



Comment on **essd-2021-414**

Anonymous Referee #1

Referee comment on "A national landslide inventory for Denmark" by Gregor Luetzenburg et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-414-RC1>, 2022

The article presents relevant new data in the form of a national landslide inventory for Denmark. The new data collected and presented in the article fills a gap in the European and international panorama of landslide mapping efforts and related resources. I commend the authors for their efforts, and for making the collected data publicly available.

Overall, the landslide data collected and presented in the article are of potential scientific interest to a broad community, and are of possible practical use. The landslide data appear to be collected, treated, and organised in a manner scientifically and technically consistent with common practices to produce national-scale landslide inventories. Therefore, it is my opinion that the work deserves publication, pending some improvement concerning additional information on how the inventory was produced, the presentation of the landslide data, the text of the article that needs adjustments, and the references.

I have a few general remarks and some specific comments. I list them below, in the hope that they will help the authors improve their already interesting article.

General remarks

Denmark is a European Country with vast oversee territories, including the Faroe Islands and Greenland, in the Atlantic Ocean. For readers – like me – non entirely familiar with the political and physical geography of Denmark, the authors should explain clearly what part the Country is covered by the new inventory.

Inspection of Figure 2 reveals that the authors have used a single polygon to encompass all parts of a single landslide, including the source area and the landslide deposit. The authors should explain why they have adopted this strategy to map the landslides, including e.g., the time available to complete the mapping, the insufficient resolution and vertical accuracy of the DEM, the difficulty in separating the source area from the main deposit, systematically. To some extent, the choice made not to separate different landslide parts may limit the possibility of using the inventory to calculate the volume of the individual landslides and to infer erosion rates.

A related item has to do with the mapping of landslides partially or totally inside larger landslides. The authors should explain how they have addressed the problem. I have downloaded and inspected the shape file for the inventory, and found that overlapping landslides are treated as separate and independent polygons. This complicates the analysis of the inventory. As an example, summing the areas of the individual polygons will overestimate the total area affected by landslides in the study area – as it will count the overlapping areas twice, at least.

The inspection of the inventory in a GIS revealed long girdles of landslide polygons, each representing individual landslides, mainly along the coasts. This is a common geomorphological setting where landslides erode and modify a “*mesa*” like, or “*table top*” morphology, the result of a nearly horizontal layering of rocks of different mechanical characteristics. A problem when mapping landslides in this setting is the lateral separation of the individual landslide blocks, which may not be trivial and it may be subject to human interpretation. I recommend that the authors address the issue, and show at least one example of their landslide girdles and the accuracy of their mapping in these areas.

In sub-section 3.2, the authors provide information on how the mapping was performed by two experts, with a third expert performing an unsystematic verification of parts of the inventory. However, how this crucial part of the preparation of the national landslide inventory was performed is not sufficiently clear. The authors should show (e.g., in Figure 1) which tiles were mapped first by one expert, and which by the other expert (KSV, GL). This may outline a source of potential geographical biases that may be present in the national inventory. Second, the authors should explain what kind of verification was performed by the second expert. Was it a completely independent survey, or the second expert had access to the map of the first expert? Where disagreements emerged, how were they resolved? Similar questions arise for the validation performed by the third expert. Were the independent mappings of the two (first and second) experts available to the third expert, or only the joint (verified) result of both experts, or none of them? Again, where a disagreement emerged, how was it resolved? Was the final mapping changed based on the opinion of the third expert, or based on some form of mutual agreement? In the latter case, how this was accomplished? In general, what was the subject of the second and the third mappings. Did the second and the third experts check only the existing mapping and refined it e.g., changing the geometry of the polygons representing the landslides? Did they change the classification of the landslides? Did they added or deleted landslide polygons identified by the first (or the first and the second) expert? These are important issues that influence the quality, and hence the usability of the landslide dataset. I recommend that the authors address these issues, albeit briefly.

The authors should provide information on the time and human effort required to prepare the national inventory, including an account or an estimate of the overall time and of the persons / month or persons / year effort required to prepare the inventory, and for the single steps of the work *e.g.*, gathering and organizing the DEM and preparing the hillshades, gathering and organizing the orthophotographs, the visual inspection and digital mapping of the landslides, the validation of the mapping, the open publication of the results.

As a last general remark, I recommend that the authors check the citations they have selected. I am not convinced that all of them are fully appropriate, at least in the locations they are used in the text. Also, I encourage the authors to look for additional, mostly recent references on landslide detection and mapping methods and tools.

Specific comments

Title

I suggest the author consider the slightly different title "A national landslide inventory for Denmark", or "National landslide inventory for Denmark".

Abstract

The abstract should focus strictly on the main topic of the article *i.e.*, the new national landslide inventory for Denmark. The abstract should address the content of the inventory, focusing on the method used to compile, validate, organize, and publish openly the new landslide information. Any other consideration on the global relevance of the landslide problem, the risk it poses, and the potential (future, and therefore currently not proven) use of the landslide information for different scopes, should be removed from the abstract.

Introduction

In the journal aims & scope (https://www.earth-system-science-data.net/about/aims_and_scope.html) one reads that "*Articles in the data section may pertain to the planning, instrumentation, and execution of experiments or collection of data. Any interpretation of data is outside the scope of regular articles*". It follows that

most of the text in the Introduction is out of scope for the journal. I understand the need, and I appreciate the attempt the authors have made to frame their work in a broader perspective, but the Introduction is too long and not focused on what should be the main scope of the paper: presenting a new, valuable, national landslide inventory for Denmark. I recommend that the authors reconsider the text in the Introduction, reducing it considerably, and focusing it on the main scope of the paper

Study area

Following up on the same argument made before for the Introduction, the description of the Danish landscape and the recent geological history of Denmark is probably out of scope for an article in this journal, unless the information was instrumental to the compilation of the landslide inventory. I encourage the authors to consider the point, and change the text accordingly.

