Remote Sensing of Environment 285 (2023) 113379

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Historical Structure from Motion (HSfM): Automated processing of historical aerial photographs for long-term topographic change analysis

Friedrich Knuth [a,*], David Shean [a], Shashank Bhushan [a], Eli Schwat [a], Oleg Alexandrov [b], Christopher McNeil [c], Amaury Dehecq [d], Caitlyn Florentine [e], Shad O'Neel [f]

[a] University of Washington, Civil and Environmental Engineering, Seattle, WA, USA
[b] NASA Ames Research Center, Intelligent Robotics Group, Moffet Field, CA, USA
[c] U.S. Geological Survey Alaska Science Center, Anchorage, AK, USA
[d] Université Grenoble Alpes, CNRS, IRD, Grenoble INP, IGE, Grenoble, France
[e] U.S. Geological Survey Northern Rocky Mountain Science Center, Bozeman, MT, USA
[f] U.S. Army Corps of Engineers, Cold Regions Research and Engineering Laboratory, Hanover, NH, USA

ARTICLE INFO

Edited by Menghua Wang

Keywords:
Photogrammetry
Co-registration
Topography
Land surface
Elevation
DEM
Time series
Glacier
NAGAP

ABSTRACT

Precisely measuring the Earth's changing surface on decadal to centennial time scales is critical for many science and engineering applications, yet long-term records of quantitative landscape change are often temporally and geographically sparse. Archives of scanned historical aerial photographs provide an opportunity to augment these records with accurate elevation measurements that capture the historical state of the Earth surface. Structure from Motion (SfM) photogrammetry workflows produce high-quality digital elevation models (DEMs) and orthoimage mosaics from these historical images, but time-intensive tasks like manual image preprocessing (e.g., fiducial marker identification) and ground control point (GCP) selection impede processing at scale. We developed an automated method to process historical images and generate self-consistent time series of high-resolution (0.5–2 m) DEMs and orthomosaics, without manual GCP selection. The method relies on SfM to correct camera interior and exterior orientation and a robust multi-stage co-registration approach using modern reference terrain datasets for geolocation refinement. We demonstrate the method using scanned images from the North American Glacier Aerial Photography (NAGAP) archive collected between 1967 and 1997. We present results for two sites with variable photo acquisition geometry and overlap — Mount Baker and South Cascade Glacier in Washington State, USA. The automated method corrects initial camera position errors of several kilometers and produces accurately georeferenced, high-resolution DEMs and orthoimages, regardless of camera configuration, acquisition geometry, terrain characteristics, and reference DEM properties. The average RMS reprojection error following bundle adjustment optimization was 0.67 px (0.15 m) for the 261 images contributing to 10 final DEM mosaics between 1970 and 1992 at Mount Baker, and 0.65 px (0.13 m) for the 243 images contributing to 18 individual DEMs between 1967 and 1997 at South Cascade Glacier. The relative accuracy of elevation values in the historical time series stacks was 0.68 m at Mount Baker and 0.37 m at South Cascade Glacier. Our products have reduced systematic error and improved accuracy compared to DEM products generated using SfM with manual GCP selection. Final elevation change measurement precision was ~0.7–1.0 m over a 30-year period, enabling the study of processes with rates as low as ~1-3 cm/yr. Our results demonstrate the potential of this scalable method to rapidly process archives of historical imagery and deliver new quantitative insights on long-term geodetic change and Earth surface processes.

1. Introduction

Millions of analog photographs documenting land surface change between the 1910s and 2000s sit isolated in vaults. A small fraction of these photographs has been digitally scanned, but remains underutilized for quantitative analysis in Earth system science due to limited metadata and poor geolocation accuracy. Measuring surface elevation

* Corresponding author.
E-mail addresses: knuth@uw.edu (F. Knuth), dshean@uw.edu (D. Shean), sbaglapl@uw.edu (S. Bhushan), elilouis@uw.edu (E. Schwat), oleg.alexandrov@nasa.gov (O. Alexandrov), cmcneil@usgs.gov (C. McNeil), amaury.dehecq@univ-grenoble-alpes.fr (A. Dehecq), cflorentine@usgs.gov (C. Florentine), shad.r.oneel@erdc.dren.mil (S. O'Neel).

https://doi.org/10.1016/j.rse.2022.113379
Received 31 May 2022; Received in revised form 14 November 2022; Accepted 19 November 2022
Available online 7 December 2022

change is fundamental to many Earth science disciplines, including hydrology, glaciology, geomorphology, and volcanology (e.g., Eaton and Murata, 1960; Østrem, 1975; Schumm, 1956). Observations of surface elevation change through time help quantify the nature of many geophysical processes, such as glacier mass balance (e.g., Hugonnet et al., 2021; Shean et al., 2020), ice dynamics in response to climate forcing (e.g., Dehecq et al., 2019; Kehrl et al., 2017), sediment transport (e.g., Anderson and Pitlick, 2014), and volcanically induced deformation (e.g., Henderson et al., 2019). Scanned historical aerial image archives provide an opportunity to "go back in time" and derive accurate historical digital elevation models (DEMs) and orthoimages for long-term measurements of surface change with many applications, including Earth system modeling, climate adaptation, urban planning, and land management (Bauer et al., 2021).

Photogrammetry is the science of making precise measurements in photographs (e.g., Meydenbauer, 1867; LaChapelle, 1962; Karara, 1967; Albertz, 2001). Traditional stereo photogrammetry methods used overlapping photographs and stereoplotters to manually delineate contour lines (e.g., Ray, 1960; Van Wijk, 1967) and interpolate elevations to prepare continuous DEM grids (e.g., Chandler and Cooper, 1989; Gesch et al., 2014). These approaches have offered valuable information over the past century, yet associated errors remain high and careful manual calibration and correction of these products is required for use in modern applications (e.g., geodetic glacier-mass-balance calculations) (Fischer et al., 2015).

The Structure from Motion (SfM) photogrammetric technique (e.g., Snavely et al., 2008) draws on both traditional photogrammetry and modern computer vision theory (Westoby et al., 2012; Eltner and Sofia, 2020) to provide improved automated stereo reconstruction. Traditional photogrammetric approaches require rigorously calibrated camera systems with precisely known interior orientation parameters (e.g. Brown, 1971; Tayman, 1984) and/or manually defined tie points between images to reconstruct the relative camera exterior orientation (position and attitude) (e.g., Westoby et al., 2012). Modern SfM approaches, by comparison, rely almost entirely on information in the images and simple camera models to iteratively solve for camera interior orientation and relative camera exterior orientation using thousands of automatically detected feature matches between overlapping images (e.g., Martin-Brualla et al., 2015; Snavely et al., 2008; Lowe, 1999). This advancement provides an essential foundation toward automated processing of historical archive imagery with poorly constrained camera information.

Due to the nature of historical imagery, several labor- and time-intensive tasks often hinder processing large datasets (e.g., Walstra et al., 2007; Bjørk et al., 2012; Mölg and Bolch, 2017; Girod et al., 2018; O'Neel et al., 2019; Child et al., 2020; Roberti et al., 2021; Geyman et al., 2022). To begin, the film must be scanned by a trained operator with professional equipment. Historical aerial images were typically collected on nitrate and acetate film, which can degrade and warp with age. The film may be further stretched and distorted during scanning, especially if a non-photogrammetric scanner is used (e.g., Gruber and Leberl, 2001; Nielsen, 2017; Salach, 2017; Sevara, 2016). In some cases, the scanning process can introduce additional sources of error, such as over-exposure, irregular film rotation on the scanner, and omission of entire portions of the photo frame. Perhaps more importantly, inaccurate or unknown camera pose (extrinsic position and orientation) significantly reduces the accuracy of the resulting products in a geographic coordinate system. The current solution to this issue usually involves manually selecting features in the images with known three-dimensional (3D) world coordinates to serve as ground control points (GCPs) (e.g., Feurer and Vinatier, 2018; Mölg and Bolch, 2017; Warrick et al., 2017), which provide absolute constraints during joint optimization (bundle adjustment) of camera interior and exterior orientation (e.g., Mikhail et al., 2001; Tarolli and Mudd, 2019; Westoby et al., 2012). For historical aerial photographs taken over remote areas, these GCPs are typically persistent natural features that are identifiable in both modern and historical datasets and assumed to be static over time, such as prominent rock outcrops and large boulders. Confident image feature identification requires at least three pixels, and the error of natural GCP coordinates scales with the modern and historical image resolution. The time-intensive GCP picking process is further limited by the resolution and accuracy of the modern reference data. Finally, the lack of accurate camera pose information for historical images prohibits predetermination of overlapping image blocks needed for efficient feature matching. Here, we define a block as three or more overlapping images collected on the same acquisition date with the same camera. This limitation can be overcome with brute-force matching of all potential image pairs, which is computationally expensive and potentially error-prone for surfaces with repetitive texture.

To solve these issues, we developed an automated method that includes image preprocessing and block detection, photogrammetric surface reconstruction, and DEM co-registration to efficiently generate accurate historical DEMs and orthomosaics without manual intervention. Our goal was to develop a robust and generic approach that is scalable and applicable to any collection of historical vertical aerial photography, even when metadata records are incomplete. To that end, we focused our efforts on the challenging U.S. Geological Survey (USGS) North American Glacier Aerial Photography (NAGAP) archive. We evaluated this method for two sites in the Pacific Northwest region of the contiguous United States with mountain glacier and geomorphic change signals. Our methodology addresses several common limitations of historical aerial image processing to generate accurate DEMs, which reduces uncertainty for future studies of long-term glacier mass change.

2. Study sites

We focus our study on Mount Baker (261 NAGAP images covering ~143 km^2 area) and South Cascade Glacier (243 NAGAP images covering ~32 km^2 area), which are both located in Washington State, USA (Fig. 1). The sites differ in terms of terrain and land cover characteristics (Figs. 1A and B), with variable NAGAP image acquisition geometries (Figs. S1 and S2). Glacier outlines for each site were extracted from the Randolph Glacier Inventory (RGI) v6.0 (RGI Consortium, 2017).

2.1. Mount Baker

Mount Baker is a stratovolcano in northern Washington State. It holds the record for highest seasonal snowfall in the USA (Leffler et al., 2001) and is the source of considerable sediment transported by the Skagit and Nooksack Rivers (Czuba et al., 2011). Easton Glacier (*RGI60-02.17739*) covers an area of 2.89 km^2 (1974 extent) on the southern flank of Mount Baker (Fig. 1A) and has been monitored annually by the North Cascade Glacier Climate Project since 1984 (Pelto and Brown, 2012; WGMS, 2021). The lower glacier is flanked by Little Ice Age lateral moraines to the east and west (Osborn et al., 2012), and encroaching conifer vegetation to the south, but otherwise surrounded by exposed and unvegetated bedrock and periglacial deposits.

2.2. South Cascade Glacier

South Cascade Glacier (*RGI60-02.18778*) covers an area of 2.92 km^2 (1958 extent) in the northern Cascade Range. It is one of five benchmark glaciers monitored by the USGS to capture long-term glacier mass change for western North America and Alaska (RGI regions 1 and 2, respectively) (Meier and Tangborn, 1965; O'Neel et al., 2019). South Cascade Glacier has the longest continuous record of annual mass-balance measurements in the contiguous United States, from 1958 to present (McNeil et al., 2022b).

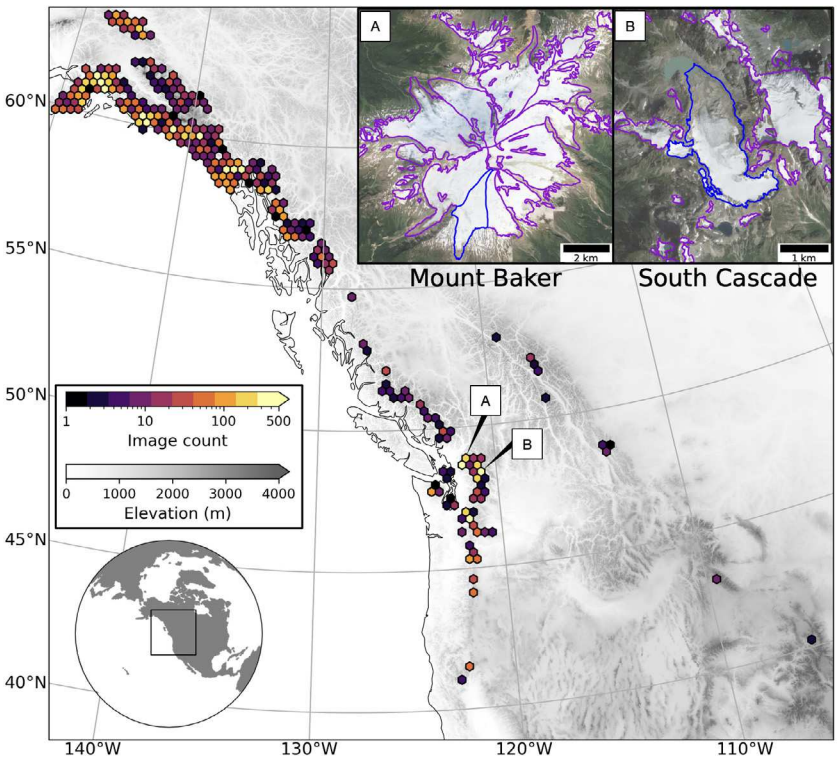

Fig. 1. Coverage map for scanned North American Glacier Aerial Photography (NAGAP) imagery acquired between 1964 and 1997 across western North America available from the NSF Arctic Data Center. Hexagon colors show the total number of frames within each ~25 km bin (~1,600 km^2 area). Basemap shows surface elevation in meters above the WGS84 ellipsoid from the 3-arcsecond Copernicus DEM (European Space Agency, 2022), resampled to a 1 km grid with Albers Equal Area projection. Annotations call out hexagon cells containing our study sites, with inset maps for: A) Mount Baker site with Randolph Glacier Inventory (RGI) v6.0 outlines from 1974 source imagery (purple) and Easton Glacier highlighted in blue, and B) South Cascade Glacier site with RGI outline from 1958 source imagery.

3. Data

The current method supports historical image and metadata retrieval from multiple archives, including the National Science Foundation (NSF) Arctic Data Center (ADC) and the USGS Earth Resources Observation and Science (EROS) Aerial Photo Single Frames archive available via the USGS EarthExplorer (EE) portal. For this study, we used historical images collected by the USGS NAGAP program, primarily hosted in digitized format by the NSF ADC (Nolan et al., 2017; Nolan, 2012), with subsets of data hosted by USGS EE.

The NAGAP program collected ~100,000 aerial photographs of glaciers in western North America and Alaska between the 1960s and 1990s (Fig. 1) (Post, 1995; Post et al., 1971). The main objective of the NAGAP program was to document regional glacier change to complement the long-term USGS Benchmark Glacier field monitoring program (Florentine and Mckeon, 2022; Berkner, 1954; Crandell et al., 1965; UNESCO, 1970; Nace, 1969). Due to the remote location and inaccessibility of many North American glaciers, aerial photography provided an effective approach to conduct regional glacier surveys and provide context for the benchmark glaciers with traditional stake mass-balance measurements (e.g., O'Neel et al., 2019; Zemp et al., 2013). Some of the NAGAP photographs were initially used to derive contour maps using traditional photogrammetric methods (Harper, 1993; LaChapelle, 1962).

