Supplementary material: Comparison of long-term albedo products against spatially representative stations over snow

Ruben Urraca[1], Christian Lanconelli[1,2], Fabrizio Cappucci[1,2], Nadine Gobron[1]

[1]European Commission, Joint Research Centre, Via Fermi 2749, I-21027 Ispra, Italy
[2]UniSystems SA, Rue du Puits Romain 29, L-8070 Bertrange, Luxembourg

1 True size of the coarsest satellite pixel validated

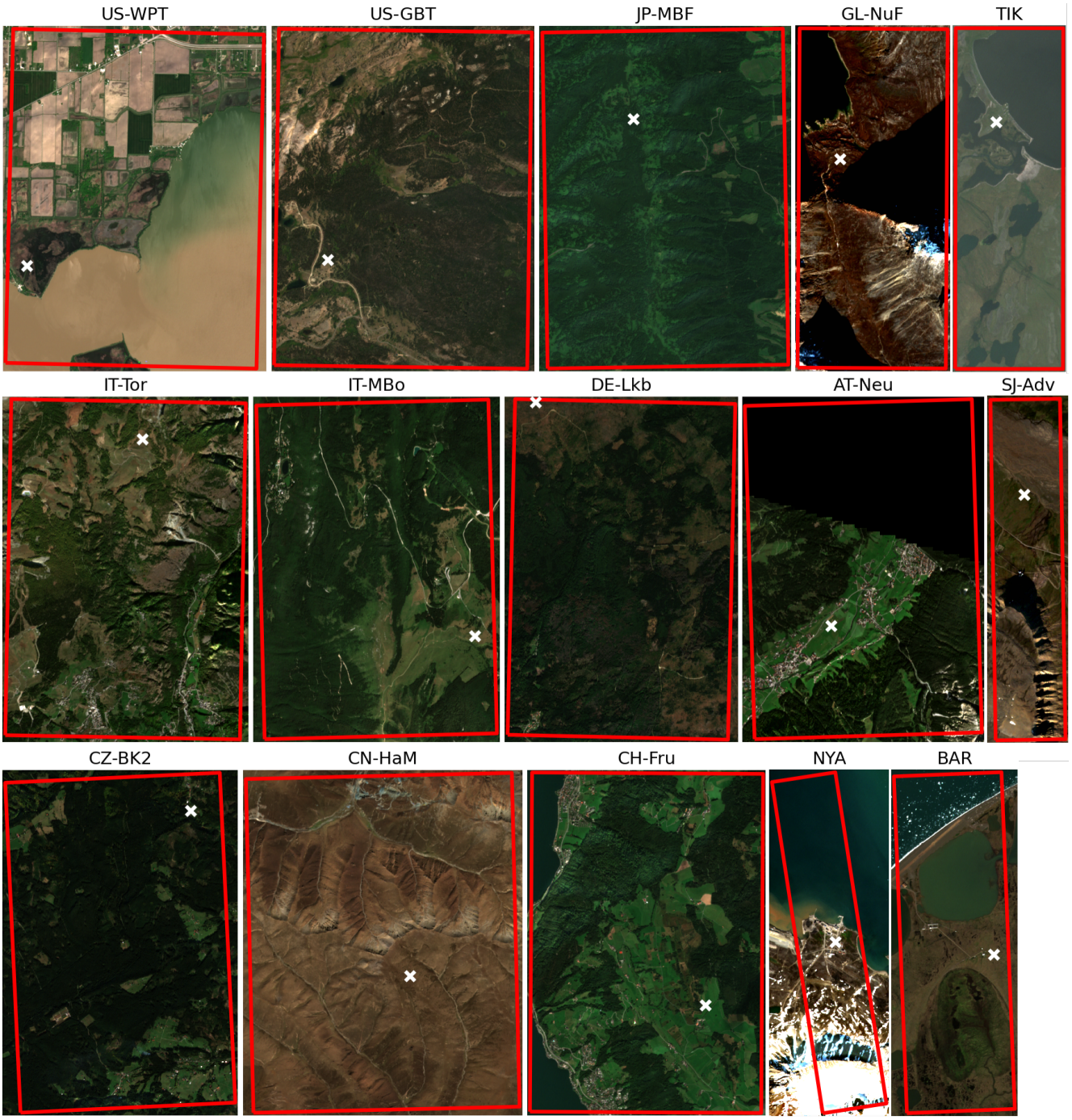

Figure S1: Sentinel-2 MSI L2A True Color Image (TCI) (10×10m) in the 0.05×0.05° pixels containing stations discarded due to an heterogeneous snow cover. The white cross represents the station location. The red rectangle shows the 0.05×0.05° pixel.

Figure S2: Same as Figure S1 but for stations discarded due to an heterogeneous land cover.

