

NWCC/WREN Webinar August 23, 2018

View online: <https://www.nationalwind.org/first-look-at-a-new-generalized-fatality-estimator-genest/>

First Look at a New Generalized Fatality Estimator, GenEst

Presented by Manuela Huso, United States Geological Survey

PAIGE JOHNSON (American Wind Wildlife Institute): Welcome to today's NWCC/WREN co-organized webinar. This webinar today will be on a new generalized fatality estimator called GenEst. Today will be presented by Manuela Huso from the U.S. Geological Survey.

This is Paige Johnson, and I'm here with Ian Evans. We're with the American Wind Wildlife Institute. Before we get going with the presentation here and give folks a few minutes to join, I'll just go through a couple of quick logistic notes.

For sound, you should be able to hear us through your computer or through the call-in number. So, if you're not able to hear us through your computer speakers, you can call into the number listed in the "Notes" box in the Webinar Room. If you have called in and you are not presenting, please make sure to mute your line so we don't get any background noise.

If you're having any technical trouble, you can e-mail Ian at ievans@AWWI.org; and his e-mail address is in the "Notes" box in the Webinar Room as well.

We'll have time for questions after the presentation. If you have a question, please enter it into the "Q&A" box at the bottom right of your webinar screen. You will only be able to see questions you submitted, not those submitted by others. What we'll do is we'll read the questions out loud after the presentation, or Manuela will respond to you as time permits. Feel free to type in questions whenever they come up. So if you have a question that you think of during the presentation, go ahead and type it in; and we'll read it when the presentation is finished.

Finally, a reminder that we will be recording this webinar; and we will post it on the NWCC and WREN websites along with the presentation slides within a few days after the webinar.

To kick things off, I'm going to turn it over to Dr. Taber Allison, AWWI Director of Research, who will give a short introduction to the topic and to our presenter.

TABER ALLISON (American Wind Wildlife Institute): Thank you, Paige. I just want to check, you can hear me?

PAIGE JOHNSON (American Wind Wildlife Institute): Yep, coming through loud and clear, thanks.

TABER ALLISON (American Wind Wildlife Institute): All right, thank you.

Well, it's my pleasure to introduce our webinar presenter, Manuela Huso, from the Forest and Rangeland Ecosystem Science Center of the U.S. Geological Survey. As Paige described at the beginning,



Manuela is going to be talking about really what I would say is the product of a multiyear, multi-stakeholder effort to develop what we're calling the Generalized Estimator. Manuela will be talking about all the reasons and goals for the development of this estimator in the course of her webinar.

But I did want to introduce Manuela and also say that one of the purposes of this webinar in introducing the Generalized Estimator, or GenEst as we refer to it, is to give you an overview and then also describe a way that you can view the software, test the software, and provide comments. So, we see this webinar as an important part of the outreach for promoting the use of this estimator that several people have worked so hard on to develop.

As I mentioned, Manuela is a Research Statistician with the USGS Forest & Rangeland Ecosystem Science Center in Corvallis, where she has been for five-and-a-half years. Prior to that, she was at Oregon State University as a statistician for 20 years; and since 2004, she has been working on the design and analysis for post-construction fatality monitoring, as well as the evaluation of deterrent and curtailment options for mitigating the impacts of wind energy on wildlife.

With that, I'm very pleased to introduce Manuela, and thank you.

MANUELA HUSO (United States Geological Survey): I'm very sorry about that little technical difficulty, but I very much appreciate being able to talk with you all today, and also for the fact that you have come to listen to this.

I'm very excited to talk to you about this new tool, GenEst. It's, as Taber said, an estimator of wildlife mortality at renewable energy facilities. That means not *just* wind, although I'm going to focus on wind today; but it also includes solar, power lines, as well as oil stills. In short, it's an estimator that can provide accurate estimates of mortality when mortality takes place over an extended period of time and the probability of discovery of all fatalities is less than 1.

But, GenEst has not completed its formal peer review, and so I need to begin with a disclaimer that says that this information is preliminary, subject to revision, and is provided to meet the need for timely best science.

Before I begin, I want to acknowledge the development team who has worked so hard over this past year-plus to bring this from mere concept to reality. I am only the one who was elected to give the talk.

Dan Dalthorp, from the USGS, headed up this project and has masterminded so many innovative approaches to this that I can't even keep count.

Juniper Simonis of DAPPER Stats is the lead on the development of the graphic user interface for GenEst.

Lisa Madsen, with Oregon State University, has two large contributions. One was a novel solution to estimating searcher efficiency and the decline in efficiency simultaneously and thereby preserving the correlation in those parameters, and then also a solution for capturing the full variance of our mortality estimates.

There's me.



Paul Rabie and his colleague, Jared Studyvin, from WEST were instrumental in developing new methods for allowing both the shape and scale parameters in the persistence models to be estimated for different covariate class combinations.

Finally, Jeff Mintz, with the USGS, was very instrumental in developing the example datasets.

These are the people who did the hard work to make this become a reality, and everyone contributed to all parts, not just the parts that I highlighted.

Although not formally part of the development team, our consultations with Robert Wolpert, of Duke University, and Fränzi Korner, of oikostat and the Swiss Ornithological Institute, were really invaluable.

Funding was provided by the U.S. Bureau of Land Management; the U.S. Geological Survey' Department of Energy; WEST, Incorporated; Bat Conservation International; American Wind Wildlife Institute; and Avangrid Renewables. We thank them all for their contributions. And guidance, advice, and support were offered by our Steering Committee and very much appreciated.

Well, we all know that wind power is developing very rapidly; and some reports are claiming an even faster rate than we originally anticipated. We all know that there's been some measured impact to wildlife. But what's not quite yet clear is how serious that impact is to certain species like cave-roosting bats that are suffering from the impact of white-nose syndrome, or migratory tree bats like the hoary bat, or golden eagles, for example.

