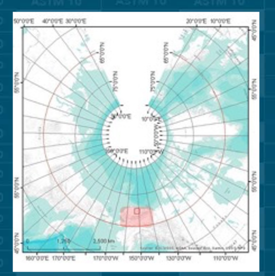

A DEEP LEARNING-BASED APPROACH FOR MAPPING TALL SHRUBS IN ARCTIC TUNDRA

Darko Radakovic[1], Mark Chopping[1], Rocio Duchesne[2], Angela Erb[3], Zhuosen Wang[4] and Crystal Schaaf[3]

[1]Montclair State University [2]University of Wisconsin Whitewater [3]University of Massachusetts, Boston [4]University of Maryland CP/ESSIC

[1]radakovicd1@montclair.edu

ABoVE 10TH SCIENCE TEAM MEETING 21-24 MAY 2024 | UCAR CENTER GREEN CAMPUS BOULDER, COLORADO

Objectives We aim to improve the mapping and quantify tall shrubs in the Arctic, using machine learning (ML) techniques over a period of ~10-18 years to assess changes in (tall) shrub succession in northern Alaska. The objective of this study is to employ semi-automated techniques to analyze high resolution (<1 m) images taken by satellites in orbit to evaluate variations in the growth of shrubs in numerous locations across the Arctic tundra regions of Alaska and Canada, spanning a decade to a decade and a half. The data produced were intended to be accessible to Arctic-Boreal Vulnerability Experiment (ABoVE) researchers for evaluating the effects on summer terrestrial albedo, comparing changes in shrub abundance in Arctic tundra from the satellite high resolution record and albedo, verifying lower spatial resolution ABoVE remote sensing data products, and initiating, driving, calibrating and validating ecological models.

Imagery The project leveraged the availability of commercial high spatial resolution satellite imagery, including QuickBird (QB) (~ 0.6 m) from around 2005 and WorldView-2 (WV02) (~0.4 m) and WorldView-3 (WV03) (~0.3 m) from around 2013 to 2022, for diverse cloud-free, summer tundra landscapes. The Maxar Technologies (then DigitalGlobe) catalog is available to NASA Earth Science investigators, at the NASA Center for Climate Simulation (NCCS).

