Clément CALENGE
Unité Données et Appui Méthodologique (UDAM)
Direction Surveillance, évaluation, données (DSUED)
Office Français de la Biodiversité (OFB)
Saint Benoist - 78610 Auffargis
e-mail: clement.calenge@ofb.gouv.fr

January 10th, 2024

Dear Editor,

Please find a revised copy of our manuscript that we wish to publish as an article in Peer Community In Infections: "*Assessing the dynamics of* Mycobacterium bovis *infection in three French badger populations*"; Calenge C., Payne A., Réveillaud E., Richomme C., Girard S., Desvaux S.).

We really would like to thank the two referees for the constructive and extensive feedbacks on the first submission. We have considered all the remarks of the referees, and we describe below how we accounted for these remarks in our revision. Note that we have sent the manuscript to AJE (https://www.aje.com/) for language editing by two native speakers.

This manuscript has been written with LaTeX and the official PCI article class.

We confirm that the work as submitted has not been published or accepted for publication, nor is being considered for publication elsewhere, either in whole or substantial part. Furthermore, all authors and relevant institutions have read the submitted version of the manuscript and approve its submission.

We hope our manuscript will meet your approval and we are looking to hear from you.

For the authors,

Sincerely,

Clément CALENGE

Response to Reviewers for the revision of the article 'Assessing the dynamics of Mycobacterium bovis infection in three French badger populations'

Clément Calenge Ariane Payne Edouard Reveillaud
Céline Richomme Sébastien Girard Stéphanie Desvaux

January 2024

Remarks of Reviewer #1

Comment 1-1: *It is not clear, but I do suspect, that the analysis will be sensitive to the trappability of badgers, and we can only assume that it will be similar for infected and non-infected badgers. This could be important as it is stated that only two badgers per commune per year are analysed.*

We agree with the referee that this issue should be discussed in the paper. We therefore added in the Methods section (line 146):

"*Note that our statistical approach assumed the ignorability of the sampling; in other words, we assumed that infected and noninfected badgers are characterized by equal trappabilities.*"

And we discuss this hypothesis in the discussion (line 577):

We assumed equal trappability between infected and noninfected badgers. However, previous studies have shown that the trappability of badgers may be influenced by factors such as weather, season or age class (Byrne et al. 2012, Martin et al. 2017). Therefore, trappability might also vary based on other individual characteristics, and particularly the infection status of the animal, although we did not find any study supporting this hypothesis. In addition, other factors related to the infection status of badgers may indirectly affect their trappability. Thus, several studies suggest that infection can lead to behavioral changes in badgers, making them more solitary and mobile, with larger home ranges (Cheeseman and Mallinson 1981, Garnett et al. 2005; Weber, Carter, et al. 2013a). In particular, greater mobility of infected animals was observed in the three clusters in our study, leading to an increased risk of being killed by cars; the proportion of infected badgers is greater in animals killed by cars collected on the side of

roads than in trapped badgers (unpublished results). This greater mobility of infected badgers may increase their exposure to traps. However, even if there was a lingering bias in the prevalence estimation, there is no indication that this bias varied among the three clusters or between years. Therefore, it is reasonable to assume that the situations can be compared consistently across clusters or between years."

Moreover, regarding the limit of two badgers per commune per year, see our response to comment 1-6.

Comment 1-2: *On page 3 (Introduction) it mentions that a commune is the smallest division of administration, however, it is on page 6 (materials and methods) that the median size of a commune is defined. I think these values should accompany the statement in the Introduction as it is an obvious question that arises.*

We corrected the manuscript accordingly, precising (line 52): "*a commune being the smallest French administrative subdivision, with median area of 12 km*^{2}"

Comment 1-3: *Perhaps some comments could be made on the consistency of trapping in the communes across the three regions. This would include the types of traps used (cages?), the number of traps used per commune, the number of trap nights etc..*

We now precise this information in the paper. We note line 120:

"*The regulatory guidelines for badger trapping are uniform across the three clusters, and are defined by the French Ministry of Ecology and Sustainable Development, through the ministerial order of January 29th, 2007. Only two types of traps are authorized in France: stop snare (i.e., snare with a mechanism that stops the noose from closing too tightly) and cage traps. Night shooting is also an option in level 3 communes. Nevertheless, with a few exceptions, French trappers predominantly utilize stop snares. Given the participatory nature of the Sylvatub program, local trappers retain the autonomy to decide on the number of traps, trap nights, and their placement. However, the Sylvatub program encourages however trapping near infected farms (technical directive from the French Ministry of Agriculture DGAL/SDSPA/2018-708).*"