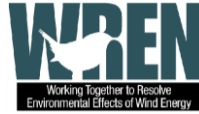
A lot of research is being conducted on ways to avoid, reduce, and mitigate these impacts. But, the fundamental tool that we really need to address these questions is an unbiased estimator of direct mortality that provides accurate measures of the uncertainty in our estimate and correctly accounts for all that we don't see.

An unbiased estimator of mortality is needed so that we can make meaningful comparisons. We commonly see comparisons of mortality rates across sites. But if estimates of mortality are not accurate – that is, they're unpredictably too high sometimes and too low other times – how can we know whether one site really did have higher mortality rates than another? How can we assess whether mortality is higher at some turbines than others?

So accurate estimates are needed to detect trends, to compare reduction strategies; and these comparisons are only possible – well, not really only possible – we can always compare apples to watermelons, but they won't be meaningful until we compare apples to apples or watermelons to watermelons through accurate comparisons.

Current practices for monitoring and reporting differ greatly among countries and often among states or provinces within countries. Some require monitoring and public access to data from all facilities; others don't require monitoring at all. Some might require monitoring, but the data can be held back as proprietary, effectively equivalent to not-at-all monitoring.

The high degree of variation in publicly-available data isn't restricted to the U.S. but appears to be something that occurs worldwide. Among these monitoring studies, approaches for accounting for non-detection varies from doing nothing and reporting raw data to using or requiring different estimators



that may make different assumptions; and then we can end up with very different estimates from the same data.

Combining information from all these sources without accounting for the large differences in quality and meaning of the estimates will lead us to misguided perceptions, just like the blind man and the elephant. Unless we appropriately account for all the ways in which we don't see everything that's killed, and unless we have a representative sample of the wind power facilities, we will develop a very biased picture of what's happening on a national, regional, and even local level and add confusion to our discussions.

Now let's think about how we can help make our estimates of mortality more accurate and, hence, better able to inform our discussions about potential impacts and potential solutions.

Estimating mortality basically depends on counting carcasses and then somehow calculating how many other carcasses might have been out there that we might have missed. You can think of this process as guided by the gods on Mount Olympus playing a game. Each time they flip a coin, an animal dies. So, the number of flips is the number of dead animals, and the probability of detecting the animal in our search process is the probability that the coin lands heads.

So, in this example, the gods flipped ten coins; so, we know that ten raptors have been killed. Mortality, what we designate as " M ", is ten. But we're not gods, so we don't know how many were actually killed. We estimate " \hat{M} ", and the little hat on the " \hat{M} " designates that we estimated it, by dividing our count, " C ", by our probability of serving it, " g ", in this case 0.5 because we have a coin that lands heads half the time, lands tails half the time. So how many did we count?

Well, it looks like three coins landed tails; so, we've counted seven. We divide that count by the probability of detecting it, 0.5; and we get 14. Well, clearly that's wrong because we know it's 10. So, what happened?

Well, in this graph, on the x-axis we've got the possible number of flips, from 0 to 25; and on the y-axis is the probability of getting a 7, heads, when you flip 0, 1, 2, all the way to 25 times. So, the first thing to notice is that having observed 7 heads, it's really not possible for the gods to have flipped less than seven times. So, we know mortality is no less than 7.

The next thing to notice is that our estimate of 14 flips has the highest likelihood of having generated 7 heads. In fact, 14 is the maximum likelihood estimate. Some of you may have heard people talk about that in statistics. So, it's the value most likely to have generated 7 heads, but it's not actually all that likely; and in fact, had we flipped only 13 times, the probability of 7 heads is the same. And, flipping 12 or maybe 15 times isn't all that much less likely.

The point here is that point estimates are always – well, almost always – wrong. So good statisticians always bracket their estimates in confidence intervals or credible intervals; an interval within which if all assumptions hold true, we would expect true value to lie that confidence-level percent of the time. The true value, 10, is indeed within this 95% confidence interval. Because we use a process that correctly estimates variance or uncertainty, then although our point estimate might not be exactly correct, our confidence interval will contain the true values the designated percentage of time. This is one of the very important properties that GenEst has.



So, as we go through this, it's important to remember what I just showed you; that even knowing the probability of a head, like we do with a fair point, doesn't guarantee we can know how many times a coin was flipped from how many heads we observe. So even if we knew the detection probability for every carcass exactly, we would not be able to know exactly how much mortality occurred.

But in fact, " g " is not known; and it must be estimated. And much of what we do in mortality monitoring is to try to accurately estimate that probability of detecting the carcass and the uncertainty around that probability.

So, estimating " g " is notoriously difficult, and it's what's led to differences among our estimators. Different estimators make different assumptions about persistence patterns and probability of the searcher seeing a carcass, the fraction in the searchable area, all kinds of things. And these assumptions aren't necessarily wrong, it's just that they're not always right.

So, our current protocol for collecting the data, we might...let's look at this, the current protocol for collecting data that we need to estimate detection probability. That is, to account for those carcasses we might have missed and how GenEst processes those data.

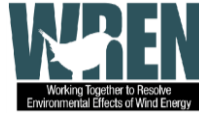
First of all, let's look at " C ", the number that we counted. Typically, we send search crews out on a regular basis to search for carcasses beneath the turbines; and they're essentially counting the number of heads. This search protocol is similar for all estimators and doesn't have to change because we're using GenEst rather than something else.

In GenEst, we ask the user to input those carcass observation data; and that can be done in the form of a CSV file. Then GenEst repeats that data to the user and shows what was uploaded, and the columns can be sorted and viewed pretty much just like you would use any spreadsheet program. But what we count isn't true mortality. So first, we don't search below every turbine. Typically, the turbines that are selected to be searched are a random sample of turbines at the site; so, it's pretty safe to assume that if you search 30% of the turbines, you're going to miss 70% of the carcasses.