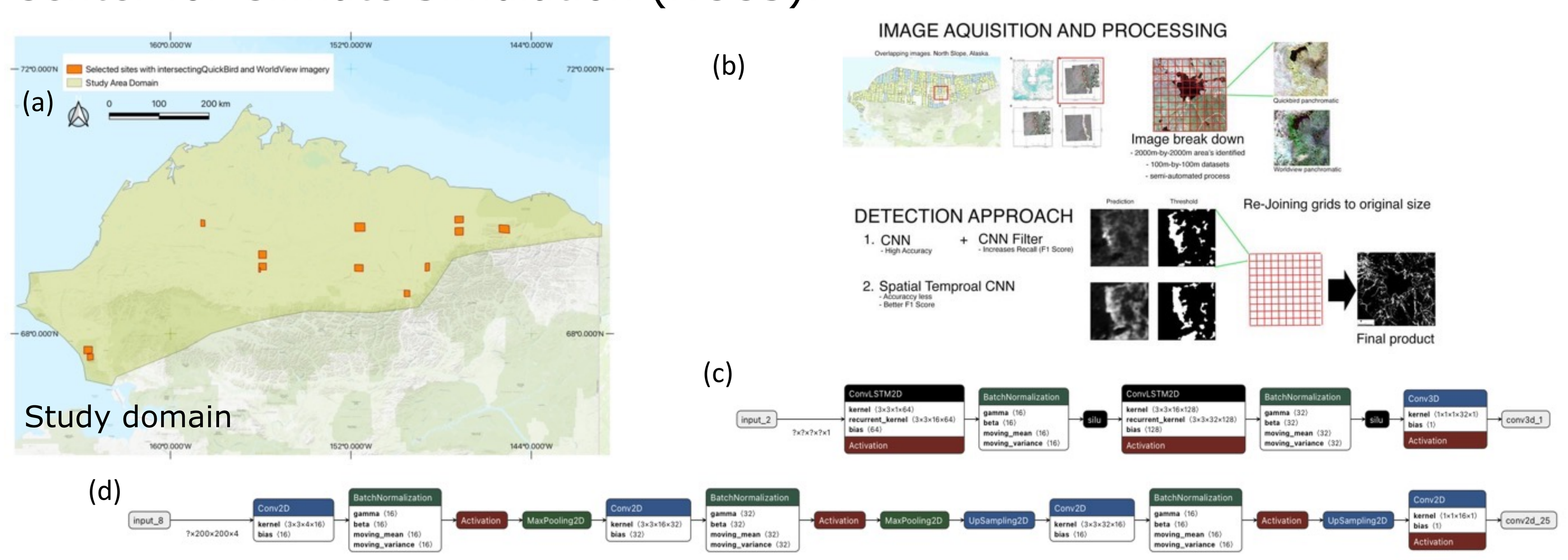

Fig 1. (a) Study Domain (b) Image processing (c,d) Convolutional Neural Network (CNN) and Spatial-Temporal CNN model architecture including ConvLSTM2D layers and batch normalization followed by swish activations, is optimized to process and analyze complex spatio-temporal data efficiently, offering moderate accuracy and recall in shrub detection.

Models The use of Convolutional Neural Network's (CNN) and Spatial-Temporal CNN's (ST-CNN) extend the concept of CNNs to analyze data across both spatial and temporal dimensions allowing to detect changes in vegetation and identify shrubs in peak growth time (Fig. 1c). The "ConvLSTM2D" layer in ST-CNNs, are adept at handling sequential data with spatial context, potentially capturing dynamic changes in vegetation over time and providing a nuanced understanding of how shrubbery expands or contracts across different years (Fig. 1d).

Preprocessing 2-by-2km areas from 13 study sites were subdivided in 100x100m x-train datasets from Pansharpened QB and WV (Fig. 1), y-train data was obtained from the Vegetation Community Map, Toolik Lake Area, Alaska, 2013-2015," which were created using high resolution UAS imagery and lidar data (± 0.02m) (Greaves et al. 2018). A relatively small training and validation sets were used for various experiments aiming to improve validation outcomes. Images with clouds and/or no data were filtered out. Various automated thresholding techniques, including Otsu (Otsu, 1979), Yen (Yen et al. 1995) and mean percentiles, showed similar metrics (Table 1).

Calibration All imagery was orthorectified to the ABoVE Albers Conic Equal Area (Canada) grid (0.5 m) and converted to calibrated spectral radiances using the Polar Geospatial Center's pgc_ortho.py code and Alaska DEM (alaskaned_mosaic_wgs84). In addition, fine-tuning of the output grid upper left XY was needed to match the Toolik Lake Vegetation Community Map.

