Supplement of Biogeosciences, 23, 2235–2259, 2026
https://doi.org/10.5194/bg-23-2235-2026-supplement

Biogeosciences Open Access EGU

Supplement of

Uncertainty Assessment in Deep Learning-based Plant Trait Retrievals from Hyperspectral data

Eya Cherif et al.

Correspondence to: Eya Cherif (eya.cherif@uni-leipzig.de)

Supplement

Table S1: Detailed information on the collected Datasets (DSs) for training the deep learning model and their corresponding studies. For some of the data sets, the link between spectra and traits was established based on the dominant species, the sampling plot identifier or the date of measurements according to the data set's available metadata. For the data set with asterisks (DS 17) the weighted species density is used to associate trait values. This table was adopted from Cherif et al. (2023).

Data sets	Instrument	Method	Number of bands	Land Cover	Number of samples	Citation
DS1	ASD Fieldspec3 + SVC HR-1024TM	Proximal canopy measurements	2151	Grassland	73	Van Cleemput et al. (2020)
DS2,3	ASD Fieldspec3	Proximal canopy measurements	2101	Grassland	658	Kattenborn. (2017) and Kattenborn et al. (2019)
DS4	ASD Fieldspec3	Proximal canopy measurements	2151	Grassland	549	Van Cleemput et al. (2019)
DS5	SVC HR-1024i	Proximal canopy measurements	2151	Tundra	511	Serbin et al. (2017)
DS6	SVC HR-1024i	Proximal canopy measurements	2151	Tundra	129	Serbin et al. (2019a)
DS7 & 30-42	AVIRIS next-gen	Airborne	426	Mix	1126	Wang. (2020) and Wang et al. (2020)
DS8	SVC HR-1024i	Proximal canopy measurements	2151	Tundra	43	Serbin et al. (2016)
DS9	SVC HR-1024i + SE PSR+	Proximal canopy measurements	2151	Tundra	178	Serbin et al. (2019b)

DS10	SVC HR-1024i	Proximal canopy measurements	2151	Shrubland	22	Serbin et al. (2018)
DS11	SE PSR+	Proximal canopy measurements	2151	Crops	174	Burnett et al. (2020)
DS12&15	AVIRIS-Classic	Airborne	223	Forest (broad- and needle-leaf)	304	Singh. (2015) and Singh et al. (2015)
DS13	ASD FieldSpec Pro FR	Proximal canopy measurements	2011	Crops	330	Pimstein et al. (2004 and 2005) and Herrmann et al. (2011)
DS14	NEON AOP	Airborne	351	Forests (temperate broadleaf)	59	Chlus et al. (2020)
DS16	FieldSpec3	Proximal canopy measurements	1950	Grassland	96	Cerasoli et al. (2018)
DS17 *	AISA Dual system (Specim, Ltd. Finland)	Airborne	75	Grassland	70	Pottier et al. (2014)
DS18	HySpex	Airborne	332	Crops	32	Hank et al. (2016)
DS19	HySpex	Airborne	332	Crops	91	Hank et al. (2016)
DS20	HyMap	Airborne	125	Crops	147	Hank at al. (2015a)
DS21	HyMap	Airborne	125	Crops	127	Hank at al. (2015a)
DS22	Apex	Airborne	288	Crops	106	Hank at al. (2015b)
DS23	ASD FieldSpec 3 JR	Proximal canopy measurements	2151	Crops	18	Wocher et al. (2018)
DS25	APEX	Airborne	248	Forest	50	Ewald et al. (2018)
DS26	APEX	Airborne	233	Grassland	90	Ewald et al. (2020)

DS27	NEON AOP	Airborne	187	Forest	80	Chlus and Townsend, (2020)
DS28	HrHSI	Airborne	2150	Grassland	60	Dao et al. (2021)
DS29	ASD FieldSpec 3	Proximal canopy measurements	2150	Mix	173	McCarty and Grimm. (2017)
DS43	ASD	Proximal canopy measurements	2150	Forest, Crops	24	Roberts et al. (2017)
DS44	AisaFENIX	Airborne	627	Crops	45	Brown et al. (2024)
DS45	AisaFENIX	Airborne	622	Forest	47	Brown et al. (2024)
DS46	AVIRIS-NG	Airborne	425	Forest	28	Brown et al. (2024)
DS47	Casi+Sasi	Airborne	220	Forest	166	Gravel et al, (2024)
DS48	AVIRIS-NG	Airborne	360	Forest	96	Zheng et al. (2023)
DS49	MSV 500 + Middleton, Spectral Vision	Proximal canopy measurements	768	Crops	94	Wang et al. (2016)
DS50	AVIRIS-NG (SHIFT)	Airborne	427	Forest	1825	Brodrick et al. (2023) and Chadwick et al. (2023)

References

- Burnett, A. C., Serbin, S. P., and Rogers, A.: Leaf and canopy spectroscopy and biochemical data of field-grown Cucurbita pepo under two stresses, EcoSIS [data set], https://doi.org/10.21232/RLmYbmE3, 2020.
- Brodrick, P. G., Pavlick, R., Bernas, M., Chapman, J. W., Eckert, R., Helmlinger, M., Hess-Flores, M., Rios, L. M., Schneider, F. D., Smyth, M. M., Eastwood, M., Green, R. O., Thompson, D. R., Chadwick, K. D., and Schimel, D. S.: SHIFT: AVIRIS-NG L2A Unrectified Surface Reflectance Version 1, ORNL DAAC [data set], https://doi.org/10.3334/ORNLDAAC/2184, 2023.
- Brown, L. A., Morris, H., MacLachlan, A., D'Adamo, F., Adams, J., Lopez-Baeza, E., Albero, E., Martínez, B., Sánchez-Ruiz, S., Campos-Taberner, M., and Lidón, A.: Hyperspectral leaf area index and chlorophyll retrieval over forest and row-structured vineyard canopies, Remote Sens. (Basel), 16, 2066, https://doi.org/10.3390/rs16122066, 2024.
- Cerasoli, S., Campagnolo, M., Faria, J., Nogueira, C., and Caldeira, M. C.: Hyperspectral reflectance, gas exchange and meteorological conditions in grassland plots undergoing different fertilization regimes in Central Portugal from March to June 2016, PANGAEA [data set], https://doi.org/10.1594/PANGAEA.892837, 2018.

- Chadwick, K.D., N. Queally, T. Zheng, J. Cryer, C. Vanden Heuvel, C. Villanueva-Weeks, C. Ade, L. Anderegg, Y. Angel, B. Baker, I. Boving, R.K. Braghiere, P. Brodrick, P. Campbell, K.C. Cushman, F. Davis, P.D. Dao, A. Dibartolo, R. Eckert, K. Grant, B. Heberlein, M. Johnson, J. Joutras, C. Kibler, M. Klope, K. Kovach, A. Kreisberg, P. Lovegreen, A.J. Maguire, C. Mcmahon, K. Miner, C. Nickles, F. Ochoa, J.P. Ocón, A. Ongjoco, E. Ordway, M. Park, R. Pavlick, A.M. Raiho, D.A. Roberts, D.S. Schimel, F.D. Schneider, K. Thompson, P. Townsend, E. Vermeer, N. Vinod, and K. Zumdahl.: SHIFT: Photosynthetic and Leaf Traits, Santa Barbara County, 2022, ORNL DAAC, Oak Ridge [data set], Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/2233, 2023.
- Cherif, E., Feilhauer, H., Berger, K., Dao, P.D., Ewald, M., Hank, T.B., He, Y., Kovach, K.R., Lu, B., Townsend, P.A. and Kattenborn, T.: From spectra to plant functional traits: Transferable multi-trait models from heterogeneous and sparse data, Remote Sens. Environ., 292, 113580, https://doi.org/10.1016/j.rse.2023.113580, 2023.
- Chlus, A. and Townsend, P. A.: Seasonal canopy spectra and traits, Blackhawk Island, WI, EcoSIS [data set], https://ecosis.org/package/seasonal-canopy-spectra-and-traits--blackhawk-island--wi (last access: 8 March 2026).
- Chlus, A., Kruger, E. L., and Townsend, P. A.: 3D LMA canopy level spectra, EcoSIS [data set], https://ecosis.org/package/3d-lma-canopy-level-spectra (last access: 8 March 2026), 2020a.
- Chlus, A., Kruger, E. L., and Townsend, P. A.: Mapping three dimensional variation in leaf mass per area with imaging spectroscopy and LiDAR in a temperate broadleaf forest, Remote Sens. Environ., 250, 112043, https://doi.org/10.1016/j.rse.2020.112043, 2020b.
- Dao, P. D., Axiotis, A., and He, Y.: Mapping native and invasive grassland species and characterizing topography-driven species dynamics using high spatial resolution hyperspectral imagery, Int. J. Appl. Earth Obs. Geoinf., 104, 102542, https://doi.org/10.1016/j.jag.2021.102542 , 2021.
- Ewald, M., Skowronek, S., Aerts, R., Dolos, K., Lenoir, J., Nicolas, M., Warrie, J., Hattab, T., Feilhauer, H., Honnay, O. and Garzon-Lopez, C.X.: Analyzing remotely sensed structural and chemical canopy traits of a forest invaded by Prunus serotina over multiple spatial scales, Biol. Invasions, 20, 2257–2271, 2018.
- Ewald, M., Skowronek, S., Aerts, R., Lenoir, J., Feilhauer, H., Van De Kerchove, R., Honnay, O., Somers, B., Garzón-López, C.X., Rocchini, D. and Schmidtlein, S.: Assessing the impact of an invasive bryophyte on plant species richness using high-resolution imaging spectroscopy, Ecol. Indic., 110, 105882, https://doi.org/10.1016/j.ecolind.2019.105882, 2020.
- Gravel, A., Laliberté, E., and Kalacska, M.: Foliar functional trait mapping of a mixed temperate forest using imaging spectroscopy, Federated Research Data Repository (FRDR) [data set], https://doi.org/10.20383/103.0922, 2024.
- Hank, T. B., Locherer, M., Richter, K., and Mauser, W.: Neusling (Landau a.d. Isar) 2012 – A multitemporal and multisensoral agricultural EnMAP preparatory flight campaign (Datasets), V. 1.2, GFZ Data Services [data set], https://doi.org/10.5880/enmap.2016.007, 2016.
- Hank, T. B., Richter, K., and Mauser, W.: Neusling (Landau a.d. Isar) 2009 – An agricultural EnMAP preparatory flight campaign using the HyMap instrument (Datasets), GFZ Data Services [data set], https://doi.org/10.5880/enmap.2015.002, 2015a.
- Hank, T. B., Richter, K., Locherer, M., Frank, T., and Mauser, W.: Neusling (Landau a.d. Isar) 2011 – An agricultural EnMAP preparatory flight campaign using the APEX instrument (Datasets), GFZ Data Services [data set], https://doi.org/10.5880/enmap.2015.003, 2015b.
- Herrmann, I., Pimstein, A., Karnieli, A., Cohen, Y., Alchanatis, V., and Bonfil, D. J.: LAI assessment of wheat and potato crops by VENμS and Sentinel-2 bands, Remote Sens. Environ., 115, 2141–2151, 2011.