Then GenEst asks what fraction of the turbines was searched. And so, the user will input 1 if 100% of the turbines were search, or some fraction – 0.7, 0.3 – if some other fraction was searched. Plot size can differ among studies. Some plots are large enough to comprise the full area where a carcass might land; usually, this means a plot with a radius pretty much over 100 meters. But others might be limited to, say, a 45-meter radius or an 80-by-80-meter plot that barely reaches beyond 50 meters. And sometimes, the area within the designated plots is just too dense or dangerous to search. So, any carcass landing outside the searched area can't be counted. Well, how many is that?

Carcass density typically changes with distance from the turbine and decreases eventually to zero; that's pretty intuitive. So, we need to model that change. We estimate the fraction of carcasses expected to land within the search area, not the fraction of the designated search area that was searched. To do this, we typically model the density of carcasses as a function of distance from the turbine. We need to use carcasses with equal probability of detection or definitely need to account for detection differences in our models.

From those models, we calculate relative density in such a way that we can make this total volume under the curve integrate to one; and then we use these models and superimpose basically any



configuration of searched area, and from it can derive an estimate of the proportion of the carcasses that we expect to land within it. We call this the density weighted proportion or DWP.

In GenEst, we ask the user to upload that Density Weighted Proportion File. So, it's a file that for every turbine gives an estimate of the fraction of each of the carcasses of each different size would be expected to land within the searched area. In this example, all of these data are the same for all bats, which means that we searched exactly the same area under each turbine; and it was all totally searchable. But these numbers don't have to be constant. They can change with turbines.

When a carcass lands at a site, we are definitely not the only ones that are looking for it. There are varying numbers of "N" types of scavengers among sites, so it's really not possible to use the universal estimate of carcass persistence patterns.

Typical protocol calls for placing fresh carcasses in the field and monitoring how long they last. Some areas might have little scavenging pressure or scavenging populations might change with seasons, maybe being very low in the summer and high in winter or vice versa. So, it's important to measure the persistence pattern, not just the average time that a carcass is expected to persist.

The proportion of carcasses that persist can actually change quite a bit, depending on the type of model that you use even with exactly the same average persistence time. Some current estimators assume that carcasses are always removed at a steady rate; that is, a constant proportion is removed every day. Others allow the distribution to differ and more closely reflect the patterns that are observed in trials. But this difference in assumptions can lead to really markedly different estimates of proportion persisting, as you can see in this graphic, and ultimately the different estimates to mortality. So again, GenEst asks the user to upload the file and then allows the user to explore the file like a spreadsheet.

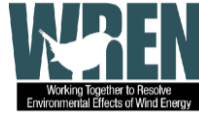
It's at this point that the user will move into the analysis phase. The user selects, so rather than data input, we go into analysis; and now we're going to look at carcass persistence analysis. The user selects variables that indicate what we call the left and the right. It doesn't have to be; that's just our name in this particular dataset. But it is a case that last time the carcass was known to be present and the first time it was known to be absent. This represents an interval during which the carcass was somehow disappeared, was removed.

Other estimators don't allow for an interval estimation but make an assumption that we know when the carcass was removed, and that again can lead to significant differences between estimates of persistence and ultimately estimates of mortality.

So, in addition, the user can input predictor variables by which the persistence is expected to change. Finally, the user has the ability to select the kinds of distributions that are of potential interest and can select from among four of them.

Once the user says, okay, go ahead and analyze, the models are presented in a table that allows for comparison using AICc. It's a statistic that's often used for model selection; and typically, the one with lowest AICc is considered the best model, although those with very near differences in AICc can be considered highly-competing models.

Then, the graphic display can help the user evaluate the differences among models. In this case, the model that was selected is highlighted in the thickest line; but the other models for those classes –



seasons in this case – are also presented so that you can get a visual comparison among them. The user selects the final models, in this screen. By default, GenEst selects the best model in terms of AICc; but the user can override that very easily with these pulldown menus.

The next fourth major source of imperfect detection is the searchers themselves. Whether we're using trained scent dogs or we're using human searchers, they will miss some; and how many? Well, that's for us to estimate. Under current protocol, trial carcasses are placed in the field, and often the same carcasses are used for persistence in soil search or efficiency. But they're placed in locations and times unknown to the users.

The trial manager then, after the search notes if the carcass was found or not, and whether it actually was present and available to be found. If those carcasses that are missed are left in the field for a next search, GenEst can use those data to calculate not only the probability of a searcher finding the carcass but how that probability declines with each search. This is unique to GenEst but can again have important consequences on ultimate estimates of mortality.

So, searcher efficiency data are uploaded, the viewer can review them, and columns indicating whether the carcass was found on the first, second, third, fourth, fifth, et cetera, search are input. Well, if the data are collected for more than one search, GenEst can calculate "k". I'm trying to find my pointer. There we go. GenEst can calculate "k," and the decline in searcher efficiency with each search. Alternatively, the user can fix "k" and enter their own value themselves.

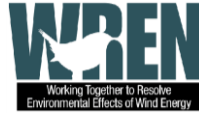
So once the user runs the model, again, a graphic display helps the user visualize how the searcher efficiency changes with each subsequent search and also compares the selected model with the hypothetical best or most complicated model. So, the grey scale estimates here are from the hypothetical model that we're comparing to. We can see that our estimates aren't exactly the same, but they're not far off. So, the extra complication in a more complicated model may not be worth the parameters that you have to estimate; worth the loss in degrees of freedom.

The models can be compared one against another using AIC again in the model comparison window. Then finally, the user again selects the model that they feel is the most appropriate for each of the size classes or each of the class categories for the carcasses.

