Supplemental material for 'Evaluating Greenland Surface-Mass-Balance and Firn-Densification Data Using ICESat-2 Altimetry'

S1. ATL11 data product:

The ATL11 product [*Smith et al.*, 2020] is generated from the ATL06 along-track-height product that identifies locations where ICESat-2 has made multiple measurements for the same location on the ice sheets from different cycles of ICESat-2's orbit. Because ICESat-2 measures the surface height using pairs of beams separated by 90 m, each cycle's measurements can help define the shape of the surface in a region around any point that it measures. ATL11 processing uses multiple measurements from different cycles to define a reference surface for local (~100-m) regions around *reference points*, located every 60 m along ICESat-2's reference pair tracks (RPTs). By subtracting the reference surface values from the height measurements from different cycles for the same repeat track at the same reference surface, the algorithm obtains a set of *corrected* height measurements that are directly comparable with one another, even though ICESat-2 does not exactly measure the same locations on each repeat cycle. Further, by combining multiple measurements from each cycle at each reference point, the ATL11 algorithm can identify some ATL06 data with large errors, and either remove them from the analysis or provide flag values that indicate that these points should be used with caution. Corrected height measurements for which the estimated error is larger than 10 m are not included in the product, and the *fit_quality* flag is used to indicate reference points that have excessively large (any component magnitude larger than 0.2) slope estimates or that have large error estimates (implying correction-magnitude errors greater than 1 m) for any component of the reference-surface polynomial.

The reference points are defined for points along the RPTs, which are the pair tracks that were designed to be measured by ICESat-2 every 91 days. Measurements of these RPTs began with the start of ICESat2's third cycle, in April of 2019. Measurements from cycles 1 and 2, which were displaced from the RPTs by 1-5 km, are included in the data at locations where cycle 1 and 2 measurements intersect the RPTs. This lets the algorithm correct these measurements for local topography using the reference surfaces defined by cycles 3-8.

We read the along-track elevation values for all tracks intersecting the Greenland ice sheet, and, to make the data more manageable, remove every second reference point from the data, to leave one reference point every 120 m along the RPTs. We then calculate height differences between subsequent cycles from the corrected height field (*h_corr)*, to give up to 5 difference measurements for each reference point. We use the *fit_quality* flag to remove any reference points suspected of having low-quality reference surfaces (we use only those reference points for which *fit_quality*=0), and filter the remaining data to eliminate any points that are not over the ice sheet, based a low-resolution mask that distinguishes the ice sheet from

outlying ice caps [*Shepherd et al.*, 2012], and a higher-resolution mask that distinguishes ice cover from rock and water [*Howat*, 2017]).

35 To include cycles 1 and 2 in our analysis, we read the crossover data for cycles 1 and 2 from the same tracks from which we read the along-track data. These data undergo the same filtering In In reading the crossover data we follow the same steps as for the along-track data, and we also remove any points for which the cycle 1 or cycle 2 measurement had an ATL06 flag value indicating a potentially problematic measurement (we require that *atl06_quality_summary=0*), and we eliminate any points collected during an ICESat-2 calibration maneuver. This last step is necessary because during the calibration maneuvers, the

40 ATLAS beams are pointed up to 5 degrees off nadir, and the spacecraft is rapidly changing its orientation, so the accuracy of the measurements is noticeably inferior to those collected close to nadir. The filtering is necessary for crossover measurements but not for the along-track measurements because the large off-nadir pointing displaces the beams outside the across-track search window, so the affected measurements are never included in the along-track solution, but the off-nadir measurements are still available for crossovers.