- Kattenborn, T., Schiefer, F., and Schmidtlein, S.: Canopy reflectance plant functional gradient IFGG/KIT, EcoSIS [data set], https://doi.org/10.21232/krt4-6x67, 2017.
- Kattenborn, T., Fassnacht, F. E., and Schmidtlein, S.: Differentiating plant functional types using reflectance: Which traits make the difference?, Remote Sens. Ecol. Conserv., 5, 5–19, 2019.
- McCarty, J. and Grimm, A.: Hawaii Volcanoes National Park February 2017 Spectra, EcoSIS [data set], https://ecosis.org/package/hawaii-volcanoes-national-park-february-2017-spectra (last access: 8 March 2026), 2017.
- Pottier, J., Malenovský, Z., Psomas, A., Homolová, L., Schaepman, M. E., Choler, P., Thuiller, W., Guisan, A., and Zimmermann, N. E.: Data from: Modelling plant species distribution in alpine grasslands using airborne imaging spectroscopy [Dataset], Dryad, https://doi.org/10.5061/dryad.n13hn, 2014.
- Pimstein, A., Bonfil, D. J., and Karnieli, A.: Wheat canopy spectra collected throughout growing season, EcoSIS [data set], https://ecosis.org/package/wheat-canopy-spectra-collected-throughout-growing-season (last access: 8 March 2026), 2004/2005.
- Roberts, D.: Canopy-level reflectance spectra and leaf area for poplars in Washington, EcoSIS [data set], https://data.ecosis.org/dataset/canopy-level-reflectance-spectra-and-leaf-area-for-poplars-in-washington (last access: 8 March 2026), 2017.
- Serbin, S. P., Rogers, A., and Ely, K.: NGEE Arctic 2016 averaged canopy spectral reflectance Seward Peninsula Alaska, EcoSIS [data set], https://ecosis.org/package/ngee-arctic-2016-averaged-canopy-spectral-reflectance-seward-peninsula-alaska (last access: 8 March 2026), 2016.
- Serbin, S. P., Yang, D., Meng, R., McMahon, A., Hantson, W., Hayes, D., and Ely, K.: NGEE Arctic 2017 canopy spectral reflectance Seward Peninsula Alaska, EcoSIS [data set], https://ecosis.org/package/ngee-arctic-2017-canopy-spectral-reflectance-seward-peninsula-alaska (last access: 8 March 2026), 2017.
- Serbin, S. P., Yang, D., Meng, R., McMahon, A., Hantson, W., Hayes, D., and Ely, K.: NGEE Arctic 2018 canopy spectral reflectance Kougarok Watershed Seward Peninsula Alaska, EcoSIS [data set], https://ecosis.org/package/ngee-arctic-2018-canopy-spectral-reflectance-kougarok-watershed-seward-peninsula-alaska (last access: 8 March 2026), 2018.
- Serbin, S. P., Yang, D., Hantson, W., and Ely, K. S.: NGEE Arctic 2019 canopy spectral reflectance Seward Peninsula Alaska, EcoSIS [data set], https://ecosis.org/package/ngee-arctic-2019-canopy-spectral-reflectance-seward-peninsula-alaska (last access: 8 March 2026), 2019a.
- Serbin, S. P., Lieberman-Cribbin, W., Ely, K., and Rogers, A.: NGEE Arctic leaf spectral reflectance and transmittance data 2014 to 2016 Utqiaġvik (Barrow) Alaska, EcoSIS [data set], https://doi.org/10.5440/1437044, 2019b.
- Singh, A.: Mapping canopy foliar chemical and morphological traits using imaging spectroscopy, EcoSIS [data set], https://ecosis.org/package/mapping-canopy-foliar-chemical-and-morphological-traits-using-imaging-spectroscopy (last access: 8 March 2026), 2015.
- Singh, A., Serbin, S. P., McNeil, B. E., Kingdon, C. C., and Townsend, P. A.: Imaging spectroscopy algorithms for mapping canopy foliar chemical and morphological traits and their uncertainties, Ecol. Appl., 25, 2180–2197, 2015.
- Van Cleemput, E., Roberts, D. A., Honnay, O., and Somers, B.: Individual forb and grass species spectra measured on field patches and on a black table, EcoSIS [data set], https://doi.org/10.21232/ntk7-zn04, 2019.
- Van Cleemput, E., Helsen, K., Feilhauer, H., Honnay, O., and Somers, B.: Canopy spectra of individual herbaceous species measured on black table, EcoSIS [data set], https://doi.org/10.21232/ss03-g783, 2020.

- Wang, D. R., Wolfrum, E. J., Virk, P., Ismail, A., Greenberg, A. J., and McCouch, S. R.: Robust phenotyping strategies for evaluation of stem non-structural carbohydrates (NSC) in rice, J. Exp. Bot., 67, 6125–6138, 2016.
- Wang, Z.: Canopy spectra to map foliar functional traits over NEON domains in eastern United States, EcoSIS [data set], https://doi.org/10.21232/e2jt-5209, 2020.
- Wang, Z., Chlus, A., Geygan, R., Ye, Z., Zheng, T., Singh, A., Couture, J. J., Cavender-Bares, J., Kruger, E. L., and Townsend, P. A.: Foliar functional traits from imaging spectroscopy across biomes in eastern North America, New Phytol., 228, 494–511, 2020.
- Wocher, M., Berger, K., Danner, M., Mauser, W., and Hank, T.: Physically-based retrieval of canopy equivalent water thickness using hyperspectral data, Remote Sens., 10, 1924, https://doi.org/10.3390/rs10121924, 2018.
- Zheng, T., Ye, Z., Singh, A., Desai, A. R., Krishnayya, N. S. R., Dave, M. G., and Townsend, P. A.: Plot-level traits and their corresponding image spectra from the Western Ghats of India, EcoSIS [data set], https://doi.org/10.21232/5nt5bd9p, 2023.

Table S2: Number of training data samples by vegetation type.

LandCover	numSamples
Crops	947
Forest	2639
Grassland	1403
Mix	202
Shrubland	312
Tundra	673
soil	12

Table S3: The distribution of trait observations used for the deep learning model. List of traits and units: Leaf Mass per Area (g/m²) = LMA, Leaf Area Index (m²/m²) = LAI, Nitrogen content (mg/cm²) = N, Chlorophyll content (µg/cm²) = Chl, Equivalent Water Thickness (mg/cm²) = EWT, Carotenoid content (µg/cm²)= Car.

	LMA	Nitrogen	LAI	Chl	EWT	Car
count	5773	4483	1775	3348	4175	2025
mean	117.352061	0.217665	3.427870	40.040327	16.695106	8.406383
std	89.580332	0.120803	1.738253	14.363197	14.506370	2.914102
min	0.136667	0.000707	0.063333	4.448305	0.048333	1.182576
0.25	60.017268	0.146366	2.080000	29.284488	10.956678	6.580399
0.5	89.911887	0.205885	3.400000	39.013221	14.156210	8.319308
0.75	156.767349	0.268898	4.730611	50.302665	18.365852	10.172243
max	682.176000	1.132215	8.770000	229.497477	513.806360	40.443217

S1. Preprocessing pipeline

All 50 compiled datasets were pre-processed using the same standardized pipeline, without any dataset-specific deviations. The procedure followed here is based on the analysis of Cherif et al. (2023) and summarized as follows:

First, reflectance spectra were quality-checked. Reflectance values outside the physical range were masked: values below zero were set to missing, and values greater than one were treated as spurious spikes. These missing values were then replaced by the mean of the nearest valid neighbors.

Second, all datasets were resampled to a common 1 nm resolution to harmonize the diverse measurements from different sensors (from proximal and airborne), which varied in spectral sampling and band centers. This resampling was not intended to enhance spectral resolution or recover fine-scale features absent in coarser sensors, but to provide a uniform input representation required for the deep learning model. Most datasets were already acquired at 1 nm resolution, so upsampling was preferred over downsampling to minimize manipulation of the data and avoid loss of information from higher-resolution sensors.

Third, spectral intervals strongly affected by atmospheric water absorption were excluded uniformly across all datasets. In the implementation, the following wavelength ranges were removed: 1351–1430 nm, 1801–2050 nm, and 2451–2500 nm.

Finally, the remaining reflectance data were smoothed using a Savitzky–Golay filter applied independently to three contiguous segments of the spectrum: 400–1350 nm, 1431–1800 nm, and 2051–2451 nm. Each segment was filtered with a window size of 65 nm and a polynomial order of one.

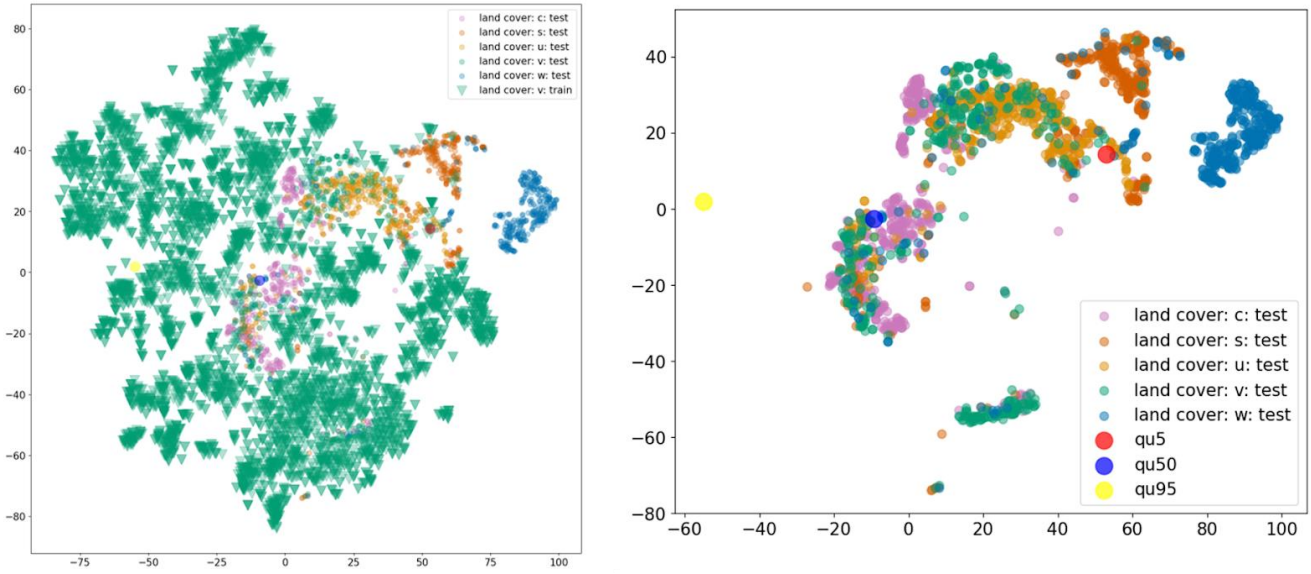


Figure S1: T-SNE visualization of the last embedding layer of training vegetation data and out-of-domain (OOD) data. The green triangle points are vegetation points from the training data and the dotted points are samples from an EnMAP scene from different scene components.

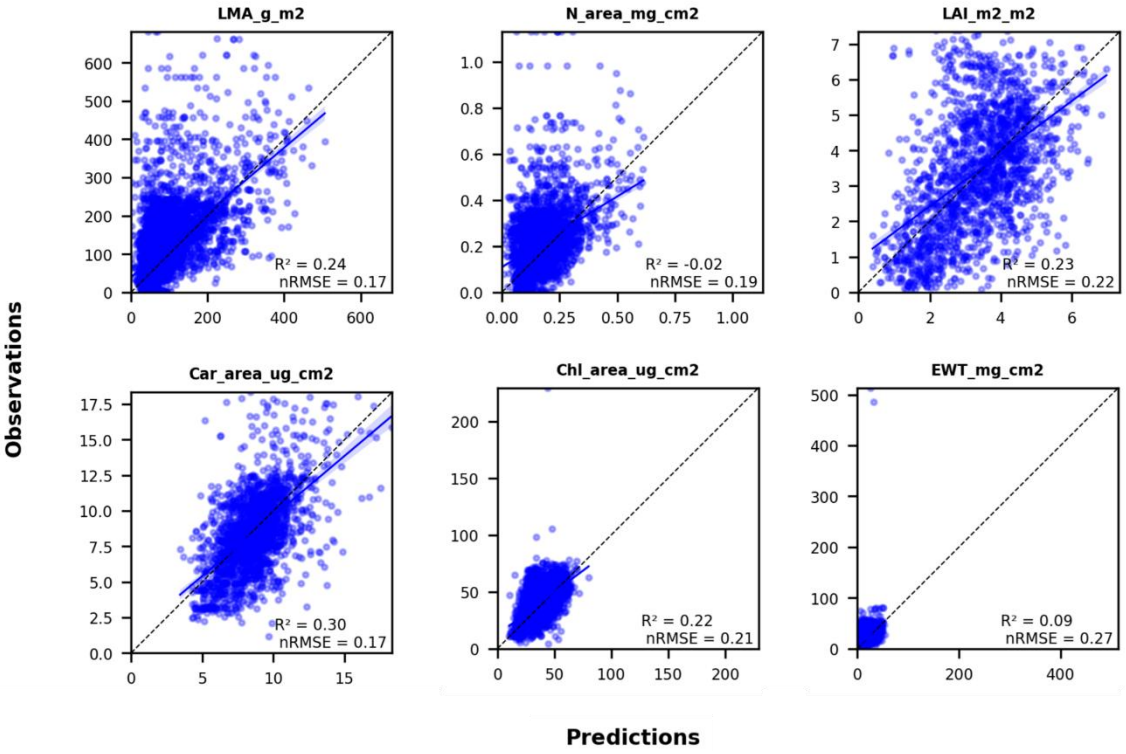


Figure S2a: Leave-one-dataset-out cross-validation (LODO-CV) results for six plant traits: Leaf Mass per Area (LMA), Nitrogen content (N), Leaf Area Index (LAI), Chlorophyll content (Chl), Equivalent Water Thickness (EWT), and Carotenoid content (Car). Each panel shows observed values (y-axis) versus model predictions (x-axis), with the 1:1 line (dashed) as reference. The perfoemance of the model is summarized by the coefficient of determination (R^2) and normalized root mean square error (nRMSE).

