Supplemental materials to Van der Sluijs et al., *Allometric scaling of retrogressive thaw slumps*

S1: Additional elevation dataset descriptions

At the smallest geographic extent and within the Inuvik-to-Tuktoyaktuk (ITH) and Dempster Highway (DH) corridors, airborne stereo-photogrammetry and LiDAR surveys were acquired in 2011. For detailed descriptions of acquisition and data processing readers are referred to Van der Sluijs et al., (2018).

At a larger geographic extent airborne stereo-photogrammetry from 2004 was used to derive 1-m contour lines which were interpolated to 3-m hydrologically correct DEMs using ArcGIS 10.6.1 "Topo-to-Raster" tool (based on ANUDEM). The 2004 DEM tiles were compared to ITH and DH LiDAR elevations based on 10,000 samples per 10 km^2 tile where the datasets overlapped. This resulted in a R^2 of 0.98-0.99, mean differences of -0.91 to -1.06 m and RMSEs of 1.28-1.46 m (i.e., sub-pixel vertical uncertainties without correcting for vertical datum differences).

Finally, seamless mosaics of the pan-Arctic data product ArcticDEM were used to enable work at the largest geographical extent possible. The DEM has been developed by the Polar Geospatial Center, based on very high stereo-satellite imagery, and the most recent product (version 7) has undergone additional post-processing steps to ensure data consistency among tiles. The mosaiced dataset has a median acquisition year of 2016 for either study region (based on n=66 and 132 strips for PP and APTC, respectively). The mosaic tiles are the best currently available ArcticDEM data, yet they are missing patches of landscape due to cloud cover or lack of stereo-coverage. The imagery used for the regional scale ortho-mosaic was sourced from the ESRI World Imagery Layer. Metadata for this base imagery indicated a median acquisition year of 2017 for the blended tiles (n = 39 and 20 for ATPC and PP, respectively).

The DEMs were not differenced to obtain episodal estimates of thawed volume, avoiding common challenges such as positional accuracy, co-registration, resolution differences as well as vertical datum differences.

Table S1 Overview of DEM datasets

DEM	Year	Technology	Spatial Resolution	Hor./ Vert. accuracy	References
MVAP [1]	2004	Airborne stereo-photogrammetry	3 m	± 1.5 m	NWT Centre for Geomatics; Lantz and Kokelj (2008); Segal et al., 2016)
LiDAR	2011	LiDAR	1 m	± < 0.2 m	Van der Sluijs et al., (2018)
ArcticDEM [2]	2016 (circa)	Satellite stereo-photogrammetry	2 m	n/a	Porter et al., (2018)

[1] Mackenzie Valley Airphoto Project (MVAP) of Indian and Northern Affairs Canada, now Government of Northwest Territories. The MVAP dataset in this study represents a small subset of the entire dataset.

[2] DEM(s) were created from DigitalGlobe, Inc., imagery and funded under National Science Foundation awards 1043681, 1559691, and 1542736.

S2: Multisource Slump Inventory delineation and harmonization techniques

Several slump delineation and harmonization techniques were carried out to acquire temporally consistent boundaries for each digitized slump and to reduce the chance of undetected small or inactive slumps.

Firstly, all raster and vector data were reprojected to the NAD83 CSRS UTM Zone 8N coordinate system. The 2004 and 2011 ortho-mosaics and hillshade DEMs were used to identify the stable periphery of scar zones, indicated by the presence of tall shrubs (in orthophotos) and sparse ground-classified points (in LiDAR point cloud data), and the extents of these boundary polygons were confirmed among available data layers.

The previously published delineations of slump scar zones (e.g., Segal et al., 2016) were updated to reflect headwall positions relative to hillshade DEMs (modern or ancient) and to ensure subsequent slump delineations could not be of a smaller extent than the previous observation. Through this validation process, any detected increase in scar area in subsequent epochs could be more confidently attributed to active headwall retreat and not digitization differences. In other words, circa 2004 delineations were copied in the dataset and attributed as year = 2011, and only spatially updated in cases of active headwall retreat in the hillshade DEM. Additional slumps discovered in the 2011 datasets were added to the 2004 inventory if these occurred in previously unnoticed stabilized or ancient slump scars. Similarly, the 2004/2011 delineations were copied and attributed as year = 2016, and only spatially updated in cases of active headwall retreat using the circa 2016 ArcticDEM and ESRI basemap imagery. Additional slumps discovered in the 2016 datasets were added to the 2004 and 2011 features if these occurred in previously unnoticed stabilized or ancient slump scars.

