Supporting Information for

GERALDINE (Google earth Engine supRaglAciaL Debris INput dEtector) - A new Tool for Identifying and Monitoring Supraglacial Landslide Inputs

W. D. Smith[1], S.A. Dunning[1], S. Brough[1,2], N. Ross[1], J. Telling[3].

[1]School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne, UK.
[2]Department of Geography and Planning, School of Environmental Sciences, University of Liverpool, UK
[3]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, UK.

Corresponding author: Will Smith (w.d.smith2@newcastle.ac.uk)

1. Randolph Glacier Inventory (RGI) v6.0 errors

The Randolph Glacier Inventory v6.0 (Pfeffer et al., 2014) is a global dataset of digitised glacier outlines, excluding the ice sheets of Greenland and Antarctica. These outlines were digitised from images acquired between 1943 and the present day. This large temporal array of images introduces a variety of dataset errors, due to rapid glacial thinning and retreat, in response to climatic change over the last century. GERALDINE's delineation of new debris additions on glaciers is impacted by these discrepancies.

We run GERALDINE over the Columbia glacier, a glacier which has undergone unprecedented retreat over the last three decades, to demonstrate RGI impacts (Figure S1). RGI outlines in this region were digitised from imagery acquired in 2009 but the glacier has retreated ~8 km since then, losing 46.9 km^2 at its tongue. Clean ice is 'found' in the now ice-free area, owing to ice melange/icebergs, and debris is 'found' due to the dark, unreflective ocean surface. We advise caution with the delineation of new debris additions near the terminus of marine-terminating glaciers, if they have recently undergone substantial retreat. Glacial thinning/down-wasting can also cause discrepancies in GERALDINE outputs. All tributaries of the western trunk of the Columbia glacier have undergone down-wasting, exposing valley walls. GERALDINE delineates this down-wasting and subsequent valley wall exposure as surficial debris, represented by each glacier having thick lateral 'debris' bands (Figure S1B). These are likely to expand with further glacial mass wastage, until RGI outlines are updated in v7.0.
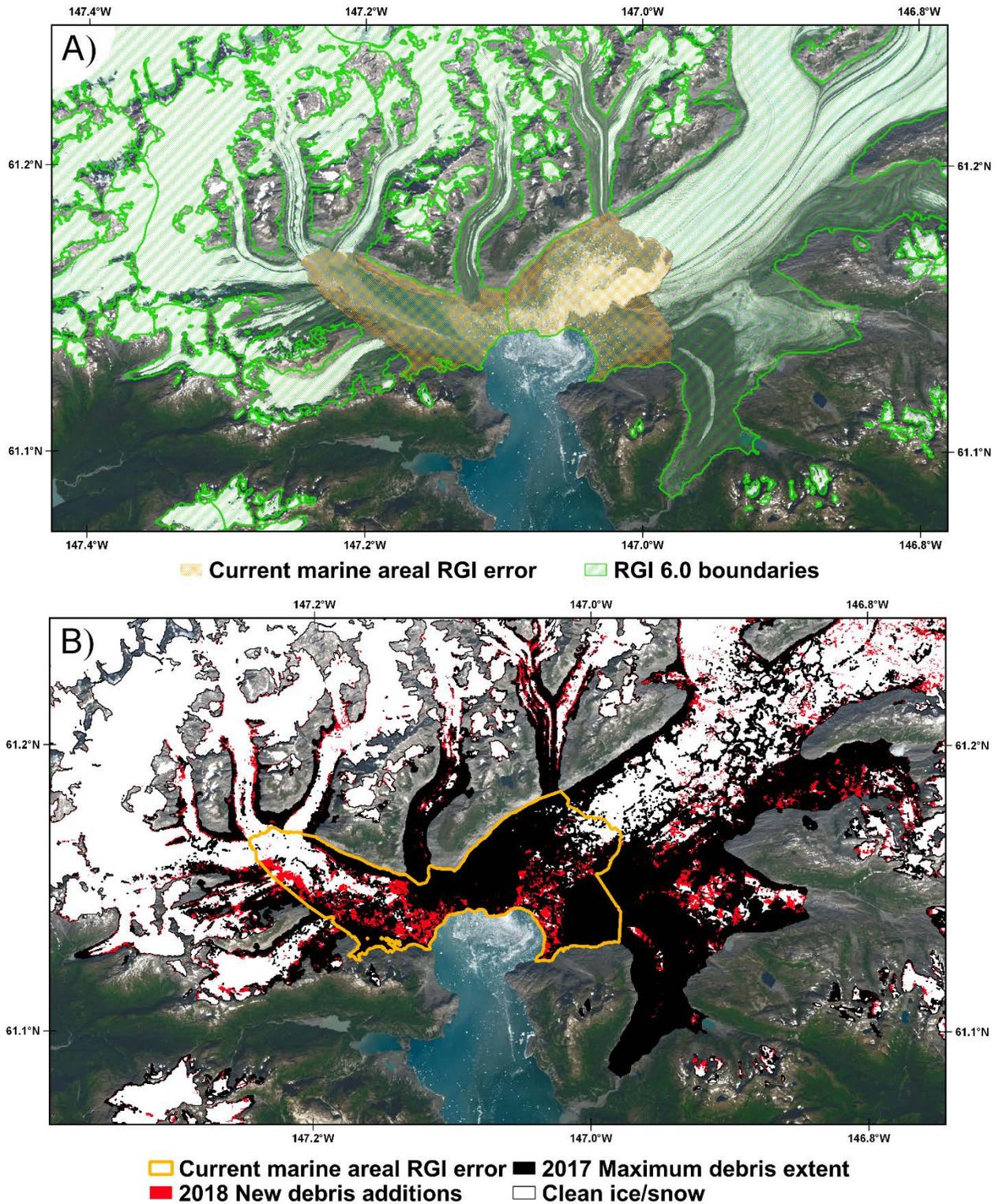
Figure S1: Retreat of the Columbia Glacier, Alaska and the impact on RGI v6.0 outline accuracy. A) RGI v6.0 glacier outlines (green) and the area, which is no longer glaciated (orange). B) GERALDINE new debris results in this area for 2018. Landsat 8 background image from 2019-08-20.

