Supplementary material

1 Mass balance observations

Mass balance measurements is measured by the Geological Survey of Canada (GSC) since 2012 under the joint GSC-Parks Canada initiative *Columbia Icefield-Water For Life*. Employing the glaciological method (Cogley et al., 2011), end-of-winter mass balance observations (*bw*) were derived from snow depth soundings at, and between ablation stakes along the glacier centerline (Figure 1c). Snow depths were converted to snow water equivalent (SWE) using snow density measured at a network of reference snowpits dug in the accumulation zone, near the ELA and in the ablation zone, and complemented by snow cores. End-of-summer ablation (*bs*) was measured at a network of 13 stakes along the glacier centerline (Figure 1c). The number of *bs* observations varied between years due to some stakes emerging completely from the ice before field visits, or because stakes in the upper part of the glacier were, on occasion, not accessible during field visits. *bw* observations are more numerous because the upper glacier was accessed by helicopter at the end of winter and the surveys conducted on skis; and because additional snow soundings were made between ablation stakes. The annual mass balance (*ba*) was calculated by summing the winter and summer balance data (see Demuth and Horne, 2018; Ednie et al., 2017). The data obtained over these five years were used to validate the mass balance model.

An independent model validation was performed by comparing the mass balance reconstructed by the model with cumulative geodetic mass changes from 1979 to 2016. Tennant and Menounos (2013) provided geodetic mass balances for the entire Columbia Icefield and main outlet glaciers for 1979-2009. Several discrepancies and shortcomings prompted us to re-calculate the geodetic mass balance: (i) the glacier outlines used in the mass balance model excluded two disconnected ice masses and moraines included in TM2013; (ii) the 1999 SRTM DEM was not bias-corrected in TM2013, resulting in a probable bias in geodetic mass change from 1999 onward; (iii) missing data were crudely interpolated in TM2013, possibly causing further bias and explaining part of the large errors found by TM2013; (iv) the 2010 WV2 DEM was used instead of the lower quality 2009 SPOT DEM, and the 2016 Pleiades DEM was used to complement the cumulative geodetic balance series. DEM processing and uncertainty analysis on topographic changes are described in the next sub-sections.

2 Horizontal registration of2016 Pleiade and 2010 WordView 2 and 2016 Pleiades DEMs

Tennant and Menounos (2013) ('TM2013') horizontally coregistered their DEMs to the 1986 reference DEMs using tie points between air photos. We followed the same procedure for the 2016 Pleiades and 2010 WordlView2 (WV2) DEMs. The 2016 DEM was coregistered horizontally to the reference 1986 DEM using 35 tie points between the 2016 0.5 m resolution panchromatic Pleiades image and the 1986 orthophoto. Ties points were chosen in the same stable areas identified by Tennant and Menounos (2013). Mean horizontal biases found for the 2016 Pleiades image relative to the 1986 orthophoto were 0.8 m in X and 8.2 m in Y, with a 2D RMS error of 11.8 m. An affine transformation

was used to reference the 2016 orthoimage and DEM to the 1986 reference orthophoto, with a residual RMS error of 7.9 m. The same procedure was used for the horizontal co-registration of the 2010 2 WV2 orthoimage and DEM with the 1986 orthoimage. 33 ties points were used, yielding a mean error of -2.4 m in X and -1.3 m in Y, with a 2D RMS error of 12.1 m. A residual RMS error of 11 m remained after the affine 2D registration.

3 3D coregistration of SRTM DEM

TM2013 did not register the SRTM to the reference DEM, which could have caused a non-negligible bias in calculated glacier elevation changes. Here we used Nuth and Kaab (2011) to coregister the SRTM DEM, using stable areas (barren lands) from a 30 m resolution 2010 landcover map (Latifovic et al., 2017). Because the reference 1986 DEM was only sampled at 100 m, we chose to reference the SRTM DEM to the previously registered WV2 DEM (c.f. section 1), which had the least missing values and best overall quality. The calculated displacement vector was dx = -4.6 m , dy = -2.3m , dz = -1.71 m . The residual error, estimated with the normalized median absolute deviation (NMAD), a robust estimated of random error used in previous studies (Berthier et al., 2004; Höhle and Höhle, 2009), was 12.0 m.