Lastly an Northwest Territories-wide change detection product based on Landsat Tasseled Cap trend analysis spanning 1985-2019 (based on methods by Fraser et al., 2014) was consulted to identify any additional slump locations previously undetected. Upon manual discovery of these slump locations the three time periods of data (2004, 2011 and 2016) were assessed to delineate and attribute the slump features accordingly. Together these harmonization procedures captured previously undocumented small or inactive slumps and ensured slump delineations between hillshade DEMs were acquired with temporal consistency (i.e., slump-affected disturbed areas can only increase through time).

For the PP study area the MSI procedures were the same for 2011 (LiDAR) and 2016 (ArcticDEM), whereby additional slumps discovered using the ArcticDEM dataset were added retroactively to the 2011 population (within the LiDAR confines). No circa 2004 high-resolution dataset existed for the PP study area, and the two DEM datasets used had varying spatial coverage (e.g., LiDAR was restricted to within 9 km of Dempster Highway; ArcticDEM contained occasional data voids; manuscript Fig. 2).

S3: Interpolation methods

Table S2: Overview of interpolation methods and their performance to reconstruct pre-erosional terrain.

ID	Suite	Main Parameter	Abr.	RMSD (m)			
				Median	IQR	90th %	95th%
1	IDW	Power of 0.5 [1]	IDW05	0.82	1.00	2.80	4.04
2	IDW	Power of 1.0 [1]	IDW10	0.82	1.00	2.80	4.04
3	IDW	Power of 1.5 [1]	IDW15	0.82	1.00	2.80	4.04
4	IDW	Power of 2.0 [1]	IDW20	0.82	1.00	2.80	4.04
	IDW			0.82	1.00	2.80	4.04
5	**TR**	**Linear**	**LIN**	**0.26**	**0.39**	**1.02**	**1.42**
6	**TR**	**Natural Neighbour**	**NN**	**0.28**	**0.38**	**1.03**	**1.40**
	TR			**0.27**	**0.38**	**1.03**	**1.42**
7	RSL	Weight of 0.000 [2]	RSL0	0.55	0.73	2.16	3.30
8	RSL	Weight of 0.001 [2]	RSL0001	0.55	0.71	2.13	2.98
9	RSL	Weight of 0.005 [2]	RSL0005	0.57	0.74	2.17	3.14
10	RSL	Weight of 0.01 [2]	RSL001	0.60	0.77	2.28	3.32
11	RSL	Weight of 0.05 [2]	RSL005	0.76	0.99	3.07	4.44
12	RSL	Weight of 0.10 [2]	RSL01	0.89	1.13	3.31	4.67
	RSL			0.64	0.86	2.52	3.61
13	RSH	Weight of 0.33 [2]	RSH03	1.21	1.49	4.56	6.37
14	RSH	Weight of 0.50 [2]	RSH05	1.36	1.69	5.06	7.19
15	RSH	Weight of 1.00 [2]	RSH1	1.64	2.09	6.21	8.83
16	RSH	Weight of 5.00 [2]	RSH5	2.24	3.08	8.85	12.65
	RSH			1.55	2.10	6.15	8.82
17	ST	Weight of 0.0 [3]	ST0	66.49	200.62	646.19	1332.49
18	ST	Weight of 1.0 [3]	ST1	0.72	0.98	2.89	4.06
19	ST	Weight of 3.0 [3]	ST3	0.75	1.00	2.90	4.07
20	ST	Weight of 5.0 [3]	ST5	0.76	1.01	2.91	4.08
21	ST	Weight of 7.0 [3]	ST7	0.77	1.01	2.91	4.08
22	ST	Weight of 9.0 [3]	ST9	0.77	1.00	2.91	4.08
23	ST	Weight of 10.0 [3]	ST10	0.78	1.01	2.91	4.08
	ST			0.93	2.03	26.90	136.30
24	EBK [4, 5]	No transform, Power	EBK-POW	0.51	0.73	2.13	2.98
25	EBK	No transform, Linear	EBK-LIN	0.71	0.95	2.70	3.93
26	EBK	No transform, Thin Plate Spline	EBK-TPS	0.55	0.80	2.49	3.47
	EBK			0.59	0.84	2.44	3.41
27	EBK-EMP	Empirical transform, Exponential	EBK-EXP	0.84	1.09	2.97	4.27
28	EBK-EMP	Empirical transform, Whittle	EBK-WHIT	0.81	1.08	2.97	4.28
29	EBK-EMP	Empirical transform, K-Bessel	EBK-KBES	0.79	1.12	2.98	4.30
	EBK-EMP			0.81	1.09	2.98	4.28
30	EBK-EMPD	Empirical transform, Exponential detrended	EBK-EXPD	0.38	0.57	1.71	2.46
31	EBK-EMPD	Empirical transform, Whittle detrended	EBK-WHITD	0.38	0.60	1.75	2.50
32	EBK-EMPD	Empirical transform, K-Bessel detrended	EBK-KBESD	0.38	0.60	1.77	2.61
	EBK-EMPD			0.38	0.59	1.75	2.54

