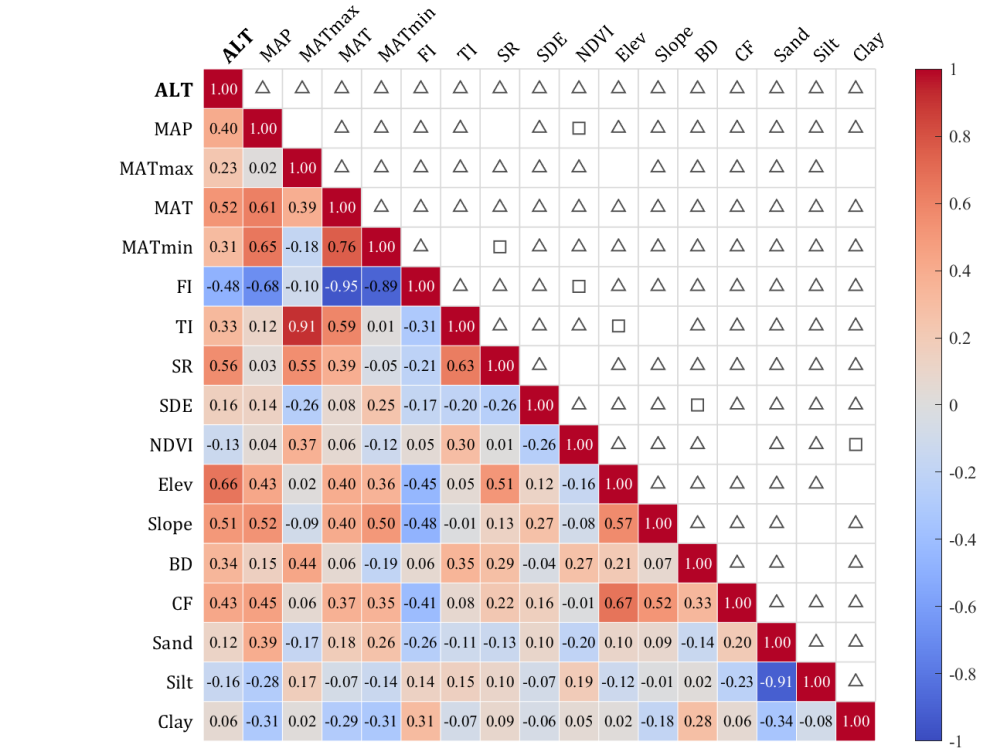


Figure S1. Pearson correlation matrix among ALT and candidate continuous environmental predictors used for preliminary variable screening.

