**MEE-14-09-583 (Lucas et al.)**

**Responses to Referee’s Comments**

We thank both Referees for their valuable and constructive comments. Please see our detailed response below. Line numbers refer to the revised manuscript unless otherwise stated. Citations in square brackets are to papers listed at the end of the current document.

**Referee 1**

*Comments to the Corresponding Author*

*It is rare that I have an opportunity to review such an potentially important paper. My interest is in obtaining high quality monitoring data through large-scale citizen science. There are a number of challenges for bats and other taxa, but there have been some important and exciting developments in recent years for bats, including through other work by the authors here, that means that it is now possible to collect and analyse large volumes of acoustic data. With these data, an increasing number of projects are collecting such data, and there is challenge to do more with the information collected on bats passes, which are often used as a measure of relative abundance. For me, the most important development, would be to reliably infer true abundance from such data. This paper provides a novel approach to this problem.*

**Referee 1 Comment 1:** *The main thing that I would like to see in this paper is discussion of other analytical approaches that could be applied to this problem. Practically, I would want to know whether the approach proposed here has significant advantages over other emerging alternatives. In particular, extensions of occupancy modelling have been applied to similar data where it is not possible to uniquely identify individuals, with the aim here as well of estimating true abundance, or at least to obtain a better understanding of relative abundance.*

*Stanley & Royle 2005. Estimating site occupancy and abundance using indirect indices. Journal of Wildlife Management 69, 874-883.*

*Guillera-Arroita & Morgan (2011). Species occupancy modelling for detection data collected along a transect. Journal of Agricultural, Biological, and Environmental Statistics 16, 301-317.*

We also copy **Referee 2 Comment 5** here as it covers many of the same issues and respond to both comments together below.

**Ref 2 Comment 5**. *Comparing gREM with existing methods: could other methods for acoustic data be better recognised within the paper? Furthermore, could the comparison between gREM and existing methods be better explained? For example, in lines 87–100, existing methods are described as “require information that is often unavaliable”; however, the gREM also requires knowledge of mean speed and detection radius which may be as*

*difficult to obtain. Also, existing methods like distance sampling or capture-recapture allow for uncertain detection of animal signals whereas gREM assumes certain detection within the detection zone. It would also be useful to include description of other current methods applied to acoustic data (Marques et al., 2013), for example, in lines 422–424; and to acknowledge the applicablity of these methods, for example, in lines 431–432 by saying gREM is another possible method one can apply instead of “gREM will make it possible”.*

**Response:** [a]We have now added discussion of occupancy modelling for estimating densities of non-uniquely marked individuals in the Introduction (lines 92-100 – check line numbers in fv & tighten up the paragraph) which makes it clearer why the gREM is a significant advance.

[b] We have added more information on other methods to detect acoustic data (lines xx-xx this hasn’t been done - maybe put in paragraph starting on lines 128?, also perhaps add the suggested reference to old lines 422 etc, line 472)

[c] We have added further clarification that the information required by gREM parameters are biological species traits which only need to be estimated once (and often can be obtained from the literature) (Methods: lines xx-xx) Suggestion would be to mention that the parameters need to be measured in methods (examples of how to do that) and that we check here for error in estimates, bring up in the results and then discuss in the simulation section of the discussion).

[d] The gREM does not ignore detection uncertainty but does deal with it in a different way to existing methods which attempt to infer a detection probability from the animal count data. The gREM explicitly models some of the processes that lead to detection uncertainty (in particularly even if an animal is present, it will only be detected if it is moving towards the sensor). We have now added more explanation of this to the ms (Discussion: lines 375-381 check line numbers in fv, this needs some tightening to make it sound like an extension to model to fit into the section of the discussion better) to make this difference between methods clearer.