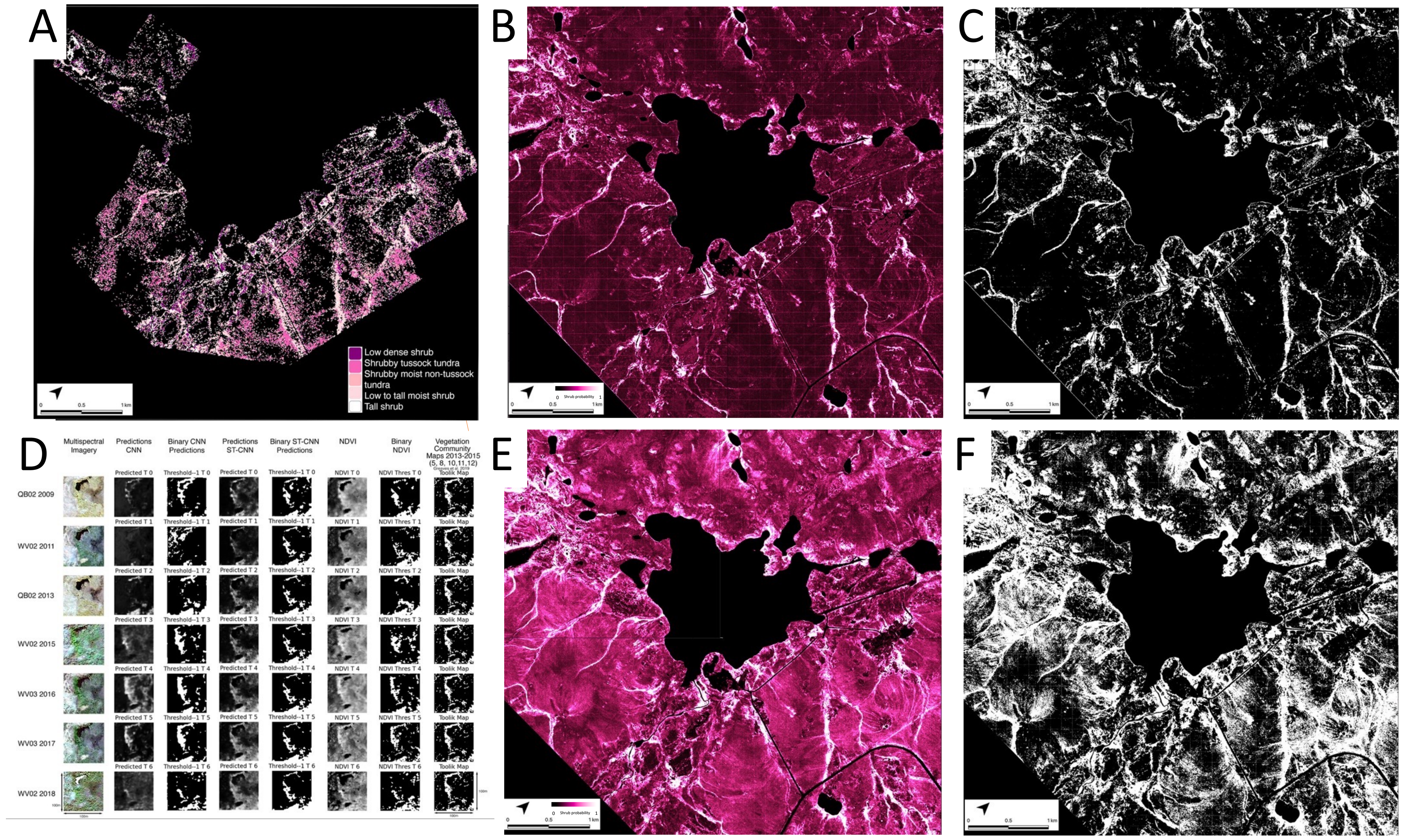

Fig. 2. Toolik Lake site in Alaska (a) Vegetation Community Map, Toolik Lake Area, Alaska, 2013-2015 (Greaves et al. 2018) (b) Shrub cover estimates multisite 6, areas 1, 2, 4 and 5, derived from QuickBird (QB), July 18, 2009 (QB02_20090718220421_1010010009F45F00_09JUL18220421-P1BS-500071841070_01_P001) with the CNN model (LL: 149.604700°W, 68.632583°N) (c) Thresholding CNN shrub cover estimates based on the 'signature area method' (d) overview of thresholding methods applied on predictions to individual small subset timeseries in multi-site6 area 2 where the most recent image is the WorldView-2 (WV02), July 12, 2018 (WV02_20180712063446_1030010080B4BF00_18JUL12063446-P1BS-502391301070_01_P002).(e) Shrub cover estimates multisite6 areas 1, 2, 4 and 5, derived from the same QB as in Fig. 2c. with the ST-CNN model (f) Thresholding ST-CNN shrub cover estimates based on the 'signature area method'.