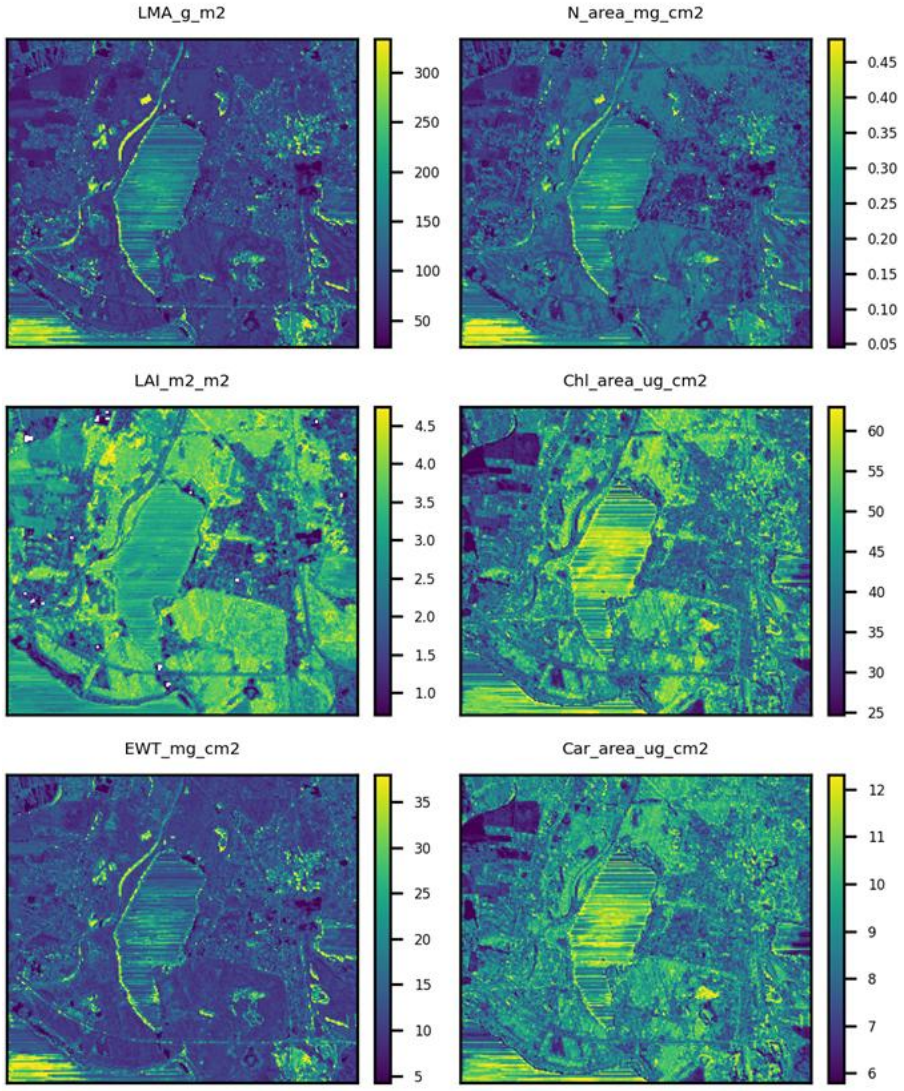


Figure S2b: The predicted trait values from the multi-trait for six traits (Leaf Mass per Area: LMA, Nitrogen content: N, Chlorophyll content: Chl, Equivalent Water Thickness: EWT, Leaf Area Index: LAI, and Carotenoid content: Car) when applied on the EnMAP scene.

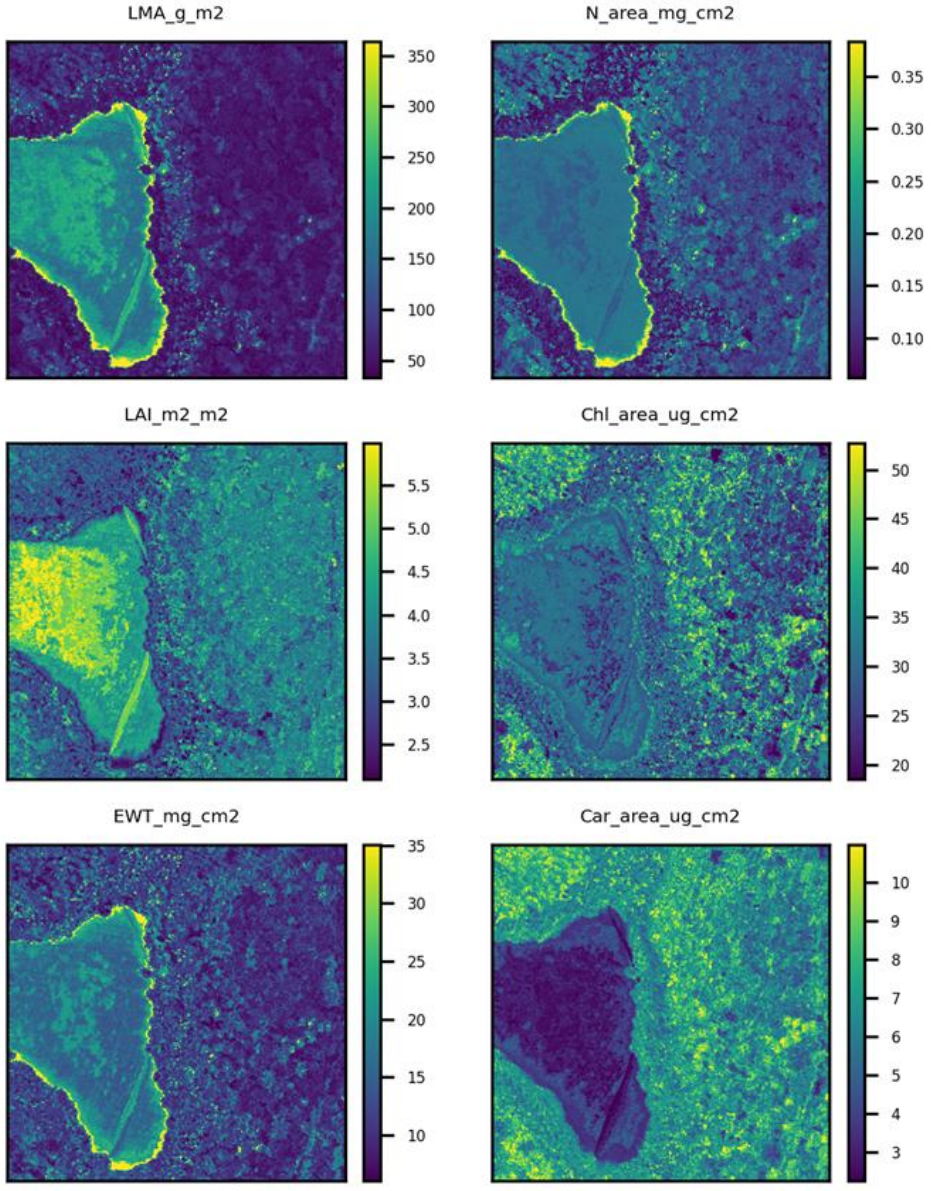


Figure S2c: The predicted trait values from the multi-trait for six traits (Leaf Mass per Area: LMA, Nitrogen content: N, Chlorophyll content: Chl, Equivalent Water Thickness: EWT, Leaf Area Index: LAI, and Carotenoid content: Car) when applied on the NEON scene.

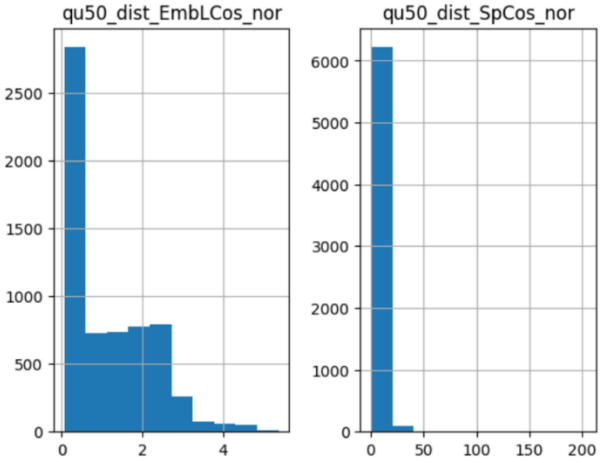


Figure S3: Distribution of the two predictors of uncertainty: normalized median of the cosine distance of 50 neighboring samples in feature space (qu50_distEmbLCos_nor) and the last embedding space (qu50_distSpCos_nor).

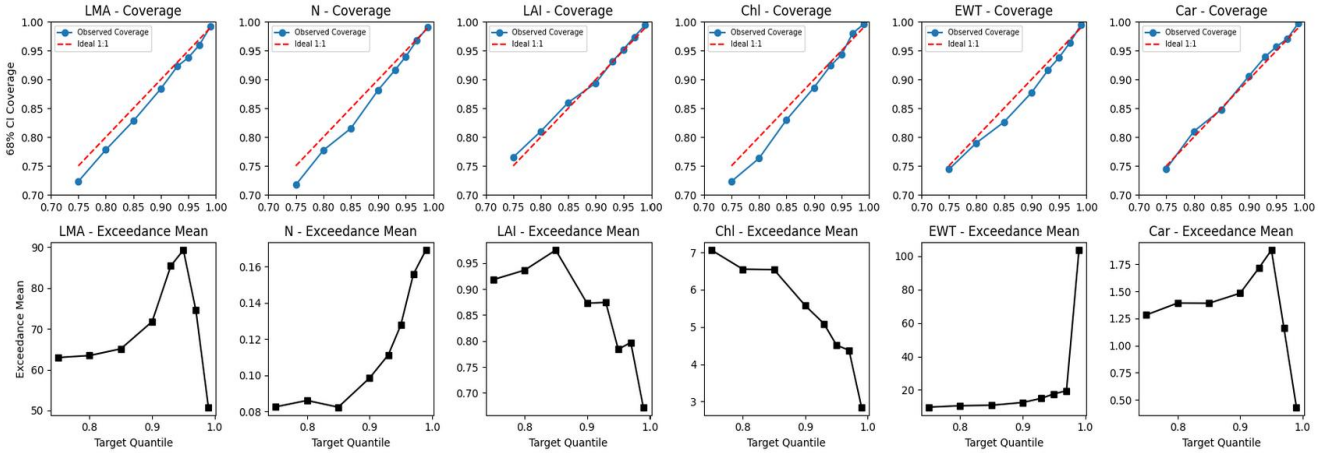


Figure S4: Validation of Dis_UN calibration and sensitivity analysis across quantile levels (τ = 0.75–0.99) for six plant traits: Leaf Mass per Area (g/m²) = LMA, Leaf Area Index (m²/m²) = LAI, Nitrogen content (mg/cm²) = N, Chlorophyll content (µg/cm²) = Chl, Equivalent Water Thickness (mg/cm²) = EWT, Carotenoid content (µg/cm²)= Car. Top panels: Empirical coverage (proportion of residuals falling within predicted bounds) with 68% Wilson confidence intervals, compared against target quantile (1:1 line). At τ=0.95, coverage reaches ≈94-96% across all traits, validating model calibration. Bottom panels: Exceedance mean (average magnitude of errors exceeding predicted bounds), reflecting worst-case error severity. Smooth behavior up to τ=0.95 transitions to abrupt instability at τ≥0.97, particularly for EWT and Car, indicating entry into unstable tail regime. Together, these results support τ=0.95 as a robust, data-driven choice that achieves high coverage while avoiding excessive conservatism and instability at extreme quantiles.

Table S4: The ratio coefficient (QuRatio) of all traits expressed in percentage as calculated with Eq. (6): predicted ranges over the observed residuals of four uncertainty models with OOD vegetation data: the distance-based method (Dis_UN), deterministic (Ens_det_UN), probabilistic deep ensemble (Ens_prob_UN) and monte carlo drop-out method (MCdrop_UN). The variance-based methods were calibrated by a scaling of 1.96 × standard deviation. List of traits and units: Leaf Mass per Area (g/m²) = LMA, Leaf Area Index (m²/m²) = LAI, Nitrogen content (mg/cm²) = N, Chlorophyll content (µg/cm²) = Chl, Equivalent Water Thickness (mg/cm²) = EWT, Carotenoid content (µg/cm²)= Car.

Traits	Dis_UN	Ens_det_UN	Ens_prob_UN	MCdrop_UN
LMA	107.446	56.480	87.791	17.439
N	86.692	52.009	119.086	13.063
LAI	22.554	54.078	46.743	16.986
Chl	65.518	43.170	84.555	15.185
EWT	77.913	60.066	103.558	17.840
Car	65.443	48.541	85.156	20.835
AVG	**70.928**	**52.724**	**87.482**	**16.891**

Table S5: The Expected Normalized Calibration Error (ENCE) values for the four compared uncertainty methods: the distance based method (Dis_UN), deterministic ensemble (Ens_det_UN), probabilistic ensemble (Ens_ prob_UN), and Monte Carlo dropout (MCdrop_UN). The variance-based methods were calibrated by a scaling of 1.96 × standard deviation. List of traits and units: Leaf Mass per Area (g/m²) = LMA, Leaf Area Index (m²/m²) = LAI, Nitrogen content (mg/cm²) = N, Chlorophyll content (µg/cm²) = Chl, Equivalent Water Thickness (mg/cm²) = EWT, Carotenoid content (µg/cm²)= Car.