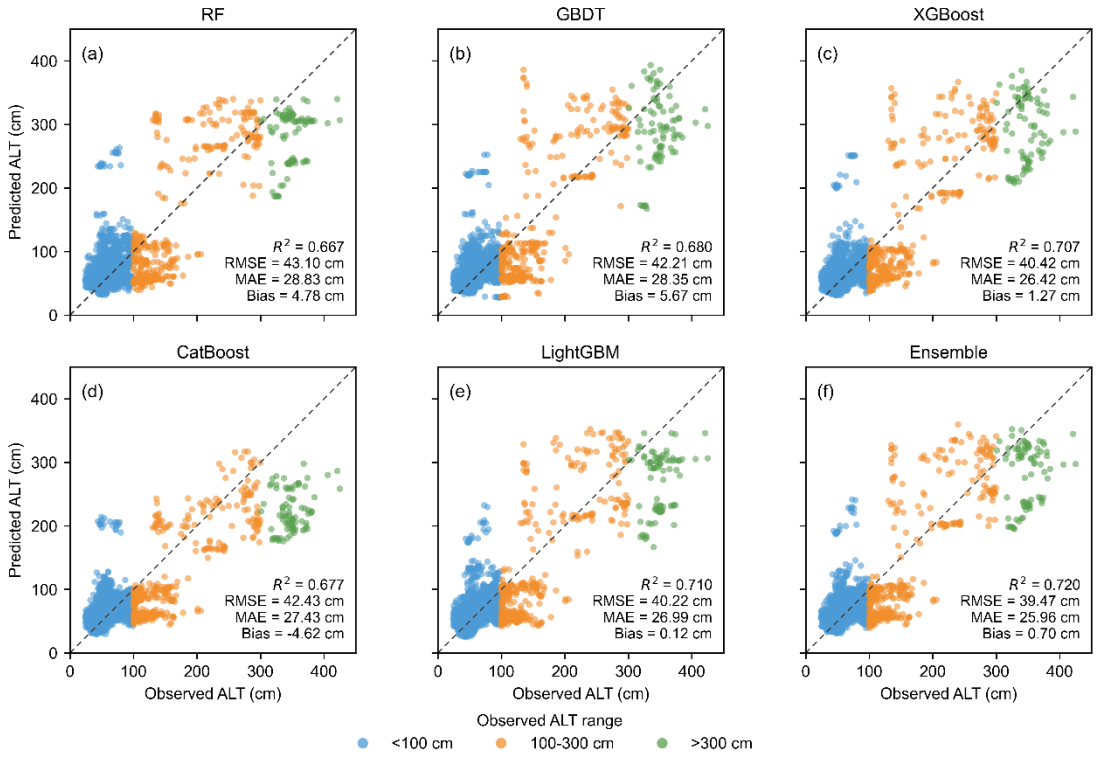


Figure S2. Spatial block cross-validation predictions stratified by observed ALT range. Points are colored by the observed ALT classes (<100 cm, 100–300 cm, and >300 cm). The dashed line indicates the 1:1 relationship.

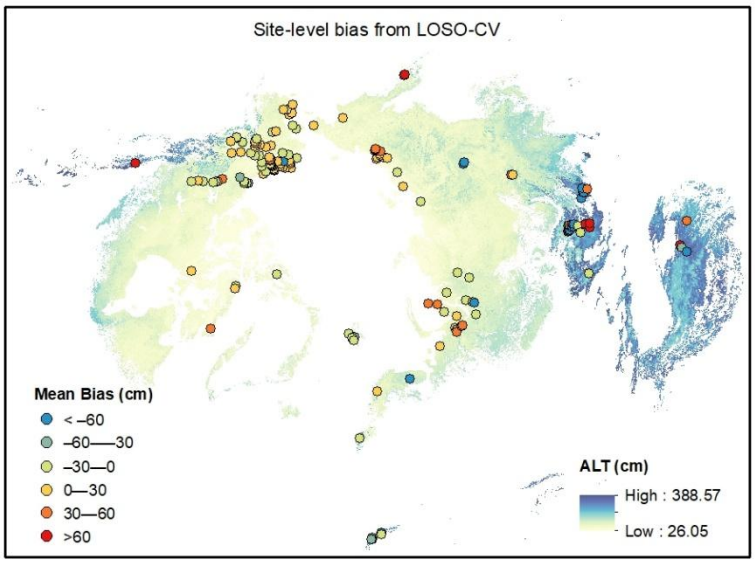


Figure S3. Spatial distribution of site-level mean bias based on leave-one-site-out cross-validation (LOSO-CV). Site-level bias was calculated as the mean difference between predicted and observed ALT (Predicted − Observed) across all available annual observations at each monitoring site. Positive values indicate overestimation and negative values indicate underestimation. The background shows the reconstructed mean ALT distribution during 2000–2024 for spatial reference.