2.0 Cloud mask threshold validation

To determine the optimum threshold for cloud masking we chose 13 rock avalanches (RA) from the validation dataset, to validate different thresholds against. The high prevalence of RAs in Alaska meant all events chosen for cloud mask threshold validation occurred in the region, with nine occurring in Glacier Bay National Park (Coe et al., 2018) and four occurring in the eastern Alaska Range, in particular the area around Mt Hayes (Jibson et al., 2006) (Table 1). This selection incorporated two areas with different climatic regimes (marine vs continental climate) and a wide temporal coverage incorporating all Landsat satellites, with particular focus on the current constellation, due to the tools main use for aiding RA detection in the present day.

Table 1: Rock avalanches (RA) used for cloud mask threshold validation

| RA Name	Region	Year	Latitude	Longitude	Satellite	Area (km2)
2	Glacier Bay National Park	1986	58.794	-137.354	Landsat 4/5	0.025
10	Glacier Bay National Park	1986	58.830	-137.502	Landsat 4/5	1.371
14	Glacier Bay National Park	1986	58.647	-137.058	Landsat 4/5	0.544
Black Rapids Glacier W	Eastern Alaska Range (Mt Hayes)	2002	63.472	-146.263	Landsat 5/7	1.136
McGinnis Peak Glacier S	Eastern Alaska Range (Mt Hayes)	2002	63.547	-146.198	Landsat 5/7	1.321
Black Rapids Glacier E	Eastern Alaska Range (Mt Hayes)	2002	63.461	-146.167	Landsat 5/7	1.612
Black Rapids Glacier M	Eastern Alaska Range (Mt Hayes)	2002	63.466	-146.226	Landsat 5/7	1.950
7 - John Hopkins	Glacier Bay National Park	2015	58.769	-137.269	Landsat 7/8	3.375
11	Glacier Bay National Park	2015	58.626	-137.281	Landsat 7/8	1.223
23	Glacier Bay National Park	2015	58.622	-137.275	Landsat 7/8	0.586
24 - Lamplugh	Glacier Bay National Park	2016	58.779	-136.888	Landsat 7/8	19.174
26	Glacier Bay National Park	2016	58.635	-137.018	Landsat 7/8	0.231
27	Glacier Bay National Park	2016	58.852	-137.247	Landsat 7/8	0.146
28	Glacier Bay National Park	2016	58.844	-137.345	Landsat 7/8	0.707