In Figure 1, the authors use colours to show surface and submarine terrain elevation, but they do not provide a legend for the colours used. The dashed grey line shows the maximum advance of the ice sheet during the Weichsel glaciation. However, it took me a while to understand what side of the line was covered, and what side was not covered by the ice. For the readers who are unfamiliar with the geography of the region, it would be good to show the boundary using a line with a different, asymmetric symbol. The acronym LGM is not explained in the Figure caption. The locations of Fig. 2a, 2b, 2c, are not easy to spot, at first sight. The authors should consider using a larger font, bold characters, or a different text colour. Figure 2b is similar (albeit not identical) to Fig. 2a in Svennevig et al., 2020, GEUS Bulletin 44, 5302. This should be clarified in the Figure, by writing *e.g.*, "modified from", or the like.

Methodology

In line 127, the authors write "*visual validation of landslide features in the landscape*". What does it mean, precisely?

Sub-section 3.2, Landslide mapping, is too concise. I understand the authors point the reader to the – freely available – work by Svennevig et al. (<https://doi.org/10.34194/geusb.v44.5302>); but some description on the method and tools used to collect the landslide data is important to assess the quality of the data, and to decide on the use of the data. I recommend that the authors expand this sub-section.

Sub-section 3.3, Quality control is important and interesting, but it also too concise. See my general remark on this topic.

The landslide inventory

The separation of landslides into "coastal" (or "coast") and "inland" landslides is not fully clear. Do coastal landslides have (currently) their toe in the sea? Or are coastal landslides slope failures that affect slopes that have (currently) their toe in the sea? Are "inland" landslides at a minimum distance to the sea? Are landslides on the slope of a lagoon or lake (if any) classified as "coastal" landslides? In general, I recommend that the authors explain why they have made the separation between "coastal" and "inland" landslides, and that they provide a clear definition for a "coastal" landslide and for an "inland" landslide.

In line 155, the authors specify that the area (m^2) and perimeter length (m) are given in the inventory. They should specify that these are planar figures *i.e.*, they represent the area and perimeter of the landslide as shown in the map, and not their true area and perimeter length in the field. The latter are larger and longer, given the fact that the landslides form and develop in a sloping terrain. It would be good if the authors could calculate and provide this additional information with their inventory *i.e.*, with the data and metadata in the shape file. This would allow users to analyse the inventory without having to download the DEM from which the inventory was obtained.

In line 156, the authors give the area of the largest mapped landslide as 327,001 m^2 . Clearly the one m^2 is somewhat "fictitious", in the sense that a very small change in the mapping may have resulted in a different total area for the landslide. Given the fact that the authors have not mapped landslides with a (planimetric) area smaller than 25 m^2 , I recommend that the authors present their data to the nearest 25 m^2 .

In line 171, the authors write "*In most cases, the mapped landslides record single events with process durations that span from an instantaneous event to several decades or even centuries and thus some are still active while others are inactive landforms today*". The sentence appears contradictory, and needs some clarification. I recommend that the authors provide a clear definition of what they consider an "event", or a "landslide event". A "process duration" of several decades or centuries implies that the same landslide has been active for several decades or centuries; or not?

In lines 199 and 200, the authors write "*Based on the careful observation of the entire study area and the implemented quality control, the landslide inventory can be considered 87% complete with a confidence level of 90% and an error of 5% for the 2015 DEM.*" These are clear and important figures given. I am sorry, but from the previous discussion and the presentation of the inventory, I do not understand how the figures were calculated, or estimated. What does it mean that the inventory is 87% complete? That it misses 13% of the total number (or area) of the landslides? And what landslides? All the landslides that are (in principle) visible in the DEM and the orthophoto maps used for the mapping, and that for whatever reason were not detected and mapped? Or all the landslides that have occurred in the study area since the last glaciation? The difference may be significant, and can possibly be sized from frequency-density plots of the landslide areas (see *e.g.*, Malamud et al., 2014). How was the 90% confidence level calculated, and what does it mean, precisely. Ultimately, how the 5% error for the 2015 DEM was

calculated, and how this apparently small error has affected the visual detection of the landslides from the hillshades?

Significance of the dataset

In this section, I am not convinced the authors do justice to their important work.

The authors provide two main motivations for the work. The first is a step towards "a *more comprehensive hazard and risk framework for Denmark*". This may be the case, but before embarking in a comprehensive hazard and risk framework for Denmark, it is plausible that the data can be used by landslide scientists and practitioners to construct landslide susceptibility and hazard models for Denmark. Discussing the same motivation, the authors suggest that a use of the new dataset is to "*develop effective risk reduction strategies to protect human lives and property*". Again, this may be the case, but it is not clear the extent to which landslides in Denmark threaten human lives and property. Since landslide activity depends on climate, it seems to me that a potential use of the dataset will be investigating and monitoring the effects of the changing climate along the high coasts of Denmark, confronting them with the similar effects on the "inland" landslides.

The second motivation is to provide landslide information "*to the machine and deep learning research community*." Although I recognize the scope and potential of AI-based methods in several fields of science, including landslide modelling and landslide hazard and risk assessment, I would not limit the use of this new national landslide inventory to the machine and deep learning research community. Several other promising research can be attempted by exploiting this dataset that do not require AI, including machine and deep learning, for modelling and predictions.