**Ref1 Comment 2** - *As discussed by the authors here, it would not be easy to field test how well this approach performs for species, such as bats, where this approach would be most valuable. I think that it is important to get this paper out now, but I would now like to see more extensive simulations next that do not assume an equal density across the area of interest, and make use of field data collected through tracking studies to consider how well the approach is likely to work in practice for different species.*

**Response:** Unequal animal densities will not affect the accuracy of a gas model as long as sensors are placed randomly in the environment, and, as long as a number of sensors are used, the precision will not be greatly affected (Hutchinson & Waser 2007). In short, having multiple sensors will simply average out the differing animal densities. Therefore while this is certainly an important consideration for study design, we do not think additional simulations to explore the impact of different densities are required in the manuscript. We have now clarified these points in lines 451-453 (CHECK line numbers in fv). We completely agree that an important next step to validate the gREM and quantify how effective it is for different species, but this is beyond the scope of the current manuscript. We have suggested it for further work in lines 411-412 (repetitive here – delete and include in the limitation section where it is mentioned) and 445-447 (CHECK line numbers in fv) The other stuff on limitations of the model in the rest of this paragraph seem a bit bloated and could be cut down (starting line 433).

**Ref1 Comment 3** - *The only minor typo I can find it on page 13 (particularly).***Response:** Fixed (line 364 – CHECK line numbers in fv).

**Referee 2**

*General Comments*

*The paper presents a natural extension of the random encounter model (REM) (Rowcliffe et al., 2008) and ideal gas model (Hutchinson and Waser, 2007) of animal encounters. Here, the generalised random encounter model (gREM) is proposed to allow inference of absolute density from passive acoustics and camera trap surveys. The incorporation of a variable detection zone and directionality in animal signals is a useful development, especially for analysis of acoustic survey data. However, the gREM does not account for uncertain detection of animal signals (as other methods do) and assumes knowledge of mean animal speed and directionality—requiring extra field work. Further, the gREM does not account for possibly detecting the same animal multiple times, a common occurence with acoustic surveys. I think this paper provides a good step toward further generalising encounter models to better reflect the reality to which they are applied. Yet, I also think, due to the method’s limitations, that its applicability is still to be realised. I have included some suggestions that I hope may help improve the paper.*

**Response:** We respond fully to these points below.

**Ref 2 Comment 1.** *Increasing sample size improves precision: in lines 57–60, 290–295, 301–302 and 318–320, you describe how precision improves as the number of captures increases or the signal/sensor width increases. For any reasonable estimator, one would expect its precision to improve as sample size increases; and so, could you make it clear that greater precision for a higher number of captures is a familiar result? Furthermore, you say that greater sensor/signal widths improve precision as if this were a seperate finding, yet it again is due to the number of captures increasing—could this be explicitly stated? The equivalency of the two results is implied in lines 371–372.*

**Response:** We have now made it clearer that the relationship between precision and number of captures is to be expected (lines 313-314 CHECK line numbers in fv) and added clarifications that greater sensor/signal widths improve precision because of their increased capture rate (lines 306-307 and 331-332 CHECK line numbers in fv).

**Ref 2 Comment 2.** *The effect of animal movement: in lines 60–61, you claim that different animal movement models have no effect on precision or accuracy of the gREM; however, this is contingent on perfect knowledge of animal directionality and average speed. Furthermore, it has been shown that animals with highly tortuous paths reduce precision (Gurarie and Ovaskainen, 2012). Could this claim be better qualified?*

**Response:** Our claim (original ms lines 60-61) was that we found no effect of different movement models on accuracy or precision in our simulations. We have clarified (lines 407-408 [not sure this is the best place for this – move to Discussion: limitations]) that this result is specific to the simulations we ran and that other movement models may impact accuracy and precision differently. As discussed more fully in Ref 2 Comment 4, imperfect knowledge of average speed does bias results (HOW?), but not excessively (discussed in ms at lines Results line xx-xxx, Discussion lines xx-xxx – add this in). Although we have not studied the impact of directionality error, as long as a number of sensors are used and the direction of the sensors are random as well as their locations, imperfect knowledge of animal directionality will only decrease precision, not accuracy. We maintain that tortuous movement should not affect precision (WHY NOT???) (lines 388-389 what? This is not correct line number (406-407) but maybe this should be moved with the rest of this to Discussion; limitations).