Validation The Toolik Lake Vegetation Community Map was reprojected onto the same 0.5 m. The classes 0_No_Data, 5_Low dense shrub, 8_Shrubby tussock tundra, 10_Shrubby moist non-tussock tundra, 11_Low to tall moist shrub, and 12_Tall shrub were recoded as a binary map of one class (Fig. 2 and 6). The CNN and ST-CNN shrub estimates were similarly recoded to the binary classes 'nonshrub' and 'shrub' to calculate confusion matrices. To address inconsistencies and assess the impacts of using imagery acquired at differing solar and viewing angles between time-series imagery, low solar and viewing angles were excluded. Cover sequences show anomalous, often infeasible trajectories between timeseries (Fig 3.). Different sensor sun and view angles yielded an adjusted R^2=0.04, N=789, significant at p=0.05, indicating a weak dependence with multiple regression (Chopping et al. 2023). Furthermore, our non-ML approach using Roughness Maps strongly related to aboveground biomass determined from 30 m Landsat imagery (Fig. 6) (Chopping et al. 2023).

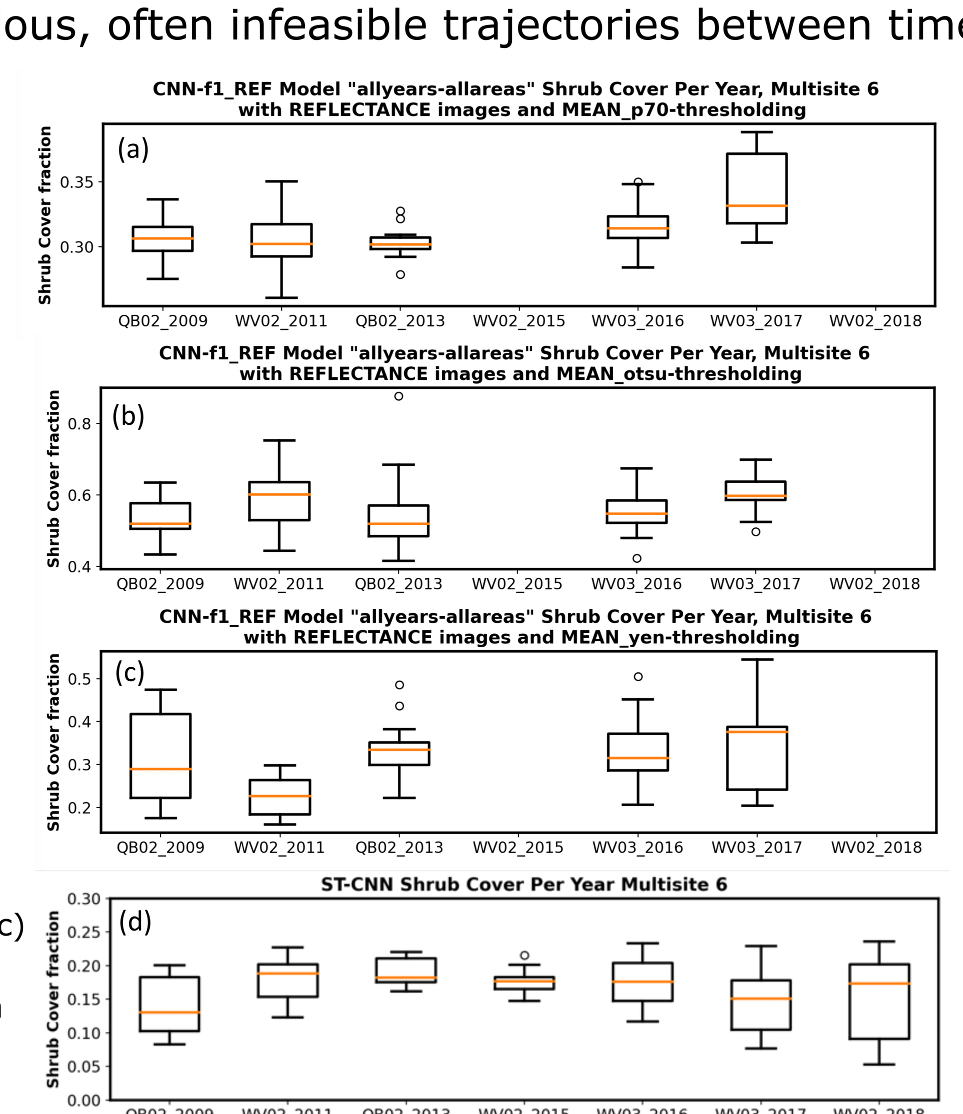

Fig 3. (a) CNN estimated shrub cover for Toolik Lake multi-site 6 all 14 areas with low solar and viewing angles excluded for the mean percentile-70 thresholding (p<0.001), (b) Otsu-thresholding, (c) Yen-thresholding. (d) ST-CNN estimated shrub cover based on a specific manual subset image thresholding technique.

