating *et al.* (2002) and Cherry *et al.* (2007) explicitly recognized the potential negative bias of the Chao2 estimator and Schwartz *et al.* (2008) demonstrated bias due to the second source. The simulation results of Doak & Cutler are

consistent with these previous reports but not discussed (all estimates of Chao2 are well below the known population size in their Figs. 3–5). The IGBST has long held that this underestimation was appropriate during recovery stages of this population (e.g., Keating *et al.* 2002, Schwartz *et al.* 2006e, Cherry *et al.* 2007). Workshops with external experts during 2011–2012 led to a revision of that philosophy and a desire to correct this underestimation bias (IGBST 2012). This resulted in the development of a mark-resight estimator using a latent multinomial model based on data from the observation flights (Higgs *et al.* 2013). These mark-resight estimates indicate the number of F_{COY} is greater than those based on Chao2, although the trend tracks that of the Chao2.

Finally, Doak & Cutler cite Link (2003) in arguing that population size is nonidentifiable for estimators using observation frequencies when observation 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 the true population size. Cherry *et al.* (2007) actually referenced this issue (p. 198) and clearly demonstrated that estimates were close to truth, although still biased low.

Demographic Analyses

Incorporating Survival Senescence in Population Projections.—Doak & Cutler stated that IGBST’s data for “mean adult female survival estimates of 0.922 and 0.950 would result in 10% or 24% of adult bears surviving to age 30, and between 25 and 42% dying past the age of 20” and provide these surprisingly large percentages of old bears as an argument that survival senescence should be taken into account. However, the calculations of Doak & Cutler are incorrect: using cub survival of 0.640 and yearling survival of 0.817 (Schwartz *et al.* 2006b:27),

these estimates would be 5% surviving to age 30 (i.e., $0.640 \times 0.817 \times 0.922^{28}$), not 10%, for the low survival rate and 12% (i.e., $0.640 \times 0.817 \times 0.950^{28}$) instead of 24% for the high survival rate. That is, Doak & Cutler failed to include survival of cubs and yearlings (i.e., $0.640 \times 0.817 = 0.51$) into their calculations and thus wrongly claimed the estimates of population growth of Harris *et al.* (2006, 2007) required there be twice as many old-aged females as in fact were in those models.

Doak & Cutler incorporated survival senescence into their simulation models by applying estimates from the Anderson-Gill mortality model (proportional hazards model) of Johnson *et al.* (2004) based on covariates for age and age² to age classes of 20 and greater, but ignored all other covariates in the model (i.e., sex, in Yellowstone National Park, number of management actions, distance to development, road density, trail density). This is equivalent to setting values of all other covariates to zero or assuming that all are independent of age. Additionally, the data to develop the age-specific functions were from 1975–1997 (Johnson *et al.* 2004), which included the time period of the late 1970s and early 1980s during which the GYE grizzly bear population remained in decline; this was the primary reason that Schwartz *et al.* (2006d:10–11) used 1983 as the starting year for their analyses. When Doak & Cutler applied these coefficients, it resulted in a biologically unrealistic decrease in mean annual survival of females from 0.949 at age 19 to 0.044 at age 20 and essentially no survival (<0.0012) for ages ≥ 21 (Fig. S5, gray line).

To examine the influence of survival senescence as modeled by Doak & Cutler versus empirical data, we used data for 1983–2011 to estimate age-specific survival. We followed procedures of Haroldson *et al.* (2006) but instead of generating a single mean survival rate for the entire period, we estimated age-specific survival of independent females and estimated coefficients for age and age², similar to the analyses of Johnson *et al.* (2004; Fig. S5, solid black

line). We then used these age-specific, empirical data to repeat the survival senescence population projections of Doak & Cutler. In contrast to Doak & Cutler's reduction of 1.1% population growth for both high and low survival scenarios (Tables 1A and 1B), we observed a 0.1% reduction (high survival, Table 1A) and a 0.4% increase of population growth (low survival, Table 1B) compared with projections using the age-constant survival rates of Harris *et al.* (2006, 2007; Fig. S5, dashed line). These findings suggest little contribution of age-specific survival on the population projections and confirm those of Haroldson *et al.* (2006), who found that age class did not explain much variation beyond other covariates in the model. We do not dismiss Doak & Cutler's point that survival senescence may become biologically 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. Also, the practical challenges of basing sustainable mortality limits, a requirement under the Grizzly Bear Recovery Plan for the GYE, would be daunting: wildlife managers would have to allocate mortality limits across age classes, which, given the nature of grizzly bear mortalities, is practically impossible.

Incorporating Reproductive Senescence in Population Projections.—Doak & Cutler point out that the population projections conducted by Harris *et al.* (2007) did not account for reproductive senescence of bears until they reached age 30. This is true but misleading. The stochastic projections contained in Harris *et al.* (2007) were designed to address issues remaining from the earlier work of Harris *et al.* (2006). Contrary to the assertion of Doak & Cutler, Harris *et al.* (2006:47) explicitly modeled reproductive senescence of adult female grizzly bears and examined an alternative reproductive function that incorporated influences of mother's age on reproduction that were well-supported by GYE-specific empirical data during 1983–2001.

Harris *et al.* (2006) also used data from Schwartz *et al.* (2006a:21) and Schwartz *et al.* (2006b:29) to develop a well-supported alternative to a constant function, accounting for both the lower litter size and cub survival typical among offspring of primiparous females (Schwartz *et al.* 2006b:30) and lower production among older females (Schwartz *et al.* 2003). Projections using this alternative reproductive function generated mean lambda and probabilities of decline within a 15-year time horizon that were within 1% of those using the constant reproductive function (Harris *et al.* 2006:51). Given that considerably more uncertainty than this was present due to the unresolved fate of some adult females, not to mention sampling uncertainty (treated in greater detail by Harris *et al.* [2007]), we viewed this difference as sufficiently small and thus focused on the simpler model in subsequent work. Thus, the conclusion of Doak & Cutler that our demographic analyses “ignored senescence” or paid “inadequate attention to the life history characteristics” of GYE grizzly bears is unfounded.