45

S2. Velocity-based data masking
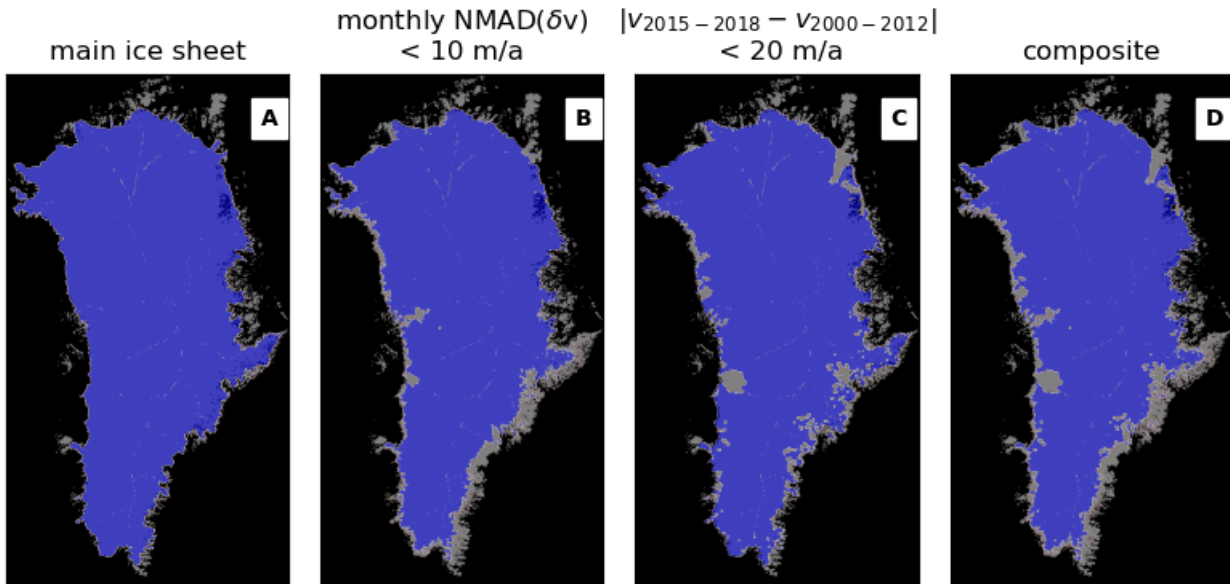
Figure S1. Masking steps to isolate the effects of SMB changes from those of velocity changes. Blue areas indicate data to be used based on each step, gray areas indicate ice identified based on the GIMP ice mask [Howat, 2017]

50 To identify height-difference measurements from the ice sheet that are likely not strongly affected substantially by ice-dynamic changes, we begin with a mask that separates the ice sheet from rock and ocean [*Howat*, 2017]. To isolate the main ice sheet

from less well sampled outlying ice caps, we combine this with a mask that only includes the main ice sheet [*Shepherd et al.*, 2012].

We assume that we can identify areas affected by seasonal or interannual ice-dynamic changes by estimating changes in ice flow parallel to the mean ice-flow direction. To identify areas that have substantial seasonal velocity variability, we use a set of monthly velocity maps spanning 31 December 2014 to 30 November 2019 [*I Joughin*, 2020] derived from synthetic aperture radar (SAR) and optical feature tracking, combined with the error-weighted mean of a set of annual velocity mosaics from years between 2014 and 2018 [*I. Joughin et al.*, 2015, updated 2018.]. For each pixel in each map, we calculate the along-flow component of the velocity difference relative to the 2014-2018 map. We then calculate the pixel-wise median of these differences for each calendar month, for all years available (e.g. the median velocity difference from all September mosaics between 2014 and 2019). We then calculate the pixel-wise NMAD (normalized median absolute difference) of the along-flow component differences for all the months, to give a total per-month variability for each pixel on the ice sheet (figure S2A). We create an initial binary monthly-variability mask from the set of all points that have NMAD values less than 10 m yr^{-1}. This mask contains a granular scatter of point classifications around its boundaries, so we first consolidate the areas that have a large concentration of variability values greater than the threshold using a binary opening with a scale of 1 km (i.e. remove points to the mask that are within 1 km of the mask boundaries, then add points that are within 1 km of the reduced mask boundaries), then remove isolated, patches with large NMAD values using a binary closing with a scale of 3 km (a binary closing performs the same steps as a binary opening, in the opposite order). Because the velocity estimates in the interior of the ice sheet, where SAR measurements are sporadic, are noisy, we use a manually drawn polygon to add points in the interior of the ice sheet to the mask.