Five different cloud thresholds were tested: 10%, 20%, 30%, 50% and 90%, to investigate their influence on highlighting new RA events. GERALDINE was run for the year of the event, or in the case of the eastern Alaska Range RAs, the year after the event, because these RAs occurred in November 2002 and therefore appeared in no Landsat imagery during that year. New debris layers generated by GERALDINE were downloaded, and the area of new debris detected in the location of RAs, was compared to digitised RA outlines from the same year. A cloud mask threshold of 20% highlighted the largest area of new RAs, delineating 60.6% of RA area (Figure S2). The 10% cloud threshold masked too much of each image, inhibiting its ability to highlight new debris. Higher thresholds did not mask enough cloud discrepancies out of images resulting in misclassification of cloud as debris in previous year debris extents. This misclassification prevented GERALDINE from highlighting new debris because debris was already present in the previous year. GERALDINE therefore utilises a 20% cloud threshold by default.
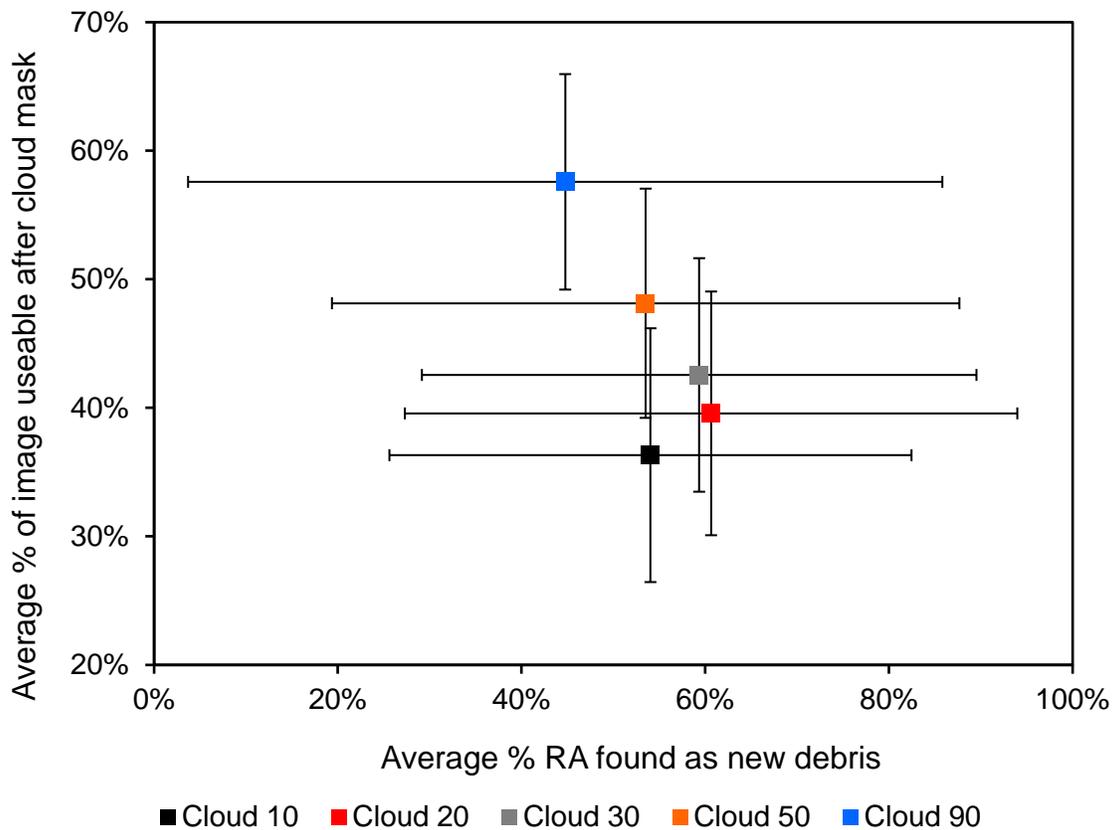
Figure S2: Mean area of RA highlighted as new debris by GERALDINE, and mean area of each Landsat image available after cloud masking, for five different cloud thresholds. Error bars represent standard deviation of the mean.

