# A review of the statistical methods employed in the article "The impact of free-ranging domestic cats on wildlife of the United States"

Gregory J. Matthews, Ph.D.
March 4, 2013

## 1   Summary of Loss et al.

On January 29, 2013 the peer-reviewed journal article titled "The impact of free-ranging domestic cats on wildlife of the United States" and written by Scott R. Loss, Tom Will, and Peter Marra was published in the journal Nature Communications. The authors state that there are many anthropogenic threats

to wildlife and argue that free-ranging domestic cats are among the most deadly. However, they note that the magnitude of the mortality caused by cats is not well understood. Therefore, they set out to perform a systematic review of the scientific literature in order to attempt to quantify the number of deaths caused by free-ranging domestic cats. The results of their study give an estimated range for the number of birds killed by cats annually at between 1.4 and 3.7 billion birds per year and the number of mammals killed by cats annually at between 6.9 and 20.7 billion.

## 2   General conclusion

If I were asked to review this article for admission to a peer-reviewed journal based on the statistical validity of the methods used, my decision would be to reject it. There are numerous major flaws in the statistical arguments made by the authors of this article that, I believe, even upon major revision of the article, would still be unacceptable for publication based on the merits of the statistical methods used by the authors. Below, I have included several sections describing some of the specific flaws that I see in the article.

# 3 Specific Areas of Concern

## 3.1 No meta-analysis was performed

A common type of analysis involved in a systematic review is a procedure referred to as meta-analysis. The general idea of meta-analysis is to combine results from similar, individual studies to calculate an estimate of some effect size or parameter. The benefit of a meta-analysis is that while each individual study may only have a small sample size, the estimate based on the meta-analysis is based on all the results across all of the studies. This effectively results in a larger sample size and, thus, usually produces a smaller estimation error. While there are certainly many complications involved in implementing a meta-analysis, I believe the authors should have at least addressed the issue of why they chose to not use a meta-analysis procedure for this particular study, since they attempt to perform a systematic review of many sources.

## 3.2 Extrapolation is easily misused

In mathematics and statistics, extrapolation involves the projection of some value outside of the studied interval based on the data collected within an interval under study. To put this more simply, the idea of extrapolation involves studying some subset of a larger population and then assuming that what was discovered based on the subset is applicable to the entire population. This idea is often misused and abused. One good example of where extrapolation fails is in projecting the height of a human being based on his or her age. If one were to collect heights and weights of children and plot them against one another, it would be clear that, on the average, older children are taller than younger children. Based on this data alone, and using the idea of extrapolation, one might infer that this trend would continue on through adulthood. However, common sense tells us that this is not the case. The problem here arises because of extrapolation. The observation that age is associated with height only applies to the group that is studied in this case, children. While extrapolation is not always incorrect, one must think very hard about using it.

Loss et al. seems to use extrapolation in many instances where it is likely unwarranted. For instance in Supplementary Table S1, the authors present predation estimates. One estimate, based on [31], estimates 17.32 mammals and 1.64 birds are killed per cat per year. I believe that they arrive at these numbers based on Table 3 in [31], which estimates the "average kills per cat per summer month" to be 1.67. 1.67 is then multiplied by 12 to yield an estimate of predation per cat per year of 20.04. This number is then partitioned into birds and mammals based on the percentages of each observed in the study. In the study, 59 total prey were observed. 51 of these were mammals and the remaining 8 were birds. Multiplying $\frac{51}{59}(20.04) = 17.32$, which is used as one of the estimates of cat predation per year for mammals. (Likewise, I believe the calculation for birds would follow in a similar fashion with $\frac{8}{59}(20.04) = 2.71$; however, this is not the number that is presented in the Supplemental Table S1

2

from the paper.)

However, to arrive at these numbers, the authors are extrapolating in a way that is, at the very least, dubious. The numbers presented in Table 3 from [31] are based on 12 cats over the months June to August 2001. Further, [31: Page 2] states that "This work was conducted in the forest fragments and neighborhoods in and around Albany Pine Bush Preserve (APBP), Albany, NY – a small protected area completely surrounded and bisected by suburban development." Based on a small sample of cats over three summer months in one specific geographic area, the authors see fit to extrapolate this predation rate to all cats at all times of the year in all geographic regions in the United States. This is almost surely an instance where extrapolation is not warranted, as predation rates likely drop in the winter months as mammals and birds become scarcer. Furthermore, during the summer months, cats that spend time near wooded areas likely have higher predation rates than cats not near wooded areas, as they have opportunities to hunt when living close to a protected habitat of many birds and small mammals. In fact, the authors of [31] note in the last sentence of their abstract that "Cold weather and healthy cat predator populations are speculated to minimize the ecological impact of IOHC [indoor/outdoor house cats] on this area."