To identify areas that have undergone substantial velocity change in the last two decades, we use a set of annual velocity mosaics spanning winters between 2000 and 2012 [*I. Joughin et al.*, 2015, updated 2018.]. We calculate the difference between annual velocity estimates from between 2000 and 2012 and our 2014-2018 velocity composite. Our long-term variability mask (Fig. S1c) is the set of pixels that do not have an absolute velocity difference greater than the quadratic sum of the per-pixel error estimate and 20 m yr^{-1} for any year between 2000 and 2012; because the maps for a few years contain substantial velocity errors in the interior of the ice sheet, we manually removed large isolated patches velocity differences from interior of the ice sheet from this mask, so that the maxed pixels are only around the edge of the ice sheet. We applied a binary opening on this mask to remove isolated pixels.

Our composite mask (Fig. S1d) is the intersection of the ice-only mask, the main-ice-sheet mask (Fig. S1a), the short-term variability mask (Fig. S1b), and the long-term variability mask (Fig. S1c).
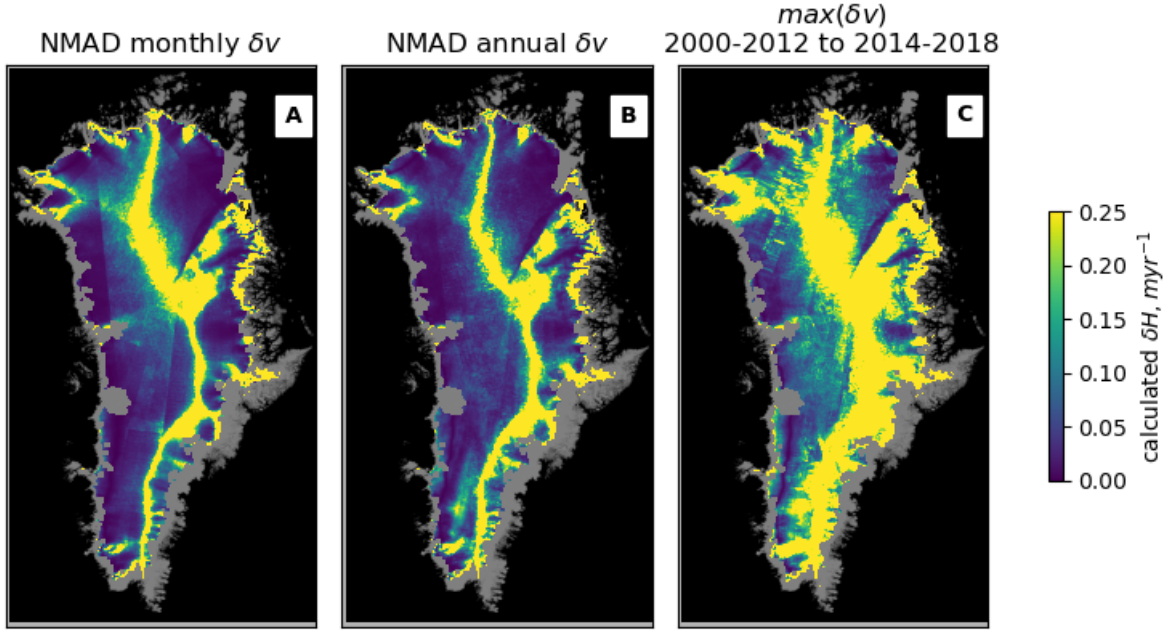
Figure S2. Predicted thickness change based on the steady-state approximation. Thickness-change rates are calculated with Equation S2, based on A: the NMAD of monthly velocity fluctuations, B: The NMAD of year-to-year fluctuations, C: the maximum difference between any annual velocity between 2000 and 2012 and the 2014-2018 mean velocity. All maps are set to zero in masked areas