3.0 NDWI mask

To evaluate NDWI thresholds we utilised the 2018 glacial lake inventory of (Wang et al., 2020). We chose two glacial regions with a high frequency of glacial lakes, the Bhutan Himalaya and Sagarmatha National Park. Lakes were clipped to RGI boundaries for analysis and any lakes <0.005 km^2 were omitted, resulting in 82 and 98 lakes respectively (Figure S3). The average NDWI values of all lakes were mapped across 2018 (Figure S4) along with average NDWI values of supraglacial debris (NDSI debris masks from Scherler et al., 2018) and clean ice/snow, calculated by differencing debris masks from RGI boundaries.
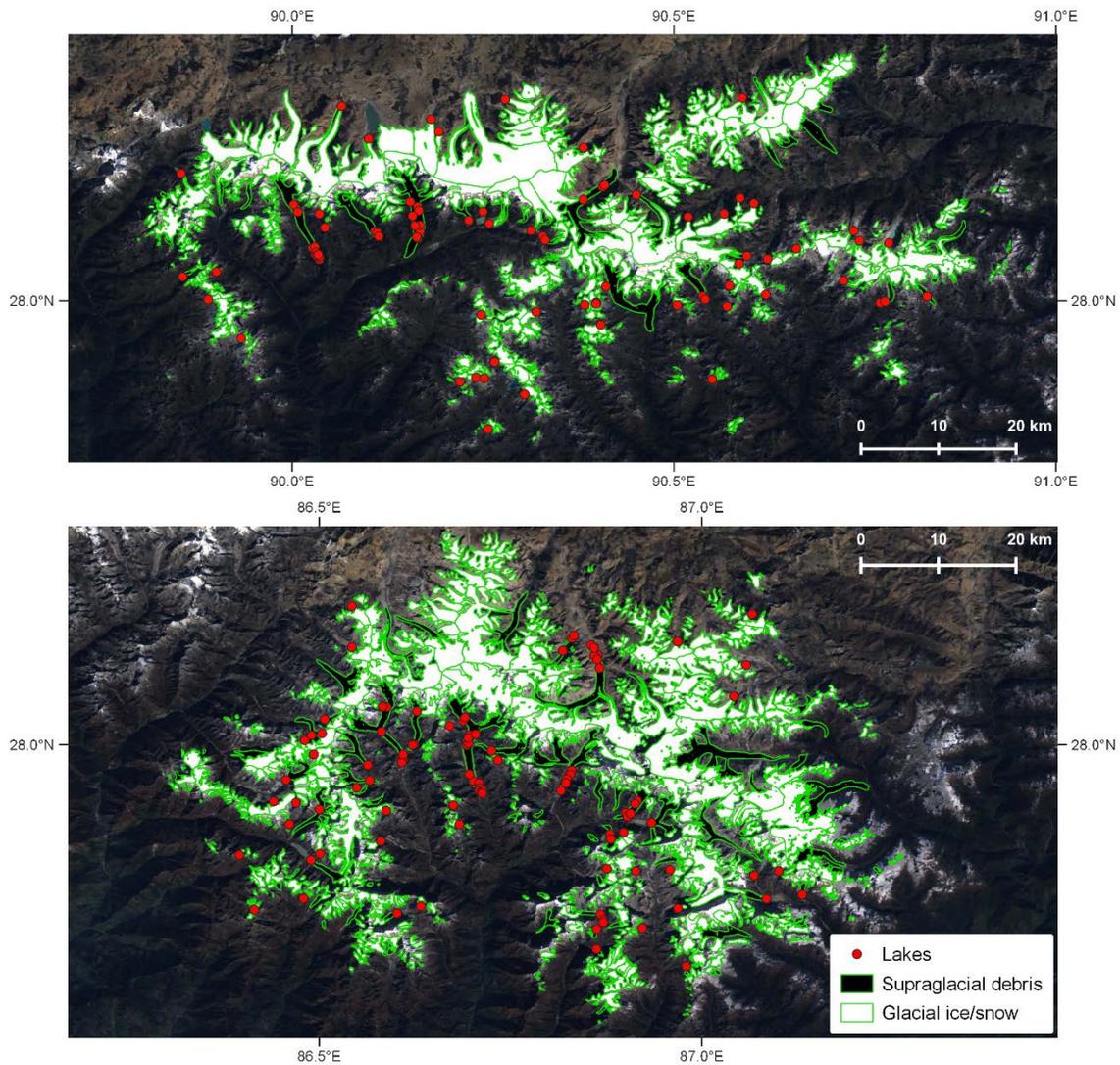
Figure S3: Glacial lake locations in the A) Bhutan Himalaya and B) Sagarmatha National Park. Glacial outlines are from the RGI v6.0 and supraglacial debris is courtesy of Scherler et al. (2018). Background images are 2018 Landsat 8 cloud free mosaics created using Google Earth Engine.

NDWI values follow a seasonal trend over 2018 with lower values between February and May (Figure S4), potentially due to the ephemeral state of water within them i.e. frozen in winter months, modifying their spectral signature to that similar of clean ice and debris. These large fluctuations in values between -0.056 and 0.717 inhibits any threshold from masking these features year-round. GERALDINE employs a 0.4 threshold to mask water bodies, as debris and clean ice do not get close to this value (highest observed was in Sagarmatha National Park on 02/12/2018 at 0.287) and it has been widely used in other studies in mountainous regions (Acharya et al., 2018; Zhao et al., 2018). We recognize that a solitary threshold is sub-optimum (Li and Sheng, 2012), particularly for global scale analysis but the user can tune this value (code line 245 and 255) for their specific ROI if there is a high prevalence of water bodies.

