Supplementary Materials

Table S1 shows the formulas of the narrow-band indices used in this research.

Table S1. Formulas of the narrow-band indices used in this research.

Index	Formula
NDVI	$\frac{\rho_{800}-\rho_{660}}{\rho_{800}+\rho_{660}}$
NDVI705	$\frac{\rho_{750}-\rho_{705}}{\rho_{750}+\rho_{705}}$
mNDVI705	$\frac{\rho_{750}-\rho_{705}}{(\rho_{750}+\rho_{705}-2\cdot\rho_{445})}$
EVI*	$G\cdot\frac{\rho_{800}-\rho_{660}}{\rho_{800}+C_1.\rho_{660}-C_2.\rho_{480}+L}$
NDWI	$\frac{\rho_{860}-\rho_{1240}}{\rho_{860}+\rho_{1240}}$
NDII	$\frac{\rho_{819}-\rho_{1649}}{\rho_{819}+\rho_{1649}}$
CAI	$0.5\cdot\frac{\rho_{2005}-\rho_{2203}}{\rho_{2106}}$

*L=canopy background adjustment for correcting nonlinear, differential NIR and red radiant transfer through a canopy; C1 and C2 = coefficients of the aerosol resistance term (which uses the blue band to correct for aerosol influences in the red band); and G = a gain or scaling factor. Coefficients in EVI are, L=1, C1=6, C2=7.5, and G=2.5.

Table S2: We tested the relationships among all pairs of LiDAR metrics, measured by Pearson's R coefficient. Results show there are strong inter-correlations among LiDAR metrics in the three study sites, highlighted in red. Metrics were clumped around a few variables such as those (1) related to biomass, representing canopy foliage and stems (VVPint,, LAI, FC, FC_1ret), (2) related to canopy height (Hmean, Hmedian and Hmax), (3) related to vegetation spatial heterogeneity or canopy roughness (Hstd and CHMstd), and (4) related to spatial clumping of vegetation, which had the lowest correlations with other structural metrics.

Table S2. Relationships among all pairs of LiDAR metrics using Pearson's R coefficient for the three study sites (SJER, SOAP and TEAK). (VVPint: vegetation vertical profile, LAI: leaf area index, FC: fractional cover, FC_1ret: fractional cover from the first returns, Hmax: maximum height, Hmean: mean height, Hmedian: median height, Hstd: standard deviation of height, CHMstd: standard deviation of canopy height model). Colored coefficients represent best correlations for biomass (red-bold), height (red-underlined) and vegetation heterogeneity (red-italic).

Site		Clumping	VVP_{int}	LAI	FC	FC_1ret	Hmax	Hmean	Hmedian	Hstd	CHMstd
SJER	**Clumping**	1.00	0.56	0.56	0.56	0.58	0.42	0.35	0.34	0.43	0.38
	VVP_{int}	0.56	1.00	**0.98**	1.00	**0.99**	0.71	0.69	0.67	0.65	0.56
	LAI	0.56	**0.98**	1.00	**0.98**	**0.98**	0.67	0.65	0.63	0.60	0.51
	FC	0.56	**1.00**	**0.98**	1.00	**0.99**	0.71	0.69	0.67	0.65	0.56
	FC_1ret	0.58	**0.99**	**0.98**	**0.99**	1.00	0.69	0.67	0.65	0.62	0.53
	Hmax	0.42	0.71	0.67	0.71	0.69	1.00	0.95	0.91	0.94	0.90

SOAP	**Hmean**	0.35	0.69	0.65	0.69	0.67	0.95	1.00	0.99	0.84	0.78
	Hmedian	0.34	0.67	0.63	0.67	0.65	0.91	0.99	1.00	0.79	0.74
	Hstd	0.43	0.65	0.60	0.65	0.62	0.94	0.84	0.79	1.00	*0.96*
	CHMstd	0.38	0.56	0.51	0.56	0.53	0.90	0.78	0.74	*0.96*	1.00
	Clumping	1.00	0.61	0.73	0.70	0.69	0.24	-0.02	-0.07	0.32	0.26
	VVP_{int}	0.61	1.00	**0.91**	**0.98**	**0.97**	0.64	0.51	0.47	0.57	0.53
	LAI	0.73	**0.91**	1.00	**0.92**	**0.91**	0.47	0.28	0.24	0.46	0.40
	FC	0.70	**0.98**	**0.92**	1.00	**0.99**	0.59	0.41	0.37	0.55	0.51
	FC_1ret	0.69	**0.97**	**0.91**	**0.99**	1.00	0.55	0.39	0.35	0.51	0.46
	Hmax	0.24	0.64	0.47	0.59	0.55	1.00	0.89	0.81	0.93	0.91
TEAK	**Hmean**	−0.02	0.51	0.28	0.41	0.39	0.89	1.00	0.97	0.74	0.74
	Hmedian	−0.07	0.47	0.24	0.37	0.35	0.81	0.97	1.00	0.66	0.66
	Hstd	0.32	0.57	0.46	0.55	0.51	0.93	0.74	0.66	1.00	*0.98*
	CHMstd	0.26	0.53	0.40	0.51	0.46	0.91	0.74	0.66	*0.98*	1.00
	Clumping	1.00	0.78	0.79	0.81	0.79	0.71	0.68	0.67	0.67	0.64
	VVP_{int}	0.78	1.00	**0.97**	**0.99**	**0.98**	0.92	0.91	0.89	0.86	0.84
	LAI	0.79	**0.97**	1.00	**0.97**	**0.99**	0.87	0.88	0.86	0.81	0.77
	FC	0.81	**0.99**	**0.97**	1.00	**0.98**	0.91	0.89	0.87	0.86	0.83
	FC_1ret	0.79	**0.98**	**0.99**	**0.98**	1.00	0.87	0.86	0.84	0.81	0.78
	Hmax	0.71	0.92	0.87	0.91	0.87	1.00	0.97	0.95	0.97	0.96
	Hmean	0.68	0.91	0.88	0.89	0.86	0.97	1.00	0.99	0.92	0.89
	Hmedian	0.67	0.89	0.86	0.87	0.84	0.95	0.99	1.00	0.90	0.87
	Hstd	0.67	0.86	0.81	0.86	0.81	0.97	0.92	0.90	1.00	*0.99*
	CHMstd	0.64	0.84	0.77	0.83	0.78	0.96	0.89	0.87	*0.99*	1.00