In generating these masks, we adopted 20 m yr^{-1} as a threshold representing substantial interannual velocity variations and 10 m yr^{-1} as a threshold representing substantial monthly variations. We can evaluate how well this standard removes height-change signals caused by dynamic (velocity) variations by considering a vertically-averaged description of local ice-sheet mass balance :

$$\nabla \cdot \mathbf{u_{xy}}fH + \frac{1}{\rho_i}\frac{dM_{SMB}}{dt} + \frac{dh_{air}}{dt} - H_t = 0 \qquad\qquad S1$$

Where $\mathbf{u_{xy}}$ is the surface horizontal velocity field, f is the ratio between the surface and depth-averaged velocities, H is the ice thickness. ρ_i is the density of ice, $\frac{dM_{SMB}}{dt}$ is the net contribution of surface-mass-balance processes, $\frac{dh_{air}}{dt}$ is the rate of change in the firn air, and H_t is the rate of change in the ice thickness. In most parts of the ice sheet, we do not know the details of the ice thickness well enough to evaluate the velocity-dependent term, $\nabla \cdot \mathbf{u_{xy}}fH$, accurately at short scales, and errors in velocity estimates tend to make large apparent contributions to this term. However, we note that for an ice sheet in steady state, the first two terms should balance, so that the net ice flux divergence is equal to the ice accumulation or ablation rate. This means that fractional changes in the mass-flux divergence should result in surface-height change rates that are a similar fraction of the steady-state surface-mass-balance rate. Further, as long as changes in the spatial gradients of $\mathbf{u_{xy}}$ are

proportional to the magnitude of $\mathbf{u}_{xy}$, the fractional change in thinning due to ice-flux divergence should approximately equal the fractional change in velocity. Based on this, we can approximate the magnitude of the thickness-change signal:

$$\left|d\mathrm{H}_{t,vel}\right| \approx \frac{|\delta u|}{|\boldsymbol{u}|}\left|\frac{1}{\rho_i}\frac{d\mathrm{M}_{SMB}}{dt}\right| \qquad \text{S2}$$

105

The liberal use of absolute-value signs here is intended to convey that we are mostly interested in the magnitudes of the components of each side of the equation. Figure S2 shows the height-change magnitudes calculated from equation S2 based the MERRA-2 mean SMB field, and the NMAD of monthly velocity differences (Fig. S2A), the NMAD of 2014-2018 annual differences (Fig. S2B), and the maximum magnitude calculated based differences between velocity fields from 2000 through

110 2012 and the 2014-2018 average velocity (Fig. S2C), after the combined mask has been applied. For the monthly and annual maps (Figs S2A-B), the largest calculated height differences are near the divides, where equation S2 divides small velocity errors by smaller velocity magnitudes, resulting in large height-change values. We do not feel that these areas should be masked, because we have no evidence that there should be substantial dynamic change in these areas. The larger values to the east of the flow divide in Southeast Greenland in figs S2A and S2C are also likely related to errors in the velocity maps,

115 because both the monthly maps and the pre-2010 velocity maps appear to have somewhat noisy values in these areas, where clouds and heavy snowfall interfere with optical and SAR measurements, respectively; when generating the mask, we took these errors into account, but did not in generating the error-estimate map for S2C. For the remainder of the ice sheet, calculated height-change rates are on the order of a few cm yr^{-1}, and are not clearly in excess of those expected due to errors in the velocity measurements.