The fifth major source of imperfect detection is the period when the monitoring did not occur. Well, some studies we design to cover an entire year; some are designed for less. For example, some bat mortality studies in the Northern U.S. are conducted only in the summer and the fall. A majority of the bat fatalities appear to occur during that time, but several studies have noted that fatalities do occur outside that period. So, in order to standardize our estimates, we would have to somehow extrapolate to that unmonitored period.

Unless there are mortality estimates expanded and standardized, estimates from different sites and regions might not be able to be meaningfully compared. This is an issue that people deal with. But we really don't have any way of knowing what fraction or how to adjust for that because we can only estimate for the monitored period. So, this adjustment to some standardized period is incumbent on the user.

A simple model here is one that says that the count that we observe can be calculated as the true mortality times the fraction of turbines that we searched times the fraction of carcasses in the



searchable area beneath those turbines times the proportion of carcasses remaining un-scavenged times the proportion we're lucky enough to find and, finally, times the proportion of annual mortality that we believe occurred during our monitoring period to extrapolate to an entire year.

So, the product of all those factors forms " g ", our detection probability. But this is an extreme oversimplification. It actually looks a lot more like this and is a lot more complicated, but that's why that we developed the software, so that we do the calculations and you don't have to.

Now that we have detection probability, let's estimate mortality. Variance estimates in GenEst come from parametric bootstrapping, so we need to say how many bootstraps we want and what level of confidence we want for our confidence intervals. In addition, we indicate a variable of the carcass data that distinguishes carcasses; that is, it's something that it's very likely that detection probabilities will differ among these classes. Size is a common characteristic for this.

One last input file that's required before we can estimate mortality is an indicator of which turbines were searched on which day. We call it the Search Schedule File. In that file we have dates, we have a list of turbines across the top, date and turbine combination we have an indicated we have a zero for not being searched, a 1 for being searched. In addition, we can add co-variants to this file that are associated with date; co-variants like seasons or rainfall or average temperature, for example, anything that has to do with that particular date but is common to all turbines.

Finally, when we hit "Estimate," we're given a graphic with the point estimate – that little circle on the top – and a characterization of the uncertainty of our estimate in this histogram. GenEst assesses each carcass's contribution to the estimated whole and then sums them up. What we see here is the total; but because we assess each carcass separately and we sum them, you can split a number of different ways, by size, species, all kinds of things. The idea is really simple, but the statistics for making this work correctly are enormously complicated.

What sets GenEst apart is that we've gone through extensive trouble to make sure that the models are correct, and the confidence intervals are accurate not only for the overall total, like you see here, but also for the split. And, we offer a very simple and flexible interface to these complicated models.

For example, we take these data and the user might say, "I want to split on species groups." If in the carcass observation file, I'm sorry, but I don't have access to my pointer or I would be pointing to this down at the lower left. So, if in the Carcass Observation File there happens to be a variable that is called "species group" and in it is an indicator of whether, let's say, the species is a passerine versus a raptor or a shorebird, then we can ask for splits. And now, GenEst has taken those total mortality data and split them into the groups that we asked for. This graphic can be downloaded to be used in reports. Oh, look, there's my pointer. But it's also presented in a tabular form which, again, can be downloaded.

You could ask for estimates that occur through time. The ones that are – those are covariates that might be found in the Search Schedule File. So, in this example, I've used search date; and you can see that we've got an estimate for every search date, we have an estimate of the mortality that occurred. So, quite a bit higher in general in the early part and then a little bit of a lower hump in the latter part. I want to emphasize that all of the data that you look at here are all completely created, made up. They have no resemblance to any particular site or any real estimate of mortality; but they are very, very representative of the type of mortality data that we might see in reality.



You might ask for combinations of variables. So, for example, in this case we used two covariate combinations; species group and turbine types. So now we've got estimates for each species group but then subdivided by turbine type x , y , and z .

Finally, you might be interested in estimates of detection itself. That is a very, very difficult thing to actually talk about and to represent. GenEst gives you a generic representation of the detection probability for the factors that were relevant in the model that you selected. I say "generic" because practically each carcass has a unique detection probability based on its particular conditions; where it lands, when it was found, the search schedule it was found under, the carcass arrival pattern, like do carcasses arrive early or late in the monitoring season, when search conditions might differ, et cetera. So this estimate of detection probability will apply really to an idealized search schedule, and it can be very useful for planning future monitoring studies. But it doesn't mean that every carcass that was in the difficult visibility class that was a bat had a .224 median probability of being detected.

Okay, I'm going to start wrapping things up. We've seen that the way in which an estimator deals with each of the factors that lead to imperfect detection and what assumptions are made and how they put those together in their equations to calculate uncertainty in estimates is what distinguishes them from one another. At times, these differing assumptions can have very little consequence, as in this example where these two estimates from different estimators are more or less the same or very close to each other. But sometimes, the assumptions can have very large consequences and lead to very, very different estimates from the same data.

We created GenEst very much with the motivation to try and avoid these situations. Over the past several years, there has been a natural evolution of methods; and each new estimator that comes along is contributing some improvement over the previous one. The earliest estimators didn't have special software created to carry out the calculations, so to make them accessible to users, they sort of necessarily had to be fairly simple.

GenEst takes a huge leap both in complexity of calculation and in providing users an easy-to-use tool with which to carry them out. It relaxes several of the restrictive assumptions that previous estimators have had with respect to persistence patterns, the proportions found on each search, and the proportion in the search area.

It provides the user with graphical and tabular information with which to evaluate the assumptions and determine the best model. And, most importantly, it provides accurate estimates of uncertainty because, as we all know, point estimates are almost always wrong. So, we have all these advancements with GenEst, and yet the data collection process doesn't have to change. We don't have to modify what we're used to doing.

In this webinar, I have focused primarily on the graphic user interface implementation of GenEst; but it's really much, much more. GenEst is, or will be, an R-package that's going to be available for Windows, Mac, and Linux both in the GUI and the command line implementation. It's open source so that anyone can see exactly what we're doing with the data. None of our equations or calculations will be hidden to anyone. And, of course, if some mistake were to be found, we would love to hear about it.