Traits	Dis_UN	Ens_det_UN	Ens_prob_UN	MCdrop_UN
LMA	0.519	0.362	0.494	7.861
N	0.441	0.440	0.499	6.223
LAI	0.531	0.117	0.439	3.300
Chl	0.503	0.205	0.550	3.749
EWT	0.533	0.652	0.585	5.377
Car	0.500	0.227	0.555	3.073

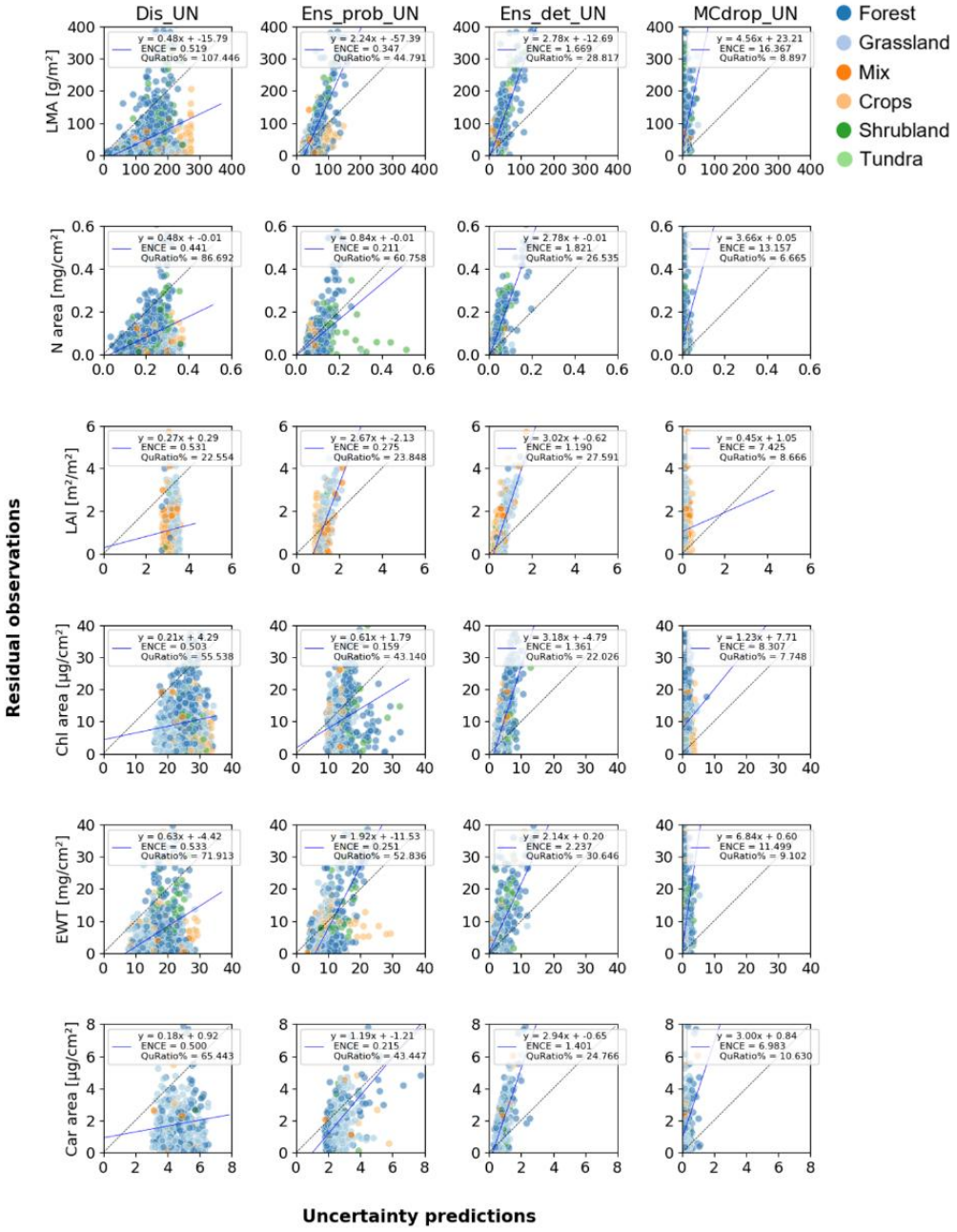


Figure S5: Scatter plots comparing the predicted uncertainties (x-axis) from four methods—distance-based (Dis_UN), probabilistic ensemble (Ens_prob_UN), deterministic ensemble (Ens_det_UN), and Monte Carlo dropout (MCdrop_UN)—against observed residuals (y-axis) of the multi-trait models across six traits: Leaf Mass per Area (LMA), Nitrogen content, Chlorophyll content (Chl), Equivalent Water Thickness (EWT), Leaf Area Index (LAI), and Carotenoid content (Car). Each point represents a sample, colored by vegetation type. For each trait–method combination, the regression line (blue) is compared to the 1:1 line (black) to visualize alignment between predicted and

observed errors. Model calibration is quantified by the Expected Normalized Calibration Error (ENCE), while the ratio of predicted to observed uncertainty ranges (QuRatio) indicates the coverage of residual variability. This shows that the unscaled version of the variance-based methods underestimates the uncertainty estimation.

Table S6: The regression coefficients of each scaled predictor for all the six traits: Leaf Mass per Area (LMA), Nitrogen content (N), Leaf Area Index (LAI), Chlorophyll content (Chl), Equivalent Water Thickness (EWT), and Carotenoid content (Car).

	LMA	**N**	**LAI**	**Chl**	**EWT**	**Car**
Intercept	139.1837	0.211816	3.204911	24.83733	17.99426	4.691486
Distance embedding space	7.071657	0.001075	-0.13865	1.821151	0.86472	0.104252
Distance feature space	53.74244	0.063145	-0.12674	3.461149	4.355728	0.812609

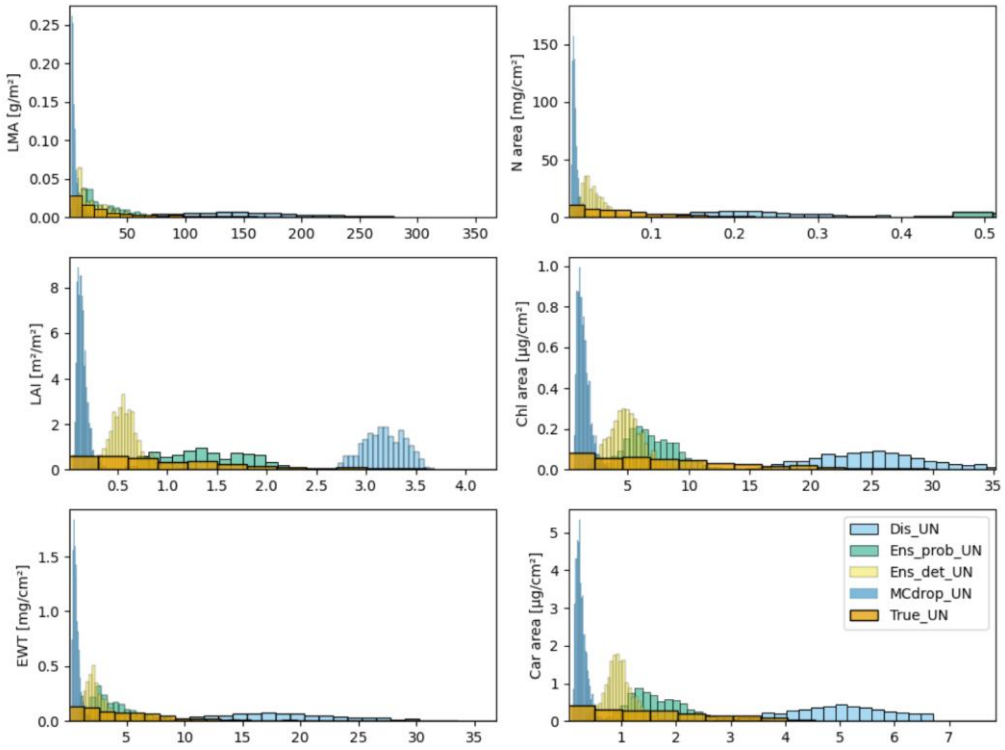


Figure S6a: Density distribution of the uncertainty prediction values of four uncertainty models with OOD vegetation data: the distance-based method (Dis_UN), deterministic deep ensemble (Ens_det_UN), probabilistic deep ensemble (Ens_prob_UN) and Monte Carlo dropout method (MCdrop_UN). List of traits and units: Leaf Mass per Area (g/m²) = LMA, Leaf Area Index (m²/m²) = LAI, Nitrogen content (mg/cm²) = N, Chlorophyll content (µg/cm²) = Chl, Equivalent Water Thickness (mg/cm²) = EWT, Carotenoid content (µg/cm²)= Car.

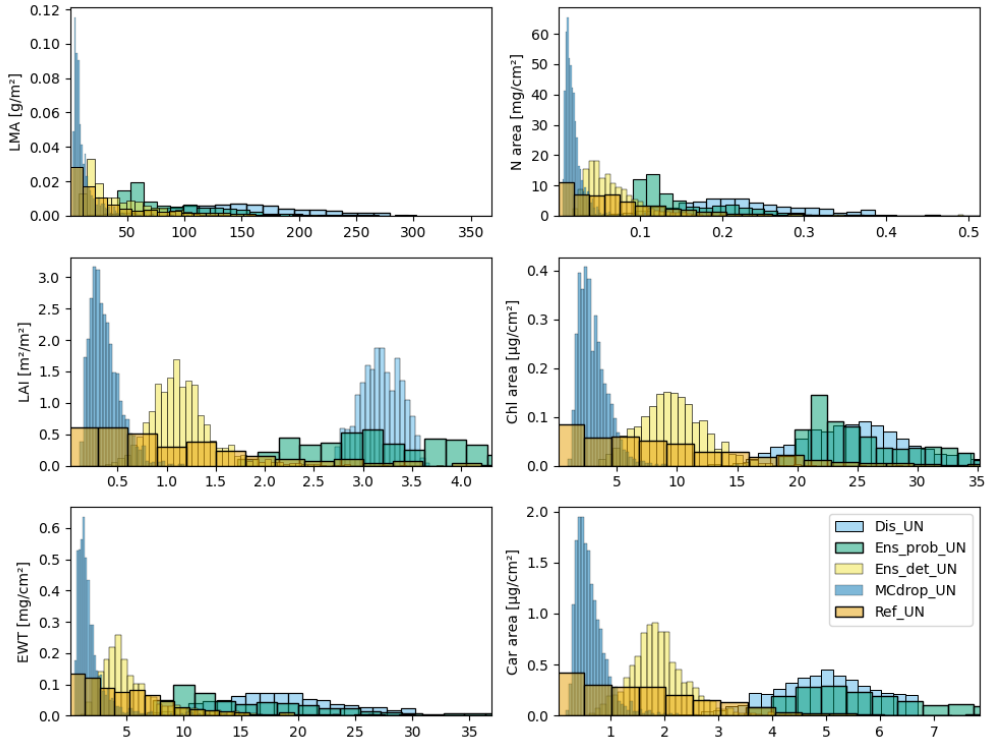


Figure S6b: Density distribution of the uncertainty prediction values of four uncertainty models with OOD vegetation data: the distance-based method (Dis_UN), deterministic deep ensemble (Ens_det_UN), probabilistic deep ensemble (Ens_prob_UN) and Monte Carlo dropout method (MCdrop_UN). The variance-based methods were calibrated by a scaling of 1.96 × standard deviation. List of traits and units: Leaf Mass per Area (g/m²) = LMA, Leaf Area Index (m²/m²) = LAI, Nitrogen content (mg/cm²) = N, Chlorophyll content (µg/cm²) = Chl, Equivalent Water Thickness (mg/cm²) = EWT, Carotenoid content (µg/cm²)= Car.