Further, Doak & Cutler’s analysis of reproductive senescence hinges on the argument that by not accounting for reproductive senescence, IGBST’s projection models overestimate population growth. They support this argument by comparing projection models incorporating age-specific fecundity ($m_{x\ age}$; Doak & Cutler refer to this as reproductive senescence) based on Schwartz *et al.* (2003: model D) to a baseline projection model with constant fecundity across all ages, but varying annually, ($m_{x\ pop}$) based on Harris *et al.* (2007) (Fig. S6A). However, the baseline fecundity (Harris *et al.* 2007) and age-specific fecundity (Schwartz *et al.* 2003) come from very different sources (Fig. S6B) and were constructed for very different reasons. Therefore, they should not be compared to demonstrate the influence of reproductive senescence on population projections; mean fecundity of females ≥ 4 years ($\bar{m}_{x\ pop}^* = 0.2635$; the asterisk indicates a weighted mean of age-specific fecundity [$m_{x\ age}$] using the age distribution from each

simulation year as the weight) was substantially smaller than the mean fecundity estimate ($\bar{m}_{x\ pop} = 0.322$; Fig. S6) they used for their baseline. Doak & Cutler acknowledge this difference and its importance to population projections but seem to rescue their argument in their Supplemental Information: “This difference could conceivably be partially due to higher reproductive rates in GYE than elsewhere (although this possibility is not alluded to in Schwartz *et al.*’s discussion of reproduction in different regions), or, as seems more likely, from the non-representative age sample of collared bears used to arrive at average adult reproductive rates (as discussed under survival senescence).” That is, Doak & Cutler center their argument on the premise that non-representative sampling biased the estimate of fixed fecundity used in IGBST population models and simply dismiss potentially higher reproductive rates of GYE grizzly bears. However, the sample of independent females would have to be extremely biased and restricted to a few age classes with highest fecundity from Schwartz *et al.*’s (2003) model D to obtain a $\bar{m}_{x\ pop*} = 0.322$ (Fig. S7A). The age distribution of independent females monitored by IGBST for vital rate analyses clearly demonstrates no such extreme bias exists (Fig. S7A). More importantly, the data used by Schwartz *et al.* (2003) indeed indicate that GYE bears had higher reproductive rates than most others, with proportion of females with cubs 0.308 for age classes 4–30 (6th ranking among the 20 populations; C. C. Schwartz, U.S. Geological Survey (Emeritus), unpublished data), which, after multiplying by mean number of female cubs/litter results in a fecundity of 0.314 (Fig. S7B). We show these data to establish that GYE bears have had high reproductive rates compared with other populations, but we recognize these data were not presented in Schwartz *et al.* (2003). Regardless, Doak & Cutler should have known their age-specific m_x was low and not representative of GYE based on previous studies (e.g., Eberhardt *et al.* 1994:361, Craighead *et al.* 1995:420, Schwartz *et al.* 2006a:20).

To demonstrate the importance of which baseline fecundity is used for comparison, we altered Doak & Cutler's simulations to develop a reproductive senescence-only scenario, which they did not include. A hypothetical population with a comparable baseline of $\bar{m}_{x\ pop*} = 0.2635$ would undergo only a 0.5% (high survival, Table S1A) and 0.5% (low survival, Table S1B) reduction in population growth during 1983–2001 upon inclusion of reproductive senescence, compared with a 2.5% and 2.6% reduction, respectively, using the higher, incompatible baseline of $\bar{m}_{x\ pop} = 0.322$ (Table S1). The lack of a strong reproductive senescence effect confirms findings from our earlier work, as discussed above. Clearly, the choice of fecundity baseline changes interpretation regarding impacts of reproductive senescence, which Doak & Cutler did not acknowledge.

Incorporating Survival and Reproductive Senescence in Population Projections.—How Doak & Cutler's choice of extreme survival senescence and incompatible baseline for reproductive senescence influenced their simulation results is evident from Table S1. Using the hypothetical reference data for inference, we see that reductions in population growth due to Doak & Cutler's combined reproductive and survival senescence is 1.6% (Table S1A) and 1.5% (Table S1B); these reductions are the same as those for survival senescence only (Tables S1A and S1B) because when both senescence functions are combined and an appropriate reference model is used, the extreme survival senescence function of Doak & Cutler (i.e., almost no survival past age 20) overrides any effects from reproductive senescence. We point this out because Doak & Cutler imply a strong additive effect from reproductive senescence in their Fig. 6, but how could this be if their simulations allow no females to survive long enough to reach reproductive senescence age? The answer is that their additive effect from reproductive senescence is a function of an inappropriate fecundity reference, rather than senescence itself.

Finally, when using the hypothetical reference model for fecundity, IGBST estimates of age-specific survival resulted in only a 0.2% reduction (Table S1A) and a 0.2% increase (Table S1B) in population growth with a slightly additive effect when combined with Doak & Cutler's reproductive senescence (0.4% [Table S1A] and 0% [Table S1B] reduction of population growth, respectively). These minor differences are not biologically relevant and do not suggest the importance, based on these simulations, of incorporating survival and reproductive senescence. We conclude that Doak & Cutler'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 erroneous conclusions.

Discussion

Doak & Cutler conclude that “we actually know very little about the past trends of this population, and hence about their likely future fate” and that “with rapidly accelerating impacts, and with Chao2 estimates flattening in the last decade, even as search effort has continued to increase, it is quite likely that the population is now, in fact, declining.” (p. 9). Doak & Cutler further assert that the comprehensive population studies of IGBST (e.g., Schwartz *et al.* [2006d], Keating *et al.* [2002], Cherry *et al.* [2007], Harris *et al.* [2006, 2007], and IGBST [2006, 2012]) 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 incorrect use of

existing data, and fail to recognize the broader context provided by available data and other published information.

We recognize that no two investigators are likely to make the exact choices in dealing with complex data. It can be tempting, from afar, to imbue such choices with intent, when in fact, it merely results from the inevitable differences in choice made by equally qualified and objective investigators. In fact, 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 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, and provided complete transparency of the process (IGBST 2012:20).

For example, Harris *et al.* (2006) recognized that a constant age model, while best supported by our model-selection approach, yielded survival rates of adult females that generated proportionally more reproductively-active old animals than we suspected actually existed while ignoring evidence that experienced females recruited more offspring than naïve ones. They thus conducted sensitivity analyses using a second reproductive function that provided an alternative balancing of complexity and precision, finding it generated results similar to our simpler one. As well, even the 19-year long data set of Schwartz *et al.* (2006d) lacked potentially important information regarding survival of yearlings. Schwartz *et al.* (2006b) were only able to infer the fate of dependent young by following their mothers. Cubs and yearlings were both assumed to have died when continued observation flights failed to find them in the presence of their mother. But unlike cubs, some yearlings may have been weaned and survived (Schwartz *et al.* 2006b:26).

Rather than account for this complexity where no data existed to allow its quantification, Schwartz *et al.* (2006b) made a reasonable decision to treat all such yearlings as having died, despite this likely inducing a negative bias in survival rate, rather than allowing for positive bias. Another likely source of negative bias occurred in stochastic projections. Because Schwartz *et al.* (2006b) were unable to separate process from sampling variance in estimates of cub and yearling survival, Harris *et al.* (2006) thus projected populations with both sources of variation, inducing a slight negative bias. Whenever decisions such as these had to be made, IGBST scientists invariably chose the conservative alternative.