4 DEM vertical bias correction

Since TM2013 concentrated on extracting elevation points on the glacier, only limited areas of stable ground were available around the glacier in their DEMs. As such, only one patch of stable, low angle ground was available south of the glacier to derive elevation uncertainties and correct biases between DEMs (49-132 pixels per DEM. Outliers in elevation change values (Δh) were first removed by excluding values smaller or larger than ±3 standard deviation from the mean (e.g. Berthier et al., 2004; Gardelle et al., 2013). Absolute vertical biases were then estimated by calculating the median elevation difference between each DEM and the highest quality reference 2010 WV2 DEM over the stable ground for slopes <32° (e.g. Berthier et al., 2016; Dusaillant et al., 2019) (Table S1). The absolute vertical biases were then removed from each DEM.

5 Uncertainty estimates of geodetic mass balance

Uncertainties on geodetic mass balances were recalculated for the TM2013 DEMs and for the 2010 and 2016 DEMs. The uncertainty on elevation differences between two successive DEMs, $\varepsilon_{\Delta h}$, is estimated according to McNabb et al. (2019), following Eq. (1):

$$\varepsilon_{\Delta h} = \sqrt{\frac{\varepsilon_{\mathrm{rand}}^2}{\sqrt{n/(L/r)^2}} + \varepsilon_{\mathrm{bias}}^2} \quad (1)$$

Where $\varepsilon_{\mathrm{rand}}$ and $\varepsilon_{\mathrm{bias}}$ are respectively the standard deviation and median of elevation changes over stable ground, n is the number of pixels falling into the glacier outline, L is the autocorrelation distance, assumed here to be 500 m (e.g. Rolstad et al., 2009) and r is the pixel size (100 m). $\varepsilon_{\mathrm{rand}}$ and $\varepsilon_{\mathrm{bias}}$ were estimated on the patch of stable ground,

excluding slopes larger than 32° (e.g. Berthier et al., 2016; Dussaillant et al., 2019), giving and average of 49 gridpoints in the control patch. The mass-balance error was then computed following Eq. (2):

$$\epsilon_{\mathrm{mb}} = \varepsilon_{\Delta h}\left(5g + (1-g)\right)\left(\frac{\bar{\rho}+\epsilon_{\bar{\rho}}}{\rho_w}\right) \quad (2a)$$

$$\bar{\rho} = \rho_{\mathrm{firn}}(A_{\mathrm{firn}}) + \rho_{\mathrm{ice}}(1 - A_{\mathrm{firn}}) \quad (2b)$$

$$\epsilon_{\bar{\rho}} = 100 A_{\mathrm{firn}} \quad (2c)$$

Where g if the fraction of missing data in the accumulation zone, $\bar{\rho}$ is the average density, ϵ_{ρ} the uncertainty on the mean density, ρ_{firn} is the firn density (550 kg m^{-3}), ρ_{ice} is the ice density (900 kg m^{-3}), ρ_{w} is the density of water (1000 kg m^{-3}), and A_{firn} is the relative accumulation zone area, estimated as the fraction of the total glacier area above the mean, late-summer snowline altitude as estimated by TM2013. An uncertainty of 100 kg m^{-3} for the firn density in the accumulation zone is assumed to estimate $\epsilon_{\bar{\rho}}$. Errors in glacier outline delineation were not considered. A conservative multiplicative factor of 5 is used in equation (2) for the uncertainties in the unsurveyed areas (Berthier et al., 2014; Dussaillant et al., 2019), akin to the ± 5 m w.e. error originally ascribed to the accumulation areas by TM2013. Results of error calculations are shown in Table S2.

6 Calculation of elevation changes

Elevation changes were calculated on a pixel-to-pixel basis between two successive DEMs, after excluding outliers as previously described. In order to account for significant data voids in the accumulation zone in certain years, Δh values were spatially interpolated using bilinear interpolation, which was recently found to be one of the best methods to correct biases in geodetic balance estimates due to data voids (McNabb et al., 2019). Following the procedure of TM2013, Δh values were converted to water equivalent using a mean firn density of 550 kg m^{-3} for the accumulation area and an ice density of 900 kg m^{-3} for the ablation area, and then averaged to yield the mean geodetic mass balance. The mean snowline altitudes mapped by TM2013 were used to separate the accumulation zone from the ablation zone. We replaced the SPOT2009 DEM from TM2013 by the WV2 2010 DEM for mass balance calculations as the SPOT2009 DEM suffered from inaccuracies due to numerous voids in the accumulation zone (Table S1, Table S2).