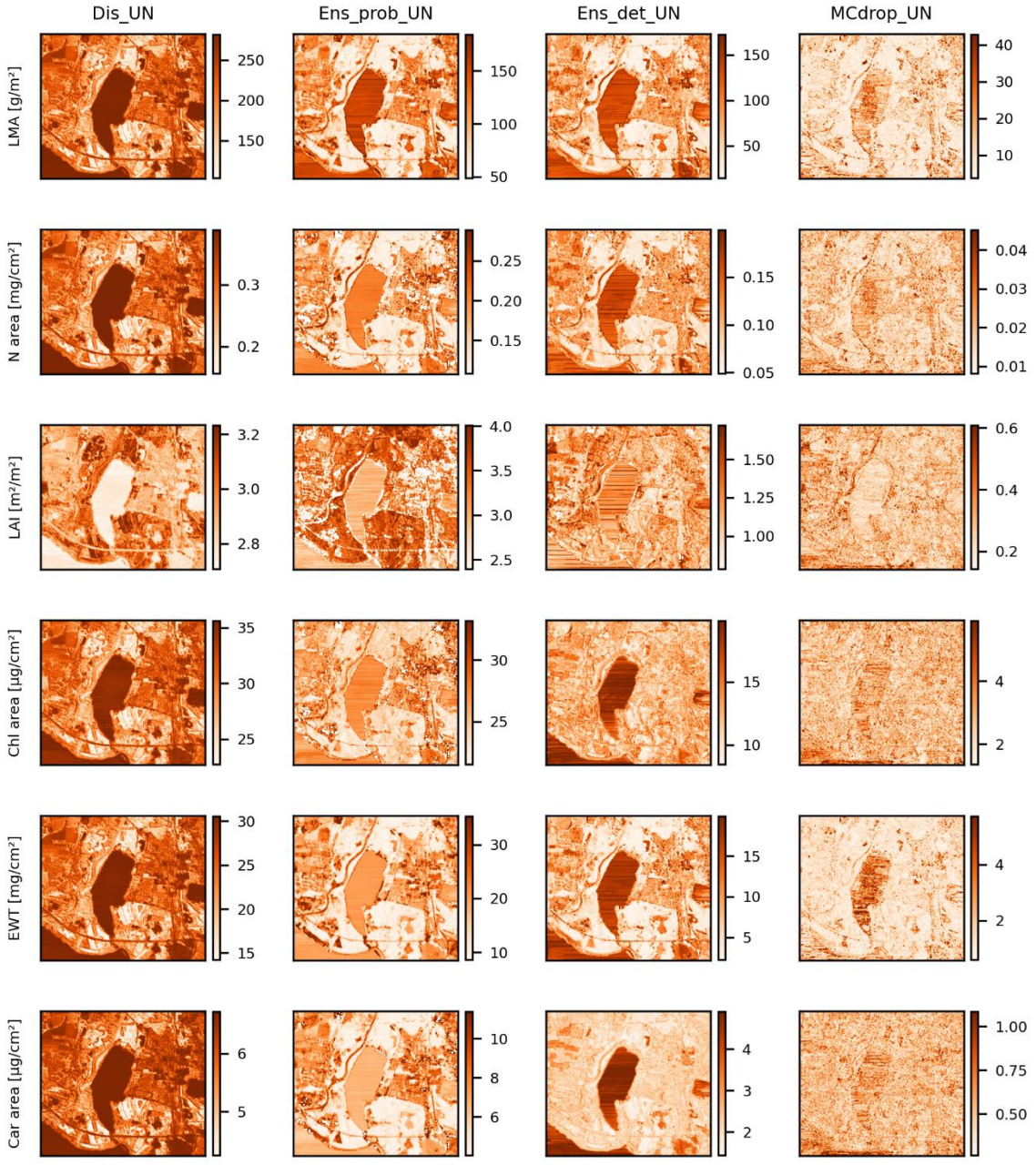


Figure S7a: EnMAP scene: Comparison of uncertainty estimations for six traits Leaf Mass per Area (LMA), Nitrogen content, Chlorophyll content (Chl), Equivalent Water Thickness (EWT), Leaf Area Index (LAI), and Carotenoid content (Car) using four methods: Distance-based (Dis_UN), probabilistic ensemble (Ens_prob_UN), deterministic ensemble (Ens_det _UN) and Monte Carlo dropout (MCdrop_UN). The variance-based methods were calibrated by a scaling of 1.96 × standard deviation. Spatial maps show the uncertainty distribution for each method.

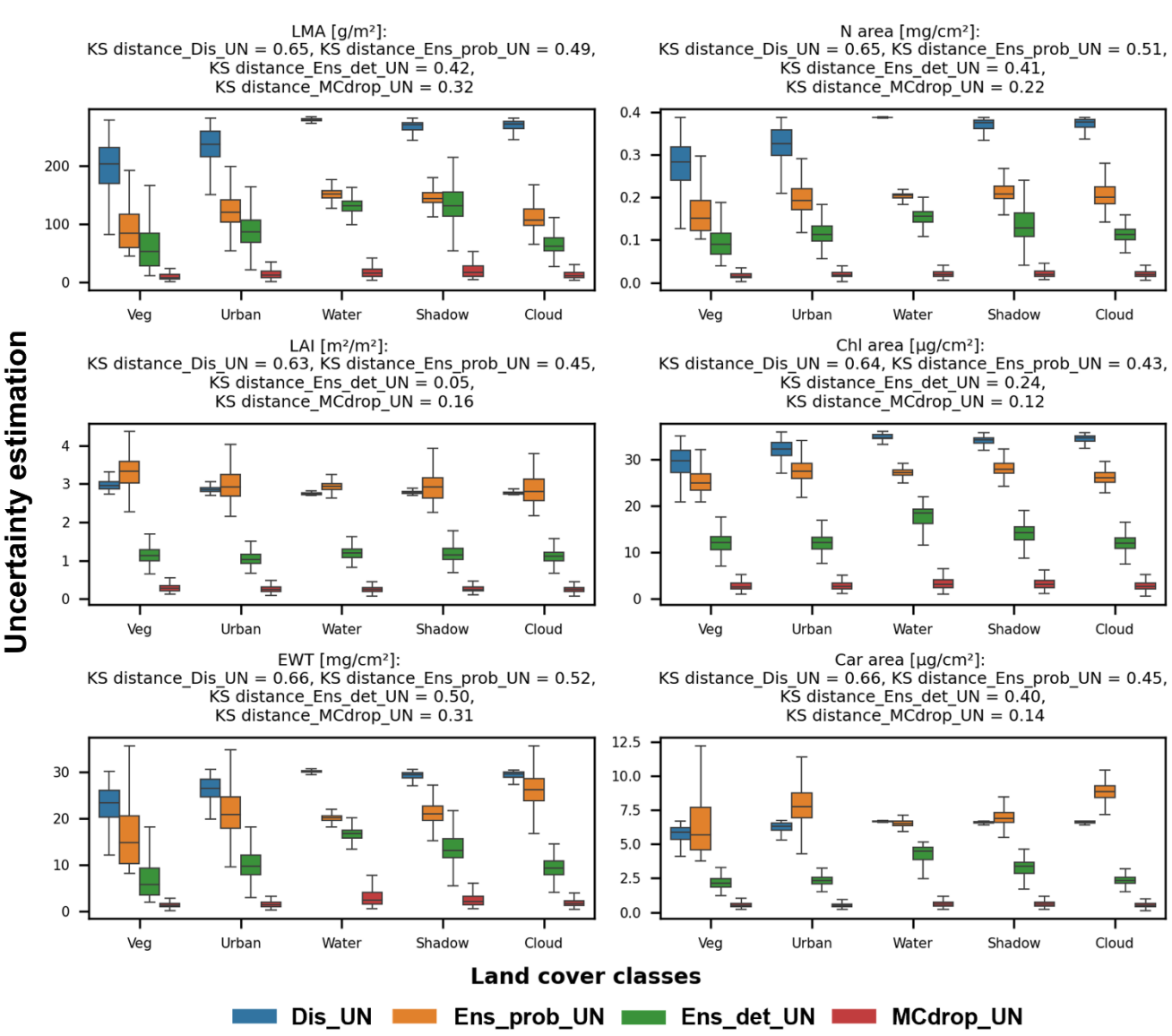


Figure S7b: ***EnMAP scene:*** Comparison of uncertainty estimations for six traits Leaf mass per area Leaf Mass per Area (LMA), Nitrogen content, Chlorophyll content (Chl), Equivalent Water Thickness (EWT), Leaf Area Index (LAI), and Carotenoid content (Car) using four methods: Distance-based (Dis_UN), probabilistic ensemble (Ens_prob_UN), deterministic ensemble (Ens_det_UN) and Monte Carlo dropout (MCdrop_UN). The variance-based methods were calibrated by a scaling of 1.96 × standard deviation. Box plots show uncertainty distributions across different scene components (Vegetation, Urban, Water, Shadow and Cloud) for each method. The Kolmogorov–Smirnov (KS) distance values quantify the separability of vegetated pixels with non-vegetated classes.

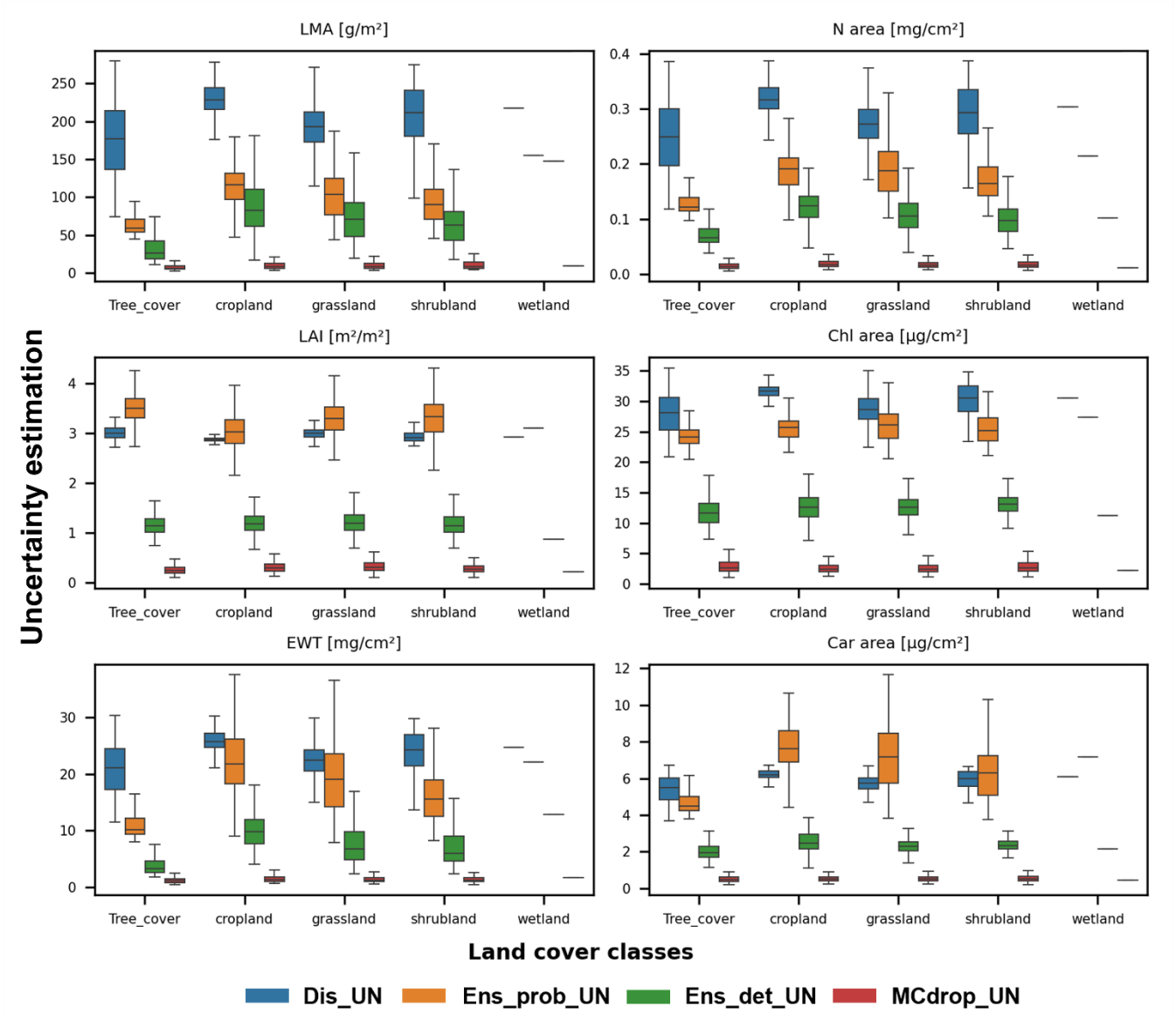


Figure S7c: ***EnMAP scene:*** Comparison of uncertainty estimations for six traits Leaf mass per area (LMA), Chlorophyll (Chl), Carotenoids (Car), Nitrogen (N) content, Leaf area index (LAI) and Equivalent water thickness (EWT) using four methods: Distance-based (Dis_UN), probabilistic ensemble (Ens_prob_UN), deterministic ensemble (Ens_det_UN) and Monte Carlo dropout (MCdrop_UN). The variance-based methods were calibrated by a scaling of 1.96 × standard deviation. Box plots show uncertainty distributions across different Vegetation types for each method.