120 **S3. Regional statistics of GSFCv1.2 high-melt height differences.**

In the main text, we presented statistics for the high-melt subsamples of GSFCv1.1 and MARv3.11.5. Figure S3 shows the corresponding histograms for GSFCv1.2.
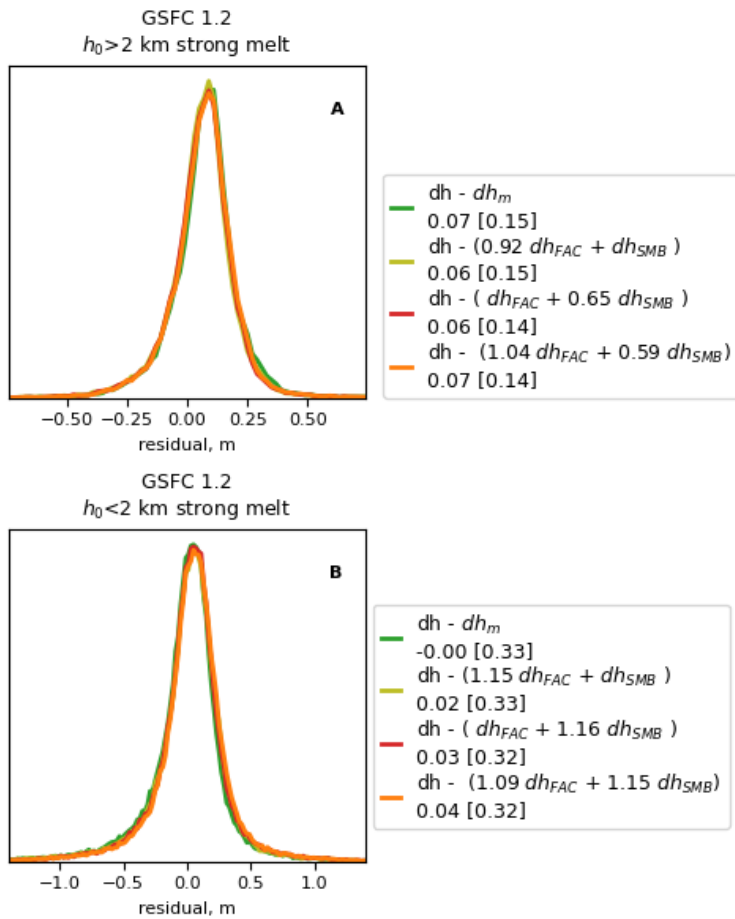
Figure S 1 . Regional model-data difference histograms for GSFCv1.2. (A, C) and GSFCv1.1 (B, D) models, for the high-elevation (A, B) and low-elevation (C, D) subsamples of the high-melt ($z_{melt} > 0.2$ |dh_{SMB}|) data. Model components listed in the captions are the firn-air-column change (dh_{FAC}) and the surface-mass-balance height change (dh_{SMB}). The mean and [standard deviation] of the residuals are given in each caption.

