Editor decision

Dear Kees C.H. van Ginkel and co-authors,

thank you very much again for submitting your interesting manuscript 'Direct flood risk assessment of the European road network: an object-based approach' to NHESS. After the evaluation of the major revisions of your manuscript, three referee reports are now available for this version of the manuscript.

As you can see from these reports the reviewers are mostly satisfied with your revisions. However, some minor points remain open as stressed by the reviewers #3 and #4. These comments include a number of useful suggestions to improve your manuscript and make it stronger.

I kindly ask you to take the specific comments of both reviewers into thorough consideration and rework the manuscript accordingly. Therefore, I decide on minor revisions.

Please provide a revised marked up (track changes) version of your manuscript to make clear how you include the changes in response to the referee reports.

Please accept my sincere apologies for the delay in completing the review process for your paper. Not all referees of your original submission were available to review the revised manuscript and it was challenging to secure two new referees during complicated times and a busy period of the year.

I look forward to receiving your revised manuscript for review by editor only.

Yours sincerely

Kai Schröter

Dear dr. Schröter,

Thank you for your supervision of the review process. Below you will find our point-to-point replies to the requests and remarks made by the referees. The numbers between brackets [] refer to the track-changes version of the manuscript, where you can find how we processed each comment.

We are looking forward to your response on this revised version of the manuscript.

Best regards,

On behalf of all the co-authors,

Kees van Ginkel

Anonymous referee #3 (Report #1)

Overall, this is a very interesting and complex case study. I recognise the multi-institutional and -discipline effort, as well as the European scale of it. However, I find also major issues that I outline below. Despite I value the work in its nature, my review is critical because it identified these flaws that appeared straightforward to me.

We want to thank the reviewer for the careful consideration and the in-depth review of the paper. Below we will provide point-by-point replies to the issues raised.

1) What is has been defined as "object-based" approach is NOT a new approach. Many works in literature (included some of those cited by the paper) assessed the water depth by overlaying road links and flood footprint; I would say this is a "normal procedure" of spatial analysis at urban scale. Perhaps, what is new here is the application of the method at European scale and the damage curves. So, I would revisit the paper and omit the "object-based" approach is new. *[1]*

We agree with the reviewer. The object-based approach itself is indeed not new, but the examples we found in the literature all have a limited spatial scope; the continental-scale application is new. We have more carefully phrased this throughout the manuscript, by omitting the word 'new' when discussing the object-based approach (e.g. line 117). We also explicitly added that object-based approaches are already frequently used at the urban scale, with some additional references (line 79-80, also line 105).

2) the resolution of the hazard domain is 100m (!). I am wondering if this is compatible with an assessment of flood impact to roads. Usually for "object-based" assessment analysis to roads in cities the resolution is ~5m (or less). This because the uncertainty of the model simulation (cm? dm? m?) should be compatible with the road dimension (two lanes road, ~6m) and the scale of the object. So perhaps, applying such a coarse flood hazard to a detailed exposure could not make sense. This point is even not discussed or considered in the paper, and I expected authors to address it properly. *[2]*

*We agree that due to the more precise exposure data, the resolution of the hazard data has now become the new bottleneck for continental-scale assessment. This on itself is a step forward, because typically it is the exposure and vulnerability that is relatively underdeveloped (see e.g. lines 67-74). Indeed it would be much better to use higher-resolution flood data, such is available for smaller spatial scales. Unfortunately, such data is currently not available on the continental scale, where arguably 100*100 m is state-of-the art; we have elaborated this point in the response to referee #2 in the first review round (comment 26).*

We already briefly touched upon this point in lines 577-579, 611-615, and have now elaborated this on lines 544-549.

On the other hand, there is also a benefit in the use of the rather coarse hazard dataset, because it enables a fair comparison to the grid-based approaches which are carried out on the same resolution, and which has been extensively used to inform European flood risk policy.

3) I don't understand, and I invite authors to explain, why Sec. 5 is about areas that are different compared to Fig. 5. *[3]*

We don't fully understand to which section the referee is referring. Figure 5 (now figure 6) covers almost all European countries (footnote 1 gives a complete list). We don't discuss areas that are not in this figure.