Other examples of extrapolation include [25] and [26], which are both cited as providing estimates of predation rates among un-owned cats. It is not immediately clear to me how the authors arrived at the predation rates of 110.35 and 305.58 for birds and mammals, respectively, based on [25], but even assuming that these numbers are correct, applying these estimates to all cats across the country is highly questionable. [25] was based on an examination of the stomach contents of 86 cats who were hunted and killed for examination in Willamette Valley, Oregon in the late 1930s. However, the authors of Loss et al. are assuming that these estimates apply in the present and over all geographic regions of the United States throughout all parts of the year. (It should be noted here that the hypothesis of this study was that cats are the greatest known predator of birds and the major conclusion after examination of the data was that there is no evidence that this belief is true. In the words of the authors of [25], "The most important general conclusion reached in this investigation is that the stomach analyses of the feral and rural cat do not bear out the contention that the domestic cat in the Willamette Valley is a confirmed game bird consumer." Only 3.37% of the cats investigated contained game birds.)

Based on [26], Loss et al. cites the predation rates of 63.88 and 355.88 per cat per year for birds and mammals, respectively. It is again unclear to me what calculation the authors of Loss et al. perform to arrive at these numbers. Even so, the authors are extrapolating. The study presented in [26] is based on 15 digestive tracts and 25 scats (40 total) in Patuxent Research Refuge, Maryland and a majority of the samples (33 out of the 40) were collected either in November or between December and March. In spite of the specific geographic location and specific season of data collection, the authors extrapolate these numbers to all cats everywhere across the United States.

## 3.3   Ad hoc analysis

Loss et al. argues in the supplemental material that "we defined uniform probability distributions because there is not sufficient data to ascribe greater likelihood to any particular value." I do not believe this adequately addresses why they have chosen to use a uniform distribution. The uniform distribution is characterized by an upper and lower limit where all values between these two bounds are all exactly equally likely to occur. Assuming that we allow for a uniform distribution to be used where they have used it, the decisions that the authors make in defining the distribution parameters are entirely ad hoc. As far as I can tell, the authors use no formal statistical method for estimating the upper and lower bounds of each particular uniform distribution.

For example, Loss et al. cites 8 studies that offer some estimate of the proportion of owned cats that have access to the outdoors. Those estimates are 0.66, 0.5, 0.65, 0.4, 0.43, 0.77 0.36, and 0.56. Then based on these 8 estimates they decide, without any formal calculation, that the truth is somewhere between 0.4 and 0.7 and that all values between this minimum and maximum are equally likely. This ignores many important pieces of information, including the fact that all of these estimates are based on different sample sizes, and thus have differing amounts of error. Furthermore, three of these estimates are based on nationwide surveys while the remaining five are based on estimates from specific states (New York, Kansas, California, Michigan, and Florida). To claim that all 8 of these estimates should receive equal weight ignores many basic principles of statistics including the concept that statistics based on larger sample sizes, when all else is equal, will have less estimation error. Therefore, it seems like it would have made more sense for the authors to propose some scheme for weighting the relative importance of these 8 different estimates.

## 3.4   Mischaracterization of the error involved in each individual estimate

The authors fail to acknowledge that each individual estimate itself includes error. For example, consider proportion of cats that are allowed outdoors. The authors of Loss et al. have decided that this should follow a uniform distribution from 0.4 to 0.7. 8 estimates are used from various papers. These estimates, as before, are (from supplemental material) 0.66, 0.5, 0.65, 0.4, 0.43, 0.77, 0.36, and 0.56. From these estimates, they have determined that the absolute minimum that the proportion of outdoors cats could possibly be is 0.4. However, this estimate is based on a sample size of only 53 households (in New York only) and a 95% confidence interval for the true proportion of cats with access to the outdoors is from 26.43% to 52.77%. Therefore, even if the choice of a uniform distribution is found to be acceptable, the authors ignore the fact that there is quite a bit of variability in the individual estimates, not just variability between the estimates.

4

## 3.5 The authors cite sources that are not peer reviewed

Specifically [43], [44], and [45] were produced by Marketing and Research Services, Inc., the American Pet Products Manufacturers Association, and the American Bird Conservancy. As far as I can tell, none of these articles have been peer-reviewed.

## 3.6 Counting a study twice

[43] is based on a 1997 survey from what the authors cite as Marketing and Research Services, Inc. However, I believe the correct name for this company is Marketing and Research Resources, Inc. (MRR), which is located in Frederick Maryland. The numbers quoted in [45] are also based on a survey from MRR from 1997. It is likely that these are exactly the same sources. This was pointed out on the blog VoxFelina, which can be found at VoxFelina.com

## 3.7 Questionable interpretation of results

[43] and [45] contain results that 35% of respondents keep their cats "indoors all of the time" and 31% keep their cats "indoors mostly with some outside access". The authors interpret this to mean that 65% of cats have access to the outdoors. They fail to make any distinction between cats that are "indoors mostly with some outside access" and cats that may be completely outdoor cats. It is also unclear what the authors of [43] (and [45]) mean by "indoors mostly". This was also pointed out by VoxFelina.