**Ref 2 Comment 3.** *Double counting with acoustic detectors: in lines 388–395, you mention the possibility of detecting the same animal multiple times and claim that this would not affect accuracy of the gREM. Multiple detections of the same animal is a common problem when estimating density from acoustic data. The gREM is derived assuming each detection is with a different animal, and so multiple detections of a single animal causes positive bias in the density estimate, as well as less precision. Not being able to account for multiple detections, gREM may not be applicable to many acoustic surveys, calling into question the claims made in lines 62, 318, 325–332. Could this limitation of the gREM be discussed in the paper? Could the model be improved?*

**Response**: As the gREM models the number of encounters rather than the number of individuals encountered, it does not assume that each encounter is with a different individual (p. 340 Hutchinson & Waser) and therefore can account for multiple detections. As an example, consider a population of animals with a density of 1/km2, moving 1 km/day, surveyed using sensors with a detection radius of 0.1 km and a total sensor time of 100 days. Using the ideal gas model, we would expect 20 contacts, and calculating density from this contact rate would correctly yield 1/km2, regardless of whether the population consisted of 20 individuals in 20 km2 (in which case the average number of captures per individual would be one, but with variation), or 1 individual in 1 km2 (in which case all 20 captures would be of the same individual). Therefore multiple detections are not a problem and are integrally dealt with by the model. We have made this property of the model more explicit in the ms (lines 178-180 and 328-330 CHECK line numbers in fv).

**Ref 2 Comment 4.**  *Required knowledge of animal directionality and the detection zone: in lines 419–421, you state that perfect knowledge of animal directionality, animal speed and sensor width and sensor radius has been assumed. As such knowledge is often imperfect; it would be useful to know how misspecification would affect accuracy. The average profile can quickly change across boundaries (shown in Figure 4 of the paper) suggesting estimates would be sensitive to knowledge assumed. Also, you do not discuss how such measurements should be obtained as best survey practice. I think some practical advice would be useful if this method is to be applied.*

**Response:** We have now added results from new simulations using parameter values of animal directionality [NOT DIRECTIONALITY?], animal speed, sensor width and sensor radius that are 1% and 10% incorrect (Appendix S6) and refer to our new analysis in the ms (Methods - lines xx-xx, Results – lines xx-xx, Discussion – lines xx-xx Add in this). A 10% parameter inaccuracy gives a less than 15% error in abundance estimate which we think is a reasonable level of precision. An incorrect parameter estimate will give an inaccurate density estimate, but the direction of the parameter error is not important (Appendix S6) so for example using a number of noisy but unbiased parameter estimates would give an accurate but less precise estimate. CHECK THIS IS CORRECT which parameter?! HAVE ALL THE PARAMETERS BEEN ASSESSED THIS WAY? The average profile does not quickly change across boundaries as any two models equal each other along a boundary (see Figure R1 below). This equality along boundaries is mentioned in lines 229-230 and 296 (CHECK line numbers in fv). We have added new text on how best to obtain parameter values for the gREM (Methods - line xx-xxx move this from discussion).

Figure R1. The value of r evaluated across a 100 x 100 grid with light to dark blue corresponding to low to high values of r, respectively. It can be seen that the changes in the function are smooth, despite the non-continuous changes in the equations.

**Ref 2 Comment 5.** *Comparing gREM with existing methods: could other methods for acoustic data be better recognised within the paper? Furthermore, could the comparison between gREM and existing methods be better explained? For example, in lines 87–100, existing methods are described as “require information that is often unavaliable”; however, the gREM also requires knowledge of mean speed and detection radius which may be as*

*difficult to obtain. Also, existing methods like distance sampling or capture-recapture allow for uncertain detection of animal signals whereas gREM assumes certain detection within the detection zone. It would also be useful to include description of other current methods applied to acoustic data (Marques et al., 2013), for example, in lines 422–424; and to acknowledge the applicablity of these methods, for example, in lines 431–432 by saying gREM is another possible method one can apply instead of “gREM will make it possible”.*

We have responded to these comments in full in Ref 1 Comment 1 and made corresponding changes to the ms.

