Supplementary Materials

Assmann et al. - Freshwater ponds create highly dynamic Arctic tundra landscapes

Jakob J. Assmann[1], Cengiz Akandil[1], Elena Plekhanova[2], Alizée Le Moigne[3], Sergey V Karsanaev[4], Trofim C. Maximov[4], Gabriela Schaepman-Strub[1]

[1] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Switzerland
[2] Swiss Federal Research Institute WSL, Birmensdorf, Switzerland
[3] Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, Québec, Canada
[4] Institute for Biological Problems of the Cryolithozone, Siberian Branch Russian Academy of Sciences, Yakutsk, Russia

Content

Methods

Climate Data

The climate data presented in Figs. S1 and S2 is from the Chokurdakh weather station (WMO 21946) at 30 km distance from the study sites. We retrieved climate data (daily temperature and precipitation) for the station for the study period from http://www.pogodaiklimat.ru (accessed 28 November 2022 17:00 CET) and then aggregated the variables to monthly and annual measures (mean temperature, precipitation sum). In addition, we calculated summary statistics (mean temperature, precipitation sum) for each year based on the following hydrological seasons. Autumn of the previous year: September, October and November of the preceding year (snow fall). Winter: December of the preceding year, as well as January and February of the current year. Spring: March, April and May of the current year. Summer: June, July and August of the current year.

Determination of site boundaries

We determined the exact boundaries for each site based on the minimum shared extent of the drone imagery available for all surveys above each site location. We reduced this minimum shared extent by a further buffer of 50 m ("high" and "med") and 60 m ("low) to ensure the resulting cropped area had a consistent overlap of at least five individual drone images throughout the area. This allowed us to avoid edge effects and artefacts in the final drone mosaics and digital surface models generated by low image density. Corner point coordinates of the resulting site boundaries are provided in Table S1.

Geolocation and accuracy assessment

We followed the procedure for time-series alignment in the absence of ground control points recommended by the software manufacturer Pix4D (https://support.pix4d.com/hc/en-us/articles/204373409, last accessed 12 July 2024). As such we designated the 2021 survey for each site as the reference project as high-precision dGNSS geolocation was available for these. We then co-aligned all preceding years using a minimum of five manual tie-points in Pix4D.

To assess the accuracy of our alignments, we exported the RGB mosaics from Pix4D at their native resolution (see Table S2) and marked an additional five control points in each mosaic. We then calculated the absolute distance of each control point to their corresponding reference point in the 2021 mosaic. The mean distance for all control points to their corresponding reference point did not exceed 48 cm in any survey (Fig. S6). This distance is equivalent to three-times the ground sampling distance of the aggregated products (12 cm) used in the subsequent analysis. We believe that this represents a realistically achievable accuracy given the constraints of our methods and data quality of the early drone surveys.

Training annotations & accuracy assessment

We generated the manual annotations based on the RGB drone imagery in QGIS v. 3.36.2 (QGIS Development Team 2024). We started using a systematic approach, and then added additional training annotations in areas where the classification by thresholding performed badly, optimising the classifier iteratively. To this end, we first placed a regular grid (50 m x 50 m cells) over the area of interest of each site (origin = centre point) and then selected 6 random cells from the grid. In these cells we annotated one whole pond (class "water") and one area containing other surfaces (class "other") of varying sizes. Sometimes these randomly chosen cells did not contain any ponds. In such cases, we still annotated a polygon in the "other" class in this cell and then sampled another cell at random where we annotated both "water" and "other". Finally, we iteratively added additional polygons in areas where the classifier performed badly following visual inspection.

The resulting site- and year-specific training datasets were not balanced between the two surface cover classes ("water" and "other"), neither in terms of polygons per class, nor in terms

of the number of pixels per class (Table S3). Despite the unbalanced sample, we decided to determine the optimal BCC threshold for each site-year combination based on the overall accuracy of predictions using all annotated pixels (Table S4) rather than, for example, maximising sensitivity and specificity, as this approach yielded the best results based on visual inspection of the resulting surface water maps. To allow for a meaningful comparison of the classification accuracies amongst the site-year combinations, we further obtained a standardised random subsample of 10000 pixels per class for each site-year combination. We then split this sample into 8 random groups of 1250 pixels per class - the approximate number of pixels in an average-sized pond - and carried out a “leave-one-out” accuracy assessment to test for the robustness of the accuracy estimates in respect to the sample. Here, we calculated the mean accuracy, sensitivity and specificity of the surface water classifications for all possible combinations of seven groups while withholding one group. The results of this analysis indicate a very good performance of the classifications, exceeding 90% in all cases, except for the two 2018 surveys for the high and medium sites, which had an accuracy of 87% (Table S4).

Surface water trend and climate analysis

We tested for trends in the surface water for each site using ordinary least squares regression (OLS). We conducted these analyses in R, fitting the simple linear models using the lm() function. For each time-series, we fitted a model with the calendar year as predictor and the surface water proportion as the response variable. As OLS can be highly sensitive to outliers we also fitted the same models excluding the surface water proportions from 2017. Based on the initial visual assessments of the data, we did not suspect any trends to be present in the time-series. However, as surface water trends are commonly reported in the literature, we believe that this is part of the analysis and might nonetheless be of interest to the reader. Estimated model coefficients are shown in Table S6, S7, S8, S9, S10 and S11.