[1] Power describes the exponent of distance parameter, which controls the significance of surrounding points on the interpolated value. A higher power results in less influence from distant points.
[2] Weight is the square of the parameter referred to as tau (t). Higher values produce smoother surfaces (typical range: 0 to 0.5).
[3] Weight is the square of the parameter referred to as phi (Φ). Higher values produce coarser surfaces and more closely conform to input points (typical range: 0 to 10).
[4] Empirical Bayesian Kriging (EBK) was selected as an alternative over other kriging methods (e.g., ordinary kriging) to overcome challenges in the proper manual selection of the type of theoretical variogram model and semivariogram parameters.
[5] EBK offers both settings for data transformation (*none* or Multiplicative Skewing transformation with *Empirical* base function) and various *semivariogram* models.

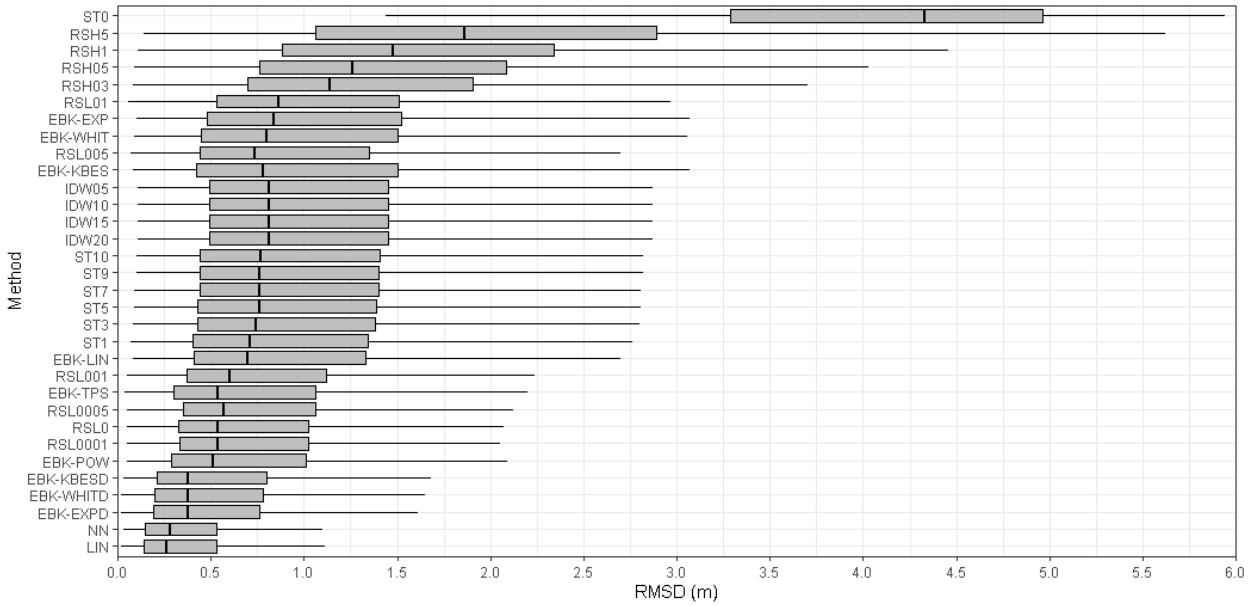

Figure S3: Boxplot of deviations between actual and modelled elevation (expressed as Root Mean Square Difference; RMSD) for each interpolation method. Due to outliers in ST with a weight of zero (ST0) the X-axis was limited to 6 m (representing triple the upper quartile (75th percentile) of the entire dataset), removing 1,668 model observations from presentation.