The IGBST has played an important role in the management of GYE grizzly bears because of its unique role bridging science and policy (Lynch *et al.* 2008:826). Management decisions regarding Yellowstone grizzly bears are not based on any single data source or analysis but rather on the totality of the data, rigorous analyses, and knowledge of the system by a team of experts. To that regard, we point out that two independent datasets (Chao2 based on unduplicated counts of F_{COY} and population projections based on demographics analyses of intensive known-fate data) have shown very similar population growth trends (Harris *et al.* 2006, 2007; Haroldson *et al.* 2011). Moreover, our interpretation of a positive population trend during 1983–2001 and diminished growth in the last decade (IGBST 2012) is supported by additional indicators, including: (1) standardized 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 the number of bears observed per hour increased from less than 1/hr in the mid-1990s to around 3/hr in late 2000s (Fig. S8); (2) proportion of unmarked bears in the capture sample (i.e., first capture) remained constant and high (50–70%) while the total individuals captured/yr increased (Fig. S9); and (3) during the 1970s and into the 1980s,

grizzly bear range was likely contained within the recovery zone (23,828 km²) but expanded to 34,416 km² by the end of the 1990s (Schwartz *et al.* 2002) to 50,280 km² in 2010 (Bjornlie *et al.* 2013).

We agree with Doak & Cutler that even long-term data sets often have deficiencies and contain internal inconsistencies. In such cases, rather than there being a single, unambiguous ‘correct’ approach, decisions must be made on how to best apply statistics and develop models in an objective manner. When unambiguous data are lacking (as they often are), investigators must decide upon and justify what assumptions to make, and accept the responsibility to clearly articulate them and examine the consequences of reasonable alternatives. The approach adopted by Schwartz *et al.* (2006d) was to adhere as closely as possible to empirical data, and to prefer simpler to complex models when they provided satisfactory fits. This is not an uncommon approach, and the wealth of papers and publicly available reports by the IGBST and collaborating scientists is testimony to their work on keeping decisions and assumptions transparent and open to peer review. Careful reading of previous papers and study team reports addresses most, if not all, of the claims presented by Doak & Cutler. We would gladly have discussed their questions with them had we been aware of their work prior to publication.

Literature Cited

- Bjornlie, D.D., D.J. Thompson, M.A. Haroldson, C.C. Schwartz, K.A. Gunther, S.L. Cain, D.B. Tyers, K.L. Frey & B. Aber. (2013). Methods to estimate distribution and range extent of grizzly bears in the Greater Yellowstone Ecosystem. *Wildl. Soc. Bull.* (in press).
- Blanchard, B.M., R.R. Knight & D.J. Mattson. (1992). Distribution of Yellowstone grizzly bears during the 1980s. *Am. Midl. Nat.* **128**, 332–338.

- Chao, A. (1989). Estimating population size for sparse data in capture-recapture experiments. *Biometrics* **45**, 427–438.
- Cherry, S., G.C. White, K.A. Keating, M.A. Haroldson & C.C. Schwartz. (2007). Evaluating estimators of the number of females with cubs-of-the-year in the Yellowstone grizzly bear population. *J. Agr. Biol. Env. Stat.* **12**, 195–215.
- Craighead, J. J., J. S. Sumner & J. A. Mitchell. (1995). *The grizzly bears of Yellowstone: their ecology in the Yellowstone ecosystem, 1959–1992*. Island Press, Washington, D.C.
- Doak, D.F. & K. Cutler. (2013). Re-evaluating evidence for past population trends and predicted dynamics of Yellowstone grizzly bears. *Cons. Lett.* doi: 10.1111/conl.12048.
- Eberhardt, L. L., B. M. Blanchard & R. R. Knight. (1994). Population trend of the Yellowstone grizzly bear as estimated from reproductive and survival rates. *Can. J. of Zool.* **72**, 360–363.
- Harris, R.B., C.C. Schwartz, M.A. Haroldson & G.C. White. (2006). Trajectory of the greater Yellowstone ecosystem grizzly bear population under alternative survival rates. In: *Temporal, spatial and environmental influences on the demographics of grizzly bears in the greater Yellowstone ecosystem* (eds. Schwartz, C.C., M.A. Haroldson, G.C. White, R.B. Harris, S. Cherry, K.A. Keating, D. Moody & C. Servheen). Wildl. Monogr. 161, pp. 45–56.
- Harris, R.B., G.C. White, C.C. Schwartz & M.A. Haroldson. (2007). Population growth of Yellowstone grizzly bears: uncertainty and future monitoring. *Ursus* **18**, 167–177.
- Haroldson, M. A. (2011). Assessing trend and estimating population size from counts of unduplicated females. In: *Yellowstone grizzly bear investigations: annual report of the*

- Interagency Grizzly bear Study Team, 2010* (eds. Schwartz, C.C., M.A. Haroldson & K. West). U.S. Geological Survey, Bozeman, MT, pp. 10–15
- Haroldson, M.A., C.C. Schwartz & G.C. White. (2006). Survival of independent grizzly bears in the Greater Yellowstone Ecosystem, 1983–2001. In: *Temporal, spatial and environmental influences on the demographics of grizzly bears in the greater Yellowstone ecosystem* (eds. Schwartz, C.C., M.A. Haroldson, G.C. White, R.B. Harris, S. Cherry, K.A. Keating, D. Moody & C. Servheen). Wildl. Monogr. 161, pp. 33–42.
- Higgs, M.D., W.A. Link, G.C. White, M.A. Haroldson & D.D. Bjornlie. (2013). Insights into the latent multinomial model through mark-resight data on female grizzly bears with cubs-of-the-year. *J. Agr. Biol. Env. Stat.* doi: 10.1007/s13253-013-0148-8
- Interagency Grizzly Bear Study Team. (2006). *Reassessing methods to estimate population size and sustainable mortality limits for the Yellowstone grizzly bear: workshop document supplement*. Interagency Grizzly Bear Study Team, U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT.
- 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.
- Johnson, C.J., M.S. Boyce, C.C. Schwartz & M.A. Haroldson. (2004). Modeling survival: application of the Andersen-Gill model to Yellowstone grizzly bears. *J. Wildl. Manage.* **68**, 966–978.

