Referee comment (Referee #1)	Author's response
Page 5: The model represented by ArcGIS modelbuilder in Figure 1 is necessary. Sure, the users can run the tool of Li (2018) in this way in modelbuilder, but the simplest way to run the analysis is by double clicking the tool and specify the inputs in the GUI of this tool. In this figure, I suggest the author to show a screen shot of the GUI (panel A on the left) and then the screen shot of the attribute table of the point file (panel B on the right) to show how the optional parameters, such as dip, slope, and height, are organized. In addition, this tool accepts both shapefile and feature class as the input of the sample sites. The author may just use a GIS point file of the sampling sites (vector) here.	The figure was modified accordingly (Fig. A1). Panel (a) shows a screenshot of the graphical user interface. Panel (b) shows the attribute table of a point file for shielding factor calculations.
Page 17: Figure 6: May be better to also show the correlations between the shielding factors derived from 1 m DEM and resampled 12m and 30 m DEMs to see if there are major changes on the shielding factors because of the changes in DEM resolutions.	The figure (Fig. 4) was revised accordingly. Correlations between the shielding factors derived from the 1 m-DEM and those derived from the resampled versions are shown in panel (c) and (d)
Page 18: Figure 7: Why the number reduced to 23 in this figure? The number is 37 in the previous figures. If the reason is because only having 23 CRE ages, the author can just report the panel (c) because the previous figures already show strong correlations between GIS and field-derived values.	CRE ages are only available for 23 moraine boulders. As suggested by the reviewer, panel (a) and (b) were removed. Histograms of CRE differences are shown in one figure (Fig. 8). For clarity, the following phrases were added in the methods section: "Note that CRE ages were not available for 14 boulders in the southern Black Forest that have been selected for this study. Therefore, only 23 CRE ages were recomputed"
Page 20: Figure 9 and the texts: It will be great to discuss the correlations after removing the "outlier" (STEI-7)? I guess the field measurement for this sample site may be problematic.	R^2 rose to 0.91after removing the outlier. This information was added in Sect. 4.1.3: "After the exclusion of the potentially problematic shielding factors for the STEI-7 boulder, R^2 rose to 0.91 in both cases ($p < 0.05$)" In the discussion (Sect. 5.1), the author states that "After the exclusion of the potentially problematic shielding factors for the STEI-7 boulder, SRTM DSM-based and TanDEM-X DSM-based shielding factors were equally consistent with field data-based shielding factors. The similar fit suggests that TanDEM-X data do not have an advantage over SRTM data"
Pages 21-25: Section 4.2: Sensitivity tests. The three sensitivity experiments reported in this section are the results from the three study sites. They are apparently different from	The sensitivity tests do not differ from the sensitivity tests described in the methods section. The term 'sensitivity experiment' was removed from the whole manuscript for

the three sensitivity tests described in the method section (Page 13). Based on the descriptions in the method section, the first test is to assess the effect of the different methods for calculating topographic shielding factors on CRE ages of boulders in mountains with an intermediate elevation that have been exposed to cosmic radiation since the Late Pleistocene. The second test is to determine whether the choice of topographic shielding factors has a significant impact on the CRE ages of surfaces that have been exposed for the last few millennia. The third test is to assess the impact of topographic shielding factors on the young CRE ages of LIA or younger. Although the three tests are related to the three sites, respectively. They have different focuses. I hope the author can check the consistency of the three sensitivity tests described in the method and results sections.	clarification. Sect. 4.2 was shortened to make the manuscript more concise.
Tables 2, 3, and 4 are likely not necessary for the main text. Maybe can put these tables as the supplementary. The author can create the histograms of the CRE age difference for the three sites and put them in one figure. In this way, the main text can be shortened.	The Tables 2, 3 & 4 were moved to the appendices (Tables B1, B2 & B3). Histograms of CRE differences are shown in one figure (Fig. 8).
Pages 26-29: Section 5.1: The discussion about the vegetation-corrected or not corrected should belong to Section 5.3. There are some repeated parts in these sections. The author needs to re-organize the writing to avoid the repeated sentences and paragraphs.	The discussion section was restructured and shortened to make the text more concise. The discussion was subdivided in the following sub-sections: 5.1 Impact of the spatial resolution and quality of elevation data on shielding factors, 5.2 The role of vegetation, 5.3 Correcting for the boulder height – does it matter? 5.4 Impact on CRE ages, 5.5 Practical guidelines
It is interesting that the vegetation seems have different impacts on SRTM DEM and TanDEM-X. This can be a good point. SRTM data is collected in February 2000 (leaf off season), so that the impact of vegetation cover on topography may be not very high. When the TanDEM-X data were collected? if it was during the grown season, it may have a bigger impact. I suggest the author to check the data sources of different DEM sources and explain the vegetation impact on topography more.	Thank you for the hint that the time of the acquisition of elevation data might have an impact on the results. TanDEM-X data were obtained in the 2010-2015 period and multiple acquisitions were averaged. This could partly explain the relatively poor performance of TanDEM-X data in forested areas. The following phrases were added: “It should be mentioned that SRTM data were acquired in February 2000, i.e. during the leaf-off period in the northern hemisphere, whereas TanDEM-X data were obtained by averaging data from multiple acquisitions. Data collection during the leaf-off period could be one explanation for the better performance of SRTM data“