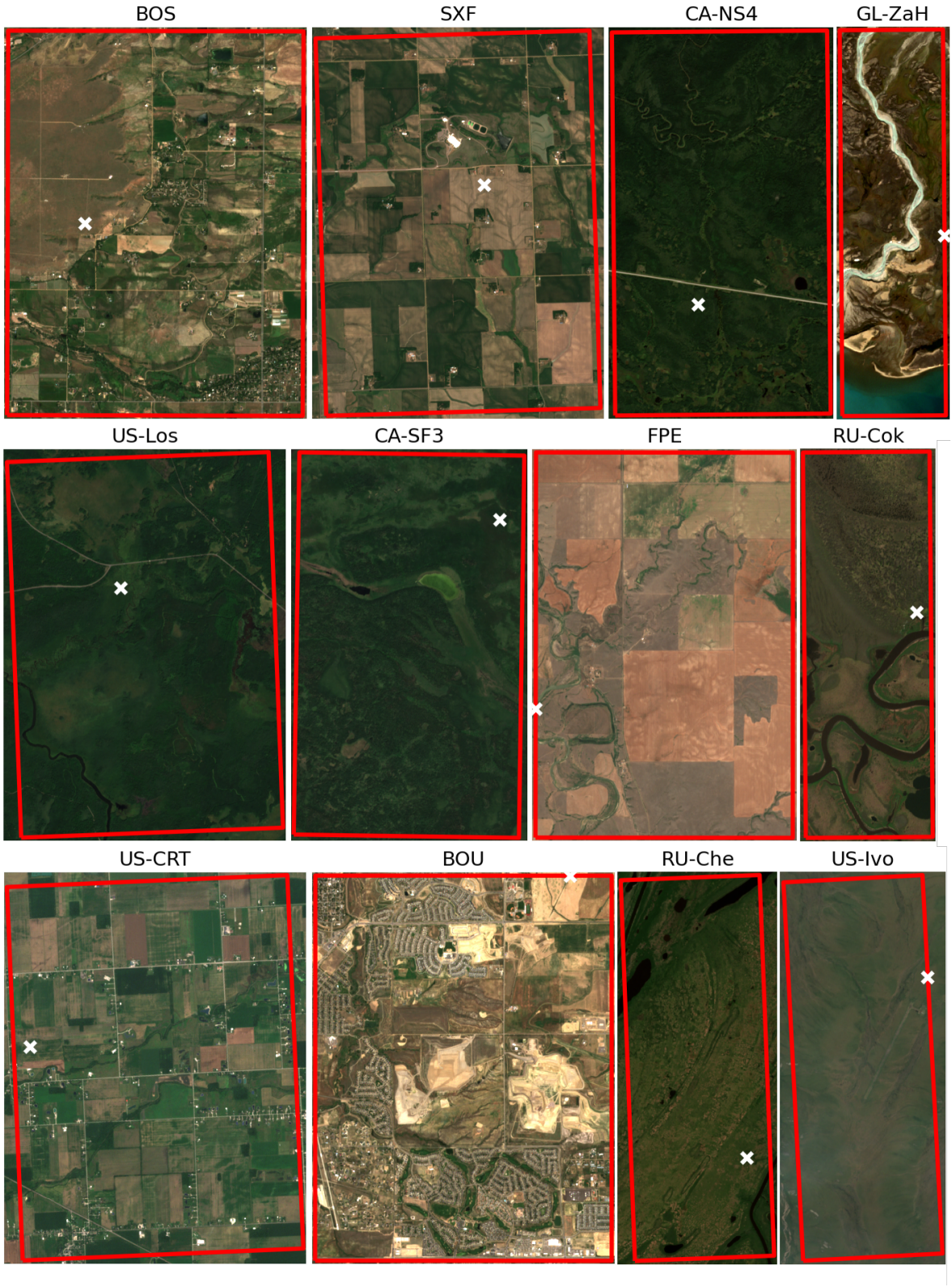

Figure S3: Same as Figure S1 but for stations with homogeneous land and snow cover.

2 Analysis of variogram models based on Sentinel-2 reflectance

Concordia (DOM) and Neumayer station (GVN) have the most spatially homogeneous surface reflectance, with only small heterogeneities in the outer part of the 4km subset caused by the station buildings. DOM is located at the Antarctic Plateau at 3233 m, while GVN is located near the coast. Both are snow-covered during the whole year. Both have the smallest coefficient of variation (CV = 0.009-0.018) and variograms with very small range ($a < 100m$). In both stations, the variogram sill increases from 1-1.5 km to 4 km due to the presence of the station buildings but remains low in absolute terms.

Zackenberg Heath (GL-Zah), Cherski (RU-Cok), Ivotuk (US-Ivo) and Chokurdakh (RU-Cok) are the northernmost stations ($> 65°$ N). All of them are located in low-vegetation sites. In all of them, surface reflectance is very spatially homogeneous in the largest 4×4 km region, just slightly altered by the presence of water bodies (GL-ZAh, RU-Cok, RU-Che) or snow-covered mountains (US-Ivo). Us-Ivo is located at the foothill of the Brooks Range in an area covered by tussock sedge, dwarf shrub, and moss tundra. All variograms have a small range ($a < 350m$) and sill except for the 4 km variogram that is affected by the presence of a snow-covered mountain in the east, but this mountain is not contained by the coarsest pixel validated in this study (Figure S3). RU-Che is located in an area of permanent wetlands. Surface reflectance is homogeneous NW of the station but not in the SE due to the presence of a river and several lakes. The increasing albedo from SE to NW explains why variograms do not reach a sill within the maximum lag, but again, most of these lakes are not included in the coarsest pixel validated (Figure S3). RU-Cok is surrounded by open shrublands, having a good spatial homogeneity north of the station but not in the south, due to the presence of a river and small lakes. The river is included in the three spatial subsets, explaining the good consistency between variograms at different spatial resolutions. However, this also explains why RU-Cok has the largest coefficient of variation ($CV \sim 0.27$) during snow-free conditions. GL-Zah is located in an area of continuous permafrost and includes a portion of the sea in the 4×4 km region, increasing the 4km variogram sill. The spatial representativeness of all these stations improves during the snow season when all of them show smaller CVs and smaller variogram ranges. Still, the presence of water bodies around RU-Cok and GL-Zah, but outside the albedometer footprint, could lead to an underestimation of albedo by satellite products in the point-to-pixel comparison.