- Knight, R.R., B.M. Blanchard & L.L. Eberhardt. (1995). Appraising status of the Yellowstone grizzly bear population by counting females with cubs-of-the-year. *Wildl. Soc. Bull.* **23**, 245–248.
- Keating, K.A., C.C. Schwartz, M.A. Haroldson & D. Moody. (2002). Estimating numbers of females with cubs-of-the-year in the Yellowstone grizzly bear population. *Ursus* **13**, 161–174.
- Link, W. A. (2003). Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrics* **59**, 1123–1130.
- Lynch, H. J., S. Hodge, C. Albert & M. Dunham. (2008). The Greater Yellowstone Ecosystem: challenges for regional ecosystem management. *Environmental Management* **41**, 820–833.
- Proctor, M. F., B. N. McLellan, C. Strobeck & R. M.R. Barclay. (2004). Gender-specific dispersal distances of grizzly bears estimated by genetic analysis. *Can. J. of Zool.* **82**, 1108–1118.
- Schwartz, C.C., M.A. Haroldson & S. Cherry. (2006a). Reproductive performance of grizzly bears in the Greater Yellowstone Ecosystem, 1983–2002. In: *Temporal, spatial and environmental influences on the demographics of grizzly bears in the greater Yellowstone ecosystem* (eds. Schwartz, C.C., M.A. Haroldson, G.C. White, R.B. Harris, S. Cherry, K.A. Keating, D. Moody & C. Servheen). Wildl. Monogr. 161, pp. 18–23.
- Schwartz, C.C., M.A. Haroldson & G. White. (2006b). Survival of cub and yearling grizzly bears in the Greater Yellowstone Ecosystem, 1983–2002. In: *Temporal, spatial and environmental influences on the demographics of grizzly bears in the greater Yellowstone*

- ecosystem* (eds. Schwartz, C.C., M.A. Haroldson, G.C. White, R.B. Harris, S. Cherry, K.A. Keating, D. Moody & C. Servheen). *Wildl. Monogr.* 161, pp. 25–31.
- Schwartz, C.C., M.A. Haroldson, S. Cherry & K.A. Keating. (2008). Evaluation of rules to distinguish unique female grizzly bears with cubs in Yellowstone. *J. Wildl. Manage.* **72**, 543–554.
- Schwartz, C.C., M.A. Haroldson, K.A. Gunther & D. Moody. (2002). Distribution of grizzly bears in the Greater Yellowstone Ecosystem, 1990-2000. *Ursus* **13**, 203–212.
- Schwartz, C.C., M.A. Haroldson, K.A. Gunther & D. Moody. (2006c). Distribution of grizzly bears in the Greater Yellowstone Ecosystem in 2004. *Ursus* **17**, 63–66.
- Schwartz, C.C., M.A. Haroldson, G.C. White, R.B. Harris, S. Cherry, K.A. Keating, D. Moody & C. Servheen. (2006d). Temporal, spatial, and environmental influences on the demographics of grizzly bears in the Greater Yellowstone Ecosystem. *Wildl. Monogr.* **161**, 1–68.
- Schwartz, C.C., R. B. Harris & M.A. Haroldson. (2006e). Impacts of spatial and environmental heterogeneity on grizzly bear demographics in the Greater Yellowstone Ecosystem: a source-sink dynamic with management consequences. In: *Temporal, spatial and environmental influences on the demographics of grizzly bears in the greater Yellowstone ecosystem* (eds. Schwartz, C.C., M.A. Haroldson, G.C. White, R.B. Harris, S. Cherry, K.A. Keating, D. Moody & C. Servheen). *Wildl. Monogr.* 161, pp. 57–63.
- Schwartz, C.C., K.A. Keating, H.V. Reynolds, III, V.G. Barnes, Jr., R.A. Sellers, J.E. Swenson, S.D. Miller, B.N. McLellan, J. Keay, R. McCann, M. Gibeau, W.F. Wakkinen, R.D. Mace, W. Kasworm, R. Smith & S. Herrero. (2003). Reproductive maturation and senescence in the female brown bear. *Ursus* **14**, 109–119.

Table S1. Population size, mean annual population growth (λ), and percent decline in λ for female grizzly bears in the Greater Yellowstone Ecosystem based on simulations of Doak & Cutler (2013). We compared influence of survival senescence as implemented by Doak & Cutler with IGBST estimates of age-specific survival. Doak & Cutler used fecundity (m_x) 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 provide a compatible reference dataset to assess effects of age-specific fecundity; this model is for demonstrative purposes only and is not representative of the GYE (see text 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 see footnotes for corrections.

Population projection scenarios	m_x in reference model based on Harris <i>et al.</i> (2007) ($\bar{m}_{x\ pop} = 0.322$) ^a			m_x in reference model based on Schwartz <i>et al.</i> (2003) ($\bar{m}_{x\ pop} = 0.2635$) ^b		
	N_{2001} ^c	λ 1983–2001 ^d	% reduction of λ	N_{2001} ^c	λ 1983–2001 ^d	% reduction of λ
A. High survival scenario						
Reference model	219	1.074	0	148	1.052	0
Survival senescence ^e	176	1.062	1.1	110	1.035	1.6
Survival senescence (IGBST data) ^f	216	1.073	0.1	143	1.050	0.2
Reproductive senescence ^{g,h}	136	1.047	2.5	136	1.047	0.5
Survival senescence ^e and reproductive senescence ^h	110	1.035	3.6	110	1.035	1.6
Survival senescence (IGBST data) ^f and reproductive senescence ^h	138	1.048	2.4	138	1.048	0.4
B. Low survival scenario						
Reference model	131	1.046	0	88	1.024	0
Survival senescence ^e	108	1.035	1.1	67	1.009	1.5

Survival senescence (IGBST data) ^f	142	1.050	-0.4 ⁱ	93	1.026	-0.2 ^h
Reproductive senescence ^{g,h}	81	1.019	2.6	81	1.019	0.5
Survival senescence ^e and reproductive senescence ^h	66	1.009	3.5	66	1.009	1.5
Survival senescence (IGBST data) ^f and reproductive senescence ^h	90	1.024	2.1	90	1.024	0

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

^bAs implemented by authors for additional simulations with different m_x for reference model; based on R code in Doak & Cutler (2013) Supplemental Information.

^cPopulation size at end of 19-year time series (2001).

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

^eSenescence as modeled by Doak & Cutler (2013) based on parameter estimates for age and age² from Johnson *et al.* 's (2004) proportional hazards model (Fig. 5, gray line).

^fAge-specific survival based on IGBST empirical data (see Fig. 5 [solid black line] for high survival scenario).

^gNot included in simulation models of Doak & Cutler (2013).

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

ⁱNegative value indicates population increase compared with reference data (due to greater survival among ages 6–18 years; Fig. 5).