3.1. Photograph acquisition

Many of the NAGAP photographs were acquired with near vertical (nadir) orientation, often with ~60% forelap to provide stereo coverage. Photographs were captured on 9 × 9 inch (22.86 × 22.86 cm) panchromatic film, using metric cameras with ~6 inch (151.283–151.674 mm) focal length. Cameras were mounted inside the fuselage of a fixed-wing aircraft with viewing port and manually triggered by an operator (Austin Post, personal communication, October 2011). The NAGAP photo collection flight lines typically followed the glacier centerline, starting over proglacial terrain and ending near or above the glacier Equilibrium-Line Altitude (ELA) (Figs. S1 and S2). The photographs generally offer continuous spatial coverage over the entire glacier for smaller, lower elevation sites (e.g., South Cascade Glacier), while larger, high-elevation sites, such as Mount Baker or Mount Rainier, have some gaps in photo coverage over the summit.

3.2. Image scanning

Between 2014 and 2017, (Nolan et al., 2017) scanned ~20,000 of the total ~100,000 NAGAP film negatives, giving priority to nadir panchromatic photographs. Each film was scanned at 20-micron resolution using two Vexcel Ultrascan 5000 Model US5000 flatbed scanners with a precision of ± 2 μm (Gruber and Leberl, 2001). Scanned images were stored as 8-bit grayscale TIF files with dimensions of approximately 12,000 × 12,000 pixels. The ground sample distance (GSD) of the 20-micron scans ranges from approximately 0.1 to 1.0 m for typical flight altitudes of 500 to 3000 m above ground level. In August 2017, the ~20,000 scanned images and metadata were publicly released through the NSF Arctic Data Center (Nolan et al., 2017). Some of these same images are also available via the EE archive at 25-micron resolution. The USGS offers on-demand scanning services at 7- or 14-micron resolution, which may provide additional detail depending on the original film quality and resolving power. In practice, limitations associated with film grain, motion blur, and the film development process may limit resolving power, especially for older images (1940s–1960s), and the available 20-micron scans may already oversample the true image resolution. The remaining ~80,000 NAGAP photographs, including some oblique and color photographs, are currently archived at the University of Alaska Fairbanks.

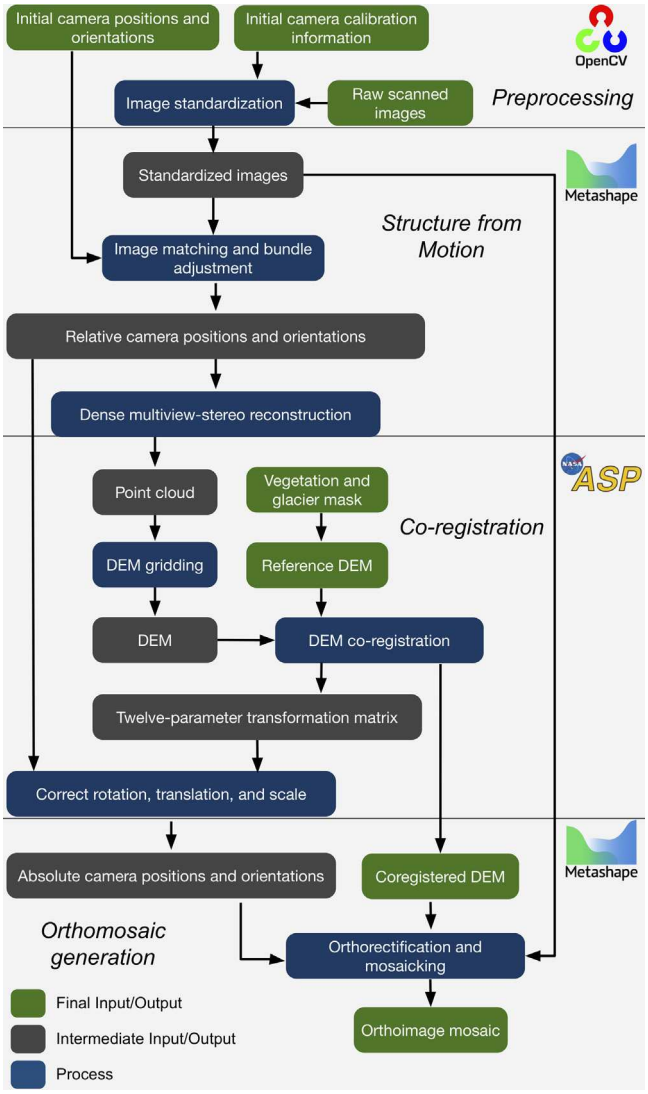

Fig. 2. Historical Structure from Motion (HSfM) automated workflow for the *single date - single block* mode.

3.3. Camera position metadata

Camera longitude and latitude positions for the scanned NAGAP images were manually estimated using basemaps in Google Earth (Nolan et al., 2017). Altitude information (assumed to be feet above home airport elevation or mean sea level), when available, was read from the photo frame and recorded in the NAGAP metadata files. The ADC archive includes metadata files in comma-separated value (csv) format containing acquisition date, longitude, latitude, and in some cases, altitude. Additional metadata files in keyhole markup language (kml) format are available for a subset of photographs from the 1970s, 1980s and 1990s, with fields for image file name, longitude, latitude, and altitude as orthometric height above the EGM96 geoid (Nolan et al., 2017). We use the positions in the kml files where available, and supplement with the csv files or EE metadata if necessary. The EE archive metadata typically includes fields for the camera center longitude, latitude, and altitude, along with estimated image footprint corner longitude and latitude coordinates. We convert all vertical altitude or elevation values to height above the WGS84 ellipsoid before processing, and evaluate the accuracy of available camera position metadata in Section 4.7.5.

3.4. Camera calibration

Camera calibration reports contain information that can be used to recover the interior orientation, including focal length, distances between fiducial markers, lens distortion, and the location of the optical center on the film (principal point of autocollimation) (Gruber and Leberl, 2001; Tayman, 1984). The cameras used for the NAGAP program were regularly calibrated (Tayman, 1984), with some calibration reports available via EE or the U.S. Department of Agriculture Remote Sensing Applications Center (USDA RSAC) database (USDA Forest Service, 2018). Calibration reports for a given film roll can be retrieved from the USDA RSAC database using the lens serial number, focal length, and acquisition date, all of which can usually be manually extracted from the analog metadata available on photo frame border. Alternatively, we implemented support for automated EE API queries to retrieve the corresponding camera calibration report under the same object identifier as the NAGAP image, if available.

3.5. Reference DEM data

We prepared high-quality reference DEM data for processing and evaluation of output products at both study sites. For the Mount Baker site, we generated a high-resolution 1 m reference digital surface model (DSM) from airborne lidar data collected for the USGS 3DEP program by Quantum Spatial on August 27, 2015, using a Leica ALS80 scanner (Ramsey, 2016) (Fig. S3). The lidar point cloud data were obtained from the Washington Department of Natural Resources (WADNR) lidar portal. The reported vertical accuracy (RMSE) of the lidar point cloud is 0.04 m (Ramsey, 2016).

For the South Cascade Glacier site, we prepared a high-resolution (2 m) composite reference DEM to use for co-registration and accuracy assessment. The composite DEM was generated from five summer Maxar WorldView (WV) stereo images acquired between 2013 and 2015 using an automated processing workflow (Shean et al., 2016). The five DEMs were co-registered and a weighted average composite was created, which offers superior absolute accuracy and improved coverage over steep slopes compared to any individual WV DEM. For analysis of change over glacier surfaces, we use a single "snapshot" WV reference DEM from October 14, 2015 (McNeil et al., 2022a) and glacier outlines for 1970, 1979, 1986, and 1992 provided by the USGS (McNeil et al., 2022a; O'Neel et al., 2019).

In addition to these high-resolution reference DEM products, we also evaluated the Copernicus DEM (Section 4.7.3), a 30 m DEM product ("GLO-30") with global coverage (European Space Agency, 2022). The Copernicus DEM has a reported absolute horizontal accuracy of < 6 m (CE90) and absolute vertical accuracy of < 4 m (LE90) (European Space Agency, 2020; NSDDA, 1998).

All reference DEMs were reprojected to UTM Zone 10N (EPSG: 32610) and adjusted to provide elevation values in meters above the WGS84 ellipsoid.

4. Methods

We developed an automated pipeline — Historical Structure from Motion (HSfM) — which enables rapid generation of DEMs and orthomosaics from archives of aerial imagery. The core workflow in the HSfM method involves the following steps: 1) image preprocessing, 2) SfM reconstruction, 3) co-registration with modern reference DEM data, and 4) orthomosaic generation using refined camera extrinsics (Fig. 2). Image preprocessing is performed with OpenCV v4.4 (Bradski and Kaehler, 2008), SfM reconstruction and orthomosaic generation with Agisoft Metashape v1.6.0 (AgiSoft Metashape, 2021), and co-registration with the NASA Ames Stereo Pipeline (ASP) v2.6.1 (Beyer et al., 2019, 2018; Shean et al., 2016) followed by Nuth and Kääb (2011) coregistration implemented in the *demcoreg* Python package (Shean et al., 2021). The HSfM package includes Python function

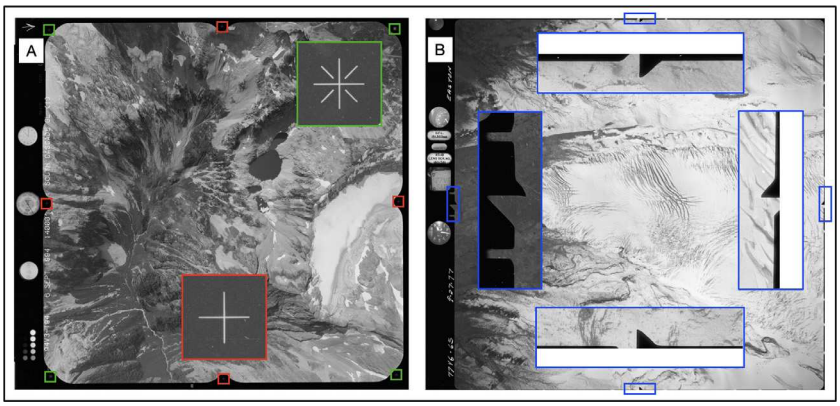

Fig. 3. Examples of fiducial and pseudofiducial markers in North American Glacier Aerial Photography (NAGAP) images. **A)** Sample NAGAP image acquired on 1994-09-06 over the South Cascade Glacier site (NAGAP_94V3_211) with insets showing detail of mid-side (red) and corner (green) fiducial markers. **B)** Sample NAGAP image acquired on 1977-09-27 over the Easton Glacier site (NAGAP_77V6_065) with insets showing detail of mid-side pseudofiducial markers (blue). Portions of the top, right, and bottom markers are missing due to problematic cropping during scanning, and left marker tip is slightly bent. See Fig. S4 for additional NAGAP fiducial and pseudofiducial templates available for marker detection.

libraries, scripts, and notebooks that provide a high-level interface for the full processing workflow. All code is open-source and publicly available (Knuth et al., 2021a), with the exception of Agisoft Metashape.

4.1. Historical image and metadata preprocessing

We developed the Historical Image Pre-Processing (HIPP) library (Knuth et al., 2021b) for scanned historical image retrieval and standardization. The HIPP package includes utilities to search the ADC and EE archives using spatial bounds, temporal bounds and keywords. The returned images are grouped into "image sets" by camera and acquisition date, then automatically downloaded and standardized in parallel. An image set is defined as all images that were acquired with the same camera on the same date. Image data can be loaded from disk or directly from image servers (e.g., ADC arcticdata.io or EE earthexplorer.usgs.gov) for remote data retrieval and in-memory processing.

4.1.1. Image standardization

After retrieval, the scanned historical images must be standardized to 1) remove any distortion introduced during film storage, handling, and/or scanning, 2) enhance contrast, and 3) remove the photo frame. The HIPP package offers three approaches to standardize images, depending on the availability of camera calibration reports and visibility of fiducial markers in the scanned images: 1) affine transformation to match calibration data, 2) affine transformation to match mean marker positions, and 3) pseudofiducial marker detection without affine transformation.

Fiducial marker detection. Fiducial markers are used to define the interior geometry of a film camera. Generally, four *mid-side* fiducial markers are located at the center of each photo frame side, along with four *corner* fiducial markers in the photo frame corners (Fig. 3). By precisely measuring the position of each marker in the scanned images, the apparent dimensions of the scanned image (pixels) can be adjusted to match the known interior dimensions of the calibrated camera (mm).

HIPP uses fiducial marker templates to automate fiducial marker detection for all images in a set. The software includes a number of predefined fiducial marker templates for the NAGAP archive (Fig. S4), and interactive tools that can be used to generate representative templates for scanned images from other sources (e.g., EE archive). When processing images from other sources, templates for one corner and one mid-side marker need to be created for each camera/lens combination. These templates are created from a user-selected sample image that is 1) representative of all other images from the same camera/lens combination and 2) does not suffer from overexposure, dust, scratches, or other issues that obscure the marker. The user selects the center position of representative fiducial markers in the sample image and HIPP generates standardized templates at both the native resolution of the scanned image and an upsampled resolution (8×) for improved subpixel precision during subsequent template matching.

Once fiducial marker templates are created, HIPP determines the sub-pixel fiducial marker coordinates for all images in the set using the OpenCV *matchTemplate* algorithm with search windows defined for the corresponding fiducial marker locations in each image. To ensure match quality, the *matchTemplate* score (range 0–1) for each fiducial marker must be within a predefined threshold of the mean score for all matches at the same fiducial marker location across all images in the set, otherwise the detected position is discarded. A default threshold of 0.01 was used in this study as it robustly excluded spurious matches for a large sample of the NAGAP images. HIPP produces quality control plots showing the subpixel position of each template match, the distance between fiducial markers, and orthogonality metrics for opposing fiducial markers. These plots can be used to evaluate the quality of fiducial marker detection and adjust the threshold if necessary.

Affine transformation to match calibration data. When camera calibration reports are available, HIPP computes a 2D affine transformation between the observed fiducial marker positions for each scanned image and the known fiducial marker positions from the camera calibration report. For this study, the marker positions were manually extracted from the available calibration reports, but the HIPP package includes experimental utilities for optical character recognition (OCR) to facilitate bulk extraction. The affine transformation is used to warp each scanned image to restore a uniform calibrated interior geometry for the image set, as well as compute an updated pixel pitch.

Affine transformation to match mean marker positions. For cases with high-quality fiducial marker matches but no available calibration information, we compute the average position of each fiducial marker (relative to the inferred principal point) across the full image set, and use these average positions to define the affine transform needed to standardize each image.

Pseudofiducial marker detection without affine transformation. The most common case for the NAGAP archive involves no available calibration information and imprecise, unusable, and/or missing fiducial markers in the scanned images (Fig. S4). In these cases, we use distinct shapes (e.g., notch patterns) on the photo frame as "pseudofiducial markers". The pseudofiducial markers are centered on the four sides of the photo frame for the NAGAP cameras. Their locations can be used to define the precise frame midpoint positions, which are then used to estimate

the principal point location (Fig. 3). The HIPP software includes predefined pseudofiducial marker templates for all images in the NAGAP ADC collection (Fig. S4), and utilities to define custom pseudofiducial templates when processing images without fiducial markers from other collections.

The pseudofiducial marker workflow uses the same template matching approach described above for the fiducial markers, only without subpixel refinement due to the inferior match precision. In addition, the frame appears physically damaged in some image sets and improperly cropped during scanning in others (e.g., Fig. 3B), which can introduce errors in the resulting pseudofiducial matches. To accommodate partially visible pseudofiducials, we pad the scanned image with a black border (DN value of 0, default width of 250 px) which provides a larger search window for template matching. To ensure match quality, the matched position of each pseudofiducial marker must be within a threshold of 50 pixels of the mean position for all matches at the same marker location across all images in the set, or it is discarded. The default value was determined by examining quality control plots produced by HIPP and identifying an appropriate threshold to exclude spurious matches for a large sample of NAGAP images.