**Ref 2 Comment 6.** *Wilcoxon signed-ranks test: in lines 256–258, 287–290, you use a Wilcoxon signed-ranks test to show that the gREM is accurate, however technically you only fail to reject the null hypothesis that the median difference between the true and estimated density is zero, you do not accept the null hypothesis or prove it. I do not think this test is necessary. Is it not sufficient to report the estimated bias of the density estimator? A relatively small bias is indicative of an accurate estimator.*

**Response:** We have removed the Wilcoxon test from the ms (lines 269-270; 300-301; 311; 320; 322 don’t think you need the line numbers as it is something you have removed – but you need to check the ‘significance’ in the sentences you have left).

**Ref 2 Comment 7.** *Reported precision: in lines 302–303, 373–378, you state that the coefficient of variation falls to 10% for all submodels at 100 captures. This result is dependent on the simulation scenario considered and, further, it ignores the uncertainty in assuming average speed, signal width and sensor width. Could it be made clearer that in practice the coefficient of variation will be much higher due to non-linear animal movement, multiple detections of a single animal and uncertain knowledge of speed and directionality?*

**Response:** The effect of nonlinear movement on the coefficient of variation is now discussed in lines 407-408 – this has been moved to Discussion check line numbers). As discussed in Ref 2 Comment 3, we do not agree that multiple detections of the same animal will have any effect. We also make it clearer that the coefficient of variation may be impacted by uncertainty in animal speed and directionality (Methods line xx-xx, Results line xx-xx, Discussion xx-xx add in).

*Specific Comments*

**Ref 2 Comment 8.1** *Line 93: “gas model”— calling it the “ideal gas model” could be clearer and a citation, useful. See (Hutchinson and Waser, 2007).*

**Response:** We have reworded as the ideal gas model and added a reference (line 102 CHECK line number in fv).

**Ref 2 Comment 8.2** *Line 160: “animals outside the zone are never captured”— could this be made clearer? Animals outside that then move into the detection zone are captured.*

**Response:** We have reworded this as “while outside this zone, animals are never captured” (line 170 CHECK line number in fv).

**Ref 2 Comment 8.3** *Line 228: what value does r take in the simulation? It would be useful to know for reproducibility..*

**Response:** We have added the value (10m) to the ms (line 240 CHECK line number in fv).

**Ref 2 Comment 8.4**  *Line 237: “individuals were assigned 11 signal widths”— this sounds as if signal widths were randomly allocated to each individual for every simulation; however, I think you mean that a separate simulation was run for each signal width? Clarify?*

**Response:** We have reworded as “11 signal widths α between 0 and π were used” (line 249-250 CHECK line number in fv).

**Ref 2 Comment 8.5** *Line 242: σ d = vT10 — why was this standard deviation chosen? Is it because a normal distribution was used to simulate a strictly positive parameter (distance)? Could a biological justification be given, as you do for the other simulation parameters chosen?*

**Response:** σ d = vT/10 was chosen so that the standard deviation of the movement scaled with the speed of the movement. The ms has now been changed to “standard deviation, σ d = vT/10 where the standard deviation was chosen to scale with the average distance travelled” (line 254 -255 CHECK line number in fv).

**Ref 2 Comment 8.6** *Line 243: “this is the largest day range”— here you refer to an average velocity while day range is a distance, might be better to say average velocity was chosen based on a day range of. . .*

**Response:** Ms is now reworded in line with this comment (line 256 CHECK line number in fv).

**Ref 2 Comment 8.7.** *Line 246: “or change direction (in a uniform distribution with maximum angle A)”— I understand you are defining a correlated random walk, but this sentence could imply you are sampling the new direction randomly from the interval (0, A). Clarify?*

**Response:** We have changed the ms to “(change in direction has a uniform distribution in the interval [-A, A])” (lines 259-260 CHECK line numbers in fv).

**Ref 2 Comment 8.8.** *Line 249: Might be best to say it is a correlated random walk for (3) as you do later on in the paper*

**Response:** Reworded in line with this comment (line 263 CHECK line numbers in fv).

**Ref 2 Comment 8.9.** *Line 251: “were counted as they moved in and out of the detection zone”— I assume animals were only counted as they entered and not counted again as they left, also animals initially inside the detection zone would also be counted.*

**Response:** Ms changed to “were counted as they moved into the detection zone” (line 264-265 CHECK line numbers in fv).