The statistical models that underlie GenEst – and many of these are relatively novel – will be published in a technical manual; and again, everything is open and will be peer-reviewed.



The GUI implementation will have a detailed user guide headed up by Juniper Simonis; and consistent with R-package development guidelines, GenEst is going to have three separate vignettes. One vignette being examples of tutorials of how to run it through an analysis; one will be with wind; one will be with solar; and a final one will be a walk-through in the command line.

So why do we need GenEst? Well, one thing, it accounts well for pretty much all – or at least almost all, not including the part of the year that's not monitored – all sources of non-detection. It's an unbiased – well, almost unbiased – estimator. In fact, no estimator of this kind is unbiased when "g" is not known; and, of course, we don't know "g", we are estimating it. We are as close as you can get, but nothing can get unbiased.

What that allows us to do then is to be able to compare among site and site properties – turbines, turbine properties, species, et cetera – to estimate trends to evaluate the efficacy of mitigation treatment. And we can estimate for different locations like power lines versus buildings, for different groups of species all from the same study and from the same software.

So, in all of this, giving a real sense of the uncertainty around any of our estimates is critical. A single point estimate is, by itself, almost certainly wrong. We need a process that gives a confidence value that actually reflects the same probability. When we give a 90% interval, we want to know that only 1 out of 10 times will the true estimate be outside of it. Well, GenEst does that.

What GenEst does not do is make decisions for the user, sorry. GenEst does not interpret results; we still have to use our heads and our knowledge of what's going on in the system. And maybe very important to many of you, GenEst is not EoA. GenEst will not give an estimate when zero carcasses are observed. However, it's important to note that when we start to observe carcasses, GenEst and EoA estimates will converge, which is what you should expect if you have an unbiased estimator.

Recently, we received the first round of beta test reviews. Thank you very much if any of the reviewers are out there. We sent it both to statisticians and to users experienced in estimating mortality or using other approaches. All reviews were anonymous unless the reviewer chose otherwise.

We're now in the process of reconciling those reviews and adding features or correcting issues that they've noted. I also want to note that there are several spinoffs from this work. We are publishing, actually, we've already published another R-package called `cbinom` for continuous binomial that we use in our modeling. We're working on a paper called \tilde{x} that has to do with the variates in "M" with uncertain "g". Then also another paper with Michael Schirmacher, who was instrumental in gathering the data for us, to simultaneously estimate searcher efficiency and the decline and preserve the covariates in those two.

Now we're at the point where we're ready to request the next round of reviews from a larger community, and we hope that means you. If there's anyone on this webinar who would be willing to provide us review comments, we invite you to go to this website. We'll hold this up; I know it's an ugly, ugly link but we'll keep this up on the next slide as well, and you can always ask us for it. So, go there, and you can ask to have access to the shiny-server hosted by WEST where you can play with GenEst as much as you'd like. We would definitely appreciate your feedback, and we will reconcile that feedback as it comes in. We'll incorporate as much as possible before we formally publish, which we're hoping to be by the end of October, if not sooner.



In the near future, we are planning to give training workshops around the U.S., trying to make it convenient for people to be able to attend and are actually looking for funding for scholarships for those people who don't have much travel money. We're also developing a proposal; it looks like we're going to be funded to do a simulation analysis in conjunction with WEST to compare GenEst to other estimators. And with funding, which we don't have, we have all kinds of ideas of what we might do to add to GenEst now to make Version 2. Perhaps some of you, after you've played around with a little bit, you might have some of your ideas of your own, which we would love to hear.

In closing, I want to emphasize the collaborative nature of this work. Most of the statisticians involved in creating earlier estimators of mortality have come together to produce a single estimator that has flexibility and strength to address the *many* nuances that are involved in accurately estimating mortality. This product is peer-reviewed. It's open source, which gives anyone the chance to see exactly what we do and provide feedback. Nothing is hidden.

The single estimator motivation for this was to alleviate that conflict that people are faced with in trying to decide which estimator to use. This estimator is unbiased or, as I said, as unbiased as can be expected and because of that will allow us to make those comparisons that we need to make. It will allow us to see the same elephant and make the decisions that we need to make based on comparable information.

With that, I close, once again thanking all of the members of the development team, our financial supporters, our Steering Committee, the reviewers, and many others who contributed in so many ways, and, of course, to you guys. With that, I'd like to open it up for questions. Thank you.

PAIGE JOHNSON (American Wind Wildlife Institute): Thanks so much, Manuela. I just want to let everyone know that we will send this link out in an e-mail to everyone who registered for this webinar. So, if you don't want to have to write it down now, we will send it out following the webinar.

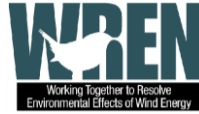
We've had a number of questions come in, so I'll just get started. The first one is: "Does GenEst calculate the DWP, or does that need to be done independently?"

MANUELA HUSO (United States Geological Survey): At this point, it needs to be done independently. However, that's something that we are looking for funding for. Where did that go? I think, yes, we would like very much to develop a module to do that; but at this moment, that's incumbent on the user.

PAIGE JOHNSON (American Wind Wildlife Institute): Thanks, the next question is: "The graphic user interface for left versus right inputs, do those need to be selected for each row of data?"

MANUELA HUSO (United States Geological Survey): I don't believe so. It's a column in the carcass persistence dataset that just indicates – for each trial carcass in the dataset, it indicates how long since having been placed. The trial conductor knows that the carcass was present and still available to be present and then the first time it was known to be absent. So those are just two columns in the dataset that the user inputs from the trial data.

PAIGE JOHNSON (American Wind Wildlife Institute): Okay, thanks. The next question is: "Is there a tool available to quickly calculate the density value?"