In the discussion and conclusions, the author argues that a relatively low-resolution DEM is better for determining the topographic shielding factors. I suggest the author using medium resolution instead. A low-resolution DEM, such as SRTM 90-m DEM, is not very accurate for the topographic shielding factors as illustrated in Figure 2 (Page 7).	The author considered a DEM with a spatial resolution of 30 m a low-resolution DEM. This is indeed confusing, as SRTM with a *xy*-resolution of 90 m is also available. Therefore, the author refers to actual *xy*-resolutions of DEMs instead of terms, such as "intermediate spatial resolution"
The manuscript is relatively long and there are some repeated sentences and paragraphs or meanings in different sections. I suggest the authors to re-organize the writing to make the manuscript more concise.	The manuscript was shortened to make the text more concise. Repetitions in the discussion section were removed and unnecessary phrases were deleted.
p. 5, line 98: Can be a feature class. Maybe just say a GIS point file of the sampling sites (vector)	The manuscript was revised accordingly ("point file of the sampling sites (vector)")
p. 5, line 100: the point file	"shapefile" was replaced by "point file"
p. 6, line 118, measurements	The sentences were rephrased as follows: "In practice, the number of measurements is usually much lower than in the GIS-based approach. If elevation data is correct, GIS-based shielding factors should theoretically be more accurate than field data-based shielding factors"

Referee comment (Referee #2)	Author's response
As mentioned above, this work in its current form is very long and overly elaborate. This work tests an existing published tool with data from three locations using 14 figures and four data tables, all of which are in the main text. In several cases, these figures do not significantly add new information and can probably be summarized into 3-4 figures. This work reads more like a literature review, with paragraphs and tables summarizing previous works without discussing new information (see examples in the detailed comments). As most readers of this type of paper are well informed in the methodology and background, I would suggest writing it as a short technical note focused on the findings.	Three of four data tables were moved from the main text to the appendices. The number of figures in the main text was reduced to 11. Why should the manuscript not be written as a review? The study aimed at comparing the output of Li's toolbox with field data-based shielding factors. Hence, it is crucial to review both approaches and elucidate how they differ. In Sect. 2.1, the author points out the differences between the methods: "Li's toolbox computes topographic shielding factors for each sampling surface with 360 pairs of azimuth and elevation angles. In the field data-based approach, the horizon is approximated by points that are linked by straight lines (Balco, 2018). The azimuth and the corresponding elevation angle is recorded for each of these points with an inclinometer. In practice, the number of measurements is usually much lower than in the GIS-based approach. If elevation data is correct, GIS-based shielding factors should theoretically be more accurate than field data-based shielding factors". To make the main text more concise, Sect. 2.1 ("Principles of the ArcGIS toolbox and validation") was moved from the main text to the appendices (Appendix A). Unnecessary information in Sect. 2.2 (e.g. why did researches use the toolbox) was removed make the text more concise.
Additionally, this work while very elaborate is a bit lacking in data. Given that the author examines an existing previously published ArcGIS tool, it should be relatively easy to also examine other shielding factors from published works. Specifically, I think this work would benefit from shielding factors from data with both high and low 10Be concentrations and production rates. I would be interested in seeing whether SRTM vs field-measured shielding calculations make more or less of a difference for different settings. As 10Be data is easily available (see ICE-D and OCTOPUS databases), this would make the results more robust and would this work more interesting.	Thank you for this comment. In the introduction and Sect. 3.3, the author emphasises that he recalculated CRE ages for the Écrins massif (Le Roy et al., 2017) and for the forefield of Steingletscher (Schimmelpfennig et al., 2014) for exactly this reason: "To assess the effect of the choice of shielding factors on CRE ages, previously published CRE ages of moraine boulders with varying ^{10}Be concentrations at three sites were recalculated. The selected sites differed in terms of ^{10}Be production rates" (introduction) "To test whether the choice of the topographic shielding factor influences CRE ages of surfaces that have been exposed for the last few millennia, CRE ages of boulders in the forefield