Table S1. DEMs used for geodetic mass balance calculations. The mean, median and standard deviation (Std) of elevation differences between each DEM and the reference 2010 WV2 DEM over stable ground are shown. N: number of gridpoints over stable ground; F: fraction of glacier surface with missing data.

year	Source	Date	Scale/ resolution	Mean (m)	Median (m)	Std (m)	n	f
1979	Federal AP (TM2013[1])	9 Jul.	1 : 70 000	2.93	2.94	1.11	49	0.39
1986	Provincial AP (TM2013)	15 Aug.	1 : 60 000	3.21	3.38	1.45	49	0.34
1993	Federal AP (TM2013)	9 Sept.	1 : 50 000	2.43	2.41	0.70	49	0.30
1999	SRTM 2000[2] (TM2013)	Feb.	90 m	-6.85	-5.06	7.21	49	0.06
2009	SPOT 5 (TM2013)[3]	20-30 Aug.	2.5 m	4.48	5.43	7.83	42	0.50
2010	Word View 2 (this study)	31 Jul./18 Sept.	1 m	0.00	0	0.00	0	0.00
2016	Pleiades (this study)	21 Sept.	2 m	-6.01	-6.13	1.12	177	0.11

[1] TM2013: Tennant and Menounos, 2013

[2] SRTM from Feb. 2000 was ascribed to 1999 end-of-summer surface, following TM2013

[3] Not used in this study

Table S2. Elevation changes (m w.e.) and relative uncertainty estimates for the different DEMs and time periods. Symbols refer to the equations given in text. *mb*: geodetic mass balance for the stated period. $\dot{mb}$ = mass balance rate.

Start year	End year	$\varepsilon_{\mathrm{bias}}$ (m)	$\varepsilon_{\mathrm{rand}}$ (m)	n	g	n	A	$\bar{\rho}$ (kg m^{-3})	$\varepsilon_{\Delta h}$ (m)	ϵ_{mb} (m w.e.)	mb (m w.e.)	$\dot{mb}$ (m w.e. a^{-1})
1979	1986	-0.21	1.07	49	0.41	3226	0.62	680	0.38	0.79	-3.71	-0.53
1986	1993	-0.09	1.55	49	0.41	3118	0.53	720	0.47	1.02	-4.49	-0.64
1993	1999	-0.31	7.12	49	0.31	3029	0.34	780	2.17	4.24	-0.65	-0.11
1999	2010	-0.07	7.06	45	0.07	3030	0.16	840	2.13	2.55	-11.78	-1.07
2010	2016	0.02	1.06	172	0.11	2893	0.16	840	0.32	0.44	-6.10	-1.02

7 Precipitation lapse rate

Figure S1 shows the mean seasonal cycle of precipitation from the homogenized record from Parker Ridge and Columbia icefield stations, as well as Lake Louise, located 100 km southeast of Saskatchewan Glacier. The monthly calculate lapse rate is low in summer months (May-August) when convective precipitation occur and increase in the fall towards maximum values in winter when frontal weather forces air masses up the mountain ranges. In consequence the orographic precipitation gradient is best defined during the winter months (November to March, r^2 = 0.66-0.84, Figure S1b), when precipitation fall as snow on the glacier (Figure S1a). As such, the mean orographic gradient between November and March (15.6 +/- 4% 100 m^{-1}) was used to distribute precipitation to the glacier surface.