**Ref 2 Comment 8.10.** *Line 253: “by assuming the number of captures per simulation” — Why would you assume the number of captures? You obtain the number of captures from the simulation. Clarify?*

**Response:** Assuming was a typo, and was meant to say “summing”, line 266 (CHECK line number in fv) has been changed to reflect this.

**Ref 2 Comment 8.11**. *Line 256: “the density in the simulation”— better to say the true density to prevent confusion with the density estimated from the simulated data?*

**Response**: We have changed this to “by comparing the true simulation density with the estimated density” as suggested (line 268-269 CHECK line numbers in fv).

**Ref 2 Comment 8.12.** *Line 266: “we calculated how long the simulation needed to run. . .” — do you mean how long on average, as the formula applies to the expected number of encounters.*

**Response**: Line 277-278 changed to: “From a random starting point we ran the simulation until a range of different capture numbers were recorded (from 10 to 100 captures), recorded the length of time this took, and estimated the animal density for each of the four sub-models.” (lines 277-280 CHECK line numbers in fv).

**Ref 2 Comment 8.13**. *Line 269: “real density” — perhaps use true density for consistency?*

**Response**: Changed (line 280 check in fv).

**Ref 2 Comment 8.14**. *Line 271: “compare the precision between capture numbers” — I think more detail is* *required: compare the precision of the density estimates from simulations with a different expected number of captures.*

**Response**: Edited as suggested (line 281-283 CHECK in fv).

**Ref 2 Comment 8.15**. *Line 275: “using different amounts of time spent stationary” — should this not be the time spent stationary on average? (This applies to all other instances of “time spent stationary” in the paper.)*

**Response**: Added 'average' in Fig.7 caption, line 288 and line 317 CHECK lines in fv.

**Ref 2 Comment 8.16**. *Line 338: “the sum of the average animal velocity and the sensor velocity” — could be clearer? It is the average animal velocity (averaged over all bearings) relative to a single (animal or sensor) focal individual (Hutchinson and Waser, 2007).*

**Response**: We changed to “by replacing animal speed v with v + v\_s where v\_s is the speed of the sensor” in line 349-350 CHECK in fv

**Ref 2 Comment 8.17**. *Lines 388–395: you begin by saying accuracy may be affected by the interaction between the movement model and detection zone, then end saying the opposite, in brackets.*

**Response**: We have changed accuracy to precision (line 412 CHECK in fv). It is correct to point out that the accuracy of the model should not be affected by this interaction, however the precision of the model would be.

**Ref 2 Comment 8.18**. *Line 411: “entire world”— better to say study area?*

**Response**: Changed (line 448 CHECK in fv).

**Ref 2 Comment 8.19**. *Line 416: “The distance travelled for an animal was assumed to be 40 km day −1 ”— technically 40 km day −1 is a speed, also isn’t this the average distance per day?*

**Response**: Changed to: “In the simulation animals moved at a speed of 40km day-1, equivalent to the largest day range of terrestrial animals” (line 455-456­ CHECK in fv).

**Editorial Comments**

1. *Line 167: Typo — “number of expected captures” should say expected number of captures.*
2. *Lines 174-176: this sentence does not make sense.*
3. *Line 226: perhaps better not to start a sentence with a number? Might be better to use*

*words: one hundred.*

1. *Line 228: “exact centre of the simulation”— does this mean the centre of the simulated study area?*
2. *Line 241: better not to start a sentence with a mathematical symbol?*
3. *Line 245: Typo — At the end of each step*
4. *Lines 255, 258: perhaps better to not start a sentence with lower-case “g” in gREM?*
5. *Line 282: this sentence does not make sense.*
6. *Line 310: “pi” — use symbol π*
7. *Line 430: “which are important reservoir of infectious disease”— reword? For example, which are an important reservoir of infectious diseases*

**Response:** We have addressed all these editorial comments.

**References for response to referee's comments**

[1] Hutchinson, J.M.C. & Waser, P.M. (2007) Use, misuse and extensions of “ideal gas” models of animal encounter. Biological Reviews of the Cambridge Philosophical Society, 82, 335–359.