

Hyperspectral Satellite Imagery for Digital Twinning of Boreal Forests

Matti Möttöus, Laura Sirro, Renne Tergujeff

VTT Technical Research Centre of Finland Ltd, P.O. Box 1000, FI-02044 VTT, Finland

matti.mottus@vtt.fi, laura.sirro@vtt.fi, renne.tergujeff@vtt.fi

Keywords: Imaging Spectroscopy, EnMAP, Sentinel-2, Forest Inventory, Convolutional Neural Network, Digital Twin

Abstract

Digital twins of the Earth require Earth observation data streams that are both spatially continuous and informative about the state of the modelled system. For forests, the standard optical data source is Sentinel-2, whose 10 m multispectral imagery trades spectral detail for spatial resolution; the forthcoming operational imaging spectroscopy missions CHIME and SBG will offer the opposite trade-off with 30-m pixels and hundreds of spectral bands. We quantified the cost of that trade-off using a near-simultaneous EnMAP and Sentinel-2 image pair acquired over the boreal forests in Hyytiälä, central Finland, on 20 August 2023, together with 213 field plots measured by the Finnish Forest Centre. A U-Net convolutional neural network was trained fully convolutionally to predict stand basal area, basal-area weighted mean diameter and mean height. Training had two stages: self-supervised pretraining (spectral reconstruction for EnMAP, masked band prediction for Sentinel-2) followed by supervised fine-tuning on the field plots. On 32 independent test plots, EnMAP resampled to 10 m matched Sentinel-2: R^2 was 0.65 vs. 0.67 for basal area, 0.80 vs. 0.80 for mean diameter and 0.79 vs. 0.79 for mean height, while mean absolute errors were 9–13 % lower for EnMAP. Predictions obtained from EnMAP at its native 30 m sampling were equally accurate. The nine-fold loss in ground sampling area that accompanies the higher spectral resolution therefore did not degrade the retrieval of forest structural variables, which supports the use of spaceborne imaging spectroscopy as a baseline information source in a forest digital twin.

1. INTRODUCTION

1.1 Digital twins of forests

Destination Earth (DestinE), a flagship European Union initiative, is building a digital twin of the Earth, “an information system that exposes users to a digital replication of the state and temporal evolution of the Earth system constrained by available observations and the laws of physics” (Bauer et al., 2021). The twin is kept synchronised with its real-world counterpart through a continuous flow of sensor data which – on the global scale – can uniformly and continuously only be provided by satellite Earth observation (EO). Within DestinE, in the context of the Digital Twin Earth programme of the European Space Agency ESA, VTT leads the consortium building the digital twin of the Earth's forests¹.

As the name implies, a forest digital twin is primarily a modelling system (Möttus et al., 2021): process-based growth and carbon balance models such as PREBAS (Minunno et al., 2016, 2019) propagate the state of the forest forward in time and respond to climate forcing, disturbances and management. Such a system is, however, inconceivable without EO data that connects it to the status (coverage, structure, species composition) and processes (e.g., photosynthetic activity, pest attacks) in real forests. Earth observation enters the twin in at least three roles: it provides the initial state of the modelled stands, it supplies spatially continuous calibration data for productivity models, and it allows the model trajectory to be corrected through data assimilation. In a digital twin, all three roles require repeated, wall-to-wall observations rather than isolated acquisitions.

1.2 Optical Earth observation of forest structure

The operational optical data source for forest monitoring in Europe is the Sentinel-2 satellite constellation of the Copernicus programme with its multispectral instrumentation (Drusch et al., 2012), which combines 10–20 m ground sampling with a five-

day revisit. Empirical models relating Sentinel-2 reflectance to forest inventory variables have been developed for the boreal zone with both classical estimators (Astola et al., 2019) and deep neural networks (Astola et al., 2021), and global canopy height products are now derived from the same data with convolutional networks (Lang et al., 2023). The accuracy of these retrievals is limited less by spatial detail than by the small number and broad width of the Sentinel-2 bands, which constrain how much of the canopy biochemical and structural signal can be separated.

1.3 Spaceborne imaging spectroscopy

Spaceborne imaging spectrometers sample the reflected solar spectrum in hundreds of contiguous narrow bands and therefore lift precisely the aforementioned separation constraint. EnMAP, launched in 2022, is the first satellite mission of this class to deliver systematically tasked, radiometrically stable data over Europe (Guanter et al., 2015; Chabrillat et al., 2024). It will be followed by the operational Copernicus Hyperspectral Imaging Mission for the Environment (CHIME; Rast et al., 2021) and by NASA's Surface Biology and Geology mission (SBG; Cawse-Nicholson et al., 2021), which together will, for the first time, make time series of spaceborne imaging spectroscopy available for forest applications (Hill et al., 2019).

Published forest applications of EnMAP have so far concentrated on canopy biochemistry and on spectral diversity: leaf and canopy trait mapping with hybrid retrieval models (Farmonov et al., 2025), spectral-heterogeneity indicators of biodiversity (Torresani et al., 2025) and, for the same image pair used here, vegetation-index-based estimation of leaf area index and canopy closure (Juola et al., 2024). Retrieval of the structural variables that forest inventory and forest growth models actually use – basal area, stem diameter and tree height – has not been assessed for spaceborne imaging spectroscopy in the boreal zone.