	of four glaciers in the Écrins massif (*n* = 24) were recomputed. Although ^{10}Be production rates at sampling sites are much higher than in the southern Black Forest due to higher elevation, the in-situ accumulated ^{10}Be concentrations in the sampled boulders are lower due to the relatively short duration of exposure. ^{10}Be concentrations in samples from these boulders range from 2,800 to 21,800 atoms $^{10}Be\ g^{-1}$ quartz and are thus much lower than in the boulders in the southern Black Forest". In addition, he recalculated CRE ages for the forefield of Steingletscher: "As CRE dating has also been applied to terminal moraines of only a few centuries in age (e.g. Schaefer et al., 2009 or Braumann et al., 2020), CRE ages (*n* = 16) of boulders on Little Ice Age (LIA) and post-LIA terminal moraines of Steingletscher (Fig. 1), were also recalculated. See Schimmelpfennig et al. (2014) for a description of the site and the interpretation of the ages. The ^{10}Be concentration in samples from the boulders vary between 2,230 and 12,220 atoms $^{10}Be\ g^{-1}$ quartz due to short durations of exposure" The author thought of including ^{10}Be data from the ICE-D database. However, as pointed out by Li (2018), measuring azimuth and elevation angles in the field is a subjective process. The author learned obtaining field data from I. Schimmelpfennig and conducted fieldwork with M. Le Roy. Therefore, the author only used shielding factors obtained by I. Schimmelpfennig, M. Le Roy and himself to ensure consistency of the field data-based shielding factors.
The data for this manuscript is only available upon request from the author (see line 524). This does not go in line with the data policy of Copernicus and can be uploaded to any one of several online repositories.	Thank you for this hint. Actually, this statement is not needed, as all data are included in the appendices and in the supplement.
Abstract: The abstract is missing a short summary of the results. It would have been helpful if the abstract had a short summary of the CRE age differences.	Results of the recalculation of CRE ages are emphasized in the abstract: "In most cases, recalculating CRE ages of the same sampling sites with different shielding factors led to age shifts between 0 and 2%. Only one age changed by 5%"
Lines 31-33: This sentence can probably be removed altogether. People reading this will already be familiar with this.	The sentence was removed.