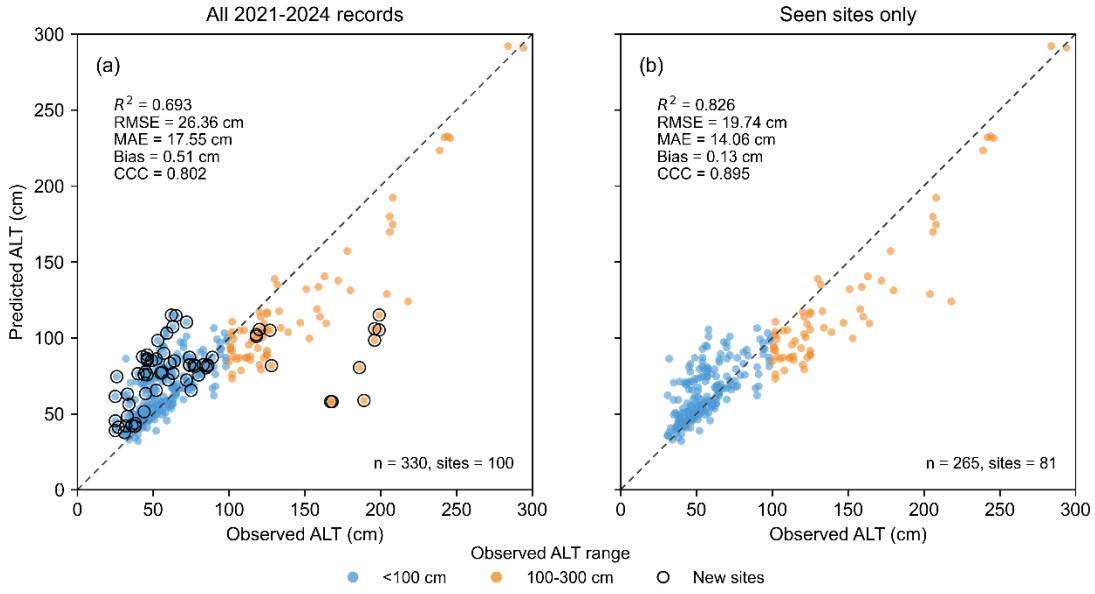


Figure S4. Observed versus predicted active layer thickness (ALT) under temporal blocked validation for 2021–2024. (a) All available observations from 2021–2024, including both previously observed sites and newly observed sites. Newly observed sites are highlighted by open circles. (b) Observations from sites that were already represented in the 2000–2020 calibration dataset. Point colors indicate the observed ALT range, and the dashed diagonal line represents the 1:1 relationship. Model performance statistics are reported within each panel. Bias was calculated as predicted minus observed ALT.