4) I understand that Sec.3.4 is about offering a sort of validation of the "new" damage curves. However, I see major criticalities in validation damage curves applied at EU level with a 7x7m reference case. In particular, because this refer to one road type only (motorway), whereas damage curves were developed for six road types (Table 1). Also, even assuming this reference case as appropriate, the model estimation ranges from €3.4m to €28.6m (!) – when the comparison value is €3.8m. I can't accept this refence case as validation of all the curves (or even the motorway one). *[4]*

We agree that reference case has serious limitations and cannot serve as a full validation of the entire European situation and also not of all the curves. We also do not claim it to be. But we do feel it is important that the readers gets a feeling of how the results in this study should be interpreted. The fundamental problem here is that better validation data is not available. Collection of such analysis is of key importance to bring the research field to the next step. This point we already stressed in the introduction (line 66); and conclusion where we emphasized that our model is 'a starting point for further analysis' (line 611) and that 'local scale case studies are required to validate the proposed damage curves. Currently, very little road flood damage case studies are described in the literature, collection of such data by road operators and academia should be a research priority because the absence of damage data hampers the validation of the models.' (line 611-618)

We have added a new paragraph in the discussion (line 590-595) with a disclaimer for using the damage curves, and an invitation to the scientific community to further calibrate and validate the curves if they have access to more data.

Perhaps, this work is trying to achieve too much in just one paper. The "new" damage curves could be a paper on their own, if well-documented and validated. *[5]*

We fully agree that a detailed study on well-documented and validated damage curves can serve as a paper on its own. Because there currently is insufficient validation data available, we have sought to present everything we could find as transparent as possible in the SI, so that researchers with access to more data can take a starting point in this study. We thought this was the best approach to advance the science in this field. We have added this to the discussion, line 590-595.

Other detailed comments (some also not minor) are included in the following list.

L19 and in general: I find the term "direct" very misleading. This because, until specified, could be misinterpreted (direct risk? Direct assessment? Direct method?). Also, to my knowledge, "direct losses" is the correct term (rather than "direct risk"). *[6]* I think a simple title "Flood risk assessment of the European road network" would be much more effective. *[7]*

We thank the reviewer for pointing this out, we decided to include a few lines on direct/indirect tangible/intangible definitions. As the boundaries of these definitions are disputed, we decided to go with the categories proposed by Jonkman et al. (2008) (line 38-43). [6]

We shortened the title in line with the referee's suggestion [7].

L23: "more precise" – is it possible to quantify it, with numbers? *[8]*

*This was already elaborately discussed in the introduction. In the continental-scale grid-based approach, the exposure data are captured in a 100*100 m grid, with rather coarse land use categories. In the grid based approach, the exposure data the geometric descriptions often deviate* *less than 1 m from their actual positions. Moreover, the metadata distinguishes between road types, # lanes, the presence of bridges and even street lighting. In section 2.2 we further elaborate this point, with a reference to Fig S2 where the differences can be very clearly seen. We invite the referee to have a look into* http://openstreetmap.org

L26: move "the Alps" after, i.e. "such as roads in the Alps and along the Sava river" *[9]*

We thank the reviewer for this textual improvement.

L38: here "direct" make sense because it is associated to losses. *[6]*

L40 and below: also vulnerability of parked vehicles
https://www.sciencedirect.com/science/article/pii/S0048969718345388

We have included this, it also appeared in the framework of Jonkman et al. (2008) (line 38-44).

L47 and below: as said, this is not true. Most of urban studies use road links to assess impact of flooding to roads. *[10]*

We think the beginning of the paragraph ("Existing continental scale studies (…)") makes very clear that we are discussing continental scale studies here. Halfway the paragraph we also mention "(…) a typical grid size for ***continental-scale*** *modelling (…)" to remind the reader about this.*

Moreover, in lines 104-105 we mention that "actual modelling has so far focused on small spatial scales" with a reference study to acknowledge that this type of modelling is already done for small geographical scales.

Besides the small-scale studies we already referenced, we have added a short remark with references to Chang et al and Suarez et al on line 80, to take away any remaining confusion.

L57: I think an overview about the type of damage/losses is needed: direct/indirect, tangible/intangible. Indirect losses should be clearly defined as e.g. "physical damages" of road surface, etc *[6]*

See earlier comments, we have included this in the introduction [6].

L68: "complete" – what does it mean? *[11]*

It means that almost all roads that exist in the real world are also present in the dataset. We clarified this in line 76-78.