Average water body size within RGI boundaries from Wang et al. (2020) was 0.027 ± 0.086 km^2 equating to just 22 Landsat pixels. Proglacial lakes are the largest water bodies within
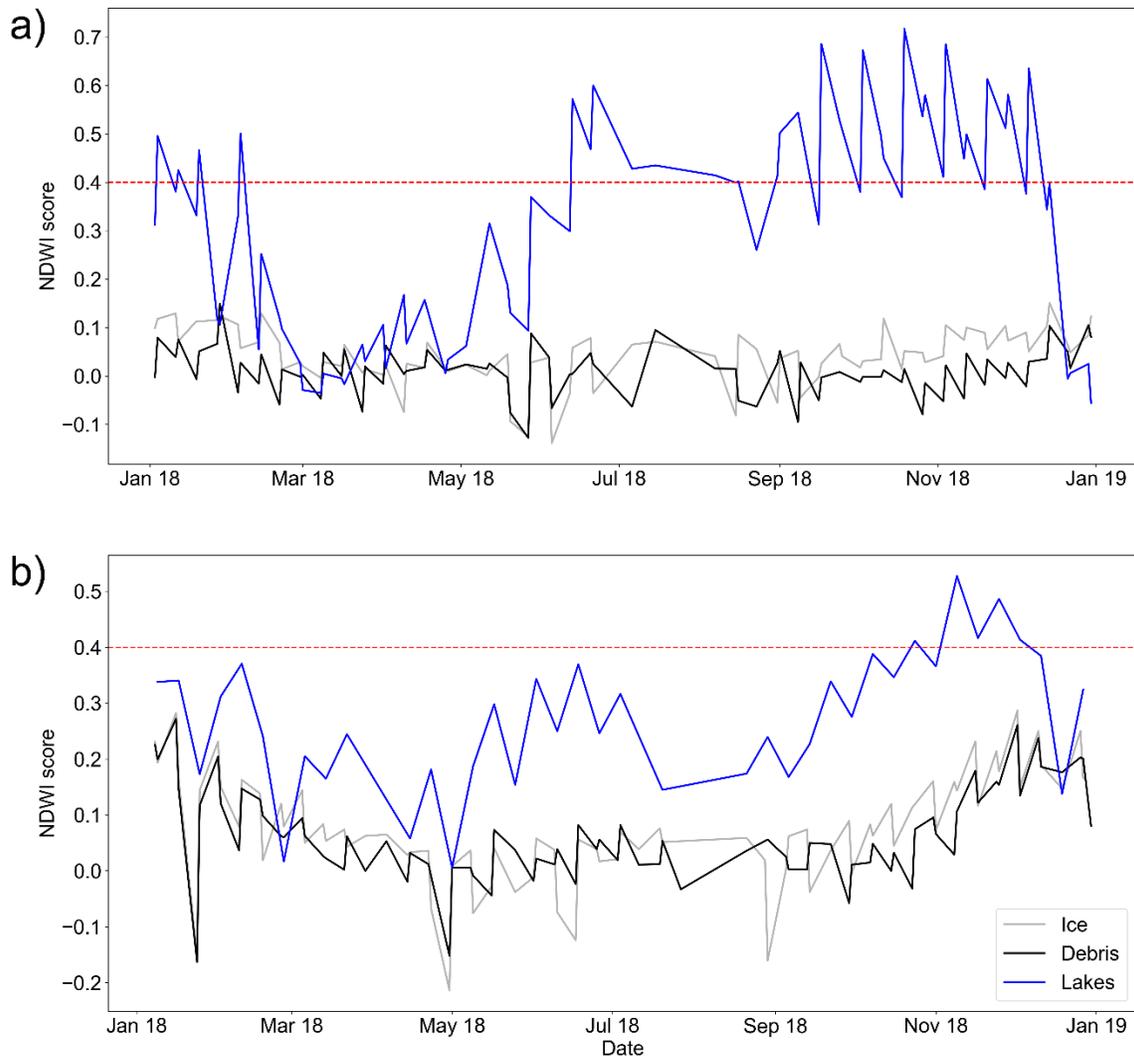
Figure S4: Mean NDWI scores in a) the Bhutan Himalaya and b) Sagarmatha National Park for all ice, debris and proglacial/supraglacial lakes within RGI boundaries. Red dashed line indicates NDWI threshold – anything greater than 0.4 is masked from analysis.

RGI boundaries due to glacial retreat since RGI boundaries were digitised. Updated glacier margins would omit proglacial lakes from analysis and can be utilized if available (see section 2.0 Method) but the RGI is the best worldwide inventory. The dark spectral nature of water bodies means they are often depicted as debris in the previous year debris extent, due to their low reflectance in the visible spectrum, so they do not influence detection of new debris additions, preventing false detections.

4.0 GERALDINE User Guide

The tool is freely available to use at (https://code.earthengine.google.com/79a0e5b27c2d7824559314e09f5971a9?hideCode=true) but requires a Google account authorized to use Google Earth Engine (GEE), which is free of charge if used for research and educational purposes (sign up for Google account here: https://accounts.google.com/signup/v2/webcreateaccount?flowName=GlifWebSignIn&flowEntry=SignUp and register for GEE access here: https://earthengine.google.com/). Exporting of tool outputs requires a Google Drive account, which is complimentary with the Gmail account required to sign up for GEE. The tool is open access and GUI (graphical user interface) driven. Tutorials on how to use Earth Engine are available at https://developers.google.com/earth-engine/ but here we will provide instructions on how to use our tool to detect supraglacial debris inputs.

Step 1:

Open v1.0 of GERALDINE (the version described in the manuscript) by clicking on this link: https://code.earthengine.google.com/79a0e5b27c2d7824559314e09f5971a9?hideCode=true or access the latest version of GERALDINE at https://doi.org/10.5281/zenodo.3524414 (if using the latest version these instructions may differ slightly).