Comment 1-4: *Among all of the badgers captures in each commune / region, how were the two study animals from each commune chosen?*

We now precise (line 132):

"*The choice of analyzed animals among trapped badgers was left to the local partners of the network. Given that the infection status of trapped badgers is seldom discernible from external observations (as most TB lesions diagnosed in badgers are internal, as noted by Reveillaud et al. 2018), we are confident that there was no sampling bias directly related to the infection status of animals. While trapping efforts were intensified near infected farms to control the density of badgers in proximity to these areas, Sylvatub guidelines encouraged the analysis of badgers distributed spatially as uniformly as possible. In practice, we observed that the badgers of a commune sent for analysis were often the first two trapped badgers.*"

Comment 1-5: *During necropsy, were the samples collected from each animal based on a pre-selected list of tissues, or did it depend on the lesion status of the animal, i.e., were more issues likely to be collected from an animal presenting with lesions.*

Yes, there is a pre-selected list of tissues, and it did not depend on the lesion status of the animal. Perhaps this wasn't explicitly clear in the previous version of the manuscript; we made an effort to enhance the clarity of the analysis protocol (line 166):

"*Following necropsy, two types of first-line tests were carried out on the animal samples. Pools of lymph nodes (retropharyngeal, pulmonary and mesenteric) and organs with gross lesions were used in the analysis. The type of analysis depended on the period: (i) from 2013 to 2015, the first-line test was the bacterial culture performed on the sampled tissues of all tested animals, following the protocol established by the French NRL (NFU 47-104) for the isolation of* M. bovis*; (ii) since 2016, the first-line test has been real-line PCR performed after DNA extraction from the sampled tissues.*"

Comment 1-6: *A key finding from the study was the low correlation between the infection status of badgers in the same commune was very low. Given that only two badgers were tested in each commune there could be considerable error in this estimate.*

We apologize for our lack of clarity regarding this recommended limit of the number of badgers trapped per commune. We now try to provide a more precise explanation in our manuscript. The guideline of two sampled badgers per commune per year was intended as a general recommendation rather than a strict fieldwork requirement. Fieldworkers generally adhered to this guideline, but approximately 25% of the communes trapped and analyzed more than 2 animals per year in average. In all three clusters, between 12% and 13% of the communes trapped and analyzed more than 21 badgers over the 7 year period, and between 3 and 4% of the communes trapped more than 36 badgers during this period; allowing a precise assessment of the intra-class correlation. As a result, the sample size exceeded two per year in many communes.

Moreover, the estimation of the intra-class correlation utilizes all the data collected in a commune, not solely the data from a single year. A key result of our analysis is that the prevalence of the infection within a commune does not vary substantially from one year to the next (it takes several years for the prevalence to change significantly). Thus, the intra-class correlation is not estimated solely with two animals per commune, but with animals trapped during several years.

Finally, even if only two animals had been trapped per commune in a single year, our extensive number of communes within a cluster would still have allowed to estimate precisely the intra-class correlation (in statistical terms, the high number of degrees of freedom facilitated the precise estimation of one parameter).

To illustrate this idea, consider a hypothetical scenario with only one year of data for the 444 communes of the Dordogne/Charentes cluster. Assuming a constant prevalence equal to 0.07 in all the communes (mean prevalence over the cluster area), a correlation parameter $\rho = 0.05$, and only two animals trapped per commune, simulations indicate that one should obtain in average 3.6 communes out of the 444 where both trapped animals are infected. In addition 95% of the simulations with this intraclass correlation are characterized by fewer than 8 communes with both sampled animals that are infected. On the other hand, if the intraclass correlation were slightly larger, i.e. $\rho = 0.2$, the mean number of communes where both sampled animals are infected would be larger (8 communes), and 95% of the simulations would be characterized by more than 4 communes where both trapped animals are infected. In other words, even in a hypothetical situation with only one year of data and only 2 animals per commune (substantially less than the 7 years available in our study), the data would be sufficient to distinguish between a very low intra-class correlation and a moderately low intra-class correlation (based on the number of communes where the two trapped animals are infected).

Therefore, our estimation of ρ is not characterized by a considerable error (on the contrary!).