L73: "object-specific attributes" – what does it means? Just…attributes? *[12]*

We formulated it like this to emphasize that the attribute values may vary from object to object. In OSM, the number of lanes is specific for each road segment; rather than one general attribute that holds for all segments of a certain type.

L106: it's not "new" *[1]*

Agree, the novelty is the continental scale on which it is applied. We will remove the word 'new', see comment [1].

L109-110: authors need to explain before this list what CORINE and LUISA are *[13]*

This we now explain in line 57.

L112-113: is "object representation" different from "object-based"? *[14]*

In this context, we wanted to emphasize that Huizinga initially provided grid-based curves, and that we transferred them to object-based functions with some assumptions. This is explained in the text below the list. We rephrased as object 'translation'.

L133: "not include…pluvial, …coastal,, ..river, ..flash"- so what does it include? *[15]*

It includes what the sentence before indicated… Maybe we caused some confusion with the comma, we will rephrase as: "These maps represent the inundation depth and extent in all river sections with upstream area larger than 500 km2.They do not include the effect of pluvial flooding and coastal flooding. They also do not include river and flash flooding in the most upstream catchments, with area smaller than 500 km^2."

L223: "motorways" – in italics with no reason *[16]*

We have corrected this.

Table 1: worth adding the labels C1, C2, etc here *[16B]*

We have added these labels in a new column. Together with the new Table 2 and new Figure 4 (see referee report #3) it should be more clear now.

L315: "same damage curves" – specify if Huzinga's, if this is the case *[17]*

We replaced "same" with "Huizinga's"

L336: "strong increase" – of which one? Specify *[17B]*

We clarified this.

L339-340: this point seems quite important to me, because there is the risk to compare "apple and oranges". Is it possible to quantify the "small" contribution (L342)? *[18]*

Quantifying the share of railway damage in LUISA maps would require a significant amount of work (i.e. calculating the land area occupied by road and railway networks, and then evaluate how each network is represented in LUISA maps). We believe that such estimation would be beyond the scope of this study. Moreover, from a visual inspection of the maps we are quite sure that most of the railways does not appear on the maps at all; with the exception of some railway stations. The area of these is however small compared to the area of the road network.

We rephrased these lines to further justify our assumption (line 399-405).

Fig. 4: I would try to use the colours in a way that it is clear what is compared between (a) and (b/c) *[19]*

We have used the recommended color schemes of CORINE and LUISA for (a), which enables a quick comparison with these maps and other studies. For panel (b/c) we use a yellow-to-red color scheme which has a gradual brightness gradient which is also is clearly visible in greyscale print. We prefer to use these colors, which are intuitive for displaying road networks. In the text, we compare multiple things, so it will be hard to find one color setting to capture everything.

L376: from here, the text seems to belong to the following section

The entire section 3.2 discusses the region-aggregated damages in the object-based approach. In section 3.3 we introduce the results for individual road segments.

Although we can see that dividing the results in a differently could also work, we don't think that it will improve the readability of the manuscript; especially not because 3.2 is a common way of presenting results in continental scale studies, whereas 3.3 is specific to the object-based approach. Also, the length of the sections is more balanced in its current form.

Fig. 5: I can't see the areas mentioned in the main text, e.g. Po area

In figure 6, we will add labels to the separate panels; but for figure 5 we do not want to make this already complex figure more crowded. We think some basic knowledge of European topography can be assumed for the readers of a Copernicus journal (we already indicated to discuss rivers flowing from the Alps, and that the Po is in Northern Italy: country boundaries can be clearly distinguished in the current image).

Fig. 6: I would specific on the top of a/b/c the location (Nl/Alps/Balkans); also a/b/c have a different style compared to Fig. 5. The colours of the figure are not clear. Although I recognised it is difficult to represent road type and risk, this picture is not effective at all since the colours do not convey the perception of risk (trivially: red for high risk, green for low risk) and in general are difficult to be seen *[20]*

We reformatted the a/b/c labels in line with Fig 5, and added the requested captions.

We have experimented with over 10 different visualisations, aiming at:

- *Separation between land and sea*
- *Indicating included and excluded countries*
- *Indicating country borders*
- *Showing all non-flooded motorways and trunks (and the differences between them)*
- *Finding a gradient color scale with sufficient contrast with all of the above, notably the non-flooded motorways.*

We don't feel we can make it much better than this. Moreover, we used pink (which is not far from red), and we perceived blue as a natural color indicating a potentially flooded road. In the first review round, referee #1 indicated to "really appreciate" the image, so maybe it also is a matter of taste.