Lines 85-86: The phrasing is confusing. The sentence would probably be clearer if it skipped the first part and started with “three sets of previously…	The sentence was rephrased as follows: “To assess the effect of the choice of shielding factors on CRE ages, previously published CRE ages of moraine boulders with varying ^{10}Be concentrations at three sites were recalculated. The selected sites differed in terms of ^{10}Be production rates”
Lines 90-95: questions 2 and 3 are similar and should probably be presented as a single question	The second question was reformulated as follows: “2. Do the *xy*-resolution and the type of the elevation data [DEM or digital surface model (DSM)] significantly influence the quality of the shielding factors?”
Line 114: How far is farthest? Is there a maximum limit?	There is no maximum limit (default setting). This information was included as follows: “The skyline function generates a skyline that represents the farthest visible points along the line of sight around a locality (default setting: no maximum distance)” (see Appendix A)
Line 288: Can you explain why this discrepancy exists?	Reasons for this discrepancy are discussed in detail in Sect. 5.2: “The TanDEM-X DSM-based shielding factors for the FS-1a, FS-2a, SW-2, SW-9 and WH-1a boulders did not agree with field-data based shielding factors (Fig. 3d). Except of the SW-9 boulder, these boulders were situated in areas covered by mixed and coniferous forests. As TanDEM-X data are not corrected for vegetation, differing canopy heights and small-sized anomalies in vegetation cover are prone to be misinterpreted as topographical obstructions by the toolbox. The SW-9 boulder, for example, was situated in open grassland close to a coniferous forest. Measuring pairs of azimuth and elevation angles in the field turned out to be straightforward, as the horizon around the boulder was even visible through the coniferous forest. Inspecting the skyline for the SW-9 boulder in ArcMap revealed that the edge of the coniferous forest was misinterpreted as a topographic barrier by the toolbox. Excluding the problematic shielding factors for the FS-1a, FS-2a, SW-2, SW-9 and WH-1a boulders led to a strong correlation with field-data based shielding factors ($R^2 = 0.88$; Fig. C1). It should be mentioned that SRTM data were acquired in February 2000, i.e. during the leaf-off period in the northern hemisphere, whereas TanDEM-X data were obtained by averaging data from multiple acquisitions. Data collection during the leaf-off period could be one explanation for the better performance of SRTM data”

Lines 434-449: this whole section should also be in a supplemental section.	This is an important paragraph of the discussion. TanDEM-X data-based shielding factors were seemingly most consistent with field data-based shielding factors for boulders in the forefield of Steingletscher. The author shows that TanDEM-X data-based shielding factors should not be considered realistic, although they fit well with field data-based shielding factors.
Lines 465-472: This section is not necessary. It is a summary of a previously published paper. The relevant results of this work should be briefly referred to with a reference	The discussion of the role of small topographic obstructions is highly relevant. One could argue that high-resolution should be used to account for small-sized topographic obstructions. The reasoning in the paragraph indicates that small-sized objects do not induce significant topographic shielding, as most of the cosmic rays are not stopped when they penetrate through topographic obstructions. This is another argument for the use of elevation data with an intermediate spatial resolution.
Figure 1: This is a published tool. This figure is not needed in this type of manuscript.	Figure 1 was moved to Appendix A.
Figure 2 is a little confusing and does not really add new information. It is probably better to move it to a supplemental section.	Fig. 1 was moved from the main text to Appendix A. Li only showed the plot in panel (c) in his 2018 paper. The remaining shielding factors are provided in tables (see Fig. 6 and Table 1 in his 2018 paper). As Li did not plot the data, it is hard to comprehend his validation study. In addition, he did not assess the correlation between GIS- and field data-based shielding factors.
Figures 8-9: These are excellent figures that make the actual differences and their consequences easy to understand. It would be helpful to add the cumulative frequency on the y-axis.	Histograms of differences in CRE ages for the southern Black Forest, the Écrins massif and for the forefield of Steingletscher are shown in Fig. 8. As suggested, cumulative frequencies were included.
Figure 11: This too should also go to a supplemental section. There should be one figure like this for all data and the rest should go in a supplemental section.	Figure 11 was moved from the main text to the appendices (Figure C1).
Table 1: This table can be moved to a supplemental section and is not needed in the main text.	Table 1 contains crucial information. Table 1 shows that elevation data with a *xy*-resolution between 0.5 and 30 m were used in previous studies. The output of the toolbox of Li (2018) has not been systematically tested with elevation data with a *xy*-resolution of ≤8 m. This was one key motivation for the study: "From 0.5 to 30 m, the *xy*-resolutions of the input-DEMs were very heterogenous (Table 1). Several authors computed shielding factors with high-resolution

	DEMs (≤5 m). As Li (2018) did not test toolbox with a DEM with a *xy*-resolution of less than 8 m, a systematic assessment of the impact of the spatial resolution on the quality of shielding factors is crucially needed”