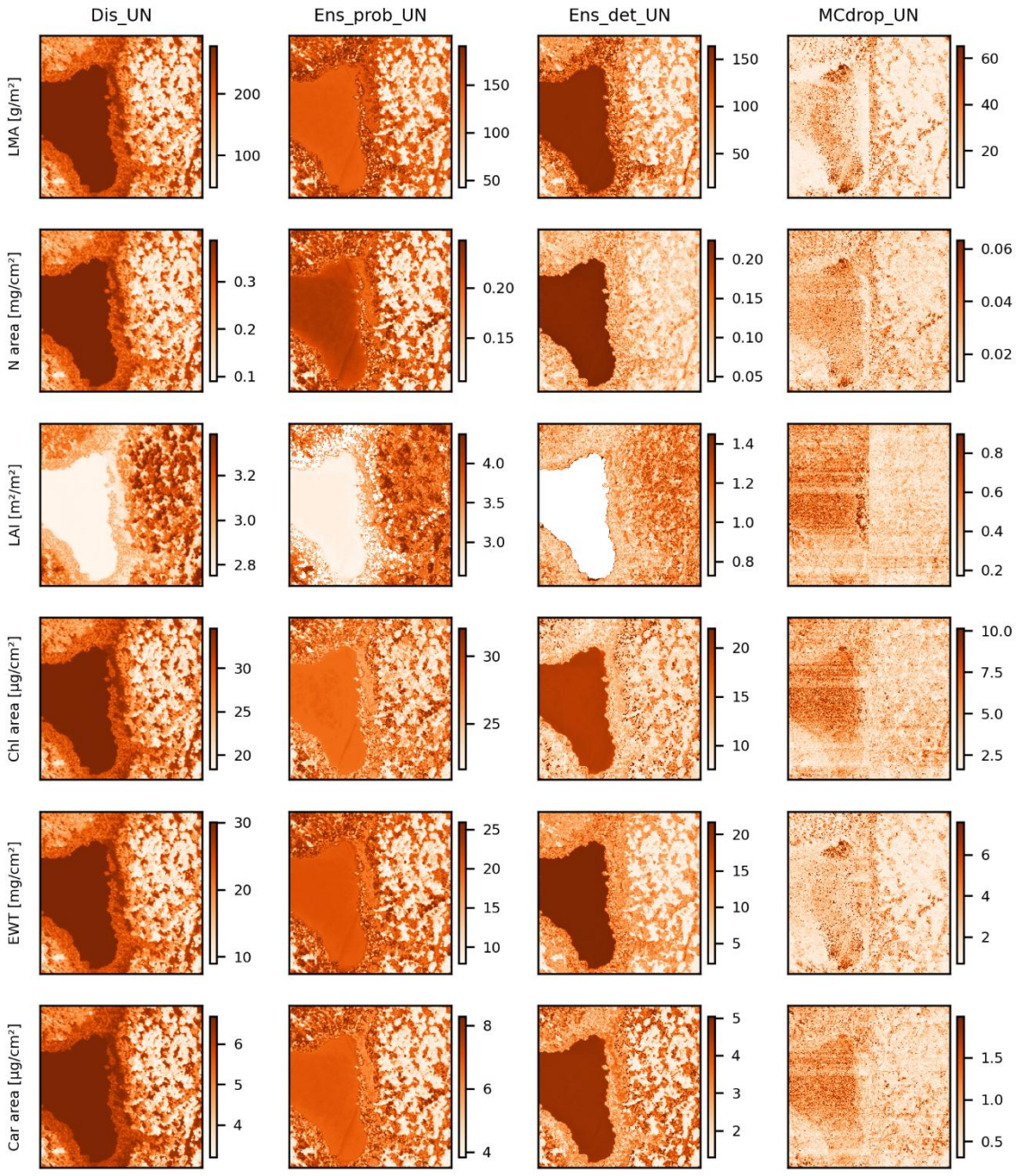


Figure S8a: ***Neon scene:*** Comparison of uncertainty estimations for six traits Leaf mass per area (LMA), Chlorophyll (Chl), Carotenoids (Car), Nitrogen (N) content, Leaf area index (LAI) and Equivalent water thickness (EWT) using four methods: Distance-based (Dis_UN), probabilistic ensemble (Ens_prob_UN), deterministic ensemble (Ens_det_UN) and Monte Carlo dropout (MCdrop_UN). The variance-based methods were calibrated by a scaling of 1.96 × standard deviation. Spatial maps show the uncertainty distribution for each method.

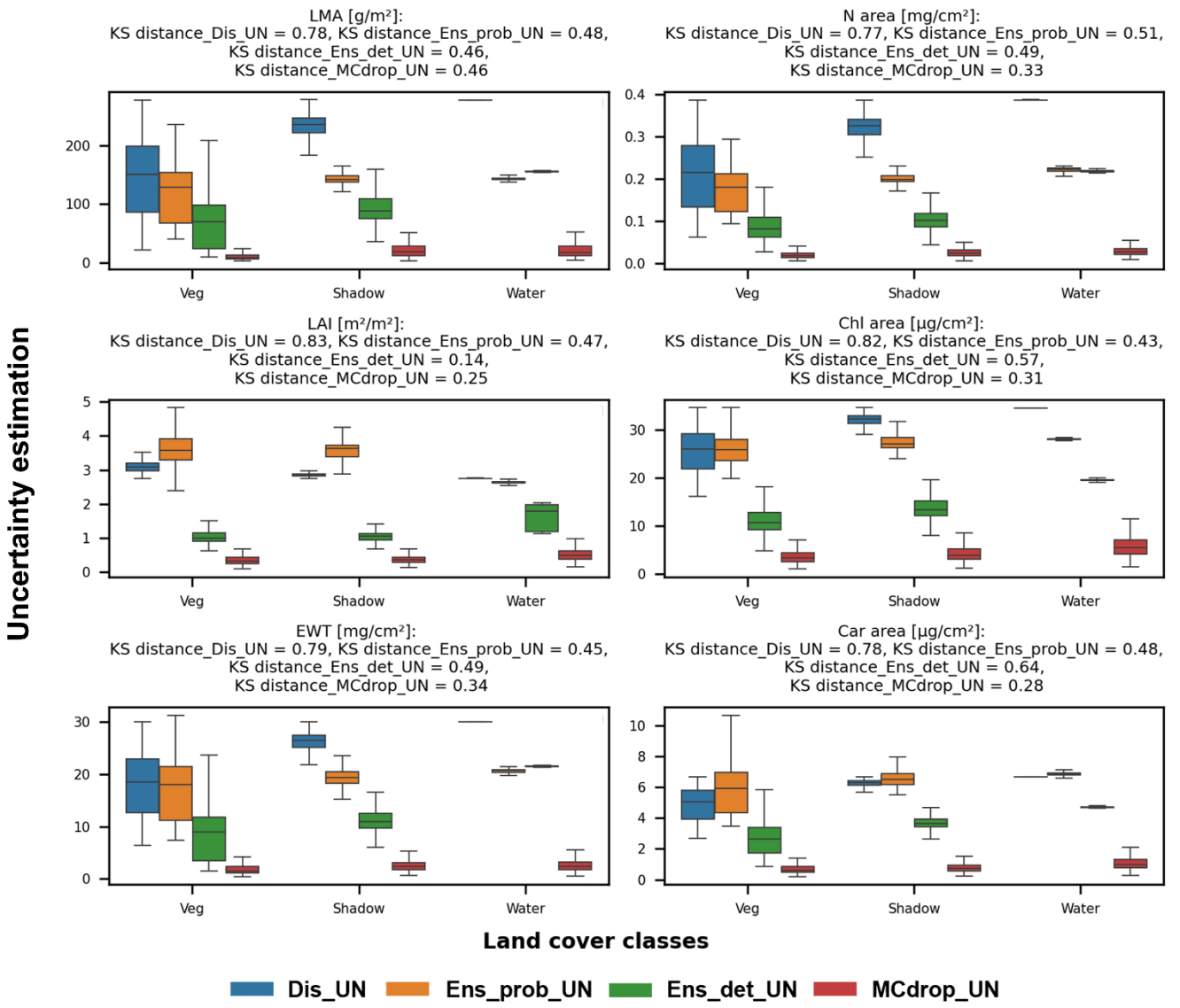


Figure S8b: ***Neon scene:*** Comparison of uncertainty estimations for six traits Leaf mass per area (LMA), Chlorophyll (Chl), Carotenoids (Car), Nitrogen (N) content, Leaf area index (LAI) and Equivalent water thickness (EWT) using four methods: Distance-based (Dis_UN), probabilistic ensemble (Ens_prob _UN), deterministic ensemble (Ens_det _UN), and Monte Carlo dropout (MCdrop_UN). The variance-based methods were calibrated by a scaling of 1.96 × standard deviation. Box plots show uncertainty distributions across different scene components (Vegetation, Shadow, Water) for each method. The Kolmogorov–Smirnov (KS) distance values quantify the separability of vegetated pixels with non-vegetated classes.

Table S7: Running time of the benchmarked methods for uncertainty estimation during the training phase.

Training	Transferability training	Distance calculation	Training
Dis_UN	~6h*50	0.34 h	~1s
Ens_prob_UN	~10h*50	-	-
Ens_prob_UN	~6h*50	-	-
MCdrop_UN	~6h	-	-

Table S8: Running time in seconds of the benchmarked methods for uncertainty estimation during the inference phase.

Inference	Ens_prob_UN		Ens_det_UN		MCdrop_UN		Dis_UN	
	Enmap	NEON	Enmap	NEON	Enmap	NEON	Enmap	NEON
Data preparation	6.64	9.62	6.64	9.62	6.64	9.62	190.9516	203.877
Model application	486.8539	556.7541	483.3784	555.3588	1567.305	1774.845	0.0889	0.0746
Total	493.4939	566.3741	490.0184	564.9788	1573.945	1784.465	**191.0405**	**203.9516**

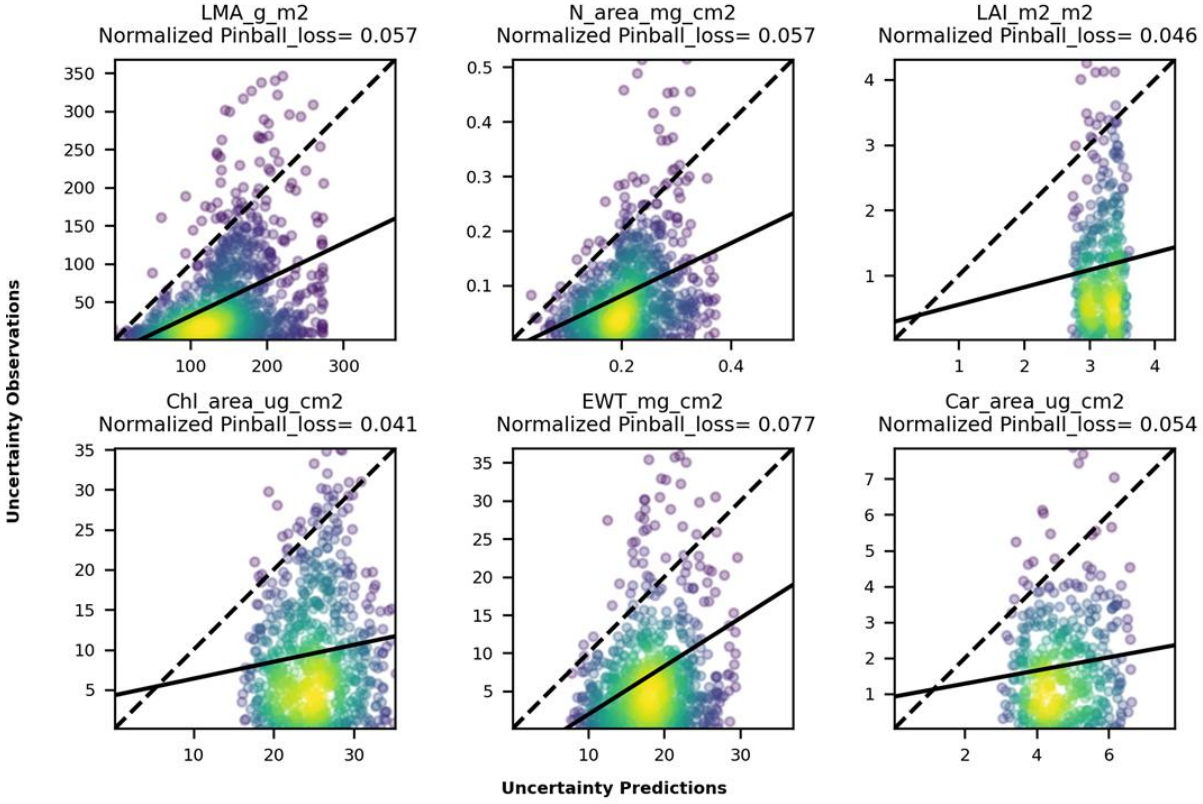


Figure S9: Correlation between observed and predicted uncertainty values of six traits: Leaf Mass per Area (LMA), Nitrogen content (N), Leaf Area Index (LAI), Chlorophyll content (Chl), Equivalent Water Thickness (EWT), and Carotenoid content (Car) from the Dis_UN model.