To test for potential climatic drivers of landscape surface water area, we calculated the correlation coefficients (Pearson's r) between the surface water area and the seasonal climatic variables derived from the climate data (see section “Climate Data” above). We did not detect any significant correlations (Table S12). However, the time-series are very short and contain very few observations. We therefore urge the reader to take caution when interpreting the resulting correlation coefficients and associated p-values.

References

QGIS Development Team, 2024. QGIS Geographic Information System. QGIS Association.

Tables

Table S1: Coordinates for the corner points of the three study sites (CRS: EPSG 4326).

Site	Corner Point	Latitude	Longitude
high	Point 1	70.8305	147.4711
	Point 2	70.8340	147.4712
	Point 3	70.8340	147.4848
	Point 4	70.8304	147.4847
med	Point 1	70.8314	147.4888
	Point 2	70.8357	147.4889
	Point 3	70.8357	147.5014
	Point 4	70.8314	147.5012
low	Point 1	70.8274	147.4573
	Point 2	70.8294	147.4574
	Point 3	70.8313	147.4637
	Point 4	70.8313	147.4680
	Point 5	70.8294	147.4705
	Point 6	70.8274	147.4675

(section continued on next page)

Table S2: Meta data for the drone surveys. Columns indicate the study site, name of processed mosaic (cbh = cloudberry hill site, tlb = drained thaw lake bed site, rdg = ridge site), associated flights dates and take-off times, drone and sensor types, geolocation methods, average ground sampling distance (GSD) in the point cloud (in cm), whether image normalisation was applied during pre-processing, as well as weather conditions (where known).

Site	Mosaic ID	Flight Date(s)	Take Off	Drone	Sensor	Geolocation	GSD (cm)	Image Pre-Processing	Weather
high	cbh_2014	2014-08-10	12:40	eBee Classic	SenseFly S110 RGB	Ordinary GPS	4.36	-	Unknown.
	cbh_2016	2016-08-12	17:06	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.02	Yes	No clouds
		2016-08-12	17:35						Partially cloudy.
		2016-08-18	10:36						Broken cloud, gusty, light winds.
		2016-08-18	11:12						Cloudy, light winds with gusts.
	cbh_2017	2017-07-16	18:00	eBee Classic	SenseFly S110 RGB	Ordinary GPS	4.14	-	Overcast, variable, broken cloud, wind 5.1-6.2 m/s.
	cbh_2018	2018-07-20	09:30	eBee Classic	SenseFly S110 RGB	Ordinary GPS	2.99	-	Cloud cover, cirrus, sometimes direct sun, strong winds at take off (8.7-9.2 m/s).
	cbh_2019	2019-07-12	12:08	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.77	-	Very light cirrus cover / haze, low sun.
	cbh_2019_b	2019-07-23	10:04	eBee Classic	SODA	Ordinary GPS	2.53	-	Unknown.
	cbh_2020	2020-07-24	18:10	eBee X	SODA	PPK	2.46	-	Obscured sun (haze/smoke from fires), light shadows wind 3 m/s.
	cbh_2021	2021-07-19	14:44	eBee X	SODA	PPK	1.82	-	Very little/no clouds wind 2.5 m/s.
med	tlb_2014	2014-07-21	16:15	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.98	-	Overcast, low cloud with some breaks, but appears to be steady over site.
	tlb_2016	2016-08-12	17:06	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.02	-	No clouds.
		2016-08-18	11:12					-	Broken cloud.
	tlb_2017	2017-07-13	16:05	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.54	-	Overcast, sun sometimes obscured, wind 7.8-8.3 m/s.
	tlb_2018	2018-07-16	16:07	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.03	-	Overcast, wind 4.7 m/s, gusts 7.1 m/s.
	tlb_2019_a	2019-07-11	NA	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.77	-	Unknown.
	tlb_2019_b	2019-07-17	NA	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.38	-	Overcast, broken clouds.
	tlb_2019_c	2019-07-23	NA	eBee Classic	SODA	Ordinary GPS	2.53	-	Overcast, broken clouds.
	tlb_2020	2020-07-24	17:25	eBee X	SODA	PPK	2.45	-	Clear Sky above working area, Wind 3.5 m/s.
	tlb_2021	2021-07-19	13:53	eBee X	SODA	PPK	2.00	-	Unknown.
low	rdg_2014	2014-07-23	17:15	eBee Classic	SenseFly S110 RGB	Ordinary GPS	4.15	-	Clear sky, otherwise unknown.
	rdg_2017_b	2017-07-16	NA	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.13	Yes	Mostly cloudy, but some sunny spots in between.
	rdg_2018	2018-07-15	13:50	eBee Classic	SenseFly S110 RGB	Ordinary GPS	2.59	Yes	Unknown.
		2018-07-15	14:50						
	rdg_2019_a	2019-07-12	NA	eBee Classic	SenseFly S110 RGB	Ordinary GPS	3.37	-	Light cloud, cirrus, inferred from accompanying photographs.
	rdg_2020	2020-07-24	16:20	eBee X	SODA	PPK	2.22	-	Obscured sun (haze/smoke from fires), light shadows wind 2 m/s.
	rdg_2021	2021-07-19	15:40	eBee X	SODA	PPK	2.04	-	No clouds, some wind (around 3 m/s ? - based on other flights from same day).