MANUELA HUSO (United States Geological Survey): I think that's the same question as earlier. Right now, there are guidelines for that. We have a paper out in *Journal of Wildlife Management* that talks about it. But there is no formal module for it as far as I know.

PAIGE JOHNSON (American Wind Wildlife Institute): Okay, thanks. Sorry for the repeat. The next question is: "Did you say that the probability of where carcasses land on a plot was determined with the dataset?"

MANUELA HUSO (United States Geological Survey): I said that what we do is using distance information that's associated with the location of each carcass, we can use that data to develop these models for density-weighted proportions. Again, that's not done in GenEst; but it can be done by the users, and there are guidelines for how to do that.

PAIGE JOHNSON (American Wind Wildlife Institute): "Does GenEst allow for background mortality to be included in the calculations?"

MANUELA HUSO (United States Geological Survey): GenEst responds to what the user provides. If a user is estimating background mortality at an off-site or at a site where turbines will be but aren't right now, they can still use GenEst for the estimation. If a user wants to add a covariate that indicates that the carcasses are background versus known mortality from the turbines, that's a possibility too. But that's more incumbent on the user and how they use it than on GenEst to determine that for anyone.

PAIGE JOHNSON (American Wind Wildlife Institute): "Is there a way to include combined detection probability rather than separate searcher and persistence?"

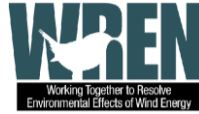
MANUELA HUSO (United States Geological Survey): That is a very, very difficult idea. The appropriate combination of those two is dependent on placing carcasses in a way that accurately mimics their arrival process. And if we knew their arrival process accurately, then we would probably be able to estimate mortality using that knowledge. So, I guess the short answer is, yes, you could probably use GenEst in that way. The long answer is it's not clear that that is an approach that isn't without some very serious considerations.

PAIGE JOHNSON (American Wind Wildlife Institute): "For calculating DWP, do you recommend using the Hull and Muir ballistics model for road and pad areas that are unsearchable?"

MANUELA HUSO (United States Geological Survey): I don't believe so. In our paper in *Journal of Wildlife Management* where we looked at the density-weighted proportion, we compared against Hull and Muir's estimate of the maximum distance at which an animal could land and then just using what we call the triangular distribution...basically, it wasn't like what I showed in this graphic where there's a sort of gradual curve to it.

That certainly helps when you compare it to just saying what proportion of the area that you searched because that assumes a uniform distribution of the carcasses within the searched area. But beyond using their max distance, I don't think that I would recommend that you use their full model.

PAIGE JOHNSON (American Wind Wildlife Institute): Okay, thanks, the next question is: "Does the proportion of carcasses persisting take into account situations such as scavenging? Would that fall under



this variable? For example, a coyote may scavenge before a surveyor has a chance to survey; and you may never know the carcass was present. Is this situation common enough to worry about?"

MANUELA HUSO (United States Geological Survey): Yes, indeed, that situation is common enough to worry about; and that's exactly what the trials are set out there to test. So, we place carcasses. We know they're out there; and when they disappear, we are pretty sure they disappeared because a coyote took them or a raven or something else. A snake, anything. So, by measuring the amount of time, that interval during which we know they persisted versus where we know they were gone and putting all those pieces together from all the different trials carcasses, that's what we use to estimate the fraction that we just won't possibly see because they've been taken away.

PAIGE JOHNSON (American Wind Wildlife Institute): Okay, great, thanks. The next question is: "With carcass persistence, it shows a location and scale value. When reporting carcass persistence, do you recommend reporting both these values? And can you explain their relevance for interpreting meaning?"

MANUELA HUSO (United States Geological Survey): (Laughing) Those are two parameters that are necessary for any of the distributions that we have available to people; the log normal, the log logistical, the Weibull. Those are all two-parameter distributions. And because of that, they're very flexible; and they can much better mimic that removal pattern than the fourth distribution, which is the exponential, which assumes that removal occurs at a constant rate. It just doesn't change with time.

For example, an olfactory predator might not detect carcasses right away until they start to smell a little bit. So, their persistence might be relatively high early on, in the first day or two; and then as they become more detectable by the predators, they are then removed much more rapidly. So, their rate changes.

Alternatively, you might have things that are fresh, and scavengers like those. When they're fresh they disappear very rapidly; and then after they start to dry out and become not very interesting, they can just last forever. Again, that is a change in their rate. So, the two-parameter models reflect that. The two parameters are what determine how those persistence graphs look like, and I would encourage reporting them. I don't think very many of us would be able to look at them and say from just looking at them what that shape or what that distribution for persistence was going to look like, but they're very important in modeling the proportion that will be expected to persist.

PAIGE JOHNSON (American Wind Wildlife Institute): Excellent, thank you. Our next question is: "If 'C' equals zero, do you still recommend using EoA?"

MANUELA HUSO (United States Geological Survey): Yes.

PAIGE JOHNSON (American Wind Wildlife Institute): Great, "GenEst doesn't calculate 'A.' It is user-defined, correct?" The person asking the question is saying they're unsure how to calculate "A" on their own; is there a tool for this?

MANUELA HUSO (United States Geological Survey): Yeah, same question. No tool. Look to our paper in 2014 *Journal of Wildlife Management* for some guidance.



PAIGE JOHNSON (American Wind Wildlife Institute): Okay, thank you: "Is there a way, or recommendation, for users with lower statistics capability to generate the required DWP files?"

MANUELA HUSO (United States Geological Survey): Same question, same answer.

PAIGE JOHNSON (American Wind Wildlife Institute): "Do you think GenEst could be used to recalculate fatalities for older studies that use different estimators?"

MANUELA HUSO (United States Geological Survey): Yes, it depends on what kind of data was collected; but it's very possible to do that.