UCI-1964 (CA-NS4), Saskatchewan - Western Boreal (CA-SF3), Lost-Creek (US-Los) are all surrounded by a mosaic of forests, natural grasslands and shrubs. CA-NS4 is located in a continental boreal forest dominated by black spruce trees. CA-SF3 burned site currently contains a mosaic of jack pine, black spruce and sparse grass. US-Los is a wetland site mostly surrounded by a deciduous shrub wetland that also contains some coniferous. The mosaic of forest with grasslands and shrub, as well as the low average albedo, explain the large coefficient of variations during the snow-free season ($CV \sim 0.25$). The spatial representativeness decreases even more during the snow season due to the large reflectance differences between snow-covered forests and snow-covered grasslands surrounding the forests. As a consequence, snow-covered images of forest stations have the largest CV (0.7-0.8) overall, and greater variogram ranges and sills than during the snow-free season.

Curtice Walter-Berger Cropland (US-CRT), Fort-Peck (FPE), Sioux Falls (SXF), Boulder (BOS and BOU) are all surrounded by mosaics of croplands and grasslands. BOS and BOU are just 21 km far apart. As in the three stations surrounded by forests, the mosaic of different types of crops increases the spatial variability around the station. However, the spatial representativeness of stations surrounded by mosaics of croplands is better than those surrounded by forests, particularly during the snow season due to the more similar reflectance of different types of crops covered by snow. The spatial representativeness still worsens during the snow season, but CV remains below 0.35, with BOS and BOU showing the best homogeneity ($CV < 0.2$). Note that in these stations the spatial representativeness will also vary with the different growing seasons of each crop.

Based on the previous analysis, we discarded the three stations surrounded by forests (US-Los, CA-NS4, CA-SF3) due to their large spatial heterogeneity, particularly during the snow season.

Table S1: Quantitative assessment of the spatial representativeness of the stations with respect to albedo based on the coefficient of variation (CV), ST_{score}, and RAW_{score} derived from variograms at 4×4 km, 1.5×1.5 km and 1×1 km.

			CV			ST_{score}	RAW_{score}	
	Snow	date	$1km$	$1.5km$	$4km$	1-$1.5km$	1-$1.5km$	1-$4km$
GL-ZaH	free	2021-08-17	0.184	0.188	0.279	0.59	28.39	0.97
	covered	2020-04-07	0.141	0.153	0.166	0.91	5.77	2.80
RU-Cok	free	2021-08-08	0.278	0.279	0.267	1.34	114.30	12.54
	snow	2020-04-04	0.069	0.063	0.055	1.77	5.59	2.39
RU-Che	free	2021-07-03	0.180	0.189	0.342	0.36	9.46	0.55
	covered	2020-04-26	0.093	0.130	0.132	0.02	1.25	1.19
US-Ivo	free	2021-07-16	0.045	0.046	0.094	1.19	18.27	0.46
	covered	2021-04-20	0.020	0.017	0.117	1.16	4.54	0.10
CA-NS4	free	2021-08-26	0.253	0.224	0.217	1.59	4.34	3.49
	covered	2020-05-06	0.394	0.387	0.463	1.60	26.24	2.86
CA-SF3	free	2021-07-01	0.183	0.179	0.189	1.01	28.29	14.89
	covered	2020-03-06	0.798	0.712	0.735	0.63	4.63	6.34
FPE	free	2021-07-23	0.141	0.133	0.119	1.26	8.96	3.26
	covered	2020-01-20	0.336	0.346	0.318	1.61	16.38	9.50
US-Los	free	2021-08-16	0.122	0.139	0.150	1.13	3.64	2.15
	covered	2020-02-23	0.610	0.743	0.776	1.12	2.31	1.84
SXF	free	2021-07-01	0.119	0.162	0.130	0.81	1.36	5.33
	covered	2021-01-02	0.296	0.319	0.261	0.96	6.50	4.21
US-CRT	free	2021-07-18	0.151	0.142	0.154	0.72	8.13	22.15
	covered	2020-02-29	0.202	0.278	0.286	0.58	1.33	1.20
BOS	free	2021-07-10	0.102	0.109	0.135	1.22	7.95	1.55
	covered	2020-02-11	0.158	0.159	0.159	1.71	80.34	76.04
BOU	free	2021-07-10	0.189	0.185	0.191	16.20	20.17	62.04
	covered	2020-02-08	0.084	0.153	0.205	1.56	0.61	0.35
GVN	covered	2020-01-04	0.012	0.012	0.018	1.46	46.17	0.97
DOM	covered	2021-01-01	0.009	0.010	0.018	0.78	3.99	0.46

Table S1 summarizes ST_{score} and RAW_{score} values for each station. The results obtained evidence some of the limitations of using RAW_{score} or ST_{score} to assess spatial representativeness. First, both are derived using relative metrics, so when the parameter measured is low, even small changes lead to a large RAW_{score} and ST_{score}. This is why Antarctic stations have a small RAW_{score} despite their good spatial homogeneity. Second, the variogram-based method places the station in the center of the pixel. Not taking into account the exact position of the station with respect to the pixel validated excludes some valid stations such as US-Ivo and RU-Che, where the heterogeneities causing the small RAW_{score} are outside the pixel validated. Third, both metrics do not evaluate the temporal changes of spatial representativeness, as the method used to assess snow cover homogeneity does. At most, some studies calculate these metrics (and the variograms) for different seasons (leaf off/on, snow-covered/free). However, as shown in the comparison between snow-free/covered conditions, spatial representativeness significantly varies temporally especially over mosaics of crops, and mosaics of forest with grasslands. All these issues should be accounted for when using ST_{score} and RAW_{socre} metrics, and generally, an automatic selection of stations based solely on these metrics should be avoided.