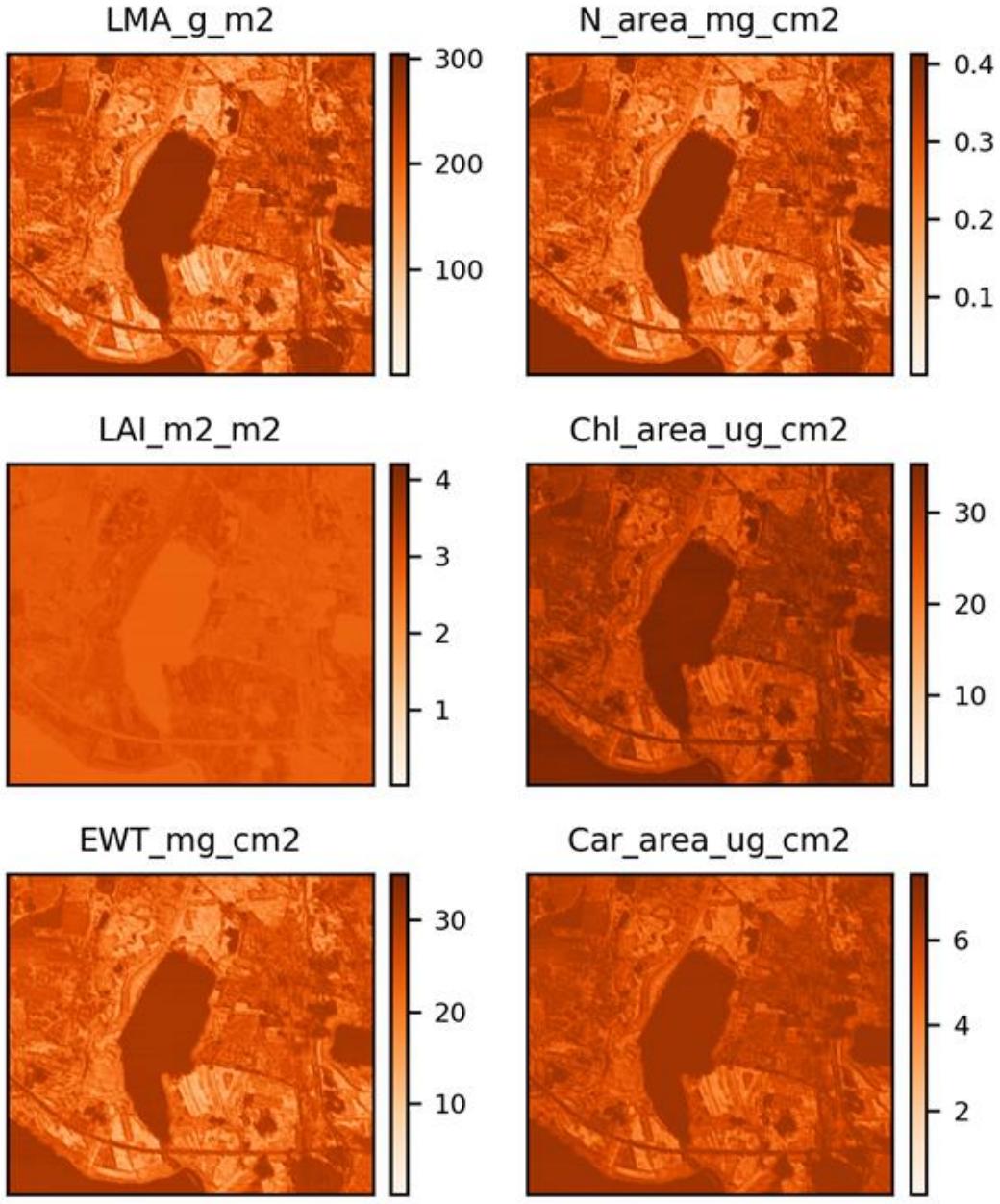


Figure S10: The predicted uncertainty values from the Dis_UN method for six traits (Leaf Mass per Area (LMA), Nitrogen content, Chlorophyll content (Chl), Equivalent Water Thickness (EWT), Leaf Area Index (LAI), and Carotenoid content (Car)) when applied on the EnMAP scene. This is a supplement to Fig. 4 where the color bar ranges use the range of training data as basis.

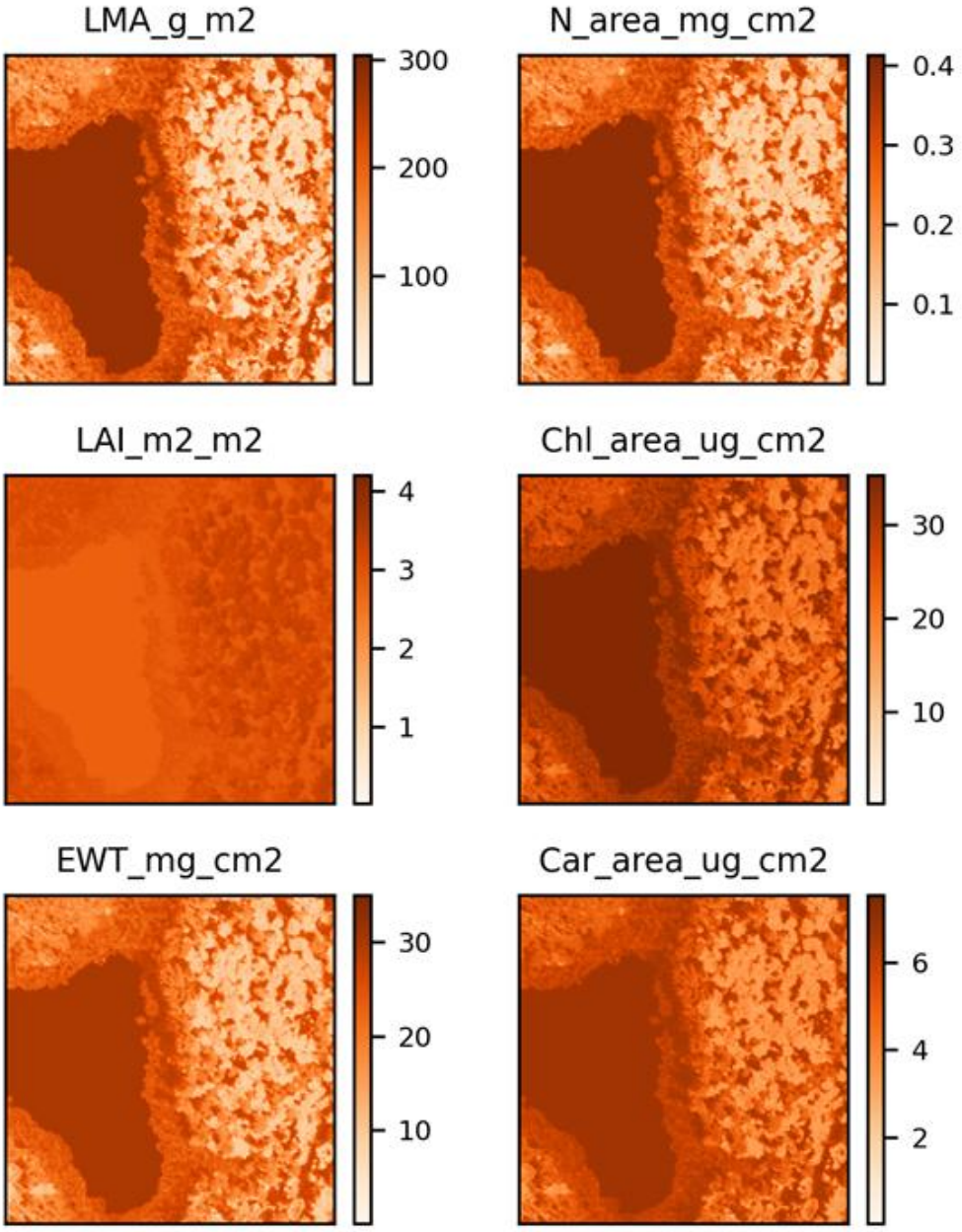


Figure S11: The predicted uncertainty values from the Dis_UN method for six traits (Leaf Mass per Area (LMA), Nitrogen content, Chlorophyll content (Chl), Equivalent Water Thickness (EWT), Leaf Area Index (LAI), and Carotenoid content (Car)) when applied on the NEON scene. This is a supplement to Fig. 5 where the color bar ranges use the range of training data as basis.

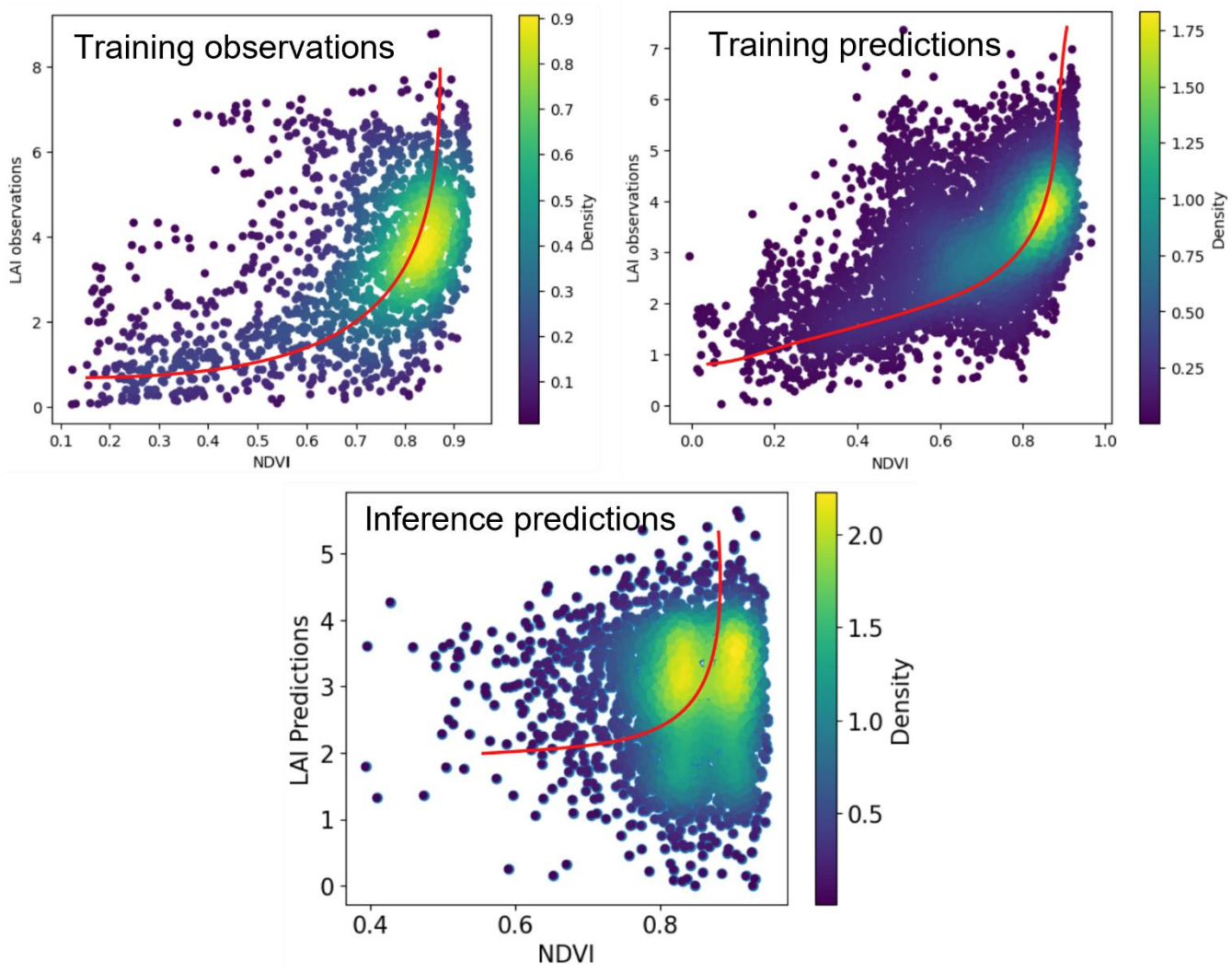


Figure S12: Scatter plots of LAI observations against NDVI within training observations (top left), within training predictions (top right), and within inference predictions (bottom). Color shading represents point density, and red lines indicate fitted trends. The training data reveal LAI saturation with NDVI, which propagates into the model predictions, leading to systematic underestimation of high-LAI values.

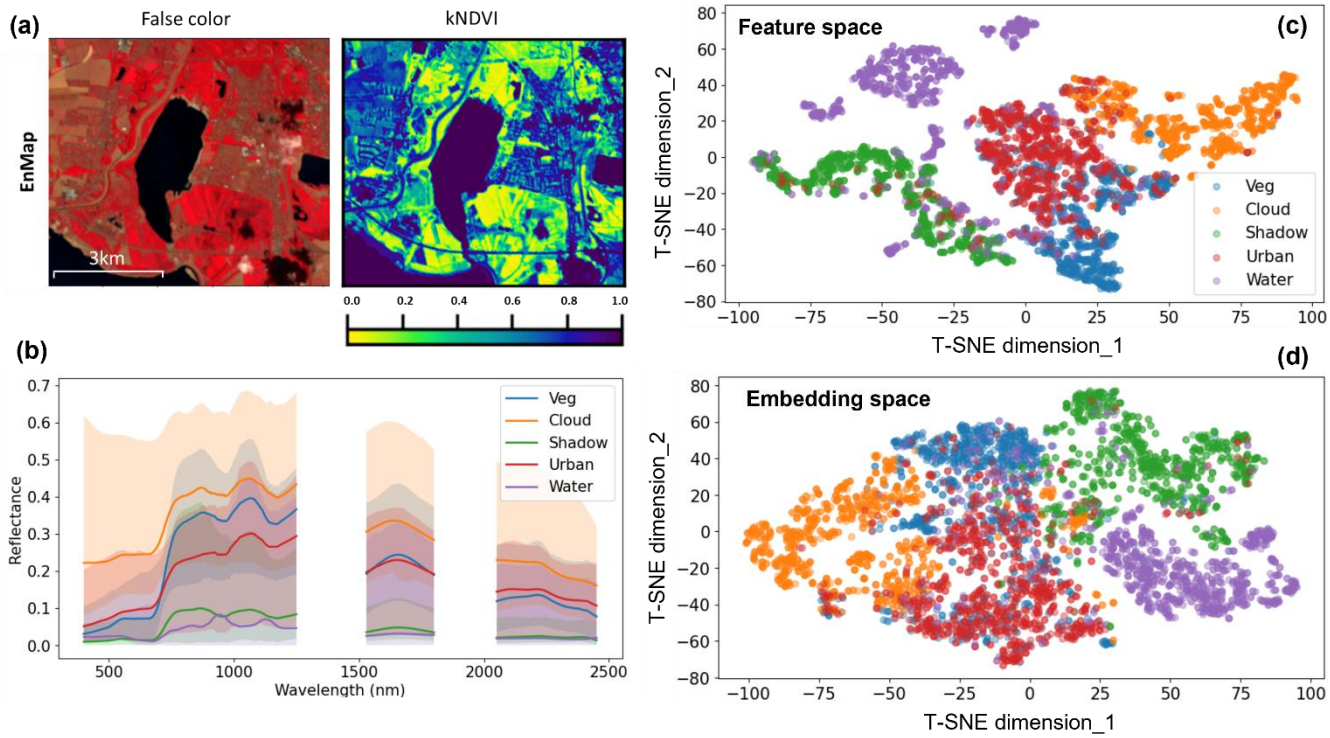


Figure S13a: Visualization of 1000 sampled pixels from each scene component from the EnMAP scene: (a) Characteristics of the scene (b) Distribution of canopy reflectance (c) T-SNE visualization of the spectra in the feature space (d) T-SNE visualization of the spectra in the embedding space of the multi-trait model.