Table S3 provides more detailed information about the statistics of the training-testing and validation datasets for the four structural variables at each study site (SJER, SOAP, and TEAK). Both datasets capture the variability within the area in term of biomass, height, heterogeneity and clumpiness.

Table S3. Statistics summary of training-testing (TT) and validation (V) datasets for SJER, SOAP and TEAK for each of the four structural variables. VVP and clumping scaled from 0 to 1, Hmean and CHMstd in meters.

SJER	**VVP_{int}**		**Hmean**		**CHMstd**		**Clumping**	
	TT	**V**	**TT**	**V**	**TT**	**V**	**TT**	**V**
Mean	0.260	0.138	5.101	3.588	1.980	1.294	0.187	0.100
Median	0.259	0.124	5.327	3.622	1.810	1.188	0.160	0.079
Standard Deviation	0.143	0.102	2.235	2.047	1.162	0.920	0.130	0.088
Kurtosis	−0.715	0.813	0.219	0.201	1.057	2.044	0.589	2.603
Skewness	0.113	0.871	−0.072	0.291	0.904	1.063	0.904	1.404
Range	0.718	0.840	15.552	15.374	7.861	6.901	0.788	0.855

SOAP	**VVP_{int}**		**Hm**		**CHMstd**		**Clumping**	
	TT	**V**	**TT**	**V**	**TT**	**V**	**TT**	**V**
Mean	0.623	0.675	13.878	10.176	7.669	5.417	0.515	0.648
Median	0.643	0.695	13.899	9.919	7.588	5.340	0.524	0.685
Standard Deviation	0.139	0.151	5.103	4.672	2.723	2.536	0.168	0.165
Kurtosis	1.680	1.461	0.172	−0.390	−0.184	−0.409	−0.444	0.235
Skewness	−1.008	−0.974	0.170	0.313	0.076	0.346	−0.271	−0.834
Range	0.948	0.981	34.092	28.267	17.125	15.171	0.908	0.947

TEAK	VVP_{int}		Hm		CHMstd		Clumping	
	TT	V	TT	V	TT	V	TT	V
Mean	0.569	0.54	15.793	15.50	7.977	7.98	0.463	0.41
Median	0.591	0.56	15.812	15.56	8.149	8.02	0.462	0.41
Standard Deviation	0.178	0.18	6.675	7.07	2.983	3.16	0.163	0.15
Kurtosis	–0.366	–0.47	–0.687	–0.64	–0.521	–0.47	–0.437	–0.41
Skewness	–0.476	–0.46	0.052	0.07	–0.063	0.02	0.048	0.03
Range	0.996	0.94	38.795	40.20	20.213	19.87	0.946	0.90

Table S4 showed the coefficients of determination (R^2) for correlations between the four LiDAR structural variables and the 10 most important optical metrics, based on ranking from the Random Forest regression. Results show that there is not a linear relationship between spectral metrics and structural variables.

Table S4. Relationship measured by R^2 between the four structural variables (i.e. biomass (VVP_{int}), canopy height (Hm), vegetation heterogeneity (CHMstd) and clumping) and the most important optical metrics according to the ranking of Random Forest.