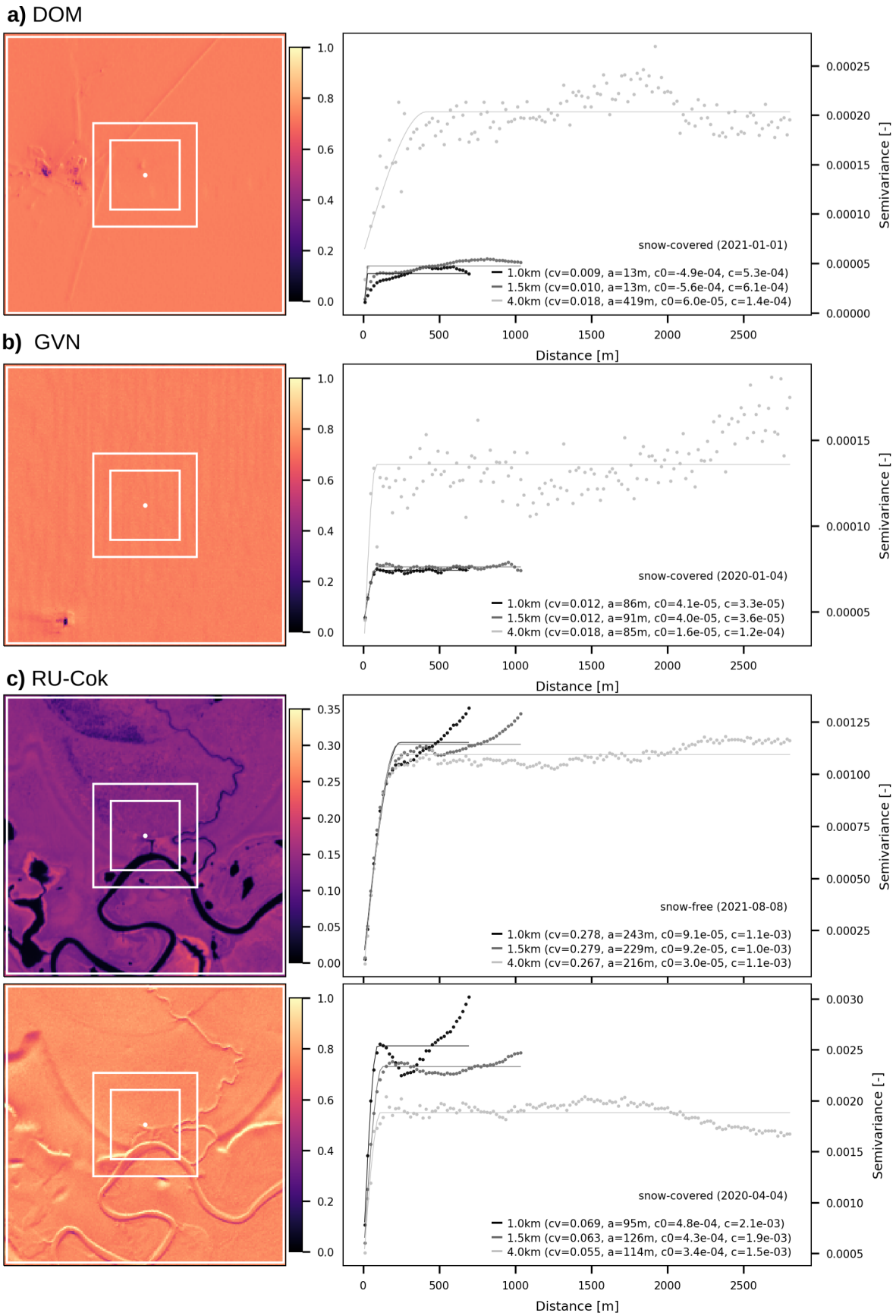

Figure S4: Variogram estimator (dots) and spherical models (lines) derived from S2 MSI L2A shortwave reflectance [-] (20×20m) within 4×4 km, 1.5×1.5 km and 1×1 km regions around the station: (**a**) DOM - Concordia Station, Dome C, (**b**) GVN - Georg von Neumayer (**c**) RU-Cok - Chokurdahk. CV = coefficient of variation, a = range, $c0$ = nugget, c = partial sill.