4.1.2. Principal point determination and frame removal

The principal point is computed as the center point between all detected markers that passed quality control checks. A minimum of two markers is necessary to estimate the principal point, while including additional markers provides redundancy that will improve accuracy. If the principal point cannot be determined automatically from the marker template matching, the HIPP *pick_fiducials* utility can be used to manually select mid-side or corner points to determine the principal point.

After the principal point position is determined, each image is cropped to a fixed square area around the principal point to remove the photo frame. The output image dimensions are set based on the minimum distance between the fiducial or pseudofiducial markers and the principal point minus an offset that removes the image frame from all images in the set. At this stage, we consider the image set to be standardized with common interior orientation, and ready for photogrammetric processing.

4.1.3. Image enhancement

HIPP includes several options to improve cropped image contrast as a final preprocessing step. In this study we used the Contrast Limited Histogram Equalization (CLAHE) algorithm implemented in OpenCV (*createCLAHE* with a *clipLimit=2* and *tileGridSize=(8,8)*), followed by a linear histogram stretch with minimum and maximum values set to the 0.1 and 99.9th percentile of cropped image values. This image enhancement improved feature matching for under- or over-exposed surfaces (e.g., shadows, snow) and provided more consistent orthoimage mosaics.

4.1.4. Camera position estimation

As described in Section 3.3, we extract camera longitude, latitude, and altitude (converted to geodetic height) from available metadata whenever possible. For photographs with horizontal position data but no altitude or elevation information, HIPP samples the approximate surface elevation beneath each initial horizontal camera position using the USGS elevation data API, which queries the $\frac{1}{3}$ arc-second National Elevation Dataset (NED) (U.S. Geological Survey, 2019). HIPP then converts the NED orthometric heights (above the NAVD88 vertical datum) to geodetic height above the WGS84 ellipsoid. Finally, we assume a flight altitude of 1,000 m above this surface elevation and determine the corresponding camera position elevation above the ellipsoid. The estimated flight altitude of 1,000 m above ground level was determined by iteratively examining the true positions for a large sample of NAGAP images processed with HSfM. Since the NAGAP photos were generally acquired at a constant elevation (above vertical datum) for each site rather than a constant altitude above ground, HIPP determines the maximum estimated camera elevation for the image set and assign this value to any cameras in the set lacking elevation information.

4.2. Structure from Motion

The HSfM method supports three processing modes for a range of potential input image configurations: (1) *single date - single block*, (2) *single date - multiple blocks*, (3) *multiple dates - multiple blocks*. Each mode builds upon the functionality provided by the previous mode. As mentioned in Section 1, we define a block as three or more overlapping images acquired on the same acquisition date with the same camera. Multiple separate blocks may exist for the same image set (e.g., blocks for non-overlapping flightlines covering adjacent glaciers). When individual blocks from multiple acquisition dates (and potentially multiple cameras) overlap at the same site, these form a single "multi-temporal block" which we processed with the *multiple dates - multiple blocks* mode. These multi-temporal blocks are used to improve intrinsic and extrinsic parameters for all cameras prior to processing each block with the *single date - single block* mode. We now briefly describe the critical processing steps in the *single date - single block* workflow (Fig. 2), which forms the core of all processing modes.

4.2.1. Image matching

The HSfM workflow begins by identifying sparse feature matches (tie points) between all possible standardized image combinations in the block using the Metashape *matchPhotos* function with *high* matching accuracy (native image resolution). See Table S1 for additional details on feature matching settings.

4.2.2. Bundle adjustment

Bundle adjustment optimization relies on the fundamental collinearity equations to optimize interior and/or exterior camera orientation parameters using these feature matches (e.g., Fraser, 1997; Ackermann, 1981). The HSfM workflow performs a bundle adjustment optimization (Metashape *alignCameras* function) using the sparse feature matches and known values for focal length, pixel pitch, and initial camera position. We assume an initial accuracy of 1,000 m for the camera position estimates described in Sections 3.3 and 4.1.4. While estimates for flightline heading are available for some photos, we assume an initial nadir orientation and north heading for all cameras (pitch, roll, and yaw of 0°) and assign estimated accuracy of 10° for pitch and roll, and 180° for yaw. We use the Metashape default Brown–Conrady lens distortion model (Brown, 1971), and solve for a subset of intrinsic parameters during bundle adjustment, including focal length (F), principal point coordinates (Cx, Cy), as well as radial (K1, K2, K3) and tangential (P1, P2) lens distortion coefficients.

4.2.3. Dense multi-view stereo reconstruction

Once interior and exterior camera parameters are refined, HSfM performs dense multi-view stereo reconstruction. This process begins with dense image correlation for pairs of overlapping images using the Metashape *buildDepthMaps* function with *high* quality (downscale factor = 2) and *AggressiveFiltering* to remove outliers (Table S1). Finally, HSfM uses the *buildDenseCloud* Metashape multi-view triangulation function to generate a dense point cloud, which is exported in LAZ format.

4.3. DEM co-registration

The output historical DEM will have absolute position, rotation, and potentially scaling error, due to inaccurate initial camera positions and bundle adjustment without GCPs. To correct these issues, HSfM includes a multi-stage co-registration process to align the initial DEM to a modern reference DEM.

4.3.1. DEM gridding

The HSfM workflow generates a DEM from the dense point cloud using the ASP *point2dem* utility, which uses an Inverse Distance Weighting (IDW) algorithm with 1-grid-cell search radius to improve prediction of true elevation values for a given cell, while avoiding interpolation artifacts over large data gaps. We conservatively set the output DEM grid cell size using a multiplicative factor of 2–3 times the native image GSD to reduce noise/artifacts and data volume, while preserving sufficient detail for geodetic analysis. The native GSD values for the NAGAP images generally range from 0.1–0.3 m, and we use a standard posting of 1 m for all DEMs for our study sites.

4.3.2. Stable surface mask preparation

Successful co-registration of historical DEMs with evolving surfaces (e.g., glaciers) requires careful identification of nearby control surfaces that are stable and static over time. Fortunately, most of the NAGAP datasets were acquired over high alpine sites above the treeline with a good distribution of stable, exposed surfaces to use for co-registration. We assumed that all non-forested and non-glacierized surfaces in our study areas were stable for the time period between the historical photo acquisition and the modern reference DEM acquisition. While irrelevant for local relative elevation change analysis following co-registration, this assumption is supported by estimated regional erosion (< 0.3 mm/yr) (Reiners et al., 2003) and vertical deformation (< 2 mm/yr) (Lau et al., 2020) rates for Washington State, which will offset each other, and are negligible compared to our DEM elevation uncertainty and elevation change rate uncertainty.

We use the *dem_mask.py* utility implemented in the *demcoreg* Python package (Shean et al., 2021) to mask the reference DEM over glacier polygons from the RGI v6.0 (RGI Consortium, 2017) and forest land-cover classes (41–43) from the 30-m National Land Cover Dataset (NLCD) (Dewitz, 2016). The user can also manually define polygons over stable control surfaces and apply the inverse to mask the reference DEM for use during co-registration.

4.3.3. Iterative Closest Point (ICP) co-registration

The first stage of our multi-stage co-registration procedure uses a three-step Iterative Closest Point (ICP) alignment routine implemented in the HSfM Python function *pc_align_p2p_sp2p*, which wraps the ASP *pc_align* command-line tool (Beyer et al., 2018; Shean et al., 2016). The *pc_align* tool provides various implementations of the ICP algorithm (Pomerleau et al., 2013) along with robust outlier filtering and other configurable parameters, including the maximum expected offset between the two input point clouds. The first ICP co-registration step uses the *pc_align point-to-plane* ICP method to compute a rigid body translation and rotation needed that minimizes the large initial offset between the historical and reference DEM. The second step refines the resulting transform by correcting for scale, along with residual translation and rotation offsets, using the *similarity-point-to-point* ICP method. The final step further improves this transform, again using the *similarity-point-to-point* method, but with a masked version of the reference DEM and tighter constraints to refine the final transformation over a subset of exposed, stable surfaces. The maximum expected three dimensional displacement decreases for each of the three steps in the HSfM co-registration process as alignment improves. These values are configurable, with default values set to 2,000, 1,000, and 100 m, respectively, based on observed offsets for a large sample of NAGAP DEMs.

4.3.4. Nuth and Kääb (2011) co-registration

The second stage of our multi-stage co-registration uses the aspect- and slope-dependent method developed by Nuth and Kääb (2011) and implemented in the *demcoreg dem_align.py* utility. In addition to the reference DEM masking described above, *dem_align.py* applies a filter to mask surfaces in the reference DEM that have a slope less than 0.1° or greater than 40°, an absolute elevation offset greater than 100 m, or elevation offset that exceeds a robust 3-sigma outlier threshold (median ± 3 × NMAD). Generally, this second stage does not significantly change results following successful ICP co-registration, but it can provide sub-pixel co-registration accuracy in mountainous terrain and remove residual systematic offsets for non-optimal ICP runs.

4.4. Updating camera pose and orthomosaic generation

After successful co-registration, the final 12-parameter ICP transform and the 3D translation computed by *dem_align.py* are applied to the self-consistent camera positions from the previous bundle adjustment, which improves their absolute geolocation accuracy. These updated camera positions are imported into Metashape with tight constraints on geolocation accuracy (0.01 m) and the global *chunk.updateTransform* function is run to update camera orientations for the block. The updated camera models and the co-registered historical DEM are then used to orthorectify each input image. Finally, a seamless, blended orthomosaic is produced at the median native GSD of the input images, with geolocation accuracy similar to the co-registered DEM.

4.5. Single date - multiple blocks

The second mode supported by HSfM is *single date - multiple blocks*, which is used if there are multiple, non-overlapping blocks in the input image set. The first step in this mode attempts to identify all blocks by passing the standardized images and initial extrinsic/intrinsic metadata to the bundle adjustment and dense cloud reconstruction functions, as described in the *single date - single block* approach. Here, *matchPhotos* and *buildDenseCloud* are run with *medium* matching accuracy and quality parameters to expedite processing (Table S1). The resulting list of overlapping images with shared feature matches ("aligned images") is then parsed by HSfM to identify separate blocks. The block detection algorithm iterates on unaligned images until no additional blocks with 3 or more images can be identified. Each block is then processed using the *single date - single block* mode outlined above, and all co-registered DEMs and orthomosaics are mosaicked using the ASP *dem_mosaic* utility. The refined camera model for each block will be slightly different in this mode. While this was acceptable for our test cases, a better approach is to solve for a single set of intrinsic parameters shared by all blocks acquired with the same camera on a single date, as described in the *multiple dates - multiple blocks* mode below.

4.6. Multiple dates - multiple blocks (4D SfM)

The most sophisticated *multiple dates - multiple blocks* (*4D SfM*) mode supported by HSfM performs a large joint optimization ("multi-temporal bundle adjustment") using all available images across all acquisition dates for a given site. Previous studies refer to this general concept as *multi-epoch* or *multi-temporal* reconstruction (e.g., Feurer and Vinatier, 2018; Korpela, 2006; Vargo et al., 2017; Warrick et al., 2017; Zhang et al., 2021), but we prefer the shorthand *4D SfM* (four-dimensional Structure from Motion), for our approach, which includes both multi-temporal bundle adjustment and multi-temporal composite DEM co-registration. Our approach differs from previous implementations in that we: (1) do not use manually selected GCPs during optimization, (2) use a multi-stage co-registration method, and (3) perform a final optimization for each individual block via the *single date - single block* mode.

This mode begins with the block preparation routine outlined in the *single date - multiple blocks* mode, except that we include all images from all acquisition dates, with careful accounting of images that share the same camera intrinsic parameters. We use *medium* quality settings to expedite processing and produce a coarse-resolution (10 m) multi-date composite DEM that is co-registered to the reference DEM to improve initial geolocation accuracy for all optimized images. We then split the optimized images by acquisition date and run the *single date - multiple*

blocks routine to process all blocks collected on the same date with the shared and optimized intrinsic parameters held fixed. Any images that were not aligned during the large joint optimization are re-introduced in the appropriate block with original pose estimates, and each block is reprocessed with *high* quality settings to further refine all poses. Though more complex, this robust approach offers improved, self-consistent results for the full time series of historical images, and we used this approach to prepare all products for our study sites.

4.7. Accuracy assessment

We use the following criteria to evaluate the performance of the HSfM method and quality of output products.

4.7.1. Absolute and relative DEM accuracy

Co-registration is deemed successful if there is no systematic error in the final elevation difference map between the co-registered historical DEM and reference DEM. Specifically, the distribution of difference values over stable surfaces should appear random with near-zero bias and limited spread, providing a measure of absolute DEM accuracy in a geodetic coordinate reference system. To make this assessment of absolute accuracy, we compute the median and Normalized Median Absolute Deviation (NMAD) of residual elevation difference values over the masked, filtered, stable reference surfaces. These robust estimators of the residual bias and spread (Höhle and Höhle, 2009; Shean et al., 2016) are defined as:

$$m_{dz} = median(dz) \tag{1}$$

$$NMAD = 1.4826 \cdot median(|dz_i - m_{dz}|) \tag{2}$$

where dz denotes the full sample of per-pixel difference values between the historical and masked reference DEM surfaces after co-registration.

While the NMAD of these absolute difference values can be used to estimate relative accuracy (precision) of the historical DEM, it will potentially include relative errors introduced by the reference DEM. To avoid this issue, we assess the relative accuracy among the co-registered historical DEMs by computing the per-pixel standard deviation (SD) and NMAD of all elevation values in the historical DEM time-series stack. We only consider pixels over masked, filtered, stable surfaces with three or more observations in the DEM time-series stack. We then compute summary statistics for each site as the mean and median of the per-pixel SD and NMAD for all valid pixels.

4.7.2. Spatially autocorrelated error

To assess the spatial autocorrelation of residual error in the historical DEM products, we sampled empirical variograms for elevation difference maps between the historical and reference DEMs over stable surfaces. We used the *sample_empirical_variogram* function implemented in the xDEM package (xDEM contributors, 2021; Hugonnet et al., 2022), which relies on the SciKit-GStat Python package (Mälicke et al., 2022). For each difference map, we sample and average five unique variograms with sample size of 10,000. We then use weighted least-squares to fit a double-range nested spherical variogram model to the empirical variograms, which captures both the short and long-range spatial autocorrelation of the errors (Hugonnet et al., 2022). The long-range sill is reported as additional variance above the short-range sill, and the total variance of the nested model is the sum of the short and long-range sill components.

4.7.3. Reference DEM sensitivity testing

As outlined in the *single date - single block* description, the HSfM method uses reference DEM data to improve the geolocation accuracy of the historical DEMs. To evaluate the sensitivity of this approach to reference DEM source and resolution, we processed the Mount Baker historical DEMs using multiple reference DEMs, including the high resolution (1 m) product derived from airborne lidar and the medium resolution (1-arcsecond, 30 m) Copernicus DEM (Section 3.5). To assess the relative accuracy of time series stacks produced using each reference DEM, we computed the corresponding per-pixel standard deviation and NMAD of elevation values as described in Section 4.7.1.

4.7.4. Comparison of automated HSfM method with manual GCP method

An earlier historical image processing effort by the USGS produced historical DEMs from NAGAP images collected on 1970-09-20, 1979-10-06, 1986-09-05, and 1992-10-06 over South Cascade Glacier using Metashape with manually identified natural GCPs as control during bundle adjustment (McNeil et al., 2022a; O'Neel et al., 2019). We processed the same input images using the HSfM approach, which enabled a direct comparison of output products. We co-registered all DEMs to the same 2013–2015 composite WV reference DEM using the *dem_align.py* utility and then analyzed residuals and spatial autocorrelation over stable surfaces in the resulting elevation difference maps. We then prepared difference maps between DEMs generated using the HSfM and manual GCP method for each date, and analyzed residuals over the glacier surface.