PAIGE JOHNSON (American Wind Wildlife Institute): Excellent, thank you. "Is there currently a module for solar projects, or is that something planned for future releases?"

MANUELA HUSO (United States Geological Survey): I think we designed this in such a way that we don't need to have a different module. What I recommend is that you follow – or if someone is interested in solar – to look at the example or the vignette that we've developed for solar. We have, I think, two or three different solar example datasets that people can use; and we give guidance on how to go about using GenEst in a solar context as opposed to a wind context.

PAIGE JOHNSON (American Wind Wildlife Institute): "Will there be spreadsheet templates available to facilitate data entry and collection that will be compatible as input files for GenEst?"

MANUELA HUSO (United States Geological Survey): GenEst comes with example datasets. Many example datasets, and any of those could be used as templates.

PAIGE JOHNSON (American Wind Wildlife Institute): "Are the data collection needs inputs for GenEst and EoA identical; that is, can one collect the data and then later choose which analysis to use depending on whether or not the carcasses were observed?"

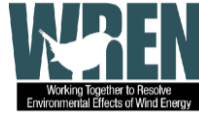
MANUELA HUSO (United States Geological Survey): If by that you mean one conducts a monitoring study as one would typically and then one discovers that there were none of a particular species of interest that they found, then, yeah, the data that would be used and inputted to GenEst could then contribute to the inputs that you need for EoA for that species.

Wait, hang on, Dan wants to say something.

DANIEL DALTHORP (United States Geological Service): The data are the same largely, but the formatting is slightly different.

PAIGE JOHNSON (American Wind Wildlife Institute): Okay, thank you: "Can detection probability be modeled in GenEst using carcass trial data where carcasses are placed, and the only data collected is when and where they are detected?"

MANUELA HUSO (United States Geological Survey): I think we had that question already. I said that, yes, we can use it in GenEst; but I think one needs to think very, very, very hard about how those carcasses are placed and how they reflect the arrival process or not.



PAIGE JOHNSON (American Wind Wildlife Institute): "Over what temporal scale can GenEst run to given that wind turbine sites could lead to a targeted area for predation if carcasses are lying around, so carcass persistence would decrease over time?"

MANUELA HUSO (United States Geological Survey): I'm sorry, I didn't understand that one. Can you ask it again?

PAIGE JOHNSON (American Wind Wildlife Institute): Sure, it's asking over what temporal scale can GenEst cover. And it says, "Given that wind turbine sites could lead to a targeted area for predation if carcasses are lying around?"

MANUELA HUSO (United States Geological Survey): GenEst covers the temporal period indicated in the search schedule, basically. You input a search schedule file; and then GenEst says, okay, this is the information that I have, and so it won't extrapolate beyond that.

PAIGE JOHNSON (American Wind Wildlife Institute): Okay, thanks: "Is the basic formula behind GenEst the same as the previous Huso estimator with improvements? Can it incorporate different estimators as well or simultaneously give results using different estimators?"

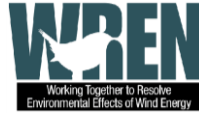
MANUELA HUSO (United States Geological Survey): GenEst is very, very flexible. So, if I forced it to have " k equals zero or k equals 1," let's say, and if I forced it to have an exponential distribution, then I might be able to get the same answer out of the same data as one might expect out of Schoenfeld, for example. But one of the most important advancements that I think GenEst is making is not in the estimate itself – although that's a very important one in the accuracy of the estimate. What's the variance? The variance that GenEst gives you and the coverage that it gives you is what it says. If we say we've got 90% confidence interval, we have a 90% confidence interval. The other estimators are not as accurate as that in terms of their variance.

PAIGE JOHNSON (American Wind Wildlife Institute): Okay, thank you: "Are there parameters separate in GenEst for birds and bats?"

MANUELA HUSO (United States Geological Survey): Certainly, we can ask GenEst to estimate separately for birds and bats; and that's what that size class variable or input that we suggested that people use. It's not a requirement, but it's exactly the way you would go about doing it. If you have a dataset with birds and bats in it, that's probably one of the things you want to include is different parameters for bats and birds and for birds of different sizes.

PAIGE JOHNSON (American Wind Wildlife Institute): "What are the minimum number of searcher efficiency trials and carcass removal trials that need to be completed for the model to run?"

DANIEL DALTHORP (United States Geological Service): It depends on how complex your model is. If you just have one type of area, like roads and pads for one season, then the required number of carcasses is that you have enough so that you find some and miss some. Unfortunately, in a lot of cases you've got different types of area that you're monitoring in; and your detection policy is very different. You have different types of carcasses that you're interested in. Large birds, small birds, bats. You'll need a sufficient number of carcasses for each of those types of carcasses *and* each of the types of areas that you're surveying. So, it could be a large number; it depends very much on how complex your situation is and what the demands of your study are.



MANUELA HUSO (United States Geological Survey): I want to add a little bit to that. In the Huso estimator, there was a lower limit of 10. In this one, we haven't explicitly put that lower limit in; but I want you to think sort of intuitively about it. Remember the example with the 10 dead raptors and the gods flipping the coins. If those had been our trial carcasses and we are measuring searcher efficiency, the true searcher efficiency would have been 0.5; and yet we would have measured it at 0.7. So, the fewer that you have, the more likely your estimates of the different factors that go into detection rates will be not quite so accurate.

PAIGE JOHNSON (American Wind Wildlife Institute): Thank you. A few folks are asking, Manuela, if we can send around your 2014 paper when we send this follow-up e-mail. Would that be okay?

MANUELA HUSO (United States Geological Survey): I don't know if that's okay with the publisher. We can certainly send the link to it, and I don't remember if it's open source or not. I would have to make sure that that's okay with the publisher.

PAIGE JOHNSON (American Wind Wildlife Institute): Okay, well, we can at least send a link to where it is published on the journal page. Then if it's not open source, then if folks have institutional access then they can get it there.