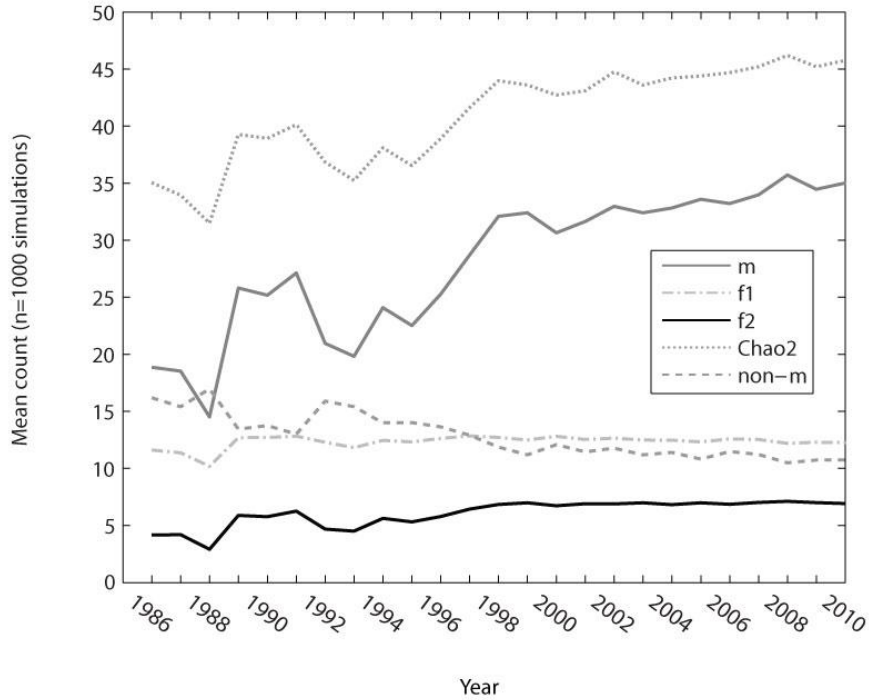


Fig. S1. Mean count of female grizzly bears with cubs-of-the-year (F_{COY}) seen once ($f1$), twice ($f2$), unique F_{COY} observed (m), Chao2 estimate, and F_{COY} not observed ($non-m$; Chao2- m) based on simulations of Doak & Cutler (2013) for a known population of 70 F_{COY} , 1986–2010.

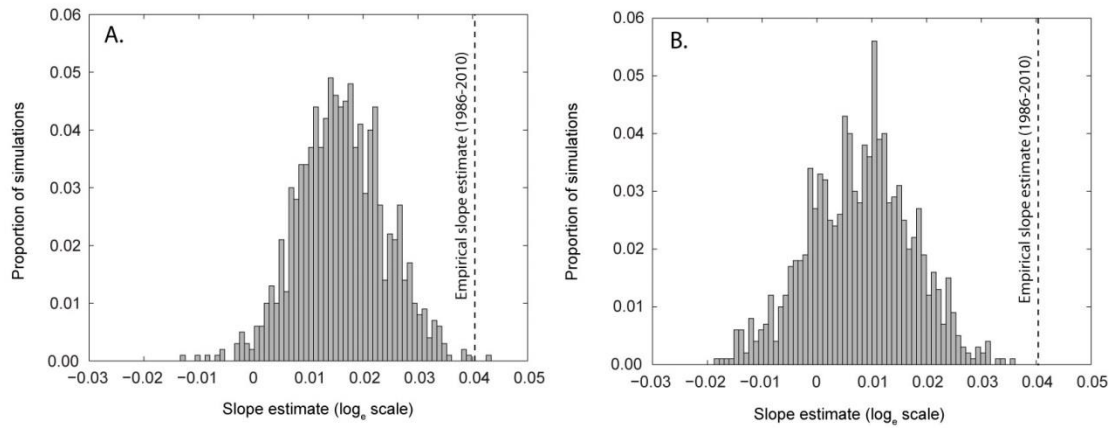


Fig. S2. Distribution of slope estimates obtained from 1,000 simulated realizations of data using linear regression of Chao2 (natural log scale) estimates of female grizzly bears with cubs-of-the-year (F_{COY}) as a function of (A) search effort (number of flight hours) and (B) area-adjusted search effort (number of flight hours adjusted for area surveyed) based on 1,000 simulations of 25-year time series. The vertical dashed line shows the analysis value of the actual estimated slope for empirical Chao2 estimates of F_{COY} during 1986–2010 (Haroldson *et al.* 2011).

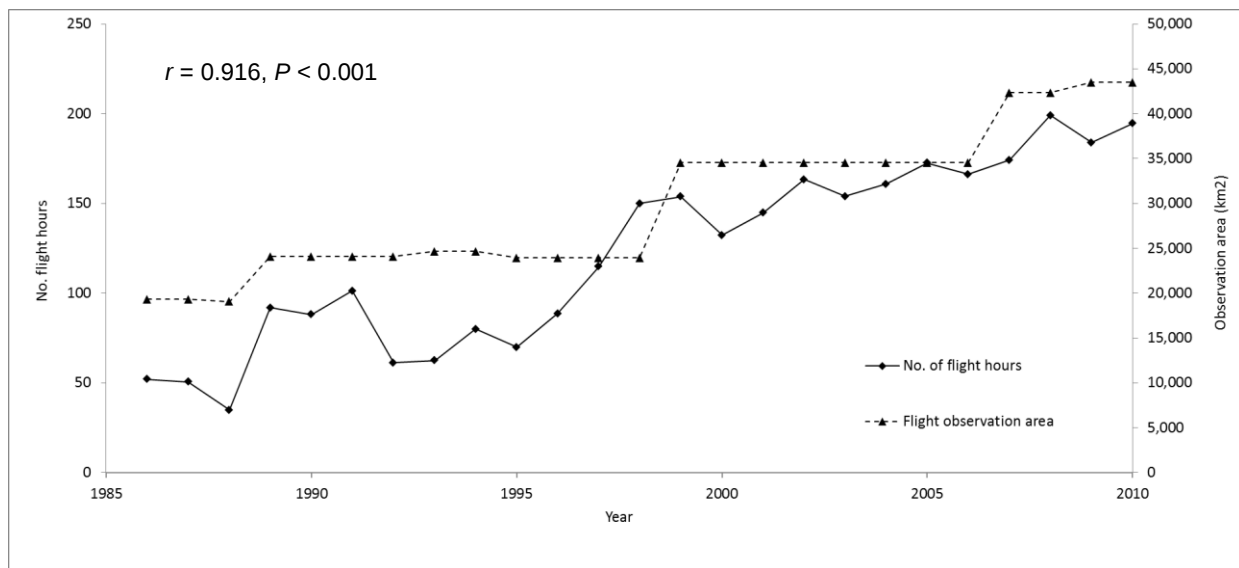


Fig. S3. Number of flight hours increased as flight observation areas for female grizzly bears with cubs-of-the-year were added to accommodate range expansion, Greater Yellowstone Ecosystem, 1986–2010.