4.7.5. Camera position accuracy

We use the final RMS reprojection error in pixels to evaluate the quality of the bundle adjustment optimization for each block. Where multiple blocks contribute to a final single-date DEM mosaic, we report the mean RMS reprojection error for all input blocks. We also evaluate the accuracy of the initial camera position estimates (Sections 3.3 and 4.1.4) by computing the translation magnitude for refined camera position after bundle adjustment and final camera position after co-registration.

4.7.6. Evaluation of image standardization approaches

We evaluated the HIPP standardization approaches for three image sets over the South Cascade Glacier site with well-defined fiducial markers and calibration reports (1994-09-06, 1996-09-10, 1997-09-23). We prepared a baseline interior orientation result for each image set using the precise fiducial marker and principal point determination approach, without subsequent affine transformation. We then examined differences between these baseline values and values for standardization using 1) affine transformation to match calibration data, 2) affine transformation to match mean marker positions, and 3) pseudofiducial marker detection without affine transformation approaches. Specifically, we considered the estimated error (standard deviation from diagonal terms in the intrinsic parameter covariance matrix) for the optimized focal length and principal point offset computed during bundle adjustment. We also considered the total block RMS reprojection error and NMAD over stable surfaces for output DEMs generated using each approach.

5. Results

We highlight results for three common spatial and temporal data configurations encountered when working with historical images. These include 1) a single block of images collected on a single date (Easton Glacier, 1977), 2) single blocks collected on multiple dates for the same site (South Cascade Glacier,1967–1997), and 3) multiple non-overlapping blocks collected on multiple dates (Mount Baker,1970–1992).

Each of these case studies was processed using the *multiple dates - multiple blocks* mode (4D SfM), which automatically handles the

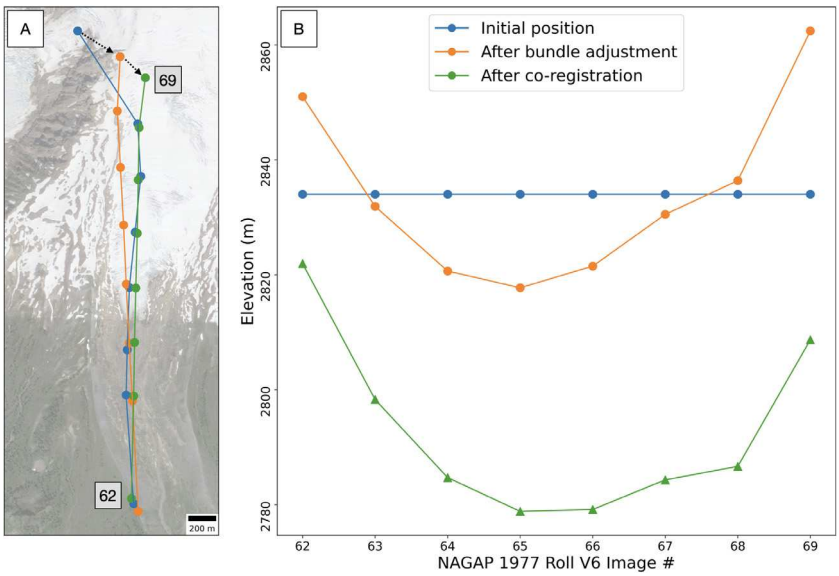

Fig. 4. Iterative camera position refinement for the 1977-09-27 North American Glacier Aerial Photography (NAGAP) images over Easton Glacier. (**A**) Google satellite basemap (imagery source date 2016-07-28) with initial camera positions (blue), positions after relative bundle adjustment (orange) and final positions after co-registration with the 2015 lidar reference digital surface model (green). Gray labels identify frame numbers for the first and last images in the flight line. Dashed arrows show translation vectors for image #69. (**B**) Corresponding camera elevation refinement in meters above the WGS84 ellipsoid.

respective spatial and temporal data configuration of the input images. For the Easton Glacier 1977 test case, HSfM immediately proceeded to the final *single date - single block* step, as no multi-temporal blocks could be formed for the input images from a single date. For the other two cases, the multi-temporal blocks were processed to refine shared intrinsic parameters prior to *single date - single block* processing.

5.1. Easton Glacier, 1977

We processed the eight available NAGAP images collected over Easton Glacier on 1977-09-27. As outlined in Section 4.1.4, camera altitudes were estimated at a constant 1000 m above the highest sampled ground elevation for the image set (Fig. 4). Camera positions were then iteratively refined, with final positions shifted by up to 600 m in the XY direction and 60 m in the Z direction following the bundle adjustment and co-registration procedures (Fig. 4).

The RMS reprojection error for the Easton Glacier block was 0.67 px (0.15 m) and the absolute accuracy over stable surfaces in reference DEM was 0.00 m, with NMAD of 0.32 m (Fig. 5D). The long-term elevation difference map (Fig. 5G) reveals large glacier surface elevation change signals between 1977 and 2015, with magnitude of over 50 meters in places. The difference map also shows meter-scale geomorphic change (Fig. 6) associated with lateral moraine mass wasting, stream erosion and moraine deposition following the glacier advance between 1956 and 1987 documented by Harper (1993).

5.2. South Cascade Glacier, 1967–1997

We successfully processed 243 images over the South Cascade Glacier site to produce 18 DEMs for 18 acquisition dates between 1967-09-21 and 1997-09-23 (Table S2, Fig. S2). The block per-pixel image count was typically 2–3 (Fig. S5), with additional overlap from multiple flight lines at different altitudes in some years. The mean RMS reprojection error for all blocks was 0.65 px (0.13 m) and the NMAD of difference values over stable surfaces ranged from 0.27 to 0.75 m (Table S2), with limited spatial autocorrelation (Fig. S9). The final historical DEMs have no bias and randomly distributed residuals relative to the reference DEM (Fig. 7D and S6). We do not observe a relationship between the historical DEM accuracy and the stable surface area used during co-registration. The relative accuracy (NMAD) of the historical DEM time series stack was 0.37 m over stable surfaces (Fig. 7C and E), while some portions of the thinning glacier surface (Fig. 8A) with high DEM count display elevation change variability of > 25 m over the 30 year period (Fig. 7C).

The historical DEM time series stack offers excellent temporal coverage (multiple DEMs per decade), especially over the lower glacier (Fig. 7B). Six years (1967, 1970, 1974, 1979, 1994, 1996) have nearly complete coverage of the glacier, four of which (1967, 1974, 1994, 1996) were not processed for previous reanalysis efforts (e.g., O'Neel et al., 2019; McNeil et al., 2022b).

5.2.1. Comparison with manual GCP method

The HSfM DEM products display improved absolute and relative accuracy compared to corresponding DEMs generated using the manual GCP method (Fig. 9). The DEMs generated using the manual GCP method displayed residual horizontal and vertical offsets of 1–2 m relative to the reference DEM, likely due to natural GCP position uncertainty. After co-registering all DEMs to the same reference DEM, the HSfM residuals over stable surfaces had reduced spread compared to those from the manual GCP method (Fig. 9). The DEMs prepared with the manual GCP method contain large spatially correlated errors (Fig. 9) and appear biased over glacier surfaces, with mean difference ranging from −1.61 to −0.22 m compared to the HSfM DEMs (Fig. S7). These errors will bias subsequent geodetic mass balance calculations, though the actual impact for South Cascade Glacier is relatively small considering the magnitude of the mass loss signal and long time interval.

5.2.2. Image standardization

Three NAGAP image sets for the South Cascade Glacier site (1994-09-06, 1996-09-10, 1997-09-23) had well-defined corner and mid-side fiducial markers with corresponding calibration reports, which provided an opportunity to investigate the impact of input image standardization approach on reconstruction error and output DEM quality (Section 4.7.6). We observe improvement for image standardization with affine transformation to match calibrated values for these test cases (Table S3), but overall results are similar for the three approaches.

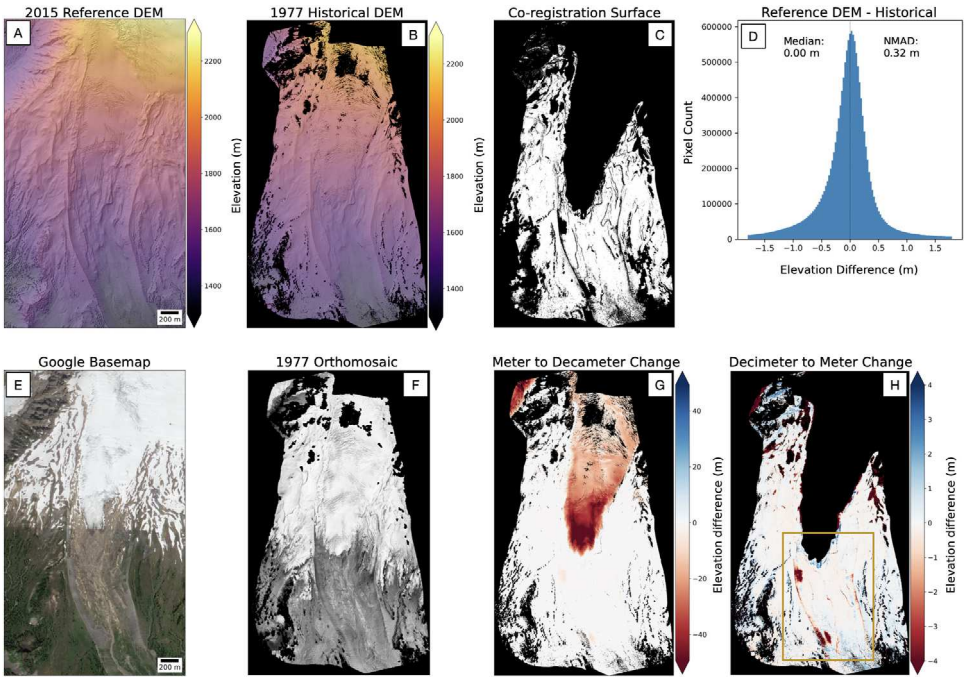

Fig. 5. Results for Historical Structure from Motion (HSfM) processing of the eight North American Glacier Aerial Photography (NAGAP) images collected on 1977-09-27 over Easton Glacier. (**A**) 2015-08-27 airborne lidar reference digital surface model (DSM) with 1 m posting. (**B**) 1977-09-27 historical NAGAP digital elevation model (DEM) with 1 m posting. (**C**) Mask defining stable surfaces used during co-registration. (**D**) Histogram of final elevation difference values over stable surfaces (reference DEM minus historical DEM). (**E**) Google satellite basemap (imagery source date 2016-07-28). (**F**) 1977-09-27 historical orthomosaic. (**G**) Final elevation difference map showing glacier thinning and retreat, including a portion of the adjacent Deming Glacier (RGI60-02.17738) to the northwest. (**H**) Same as in G, but with a stretched color ramp to bring out meter-scale elevation change, highlighting changes in forefield geomorphology. Gold outline shows extent of Fig. 6.

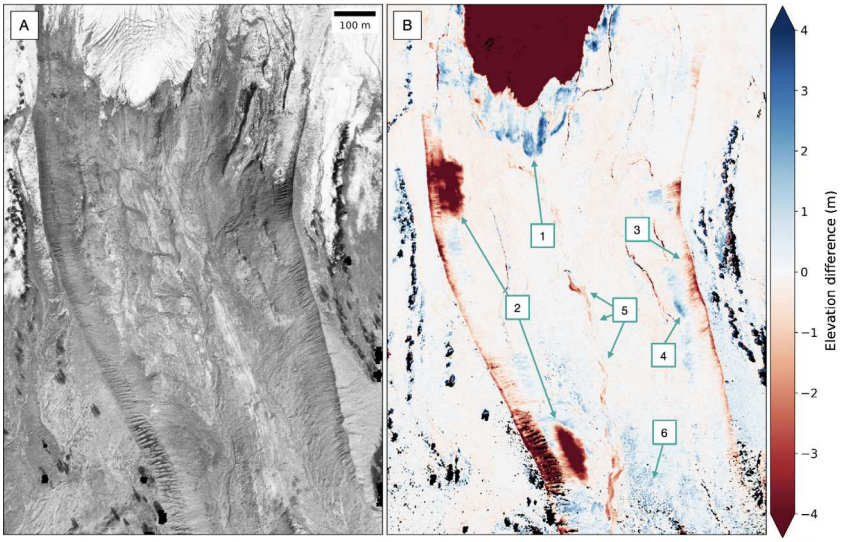

Fig. 6. (**A**) Detail of the 1977-09-27 North American Glacier Aerial Photography (NAGAP) orthomosaic over the Easton Glacier terminus and forefield. (**B**) Elevation difference between the 1977-09-27 NAGAP digital elevation model (DEM) and 2015-08-27 reference lidar digital surface model (DSM) (~38 years). Numbered labels show (**1**) moraine deposition, (**2**) large thinning signals at the base of the western little ice age moraine, (**3**) lateral moraine mass wasting, (**4**) debris deposition at the base of the eastern little ice age moraine, (**5**) stream erosion, and (**6**) vegetation growth.

As expected, larger image sets with more overlap had the best results for all three standardization approaches. Unfortunately, because these images include true corner and mid-side fiducial markers, the photo frame does not contain the same notched pseudofiducial markers present for the majority of the NAGAP archive images (Fig. S4), and these image standardization results are not directly transferable to other pseudofiducial types. The excellent absolute and relative accuracy observed for the full historical DEM time series at our study sites (Figs. 7 and 10) provides a better assessment of the generalized pseudofiducial marker standardization approach.

5.3. Mount Baker, 1970–1992

We successfully processed 261 images over the Mount Baker site, generating 21 DEMs for 10 dates between 1970-09-09 and 1992-09-18 (Table S4, Fig. S2). This is the most complex case study, with multiple non-overlapping blocks for some dates (e.g., 5 DEMs on 1977-09-27, Table S4), with DEMs that were individually co-registered prior to being combined into a final DEM mosaic for that date. During the initial multi-temporal bundle adjustment process, camera positions

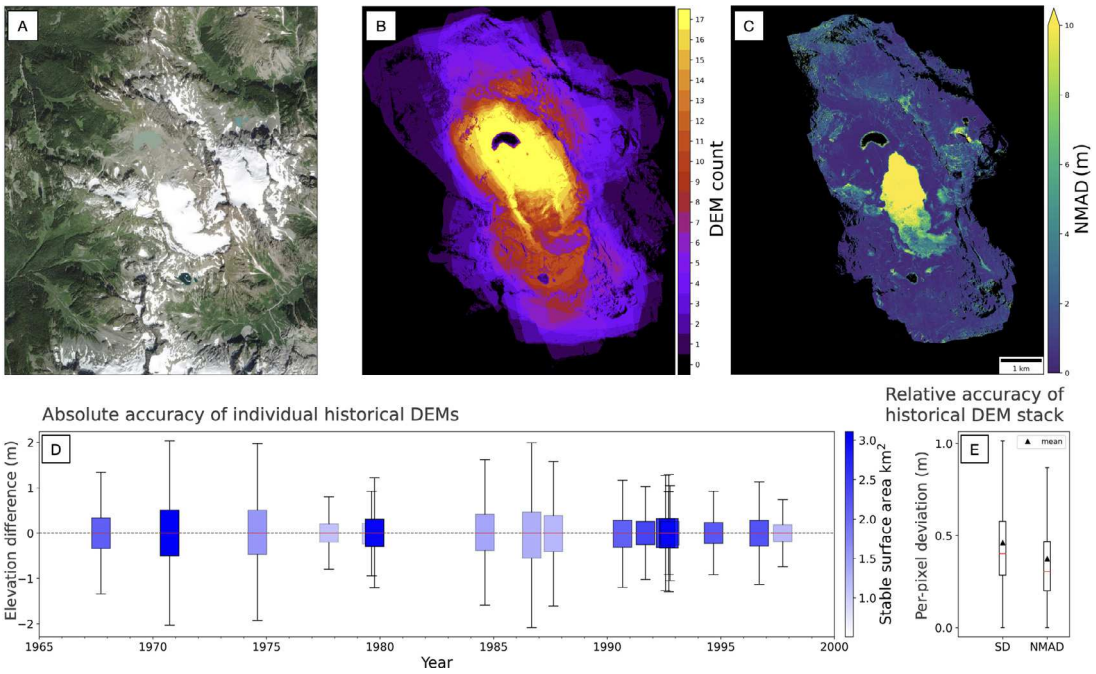

Fig. 7. (**A**) Google satellite basemap for the South Cascade Glacier site (imagery source date 2016-08-19). (**B**) Per-pixel digital elevation model (DEM) count and (**C**) per-pixel normalized median absolute deviation (NMAD) for all 18 historical DEMs in the 30-year time series stack (1967-09-21 to 1997-09-23). (**D**) Absolute accuracy metrics over stable surfaces for each historical DEM (historical DEM minus reference DEM), with box color representing stable surface area coverage. (**E**) Relative per-pixel accuracy metrics over stable surfaces for the full DEM time series stack, with mean standard deviation (SD) of 0.46 m and mean NMAD of 0.37 m. The NMAD box and whisker distribution summarizes of values shown in panel C. Boxes in (D) and (E) show interquartile range, with median values in orange and whiskers at 1.5 times the interquartile range.