We now precise that the limit of two sampled animals per commune was more

intended to be understood as as a general recommendation rather than a strict fieldwork requirement, as this limit was misunderstood by the two referees (see also comment 2-2). We added (line 141):

"*This guideline to test at most two animals was intended more as a general recommendation rather than a strict fieldwork requirement. Although fieldworkers generally adhered to these guidelines during the study period, approximately 25% of the communes trapped and analyzed more than two animals per year on average. In all three clusters, between 12% and 13% of the communes trapped and analyzed more than 21 badgers over the 7-year period, and between 3 and 4% of the communes trapped more than 36 badgers during this period.*"

We also insisted on the precision of estimation (line 412):

"*The abundance of data available in the three clusters results in precise estimations, as evident from the narrow width of the credible intervals for all parameters in this table.*"

Comment 1-7: *Although some potential reasons were given, is there any information available on the variation in estimated mean social group size among, for example, neighbouring communes (where they might be predicted to be similar). Do the results provide any information on social group size, density between the regions? The cited paper by Delahay (2000) describes infection in a very high-density population in the UK, which is likely not applicable to most populations (though I acknowledge the point the authors are making here).*

We now include a paragraph in the Material and Methods section to give the information that we could gather on these aspects (line 157):

"*Note that we lack precise information regarding the social group size of badgers in the three clusters. Jacquier et al. (2021) employed a standard methodology, utilizing camera traps and genetic identification, to estimate badger density across multiple sites in France, including the three clusters of interest. These authors showed that the badger density was highest in the Dordogne/Charentes cluster (6.18 badgers / km^2), followed by the Bearn cluster (5.39 badgers / km^2), and the Burgundy cluster (two sites were studied by these authors in this cluster and were characterized by a density of 4.08 and 4.22 badgers / km^2). For comparison, the mean density across the 13 sites studied by these authors, distributed across the entire metropolitan region of France was 5.85 badgers / km^2 – SD = 3.25 badgers / km^2).*"

Comment 1-8: *There are parts where the syntax was not clear and would benefit from a review by a native English language speaker with a good knowledge of this field of research.*

This concern was also raised by referee 2 (comment 2-7). In response, we have sent the manuscript to American Journal Expert for a review and improvement of its English quality. We hope that the English quality now meets the required standards of English language proficiency.

Remarks of Rowland Kao

Comment 2-1: *It would be very useful to have a more extended qualitative description of the data, which would help the reader to understand better what is actually being fitted. The maps themselves are helpful, however temporal and within commune descriptions would also be useful.*

We have added a new paragraph at the beginning of the results section, introducing a table that illustrates the variation in apparent prevalence (i.e., the proportion of trapped animals diagnosed with *M. bovis*) across the study period, to provide these temporal descriptions. Line 403:

"*The number of animals trapped in each* M. bovis *cluster during each year of the study period is presented in Tab 1, as is the proportion of these animals diagnosed as infected with* M. bovis*. Note that even though it is challenging to interpret the observed temporal changes in prevalence (as this proportion does not account for all the factors that influence the prevalence, i.e., inhomogeneous prevalence patterns in space, sensitivity of the tests used increasing with time, etc.), this table clearly demonstrates the overall temporal change observed in each cluster, i.e. a strong increase in Dordogne/Charentes, a decrease in Burgundy, and a moderate increase in Bearn.*".

As noted in our response to comment 1-6 from referee 1, we have also enhanced the description of the available data at the commune resolution, providing a more detailed account of the number of animals captuured per commune and per year.

Comment 2-2: *Line 172 - Not necessary for the study but as a suggestion - the limitation of two tests per community per year seems very low and so I would think this would result in very low confidence in some statistical outcomes– is this true? Of course the authors are limited to the data available to them, however a useful outcome of this study might be an assessment of the utility of increasing this.*

This concern was also raised by referee 1 (comment 1-6). As noted in our response to comment 1-6, we acknowledge the lack of clarity regarding the recommendation not to exceed two sampled animals per commune and per year. As noted above, this was intended as a general guideline rather than a strict fieldwork requirement. Fieldworkers generally adhered to this guideline, but approximately 25% of the communes trapped and analyzed more than 2 animals per year in average. In all three clusters, between 12% and 13% of the communes trapped and analyzed more than 21 badgers over the 7 year period, and between 3 and 4% of the communes trapped more than 36 badgers during this period. As a result, the sample size exceeded two per year in many communes (refer to comment 1-6 for an updated explanation in the paper).