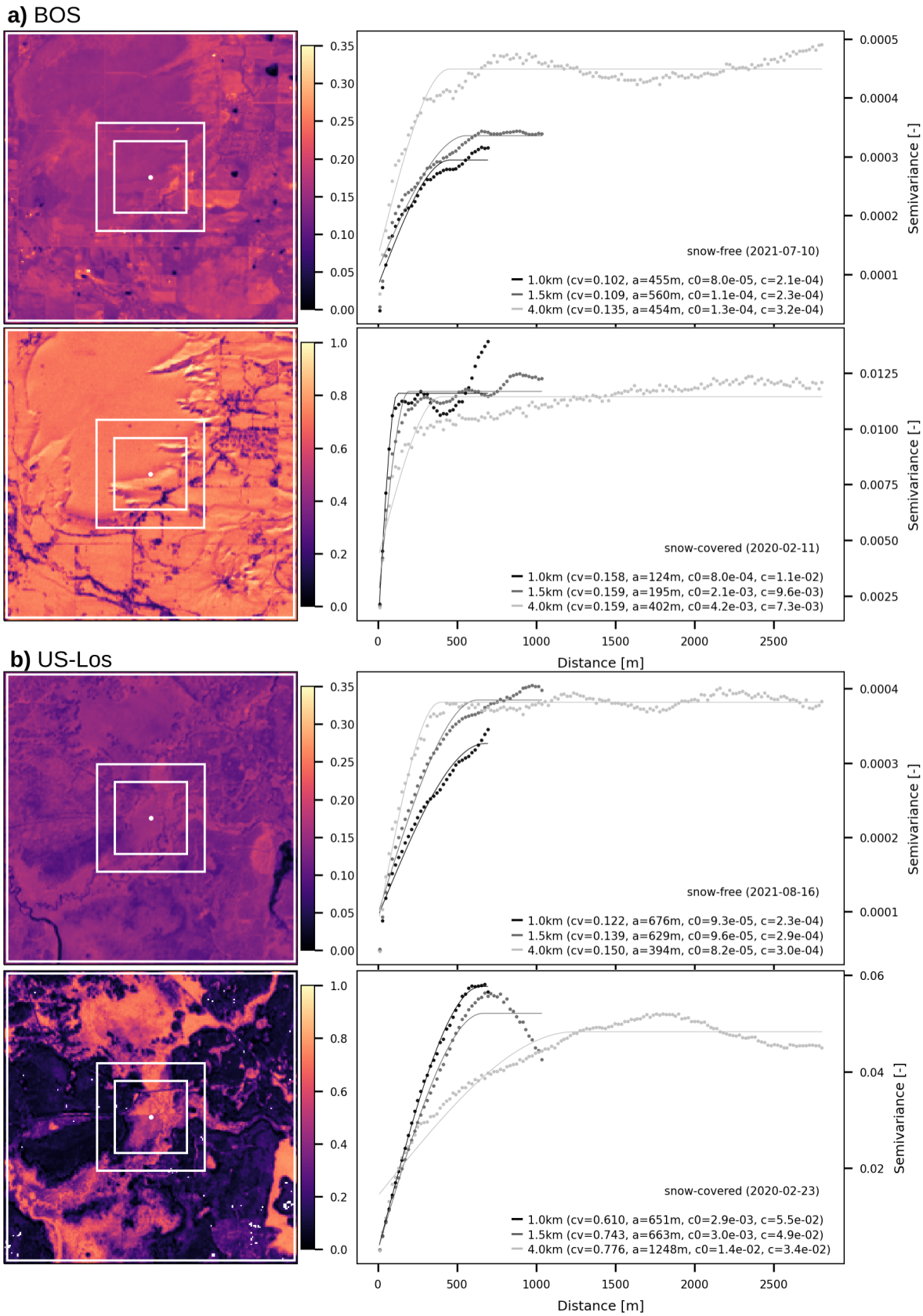

Figure S5: Same as Figure S4 but for:(**a**) BOS - Boulder and (**b**) US-Los - Lost Creek.