Spectral Metric	VVP_{int}	Spectral Metric	Hmean	Spectral Metric	CHMstd	Spectral Metric	Clumping
CAI	0.31	**CAI**	0.12	**NDII**	0.32	**PC2**	0.01
Wtr1EdgeWvl	0.38	**EVI**	0.39	**CAI**	0.16	**PC1_NIR**	0.02
SOIL	0.06	**SOIL**	0.04	**NDWI**	0.46	**CAI**	0.37
PC2	0.00	**PC1_visible**	0.14	**SOIL**	0.02	**SHADE**	0.36
PC1_NIR	0.06	**PC1_NIR**	0.09	**PC1_visible**	0.10	**PC1_SWIR1**	0.18
SHADE	0.45	**PC2**	0.00	**PC1_NIR**	0.07	**Wtr1AbAr**	0.20
PC1	0.13	**PC1_SWIR2**	0.26	**PC2**	0.00	**SOIL**	0.02
GV	0.20	**SHADE**	0.27	**SHADE**	0.31	**PC1_visible**	0.05
Wtr1EdgeMag	0.18	**PC1_SWIR1**	0.19	**PC1_SWIR1**	0.32	**EWT**	0.48
EWT	0.19	**Wtr1EdgeMag**	0.26	**EVI**	0.45	**PC1_SWIR2**	0.17

Table S5 shows the highest ranking optical metrics from the Random Forests regression, determined by the increase in mean square error when the metric was removed. Results showed that all parts of the AVIRIS-classic spectrum play a role in the model.

CAI (Cellulose Absorption Index) was in the top four metrics for all structural variables, and the percentage of SOIL chosen for biomass, height and vegetation heterogeneity. At least one water-sensitive metric is within the top 8 most important metrics for all structural variables. The albedo was estimated as a proxy of the first principal component from the full spectrum and is one of the most important metrics for height and vegetation heterogeneity. All four of the PCAs that were run on all bands, only visible, only NIR, and only SWIR were in this list, indicating that bands across the spectrum contributed to the structural metrics.

Table S5. The rank of the highest 10 optical metrics for prediction based on the increase of mean square error (parentheses) when that metric was removed from the structural variable.

Biomass (VVP_{int})	**Height (Hm)**	**Vegetation Heterogeneity (CHMstd)**	**Clumping**
CAI (53.77)	CAI (52.84)	NDII (52.59)	PC2 (40.47)
Wtr1EdgeWvl (36.56)	EVI (45.78)	CAI (46.80)	PC1_NIR (37.92)
SOIL (34.41)	SOIL (37.74)	NDWI (38.78)	CAI (36.54)
PC2 (33.74)	PC1_visible (26.51)	SOIL (37.51)	SHADE (36.10)
PC1_NIR (25.34)	PC1_NIR (25.53)	PC1_visible (24.94)	PC1_SWIR1 (27.45)
SHADE (24.05)	PC2 (21.82)	PC1_NIR (21.94)	Wtr1AbAr (23.19)
PC1_visible (21.04)	PC1_SWIR2 (20.37)	PC2 (21.61)	SOIL (22.30)
GV (20.42)	SHADE (19.69)	SHADE (20.83)	PC1_visible (22.24)
Wtr1EdgeMag (1597)	PC1_SWIR1 (18.13)	PC1_SWIR1 (18.72)	EWT (21.29)
EWT (14.92)	Wtr1EdgeMag (16.36)	EVI (17.79)	PC1_SWIR2 (21.22)

Table S6 shows the four structural variables for each site when grouped into classes of low, medium and high values of each variable. We show the number of pixels per class for both LiDAR and IS data, and both the number of pixels in which both data types (LiDAR and IS) are in agreement for each class, and the percent of agreement for each data type.

Table S6. Similarity between classes of VVP_{int}, Hm, CHMstd and Clumping of LiDAR derive data and IS-derive data for SJER, SOAP and TEAK.

SJER		N pixels per class		N pixels overlapping agreement	% overlapping agreement	
		LiDAR	IS		LiDAR	IS
VVP_{int}	Low	46509	49021	36702	78.91	74.87
	Medium	51297	44031	29786	58.07	67.65
	High	49716	54470	43770	88.04	80.36
Hm	Low	45863	53937	36862	80.37	68.34
	Medium	48490	41188	24382	50.28	59.20
	High	51439	52397	39728	77.23	75.82
CHMstd	Low	66543	80116	60472	90.88	75.48
	High	80979	67406	61335	75.74	90.99
Clumping	Low	145285	146859	145257	99.98	98.91
	High	2237	663	635	28.39	95.78
SOAP		**N pixels per class**		**N pixels overlapping agreement**	**% overlapping agreement**	
		LiDAR	**IS**		**LiDAR**	**IS**
VVP_{int}	Low	7249	18098	6750	93.12	37.30
	Medium	25893	17973	11869	45.84	66.04
	High	21217	18288	13839	65.23	75.67
Hm	Low	16738	12098	9352	55.87	77.30
	Medium	18272	16750	8928	48.86	53.30
	High	19349	25511	16326	84.38	64.00
CHMstd	Low	26003	20074	17212	66.19	85.74
	High	28356	34285	25494	89.91	74.36
Clumping	Low	18238	21759	13476	73.89	61.93
	High	36121	32600	27759	76.85	85.15
TEAK		**N pixels per class**		**N pixels overlapping agreement**	**% overlapping agreement**	
		LiDAR	**IS**		**LiDAR**	**IS**
VVP_{int}	Low	73206	92348	54547	74.51	59.07
	Medium	134917	136923	98847	73.27	72.19
	High	142290	121142	105315	74.01	86.94
Hm	Low	73601	101069	62518	84.94	61.86
	Medium	135881	122670	90687	66.74	73.93
	High	140931	126674	104223	73.95	82.28
CHMstd	Low	143667	161123	121001	84.22	75.10

