Point-by-point response, Review I - Bretwood Higman

Reviewer Comments
Author response
Changes made in manuscript

Overall this seems like a useful step forward in characterizing the distribution of hazardous materials sites in Alaska, and the meshing of OSM data with other data sets is clever. Some minor revisions, especially to make the figures more clear, would benefit the paper (see comments in the attached pdf.)

Copied from .pdf:

Abstract, line 2: decreased rather than "a loss"
Thank you for recommending this. We changed the wording accordingly.
--
Introduction: It would be good to note somewhere that impacts of permafrost degradation can sometimes be in downstream areas - e.g. from mass wasting or increased river sedimentation or erosion.
We agree that this is valuable information and added a section on this matter.

Page 2, line 34:

Some of these processes such as thaw slumps have impacts not just locally but even far away in downstream areas as sediments, solubles, and organic matter are eroded from thaw features and may follow different trajectories of transport, biogeochemical processing, and sedimentation depending on environmental conditions (Lamhonwah et al., 2016; Keskitalo et al., 2021; Kokelj et al., 2013) and can also impact ecosystems in these downstream areas (Levenstein et al., 2020).

--

Line 23: Presumably there is warming prior to this time-period too, even if the citation does not explore that.

Thank you for pointing this out. For this statement we relied on the Global Terrestrial Network for Permafrost, of which most of the boreholes were established during the International Polar Year (2007-09). The earliest borehole setups - at least for Alaska - are from the 1970s (Biskaborn et al., 2019; Smith et al., 2022). We adapted the sentence accordingly - pointing out the data basis - and added information on modeled permafrost evolution.

Page 2, line 23:

These increasing air temperatures led to a warming and thawing of permafrost since the 1980s, as borehole measurements across the Arctic demonstrate (Biskaborn et al., 2019; Smith et al., 2022). Modeling studies indicate that the initiation of permafrost warming can be traced back to as early as 1900 (Langer et al., 2024).

--

Line 32: I would also include deep-seated bedrock landslides - I think this might be one of the better citations to support that: https://doi.org/10.3389/feart.2020.00293

We agree that the previously mentioned land surface changes were mainly focused on tundra lowlands. Looking at the entire state of Alaska, mountainous regions must be included. We adapted the enumeration to include rock avalanches, inserted the suggested reference and added an Alaskan study, which used GERALDINE to detect rock avalanches.

Page 2, line 29: Numerous studies demonstrate intensifying land surface changes in the permafrost region which encompass for example processes such as thaw slumping (e.g. Runge et al., 2022; Ramage et al., 2017; Leibman et al., 2021), the development of thermokarst ponds and lakes (e.g. Muster et al., 2017; Jones et al., 2011), thermo-erosional gullying (e.g. Fortier et al., 2007; Godin et al., 2012), ice wedge degradation (e.g. Liljedahl et al., 2016; Jorgenson et al., 2006) and rock avalanches in mountainous regions (e.g. Bessette-Kirton and Coe, 2020; Smith et al., 2023) all pointing to an increasing loss in ground stability.

--

Figure 1: Use more intuitive boundaries that are more usefully distributed - perhaps < 100, 100-1000, 1000-10,000, and >10,000
Also the smallest symbol is essentially invisible on the map. Perhaps make the smallest size larger, and use a dark stroke with a light fill to create contrast against all backgrounds.

Yes, thank you. We changed the scale to the proposed boundaries and agree that it is more intuitive than the former one, which was based on the data’s statistics (quantile and median values). We also increased the size of the smallest symbol.

--

Line 79: Provide in per square kilometer units for SI consistency.

Thank you for pointing this out. We agree that it is important to stay consistent. We now provide the population density in square kilometers and keep the square mile details in parentheses as a reference for readers more used to the imperial system:

Page 4, line 83: With a population of over 733,000 inhabitants and a land area of approx. 1.7 million km2 (The Information Architects of Encyclopaedia Britannica, 2023), it is also the least densely populated state in the U.S., with a population density of 0.5 people per square kilometer (1.3 people per square mile), compared to the rest of the U.S. with a density of 35.9 per square kilometer (93 people per square mile) (Department of Labor and Workforce Development, 2020; World Bank, 2024).”

--

Line 85: I don't know this word. Would "includes" work?

Yes, thank you, “includes” works well. We changed it accordingly.

--

Figure 2: I am not sure this necessitates a large figure like this. If critical, it might be relegated to an appendix or supplement.

Thank you for the suggestion. We agree that the figure is better placed in the appendix. We moved it accordingly (Figure A1).

--

Figure 3: These two colors are too similar, especially for colorblind readers (like me.) Simply making the value of the building color darker would likely solve the problem.

Thank you for making us aware of this color issue. We adapted the figure (now Figure 2) using color-vision deficiency friendly color maps by Crameri [https://zenodo.org/records/8035877].

--

Figure 4: I think this is too many layers for a single map - some are difficult to distinguish. For instance, while the ARCADE watershed boundaries are clear against the dark color for continuous permafrost, it gets lost in some of the lighter colors. Consider making several smaller maps with simpler symbology and fewer distinctions - perhaps it would be useful to have each one show the contaminated sites, but against different backdrops. One for permafrost, one for watersheds, and one for the IS & HI data.

We agree that some layers are hard to distinguish and followed your recommendation to split them in several single maps: The figure (now Figure 3) includes subfigures showing the contaminated sites against a) the backdrop of permafrost zones, subfigure b) blended with the infrastructure data, and c) in the foreground of the watersheds. Thank you for this suggestion.