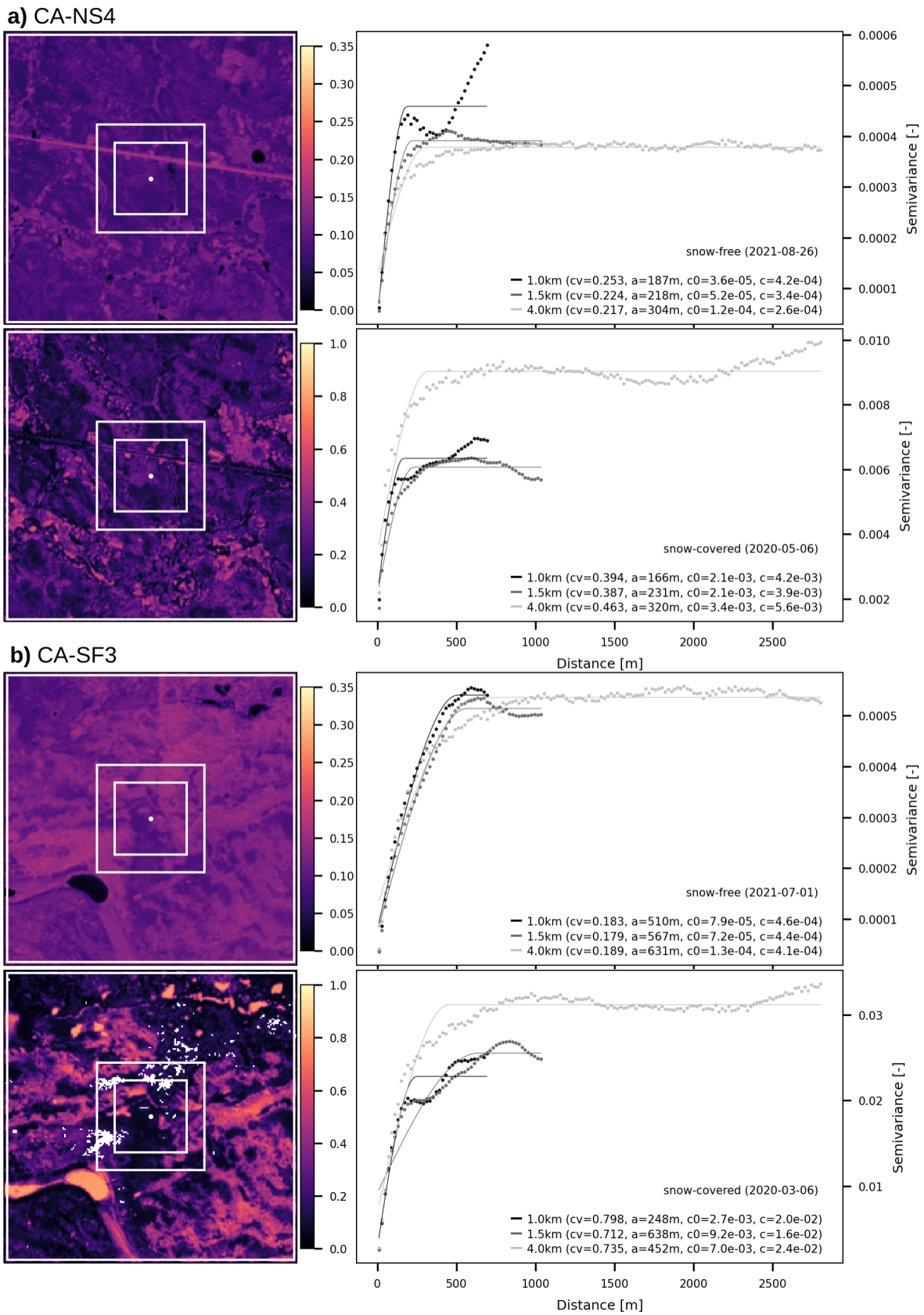

Figure S6: Same as Figure S4 but for (**a**) CA-NS4 - UXI-1964 burn site and (**b**) CA-SF3 - Saskatchewan W Boreal.

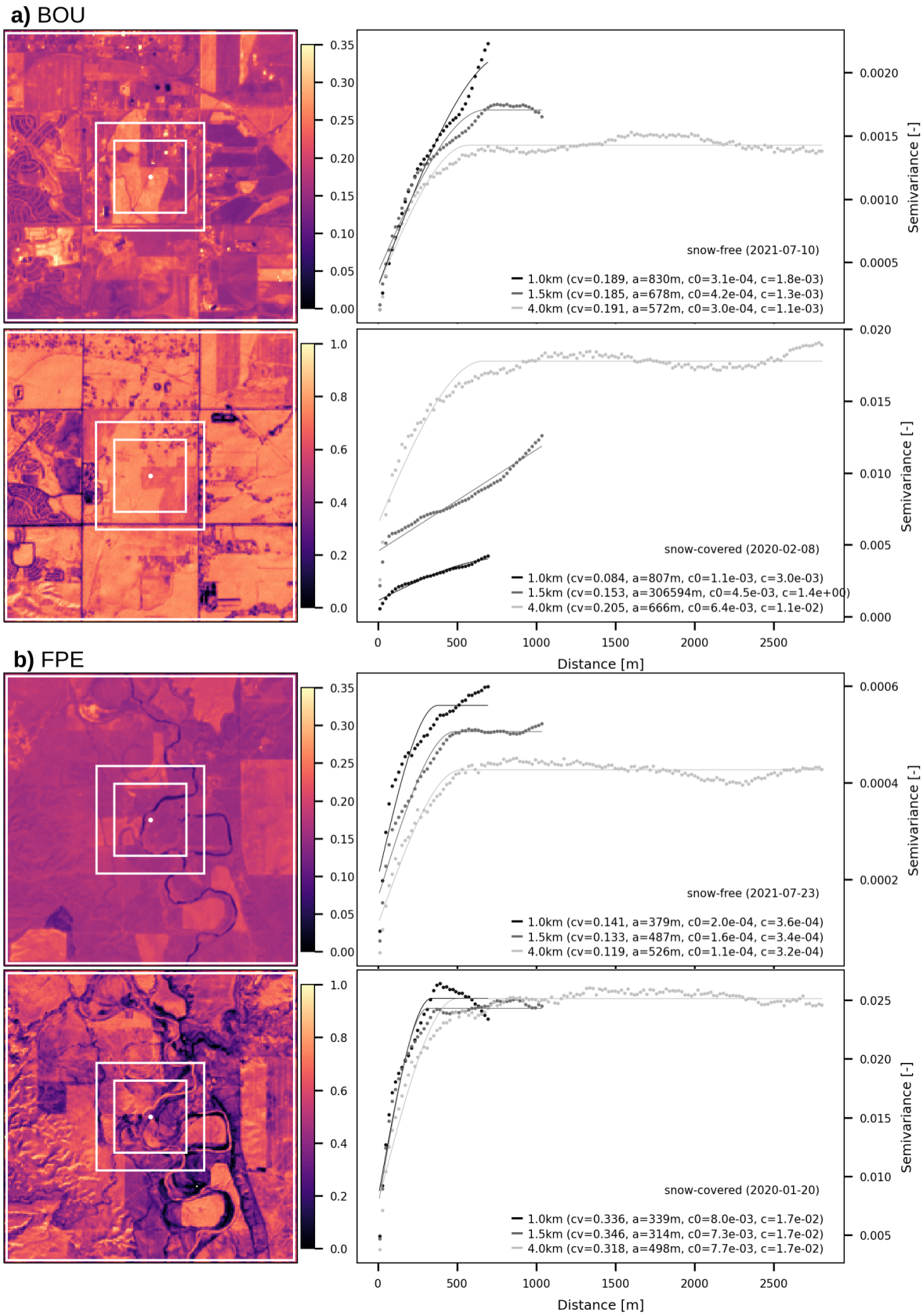

Figure S7: Same as Figure S4 but for (**a**) BOU - Bolder and (**b**) FPE - Fort Peck.