	High	206746	189290	166624	80.59	88.03
Clumping	**Low**	229724	253533	211123	91.90	83.27
	High	120689	96880	78279	64.86	80.80

Table S7 shows the CSTs defined for LiDAR and for IS data at each site (i.e. SJER, SOAP, and TEAK) and their percent of spatial agreement.

Table S7. Percentage of spatial coincidence between CSTs defined with LiDAR-derived data and IS-derive data for SJER, SOAP and TEAK.

SJER		IS		
		CST1	**CST2**	**CST3**
LiDAR	**CST1**	**81.19**	17.20	1.61
	CST2	27.75	**56.11**	16.14
	CST3	3.93	15.27	**80.8**

SOAP		IS					
		CST1	**CST2**	**CST3**	**CST4**	**CST5**	**CST6**
LIDAR	**CST1**	**88.03**	6.81	0.96	0.06	3.97	0.17
	CST2	27.96	**29.62**	12.20	7.46	16.09	6.67
	CST3	14.57	16.93	**46.08**	9.57	7.36	4.77
	CST4	4.79	8.64	17.59	**31.57**	17.28	20.13
	CST5	8.25	14.67	4.68	9.31	**53.15**	9.94
	CST6	2.76	3.28	5.37	15.23	11.30	**62.06**

TEAK		IS				
		CST1	**CST2**	**CST3**	**CST4**	**CST5**
LIDAR	**CST1**	**80.95**	3.90	14.25	0.73	0.17
	CST2	37.32	**55.74**	5.97	0.45	0.52
	CST3	16.79	1.24	**62.24**	16.37	3.36
	CST4	14.17	17.03	14.42	**25.15**	29.23
	CST5	11.45	0.23	10.72	38.49	**39.11**

Figure S1 show the agreement between the maximum heights within the 18 m LiDAR grid and field observations.

We collect field measurements between 2013 to 2016 in the lower two study sites (SJER and SOAP). We did not include data from the high elevation site (TEAK) in our analysis because it was difficult to correctly identify heights of individual trees in our measurements. Considering the two lower sites we found a correlation of $R^2 = 0.78$ between LiDAR determined maximum tree height and field measured maximum tree height. In addition, the U.S. Forest Service provided additional data from the Forest Inventory Analysis (FIA) program, mostly from the TEAK site compared to our LiDAR height data. The relationship between LiDAR maximum tree height and FIA maximum tree height measurements produced $R^2 = 0.90$. Combining both datasets (Figure 2) there is an agreement of $R^2 = 0.80$ and a RMSE of 2.38 m

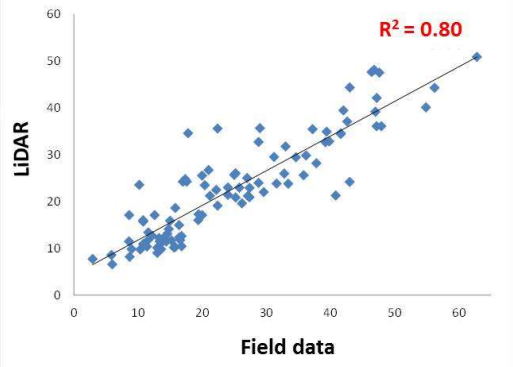
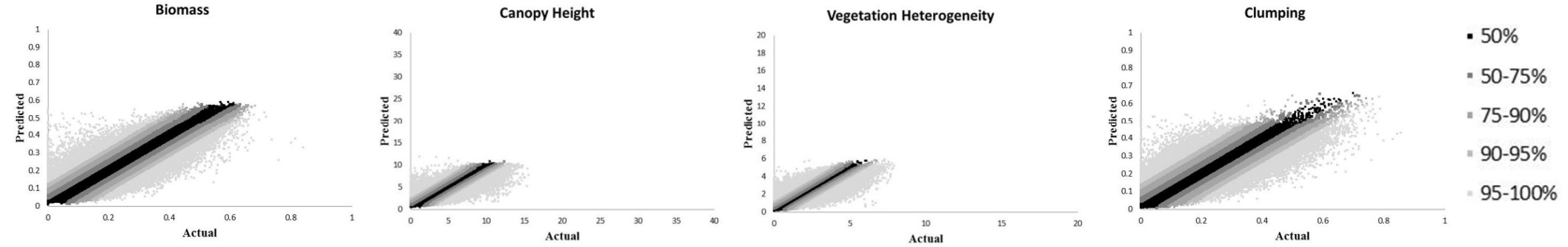