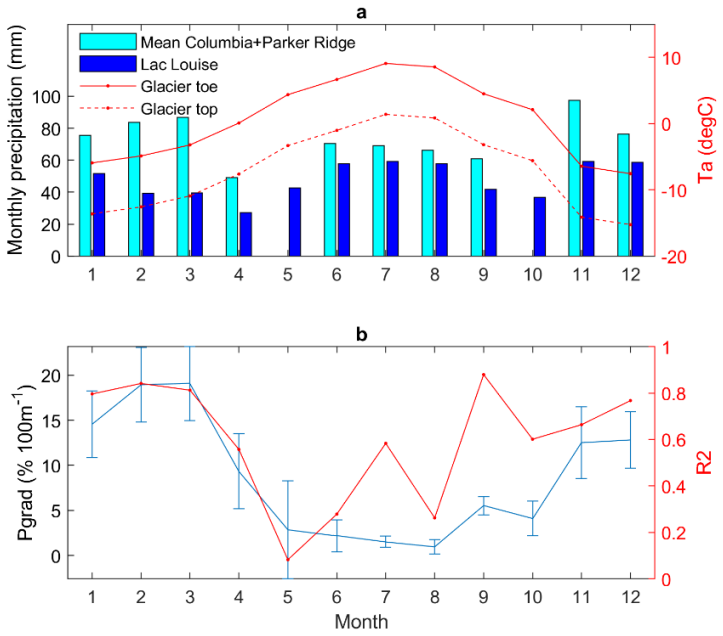

Figure S1. Seasonal cycle of precipitation and precipitation lapse rate derived from meteorological station network. (a) Monthly precipitation from Columbia and parker Ridge stations (2000 m a.s.l.) and Lake Louise (1524 m a.s.l) with mean air temperature extrapolated from glacier AWS (2193 m a.s.l.) to glacier top (3322 m a.s.l.) and glacier toe (1784 m a.s.l.). (b). Monthly precipitation lapse rate with standard error bars (blue) and r^2 coefficients.

8 Downscaling NARR

Table S3. Results of statistical downscaling of hourly NARR data using a two-year AWS record and the 29-year homogenized historical precipitation record. The daily validation is first done over the entire period before bias correction ('no downscale'). Validation statistics for each bias correction methods are the mean of the two left-out sub periods (see methods). R = Pearson correlation coefficient; RMSE = root mean square error; Bias = mean error (NARR minus observations); MAE = mean absolute error. Precipitation is not included as only daily data was available.

Method	Statistic	Ta (°C)	WS (m s^{-1})	RH (%)	G (W m^{-2})
No downscale	R	0.95	0.26	0.76	0.91
	RMSE	2.57	2.90	11.12	111.17
	Bias	-1.26	-1.13	-0.11	29.94
	MAE	1.99	2.32	8.52	60.84
Scaling	R	0.95	0.26	0.76	0.92
	RMSE	2.25	2.96	11.13	101.08
	Bias	0.00	-0.01	0.00	-0.74
	MAE	1.72	2.33	8.51	48.60
EQM	R	0.95	0.27	0.77	
	RMSE	2.08	3.18	11.57	
	Bias	0.03	0.00	-0.13	
	MAE	1.58	2.50	8.68	

9 Glacier albedo

Ice albedo in the ablation zone of Saskatchewan Glacier was mapped using 17 of the 18 cloud-free, end-of-summer Landsat images used to delineate the mean snowline position (cf. Methods section). Atmospherically corrected surface reflectance from the Landsat 5 ETM and Landsat 7 ETM+ sensors were used and converted to broadband albedo following equation (3) from Knap et al. (1999) .

$$\alpha = 0.726B2 - 0.322B2^2 - 0.051B4 + 0.581B4^2 \quad (3a)$$

$$\alpha = 0.782B4 + 0.148B4^2 \quad (3b)$$

Where B2 and B4 are respectively the green and near-infrared bands. Equation 3b was used for pixels where the green band was saturated, which frequently occurs over snow surfaces. The last image, from the Landsat 8 OLI sensor, was not used because bands are different than those in the ETM and ETM+ sensors. A median albedo map was produced (Figure S2), from which the distribution of ice albedo values was extracted in a region of interest extending below the mean snowline and excluding the glacier margins where shade effects were noticed (Figure S3). The median of the distribution (0.24) was used as the representative ice albedo in the mass-balance model and the standard deviation (0.03) used as uncertainty range for the sensitivity analysis.