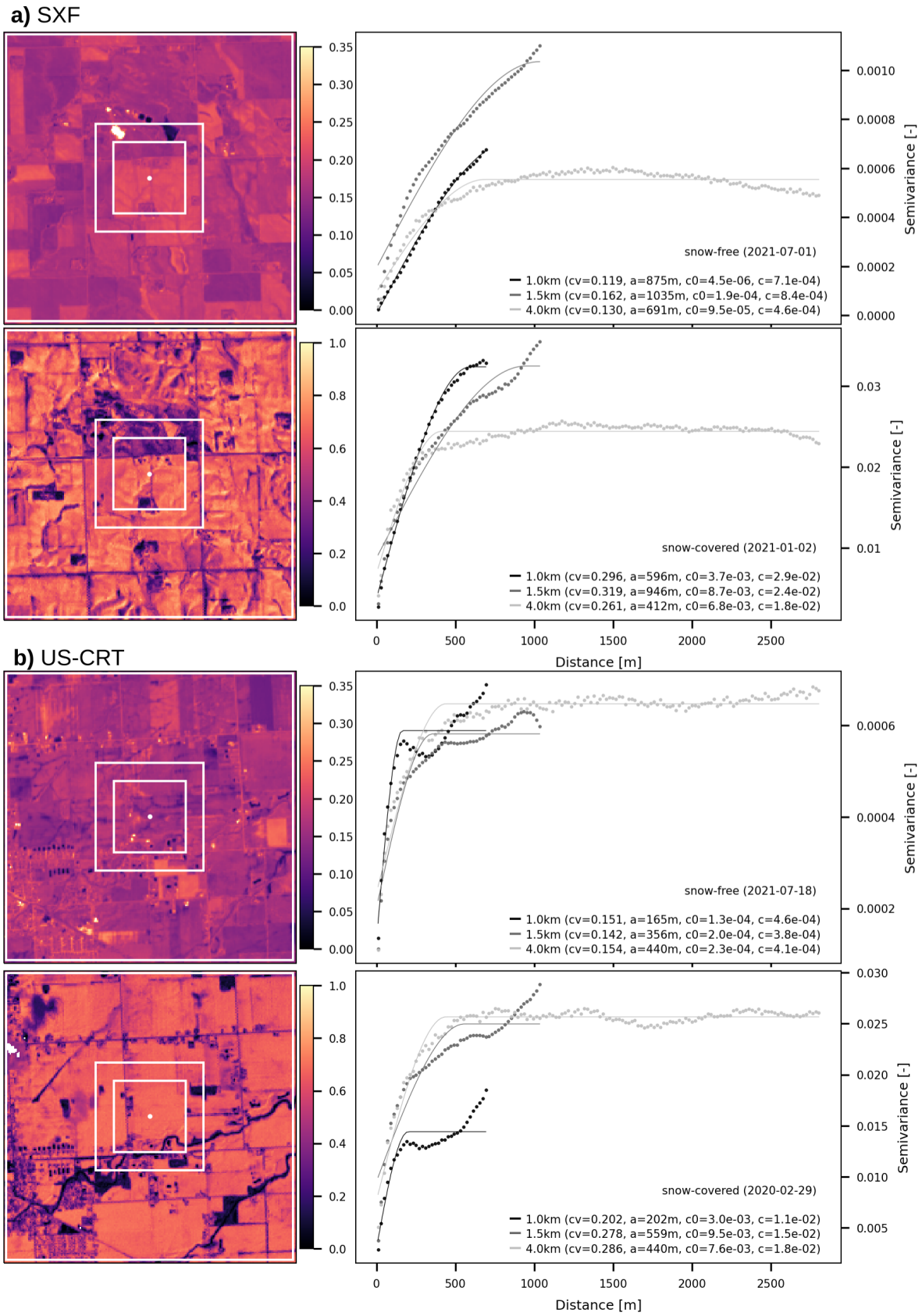

Figure S8: Same as Fig S4 but for (**a**) SXF - Sioux Falls and (**b**) US-CRT - Curtice Walter-Berger cropland.