Table S3: Overview of the training annotations, for each site-year combination (Mosaic ID), including the number of annotated polygons in each class (“other” and “water”) and the number of pixels in all polygons combined.

Site	Mosaic ID	Number of polygons		Number of pixels	
		other	water	other	water
high	cbh_2014	25	17	244909	47958
	cbh_2016	10	9	137182	30952
	cbh_2017	6	9	63151	61465
	cbh_2018	10	8	197400	22509
	cbh_2019	7	7	172343	30762
	cbh_2019_b	9	7	147093	13077
	cbh_2020	6	6	90717	12944
	cbh_2021	7	7	159549	24018
medium	tlb_2014	6	7	86887	22983
	tlb_2016	6	7	146462	12715
	tlb_2017	6	6	97944	27373
	tlb_2018	8	6	122126	10428
	tlb_2019_a	10	8	203771	16344
	tlb_2019_b	7	7	128885	26310
	tlb_2019_c	6	5	101922	23957
	tlb_2020	6	6	72247	19986
	tlb_2021	7	6	201113	16151
low	rdg_2014	6	-	598757	-
	rdg_2017_b	6	1	522567	1160
	rdg_2018	6	-	645656	-
	rdg_2019_a	6	-	464478	-
	rdg_2020	6	-	307134	-
	rdg_2021	6	-	503990	-

Table S4: Classification thresholds for the surface water detection and associated accuracy, sensitivity and specificity values based on the annotated training data for each RGB drone mosaic. A blue chromatic coordinate (BCC) threshold of 1.00 indicates that no water was present in the drone mosaic, for these mosaics accuracy information is not available (NA). Accuracy, sensitivity and specificity values are given twice, once for all training pixels available (not balanced between classes) and for a standardised sample of 10000 pixels per class. The latter sample was further split into 8 groups and accuracy, sensitivity and specificity were determined as the mean values using leave-one-out validation on the 8 groups. For further details, see section “Training annotations & accuracy assessment” above.

Site	Mosaic ID	BCC threshold	Accuracy (all)	Sensitivity (all)	Specificity (all)	Mean Accuracy (sample)	Std. Err. Accuracy (sample)	Mean Sensitivity (sample)	Mean Specificity (sample)
high	cbh_2014	0.38	0.98	0.90	1.00	0.95	0.00010	0.89	1.00
	cbh_2016	0.39	0.96	0.91	0.97	0.94	0.00008	0.90	0.97
	cbh_2017	0.34	0.98	0.97	1.00	0.98	0.00003	0.97	1.00
	cbh_2018	0.38	0.97	0.75	0.99	0.87	0.00013	0.75	0.99
	cbh_2019	0.41	0.97	0.85	1.00	0.93	0.00003	0.85	1.00
	cbh_2019_b	0.42	0.99	0.88	1.00	0.94	0.00006	0.88	1.00
	cbh_2020	0.40	0.98	0.86	1.00	0.93	0.00010	0.86	1.00
	cbh_2021	0.41	0.96	0.77	0.99	0.88	0.00005	0.77	0.99
med	tlb_2014	0.41	0.99	0.97	1.00	0.99	0.00005	0.98	1.00
	tlb_2016	0.43	0.99	0.92	0.99	0.96	0.00004	0.92	0.99
	tlb_2017	0.40	0.99	0.98	0.99	0.99	0.00005	0.98	0.99
	tlb_2018	0.41	0.98	0.74	1.00	0.87	0.00011	0.74	1.00
	tlb_2019_a	0.44	0.99	0.91	1.00	0.96	0.00005	0.91	1.00
	tlb_2019_b	0.41	0.98	0.92	0.99	0.96	0.00005	0.92	0.99
	tlb_2019_c	0.40	0.99	0.95	1.00	0.97	0.00007	0.95	1.00
	tlb_2020	0.42	0.99	0.97	1.00	0.98	0.00003	0.97	1.00
	tlb_2021	0.40	1.00	0.97	1.00	0.98	0.00002	0.97	1.00
low	rdg_2014	1.00	-	-	-	-	-	-	-
	rdg_2017_b	0.36	1.00	0.97	1.00	-	-	-	-
	rdg_2018	1.00	-	-	-	-	-	-	-
	rdg_2019_a	1.00	-	-	-	-	-	-	-
	rdg_2020	1.00	-	-	-	-	-	-	-
	rdg_2021	1.00	-	-	-	-	-	-	-

Table S5: Surface water area (m^2) and surface water proportion (%) for each classified drone mosaic.

Site	Mosaic ID	Calendar Year	Surface Water Area (m^2)	Surface Water Proportion (%)
high	cbh_2014	2014	7746	4
	cbh_2016	2016	4498	2
	cbh_2017	2017	36644	19
	cbh_2018	2018	8106	4
	cbh_2019	2019	2820	1
	cbh_2019_b	2019	3203	2
	cbh_2020	2020	4079	2
	cbh_2021	2021	3769	2
med	tlb_2014	2014	1682	1
	tlb_2016	2016	1080	0
	tlb_2017	2017	7204	3
	tlb_2018	2018	1864	1
	tlb_2019_a	2019	998	0
	tlb_2019_b	2019	1573	1
	tlb_2019_c	2019	1334	1
	tlb_2020	2020	1159	1
	tlb_2021	2021	1060	0
low	rdg_2014	2014	0	0
	rdg_2017_b	2017	19	0
	rdg_2018	2018	0	0
	rdg_2019_a	2019	0	0
	rdg_2020	2020	0	0
	rdg_2021	2021	0	0

Table S6: We found no significant trend in the surface water proportion time-series of the **high site** including all surveyed years between 2014 and 2021. The table shows the statistics for the model coefficients (intercept and year predictor).