The reason is a trade-off. EnMAP samples the ground at 30 m against the 10 m resolution of the Sentinel-2 visible and near-

¹ <https://foresttwin.org/>

infrared bands, so a ninefold larger ground area is integrated into every spectrum. In a managed boreal landscape, where the mean stand size is of the order of one hectare, this is a non-negligible penalty: a 30 m pixel is comparable in size to a whole stand, and mixed pixels along stand boundaries are correspondingly more frequent.

1.4 Objectives

This study asks whether the spatial penalty of spaceborne imaging spectroscopy degrades the retrieval of forest structural variables relative to Sentinel-2. We use a near-simultaneous EnMAP and Sentinel-2 image pair, a single field reference data set, and one convolutional neural network framework applied to both sensors, so that the comparison isolates the effect of the sensor rather than of the retrieval method. We further examine how the size of the spatial context given to the network, and the spatial sampling at which the EnMAP data are presented to it, affect the result.

2. MATERIALS

2.1 Study site

The study area covers the surroundings of the Hyytiälä forestry field station and the SMEAR II measurement station in Juupajoki, central Finland (61°50'N, 24°17'E; Hari and Kulmala, 2005; Figure 1). The area belongs to the southern boreal vegetation zone and is dominated by managed, even-aged stands of Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) H. Karst.), with birch (*Betula* spp.) as the main broadleaved species, interspersed with drained and undrained peatlands, lakes and agricultural land. Forests are mostly managed with densities usually between 500 and 1500 ha⁻¹. The area is covered with low rolling hills with an elevation below 200 m above sea level.

2.2 Satellite data

Sentinel-2B and EnMAP (Figure 2) imaged the study area near-simultaneously on 20 August 2023, at 12:53 and 13:06 local time, respectively (Table 1). Both images are essentially cloud-free over the extent of the smaller EnMAP scene, and the same image pair has previously been used for mapping leaf area index and canopy closure (Juola et al., 2024). Both data sets were obtained as bottom-of-atmosphere reflectance: the Sentinel-2 Level-2A product and the EnMAP Level-2A land product. The ten Sentinel-2 bands at 10 and 20 m (B2–B8A, B11, B12) were resampled to a common 10 m grid; the 219 EnMAP bands remaining after removal of the bands affected by strong water vapour absorption were used as delivered.

Because the two sensors differ in ground sampling distance, the EnMAP data cube was additionally resampled by a factor of three to the 10 m Sentinel-2 grid using bilinear interpolation. This does not add information, but it lets both sensors be processed by an identical patch-based network and it gives a denser sampling of the spectral signal within a field plot. Models trained on the EnMAP data at their native 30 m sampling were run as a control (Section 4.3).



Figure 1. Location of the Hyytiälä SMEAR station in the boreal forest zone (in light green).

2.3 Field reference data

Field reference data came from the forest inventory plot data of the Finnish Forest Centre (FFC, Metsäkeskus; Finnish Forest Centre, 2026). Two FFC campaigns intersected the image pair: the Tampere data set measured in 2023 and the Jämsä data set measured in 2024. The field data include both circular plots with a radius of approximately 4.0–2.5 m and tree map plots that are stored as variable shape and size polygons. The plots cluster in the southern and eastern parts of the EnMAP scene. After restricting the plots to the spatial overlap of the two images and discarding plots with fewer than ten valid pixels in either image, 213 plots remained.

For each plot, the FFC data give stand basal area G (m² ha⁻¹), basal-area-weighted mean diameter at breast height D (cm) and basal-area-weighted mean height H (m), both for the whole growing stock and separately for the three species groups (pine, spruce and broadleaved trees). All twelve variables were predicted simultaneously by a single network.

FFC sample plots have been available as open data since 2018. However, the data used in this study were obtained directly from the Finnish Forest Centre upon request as they were not publicly available at the time of acquisition.

	Sentinel-2B	EnMAP
Acquisition date	20 Aug 2023	20 Aug 2023
Local time	12:53	13:06
Product level	L2A (BOA)	L2A (BOA)
Bands used	10	219
Spectral range	490–2190 nm	420–2450 nm
Native GSD	10 / 20 m	30 m
Grid used here	10 m	10 m / 30 m

Table 1. Satellite data used in this study.