Elev.	subset	MARv3.11.5 correction	mean	std	R2
all	all	dh	-0.024	0.246	0.00
all	all	dh - dh_m	0.02	0.192	0.39
all	all	dh - 0.60 dh_m	0.003	0.139	0.68
all	sp-su 2019	dh	-0.213	0.368	0.00
all	sp-su 2019	dh - dh_m	0.113	0.255	0.52
all	sp-su 2019	dh - 0.63 dh_m	-0.008	0.166	0.80
all	sp-su 2020	dh	-0.027	0.216	0.00
all	sp-su 2020	dh - dh_m	0.022	0.186	0.26
all	sp-su 2020	dh - 0.57 dh_m	0.001	0.13	0.64
high	weak melt	dh	0.007	0.12	0.00
high	weak melt	dh - dh_m	-0.021	0.094	0.39
high	weak melt	dh - 0.68 dh_m	-0.012	0.086	0.49
low	weak melt	dh	0.072	0.255	0.00
low	weak melt	dh - dh_m	-0.055	0.178	0.51
low	weak melt	dh - 0.77 dh_m	-0.025	0.169	0.56
high	strong melt	dh	-0.088	0.225	0.00
high	strong melt	dh - dh_m	0.134	0.214	0.10
high	strong melt	dh - 0.52 dh_m	0.027	0.116	0.74
low	strong melt	dh	-0.348	0.648	0.00
low	strong melt	dh - dh_m	0.162	0.35	0.71
low	strong melt	dh - 0.79 dh_m	0.054	0.316	0.76
all	all	dh	-0.024	0.246	0.00
all	all	dh - dh_m	0.02	0.192	0.39
all	all	dh - (0.46 dh_FAC + dh_SMB)	0.005	0.143	0.66
all	all	dh - (dh_FAC + 0.50 dh_SMB)	0.012	0.173	0.50
all	all	dh - (0.52 dh_FAC + 0.74 dh_SMB)	0.003	0.137	0.69
high	weak melt	dh	0.007	0.12	0.00
high	weak melt	dh - dh_m	-0.021	0.094	0.39
high	weak melt	dh - (0.50 dh_FAC + dh_SMB)	-0.007	0.084	0.52
high	weak melt	dh - (dh_FAC + 0.44 dh_SMB)	-0.022	0.091	0.43
high	weak melt	dh - (0.25 dh_FAC + 1.66 dh_SMB)	0.001	0.082	0.54
low	weak melt	dh	0.072	0.255	0.00
low	weak melt	dh - dh_m	-0.055	0.178	0.51
low	weak melt	dh - (0.64 dh_FAC + dh_SMB)	-0.025	0.169	0.56
low	weak melt	dh - (dh_FAC + 0.66 dh_SMB)	-0.04	0.173	0.54
low	weak melt	dh - (0.71 dh_FAC + 0.85 dh_SMB)	-0.024	0.168	0.56
high	strong melt	dh	-0.088	0.225	0.00
high	strong melt	dh - dh_m	0.134	0.214	0.10
high	strong melt	dh - (0.45 dh_FAC + dh_SMB)	0.019	0.113	0.75
high	strong melt	dh - (dh_FAC + -0.23 dh_SMB)	0.115	0.19	0.29
high	strong melt	dh - (0.46 dh_FAC + 0.88 dh_SMB)	0.021	0.113	0.75
low	strong melt	dh	-0.348	0.648	0.00
low	strong melt	dh - dh_m	0.162	0.35	0.71
low	strong melt	dh - (0.57 dh_FAC + dh_SMB)	0.059	0.299	0.79
low	strong melt	dh - (dh_FAC + 0.89 dh_SMB)	0.133	0.346	0.72
low	strong melt	dh - (0.58 dh_FAC + 0.93 dh_SMB)	0.042	0.296	0.79

Table S1: Statistics of data and residuals to corrections based on the MARv3.11.5 model. Mean and standard deviation are calculated based on the inverse-point-density weight. R^2 is calculated relative to the full dataset for each subsample. The elevation and subset columns indicate the subsample of data for which the statistics were calculated.