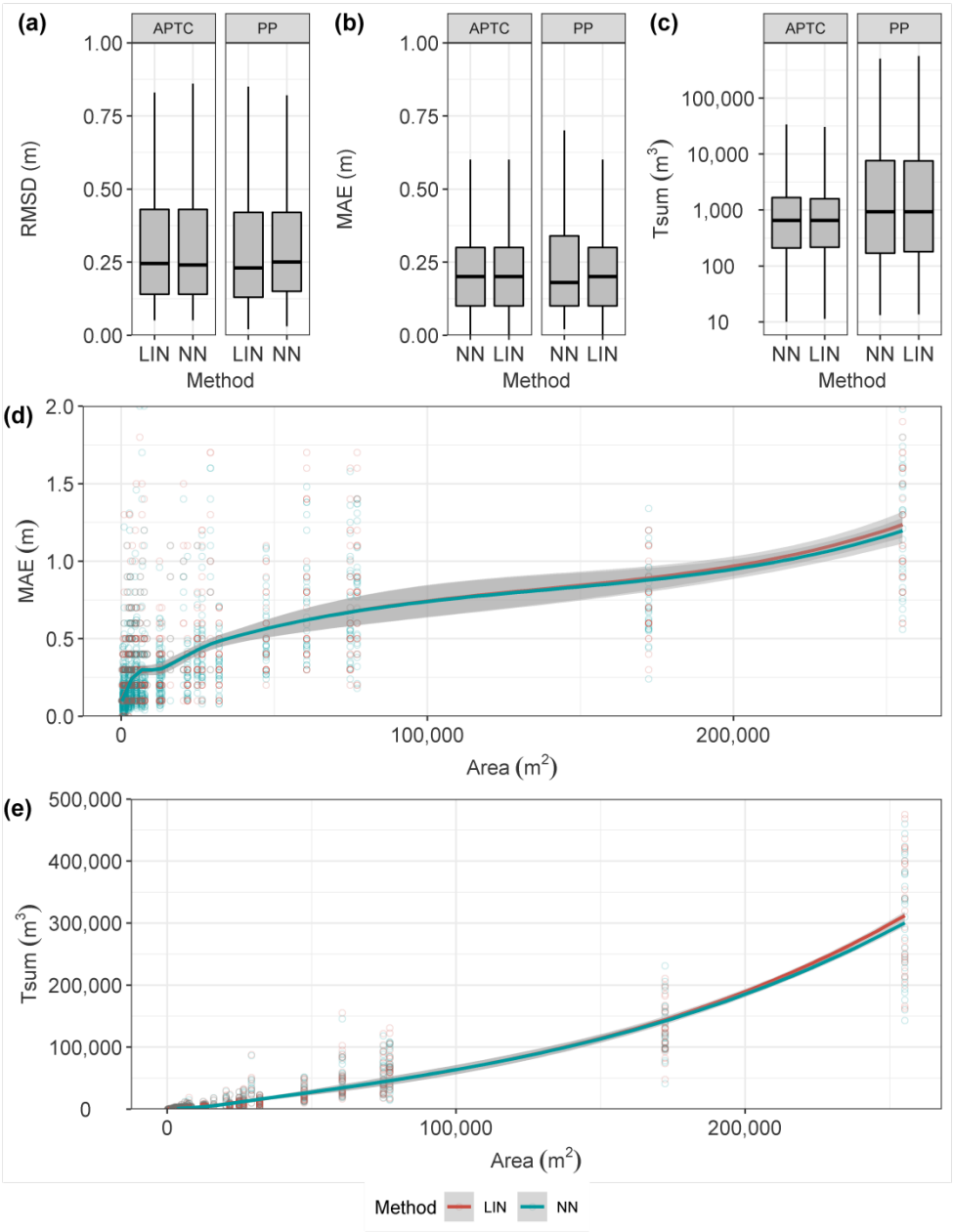

Figure S4: Deviations between actual and modelled elevation for linear (LIN) and natural neighbour interpolation (NN), expressed as boxplots based on Root Mean Square Difference (a), Mean Absolute Error (b) and summed topographic difference (c) along with two scatterplots between void surface area and MAE (d) and T_{sum} (e), respectively, with loess smoothing lines to highlight the degree to which the interpolation methods exhibit area-dependent uncertainty.