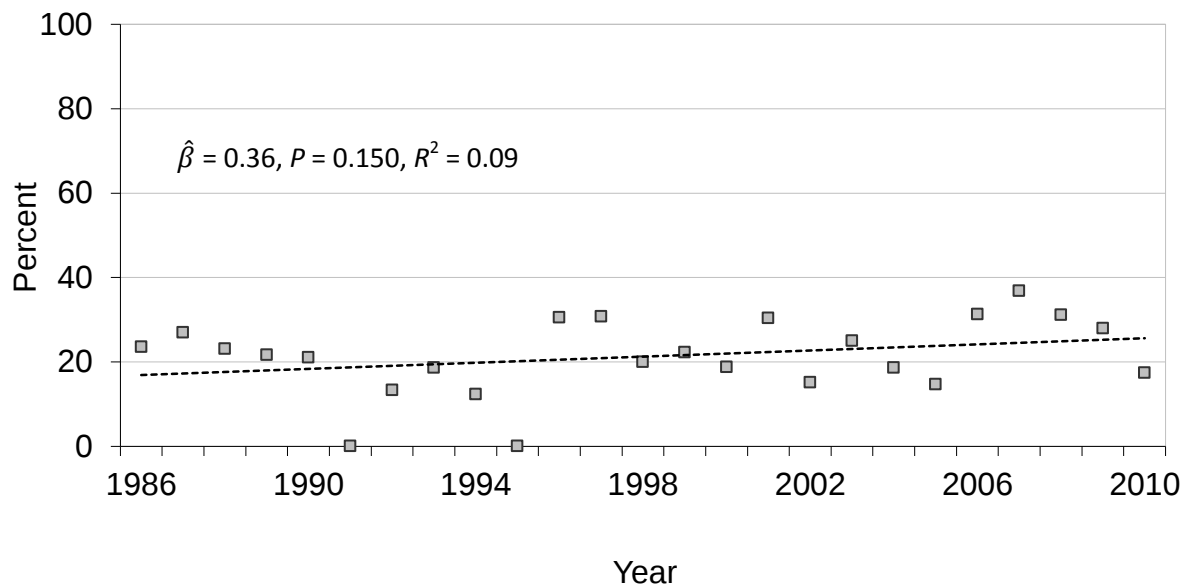


Fig. S4. Percent of locations of radio-marked female grizzly bears with cubs-of-the-year obtained during radio-telemetry flights that resulted in visual observation, Greater Yellowstone Ecosystem, May–August, 1986–2010.

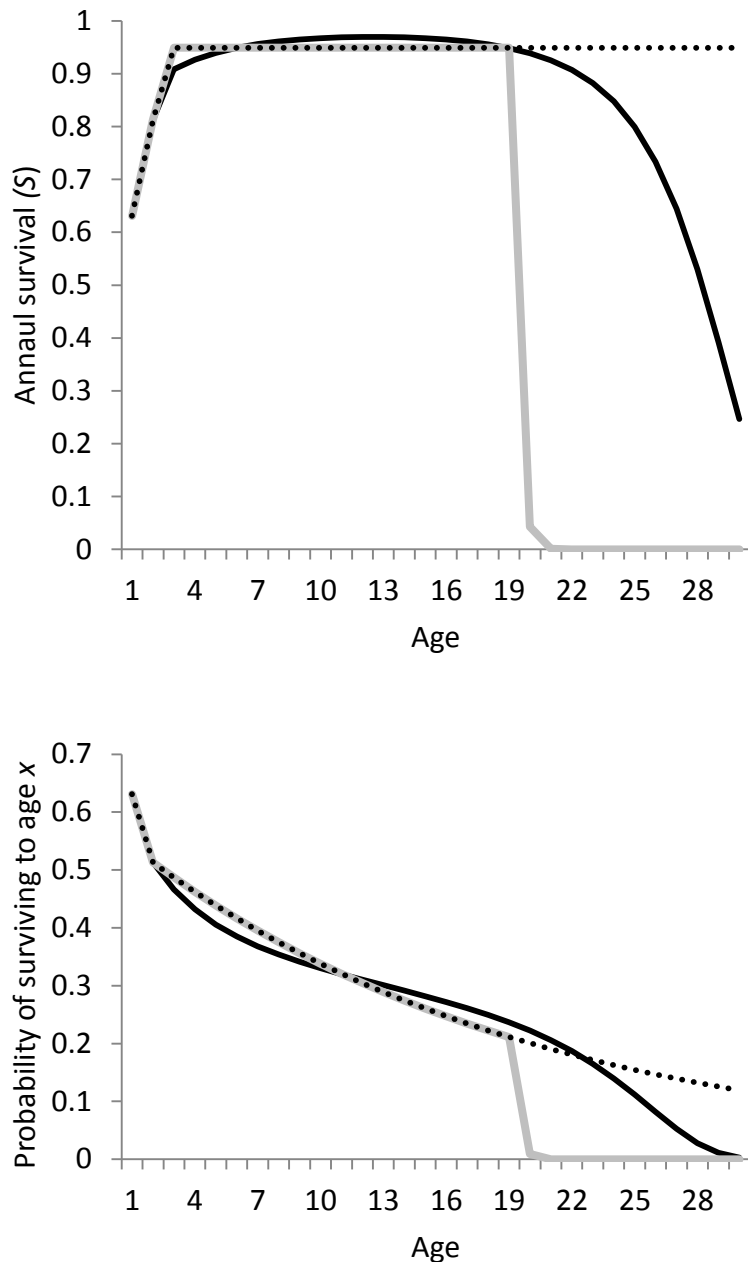


Fig. S5. Age-specific survival (A) and survival probability (B) for female grizzly bears in the Greater Yellowstone Ecosystem, 1986–2010. Dashed line is based on constant independent-aged survival from R code provided by Doak & Cutler (2013) and based on Harris *et al.* (2006, 2007); baseline annual survival ($\bar{S} = 0.949$) was the mean across all years for independent-aged bears (unknown fates censored or high survival scenario) with cub survival of 0.631 and yearling survival of 0.813. Solid gray line incorporates Doak and Cutler’s survival senescence function. Solid black line shows age-specific survival based on empirical data from IGBST for 1983–2011.