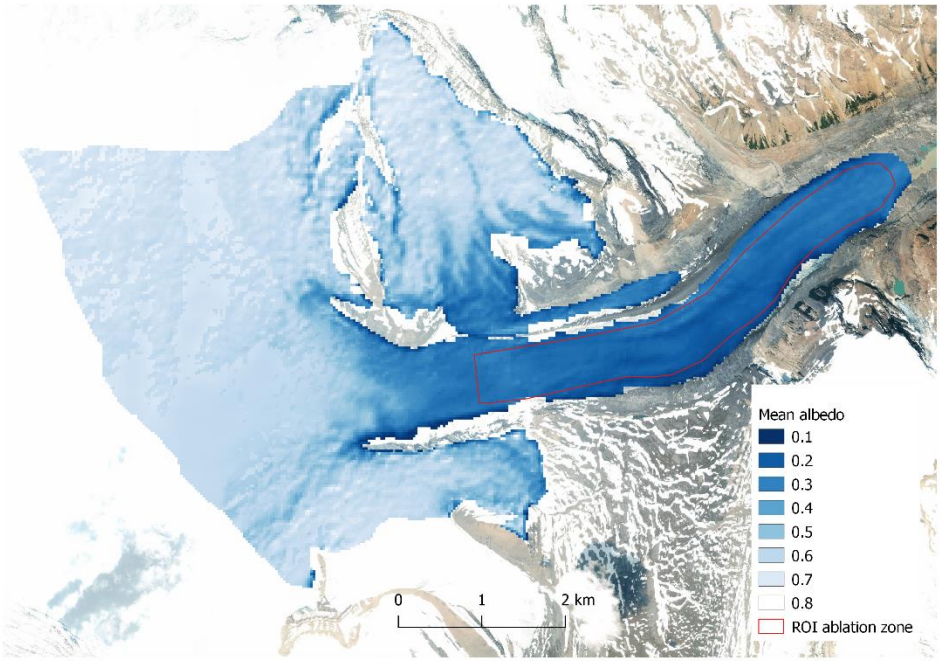

Figure S2. Mean end-of-summer albedo map from 17 Landsat scenes between 1986 and 2013.

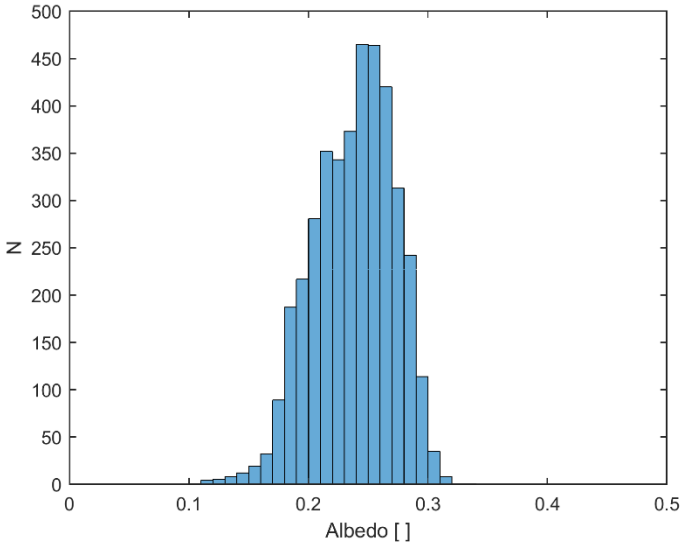

Figure S3. Histogram of mean end-of-summer albedo map in ablation zone ROI.

10 Spatial patterns of surface energy fluxes

Spatial patterns of surface energy fluxes show variations related to topography, i.e., altitude, slope, aspect and shading. SWin shows most variations, broadly increasing with altitude due to high atmospheric transmissivity, and also on south-facing slopes such as on the northern portion of the glacier (Figure S4a). Shading effects by surrounding topography are apparent in the southern portion of the glacier tongue and southern upper basin, where radiation values decrease locally. This helps to limit melting at lower altitudes where the mean annual albedo is lowest (Figure S4b). The high albedo in the predominantly snow and firn-covered accumulation area counterbalances the larger incoming solar radiation at high elevations. LW* is negative throughout the year, LWout being on average higher by 47 W m^{-2} than LWin (Figure S4c, d). Both LWin and LWout are rather uniform, varying spatially by 37 and 21 W m^{-2}, respectively, over the whole glacier. LWin increases down glacier which can be attributed to warmer air temperatures and some localized contributions from terrain emissions from rockwalls surrounding the glacier tongue. Both the latent (QL) and sensible (QH) fluxes decrease with elevation in response to negative temperature lapse rates (Figure S4e, f). The precipitation heat flux increases with elevation due to orographic precipitation, represented by the precipitation lapse rate prescribed in the model, but is otherwise negligible on average.