Step 2:

You will be greeted by the start page shown below. Click 'New project' to start analysis.
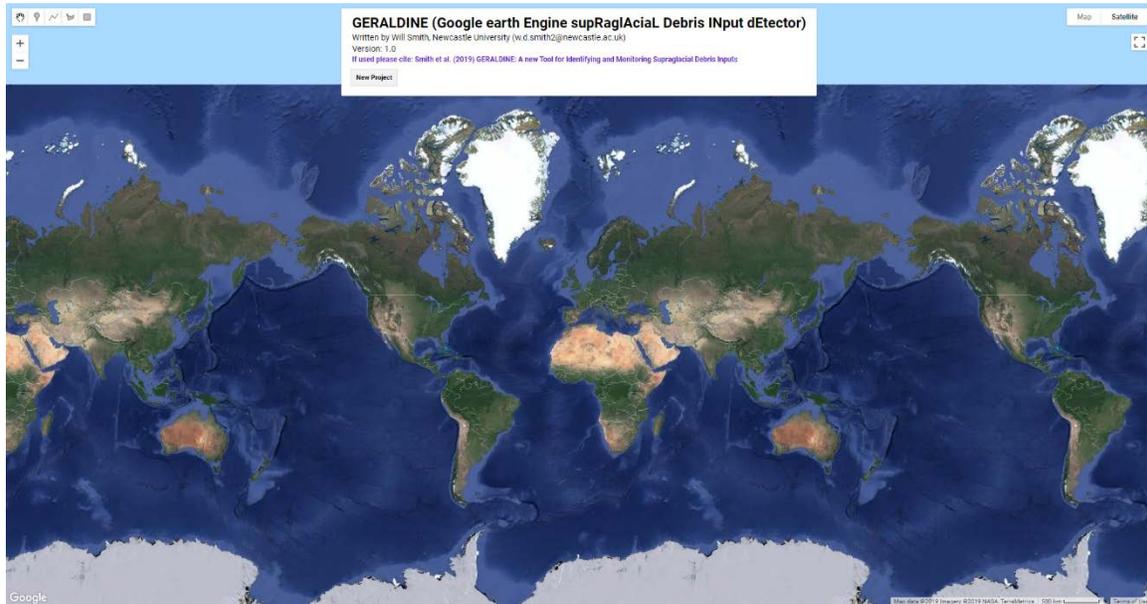

Step 3:

Draw region of interest (ROI) by zooming in and clicking around an area to draw a polygon (Note: tool areas should not exceed 5000 km² due to GEE memory capacity). Alternatively, upload a shapefile of your ROI to Google Earth Engine (see: https://developers.google.com/earth-engine/importing for more information) and specify the GEE file path, which can be found by

sliding down the top panel and navigating to the 'Assets' tab in the top left hand panel (highlighted by red box in image below). Click OK button when your ROI is defined.
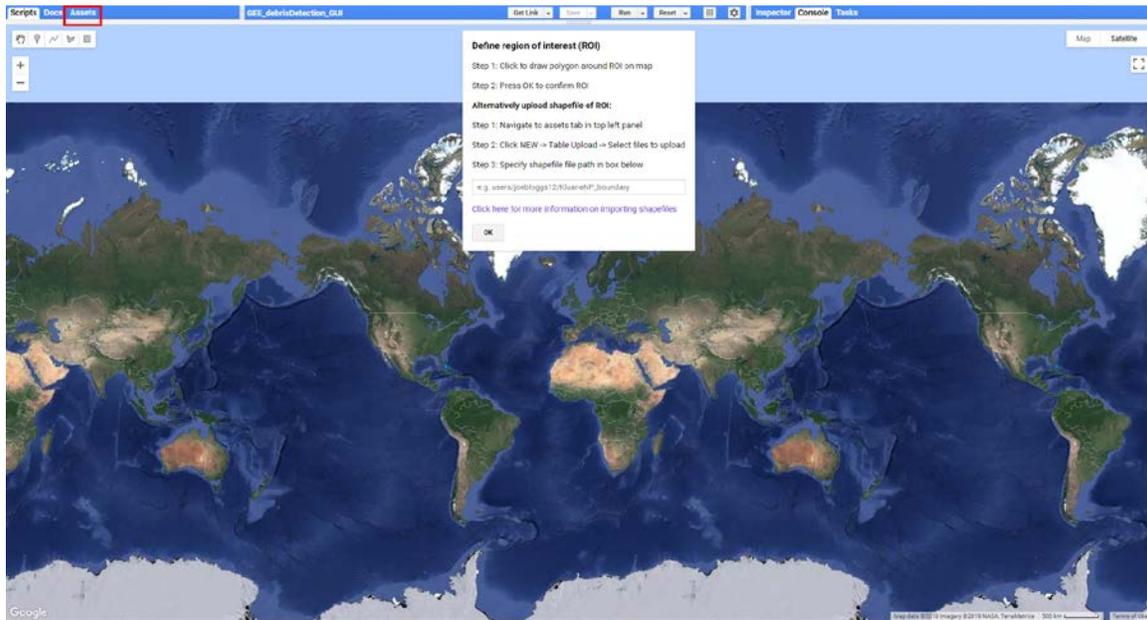

Step 4:

Specify date range from which you want the tool to detect new debris additions and select if you would like to use tier 2 and/or real time Landsat imagery in addition to the default tier 1 imagery. Tier 2 imagery is useful if minimal tier 1 imagery is available i.e. in Antarctica, and real time imagery should only be used if the event has occurred in the previous 16 days (Note: for real time imagery 'End date' must be set as todays date). Tool accuracy and speed is optimum if date ranges are annual or sub-annual and only tier 1 imagery is utilised. Date must be in the format of Year – Month – Day e.g. 2018-12-22. Press OK once start and end date are defined.
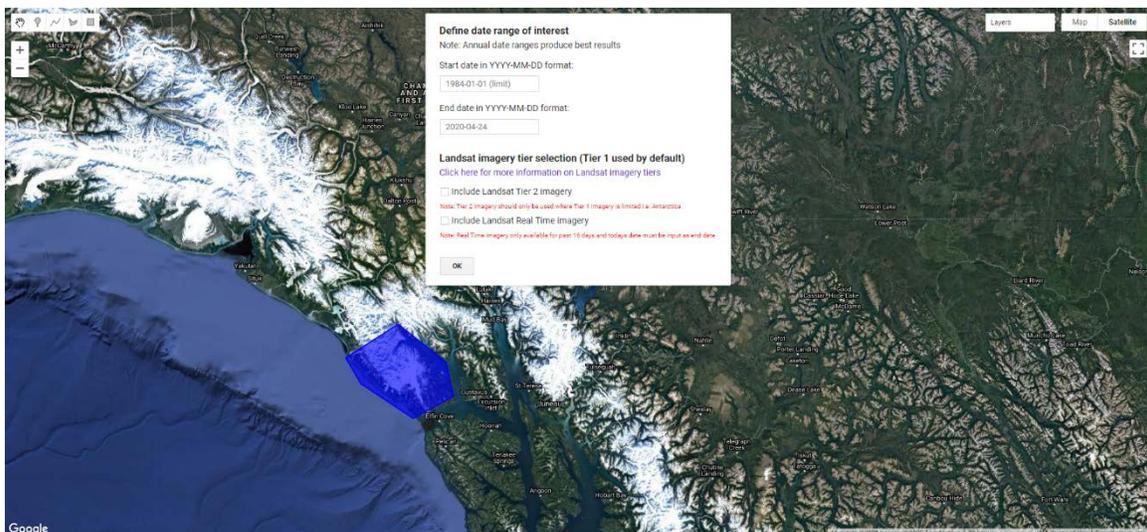
Step 5:

Tool should display results on map (it can take up to 3 minutes for layers to load if your ROI is large). Two layers are created: a previous year maximum debris cover layer and a new debris additions layer. The user can view and toggle these layers by hovering the mouse over the 'Layers' button in the top right hand corner of the map viewer (highlighted by red rectangle in below image). To export the data click on the Export data button.
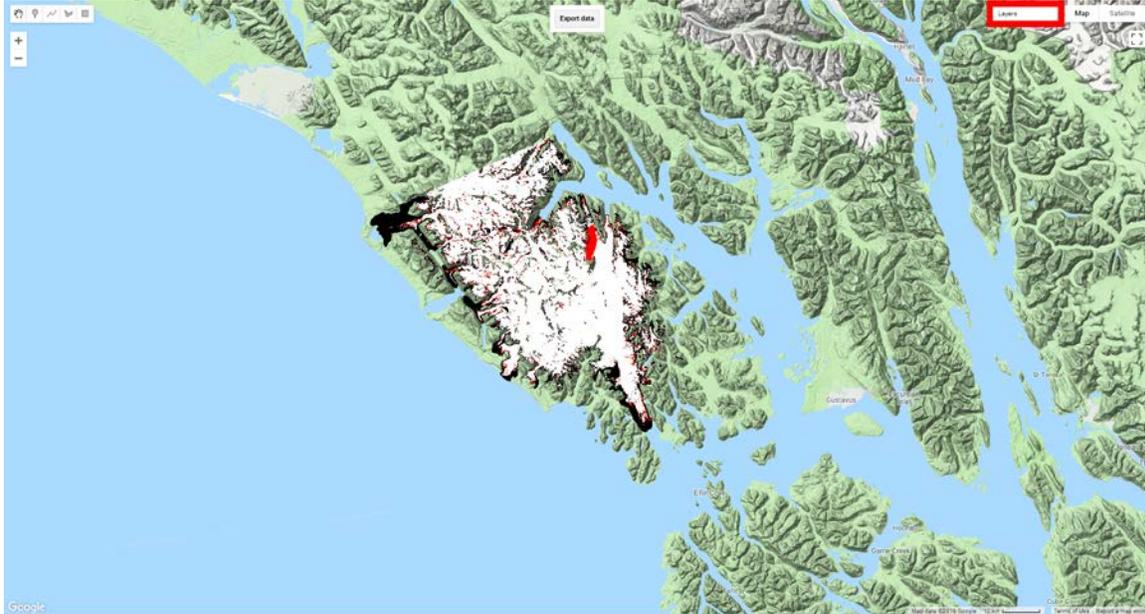
Step 6:

Instructions are displayed detailing how to export data from GEE. Once you have navigated to the Task tab in the top right hand panel and clicked 'Run' next to the layer you wish to download (note: you do not need to wait for layers to load within GEE before you export). The following window will be displayed (see image below), prompting the user to confirm or alter the filename, confirm the export format (GeoJSON is strongly recommended because it decreases export time), and confirm the save location. Once data is exported, it can be used in a GIS of your choice. Alternatively, you can save your files as an Earth Engine asset, this is particularly useful for your ROI, enabling you to call it in during Step 3, instead of redrawing it every time you use GERALDINE.
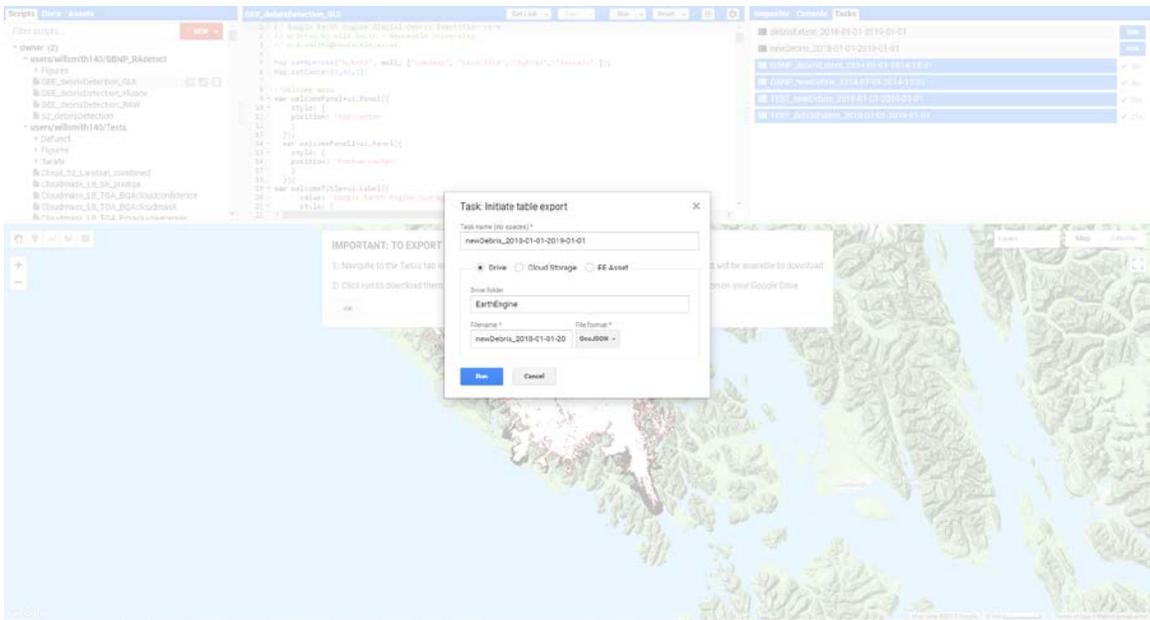
References

Acharya, T. D., Subedi, A. and Lee, D. H.: Evaluation of water indices for surface water extraction in a landsat 8 scene of Nepal, Sensors (Switzerland), 18(8), doi:10.3390/s18082580, 2018.

Coe, J. A., Bessette-Kirton, E. K. and Geertsema, M.: Increasing rock-avalanche size and mobility in Glacier Bay National Park and Preserve, Alaska detected from 1984 to 2016 Landsat imagery, Landslides, 15(3), 393–407, doi:10.1007/s10346-017-0879-7, 2018.

Jibson, R. W., Harp, E. L., Schulz, W. and Keefer, D. K.: Large rock avalanches triggered by the M 7.9 Denali Fault, Alaska, earthquake of 3 November 2002, Eng. Geol., 83(1–3), 144–160, doi:10.1016/j.enggeo.2005.06.029, 2006.

Li, J. and Sheng, Y.: An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: A case study in the Himalayas, Int. J. Remote Sens., 33(16), 5194–5213, doi:10.1080/01431161.2012.657370, 2012.

Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J.-O., Hock, R., Kaser, G., Kienholz, C., Miles, E. S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B. H., Rich, J. and Sharp, M. J.: The Randolph Glacier Inventory: a globally complete inventory of glaciers, J. Glaciol., 60(221), 537–552, doi:10.3189/2014jog13j176, 2014.

Scherler, D., Wulf, H. and Gorelick, N.: Global Assessment of Supraglacial Debris-Cover Extents, Geophys. Res. Lett., 45(21), 11,798-11,805, doi:10.1029/2018GL080158, 2018.

Wang, X., Guo, X., Yang, C., Liu, Q., Wei, J., Zhang, Y., Zhang, Y., Jiang, Z. and Tang, Z.: Glacial lake inventory of High Mountain Asia (1990 – 2018) derived from Landsat images, Earth Syst. Scence Data Discuss., (January), 1–23, doi:10.5194/essd-2019-212, 2020.

Zhao, H., Chen, F. and Zhang, M.: A Systematic Extraction Approach for Mapping Glacial Lakes in High Mountain Regions of Asia, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 11(8), 2788–2799, doi:10.1109/JSTARS.2018.2846551, 2018.