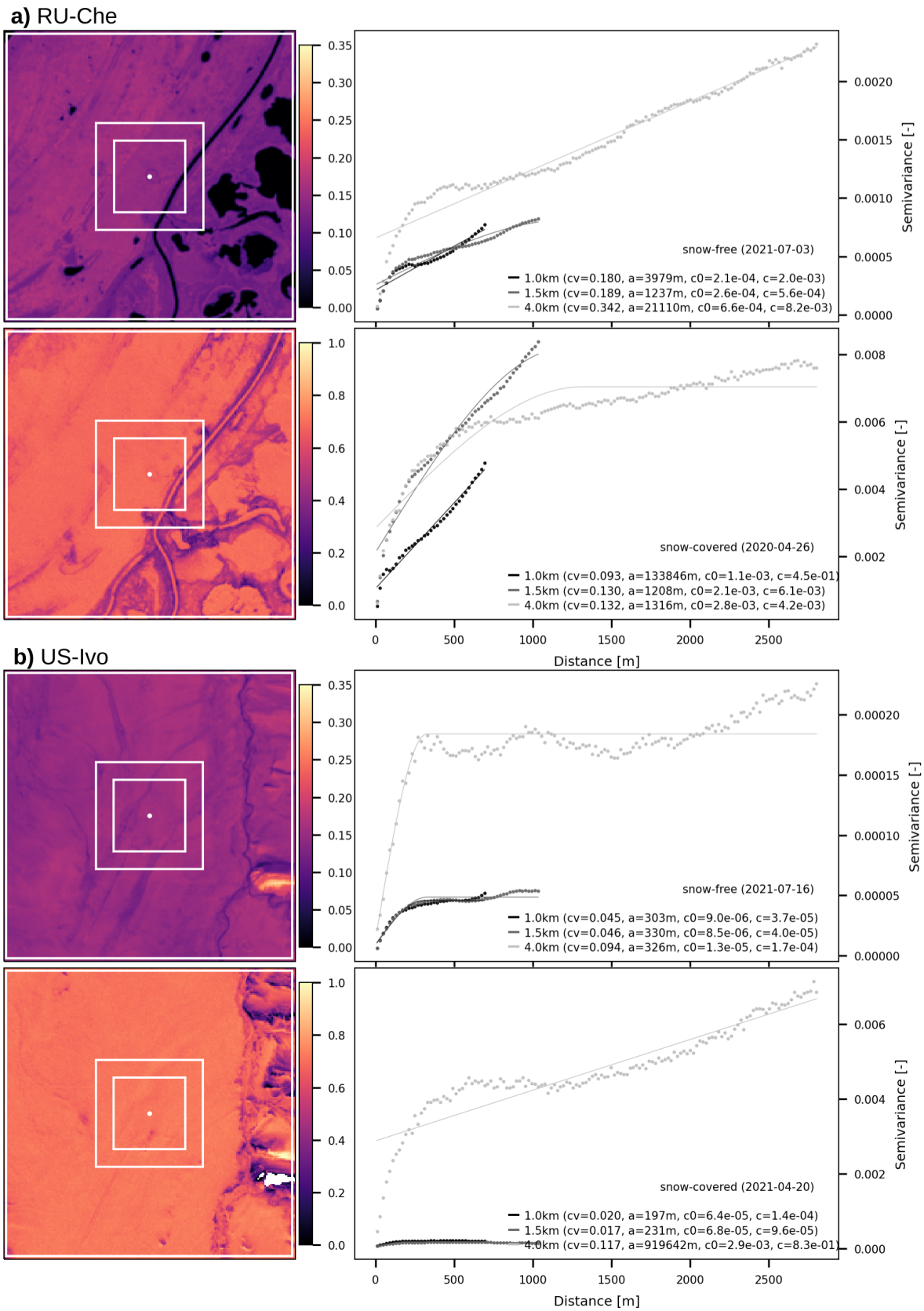

Figure S9: Same as Figure S4 but for (**a**) RU-Che - Cherski and (**b**) US-Ivo - Ivotuk.

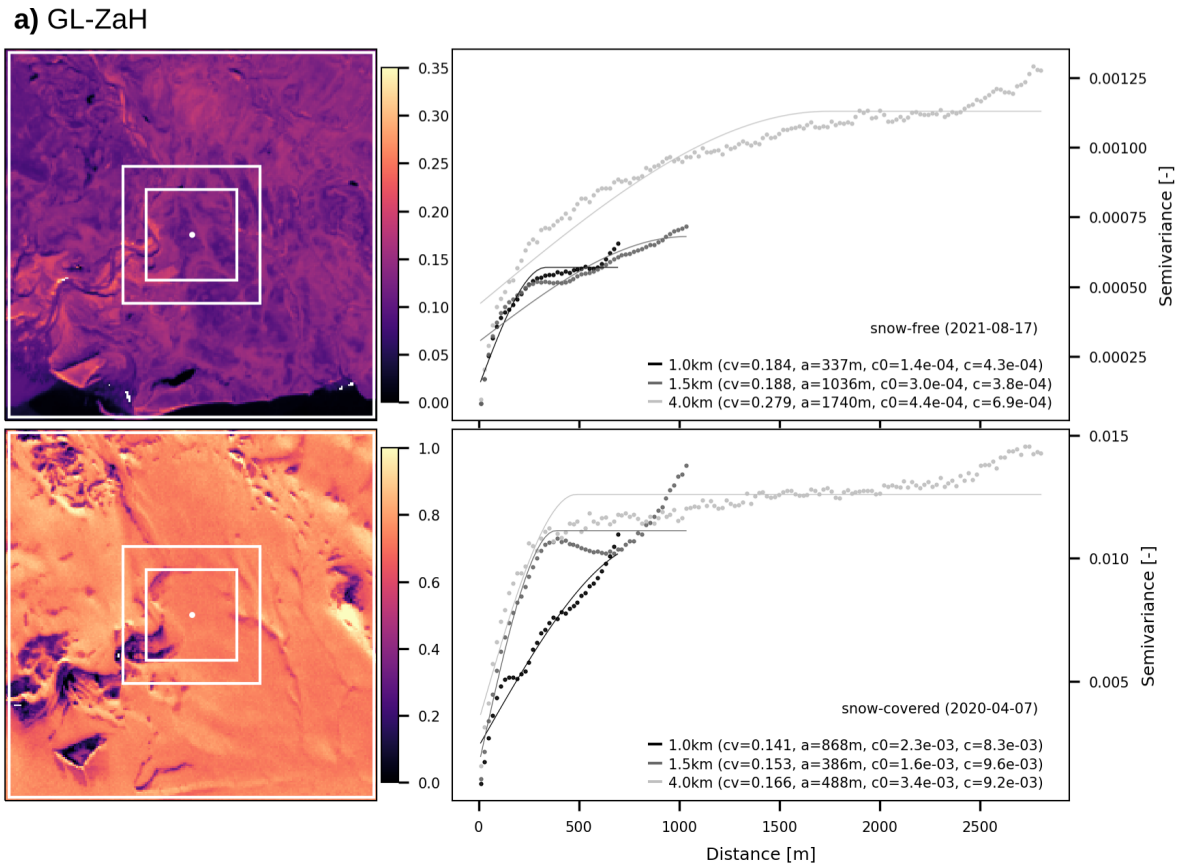

Figure S10: Same as Figure S4 but for GL-ZaH.