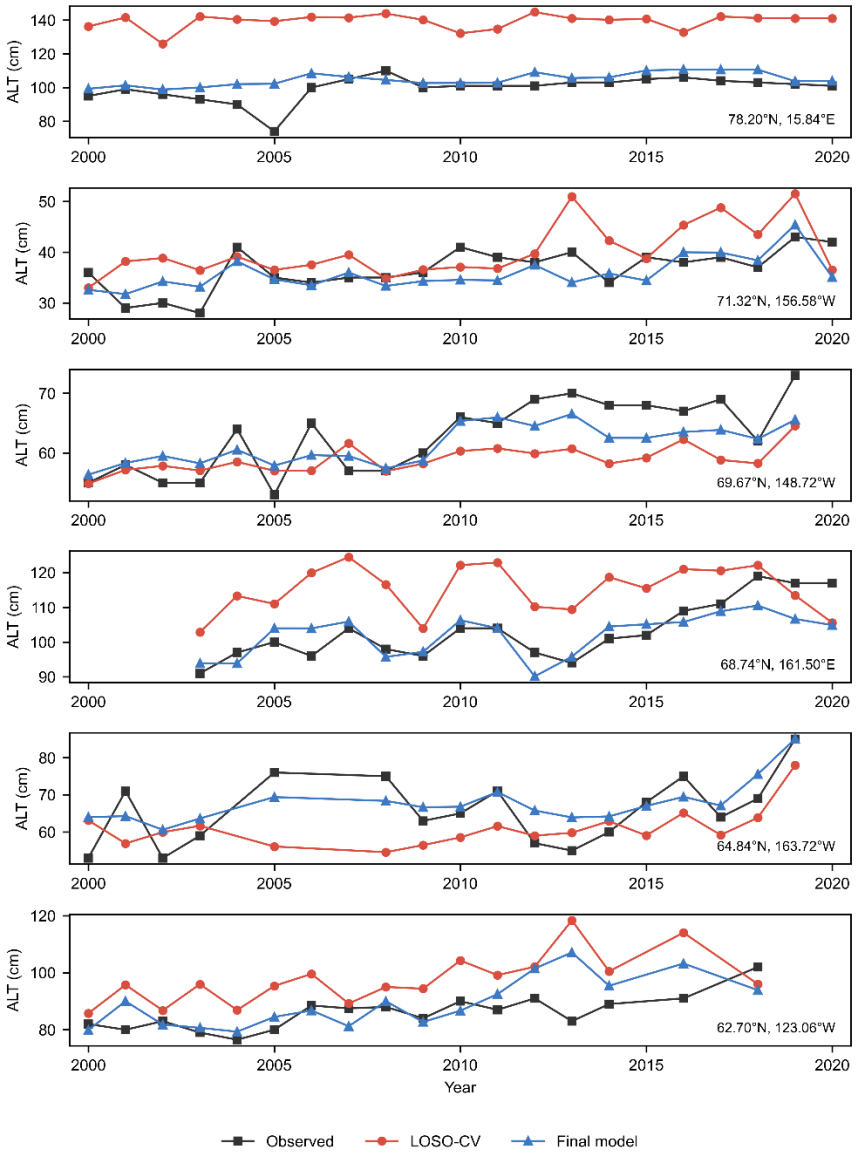


Figure S5. Observed and modelled annual active layer thickness (ALT) time series at six representative long-term monitoring sites during 2000–2020. Black squares denote observed ALT, red circles denote leave-one-site-out cross-validation (LOSO-CV) predictions obtained when all observations from the corresponding site were excluded from model training, and blue triangles denote predictions from the final ensemble model trained using the complete observational dataset. The selected sites illustrate a range of site-level prediction behaviors, including systematic offsets, year-to-year deviations, and differences in the representation of interannual variability.