Note that although the number of animals per commune and per year was limited, there are many communes in a cluster area. Consequently, the total number of animals available in each cluster is substantial (see the new Table 1, added in response to comment 2-1), especially when compared to the number of parameters to estimate in each cluster. The abundance of data contributes to the high precision of our estimations, as evident from the narrow width of the credible intervals in Table 2. We now insist (line 412):

"*The abundance of data available in the three clusters results in precise estimations, as evident from the narrow width of the credible intervals for all parameters in this table.*"

Comment 2-3: *Line 183 onwards and figure 2. It's a bit hard to tell (and in part the reason for the request for a qualitative description of the data) but it looks like the spatial distribution shows very little structure in each area. While I am not an expert in spatial statistics, I do wonder if using random effects per community (as a minor point, unless 'commune' is a standardised term I am not aware of, I think 'community' would be more usual) may be more than is required to describe the data. It may also result in model overfitting – this would be worth testing.*

First, note that the term "commune" is actually a standardised term: it corresponds to the smallest administrative district in many European countries

(Belgium, Germany, etc.) and in particular in France (see e.g. here. There is an official list of communes in France), and these communes are often used as basic statistical units in survey sampling for official statistics. We tried to precise the meaning of "commune" in the introduction of the paper, noting (line 52) "*a commune being the smallest French administrative subdivision, with median area of 12 km^2*".

Furthermore, we are uncertain about the referee's impression that the spatial distribution exhibits very little structure in each area. On the contrary, the spatial structure is highly prononced:

- in all three clusters, the set of "highly infected communes" (i.e. the set of communes characterized by a random effect greater than average) formed a perfectly connected subset of communes in the Burgundy and Bearn clusters, constituting a subregion within the cluster area. In the Dordogne/Charentes clusters, it formed a nearly perfectly connected subset of communes as well, with only two communes not immediately connected to the the main subarea but located a few kilometers away. This is expected, as the iCAR model inherently identifies a set of random effects that **are** spatially structured by definition. Therefore, it is normal for the set of "highly infected communes" to be connected to each other, defining a subarea within each cluster.
- But the spatial structure identified by the fit is not artefactual: this set of highly infected communes is characterized by a higher average number of infected animals. Comparing the proportion of animals with a diagnosed *M. bovis* infection in highly infected communes to other communes reveals a substantial difference. In the Dordogne/Charentes cluster, focusing only on the three last years (to minimize change in prevalence over time), an infection was diagnosed in 16% of trapped animals in highly infected communes, whereas it was diagnosed in only 2% of trapped animals in other communes. In Bearn over the same period, an infection was diagnosed in 11% of animals trapped in highly infected communes, whereas it was diagnosed in only 0.75% of the animals trapped in other communes. In Burgundy, focusing this time on the first three year of the study period (when the prevalence of *M. bovis* was still high), an infection was diagnosed in 10.6% of the animals trapped in highly infected communes whereas it was never diagnosed in other communes.

Thus, in each cluster, we identified a connected subarea comprising half of the cluster's communes (termed "highly infected"), where the infection was substantially more prevalent than in other parts of the cluster. This demonstrates the strong spatial structure. Maybe this point was not clear enough in the previous version of the manuscript, so that we now precise these elements (line 419):

"*Conversely, a strong spatial structure was evident in all three studied clusters, revealing distinct patterns of highly infected areas and low-risk areas within each*

cluster (see Fig 2). Specifically, the set of highly infected communes formed a connected subset of communes (i.e. a unique subarea) in the three clusters, except the Dordogne-Charentes cluster, where two highly infected communes were located only a few kilometers away from the main subarea. Furthermore, the proportion of trapped animals diagnosed as infected was greater in the highly infected communes than in the other communes (focusing on 2017–2019 to limit temporal changes: 16% in highly infected communes of Dordogne-Charentes vs. 3% in other communes; 11% in highly infected communes of Bearn vs. 0.75% in other communes; and focusing on 2013–2015 in Burgundy, when the infection rate was still noteworthy; 10.6% in the highly infected communes vs. 0% in the other communes)."