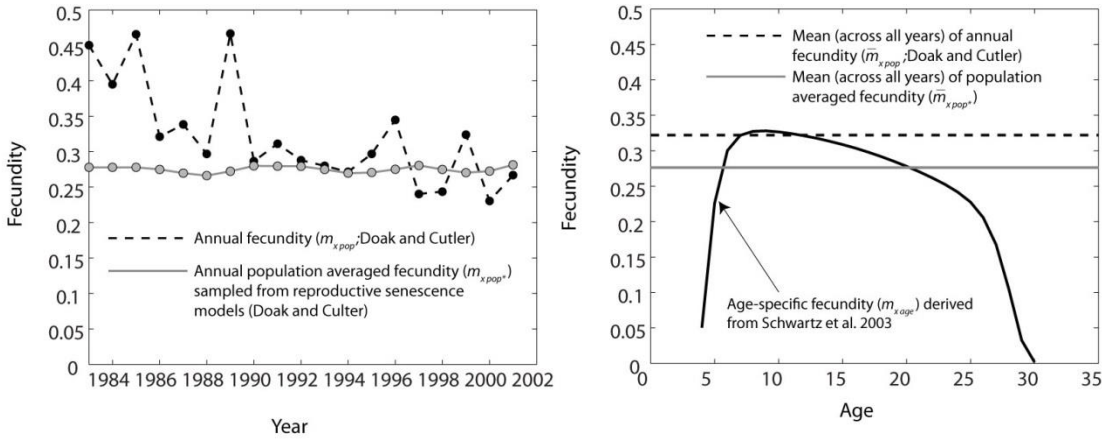


Fig. S6. (A) Doak & Cutler (2013) used population projections with annually varying fecundity values from Harris *et al.* (2007) for 1983–2001 but derived age-specific fecundity ($m_{x age}$) or “reproductive senescence” from Schwartz *et al.* (2003) (B). Based on Harris *et al.* (2007) $\bar{m}_{x pop} = 0.3220$, whereas Doak & Cutler’s simulations of age-specific fecundity using the curve from Schwartz *et al.* (2003) results in a mean $\bar{m}_{x pop*} = 0.2635$; Doak & Cutler do not provide this estimate so we calculated it for >4-year-old females using the age distribution from Doak & Cutler’s simulations for each year as a weighting factor (denoted by asterisk).

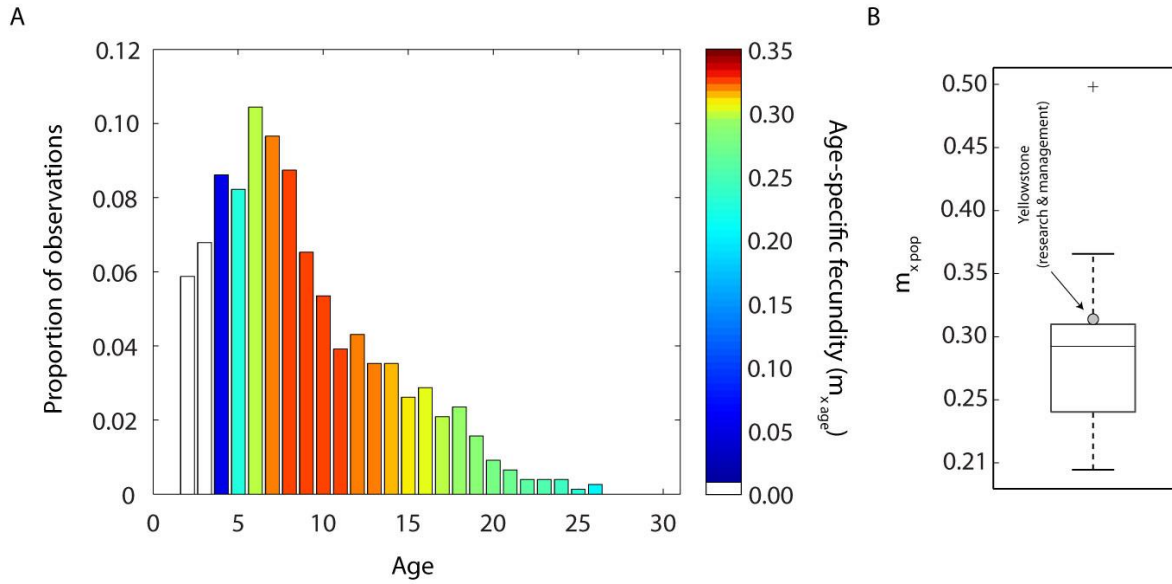


Fig. S7. (A) Empirical age distribution of independent female grizzly bears monitored via radiotelemetry by IGBST in the Greater Yellowstone Ecosystem during 1983–2010. Data from these individuals were used to obtain estimates for cub, yearling, and independent survival and reproductive rates (Schwartz *et al.* 2006a,b; Haroldson *et al.* 2006, 2007). Color gradient represents age-specific fecundity based on Schwartz *et al.* (2003) and shows no support for the argument by Doak & Cutler (2013) that biased sampling produced the higher estimate of $\bar{m}_{x,pop} = 0.322$; this level of fecundity could only have been obtained from Schwartz *et al.* (2003) if the sample of independent females had been exclusively restricted to a subset of age classes (i.e., red-colored bars) of the empirical age distribution from IGBST data. (B) Boxplot showing fecundity for GYE was high among 20 brown bear studies used in analyses of Schwartz *et al.* (2003).