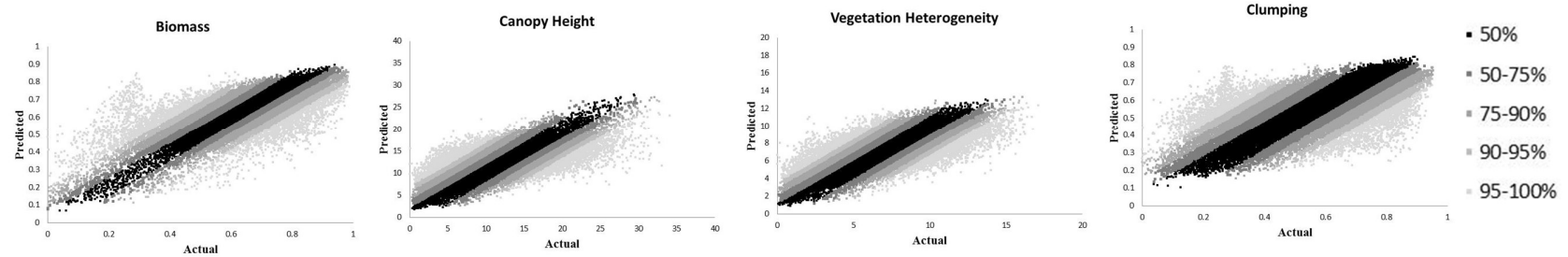
Figure S1. Relationship between maximum tree height from LiDAR measurements and field data.

Figure S2 shows the regression scatterplot relationship between LiDAR and IS-derived structural variables. The gray scale coded regions on the regression plot show >50% (■), 51%-75% (■), 76%-90% (■), 91%-95% (■), and> 96% (■) of all pixels in the three sites. The large number of pixels makes the data look continuous but each color represents a range of observations with the lowest errors. Results show how well the optical metrics non-linear models predict structural variables

SJER



SOAP



TEAK

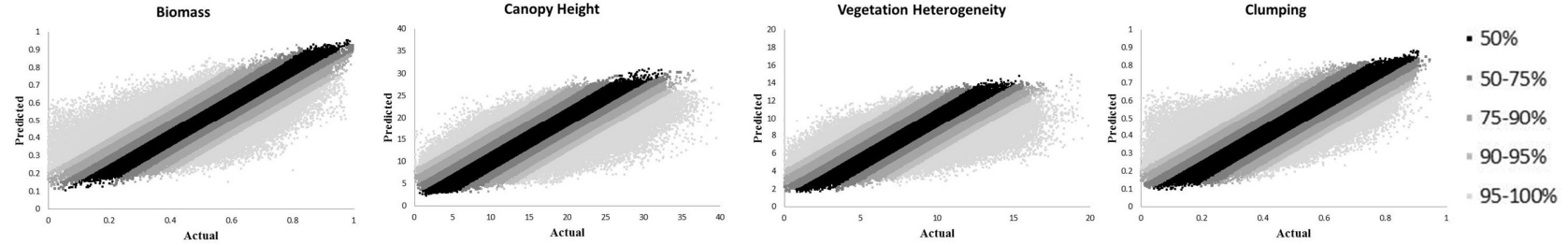


Figure S2. Relationship between LiDAR and IS-derived data for the four structural variables (i.e. biomass (VVP_{int}), canopy height (Hm), vegetation heterogeneity (CHM) and clumping), for the three study sites (SJER, SOAP and TEAK). The data looks continuous because of the high point density. The different gray levels represent the ranges >50%, 51%-75%, 76%-90% and 91%-95%, and >96% of all observations having the least errors.

Figure S3 shows the spatial distribution of the model errors for biomass (a, e, i), canopy height (b, f, j), vegetation heterogeneity (c, g, k), and clumping (d, h, l), at the three sites (SJER, SOAP, and TEAK). Gray scales show underestimation in darker to black shades, intermediate gray indicates no significant error and light gray to white represents areas of overestimations as indicated on the scale bar. Most areas have intermediate gray values and indicate low errors. The clumping had the greatest errors, both under and over estimates, seen as striping. And canopy heights at TEAK have particularly underestimated across a swath from upper right to lower left.