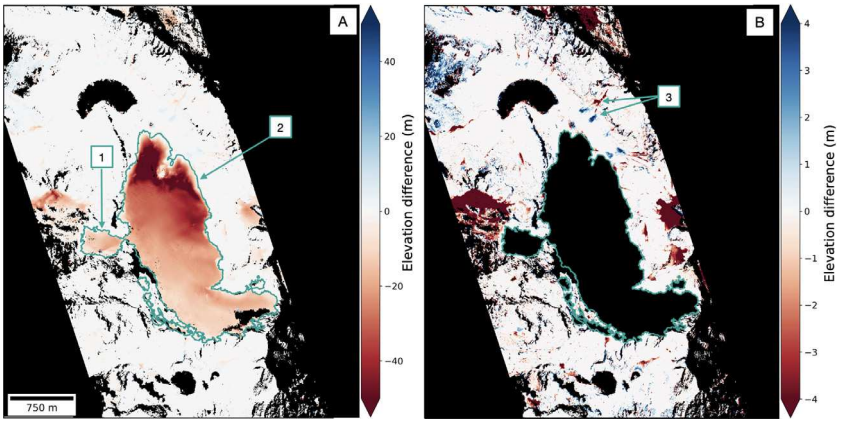

Fig. 8. (**A**) South Cascade Glacier elevation change between the 1994 historical digital elevation model (DEM) and the 2015 reference DEM. (**B**) Same as in A, but with glacier mask applied and stretched color ramp to show meter-scale change. Numbered labels show (**1**) glacier tributary near Post Peak, (**2**) 1992 glacier extent (McNeil et al., 2022a), and (**3**) negative elevation change attributed to talus mass wasting, with corresponding positive change over deposits. See Fig. S8 for additional detail of these features.

moved horizontally and vertically by up to 3600 m from their originally estimated positions. Co-registration of the multi-date low-resolution composite DEM further improved the absolute camera positions for some blocks by up to 450 m. Following these initial position adjustments, the final *single date - single block* processing step required final camera position refinement of less than 100 m.

The block per-pixel image count was generally 2–3, with overlap of > 10 images for some dates and locations (Fig. S10). The mean RMS reprojection error for individual blocks was 0.67 px (0.15 m) and the NMAD of difference values with respect to stable surfaces in the reference DEM ranged from 0.52 to 1.2 m (Table S4), with some larger DEM mosaics (e.g., 1979-10-06) displaying limited spatially correlated errors at longer lag distances (Figs. S11 and S12. The final historical DEMs have no bias and randomly distributed residuals relative to the reference DEM (Fig. 10D and S11). We observe a relationship between the historical DEM accuracy and the stable surface area used during co-registration, with larger DEMs (e.g., 1979-10-06, 143.3 km^2, Table S4) displaying higher spread of residuals over stable surfaces (Fig. 10D). The relative accuracy of elevation values in the historical time series stack was 0.68 m (NMAD) over stable surfaces (Fig. 10C and E), while thinning glacier surfaces display elevation change variability of > 10 m (Fig. 10C). The historical DEM time series stack shows the best temporal coverage for glaciers along the south and southeastern flank of Mount Baker with seven years of coverage over lower Easton Glacier (1970, 1974, 1977, 1979, 1990, 1991, 1992) (Fig. 10B–C).

5.3.1. *Intrinsic parameter refinement with 4D SfM*

We observed that refining intrinsics using the *4D SfM* process (Section 4.6) significantly reduced systematic "doming" error due to limited image overlap and poor acquisition geometry (e.g., James and Robson, 2014). A good example of this improvement can be seen for block 1 of

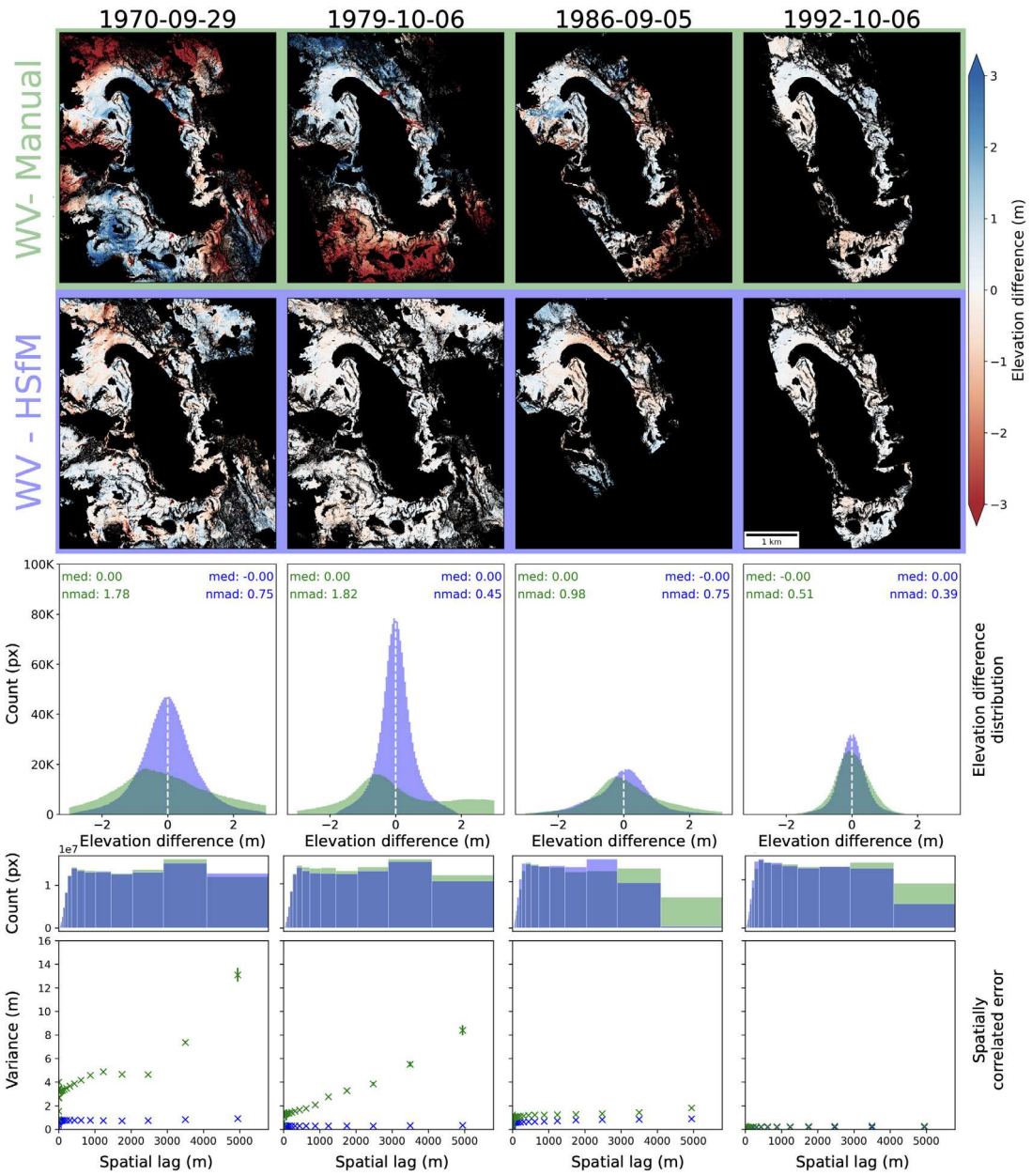

Fig. 9. Comparison between historical digital elevation models (DEMs) generated with Historical Structure from Motion (HSfM) and the manual ground control point (GCP) methods. Top row (green) shows difference maps over stable surfaces between the 2013–2015 WorldView (WV) stereo composite reference DEM and four DEMs produced with the manual GCP method. Second row (purple) shows corresponding difference maps for HSfM output. Third row shows residual distributions and final row shows empirical variograms for each difference map. The HSfM DEMs consistently show improved accuracy and precision compared to the DEMs from the manual GCP method.

the 1991-09-09 image set, which spans the southeastern flank of Mount Baker (Fig. S10). The historical DEM created using intrinsic parameters independently optimized for a single block shows systematic elevation-dependent error, with significant spatial autocorrelation at ~2–4 km lag distances (Fig. 11A). By comparison, the historical DEM created with the refined intrinsic parameters from *4D SfM* processing has near-zero difference over stable surfaces at all elevations and no significant long-range spatial autocorrelation error (Fig. 11B). This improvement stems from the additional constraints provided by overlapping image sets, even when large portions of the scene (e.g., glaciers) are changing over time.

5.3.2. Reference DEM sensitivity

The reference DEM sensitivity tests for the Mount Baker site showed that the HSfM method can produce accurate results using a global, low-resolution reference DEM. The relative per-pixel accuracy over stable surfaces for the full Mount Baker DEM time series stack was 0.68 m after co-registration with the 1 m lidar reference DEM and 1.26 m after co-registration with the 30 m Copernicus reference DEM (Fig. S3). The Copernicus DEM for the Mount Baker site contains apparent phase unwrapping artifacts over a portion of the southwestern flank (Fig. S3C), which increased errors during co-registration of smaller DEMs over these areas and led to higher, but still acceptable, errors in the DEM time-series stack.

5.4. Processing time

We ran preliminary benchmarking tests with the NAGAP images ($\sim 12,000 \times 12,000$ px) to provide a general sense of processing time for the full HIPP/HSfM workflow. These tests show that processing time scales linearly with the number of input images and co-registration area. Images were downloaded with a 1 Gigabit internet connection

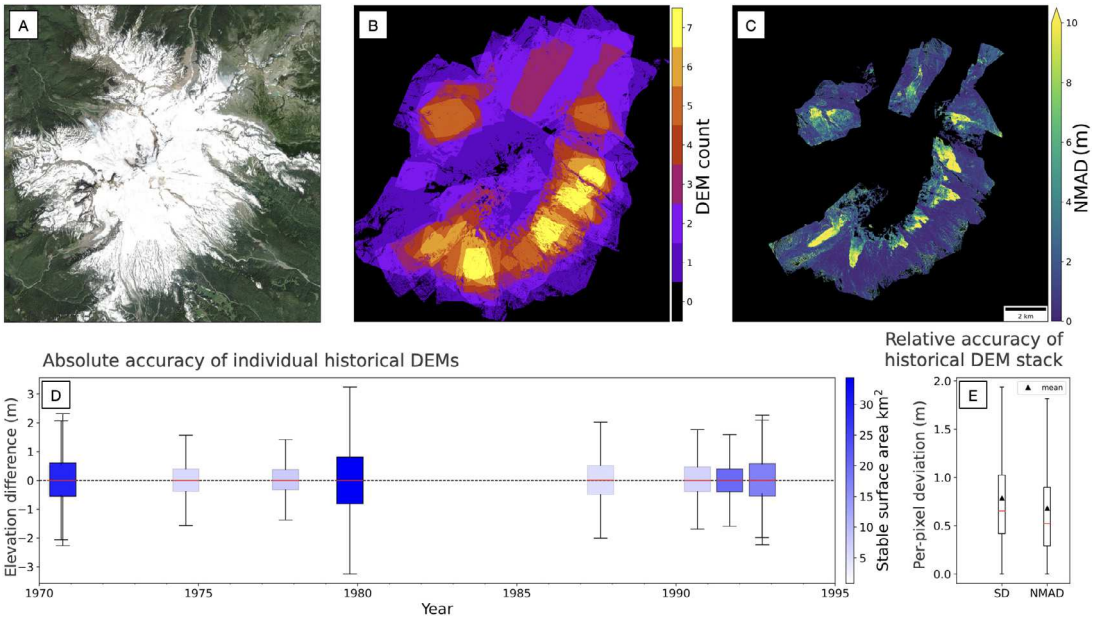

Fig. 10. (**A**) Google satellite basemap for the Mount Baker site (imagery source date 2016-07-28). (**B**) Per-pixel digital elevation model (DEM) count map and (**C**) per-pixel normalized median absolute deviation (NMAD) for all 10 historical DEM mosaics in the time series stack (1970-09-09 to 1992-09-18). (**D**) Absolute accuracy metrics for each historical DEM mosaic, with box color representing available stable surface area. (**E**) Relative per-pixel accuracy metrics over stable surfaces for the full DEM time series stack, with mean standard deviation (SD) of 0.78 m and mean NMAD of 0.68 m. The NMAD box and whisker distribution summarizes of values shown in panel C. Boxes in (D) and (E) show interquartile ranges, with median values in orange and whiskers at 1.5 times the interquartile range.