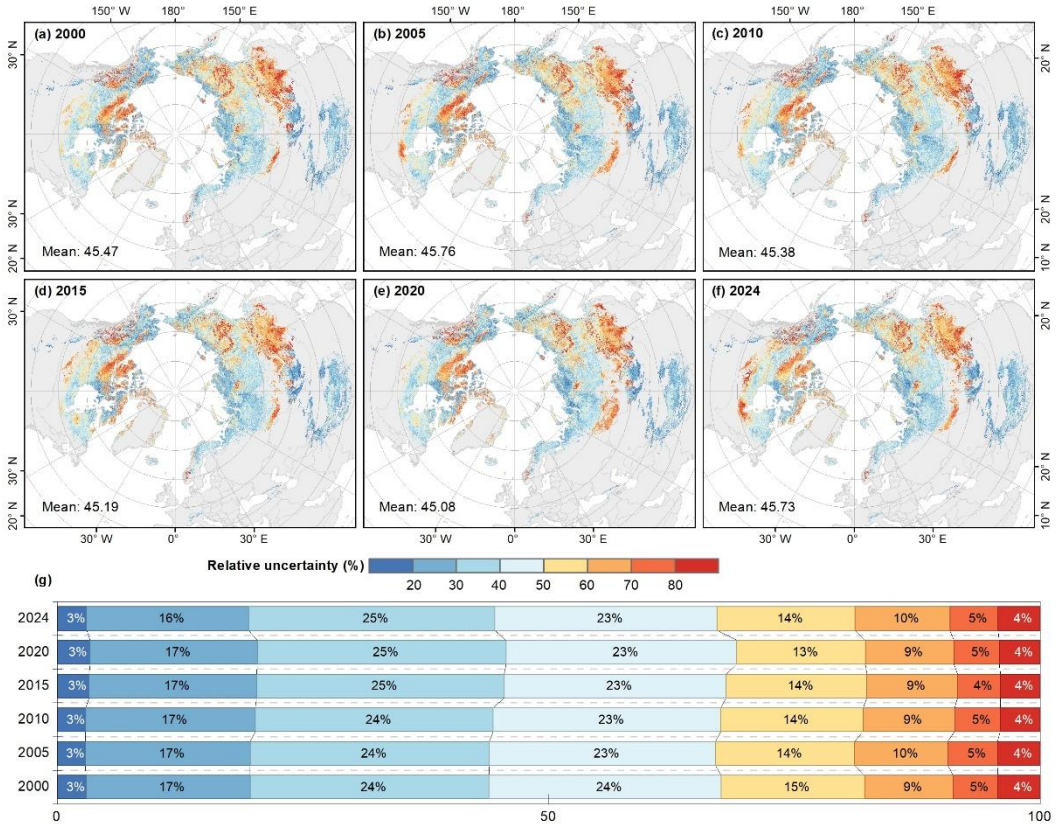


Figure S6. Spatial and temporal variation in relative site-resampling uncertainty during 2000–2024. (a–f) Spatial distributions of relative site-resampling uncertainty for 2000, 2005, 2010, 2015, 2020, and 2024, respectively. (g) Proportional distribution of mapped permafrost area among relative uncertainty classes for the corresponding years.

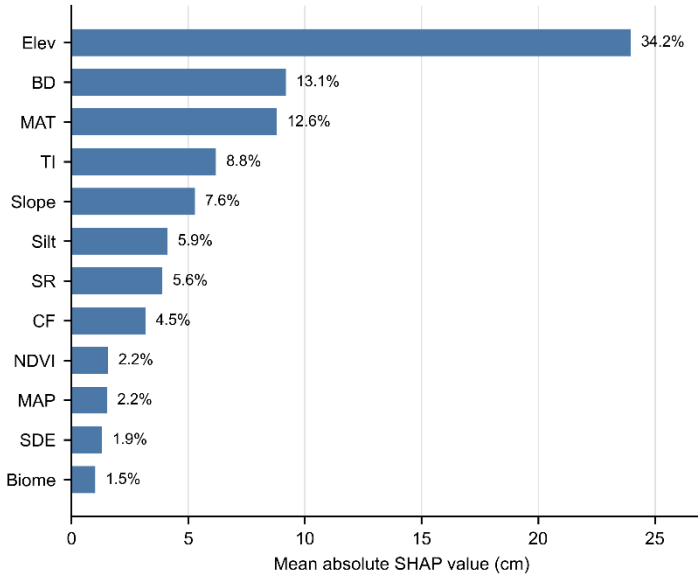


Figure S7. Relative contribution of environmental predictors to the final ensemble model based on normalized mean absolute SHAP values.

Table S1. Information on collected environmental variables.

Variable	Abbreviation	Source	Original spatial resolution
Mean annual temperature	MAT	Historical monthly weather data downscaled using WorldClim 2.1	2.5 arc-minutes
Mean annual precipitation	MAP		
Annual maximum monthly mean temperature	MATmax		
Annual minimum monthly mean temperature	MATmin		
Thawing index	TI	Calculated from monthly mean air temperature	2.5 arc-minutes
Freezing index	FI		
Solar radiation	SR	ERA5-Land	0.1 °
snow depth	SDE		
Normalized difference vegetation index	NDVI	Global 1 km Resolution 2000–2024 Annual Maximum Composite NDVI Dataset (V3)	1 km
Elevation	Elev	General Bathymetric Chart of the Oceans (GEBCO) global elevation dataset	15 arc-seconds
Slope	Slope	Derived from the elevation	
Bulk density	BD	SoilGrids	250 m
Coarse fragment content	CF		
Sand content	Sand		
Silt content	Silt		
Clay content	Clay		
Biome type	Biome	Terrestrial Ecoregions of the World: A New Map of Life on Earth	/

Table S2. Sensitivity of spatial block cross-validation performance to the SoilGrids integration depth.

Soil depth	R^2	RMSE (cm)	MAE (cm)	CCC	Bias (cm)
0–30 cm	0.696	41.189	26.772	0.838	–1.426
0–100 cm	0.722	39.364	25.741	0.853	–0.312
0–200 cm	0.720	39.473	25.965	0.849	0.696

Table S3. Model configurations used for the machine learning algorithms.