Results The accuracy of the CNN predictions with mean percentile 70, Otsu and Yen thresholding all timeseries of multisite 6, area 2 for the two classes, were 73%, 68% and 70%, and the F1 scores were 53%, 54%, 53%, respectively (Table 1). Using these thresholding techniques, the CNN predicted cover sequences show an increasing trend with, but at different rates (Fig. 3). Improving accuracies tends to lower F1 scores, while an improving the recall comes at the cost of precision and thus accuracy. Annual shrub cover for the mean p70, Otsu and Yen thresholding on the CNN estimates is 1.5% (p<0.05), 1.7% and 0.8% (p<0.05), respectively (Fig. 3, Table 1). Preliminary work shows that water depth and tundra landscape type can be related to shrub cover (Fig 5.).

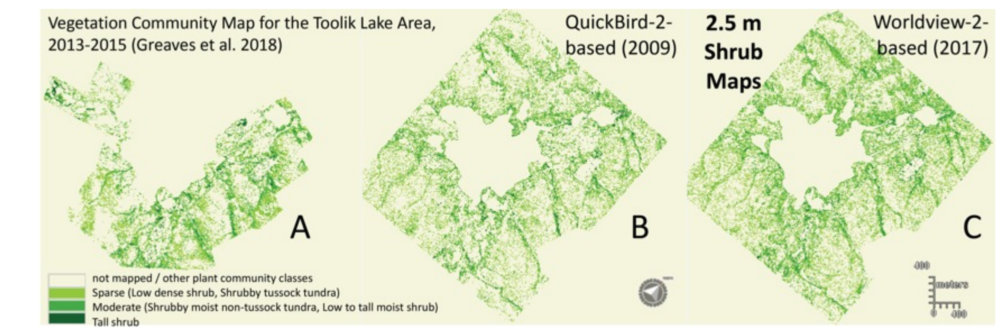

Fig. 6. Toolik Lake ABoVE Sentinel site in Alaska (a) shrub class of the Vegetation Community Map, Toolik Lake Area, Alaska, 2013-2015 (Greaves et al. 2018) (b) mapped shrub classes derived from non-ML approach with panchromatic image roughness based on QuickBird (QB), July 18, 2009 (QB021500009JUL18220421-P1BS-500071841070_01_P001) (c) the same, from WorldView-2 (WV02), August 11, 2017 (WV02_20160612215015_1030010057A6E600_16JUN12215015-P1BS-501511474060_01_P012) (Chopping et al. 2023).