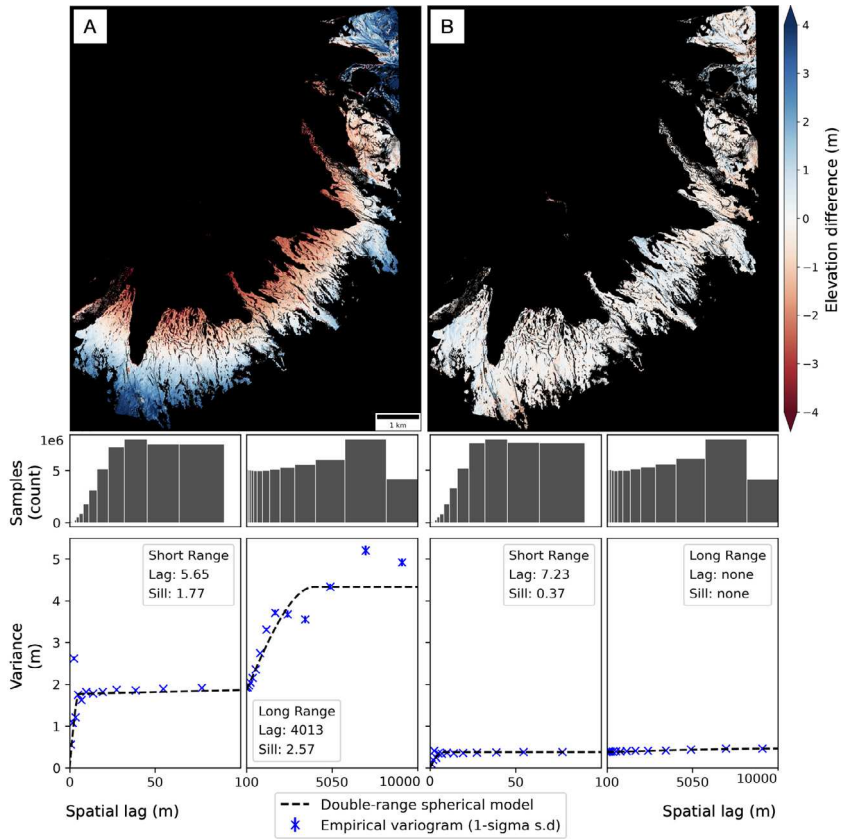

Fig. 11. Stable surface difference maps between the largest block (Fig. S10) of the 1991-09-09 historical digital elevation model (DEM) and the 2015-08-27 lidar reference DEM for the Mount Baker site processed using the (**A**) *single date - single block* approach, and (**B**) Four-dimensional Structure from Motion (4D SfM) approach. Bottom row shows corresponding empirical variograms and model results. The refined intrinsics from the 4D SfM optimization eliminate spatially correlated errors associated with "doming" artifacts present in A.

to a dedicated desktop workstation running Ubuntu Linux 18.04 with a 3.3 GHz 10-core processor, 64 GB RAM, and a 11 GB graphics processing unit with 3584 CUDA cores. The total processing time for all images and DEMs for the Mount Baker site was approximately 15 h. The total processing time for all images and DEMs for the South Cascade Glacier site was approximately 8 h. We observed average wall times of 5 s per image during preprocessing, 30 s per image during SfM reconstruction, and 12 s per km^2 during co-registration with a 1-m reference DEM.

6. Discussion

Our results show that the HSfM method produces accurate DEMs from scanned historical aerial photographs with limited metadata. Here we discuss some limitations of the input datasets and methodology, as well as broader implications and future directions for this work.

6.1. Image matching

Of the 597 total images processed over Mount Baker and South Cascade Glacier, 93 were excluded during initial block preparation (Section 4.6). There are many possible causes for this exclusion, including incorrect archive metadata (i.e., the images were not actually collected at Mount Baker or South Cascade), presence of clouds in images, or uncaught errors during image preprocessing and fiducial marker detection. Additionally, some images contain over- and under-exposed areas (e.g., snow and shadows), which is problematic for feature detection and brute-force matching. Future efforts to improve histogram equalization and noise reduction may help improve match point detection to overcome these issues. More sophisticated image overlap determination could eliminate the need for brute-force matching, including distance thresholds for image sets with sufficient initial camera position accuracy.

6.2. Availability and distribution of stable surfaces

In some cases, a historical image block is located over changing surfaces with limited or no stable surfaces (e.g., snow-covered mountain summits, glacier ice, or dense forest canopy). This is problematic for the co-registration steps, which require stable surfaces with adequate spatial distribution across the block. One potential workaround involves the traditional workflow with manual GCP selection for a limited subset of exposed stable point locations to constrain the bundle adjustment (e.g., Westoby et al., 2012). Alternatively, an initial co-registration with the unmasked reference DEM can remove gross geolocation errors, followed by a constrained ICP co-registration (e.g., a simple translation, without rotation or scaling parameters) using the small sample of stable surfaces. For cases with no stable surfaces, but accurate "before" and "after" DEMs, one can interpolate to prepare a "dynamic reference DEM" or "dynamic GCPs" for the historical image acquisition date based on observed elevation change (e.g., Shean et al., 2019).

6.3. Advantages of an automated processing approach

For some limited historical image sets over smaller sites, a trained operator can manually preprocess images and select GCPs for SfM processing to obtain results with sufficient accuracy for their application in a matter of hours to days. This approach, however, does not scale for larger processing efforts, and our results show that the automated HSfM method rapidly produced (Section 5.4) more DEMs with higher accuracy and less systematic error than the manual GCP approach (Section 5.2.1). Some of the observed systematic error in the DEMs from the manual GCP approach may be a result of different operators and evolving processing strategies. An automated and standardized processing approach like HSfM can eliminate most human error, with improved data provenance and reproducibility.

6.4. Toward fully automated historical image processing

While most of the HSfM method is automated, processing some historical datasets may require one-time manual preparation steps. In the following sections, we summarize these steps, suggest options for automation, and offer additional guidance for new HSfM users.

6.4.1. Reference DEM preparation

The HSfM method requires a reference DEM with sufficient coverage over stable surfaces within the footprint of the historical image block. While a high-resolution DEM from airborne lidar or satellite/airborne stereo is preferred, a global DEM with medium resolution (e.g., Copernicus 30 m DEM) can also be used. The results presented in Section 5.3.2 show that the historical DEM stack prepared using the 30 m Copernicus reference DEM can be used for analysis of many long-term geodetic change signals, even though overall relative accuracy is less than the stack prepared using the 1 m lidar reference DEM. Regardless of the reference DEM used, the HSfM method offers a notable improvement over workflows involving manual GCPs selected from a low-resolution reference DEM, which can introduce horizontal and vertical error. Future enhancements to the HSfM pipeline will provide automated reference DEM preparation (e.g., OpenTopography Global DEM API) when a reference DEM is not provided by the user.

6.4.2. Interior orientation preparation

The HIPP software can automatically extract essential intrinsic parameters for focal length and pixel pitch (scanning resolution) from the EE archive metadata. For the NAGAP images in the ADC archive, HIPP maps relevant intrinsic parameters to all image file names using metadata tables bundled with the software. For other archives or scanned historical image sets, the user can provide these parameters to the high-level processing routines.

The HIPP software requires fiducial or pseudofiducial marker templates to define interior orientation, standardize, and crop the scanned images. We prepared these templates for most NAGAP images in the ADC archive (Fig. S4) and included them in the HIPP software, along with metadata tables that map input image file names to the corresponding templates. Additional databases of marker templates and metadata mapping can be prepared for other historical image archives (e.g., EE) to enable fully automated preprocessing. As described in Section 4.1.1, the interactive HIPP utilities can be used to create a new representative set of templates for historical images from an unknown source. Future approaches using convolutional neural networks trained on large samples of fiducial markers from many historical image sources may eliminate the need for manual template preparation (e.g., Nielsen, 2017).

The image standardization results confirm that the scanned NAGAP ADC images do not have systematic issues (e.g., rotation or stretching) requiring more complex standardization. This result is a testament to the quality of NAGAP film preservation and scanning efforts (Nolan et al., 2017; Gruber and Leberl, 2001; Sevara, 2016). While the best possible accuracy was obtained using precise fiducial marker matching and an affine transformation to match calibrated interior orientation (Section 5.2.2), the alternative standardization approaches involving affine transformation to mean marker positions and pseudofiducial markers without affine transformation also provided high-quality results. The latter approaches enable fully-automated preprocessing of NAGAP images from the full ADC archive, as the pseudofiducial marker approach is the only option for a majority of these images.

6.4.3. Exterior orientation preparation

The HSfM method requires initial camera position estimates in a geodetic coordinate system. Fortunately, historical image archives that support spatial queries (e.g., EE) inherently contain these initial position estimates. For other image sources, initial camera position estimates can be manually extracted from photo flight line index maps

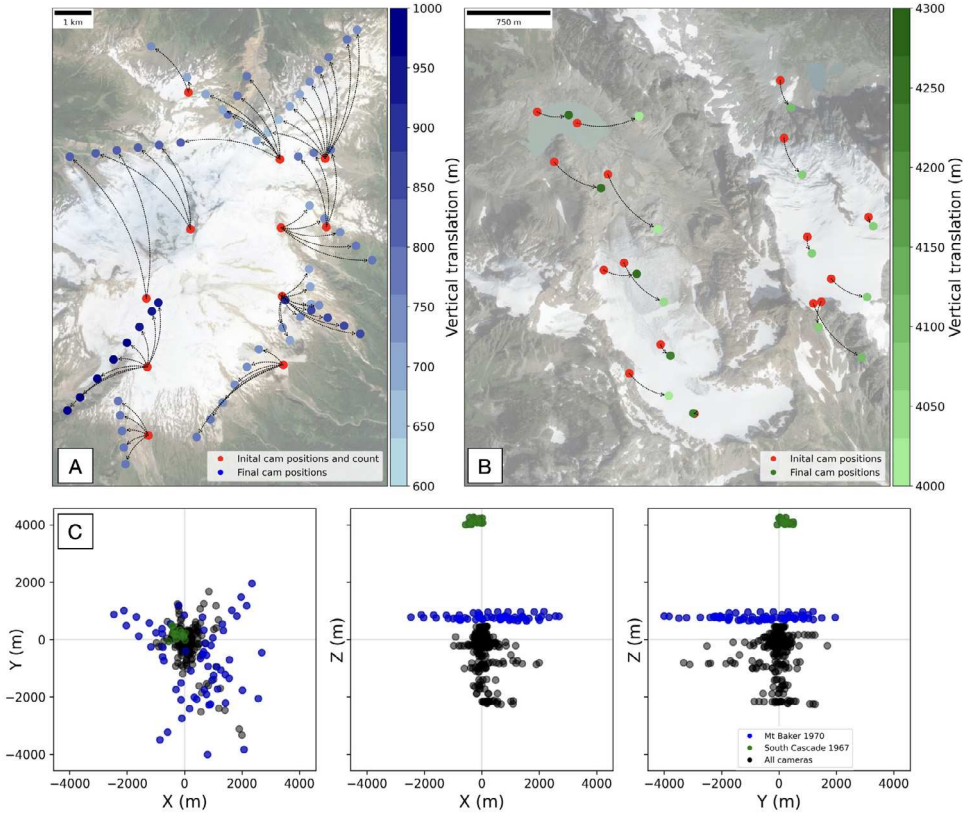

Fig. 12. Detailed examples of camera position correction during Historical Structure from Motion (HSfM) processing. (**A**) Images collected on 1970-09-29 over the Mount Baker site. The red dots show images with shared initial positions in the NSF Arctic Data Center (ADC) metadata. The blue dots show the final horizontal positions with color representing vertical translation magnitude. (**B**) Images collected on 1967-09-21 over the South Cascade Glacier site with final positions in green. A vertical translation of up to 4.3 km was required to correct the EarthExplorer (EE) archive metadata altitude of 7,500 m (24,600 ft) for these images. (**C**) Translation components required to correct camera positions for all images processed in this study (black), including the blocks shown in panels A (blue) and B (green).

sometimes archived with the images or by analyzing features in the images and manually estimating camera positions using a reference basemap, as was done for the ADC NAGAP metadata (Nolan et al., 2017).

The accuracy of the camera positions is critical, as this ultimately determines the accuracy of the triangulated points and resulting DEM. Unfortunately, estimates for historical image positions can contain horizontal and vertical errors of several kilometers. We used a large set of NAGAP and EE test cases with variable initial position accuracy to develop the HSfM workflow with appropriate weights and tolerances. Our results show that the HSfM routines can correct initial camera position error up to 5 km (Fig. 12). The HSfM approach was even able to process image sets with a single shared horizontal position and no altitude information in the archive metadata, such as 1970-09-29 images collected over Mount Baker (Fig. 12). This suggests that for smaller sites and image sets with adequate overlap, the user may only need to provide a single approximate estimate of latitude and longitude to effectively seed the bundle adjustment and co-registration steps. This approach can potentially eliminate the tedious process of manually estimating accurate position metadata during bulk scanning efforts.

An alternative approach to determine accurate camera pose during preprocessing involves feature matching between historical images and modern basemaps; then solving for the required homography or projective transform using known intrinsic parameters. Recent implementations of this approach (e.g., Zhang et al., 2021; Heisig and Simmen, 2021; Maiwald and Maas, 2021) using neural networks (e.g., DeTone et al., 2018; Sarlin et al., 2020) are promising, though challenges remain for small image footprints in rugged mountain terrain with limited stable surfaces.

6.4.4. 4D SfM for other historical archives and larger blocks

While the NAGAP camera equipment and acquisition geometry evolved over the decades, the NAGAP images in the ADC archive were systematically scanned and the block sizes are relatively small. Other historical aerial photography archives contain heterogeneous image sets acquired using a range of camera and lens combinations (different focal length, film size, lens distortion) at different altitudes, with different overlap, different spatial coverage (including large blocks spanning 10 s to 100 s of km), and different scanning resolution (e.g., Section 3.2). Though it increases complexity, we recommend the multi-temporal bundle adjustment approach (Section 4.6) for most sites with overlapping image sets, either from a single archive or multiple heterogeneous archives. Metashape will jointly refine shared intrinsic parameters using feature matches across all overlapping images, which should eliminate or reduce systematic error (e.g., “doming”) in the final DEM products (Fig. 11). The user can specify image sets acquired by each unique camera, or supervise the automated parsing done by Metashape based on unique input image dimensions. Preliminary tests using 4D SfM to process historical images from cameras with different focal lengths (132 and 152 mm) in the EE archive scanned at 25-micron resolution with NAGAP ADC images scanned at 20-micron resolution show promising results. Future work involving 4D SfM using images from other archives with larger spatial coverage will provide an opportunity to refine HSfM processing approaches and evaluate long-range spatial autocorrelation of errors.

6.4.5. User-defined thresholds and parameters

As described throughout the text, we set various thresholds, weights and tuning parameters for HIPP/HSfM using a suite of test datasets from the NAGAP and EE archives. These parameters should work well

for automated processing of the NAGAP archive, but may require tuning for other historical image datasets. The HIPP and HSfM packages generate quality control metrics and plots that can help guide parameter adjustment as needed.

6.5. Toward fully open-source software

While the HIPP and HSfM repositories are open source, the current HSfM method relies on the commercial Agisoft Metashape software. Metashape provides robust SfM reconstruction for poorly constrained historical image inputs, with excellent documentation, an intuitive Python API, and an active user community. While short-term trials offer the functionality needed to run HSfM, continued use requires the purchase of a license, and the closed source prevents detailed examination of the underlying algorithms.

We developed an experimental HSfM workflow using the open-source NASA Ames Stereo Pipeline (ASP) command-line utilities (Knuth et al., 2021a) for the *single date - single block* processing mode. Preliminary tests showed promising results for the 1977-09-27 blocks at Mount Baker (Fig. S10), but slightly larger errors than the Metashape products (NMAD increase of 0.05 m over stable surfaces compared to the reference DEM). However, the ASP workflow failed for more complex NAGAP image sets with inaccurate initial position estimates (e.g., Fig. 12A), as the current ASP bundle adjustment routines are more sensitive to initial camera pose estimates and uncertainty.

The code for the experimental ASP workflow is included in the HSfM repository, but we did not pursue further development and testing for the ongoing NAGAP archive processing or this manuscript. We are hopeful that planned improvements in the ASP bundle adjustment routines and/or support for other open-source photogrammetry packages like COLMAP (Schönberger et al., 2016; Schönberger and Frahm, 2016) will enable development of a robust, fully open-source HSfM workflow in the future.

6.6. Science applications

Here we consider some of the broader implications of this work and highlight several new science applications enabled by the HSfM workflow and data products.

6.6.1. Geodetic glacier mass-balance and velocity

Processing the NAGAP and EE archives will enable the study of long-term geodetic mass balance for glaciers in western North America with unprecedented accuracy. The increased temporal resolution and accuracy will help calibrate glacier mass-balance records from field measurements (O'Neel et al., 2019; Zemp et al., 2013) and improve estimates of historic contributions to sea level rise (e.g., Arendt et al., 2002). On interannual to decadal time scales, glaciers show both temporally and spatially heterogeneous response patterns, as a function of their local geometry and regional climate forcing (e.g., Menounos et al., 2019; Christian et al., 2022). Uncertainty associated with decadal glacier mass change remains high for western North America, compared to other regions in the world (e.g., Zemp et al., 2019). These historical DEMs can be combined with modern satellite and airborne DEM time series to better constrain regional glacier change and understand connections to climate forcing for a range of timescales (seasonal to centennial) (e.g., Dehecq et al., 2020; Hugonnet et al., 2021; Geyman et al., 2022; Mannerfelt et al., 2022).