Elev.	subset	GSFCv1.1 correction	mean	std	R2
all	all	dh	-0.024	0.246	0.00
all	all	dh - dh_m	0.012	0.18	0.46
all	all	dh - 0.66 dh_m	0	0.151	0.62
all	sp-su 2019	dh	-0.213	0.368	0.00
all	sp-su 2019	dh - dh_m	0.102	0.224	0.63
all	sp-su 2019	dh - 0.76 dh_m	0.025	0.2	0.70
all	sp-su 2020	dh	-0.027	0.216	0.00
all	sp-su 2020	dh - dh_m	-0.003	0.164	0.43
all	sp-su 2020	dh - 0.66 dh_m	-0.011	0.139	0.59
high	weak melt	dh	0.006	0.12	0.00
high	weak melt	dh - dh_m	-0.012	0.092	0.41
high	weak melt	dh - 0.79 dh_m	-0.008	0.09	0.44
low	weak melt	dh	0.07	0.252	0.00
low	weak melt	dh - dh_m	-0.083	0.188	0.44
low	weak melt	dh - 0.74 dh_m	-0.043	0.177	0.51
high	strong melt	dh	-0.088	0.228	0.00
high	strong melt	dh - dh_m	0.168	0.183	0.36
high	strong melt	dh - 0.59 dh_m	0.063	0.125	0.70
low	strong melt	dh	-0.339	0.652	0.00
low	strong melt	dh - dh_m	0.05	0.341	0.73
low	strong melt	dh - 0.98 dh_m	0.043	0.341	0.73
all	all	dh	-0.024	0.246	0.00
all	all	dh - dh_m	0.012	0.18	0.46
all	all	dh - (0.49 dh_FAC + dh_SMB)	0.001	0.148	0.64
all	all	dh - (dh_FAC + 0.59 dh_SMB)	0.006	0.17	0.52
all	all	dh - (0.52 dh_FAC + 0.90 dh_SMB)	0.001	0.147	0.64
high	weak melt	dh	0.006	0.12	0.00
high	weak melt	dh - dh_m	-0.012	0.092	0.41
high	weak melt	dh - (0.58 dh_FAC + dh_SMB)	-0.002	0.086	0.48
high	weak melt	dh - (dh_FAC + 1.34 dh_SMB)	-0.01	0.092	0.42
high	weak melt	dh - (0.31 dh_FAC + 2.18 dh_SMB)	0.013	0.079	0.57
low	weak melt	dh	0.07	0.252	0.00
low	weak melt	dh - dh_m	-0.083	0.188	0.44
low	weak melt	dh - (0.63 dh_FAC + dh_SMB)	-0.047	0.18	0.49
low	weak melt	dh - (dh_FAC + 0.59 dh_SMB)	-0.06	0.18	0.49
low	weak melt	dh - (0.76 dh_FAC + 0.72 dh_SMB)	-0.044	0.177	0.51
high	strong melt	dh	-0.088	0.228	0.00
high	strong melt	dh - dh_m	0.168	0.183	0.36
high	strong melt	dh - (0.51 dh_FAC + dh_SMB)	0.058	0.128	0.69
high	strong melt	dh - (dh_FAC + -0.04 dh_SMB)	0.135	0.157	0.52
high	strong melt	dh - (0.57 dh_FAC + 0.67 dh_SMB)	0.061	0.125	0.70
low	strong melt	dh	-0.339	0.652	0.00
low	strong melt	dh - dh_m	0.05	0.341	0.73
low	strong melt	dh - (0.74 dh_FAC + dh_SMB)	0.003	0.331	0.74
low	strong melt	dh - (dh_FAC + 1.13 dh_SMB)	0.076	0.337	0.73
low	strong melt	dh - (0.67 dh_FAC + 1.20 dh_SMB)	0.033	0.323	0.76

Table S2. Statistics of data and residuals to corrections based on the GSFCv1.1 model. Columns and subsets are the same as those in table S1.