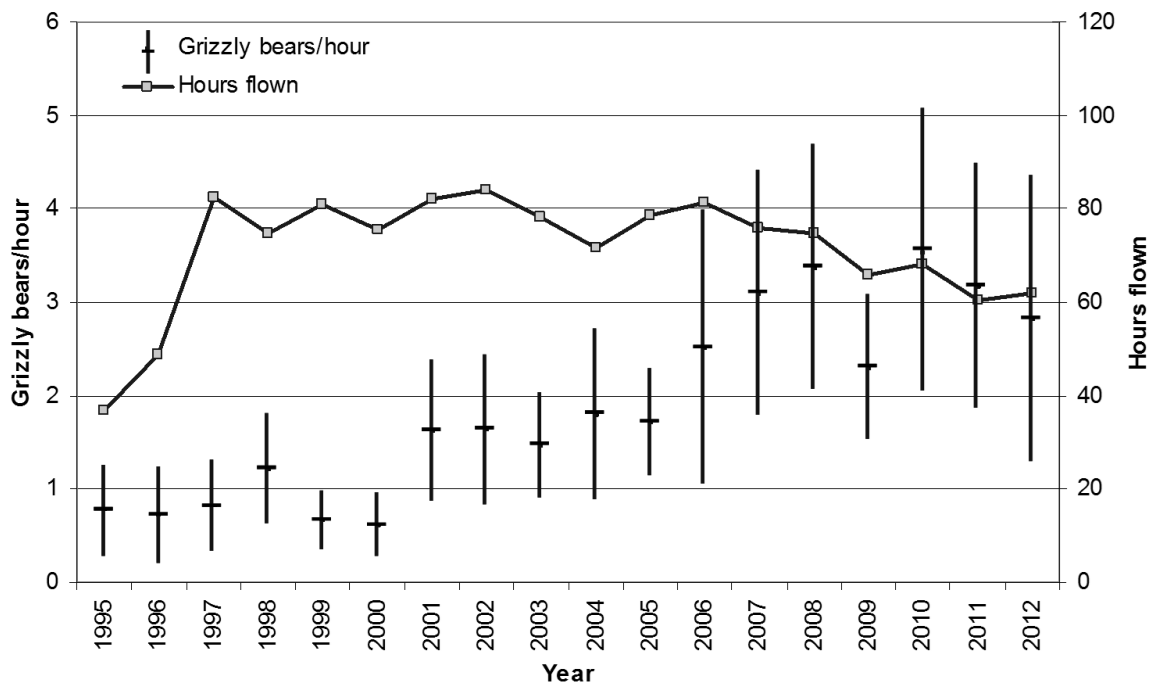


Fig. S8. Mean and 95% confidence intervals for grizzly bears observed per flight hour and total hours flown during 1995–2012 for Recovery Zone Bear Management Units, excluding those that contain army cutworm moth aggregation sites where bears congregate and are more observable.

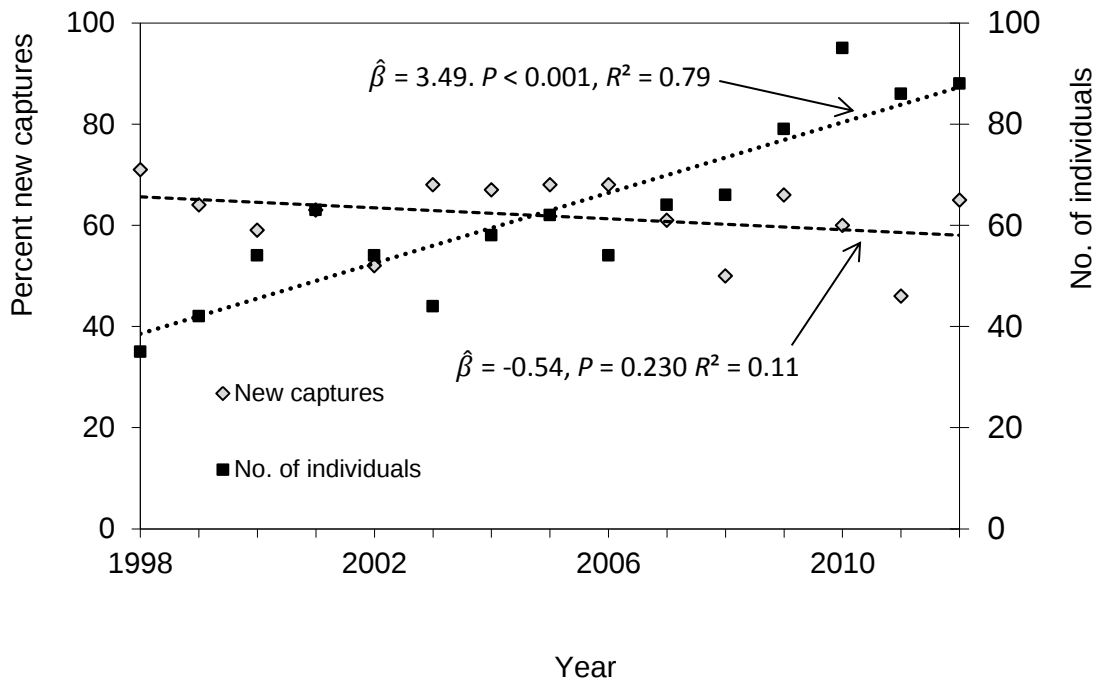


Fig. S9. Number of captures of grizzly bears and proportion of those captures representing new, unmarked individuals, Greater Yellowstone Ecosystem, 1998–2012.

Matlab Programs Used in Analyses

Repeating Simulations of Doak & Cutler and Regression Analyses

Using the same model parameters and simulation R code provided in the supplemental material of Doak & Cutler (2013), we replicated their published simulations of Chao2 for each year, or unique level of flight effort ($n = 1,000$ replicates of each year). We modified their original code to include a matrix of simulation number (1–1,000), year (1986–2010), flight hours (35–199), and the estimated Chao2 (variable `nchao` in original R code) and exported this matrix ([25,000, 4]) for subsequent statistical analysis. We used Matlab (R2013) to examine the individual time series ($n = 1,000$) using simple linear regression to test for the presence of a trend for each simulation. We used the `LinearModel.Fit` function in Matlab's statistical toolbox to fit the simple linear model:

$$\text{Ln}(\text{Chao2}) = \hat{\beta}_0 + \hat{\beta}_1(\text{year}) + \varepsilon_i$$

We used a vector of relative years (0–24) to allow the intercept term to have a biological meaning. We extracted and stored the parameter estimates for the intercept and slope, P values associated with slope for the model F test (goodness of fit), and the R^2 for each simulation. A meaningful comparison of the role of flight hours on the estimation of Chao2 is to compare the distribution of simulated slope estimates to the empirically derived slope estimate for Chao2. The slope estimates from the regression models indicate the magnitude of annual change in Chao2 across all simulations (Fig. S2). The annotated Matlab code is provided below:

```
h = waitbar(0, 'Please do somethings else...');
for i = 1:1:1000;
    subset = find(CHA02OUTb(:,2) == i);
    mdl = LinearModel.fit([0:1:24], log(CHA02OUTb(subset,4)));
    mdl2 = anova(mdl);
    chao_trackerb(i,1) = i;
    chao_trackerb(i,2) = mdl.Coefficients.Estimate(1);
    chao_trackerb(i,3) = mdl.Coefficients.Estimate(2);
    chao_trackerb(i,4) = mdl.Coefficients.pValue(2);
    chao_trackerb(i,5) = mdl.Rsquared.Ordinary(1);
    % create visual progress bar
    % loop through simulation reps
    % extract simulation(i) rows
    % fit the linear model for subset(i)
    % extract Goodness of fit p-value
    % store simulation number
    % store intercept beta estimate
    % store slope beta estimate
    % store slope p-value
    % store model r-square
```

```
        chao_trackerb(i,6) = md2.pValue(1);  
        waitbar(i /1000)  
end;  
% store model Goodness of Fit  
% update visual progress bar  
% end loop through simulation reps
```