Beyond mass balance observations, the historical orthomosaics from the same year or adjacent years can be used to prepare maps of historical glacier surface velocity using feature tracking (e.g., Wang et al., 2016; Li et al., 2017; Dehecq et al., 2022). These observations can be combined with modern glacier velocity observations to assess long-term relationships with mass balance trends (e.g., Dehecq et al., 2019).

6.6.2. Elevation change rate precision and other applications

The historical DEM accuracy was ~0.5 m for our study sites, with a total time period of ~30 years between the first and last historical DEM, and a total time period of ~50 years between the first historical DEM and the reference DEM. Depending on error correlation, these products allow for elevation change measurements with precision of ~0.7–1.0 m, enabling analysis of elevation change rates as small as ~1–3 cm/yr, with spatial resolution of <1 m. This precision is remarkable considering the data origin — degrading film rolls from aerial observation programs never intended for these applications. This elevation change precision enables analysis of many other processes beyond the relatively large glacier change signals considered here. For example, the HSfM products can be used for regional analysis of proglacial erosion and deposition (e.g., Anderson and Shean, 2022), with implications for sediment transport and downstream flood hazards.

7. Conclusions

Automated processing of scanned historical aerial photographs has remained a technically challenging, yet scientifically valuable objective across many disciplines. We developed an automated method to produce accurately georeferenced historical digital elevation models (DEMs) for science and engineering applications. The method overcomes current bottlenecks for historical Structure from Motion (SfM) processing at scale by automating image preprocessing, block detection, and DEM co-registration. Image preprocessing and standardization approaches provide accurate interior orientation for a range of fiducial marker types and detection strategies. The multi-temporal *4D SfM* approach significantly improves both intrinsic and extrinsic camera model parameters, which mitigates systematic DEM error (Fig. 11). Finally, a robust multi-stage DEM co-registration approach, instead of manual ground control point (GCP) selection, refines absolute geolocation accuracy using a modern reference DEM.

We tested the Historical Structure from Motion (HSfM) method for two study sites with variable image overlap, acquisition geometry, and reference terrain characteristics. We produced 39 DEMs from 504 historical images acquired between 1967 and 1997. While the HSfM method performed best with an accurate, high-resolution reference DEM (e.g., airborne lidar), we show that using a relatively coarse global reference DEM (30 m Copernicus DEM, (European Space Agency, 2022)) also offered precise relative elevation change measurements for historical DEM time series stacks. This implies that HSfM can be used to process historical aerial photos anywhere on Earth. The automated HSfM multi-temporal bundle adjustment and co-registration approach outperformed traditional workflows involving a limited number of manually defined GCPs (Section 5.2.1), while also improving data provenance and reproducibility. The resulting historical DEM stacks have relative accuracy of 0.37–0.68 m and an elevation change rate precision of ~1–3 cm/yr, which is negligible compared to the magnitude of observed long-term glacier and geomorphic change signals.

The HSfM method enables systematic processing of historical image archives at scale and precise surface elevation change measurements over a broad range of spatial and temporal scales. These measurements offer new insights into the underlying processes driving the observed changes. We are continuing to process the remainder of the scanned North American Glacier Aerial Photography (NAGAP) archive and suitable collections from the EarthExplorer (EE) archive with the goal of improving our understanding of regional mass balance and the climate sensitivity of North American glaciers.

CRediT authorship contribution statement

Friedrich Knuth: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **David Shean:** Conceptualization, Methodology, Resources, Writing - Original Draft, Writing

– review & editing, Supervision, Project administration, Funding acquisition. **Shashank Bhushan:** Investigation, Visualization, Writing – review & editing. **Eli Schwat:** Software, Validation, Formal analysis, Data curation, Writing – review & editing. **Oleg Alexandrov:** Software, Writing – review & editing. **Christopher McNeil:** Validation, Writing – review & editing. **Amaury Dehecq:** Investigation, Writing – review & editing. **Caitlyn Florentine:** Writing – review & editing, Funding acquisition. **Shad O'Neel:** Writing – review & editing, Funding acquisition.

Data and software availability

All scanned images and initial metadata used in this study are publicly available from the NSF Arctic Data Center and USGS EarthExplorer archives. The reference DEM data are publicly available as described in Section 3.5. The HSfM DEMs and orthoimage mosaics prepared for this manuscript are available in a Zenodo dataset archive with doi: 10.5281/zenodo.7297154. Open-source repositories for the Historical Image Pre-Processing (HIPP, doi: 10.5281/zenodo.5510876) and Historical Structure from Motion (HSfM, doi: 10.5281/zenodo.5510870) packages are publicly available on Github with MIT license, documentation, and Jupyter notebook examples.

Acknowledgments

We thank the late Austin Post and the USGS Glaciology program for acquiring the NAGAP photographs over the years. We also thank Matt Nolan and the team at the University of Alaska Fairbanks for scanning and archiving the NAGAP photographs. Scott McMichael provided additional support and guidance on using the Ames Stereo Pipeline software. Knut Christianson, Michelle Koutnik and Erin Whorton contributed to the UW Innovation Award proposal. Erin Whorton and Scott Anderson participated in early discussions on SfM workflows for historical images using GCPs. We appreciate the detailed and constructive comments from Louis Sass and anonymous reviewers, which improved this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Funding

This work was funded in part by a University of Washington Innovation Award to D.S., USGS Ecosystem Mission Area, Land Change (Climate R&D) Program cooperative agreement #G19AC00433, and NSF PREEVENTS program award: ICER #1663859.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.rse.2022.113379. The pdf contains additional information about NAGAP images, Supplemenatry Tables S1–S4 and Figures S1–S12.

References

Ackermann, F., 1981. Block adjustment with additional parameters. Photogrammetria 36 (6), 217–227. http://dx.doi.org/10.1016/0031-8663(81)90040-5.

AgiSoft Metashape, 2021. 1.6.5 Professional Edition.

Albertz, J., 2001. Albrecht Meydenbauer – Pioneer of photogrammetric documentation of the cultural heritage. In: Proceedings of the 18th CIPA Symposium, Potsdam, Germany. pp. 19–25.

Anderson, S., Pitlick, J., 2014. Using repeat lidar to estimate sediment transport in a steep stream. J. Geophys. Res. Earth Surf. 119 (3), 621–643. http://dx.doi.org/10.1002/2013JF002933.

Anderson, S.W., Shean, D., 2022. Spatial and temporal controls on proglacial erosion rates: A comparison of four basins on mount rainier, 1960 to 2017. Earth Surf. Process. Landforms 47 (2), 596–617. http://dx.doi.org/10.1002/esp.5274.

Arendt, A.A., Echelmeyer, K.A., Harrison, W.D., Lingle, C.S., Valentine, V.B., 2002. Rapid wastage of Alaska glaciers and their contribution to rising sea level. Science 297 (5580), 382–386. http://dx.doi.org/10.1126/science.1072497.

Bauer, P., Stevens, B., Hazeleger, W., 2021. A digital twin of Earth for the green transition. Nature Clim. Change 11 (2), 80–83. http://dx.doi.org/10.1038/s41558-021-00986-y.

Berkner, L.V., 1954. International scientific action: The international geophysical year 1957-58. Science 119 (3096), 569–575. http://dx.doi.org/10.1126/science.119.3096.569.

Beyer, R.A., Alexandrov, O., McMichael, S., 2018. The Ames Stereo Pipeline: NASA's Open Source Software for Deriving and Processing Terrain Data. Earth Space Sci. 5, 537–548. http://dx.doi.org/10.1029/2018EA000409.

Beyer, R., Alexandrov, O., ScottMcMichael, 2019. NeoGeographyToolkit/StereoPipeline: ASP 2.6.2. http://dx.doi.org/10.5281/zenodo.3247734.

Bjørk, A.A., Kjær, K.H., Korsgaard, N.J., Khan, S.A., Kjeldsen, K.K., Andresen, C.S., Box, J.E., Larsen, N.K., Funder, S., 2012. An aerial view of 80 years of climate-related glacier fluctuations in southeast Greenland. Nat. Geosci. 5 (6), 427–432. http://dx.doi.org/10.1038/ngeo1481.

Bradski, G., Kaehler, A., 2008. Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly Media, Inc..

Brown, D., 1971. Close-range camera calibration,Photogrammetric Engineering. Eng. Remote Sens. E 37.8, 855–866.

Chandler, J., Cooper, M., 1989. The extraction of positional data from historical photographs and their application to geomorphology. Photogramm. Rec. 13 (73), 69–78. http://dx.doi.org/10.1111/j.1477-9730.1989.tb00647.x.

Child, S.F., Stearns, L.A., Girod, L., Brecher, H.H., 2020. Structure-from-motion photogrammetry of Antarctic historical aerial photographs in conjunction with ground control derived from satellite data. Remote Sens. 13 (1), 21. http://dx.doi.org/10.3390/rs13010021.

Christian, J.E., Whorton, E., Carnahan, E., Koutnik, M., Roe, G., 2022. Differences in the transient responses of individual glaciers: a case study of the Cascade Mountains of Washington State, USA. J. Glaciol. 1–13. http://dx.doi.org/10.1017/jog.2021.133.

Crandell, D., Wright, H.E., Frey, D.G., 1965. The glacial history of western Washington and Oregon. In: The Quaternary of the United States. Princeton University Press, pp. 341–354.

Czuba, J.A., Magirl, C.S., Czuba, C.R., Grossman, E.E., Curran, C.A., Gendaszek, A.S., Dinicola, R.S., 2011. Sediment load from major rivers into Puget Sound and its adjacent waters. U.S. Department of the Interior, U.S. Geological Survey.

Dehecq, A., Altena, B., Gardner, A.S., Trouvé, E., Leinss, S., 2022. Remote sensing of glacier motion. Surf. Displ. Meas. Remote Sens. Images 339–372. http://dx.doi.org/10.1002/9781119986843.ch11.

Dehecq, A., Gardner, A.S., Alexandrov, O., McMichael, S., Hugonnet, R., Shean, D., Marty, M., 2020. Automated processing of declassified KH-9 Hexagon satellite images for global elevation change analysis since the 1970s. Front. Earth Sci. 8, 566802. http://dx.doi.org/10.3389/feart.2020.566802.

Dehecq, A., Gourmelen, N., Gardner, A.S., Brun, F., Goldberg, D., Nienow, P.W., Berthier, E., Vincent, C., Wagnon, P., Trouvé, E., 2019. Twenty-first century glacier slowdown driven by mass loss in High Mountain Asia. Nat. Geosci. 12 (1), 22–27. http://dx.doi.org/10.1038/s41561-018-0271-9.

DeTone, D., Malisiewicz, T., Rabinovich, A., 2018. Superpoint: Self-supervised interest point detection and description. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 224–236.

Dewitz, J., 2016. National Land Cover Dataset (NLCD) 2016 Products. http://dx.doi.org/10.5066/P96HHBIE.

Eaton, J.P., Murata, K., 1960. How Volcanoes Grow: Geology, geochemistry, and geophysics disclose the constitution and eruption mechanism of Hawaiian volcanoes. Science 132 (3432), 925–938. http://dx.doi.org/10.1126/science.132.3432.925.

Eltner, A., Sofia, G., 2020. Structure from motion photogrammetric technique. In: Developments in Earth Surface Processes, vol. 23. Elsevier, pp. 1–24. http://dx.doi.org/10.1016/B978-0-444-64177-9.00001-1.

European Space Agency, 2020. Copernicus DEM Product Handbook (v3.0). URL https://spacedata.copernicus.eu/.

European Space Agency, 2022. Copernicus Digital Elevation Model. URL https://registry.opendata.aws/copernicus-dem/.

Feurer, D., Vinatier, F., 2018. Joining multi-epoch archival aerial images in a single SfM block allows 3-D change detection with almost exclusively image information. ISPRS J. Photogramm. Remote Sens. 146, 495–506. http://dx.doi.org/10.1016/j.isprsjprs.2018.10.016.

Fischer, M., Huss, M., Hoelzle, M., 2015. Surface elevation and mass changes of all Swiss glaciers 1980–2010. Cryosphere 9 (2), 525–540. http://dx.doi.org/10.5194/tc-9-525-2015.

Florentine, C., Mckeon, L.L., 2022. US Geological Survey Benchmark Glacier Project. Technical Report, US Geological Survey, http://dx.doi.org/10.3133/fs20223050.

Fraser, C.S., 1997. Digital camera self-calibration. ISPRS J. Photogramm. Remote Sens. 52 (4), 149–159. http://dx.doi.org/10.1016/S0924-2716(97)00005-1.

Gesch, D.B., Oimoen, M.J., Evans, G.A., et al., 2014. Accuracy Assessment of the US Geological Survey National Elevation Dataset, and Comparison with Other Large-Area Elevation Datasets: SRTM and ASTER, vol. 1008. US Department of the Interior, US Geological Survey Sioux Falls, SD, USA, http://dx.doi.org/10.3133/ofr20141008.

Geyman, E.C., JJ van Pelt, W., Maloof, A.C., Aas, H.F., Kohler, J., 2022. Historical glacier change on svalbard predicts doubling of mass loss by 2100. Nature 601 (7893), 374–379. http://dx.doi.org/10.1038/s41586-021-04314-4.

Girod, L., Nielsen, N.I., Couderette, F., Nuth, C., Kääb, A., 2018. Precise DEM extraction from Svalbard using 1936 high oblique imagery. Geosci. Instrum. Methods Data Syst. 7 (4), 277–288. http://dx.doi.org/10.5194/gi-7-277-2018.

Gruber, M., Leberl, F., 2001. Description and evaluation of the high quality photogrammetric scanner UltraScan 5000. ISPRS J. Photogramm. Remote Sens. 55 (5–6), 313–329. http://dx.doi.org/10.1016/S0924-2716(01)00024-7.

Harper, J.T., 1993. Glacier Terminus Fluctuations on Mount Baker, Washington, U.S.A., 1940–1990, and Climatic Variations. Arctic Alpine Res. 25, 332–340. http://dx.doi.org/10.1080/00040851.1993.12003019.

Heisig, H., Simmen, J.-L., 2021. Re-engineering the past: Countrywide geo-referencing of archival aerial imagery. PFG–J. Photogramm. Remote Sens. Geoinform. Sci. 89 (6), 487–503. http://dx.doi.org/10.1007/s41064-021-00162-z.

Henderson, S.T., Pritchard, M.E., Cooper, J.R., Aoki, Y., 2019. Remotely sensed deformation and thermal anomalies at Mount Pagan, Mariana Islands. Front. Earth Sci. 238. http://dx.doi.org/10.3389/feart.2019.00238.

Höhle, J., Höhle, M., 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. ISPRS J. Photogramm. Remote Sens. 64 (4), 398–406. http://dx.doi.org/10.1016/j.isprsjprs.2009.02.003.

Hugonnet, R., Brun, F., Berthier, E., Dehecq, A., Mannerfelt, E.S., Eckert, N., Farinotti, D., 2022. Uncertainty analysis of digital elevation models by spatial inference from stable terrain. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 6456–6472. http://dx.doi.org/10.1109/JSTARS.2022.3188922.

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., et al., 2021. Accelerated global glacier mass loss in the early twenty-first century. Nature 592 (7856), 726–731. http://dx.doi.org/10.1038/s41586-021-03436-z.