TABLE 1. TALL SHRUB MEASUREMENT RESULTS, BEST RUNS

Model	Technique	Accuracy_mean	Accuracy_std	F1_mean	F1_std	ROC AUC...	ROC A...	Shrub Cover_mean	Shrub Cover_std	F1_ran
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v5_multiyears_classimbalance_f1score.tf	MEAN_p70	0.72744	0.03096	0.53262	0.05179	0.70980	0.05631	3066.58037	242.71450	12.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v5_multiyears_classimbalance_f1score.tf	MEAN_p80	0.75510	0.02495	0.49658	0.06023	0.70980	0.05631	2107.37901	156.21910	41.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v4_multiyear.h5	MEAN_p70	0.72736	0.02980	0.52858	0.04735	0.71052	0.05362	3012.86949	107.80659	16.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v5_multiyears_classimbalance_f1score.tf	MEAN_dyn_p70	0.71827	0.03153	0.53648	0.05044	0.70980	0.05631	3313.72061	269.31370	7.00000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v4_multiyear.h5	MEAN_dyn_p70	0.71873	0.03269	0.53401	0.04560	0.71052	0.05362	3260.86117	135.40737	10.0000
shrub_seg_cnn_model_ms6a2_2013_best_PCA_v5_multiyears_classimbalance_f1score.tf	MEAN_p80	0.74432	0.03516	0.50448	0.07026	0.71121	0.06541	2396.53267	160.80497	34.0000
shrub_seg_cnn_model_ms6a2_2013_best_PCA_v5_multiyears_classimbalance_f1score.tf	MEAN_dyn_p70_img60	0.73827	0.02825	0.50910	0.08213	0.71121	0.06541	2614.80063	474.27004	29.0000
shrub_seg_cnn_model_ms6a2_2013_best_PCA_v5_multiyears_classimbalance_f1score.tf	signature_img60	0.74311	0.03464	0.50341	0.07319	0.71121	0.06541	2423.67017	418.25422	36.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v5_multiyears_classimbalance_f1score.tf	MEAN_dyn_p70_img60	0.75542	0.02109	0.49093	0.08154	0.70980	0.05631	2089.76566	483.12982	53.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v4_multiyear.h5	signature_img60	0.74936	0.02060	0.49348	0.06750	0.71052	0.05362	2212.56551	360.82964	46.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v4_multiyear.h5	MEAN_dyn_p70_img60	0.74852	0.01546	0.49277	0.07836	0.71052	0.05362	2250.87182	495.36205	49.0000
shrub_seg_cnn_model_ms6a2_2013_best_PCA_v5_multiyears_classimbalance_f1score.tf	MEAN_p70	0.71795	0.03807	0.52812	0.06294	0.71121	0.06541	3211.19596	220.12195	18.0000
shrub_seg_cnn_model_ms6a2_2013_best_PCA_v5_multiyears_classimbalance_f1score.tf	MEAN_dyn_p70	0.70767	0.03827	0.53139	0.06085	0.71121	0.06541	3474.48382	274.81685	15.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v5_multiyears_classimbalance_f1score.tf	signature_img60	0.75664	0.02294	0.48754	0.07238	0.70980	0.05631	2013.98184	407.52440	60.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v5_multiyears_classimbalance_f1score.tf	MEAN_yen	0.70044	0.03039	0.53338	0.05383	0.70980	0.05631	3707.29112	854.94098	11.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v4_multiyear.h5	MEAN_p70	0.71317	0.02752	0.51648	0.04482	0.67799	0.04519	3166.93077	217.04433	25.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v4_multiyear.h5	MEAN_p60	0.68195	0.03735	0.54174	0.04280	0.71052	0.05362	4161.41276	183.83259	2.00000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v4_multiyear.h5	MEAN_p80	0.75284	0.02350	0.48392	0.05320	0.71052	0.05362	2026.04479	71.03823	64.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v5_multiyears_classimbalance_f1score.tf	MEAN_p60	0.68209	0.03723	0.54140	0.04503	0.70980	0.05631	4158.15132	281.59704	4.00000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v4_multiyear.h5	MEAN_otsu	0.68042	0.04060	0.54175	0.04294	0.71052	0.05362	4192.30856	296.15417	1.00000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v5_multiyears_classimbalance_...	