We have also included a large version of this image in the supplement (in pdf and shapefile) for further reference and exploration.

L431: specify the type of road, I assumed A3 is a highway but the specification is needed for including all the readers *[21]*

This was already elaborately described in section 2.4, and can also be seen in figure 7, but to avoid confusion we have added it in line 505 as well.

L467-471: "resulted in comparable estimates" – I expect more justification, because it seem not to me *[22]*

We had to rerun the model with a slightly altered version of the Huizinga damage function, so these values changed a bit. In the new description we have tried to be more precise in our formulation, see line 432-435.

L472-479: I don't understand this information here. If it is really informing your validation, I would create a section dedicated to validation where authors refer to the reference case(s) and this USA study (and others perhaps) *[23]*

The Vennapusa et al. (2013) report contains valuable information on this topic. However, for two reasons it could not be used as a sound validation/reference case:

- *Most importantly, although the report extensively describes repair activities after a river flood with information about the road repair costs, we could not accurately estimate the water depths from the provided information. Without information about water depths, we cannot precisely validate our water depth-damage curves.*
- *Less important, our study spatially focusses on Europe, and the typical scaling with GDP-per capita seems not to work in the USA: where the GDP per capita is relatively high, but the road design is relatively simple.*

Therefore, we decided to use this data to cross-check our assumptions on: which percentages of construction costs can be expected for different repair activities. As indicated in the discussion, we conclude from this data that our curves compare reasonably well to the data reported by Vennapusa et al.

L485, 486 and 490: references miss year

This is according to many reference style guides, where the year is only mentioned at the first instance of referring to the author(s)/study.

L490: “perceive” – not sure academic results are about perceptions

Both our results and the results of Bubeck are derived from modelling studies, where empirical data is combined with assumptions about reality to construct a full model chain. We think it is good to mention that we, on the basis of our experience (and on the basis of the reasons stated), perceive this value as high, without making a strong judgment about someone else work. Such expert judgment may also inform new studies in which a definitive answer to the question can be given.

L539-540: the reference case is local scale, isn’t it? So perhaps more cases and more road types are needed? Or validation that range from local (as the reference case) to urban to regional? *[24]*

Yes, we agree. The problem is (see also our earlier replies, also in the first review round) that such data is not available. We added an extra paragraph to the discussion section to emphasize this.

Anonymous referee #1 (Report #2)

Scientific significance: excellent

Scientific quality: excellent

Presentation quality: excellent

Accept as is.

We once again thank the reviewer for the careful consideration elaborate feedback in the first review round, and are happy to hear that the corrections and adjustments have been satisfactory.

Anonymous referee #4 (Report #3)

Summary: This paper presents many useful insights into more detailed modeling of large geographic scale transportation network (road) damages resulting from flooding in the current climate. Specific advancements include the creation of new depth-damage curves and more granular assessment of infrastructure locations and asset types using available OpenStreet Map data. Overall, the manuscript is well written, although it appears disorganized (perhaps reorganized) in the Methods and SI sections. It is clear the authors have done a lot of detailed work to create this assessment and there is very useful information presented for application in many instances. The manuscript would benefit significantly from clarifications, especially in the Methods sections, to help readers better understand exactly what assumptions and actions are taken at different steps. Some text should be moved to a 'Background' section to clarify what information informed, but did not directly contribute, to the results presented. Particularly, as noted by the authors, this kind of information (particularly the damage curves and costs) are often cited for years to come, so clarification on the specific applications and limitations of the methods will be most useful. Some specific areas for clarification are detailed below for reference.

We thank the reviewer for the in-depth review work and the nice comments on the study.

We have reordered and expanded the information in the method section, especially to make a better link between the raw data and the resulting damage curves. In the section on the new damage curve, made a subsection with background data. We decided to do this within the method-section rather than adding a complete new top-level section, because a new top-level section would have lead to additional overhead and make the article even longer.

Below will more elaborately expand on these points.

Comments:

Section 1: It would be helpful to clarify the specific meaning of "object-based" in this study – it can vary across disciplines and be helpful as a reference *[25a]*

We have added a definition on line 81-82.