James, M.R., Robson, S., 2014. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. Earth Surf. Process. Landforms 39 (10), 1413–1420. http://dx.doi.org/10.1002/esp.3609.

Karara, H., 1967. Mono versus stereo analytical photogrammetry—theoretical considerations and experimental results. Photogrammetria 22 (3), 99–113. http://dx.doi.org/10.1016/0031-8663(67)90029-4.

Kehrl, L.M., Joughin, I., Shean, D.E., Floricioiu, D., Krieger, L., 2017. Seasonal and interannual variabilities in terminus position, glacier velocity, and surface elevation at Helheim and Kangerlussuaq Glaciers from 2008 to 2016. J. Geophys. Res. Earth Surf. 122 (9), 1635–1652. http://dx.doi.org/10.1002/2016JF004133.

Knuth, F., Schwat, E., Shean, D., Bhushan, S., Alexandrov, O., 2021a. Historical Structure from Motion (HSfM) pre- release v0.1. http://dx.doi.org/10.5281/zenodo.5510870.

Knuth, F., Schwat, E., Shean, D., McNeil, C., 2021b. Historical Image Pre-Processing (HIPP) pre-release v0.1. http://dx.doi.org/10.5281/zenodo.5510876.

Korpela, I., 2006. Geometrically accurate time series of archived aerial images and airborne lidar data in a forest environment. Silva Fennica 40 (1), http://dx.doi.org/10.14214/sf.355.

LaChapelle, E., 1962. Assessing glacier mass budgets by reconnaissance aerial photography. J. Glaciol. 4 (33), 290–297. http://dx.doi.org/10.3189/S0022143000027593.

Lau, N., Borsa, A.A., Becker, T.W., 2020. Present-day crustal vertical velocity field for the contiguous united states. J. Geophys. Res. Solid Earth 125 (10), http://dx.doi.org/10.1029/2020JB020066, e2020JB020066.

Leffler, R.J., Horvitz, A., Downs, R., Changery, M., Redmond, K.T., Taylor, G., 2001. Evaluation of a National Seasonal Snowfall Record at the Mount Baker, Washington, Ski Area. Natl. Weather Digest 25 (1).

Li, R., Ye, W., Qiao, G., Tong, X., Liu, S., Kong, F., Ma, X., 2017. A new analytical method for estimating antarctic ice flow in the 1960s from historical optical satellite imagery. IEEE Trans. Geosci. Remote Sens. 55 (5), 2771–2785. http://dx.doi.org/10.1109/TGRS.2017.2654484.

Lowe, D.G., 1999. Object recognition from local scale-invariant features. In: Proceedings of the Seventh IEEE International Conference on Computer Vision, vol. 2. Ieee, pp. 1150–1157. http://dx.doi.org/10.1109/ICCV.1999.790410.

Maiwald, F., Maas, H.-G., 2021. An automatic workflow for orientation of historical images with large radiometric and geometric differences. Photogramm. Rec. 36 (174), 77–103. http://dx.doi.org/10.1111/phor.12363.

Mälicke, M., Möller, E., Schneider, H.D., Müller, S., 2022. Mmaelicke/scikit-gstat: A scipy flavoured geostatistical variogram analysis toolbox. http://dx.doi.org/10.5281/zenodo.5970098.

Mannerfelt, E.S., Dehecq, A., Hugonnet, R., Hodel, E., Huss, M., Bauder, A., Farinotti, D., 2022. Halving of Swiss glacier volume since 1931 observed from terrestrial image photogrammetry. Cryosphere Discus. 1–32.

Martin-Brualla, R., Gallup, D., Seitz, S.M., 2015. Time-lapse mining from internet photos. ACM Trans. Graph. 34 (4), 1–8. http://dx.doi.org/10.1145/2766903.

McNeil, C.J., Florentine, C.E., Bright, V.A., Fahey, M.J., Evan, E.M.C.F.L., Thoms, E., McKeon, D.E.S.L.A., Keller, R.S.M.W., Whorton, E.N., O'Neel, S., et al., 2022a. Geodetic data for USGS benchmark glaciers: Orthophotos, digital elevation models, glacier boundaries and surveyed positions. http://dx.doi.org/10.5066/P9R8BP3K.

McNeil, C.J., Sass, L., Florentine, C., Baker, E.H., Peitzsch, E.H., Whorton, E.N., Miller, Z., Fagre, D.B., Clark, A.M., O'Neel, S.R., 2022b. Glacier-Wide Mass Balance and Compiled Data Inputs: USGS Benchmark Glaciers. http://dx.doi.org/10.5066/F7HD7SRF.

Meier, M.F., Tangborn, W.V., 1965. Net Budget and Flow of South Cascade Glacier, Washington. J. Glaciol. 5 (41), 547–566. http://dx.doi.org/10.3189/S0022143000018608.

Menounos, B., Hugonnet, R., Shean, D., Gardner, A., Howat, I., Berthier, E., Pelto, B., Tennant, C., Shea, J., Noh, M.-J., et al., 2019. Heterogeneous changes in western North American glaciers linked to decadal variability in zonal wind strength. Geophys. Res. Lett. 46 (1), 200–209. http://dx.doi.org/10.1029/2018GL080942.

Meydenbauer, A., 1867. Die Photometrographie. Wochenblatt Architektenvereins Berlin.

Mikhail, E.M., Bethel, J.S., McGlone, J.C., 2001. Introduction to Modern Photogrammetry, vol. 19. New York.

Mölg, N., Bolch, T., 2017. Structure-from-motion using historical aerial images to analyse changes in glacier surface elevation. Remote Sens. 9 (10), 1021. http://dx.doi.org/10.3390/rs9101021.

Nace, R.L., 1969. Water and Man: A World View: The International Hydrological Decade. UNESCO, IHD,

Nielsen, N.I., 2017. Recovering Data with Digital Photogrammetry and Image Analysis Using Open Source Software (Master's thesis). University of Oslo.

Nolan, M., 2012. Phase one data rescue of the Austin Post air photo collection and new repeat aerial photography of Alaskan valley glaciers. NSF Award #1107737.

Nolan, M., Post, A.S., Hauer, W., Zinck, A., O'Neel, S., 2017. Photogrammetric scans of aerial photographs of North American glaciers. http://dx.doi.org/10.18739/A2KC07.

NSDDA, F., 1998. Geospatial positioning accuracy standards part 3: National standard for spatial data accuracy (no. FGDC-STD-007.3-1998).

Nuth, C., Kääb, A., 2011. Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. Cryosphere 5 (1), 271–290. http://dx.doi.org/10.5194/tc-5-271-2011.

O'Neel, S., McNeil, C., Sass, L.C., Florentine, C., Baker, E.H., Peitzsch, E., McGrath, D., Fountain, A.G., Fagre, D., 2019. Reanalysis of the US Geological Survey Benchmark Glaciers: long-term insight into climate forcing of glacier mass balance. J. Glaciol. 65 (253), 850–866. http://dx.doi.org/10.1017/jog.2019.66.

Osborn, G., Menounos, B., Ryane, C., Riedel, J., Clague, J.J., Koch, J., Clark, D., Scott, K., Davis, P.T., 2012. Latest Pleistocene and Holocene glacier fluctuations on Mount Baker, Washington. Quat. Sci. Rev. 49, 33–51. http://dx.doi.org/10.1016/j.quascirev.2012.06.004.

Østrem, G., 1975. ERTS data in glaciology—an effort to monitor glacier mass balance from satellite imagery. J. Glaciol. 15 (73), 403–415. http://dx.doi.org/10.3189/S0022143000034511.

Pelto, M., Brown, C., 2012. Mass balance loss of Mount Baker, Washington glaciers 1990–2010. Hydrol. Process. 26 (17), 2601–2607. http://dx.doi.org/10.1002/hyp.9453.

Pomerleau, F., Colas, F., Siegwart, R., Magnenat, S., 2013. Comparing ICP variants on real-world data sets. Auton. Robots 34 (3), 133–148. http://dx.doi.org/10.1007/s10514-013-9327-2.

Post, A., 1995. Annual aerial photography of glaciers in northwest North America: How it all began and its golden age. Phys. Geography http://dx.doi.org/10.1080/02723646.1995.10642540.

Post, A., Richardson, D., Tangborn, W., Rosselot, F., 1971. Inventory of glaciers in the North Cascades, Washington. http://dx.doi.org/10.3133/pp705A.

Ramsey, D., 2016. USGS Mt. Baker LiDAR Techical Report.

Ray, R.G., 1960. Aerial Photographs in Geologic Interpretation and Mapping. (373–375), US Government Printing Office, http://dx.doi.org/10.3133/pp373.

Reiners, P.W., Ehlers, T.A., Mitchell, S.G., Montgomery, D.R., 2003. Coupled spatial variations in precipitation and long-term erosion rates across the Washington cascades. Nature 426 (6967), 645–647. http://dx.doi.org/10.1038/nature02111.

RGI Consortium, 2017. Randolph Glacier Inventory 6.0. http://dx.doi.org/10.7265/N5-RGI-60.

Roberti, G., Ward, B.C., van Wyk deVries, B., Perotti, L., Giardino, M., Friele, P.A., Clague, J.J., Menounos, B., Anderson, L.S., Freschi, S., 2021. Structure from motion used to revive archived aerial photographs for geomorphological analysis: an example from Mount Meager volcano, British Columbia, Canada. Can. J. Earth Sci. 58 (12), 1253–1267. http://dx.doi.org/10.1139/cjes-2020-0140.

Salach, A., 2017. SAPC–application for adapting scanned analogue photographs to use them in structure from motion technology. Int. Arch. Photogramm. Remote Sens. Spatial Inform. Sci. 42, 197.

Sarlin, P.-E., DeTone, D., Malisiewicz, T., Rabinovich, A., 2020. SuperGlue: Learning feature matching with graph neural networks. In: CVPR. URL https://arxiv.org/abs/1911.11763.

Schönberger, J.L., Frahm, J.-M., 2016. Structure-from-motion revisited. In: Conference on Computer Vision and Pattern Recognition. CVPR.

Schönberger, J.L., Zheng, E., Pollefeys, M., Frahm, J.-M., 2016. Pixelwise view selection for unstructured multi-view stereo. In: European Conference on Computer Vision. ECCV.

Schumm, S.A., 1956. Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey. Geol. Soc. Am. Bullet. 67 (5), 597–646. http://dx.doi.org/10.1130/0016-7606(1956)67[597:EODSAS]2.0.CO;2.

Sevara, C., 2016. Capturing the Past for the Future: an Evaluation of the Effect of Geometric Scan Deformities on the Performance of Aerial Archival Media in Image-based Modelling Environments. Archaeol. Prospection 23 (4), 325–334. http://dx.doi.org/10.1002/arp.1539.

Shean, D.E., Alexandrov, O., Moratto, Z.M., Smith, B.E., Joughin, I.R., Porter, C., Morin, P., 2016. An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery. ISPRS J. Photogramm. Remote Sens. 116, 101–117. http://dx.doi.org/10.1016/j.isprsjprs.2016.03.012.

Shean, D., Bhushan, S., Lilien, D., Knuth, F., Schwat, E., Meyer, J., Sharp, M., Hu, M., 2021. Demcoreg: v1.1.0. http://dx.doi.org/10.5281/zenodo.5733347.

Shean, D.E., Bhushan, S., Montesano, P., Rounce, D.R., Arendt, A., Osmanoglu, B., 2020. A systematic, regional assessment of high mountain Asia glacier mass balance. Front. Earth Sci. 7, 363. http://dx.doi.org/10.3389/feart.2019.00363.

Shean, D.E., Joughin, I.R., Dutrieux, P., Smith, B.E., Berthier, E., 2019. Ice shelf basal melt rates from a high-resolution digital elevation model (DEM) record for Pine Island Glacier, Antarctica. Cryosphere 13 (10), 2633–2656. http://dx.doi.org/10.5194/tc-13-2633-2019.

Snavely, N., Seitz, S.M., Szeliski, R., 2008. Modeling the world from internet photo collections. Int. J. Comput. Vis. 80 (2), 189–210. http://dx.doi.org/10.1007/s11263-007-0107-3.

Tarolli, P., Mudd, S.M., 2019. Remote Sensing of Geomorphology, vol. 23. Elsevier.

Tayman, W.P., 1984. User guide for the USGS aerial camera report of calibration. Photogramm. Eng. Remote Sens. 50, 577–584.

UNESCO, 1970. Combined Heat, Ice and Water Balances at Selected Glacier Basins. a Guide for Compilation and Assemblage of Data for Glacier Mass Balance Measurements. UNESCO, Int. Hydrol. Decade.

U.S. Geological Survey, 2019. USGS 3D Elevation Program Digital Elevation Model.

USDA Forest Service, 2018. RSAC Camera Report Database - Public Edition.

Van Wijk, M., 1967. Photogrammetry applied to avalanche studies. J. Glaciol. 6 (48), 917–933. http://dx.doi.org/10.3189/S0022143000020219.

Vargo, L.J., Anderson, B.M., Horgan, H.J., Mackintosh, A.N., Lorrey, A.M., Thornton, M., 2017. Using structure from motion photogrammetry to measure past glacier changes from historic aerial photographs. J. Glaciol. 63 (242), 1105–1118. http://dx.doi.org/10.1017/jog.2017.79.

Walstra, J., Dixon, N., Chandler, J.H., 2007. Historical aerial photographs for landslide assessment: two case histories. Quart. J. Eng. Geol. Hydrogeol. 40 (4), 315–332. http://dx.doi.org/10.1144/1470-9236/07-011.

Wang, S., Liu, H., Yu, B., Zhou, G., Cheng, X., 2016. Revealing the early ice flow patterns with historical Declassified Intelligence Satellite Photographs back to 1960s. Geophys. Res. Lett. 43 (11), 5758–5767. http://dx.doi.org/10.1002/2016GL068990.

Warrick, J.A., Ritchie, A.C., Adelman, G., Adelman, K., Limber, P.W., 2017. New techniques to measure cliff change from historical oblique aerial photographs and structure-from-motion photogrammetry. J. Coast. Res. 33 (1), 39–55. http://dx.doi.org/10.2112/JCOASTRES-D-16-00095.1.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology 179, 300–314. http://dx.doi.org/10.1016/j.geomorph.2012.08.021.

WGMS, 2021. World Glacier Monitoring Service. http://dx.doi.org/10.5904/wgms-fog-2021-05, URL https://wgms.ch/data_databaseversions/.

xDEM contributors, 2021. xDEM: Analysis of digital elevation models for geosciences. http://dx.doi.org/10.5281/zenodo.4809698.

Zemp, M., Huss, M., Thibert, E., Eckert, N., McNabb, R., Huber, J., Barandun, M., Machguth, H., Nussbaumer, S.U., Gärtner-Roer, I., et al., 2019. Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016. Nature 568 (7752), 382–386. http://dx.doi.org/10.1038/s41586-019-1071-0.

Zemp, M., Thibert, E., Huss, M., Stumm, D., Rolstad Denby, C., Nuth, C., Nussbaumer, S.U., Moholdt, G., Mercer, A., Mayer, C., Joerg, P.C., Jansson, P., Hynek, B., Fischer, A., Escher-Vetter, H., Elvehøy, H., Andreassen, L.M., 2013. Reanalysing glacier mass balance measurement series. Cryosphere 7 (4), 1227–1245. http://dx.doi.org/10.5194/tc-7-1227-2013.

Zhang, L., Rupnik, E., Pierrot-Deseilligny, M., 2021. Feature matching for multi-epoch historical aerial images. ISPRS J. Photogramm. Remote Sens. 182, 176–189. http://dx.doi.org/10.1016/j.isprsjprs.2021.10.008.