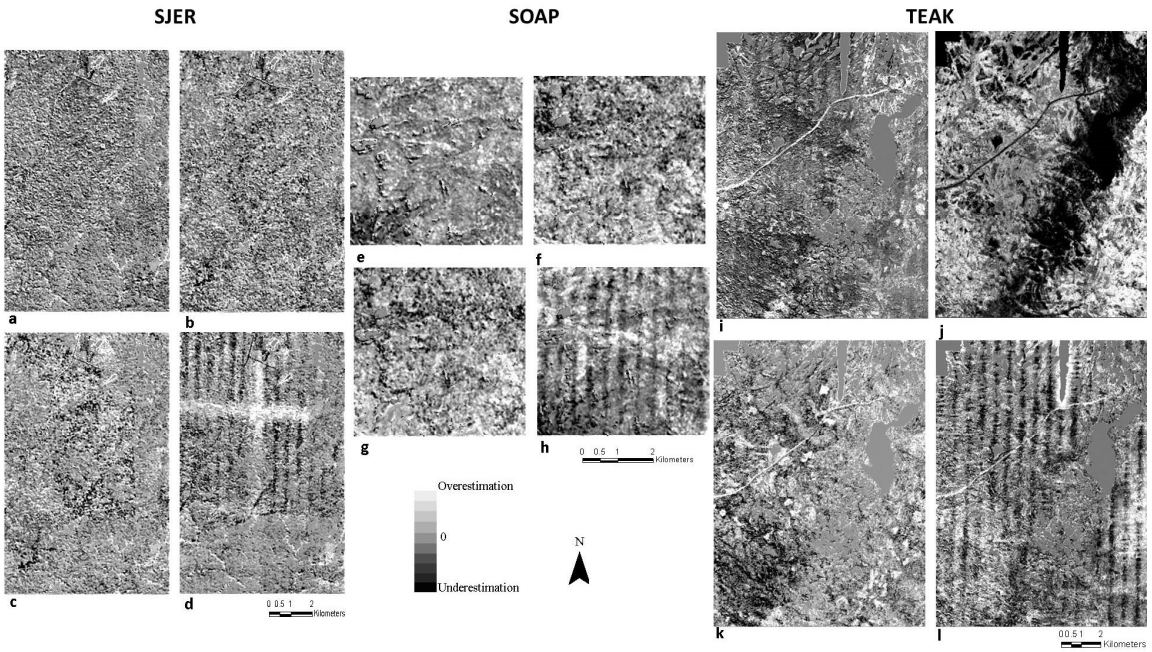
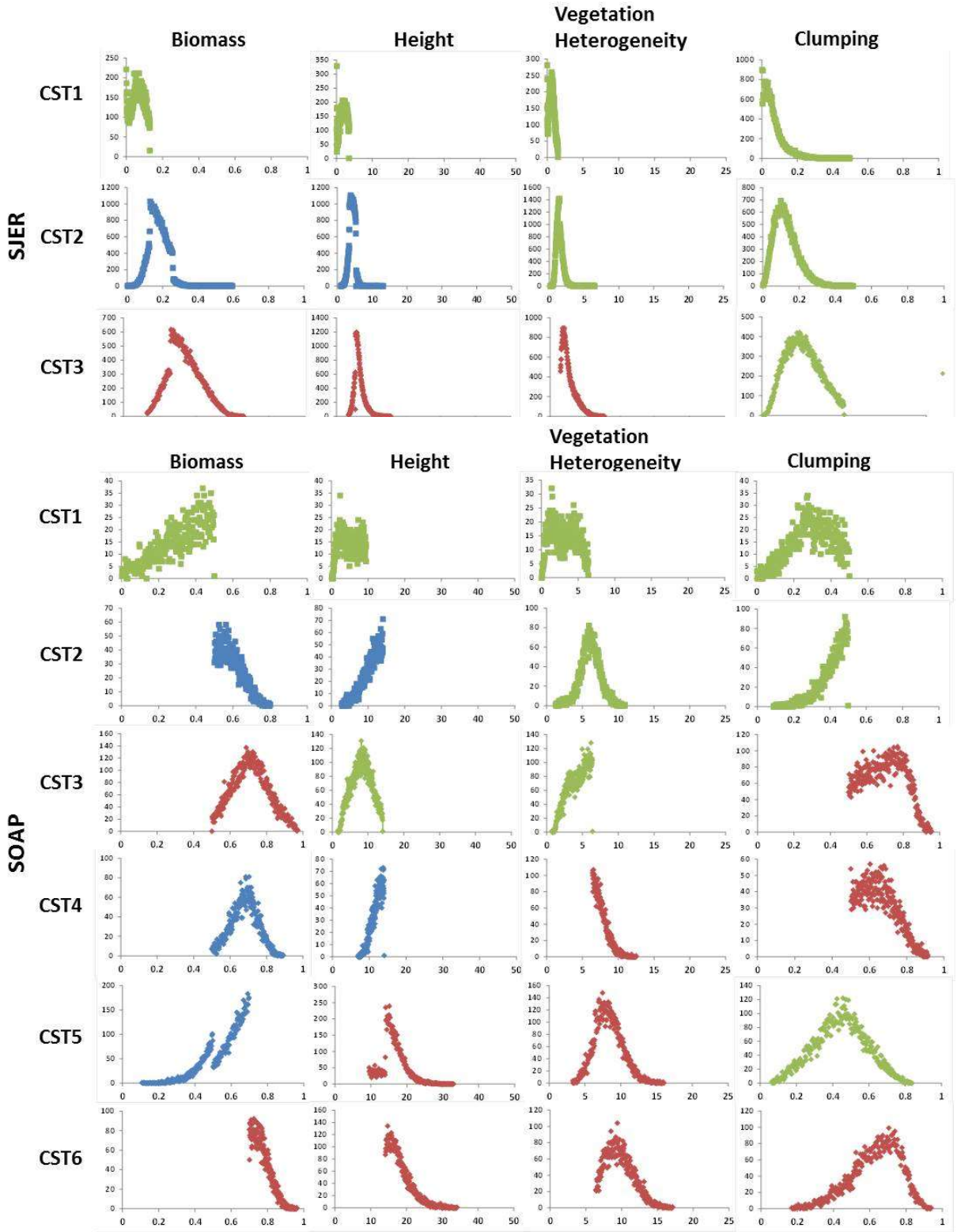