MEAN_dyn_p60	0.70732	0.03373	0.51342	0.05777	0.69380	0.06246	3252.29869	200.83740	26.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v5_multiyears_classimbalance_...	MEAN_p60	0.70206	0.03457	0.51695	0.05578	0.69380	0.06246	3402.17755	191.67615	24.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v4_multiyear.h5	MEAN_dyn_p70	0.69712	0.02986	0.52060	0.04174	0.67799	0.04519	3549.58472	244.60721	22.0000
shrub_seg_cnn_model_ms6a2_2013_best_PCA_v5_multiyears_classimbalance_f1score.tf	MEAN_p60	0.68141	0.03814	0.53546	0.05300	0.71121	0.06541	4091.32840	265.03777	8.00000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v5_multiyears_classimbalance_f1score.tf	MEAN_dyn_p60	0.67643	0.03651	0.54101	0.04430	0.70980	0.05631	4277.57004	286.51121	5.00000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v4_multiyear.h5	MEAN_p80	0.74661	0.02294	0.48416	0.05195	0.67799	0.04519	2150.29582	128.16404	63.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v5_multiyears_classimbalance_...	MEAN_p70	0.72784	0.03254	0.49340	0.06420	0.69380	0.06246	2610.48517	146.52549	47.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v5_multiyears_classimbalance_...	signature_img60	0.72628	0.03672	0.49560	0.05763	0.69380	0.06246	2650.67865	205.45569	44.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v5_multiyears_classimbalance_...	MEAN_dyn_p70	0.72909	0.03133	0.49122	0.06640	0.69380	0.06246	2567.78878	172.06010	52.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v4_multiyear.h5	MEAN_dyn_p60	0.67327	0.04073	0.54174	0.04233	0.71052	0.05362	4345.96484	232.65482	3.00000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v5_multiyears_classimbalance_...	MEAN_dyn_p70_img60	0.73025	0.03197	0.48951	0.06553	0.69380	0.06246	2523.14282	130.47819	55.0000
shrub_seg_detection_model_2_toolik.h5	MEAN_dyn_p60	0.70382	0.05302	0.50454	0.09761	0.67871	0.09679	3230.76480	350.76545	33.0000
shrub_seg_cnn_model_ms6a2_2013_best_PCA_v5_multiyears_classimbalance_f1score.tf	MEAN_dyn_p60	0.67452	0.03824	0.53539	0.05163	0.71121	0.06541	4240.00337	312.17420	9.00000
shrub_seg_detection_model_2_toolik.h5	MEAN_p70	0.72502	0.05164	0.49168	0.09952	0.67871	0.09679	2649.82834	212.30945	51.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v4_multiyear.h5	signature_img60	0.75149	0.01480	0.46432	0.07603	0.67799	0.04519	1920.12870	445.19555	79.0000
shrub_seg_detection_model_2_toolik.h5	MEAN_p60	0.69560	0.05738	0.50869	0.09372	0.67871	0.09679	3431.31943	279.97946	30.0000
shrub_seg_detection_model_2_toolik.h5	MEAN_dyn_p70	0.72790	0.04725	0.48867	0.10419	0.67871	0.09679	2582.16385	298.17217	58.0000
shrub_seg_cnn_model_ms6a2_2013_best_NDVI_REFLECTANCE_v5_multiyears_classimbalance_...	MEAN_li	0.72868	0.02437	0.48650	0.07942	0.69380	0.06246	2569.68566	494.33716	62.0000
shrub_seg_detection_model_2_toolik.h5	MEAN_dyn_p70_img60	0.71838	0.06997	0.49243	0.08017	0.67871	0.09679	2717.84098	512.50459	50.0000
shrub_seg_cnn_model_ms6a2_2013_best_REFLECTANCE_v4_multiyear.h5	MEAN_dyn_p70_img60	0.74118	0.00550	0.47945	0.08113	0.67799	0.04519	2290.23249	695.96818	68.0000
shrub_seg_detection_model_2_toolik.h5	signature_img60	0.70064	0.09098	0.49752	0.06823	0.67871	0.09679	3083.92036	1002.54020	40.0000