--

Figure 7: The scale bar does not add information - the specific values are labeled on the colors.

We agree that the color bar doesn’t add any information. We removed it (Figure 6).

--

Point-by-point response, Review II - Anonymous

Reviewer Comments
`Author response`
`Changes made in manuscript`

The combination of available infrastructure data sets to increase the semantic content is a good idea that has the potential to improve future analysis regarding contaminated sites and related hazards. Especially the increased number of usage categories is promising for many applications. The usage examples are well done and will be helpful for future users of the data set.

Below I outlined several points that would improve the manuscript in my opinion.

Figure 2 may be valuable for users of the data but could go into the appendix. Instead, a figure like a flowchart or decision tree on how the input data sets were combined may be beneficial for the reader for a better understanding of the process.

`Thank you for recommending these changes. We moved the figure to the appendix (Figure A1) and added a flowchart (Figure A2), describing the data harmonization and subsequent decision tree process.`

--

Regarding OpenStreetMap: As I understood from the text, the authors use OSM as an alternative source of linear infrastructure instead of the satellite derived Bartsch et al dataset. Previous publications including Bartsch et al 2021, which the authors heavily cite throughout the text, point out that OSM has some major drawbacks, including inconsistencies, and lacking some more recent infrastructure developments. In line 147-148 the authors note that they use OSM for the linear infrastructure, not SACHII, for buildings they describe a decision tree of combining the data sets (from line 207). Is there a reason a similar approach was not used for linear infrastructures, especially given the concerns previous publications have raised about OSM. Maybe just some additional discussion could be included on this.

Thank you for this valuable comment. Indeed, Hjort et al. (2018) report that OSM shows regional inconsistencies (with better coverage in North America and Eurasia compared to Asia) and presumably omits parts of isolated, smaller communities.
In terms of the OSM road network, however, Barrington-Leigh and Millard-Ball (2017) report that approximately 83% percent of the global road network is represented in OSM, providing a valuable data basis.

The decision to use OSM instead of SACHI for linear infrastructure was based on two major concerns:

1. Road features in the SACHI database show a low accuracy (producer’s accuracy of the final product ranging from 26-57 % compared to buildings (79.5-95 %)) (Bartsch et al., 2021).
2. Due to the relatively coarse spatial resolution of 10 m, SACHI also
   a. does not represent small/ unpaved roads
   b. struggles with distinguishing between road adjacent areas (buildings, paved areas in general, etc.) and roads itself, leading to a blurred representation of the road network (visible in Figure 2 b).

From a technical standpoint it is also beneficial to provide the road network as line features (OSM default) to enable routing.
We included these arguments in the section “Materials & Methods” and “Discussion”:

Page 6, line 148:
When examining the linear transport infrastructure, we observed some gaps in the data, particularly in settlements: extracting narrow paths or distinguishing between a linear gravel road and other human-impacted areas, such as driveways or exploration pads, were difficult with the limited spatial resolution of the Sentinel sensors (10 m). In addition, the "road" class showed a particularly low mapping accuracy compared to the "building" class (Bartsch et al., 2021). As OSM on the other hand is estimated to represent 83 % of the global road network (Barrington-Leigh and Millard-Ball, 2017), we decided to use OpenStreetMap data to represent the linear transport infrastructure.

Page 22, line 453:
This limitation may be attributed to the peripheral status of Arctic environments within the global OSM mapping network, primarily due to their sparse population. Hjort et al. (2018) report such a limitation for isolated, smaller communities and with regional variability (e.g. better coverage in North America and Eurasia compared to Asia).

Page 24, line 483:
Nonetheless, 78 % of the true road grid cells were accurately represented in the SIRIUS dataset. When accounting for the offset grid cells, this value increases to 92 %. This value underlines the effectiveness of OSM for representing linear infrastructure opposed to SACHI. OSM allows not only a clear distinction between roads and adjacent infrastructure areas but also the inclusion of narrow roads and footways. Looking ahead, it could prove beneficial to integrate official data from local or federal agencies (e.g. Alaska Department of Transportation) to evaluate the comprehensiveness of the OSM linear infrastructure data. Further, incorporating the Trans-Alaskan-Pipeline would provide a spatial context for contamination, oil exploration and transportation data.

--

Figure 6: interbal – internal?

Yes, thank you. That was a typo.

--

Figure 10 could go into the appendix.

We agree that this is - although useful for the data user - better placed in the appendix. We moved the figure accordingly (Figure A3).

--

Line 470: Is this issue not shown in Figure 8? Maybe include a reference here to this figure, just for improved understanding for the reader.

Yes, thank you for pointing this out. This is shown in the figure. We referenced it in the text.

--

Maybe I missed it in the text, but your input data sets have different extents? So would the accuracy of your output not differ between areas where both/all inputs were available and those that rely solely on OSM data?

Thank you, and yes, you are right we did not address this in the text. We added a statement in the "Discussion" section:

Page 24, line 501:
It is important to note that the high mapping accuracy of 78 % (92 %) for linear and 94 % for polygonal infrastructure in our test area of Shishmaref can likely be expected for most of the coastal regions (until 100 km inland). For inland areas (beyond the extent of SACHI), the infrastructure data relies solely on OSM, which may show the above mentioned limitations. Once again, integrating further official data sources could improve its quality.

Citation: https://doi.org/10.5194/essd-2023-393-RC2