Figure S3. Spatial distribution of errors for biomass (a), canopy height (b), vegetation heterogeneity (c) and clumping (d) in SJER. Panels (e), (f), (g), and (h) respectively for SOAP, and (i), (j), (k), and (l) respectively for TEAK. Black, white and grey shades represent underestimation, overestimation and no error respectively. The coordinates in UTM 11N of the lower left corner of SJER, SOAP and TEAK are (254502, 4104126), (295848, 4098420), (317268, 4090230) respectively. The coordinates of the upper right corner of SJER, SOAP and TEAK are (260100, 4113000), (301572, 4102650), (327186, 4103676) respectively.

Figure S4 shows the histogram of the 4 structural variables that represent each CST for the three study sites. The different colors represent low, medium and high values of each variable in relation to each study site. For instance, high values of biomass or canopy height in the blue oak savanna represent low values in a high elevation montane conifer forest.


SJER
Biomass
Height
Vegetation Heterogeneity
Clumping
CST1
CST2
CST3
SOAP
Biomass
Height
Vegetation Heterogeneity
Clumping
CST1
CST2
CST3
CST4
CST5
CST6

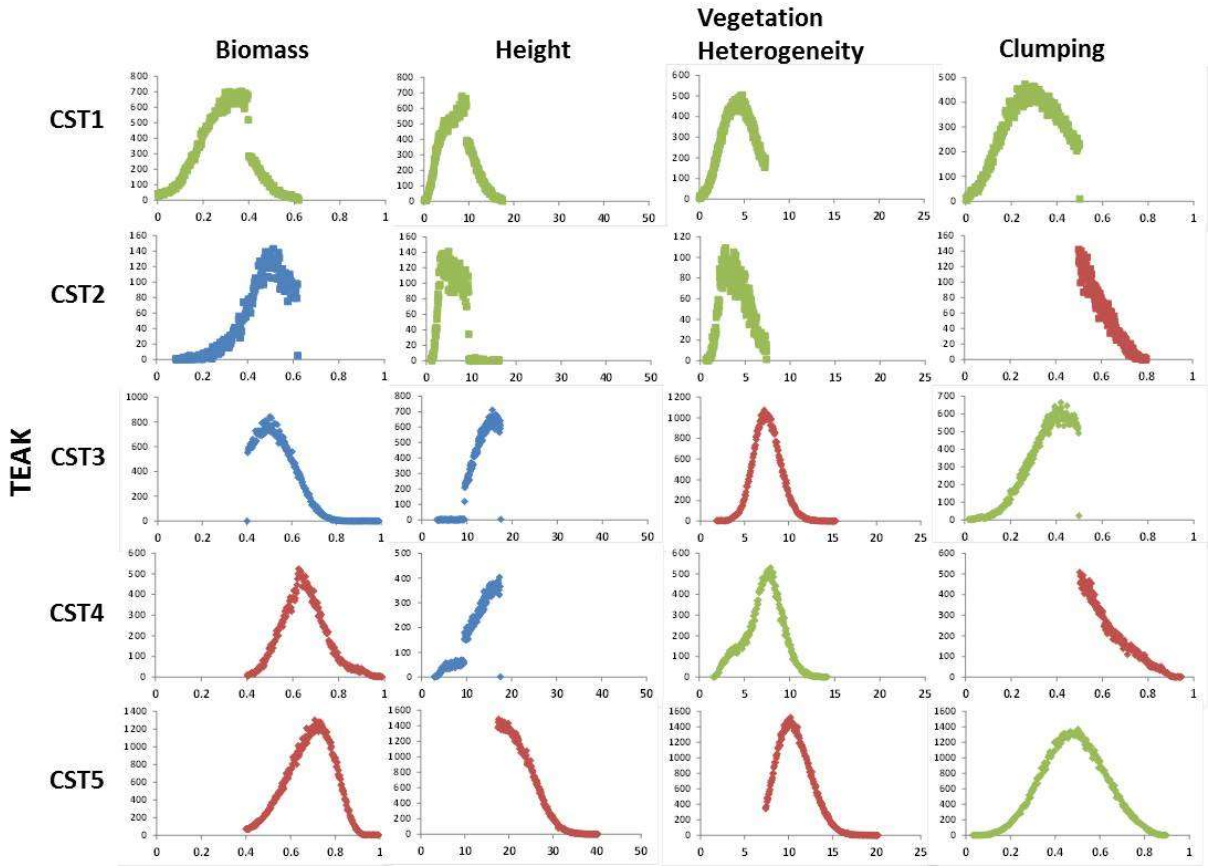


Figure S4. Histogram of the four structural variables (i.e. biomass, canopy height, vegetation heterogeneity and clumping) of each CST for SJER, SOAP and TEAK. Biomass and clumping values range from 0 to 1, height from 0 to 50 m and vegetation heterogeneity from 0 to 22 m. Red, blue and green represent high, medium and low values respectively.

