

Interactive comment on “A satellite-derived database for stand-replacing windthrows in boreal forests of the European Russia in 1986–2017” by Andrey N. Shikhov et al.

Anonymous Referee #1

Received and published: 22 June 2020

The manuscript "A satellite-derived database for stand-replacing windthrows in boreal forests of the European Russia in 1986–2017" present a GIS-database on storm events in European Russia. The data base spans more than 30 years and contains over 100,000 entries, an enormous amount of data! As such, I believe that the manuscript and database are an important addition to the literature, as information on windthrows is rare. I applaud the authors for undertaking such a great effort in compiling the data. That said, I have some issues and comments, which in my opinion need to be addressed before publication.

(1) The assessment includes a lot of subjective calls from the interpreter and there is

C1

no formal validation. I fully understand that validating the database is very challenging, but I would have at least expected some assessment on how accurate the data is. There are many steps involved in collecting the data, and some of them seem to be highly dependent on local knowledge. From this I assume that the accuracy will be higher in some regions (where there is ample local knowledge), but lower in others. In particular, I was wondering whether you searched the forest area of ER systematically or whether you applied any other “sampling strategy” to ensure you don’t miss a storm event. Moreover, the accuracy will depend on the availability of HRI. Was HRI available for each year after 2001, or only for selected years? The authors address some of these issues in the discussion, which I appreciated, but maybe they should make clear from the beginning on that the database is a subjective collection and probably far from complete and/or consistent.

(2) In the same line as the previous comment, many of the decision to group patches to windthrows, etc. are based on arbitrary thresholds. I wonder whether the authors have tested the sensitivity of their results to those thresholds. The allocation of patches to windthrow events might look very different with slight changes in threshold values.

(3) You do not explain how the Landsat data was processed. Please give some details on the processing. Moreover, how many Landsat images were available per year, on average? I guess that in Russia many winter images are covered by snow. As you use pre/post windthrow Landsat images to ensure windthrow detection, data availability is crucial. In years with only few observations, detection accuracy might be lower.

(4) The manuscript is already well written, but needs some language editing. I give some suggestions below. Also, it is quite dry to read in some parts, but I guess this is typical for a data paper.

Specific comments:

Throughout the manuscript: No “the” needed before “European Russia”.

C2

- L. 17: Sentence starting with “Additional...”: Something wrong with the sentence, please revise (e.g., “Additional information, such as . . . , is also provided.”)
- L. 21: Change to “. . . , which is in contrast to . . .”.
- L. 23: Change to “. . . can be used by both science and management.”
- L. 28: “Forests are . . .” L. 29: “Exposed to . . .” and replace “and windstorms” to “or windstorms”.
- L. 37: “. . . , and droughts . . .”
- L. 38: Remove “the” before “Western . . .”
- L. 41: Change to “. . .like increasing growing stock . . .”
- L. 43: Remove “as well”.
- L. 46: Remove “of” before “wind-related . . .”
- L. 56: Remove “substantially”
- L. 62: Replace “on the Earth” with “globally”
- L. 66: Replace “publication” with “opening”
- L. 75: Replace “the archive of Landsat images” with “the Landsat archive”
- L. 76: Remove “the” before “public map . . .”
- L. 79: Add “the” before “windthrow delineation . . .”
- L. 74-82: You rather describe where you do what in the manuscript, which is not of interest at this point. I strongly suggest to revise this paragraph to give a detailed description of what data and which specific attributes you will collect; and what is the rationale for collecting those.
- L. 88: “study region” not “study regions”. It is also not clear to me what the following

C3

sentence means. Please revise.

- L. 93: What is “the large area”? Do you mean a specific area?
- L. 100ff: Why is this written in bullet points? I strongly suggest to write the section out as proper text.
- L. 101/102: Change to “. . . forest disturbance at annual temporal resolution.”
- L. 108: Change to “. . . on forest loss classified into . . .”
- L. 122: How do you download images from Google Maps or Bing Maps?
- L. 175: This might be a personal flavor, but could you give the areas in hectare or square meters? A value of 0.0018 km² is hard to image, given all the leading zeros.
- Section 4.1.1.: How did you ensure you didn’t miss a windthrow? Was there some systematic sampling design applied? How can you be sure there are only 450 windthrows? Did you look at each and every disturbance patch?
- L. 195: Revise beginning of the sentence.
- L. 200: How was the Landsat data processed? Were any corrections or masking algorithms applied? More information needed.
- L. 211: Change “the stand-replacing” to “a stand-replacing”
- L. 212: Not sure what is meant by “However, this value may be less if these disturbances hold the substantial part of the image.”? Please explain.
- L. 230: “A similar threshold value . . .”
- L. 230ff: Here for instance I have the feeling that the choice of thresholds is very subjective and depends largely on the interpreter and his/her knowledge. How can we make sure that the data collection is systematic and not biased?

Note: I stopped marking every language issue. Please do a proper language check

C4

before resubmitting.

L. 270: How did you assess whether a storm happened in winter? My guess is that satellite data and HRI data is mostly only available for the growing period.

L. 292ff: How was the central line drawn? By hand? Or did you use an algorithm to do so? Is the line sensitive to the allocation of patches to wind throw events?

L. 389ff: This is discussion and not results. In general you mix up a lot of results reporting and discussion. In general, I think that is fine for such a paper, but then you should rename the section to Results and Discussion; and give the Discussion section a more informed title.

L. 446: You note that no winter windthrows were found. This might be a result of missing winter observations, right?

Conclusion section: You are cautious in interpreting trends in windthrow area due to inconsistencies in your data. This is great and much appreciated. However, you then conclude that the positive trend for large windthrows likely is "real". But how do you come to this conclusion? Without proper validation, we can't be sure about this. I suggest to remove this conclusion.

Comments on the database:

The dataset was downloadable and could be opened in standard GIS software (QGIS). All entries had attributes. The projection of the GIS layers was set. Units were not specific. I thus encourage the authors to add a README or similar metadata file to describe the units (i.e., are, length, etc.).

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2020-91>, 2020.