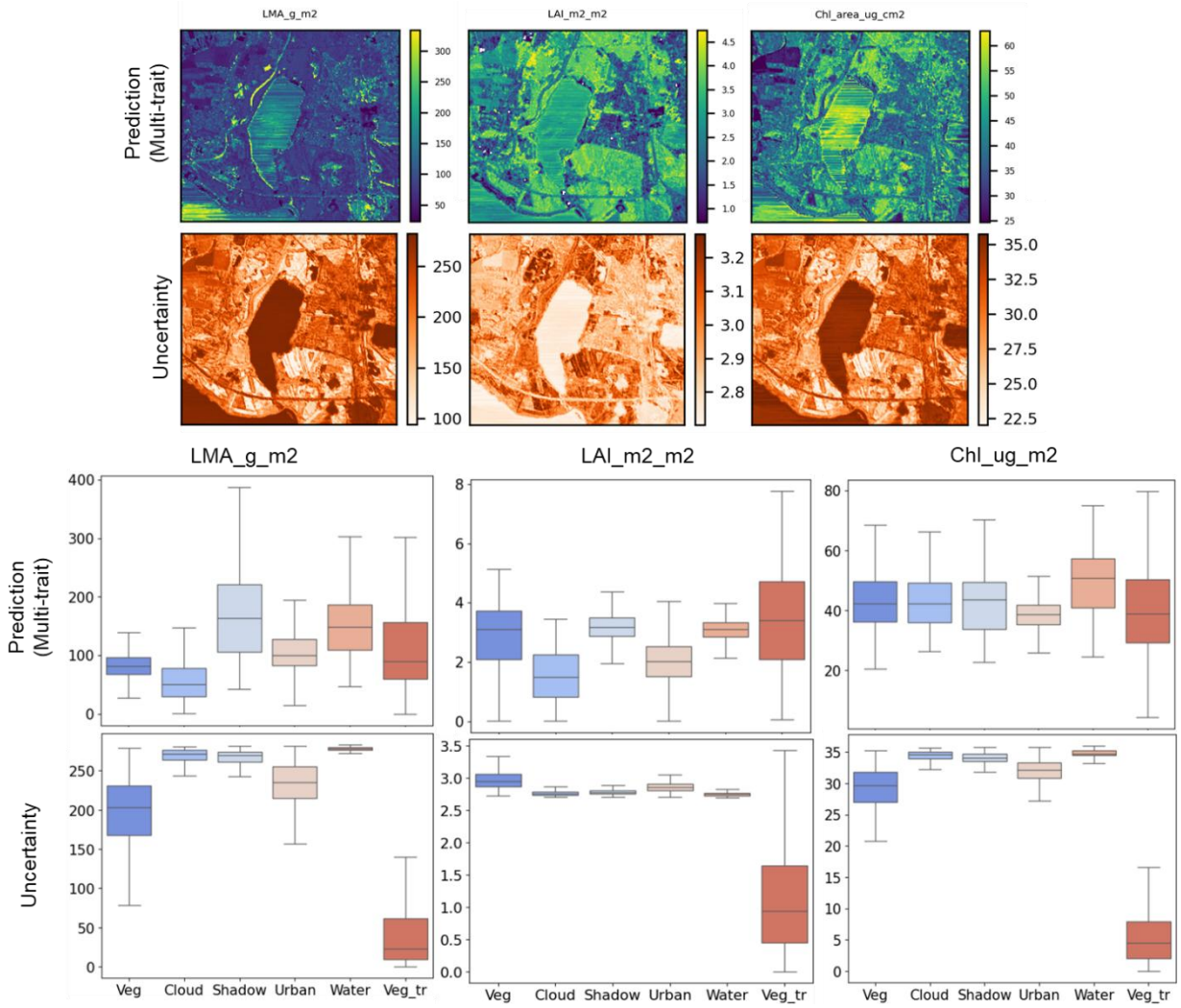


Figure S13b: Comparison of distribution in the trait predicted values from the multi-trait model and the predicted uncertainty values from the Dis_UN when applied on sampled pixels shown in Fig. 13a.

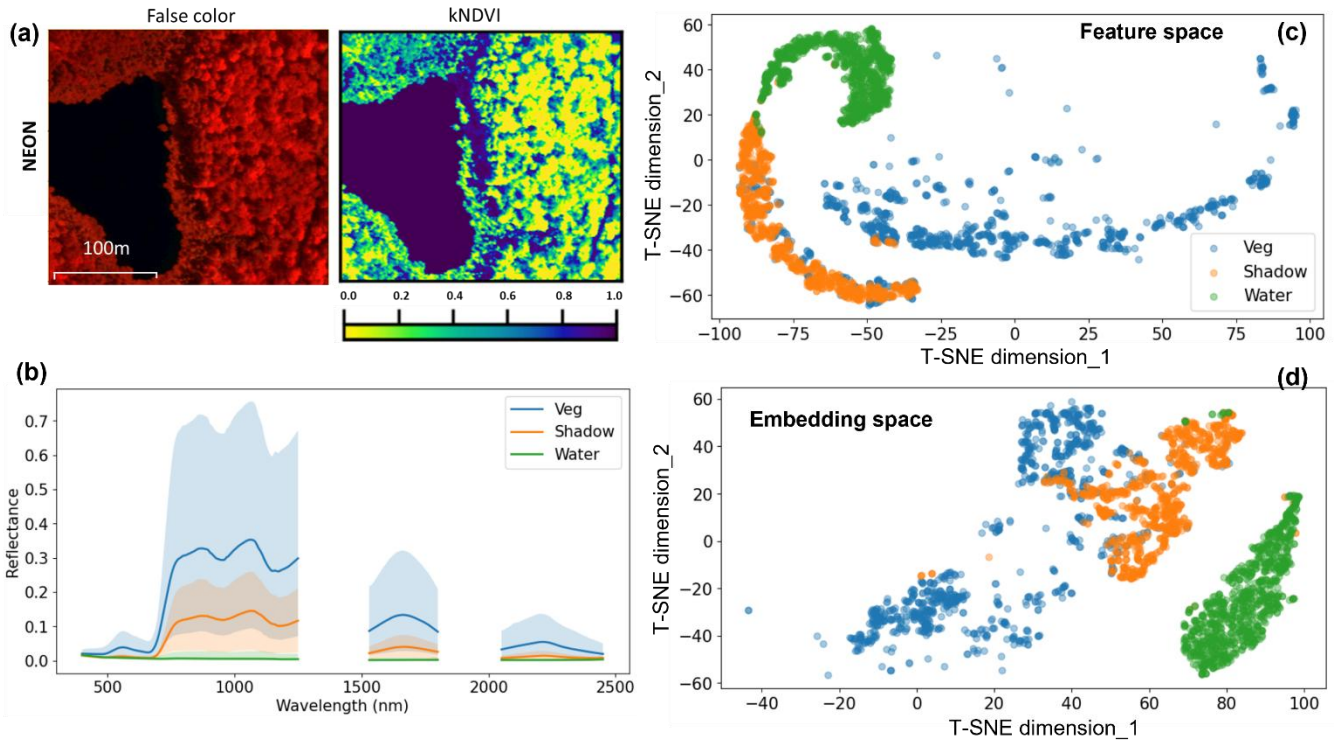
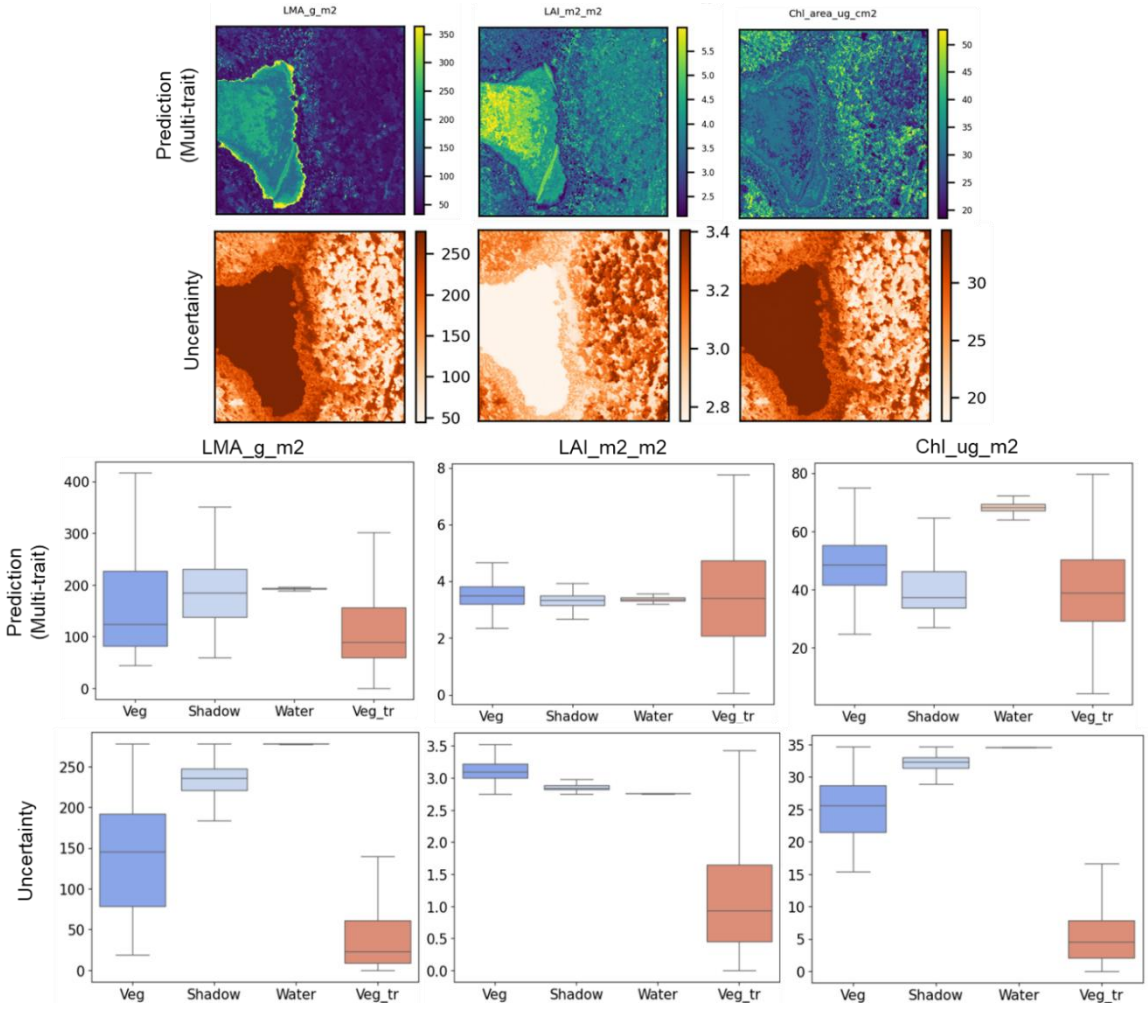


Figure S14a: Visualization of 1000 sampled pixels from each scene component from the NEON scene: (a) Characteristics of the scene (b) Distribution of canopy reflectance (c) T-SNE visualization of the spectra in the feature space (d) T-SNE visualization of the spectra in the embedding space of the Multi-trait model.



340

Figure S14b: Comparison of distribution in the trait predicted values from the multi-trait model and the predicted uncertainty values from the Dis_UN when applied on sampled pixels shown in Fig. 14a.