Figure S5 shows three color composite images for biomass, canopy height and vegetation heterogeneity (as RGB) for the three study areas (SJER, SOAP, and TEAK). The complexity of the combined spatial and color patterns are indicative of the number of CST classes for each site, with SJER having 3 CST classes, SOAP having 5 classes and TEAK having 5 classes.

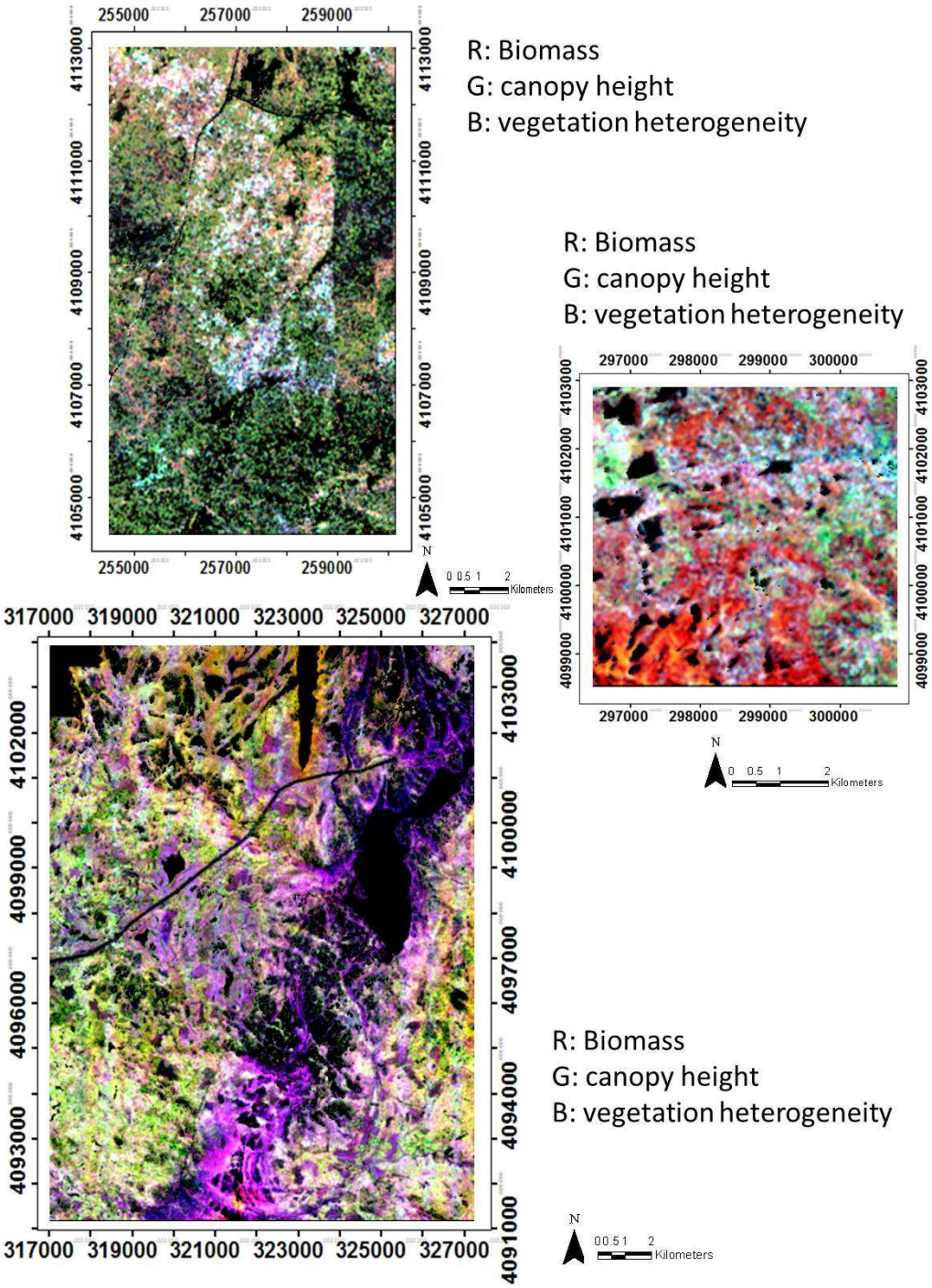


Figure S5. RBG composition based on scaled (0-1) ranges for the three most important canopy structural variables. Spatial/color patterns follow the CST classes for each site. .