Section 2.1: This study builds on previous work by some of the authors, referencing several different publications (notably Alfieri et al. 2014/2015). While full details can be referenced and found in that work, adding greater detail in this work is needed. *[25b]*

- Specifically, is this geospatial data? If not, how is it integrated into the study?

Yes, this is geospatial data. Figure 2 shows how the raster flood data is attributed to the road segments, with a description in section 2.3.1. We have elaborated our explanation both in section 2.1 and 2.3.1.

- The consideration that 'no flood protection is in place' (line 130) also could benefit from a short iteration on how this likely affects the results (it seems like potentially overestimating the damages/flooding to the most important regions of infrastructure, as those would be protected as part of the design).

We apologize for the confusion raised. We do account for flood protection, but not yet in this step. The hazard maps assume no flood protection, but the flood protection is accounted for in the risk calculation.

When discussing the flood hazard maps, we have added: "(flood protection is considered in the risk assessment step, Fig S1.)", line 143. This Figure S1 shows how we do it, this approach is common practice in continental-scale flood risk assessment studies.

We also incorporated some additional flood proofing in the shape of the damage curve for motorways, which assumes a 1-m high embankment line 288-295. The fact that in addition to this, some flood-prone areas may have additional protection was already addressed in the discussion, line 570-577.

• Additionally, from the end of line 131 to the end, it seems that this is a description of the Dottori et al (2020) dataset? Being very clear here that no new work was done to augment/change these data by the authors would be helpful

The referee is right. We specified this in the revised section 2.1.

Figure 2 mentions that the road network is "simplified" but this isn't explained in the text, unless this is referring to either the classification of the data (Table S3) or the 1498 regions it is broken into (line 179). If the latter, this seems more like a geographic binning than simplification. This is an important detail be clear about exactly what inputs were analyzed. *[26]*

We thank the reviewer for pointing this out. The simplification does not refer to the regional binning, but to a geometric operation on the vector files in OSM. In some parts of OSM, roads are mapped with a degree of detail far beyond what is useful in our study (e.g. mapping a 20-m diameter roundabout with ~500 datapoints to approach a circle). Such geometries were simplified to a resolution of $5*10^{-5}$ degree in WGS84, which is roughly 5 m in Europe. This means that 'redundant' datapoints in the linestring of a geometry (i.e. within 5 m distance of an existing datapoint) are removed to simplify the geometry. This increases the computational speed of the model.

We included this in the text (line 193).

Section 2.3.3, specifically regarding Table 1, Table S8 and S11 and the writeup for costs and

damage curves used:

• Creating new (ground-truthed) damage curves and costs is a major contribution to the literature. But, it is really hard to distill what was *actually* applied in this study from the large amount of data collected. For example, Table 1 states that "Motorway, Max Damage (high flow)" has a relative cost for A/B type road of 22%. In Figure S3, this seems to reach 90%. It may be that a cutoff value is considered (100 cm), but this also does not appear to correspond to FigureS3. It's also stated earlier that expert judgment was used. This is a valid method, but needs to be clearly explained if that is what was done, and some explanation of the discrepancies between the tables and figures should be noted. This is a key contribution of the paper, and merits more room in the manuscript to be clear for readers *[27]*

*The discrepancy between the 22% and the 90% (eventually it goes up till 100%) is a matter of definition of the damage curves. Table 1 expresses the maximum damage as percentage of the construction costs (which is 22%). In Figure S3, the damage is expressed as percentage of the maximum damage. So for example, if the 90% percentage of max damage is reached, this would correspond to 90% * 22% = 19.8% of the road construction costs.*

Also in anticipation of the referee comments below, we created a new section 2.4 'Damage curves', which first explain the follows the process that was followed to construct the curves (which includes expert judgment). It than presents all the background data, while referring to the SI (2.4.1), followed by an explanation of the different categories of damage curves (2.4.2). In the last subsection (2.4.3) we than introduce the new damage curves, relating them back to the background data.

• It appears (though is not stated in the manuscript) that Table S11 is the data actually applied to obtain the study results. Table S8, while a great collection of background data, is too disparate to provide any insight into how these results were obtained. *[28]*

We included a synthesis of the results of Table S8 in line 263-268. Indeed, there are differences between the raw data, but on the other hand there is enough agreement to gain a fair impression of the order-of-magnitude of the costs of different repair categories (such a presented in the aforementioned synthesis). These also have reasonable agreement with the data presented by Vennapussa et al. for real flood events, as discussed in line 541-544.