MANUELA HUSO (United States Geological Survey): Yeah, and if they write to me individually, I am allowed to set it individually. I just don't think I can send it en masse. Actually, I don't know that I can or can't.

PAIGE JOHNSON (American Wind Wildlife Institute): No worries, we will figure that out after the webinar; and to everyone on the webinar, we will at least send the link to where it is on the journal website.

We've come to the end of the questions so far. We do have a little bit more time if anyone has any other lingering questions. Manuela has also offered to answer questions offline by e-mail if folks have questions that they think of after the webinar or anything that they want to follow up about. You can contact me, and I can put you in touch.

Just a reminder that we will be sending a follow-up e-mail with the link that's here on this slide. We know that this is not a clickable link due to the webinar technology that we use; but we'll send that around in an e-mail, and you'll be able to go to the link and request access to GenEst.

It looks like we have another question come in here: "When is the non-beta version, the final version, expected to be available to the public?"

MANUELA HUSO (United States Geological Survey): Well, I think we're pretty close. Once we get the final reviews in from this first round and we get a few of the second round, it's going to the publisher. So, at that point, it's just a matter of how long it takes them to process it. I'm estimating October, but I hope it's before then.

PAIGE JOHNSON (American Wind Wildlife Institute): Great, very exciting. Here's another question that came in: "Do you think that carcass distribution with distance from turbines is uniform enough to produce a generalized DWP estimate with accompanying variants?"



MANUELA HUSO (United States Geological Survey): I do not think carcass distribution is uniform. Did I misunderstand?

DANIEL DALTHORP (United States Geological Service): (Inaudible).

MANUELA HUSO (United States Geological Survey): Oh, sorry, say that one again. So that was just talking about like from one year to the next?

PAIGE JOHNSON (American Wind Wildlife Institute): Well, it says: "Do you think that carcass distribution with distance from the turbines is uniform enough to produce a generalized DWP estimate?"

MANUELA HUSO (United States Geological Survey): I don't think we've explored that enough. We're pretty sure that it does – it depends on how strong the wind is on the day or at the moment when the collision occurs. Some sites are quite variable in that; some sites aren't as variable. So those that have a lot of variation in the wind regime, and then of course there's that variation in when the animals are active during the windy or not windy times. So, there's all that interaction.

I don't think we have had enough time to explore that, nor do we have the data with which to explore that yet to be able to say, sure, it's not going to make too much difference versus nope, it's pretty different from one site to the other or from one year to the next. Good thing to study though.

PAIGE JOHNSON (American Wind Wildlife Institute): Great, a next question has come in: "Can the model be run without the DWI measure?" Sorry, I don't know if DWI is what he meant to say.

MANUELA HUSO (United States Geological Survey): Probably meant DWP. It can't be run without it, but the user can do their best to estimate what they think is in that area and is a reasonable estimate for it. I don't recommend that but go ahead.

DANIEL DALTHORP (United States Geological Service): In any estimator, you're making assumptions about DWP. GenEst makes that explicit. So, if you have good information about DWP, it will handle that properly. If you have bad information about DWP, with respect to DWP, you're in the same boat that you are with other estimators and you put in bad data. The quality of your estimate will reflect how good your data are.

PAIGE JOHNSON (American Wind Wildlife Institute): Great, thanks. It looks like another question just came in. This is a follow up to one of the previous questions. They're saying: "We are creating our DWP over the entire season, so individual wind speed events don't come into it, right? If the curves are very similar across different windfarms with some known variance perhaps, a generalized DWP with some additional variance would make the calculation possible."

DANIEL DALTHORP (United States Geological Service): Currently, there's not a lot of information about how much the DWP or the distance function vary from site to site or year to year or for species types or for wind profile and in particular. So in an ideal world, yes, absolutely, we would draw up a distance function and have some variability built into that that we know; but at this stage, we just don't know. Hopefully in the future there is enough information so that we can actually have that.

PAIGE JOHNSON (American Wind Wildlife Institute): Thank you.



I just wanted to let you know that we've gotten a few comments now just thanking you, Manuela, for answering these questions and for a great presentation. So, thank you so much. Great response. Everyone is very interested, lots of questions. Again, we have a few more minutes if anyone thinks of other questions; and we can also answer questions by e-mail after the webinar is over.

I did want to remind everyone that we have recorded the webinar, and the recording will be available within a few days up on the NWCC and WREN websites.

Thank you so much to WREN for co-coordinating this webinar with us at NWCC. It was a great effort to get participants through both networks for the U.S. and North America and abroad as well.

Thanks, everybody who was able to join. We will post the links to the webinar on both sites and send the links in the follow-up e-mail that we send with the link to access GenEst as well. You'll be getting all of that information in an e-mail so that you can access everything easily.

We'll also be posting the slides on the NWCC and WREN webinars as well.

I think that probably wraps it up unless we have any more last-minute questions coming in. But it looks like we've come to the end of our questions, and we're just a few minutes away from our end time for this webinar. So, I think we can go ahead and wrap it up.

Manuela, is there anything else you wanted to share before we wrap it up?

MANUELA HUSO (United States Geological Survey): Yeah, I just want to thank everyone for your attention and for your great questions. It sounds like DWP is really high on the list for the next phase, and I agree. I think that's something that we really need to work on. So, thank you very much. If you have questions, if you are interested in being another reviewer, please contact us. We would love that.

If you have any questions or if you run into any problems with using GenEst, please contact us. Don't struggle too hard. Just contact us right away, and we'll try and help you.

PAIGE JOHNSON (American Wind Wildlife Institute): Great, thank you so much. Thanks, Manuela, and thanks to the whole GenEst Team. Thank you, everyone, for joining. We're going to sign off now, so I will say goodbye.

MANUELA HUSO (United States Geological Survey): Goodbye. Thank you.