Model	Parameter setting
RF	n_estimators=500 criterion=squared_error max_features=1.0 bootstrap=True random_state=42 n_jobs=-1
GBDT	n_estimators=100 learning_rate=0.1 max_depth=3 subsample=1.0 loss=squared_error random_state=42
CatBoost	iterations=500 learning_rate=0.05 depth=6 loss_function=RMSE
XGBoost	n_estimators = 300 learning_rate = 0.05 max_depth = 4 subsample = 0.9 colsample_bytree = 0.9 objective = reg:squarederror random_state = 42 n_jobs = -1 verbosity = 0
LightGBM	n_estimators = 300 learning_rate = 0.05 num_leaves = 31 subsample = 0.9 colsample_bytree = 0.9 random_state = 42 n_jobs = -1 verbosity = -1

Table S4. Spatial block cross-validation performance of the ensemble model stratified by observed ALT range.

ALT range	records	sites	RMSE (cm)	MAE (cm)	Bias (cm)
<100cm	1581	128	26.46	17.19	9.43
100-300cm	378	62	62.64	52.75	-20.56
>300cm	84	19	84.04	70.59	-68.12

Table S5. Record-level and site-level performance of the final ensemble model under leave-one-site-out cross-validation (LOSO-CV).

Evaluation level	R^2	RMSE (cm)	MAE (cm)	CCC	Bias (cm)
Pooled record-level	0.780	34.981	22.053	0.879	1.086
Site-level median	–	17.173	15.827	–	1.325
Site-level mean	–	31.295	29.847	–	–1.080

Note: Pooled record-level metrics were calculated using all annual out-of-fold predictions. Site-level RMSE, MAE, and bias were first calculated separately for each monitoring site and then summarized using their median and mean values, thereby giving each site equal weight regardless of record length.

Table S6. Harmonized hemispheric and regional ALT statistics for this study and two existing products during 2000–2020.

Region	Product	Common valid pixels (n)	Mean (cm)	Median (cm)	P25 (cm)	P75 (cm)
Northern Hemisphere	This study	19007126	107.871	84.909	66.165	117.689
	Westermann et al. (2024)		74.959	67.237	46.903	96.380
	Peng et al. (2024)		169.722	146.729	105.838	221.338
Alaska	This study	1100212	72.756	67.556	58.730	77.633
	Westermann et al. (2024)		68.053	65.522	51.416	82.544
	Peng et al. (2024)		97.815	97.664	86.543	108.736
Canada	This study	5014980	78.881	70.565	57.061	88.612
	Westermann et al. (2024)		60.050	54.770	38.612	78.668
	Peng et al. (2024)		145.408	137.370	102.441	155.154
Russia	This study	9476506	90.469	83.959	68.029	103.248
	Westermann et al. (2024)		78.797	71.310	51.705	101.68
	Peng et al. (2024)		155.982	143.251	105.582	181.766
Northern Europe	This study	337346	95.619	90.494	79.605	101.194
	Westermann et al. (2024)		53.578	48.726	22.090	75.398
	Peng et al. (2024)		107.929	112.410	87.020	123.548
Mongolia	This study	691342	241.104	242.106	206.638	280.682
	Westermann et al. (2024)		123.097	105.874	83.902	142.061
	Peng et al. (2024)		315.268	312.296	284.386	345.476
Tibetan Plateau	This study	1687365	232.087	237.17	195.367	268.404
	Westermann et al. (2024)		71.267	65.807	40.337	93.220
	Peng et al. (2024)		268.08	263.645	242.631	287.736

Note: Annual ALT was first averaged over 2000–2020 for each product. Hemispheric and regional statistics were then calculated using pixels valid in all three products, ensuring an identical spatial comparison domain. Regional statistics were obtained by intersecting the three-product common-valid domain with the corresponding regional masks. P25 and P75 denote the 25th and 75th percentiles, respectively.