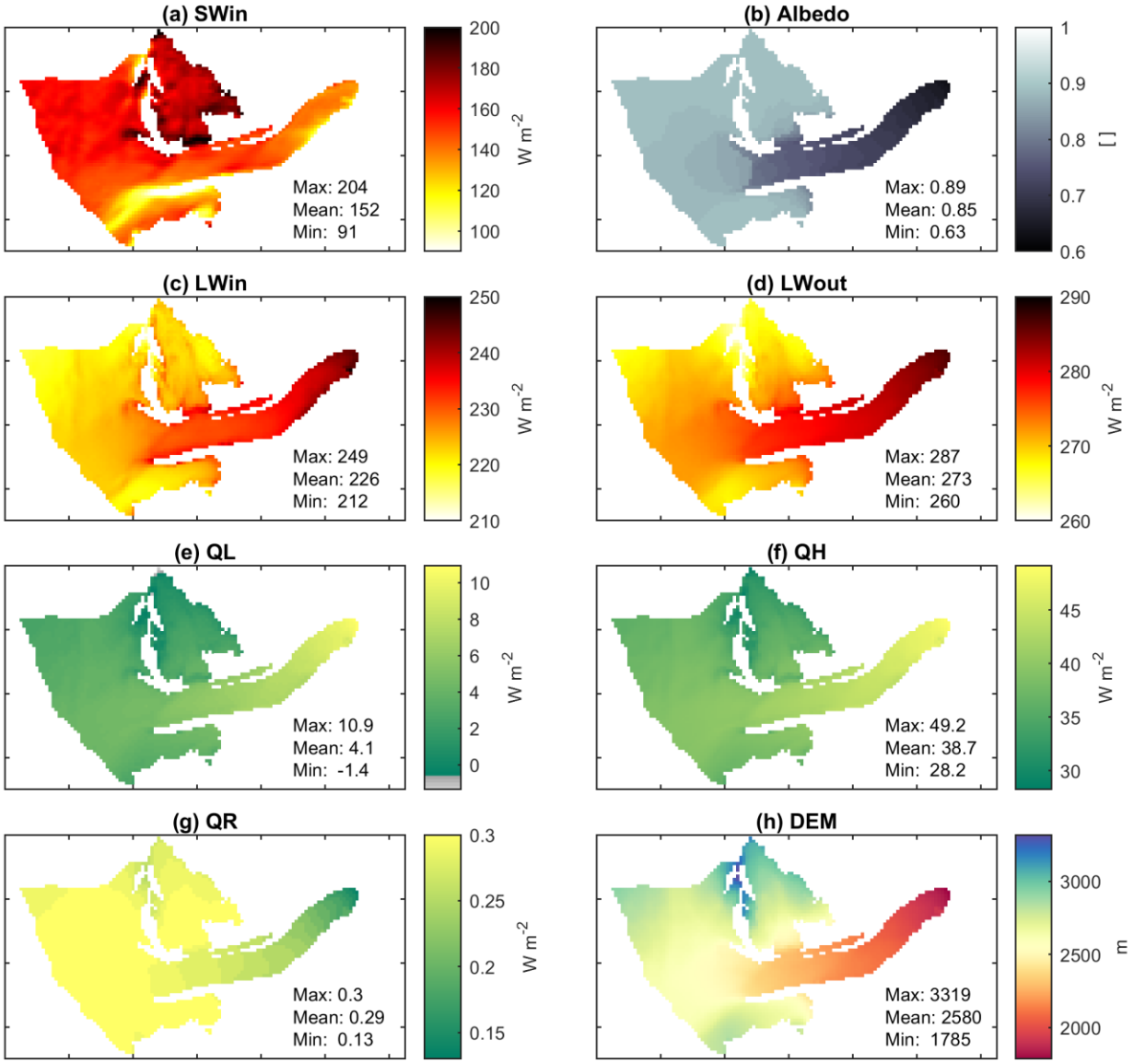

Figure S4. Mean annual spatial surface energy fluxes simulated on Saskatchewan Glacier between 1979-2016. (a) Incoming shortwave radiation; (b) albedo; (c) incoming longwave radiation; (d) outgoing longwave radiation; (e) latent heat flux; (f) sensible heat flux; (g) precipitation heat flux; (h) elevation (2010 DEM).

References

Berthier, E., Arnaud, Y., Baratoux, D., Vincent, C., and Rémy, F.: Recent rapid thinning of the "Mer de Glace" glacier derived from satellite optical images, Geophysical Research Letters, 31, 2004.

Berthier, E., Cabot, V., Vincent, C., and Six, D.: Decadal region-wide and glacier-wide mass balances derived from multi-temporal ASTER satellite digital elevation models. Validation over the Mont-Blanc area, Frontiers in Earth Science, 4, 63, 2016.

Berthier, E., Vincent, C., Magnússon, E., Gunnlaugsson, Á., Pitte, P., Le Meur, E., Masiokas, M., Ruiz, L., Pálsson, F., and Belart, J.: Glacier topography and elevation changes derived from Pléiades sub-meter stereo images, The Cryosphere, 8, 2275-2291, 2014.

Cogley, J. G., Hock, R., Rasmussen, L., Arendt, A., Bauder, A., Braithwaite, R., Jansson, P., Kaser, G., Möller, M., and Nicholson, L.: Glossary of glacier mass balance and related terms, IHP-VII technical documents in hydrology, 86, 2011.

Demuth, M. and Horne, G.: Decadal-centenary glacier mass changes and their variability, Jasper National Park, Alberta, including the Columbia Icefield region. Geological Survey of Canada Open File 8229, 2018.

Dussaillant, I., Berthier, E., Brun, F., Masiokas, M., Hugonnet, R., Favier, V., Rabatel, A., Pitte, P., and Ruiz, L.: Two decades of glacier mass loss along the Andes, Nature Geoscience, 12, 802-808, 2019.