Findings and Future Work Assessing changes in shrub cover and biomass using high resolution imagery relies not only on precise detections, but also on the numbers (recall), which can be affected by training data. Imprecision can be reduced by utilizing CNN filters or ST-CNN models to reduce false positives. Image variations (e.g., between QB and WV, sun angle etc.) challenge a one-to-all algorithm approach. This research will continue accessing the relationship of thawing permafrost and polygonal ground with tall shrub and ground water depth presence.

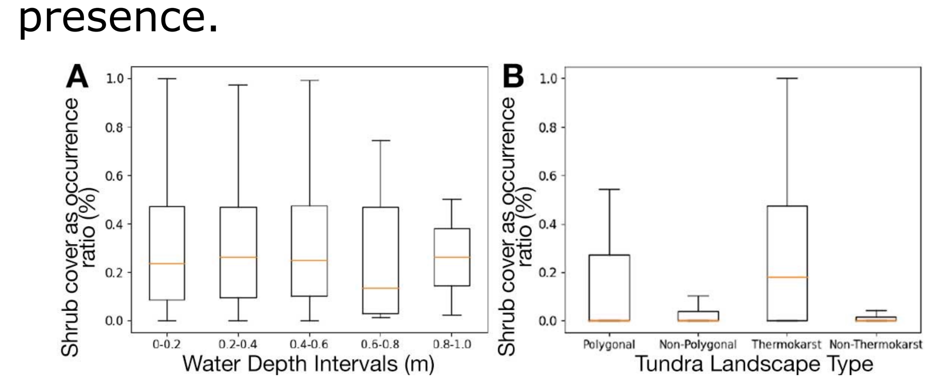

Fig. 5. (a) Combining shrub cover with water depth level and (b) landscape type.

NASA Award 80NSSC23K1559

Acknowledgments: We gratefully acknowledge the assistance of the NASA, GSFC NCCS User Services Group (support@nccs.nasa.gov); Liz Hoy (ABoVE Science Cloud Lead, NASA, GSFC); Mark Carroll (NASA, GSFC); Clare Porter (Polar Geospatial Center); Jim Shute (NASA Center for Climate Simulation (NCCS)); Bruce Cook (NASA Goddard Space Flight Center); NASA Center for Climate Simulation (NCCS);

References

Chopping, M., Radakovic, D., Erb, A., Duchesne, R., Wang, Z., & Schaaf, C. (2023). Mapping Shrub Abundance in Arctic Tundra from the Satellite High Resolution Record and Impacts on Albedo: Validation. University of Maryland College Park, Earth System Science Interdisciplinary Center; Montclair State University; University of Massachusetts, Boston; University of Wisconsin Whitewater.

Greaves, H.E., L. Vierling, J. Eitel, N. Boelman, T. Magney, C. Prager, and K. Griffin. 2018. High-Resolution Shrub Biomass and Uncertainty Maps, Toolik Lake Area, Alaska, 2013. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1573

Otsu, N. (1979). "A threshold selection method from gray-level histograms." IEEE Transactions on Systems, Man, and Cybernetics, 9(1), 62-66.

Yen, J.-C., Chang, F.-J., & Chang, S. (1995). "A new criterion for automatic multilevel thresholding." IEEE Transactions on Image Processing, 4(3), 370-378.

ABoVE 10TH SCIENCE TEAM MEETING 21-24 MAY 2024 | UCAR CENTER GREEN CAMPUS BOULDER, COLORADO