Comment 2-4: *Line 202. The simulation analysis is welcome and a useful tool. Could the authors better justify the choice of the two scenarios. Also, it might be helpful to identify ways of 'breaking' the model – a few possible considerations are greater heterogeneity in spatial risk, and examining the impact of patterns at different stages of an epidemic process – e.g. both early on and later, greater heterogeneities can be observed, depending on the process.*

We justified more the choice of the two scenarios. We also concur with the referee's suggestion that the simulations should encompass scenarios designed to break the model. Consequently, we introduced in the paper a third set of simulations representing situations anticipated to yield challenging outcomes, testing the effect of greater heterogeneity in spatial risk, spatial patterns in risk changing with time, and structure in badger sampling.

Specifically, we added in the Material and Methods (line 315):

"*We carried out three sets of simulations to assess the ability of the simpler regression model to estimate the two target quantities. The first set was designed to assess the ability of our regression model to estimate the trend over time of the prevalence in various situations that might be encountered in reality (that is, either an initially rare but increasing infection or an already widespread infection with different trends). The second set was designed to assess the ability of our regression model to estimate the mean prevalence level of the infection at a variety of actual levels (from rare to very frequent infection). The last set was designed to assess the robustness of our approach to violations of the hypotheses on which it relies (strong spatial heterogeneity remaining even in highly infected communes, spatial structure of the infection changing with time, nonrandom sampling).*"

We also added in the Material and Methods section a description of this third set of simulations (line 370):

"*In the third set of simulations, we aimed to assess the robustness of our model*

to the violation of two underlying hypotheses: (i) ignorability of the remaining spatial structure of the prevalence when the regression model is applied only to the data coming from highly infected communes and (ii) additivity of the space and time effects on the prevalence. In these two situations, we simulated the data with our Bayesian model using two different values of the intercept $\alpha = -2, 0$*, representing different mean prevalence levels. We then randomly sampled a slope* β *in a uniform distribution bounded between -0.4 and 0.4. To test the effect of the violation of the first hypothesis, we simulated random commune effects* u_i *using Equation (4), setting a very low value* $\tau = 0.1$*, corresponding to very strong spatial heterogeneity. To test the effects of violating the second hypothesis (additivity of space and time effects), we simulated the spatial structure of the infection changing with time. More precisely, we simulated two sets of commune effects,* $\{u_i^{(1)}\}$ *and* $\{u_i^{(T)}\}$*, describing the spatial structure at the start and end of the study period, respectively (using* $\tau = 0.73$ *in both cases). The set of random effects used at time t was calculated by* $\tilde{u}_i^{(t)} = ((t-1)/6) \times u_i^{(1)} + (1-(t-1)/6) \times u_i^{(T)}$*. In the two tested situations, we estimated the two parameters of interest (intercept and slope of the regression model) and compared them to the theoretical values used for simulation. In this third set of simulations, two sampling schemes were compared to demonstrate how directed sampling can exacerbate the effect of the violation of underlying hypotheses: random sampling with* $\mu_i = 2$ *and directed sampling where the mean number of animals in a commune was proportional to the mean prevalence in the commune, i.e.* $\mu_i = 2 \times M \times \exp(u_i)/(\sum_j \exp(u_j))$ *(where* M *is the number of communes)..*"

We also added in the results section a new figure and a new table showing the result of these simulations, and noted (line 457):

"*Finally, the last set of simulations showed that as long as the sample of trapped animals can be considered a random sample from the population, the model is robust to violations of the underlying hypotheses (Fig 5 and Tab 5). However, when the animals are preferentially trapped in places where the prevalence is high, the mean prevalence is overestimated (and this bias will be greater when the spatial heterogeneity is strong), and the mean proportion of animals becoming infected in one year will also be biased (although this bias is much smaller than the bias affecting the mean prevalence, and can be ignored for moderate spatial heterogeneity). Similarly, nonrandom sampling can generate bias in the estimation of the two parameters when both the spatial structure changes with time and when the sampling is directed toward highly infected communes. Note that in our study, the sampling intensity was uncorrelated with the commune random effects in the Dordogne/Charentes (Pearson correlation coefficient between the number of trapped badgers and* u_i*, R = -0.02) and the Bearn (R = 0.04) clusters, whereas the trapping effort exhibited a slight preference for the most infected communes in the Burgundy cluster (R = 0.35).*

And finally, we added in the discussion, regarding the assumption of stability of the spatial structure (line 531):

"*Simulations showed that a mild violation of this assumption does not impede its ability to assess the average situation in a cluster (mean prevalence and mean trend in prevalence), provided that the sample of trapped badgers can be considered entirely random, a condition we show to be approximately valid in our study (i.e., weak correlation between the sampling pressure and the prevalence of* M. bovis *infection).*"

and (line 562):