Elev.	subset	correction	GSFCv1.2 mean	std	R2
all	all	dh	-0.024	0.246	0.00
all	all	dh - dh_m	0.007	0.147	0.64
all	all	dh - 0.88 dh_m	0.003	0.144	0.66
all	sp-su 2019	dh	-0.213	0.368	0.00
all	sp-su 2019	dh - dh_m	0.028	0.172	0.78
all	sp-su 2019	dh - 0.96 dh_m	0.017	0.172	0.78
all	sp-su 2020	dh	-0.027	0.216	0.00
all	sp-su 2020	dh - dh_m	-0.012	0.135	0.61
all	sp-su 2020	dh - 0.89 dh_m	-0.014	0.133	0.62
high	weak melt	dh	0.002	0.122	0.00
high	weak melt	dh - dh_m	0.008	0.083	0.54
high	weak melt	dh - 1.07 dh_m	0.008	0.082	0.54
low	weak melt	dh	0.069	0.252	0.00
low	weak melt	dh - dh_m	-0.041	0.178	0.50
low	weak melt	dh - 0.91 dh_m	-0.031	0.177	0.50
high	strong melt	dh	-0.111	0.263	0.00
high	strong melt	dh - dh_m	0.072	0.148	0.68
high	strong melt	dh - 0.91 dh_m	0.056	0.146	0.69
low	strong melt	dh	-0.361	0.663	0.00
low	strong melt	dh - dh_m	-0.005	0.329	0.75
low	strong melt	dh - 1.13 dh_m	0.043	0.322	0.76
all	all	dh	-0.024	0.246	0.00
all	all	dh - dh_m	0.007	0.147	0.64
all	all	dh - (0.85 dh_FAC + dh_SMB)	0.005	0.145	0.65
all	all	dh - (dh_FAC + 0.80 dh_SMB)	0.004	0.144	0.66
all	all	dh - (0.93 dh_FAC + 0.83 dh_SMB)	0.004	0.144	0.66
high	weak melt	dh	0.002	0.122	0.00
high	weak melt	dh - dh_m	0.008	0.083	0.54
high	weak melt	dh - (0.97 dh_FAC + dh_SMB)	0.008	0.083	0.54
high	weak melt	dh - (dh_FAC + 1.47 dh_SMB)	0.012	0.081	0.56
high	weak melt	dh - (0.59 dh_FAC + 1.91 dh_SMB)	0.016	0.079	0.58
low	weak melt	dh	0.069	0.252	0.00
low	weak melt	dh - dh_m	-0.041	0.178	0.50
low	weak melt	dh - (1.04 dh_FAC + dh_SMB)	-0.043	0.178	0.50
low	weak melt	dh - (dh_FAC + 0.78 dh_SMB)	-0.03	0.176	0.51
low	weak melt	dh - (1.26 dh_FAC + 0.68 dh_SMB)	-0.042	0.175	0.52
high	strong melt	dh	-0.111	0.263	0.00
high	strong melt	dh - dh_m	0.072	0.148	0.68
high	strong melt	dh - (0.92 dh_FAC + dh_SMB)	0.06	0.147	0.69
high	strong melt	dh - (dh_FAC + 0.65 dh_SMB)	0.062	0.144	0.70
high	strong melt	dh - (1.04 dh_FAC + 0.59 dh_SMB)	0.067	0.144	0.70
low	strong melt	dh	-0.361	0.663	0.00
low	strong melt	dh - dh_m	-0.005	0.329	0.75
low	strong melt	dh - (1.15 dh_FAC + dh_SMB)	0.017	0.328	0.76
low	strong melt	dh - (dh_FAC + 1.16 dh_SMB)	0.029	0.323	0.76
low	strong melt	dh - (1.09 dh_FAC + 1.15 dh_SMB)	0.04	0.322	0.76

Table S3. Statistics of data and residuals to corrections based on the GSFCv1.2 model. Columns and subsets are the same as those in table S1.

150

References:

Howat, I. (2017), *MEaSUREs Greenland Ice Mapping Project (GIMP) Land Ice and Ocean Classification*
155 *Mask, Version 1*, edited, Boulder, Colorado, USA. National Snow and Ice Data Center Distributed
Active Archive Center.
Joughin, I. (2020), *MEaSUREs Greenland Monthly Ice Sheet Velocity Mosaics from SAR and Landsat,*
Version 2 . edited, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active
Archive Center.
160 Joughin, I., B. Smith, I. Howat, and T. Scambos (2015, updated 2018.), *MEaSUREs Greenland Ice Sheet*
Velocity Map from InSAR Data, Version 2, edited, Boulder, Colorado USA. NASA National Snow and Ice
Data Center Distributed Active Archive Center.
Shepherd, A., et al. (2012), A Reconciled Estimate of Ice-Sheet Mass Balance, *Science*, *338*(6111), 1183-
1189.
165 Smith, B., S. Dickinson, B. P. Jelley, T. A. Neumann, D. Hancock, J. Lee, and K. Harbeck (2020),
ATLAS/ICESat-2 L3B Annual Land Ice Height, Version 1 . edited, NASA National Snow and Ice Data
Center Distributed Active Archive Center.