Term	Parameter Estimate	Std.error	t-Statistic	p.value
Intercept	14.558	20.118	0.724	0.497
Year	-0.007	0.010	-0.721	0.498

Table S7: We found no significant trend in the surface water proportion (%) time-series of the **medium site** including all surveyed years between 2014 and 2021. The table shows the statistics for the OLS model intercept and year predictor of the time-series.

Term	Parameter Estimate	Std.error	t-Statistic	p.value
Intercept	2.202	3.133	0.703	0.505
Year	-0.001	0.002	-0.700	0.507

Table S8: We found no significant trend in the surface water proportion (%) time-series of the **low site** including all surveyed years between 2014 and 2021. The table shows the statistics for the OLS model intercept and year predictor of the time-series.

Term	Parameter Estimate	Std.error	t-Statistic	p.value
Intercept	0.009	0.018	0.474	0.660
Year	0.000	0.000	-0.473	0.661

Table S9: We found no significant trend in the surface water proportion (%) time-series of the **high site** including all surveyed years between 2014 and 2021, **except the outlier year 2017**. The table shows the statistics for the OLS model intercept and year predictor of the time-series.

Term	Parameter Estimate	Std.error	t-Statistic	p.value
Intercept	5.706	3.184	1.792	0.133
Year	-0.003	0.002	-1.785	0.134

Table S10: We found no significant trend in the surface water proportion time-series of the **medium site** including all surveyed years between 2014 and 2021, **except the outlier year 2017**. The table shows the statistics for the OLS model intercept and year predictor of the time-series.

Term	Parameter Estimate	Std.error	t-Statistic	p.value
Intercept	0.577	0.498	1.159	0.291
Year	0.000	0.000	-1.146	0.295

Table S11: We found no significant trend in the surface water proportion (%) time-series of the **low site** including all surveyed years between 2014 and 2021, **except the outlier year 2017**. The table shows the statistics for the OLS model intercept and year predictor of the time-series.

Term	Parameter Estimate	Std.error	t-Statistic	p.value
Intercept	0.009	0.018	0.474	0.660
Year	0.000	0.000	-0.473	0.661

Table S12: Correlations (Pearson's r) between the water area and climate variables were not significant for all climate variable and site combinations tested. Climate variables are named as follows: "temp_*_mean" indicates mean temperatures across a season (*), "precip_*_sum" indicate precipitation sums across a season (*). Seasons are indicated by the respective initial letters of the months included. Winter and autumn included months from the preceding years - See Supplementary Methods - Climate Data for details. In addition to the seasonal variables, we also tested the mean temperature and precipitation sums of the calendar year ("annual"). We did not test correlations for the low site, as water was only detected there in 2017.

Site	Climate Variable	Pearson's r	p-value
high	temp_jja_mean	-0.72	0.068
	precip_son_sum	0.56	0.187
	temp_son_mean	0.46	0.302
	temp_mam_mean	0.44	0.322
	precip_jja_sum	-0.41	0.363
	precip_dfj_sum	0.34	0.462
	precip_mam_sum	0.13	0.783
	precip_annual_sum	-0.12	0.794
	temp_annual_mean	0.07	0.877
	temp_djf_mean	0.06	0.893
medium	temp_jja_mean	-0.7	0.081
	precip_son_sum	0.51	0.239
	temp_son_mean	0.49	0.267
	temp_mam_mean	0.47	0.282
	precip_jja_sum	-0.38	0.395
	precip_dfj_sum	0.28	0.540
	precip_annual_sum	-0.16	0.739
	precip_mam_sum	0.08	0.864
	temp_annual_mean	0.07	0.881
	temp_djf_mean	0.02	0.972