"*Therefore, we recommend focusing only on highly infected communes when applying the regression model, assuming that the remaining spatial variability within this subset of communes is negligible. Note that our simulations showed that even in the presence of a substantial remaining spatial structure, there was no detectable bias in the estimation of the two focus parameters (mean proportion of the population becoming infected in one year and mean prevalence during the middle year), provided that the sample of trapped badgers could be considered completely random. When sampling is directed toward communes with the highest infection prevalence, a substantial remaining spatial structure within these highly infected communes will result in the preferential sampling of infected animals. Neglecting the spatial structure of the infection in the regression model then leads to an overestimation of the mean prevalence during the middle year and a biased estimation of the proportion of the population becoming infected in one year. Therefore, monitoring programs intending to use our regression approach should pay attention to maintaining uniform trapping pressure across a clusters' entire area. In our study, the correlation between the level of infection in a commune and the sampling effort remained low, suggesting a very limited bias in our estimation.*"

We also have corrected the supplementary material to include the code used for these simulations.

Comment 2-5: *Line 243 onwards. The rationale for the simpler approach is clear, and the authors do a good job of demonstrating its utility. However the limitation of no-spatial structure in the model is potentially an important one and so I wonder if in particular, it means that stable spatial structure (even if prevalence is changing, the relative incidence across areas may be stable) is important for this to work (see comment below as well, on proposals for working in new areas and also this reiterates the importance of the qualitative desciption - I can't really tell how the real data are varying in time and space).*

The hypothesis of a stable structure is indeed an important hypothesis in our model, and there is a large part of the discussion that discusses this assumption. However, rather than "stable structure", we talked about "the additivity of the time and space effects on the prevalence" (a more statistically inclined

terminology). We now precise that by "additivity of the time and space effects" we mean "stable structure". Line 522:

"*Our complex model identified a very marked spatial structure of the infection in the three studied* M. bovis *clusters, and both our complex model and the simpler regression approach assumed that this structure was stable in time (i.e., the areas with the highest prevalence remained the same every year; even if the mean prevalence increased or decreased in time, it changed in the same way everywhere). In statistical terms, we assumed that time and space had additive effects on the prevalence.*"

And combining the recommendations of this referee in comment 2-4 and in this comment, we also tested the effect of the violation of this assumption in the new set of simulations (see comment 2-4). Note that the violation of this assumption does not result in any bias in the estimation of the two parameters of interest. Consequently, our approach seems robust to moderate violations of this assumption (see our response to comment 2-4).

Comment 2-6: *Line 486 – the proposal for working with new areas is important if the method is to have broader applicability – a test/train approach could be used to assess the utility here – i.e. how much of a dataset for the complex model is needed to predict the future trends? (To understand the utility of the model for working with new data, it might be worth doing a test/train split (i.e. how good is the correlation between the two models based on a previous model iteration?).*

We disagree with the referee on this matter: a test/train approach is useful when the objective of the model is statistical prediction. Typically, one uses a set of predictive variables in a model to predict an outcome, and test/train approaches are employed to assess the model's ability to predict the outcome.

However, in our study, our goal is not prediction but estimation. We do not aim to predict future trends, as the referee suggests. We have two models; one is far too complex for local workers to fit, and the other is more accessible. The question we are addressing is whether these two models can be used to estimate the mean prevalence during middle year and mean proportion of animals becoming infected in one year over a given study period. Test/train approaches would not bring any useful information regarding this question.

To understand the utility of the model for working with new data, the most effective approach is still to simulate a wide variety of situations (such as widespread infection with different possible trends, the emergence of an infection expanding in a population, etc.). As the referee rightly noted in comment 2-4, the aim is to identify factors likely to result in differences between these two models (to "break the model", such as the combination of strong spatial heterogeneity

with strongly oriented sampling). We therefore left the manuscript unchanged on this point.

Comment 2-7: *A minor point – the paper overall is to me very clear in the description and the authors should be commended for this. There are however, minor examples throughout of non-standard English usage and it might be useful, for peer review to have someone do a proof read of the manuscript.*

This concern was also raised by referee 1 (comment 1-8). In response, we have sent the manuscript to American Journal Expert for a review and improvement of its English quality. We hope that the English quality now meets the required standards of English language proficiency.