Ednie, M., Demuth, M., and Shepherd, B.: Mass balance of the Athabasca and Saskatchewan sectors of the Columbia Icefield, Alberta for 2015 and 2016. Geological Survey of Canada, 2017.

Gardelle, J., Berthier, E., Arnaud, Y., and Kääb, A.: Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999-2011, The Cryosphere, 7, 1263-1286, 2013.

Höhle, J. and Höhle, M.: Accuracy assessment of digital elevation models by means of robust statistical methods, ISPRS Journal of Photogrammetry and Remote Sensing, 64, 398-406, 2009.

Knap, W., Reijmer, C., and Oerlemans, J.: Narrowband to broadband conversion of Landsat TM glacier albedos, International Journal of Remote Sensing, 20, 2091-2110, 1999.

Latifovic, R., Pouliot, D., and Olthof, I.: Circa 2010 land cover of Canada: Local optimization methodology and product development, Remote Sensing, 9, 1098, 2017.

Liang, S.: Narrowband to broadband conversions of land surface albedo I: Algorithms, Remote Sensing of Environment, 76, 213-238, 2001.

McNabb, R. W., Nuth, C., Kääb, A., and Girod, L. M. R.: Sensitivity of glacier volume change estimation to DEM void interpolation, The Cryosphere, 13, 895-910, 2019.

Nuth, C. and Kääb, A.: Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change, The Cryosphere, 5, 271-290, 2011.

Rolstad, C., Haug, T., and Denby, B.: Spatially integrated geodetic glacier mass balance and its uncertainty based on geostatistical analysis: application to the western Svartisen ice cap, Norway, Journal of Glaciology, 55, 666-680, 2009.

Tennant, C. and Menounos, B.: Glacier change of the Columbia Icefield, Canadian Rocky Mountains, 1919–2009, Journal of Glaciology, 59, 671-686, 2013.