Figures

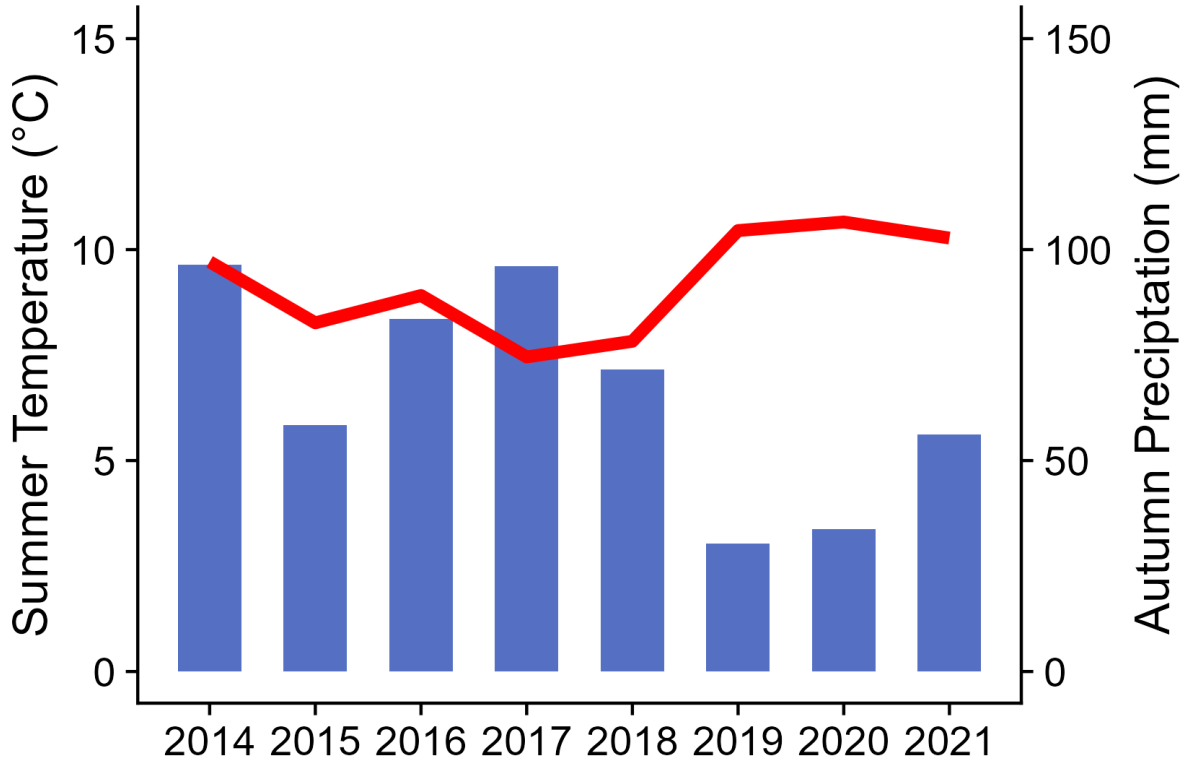

Figure S1: The variation in mean summer temperature (June-August - red line) and cumulative precipitation in autumn of the previous year (September-November - blue bars) for the time-period of our study (2014-2021). Data from the Chokurdakh Weather Station, 30 km distance from our study sites for the time-period 1956-2021 (source: http://www.pogodaiklimat.ru/). See Supplementary Methods - Climate Data for further details.

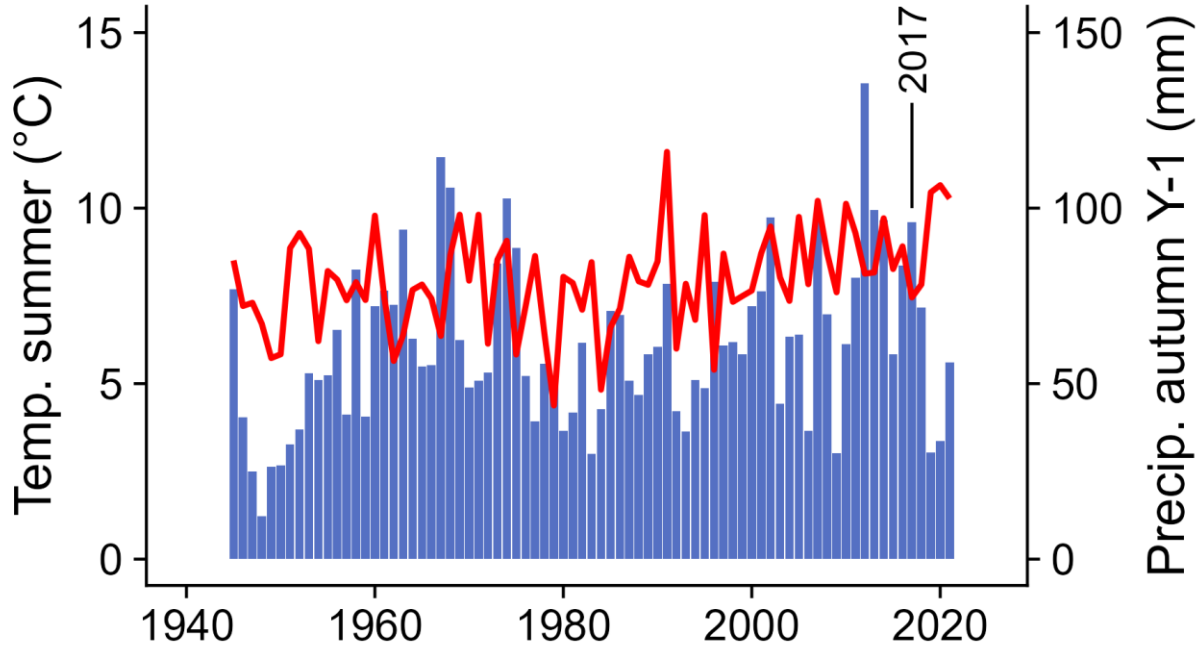

Figure S2: The variation in mean summer temperature (June-August - red line) and cumulative precipitation in autumn of the previous year (September-November - blue bars) during the time-period of our study (2014-2021) do not appear to be unusual compared to previous decades. Data from the Chokurdakh Weather Station, 30 km distance from our study sites for the time-period 1956-2021 (source: http://www.pogodaiklimat.ru/). See Supplementary Methods - Climate Data for further details.