Moreover, we need to clarify something here. Table S11, S12 and S13 are all required to obtain and reproduce the results, not just Table S13. Also, Table S11 is not based on the data in Table S8, but on the data in Table S6. We have moved Table S11 and S13 to the main manuscript to highlight their importance.

We wanted to present all background data to enable other researchers to make their own expert judgments on the curves.

• Lines 216-231 are useful background information, but seem to be more background than Methodology. Moving this to earlier in the manuscript might improve the clarity of what costs and assumptions are actually used in this study *[28]*

We have moved this to a new background section 2.4.1.

• Line 243 states that "Therefore, we assume that the predicted floods have relatively low flow velocities"… but high flow results are later considered. This merits clarification *[29]*

Indeed, the predicted river floods in the hazard dataset have relatively slow flow velocities compared to flashfloods in very steep upstream catchment. However, for these slow velocities, we can still distinguish a range of slow velocities, which we denoted with slow (that is the lower boundary of the slow range) and high (that is the upper boundary of the slow range). This we explained more clearly in line 280-281.

• Lines 247-260 are good explanation of a key contribution to this study. It might (if space permits) even be worth putting Figures S3 and S4 into the main body of the manuscript, after Figure 3. *[30A]* However, adding another paragraph to identify how the curves created (C1-C6) can be attributed to the costs presented in Tables S8, S11 and resulting in Table S13 would be really helpful to readers *[30B]*

With agreement of the editor, we have added a new figure 4 to the manuscript, in which the damage curves are visualized.

We hope that the newly structure method section (in which we separate the methodological procedure (2.4.1) from the background data (2.4.2) from the final curves (2.4.3) will help the reader to follow along with our reasoning.

• Lines 263-272 are really good background information. Moving this information to the background improve clarity of the Methods applied in the study. *[31]*

We have moved this section to the new background subsection 2.4.1

• Table S13 seems very useful to the study results. Considering a summarized version of this applicable to the Methods used to obtain the results in the body of the manuscript would be useful *[32]*

We thank the reviewer for this suggestion, after consideration with the editor we have included this in the main body of the manuscript, see table 2.

Section 2.3.5: Clarification on the type of sampling used is helpful (Was random sampling used for each road segment/flood? Or was a Monte Carlo simulation used?). Additionally, a small example here could be very useful *[33]*

We have clarified our description of the sampling procedure in section 2.5.

Line 298-299: Add a citation for the assumption of a normal distribution for the flow, or state that this assumption was made because of a lack of available reference data *[33]*

We added this to the description of the sampling procedure, line 354.

Section 2.4/Table S17 – this is a clever way to get good data on a test section. What types of roads were 'assumed' via the OSM data for this region? Does it appear accurate? Adding a short but pointed validation of the input data agreement would improve the strength of this section as a verification *[34]*

This is a good point: the OSM-representation is very accurate (which for us seemed needless to say but of course not all the readers know this), in terms of geospatial references but also in lane (and bridge) attributes. We added a short comment on line 356; it also becomes clear from Figure 8.

Section 3:

• Lines 390-396 and Figure 5: The explanation here doesn't seem to exactly correlate to the Figures. Figure 5B is the "share of regional GDP" whereas the explanatory text talks about National GDP per capita. It also seems that the explanation might be backwards – it appears that the GDP-scaled data (5B) reveals countries such as Croatia as being highrisk, where as the text states "without this GDP correction…" (line 391) *[35]*

In the text we confused panel a and b, we have corrected this.

We tried to make our description more clear. The scaling of the exposed assets to national GDP per capita is done in the entire study (see e.g. line 203-205,247-251). What is special about panel b, is that we related the damages to the total GDP (not per capita) per NUTS-3 region (so that is indeed different from the national GDP per capita). We have made this explicit in the description (line 551), so it should now be coherent with the figure.

Supplementary Information: the Supplementary Information seems out of order with what is presented in the manuscript, and some seemingly important tables are not mentioned in the main manuscript at all. I think it would benefit from a thorough organizational read. *[36]*

We slightly reordered, and added headings and a table of contents to the SI, this should make it more accessible. However, we think that the SI follows the order in which information is presented in the manuscript. We have cross-checked that all SI tables are mentioned in the manuscript and added references in the main text where missing in the initial manuscript.