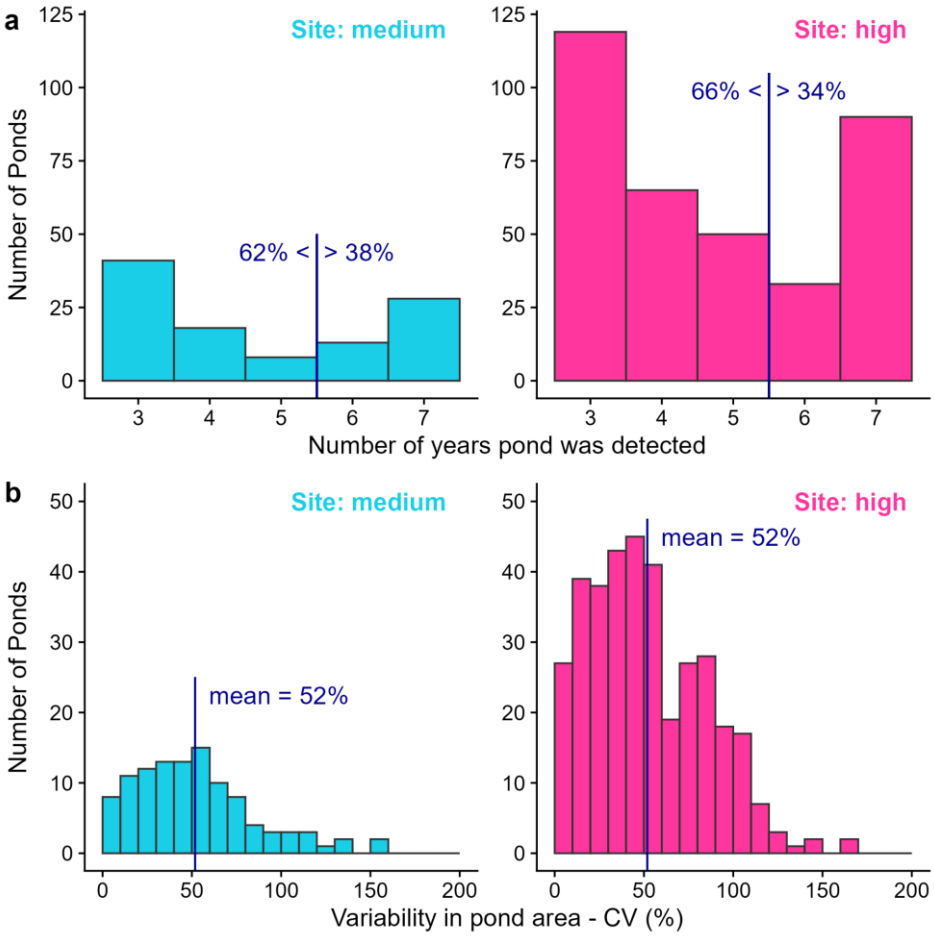

Figure S3: Ponds were short-lived and strongly fluctuated in size independent on the study site. The majority of ponds occurred for less than six years in the time-series (a) and individual ponds showed a high variability in their surface area across the time-series (b), the results were similar between the time-series of the two sites (medium = light blue, high = pink). Histograms of the detected pond-presence in each time-series (a). Here, the dark blue lines and annotations indicate the proportional split between those ponds present less than six years and those present "continuously" for six or more years (allowing for one year of failed detection). Histograms of the coefficient of variation (CV) for the pond area across each time-series are shown in (b). CV values for each pond time series were calculated excluding the extreme outlier year 2017. Here, the dark blue lines and annotations indicate the mean CV value of all ponds at the given site. In (a) and (b), we only show pond-time series where the pond occurred in more than three years (our threshold of detection), as no ponds fulfilled that condition at the "low" site, no data is shown for this site.

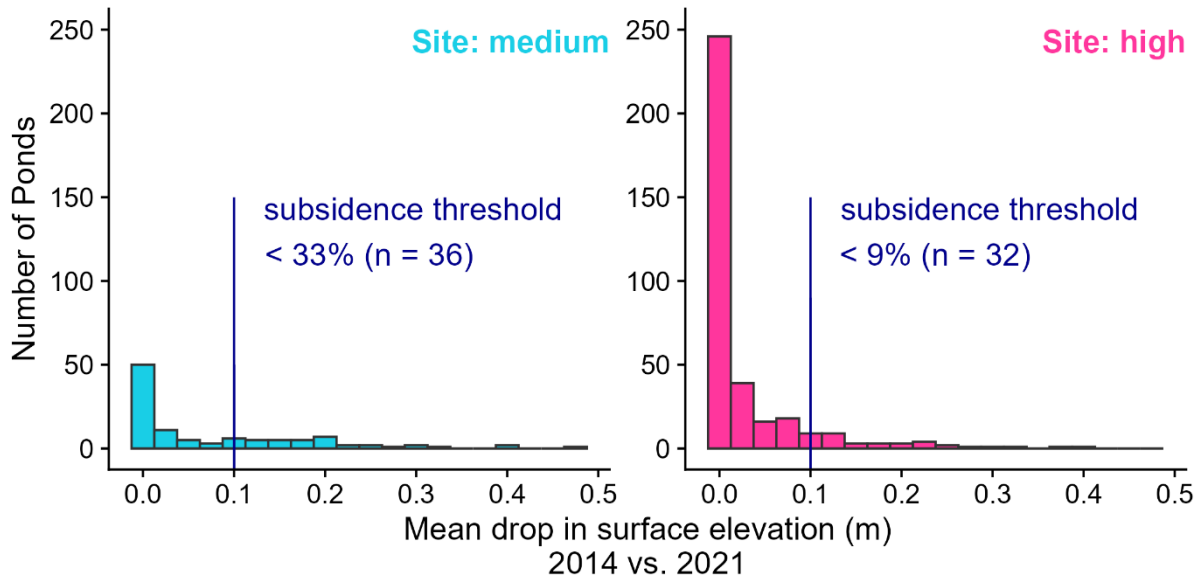

Figure S4: Thermokarst as a driver of pond change was proportionally more common at the medium site (light blue) than at the high site (pink). Histograms for the surface mean drop in surface elevation across to the land area lost (2014 vs 2021) for all pond time-series at the given study site. See also Fig. 3 in the main manuscript.

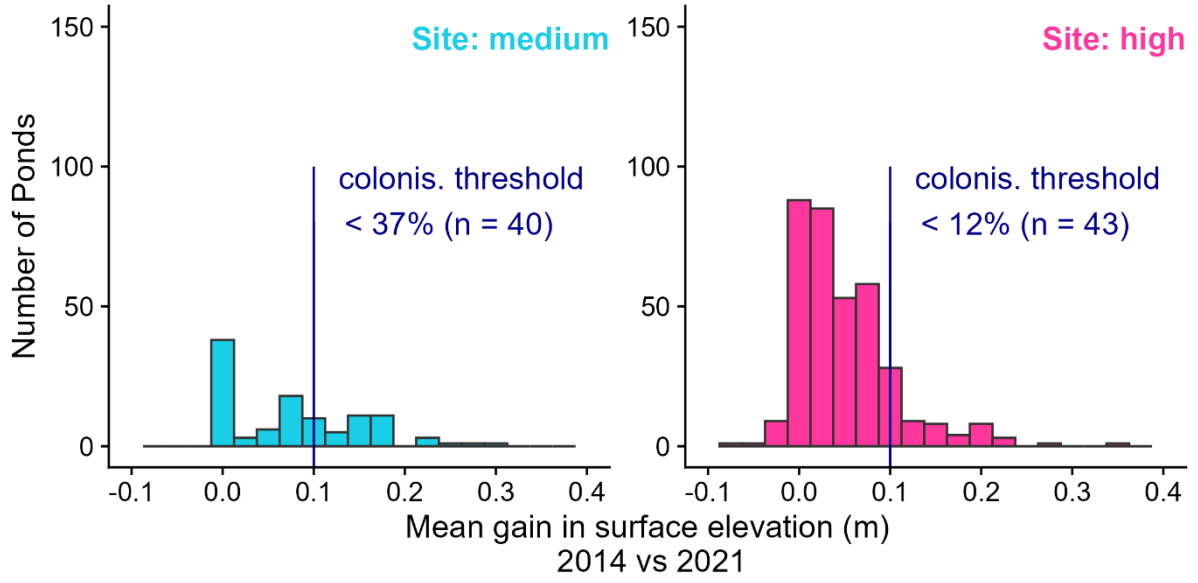

Figure S5: Vegetation colonisation as a driver of pond change was proportionally more common at the medium site (light blue) than at the high site (pink). Histograms of the mean gain in surface elevation across the land area gained (2014 vs 2021) for all pond time-series. See also Fig. 4 in the main manuscript.

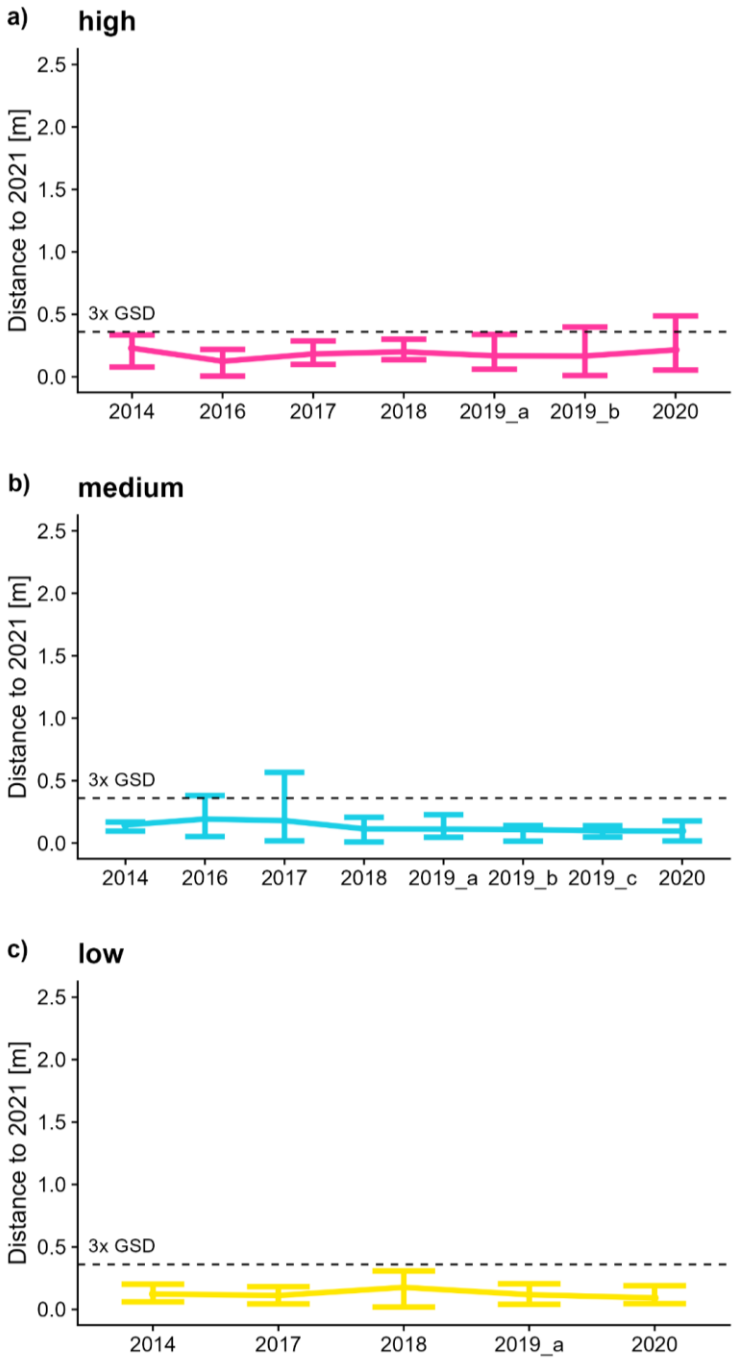

Figure S6: Mean geolocation accuracies of the final drone mosaics relative to the 2021 reference mosaic within each time-series do not exceed 46 cm (three times the raster resolution of 12 cm). Panels show the mean (points) as well as minimum and maximum (error bars) distances of five independent control points in each mosaic relative to their counterparts in the 2021 reference mosaic of the respective time-series. See also the “Geolocation” section in the Supplementary Methods above.

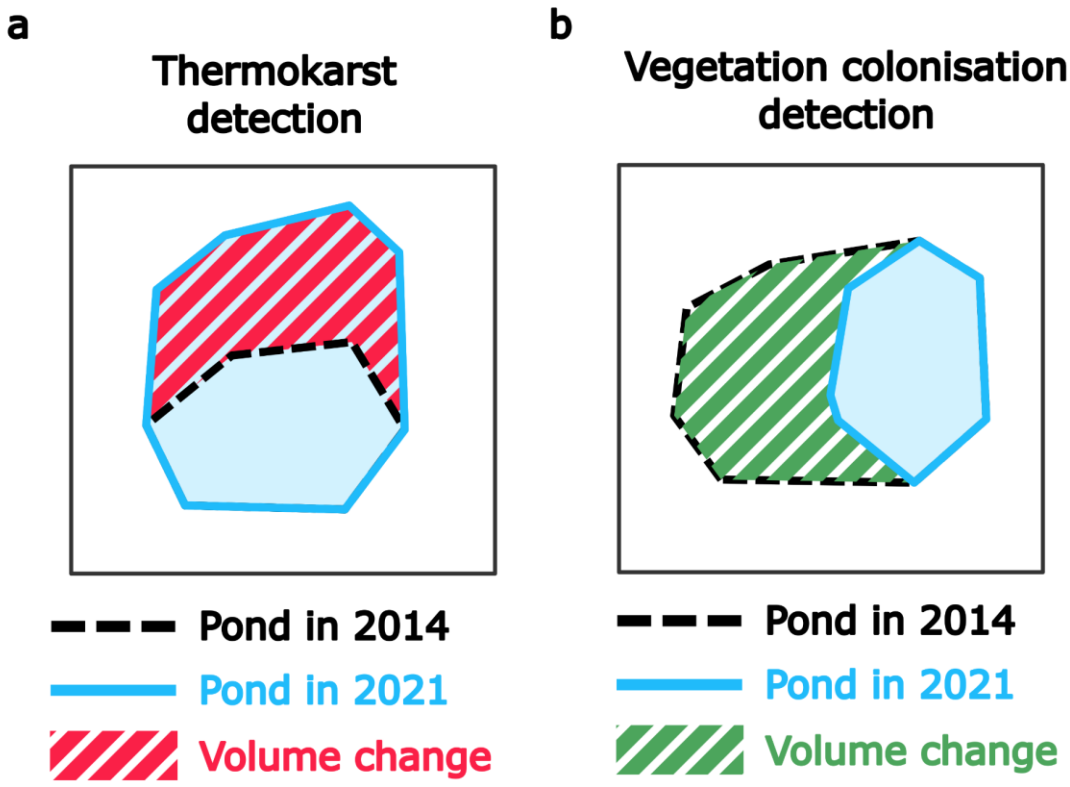

Figure S7: Schematic of the diagrams for the detection of the drivers of pond change. a) For the thermokarst detection, we calculated the mean drop in surface elevation (volume change) across the pond area gained (i.e., pond expansion - shown in hatched red) between 2014 and 2021. b) For the vegetation colonisation detection, we calculated the mean gain in surface elevation (volume change) across the pond area lost (i.e., pond shrinking - shown in hatched green) between 2014 and 2021.