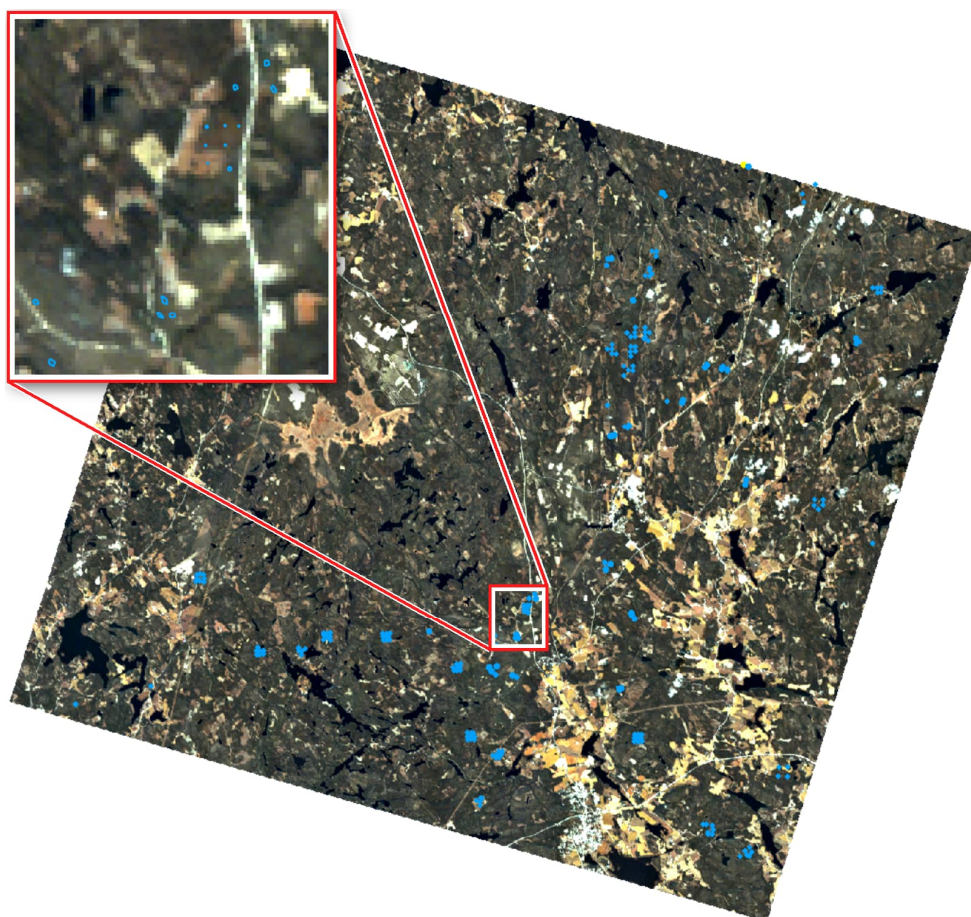


Figure 2. The 30×30 km EnMAP scene used in the study with forest plots' outlines in blue and a zoom-in showing the varying shape vectors of some of these.

3. METHODS

The complete processing chain is shown in Figure 3, with the names of the corresponding functions and classes of the analysis code given inside the boxes. The code is written in Python and uses PyTorch for the network, rasterio for raster access and GeoPandas for the plot polygons.

3.1 Preprocessing

Plot selection and splitting: Plots were read from the two FFC GeoPackages and retained only if they fell inside the spatial overlap of the two images and contained at least ten valid, non-masked pixels in both images. The resulting 213 plots were divided into training, validation and test sets in the proportions 149:32:32 (70:15:15 %). Because the sample is small, a random split can easily produce a test set whose distribution differs markedly from that of the training set; candidate splits were therefore generated with successive random seeds and the first split whose per-variable distributions passed a similarity check was accepted.

Target rasters: The plot polygons were rasterised onto the image grid. A pixel was assigned the plot values if at least 33 % of its area fell inside the polygon, the fraction being evaluated on a four-times refined grid. A separate target raster was built for each of the three splits. This is essential: with a single raster built from all plots, a training patch centred near a test plot would contain that plot's labelled pixels and the loss would be evaluated on test data during training. Keeping the rasters separate makes the three sets strictly independent.

Normalisation and patch extraction: Each band was standardised by its mean and standard deviation over the image. The final training samples were square patches of 48×48 pixels, i.e. $480 \text{ m} \times 480 \text{ m}$ on the ground, centred on labelled pixels; the choice of patch size is examined in Section 4.3. Pixels outside the Finnish Forest Centre forest-land mask were excluded.

3.2 Network architecture

The network is a U-Net (Ronneberger et al., 2015) with two downsampling levels, a bottleneck and a symmetric decoder with skip connections (Figure 4). The encoder, bottleneck and decoder are identical for the two sensors; only the input stem and the self-supervised task differ.

For EnMAP, the 219 input bands are first compressed to 32 feature channels by a 1×1 convolution. This is a learned, purely spectral projection: it exploits the strong redundancy between neighbouring narrow bands and reduces the number of parameters that the spatial convolutions have to carry. For Sentinel-2, whose ten bands are neither numerous nor redundant, the stem is instead a 3×3 convolution to 64 channels followed by a squeeze-and-excitation block (Hu et al., 2018), which learns a global weighting of the bands; squeeze-and-excitation blocks are used throughout the Sentinel-2 encoder and decoder.

The shared decoder output feeds two 1×1 convolutional heads: a reconstruction head used only during pretraining, and a regression head that outputs the twelve forest variables.

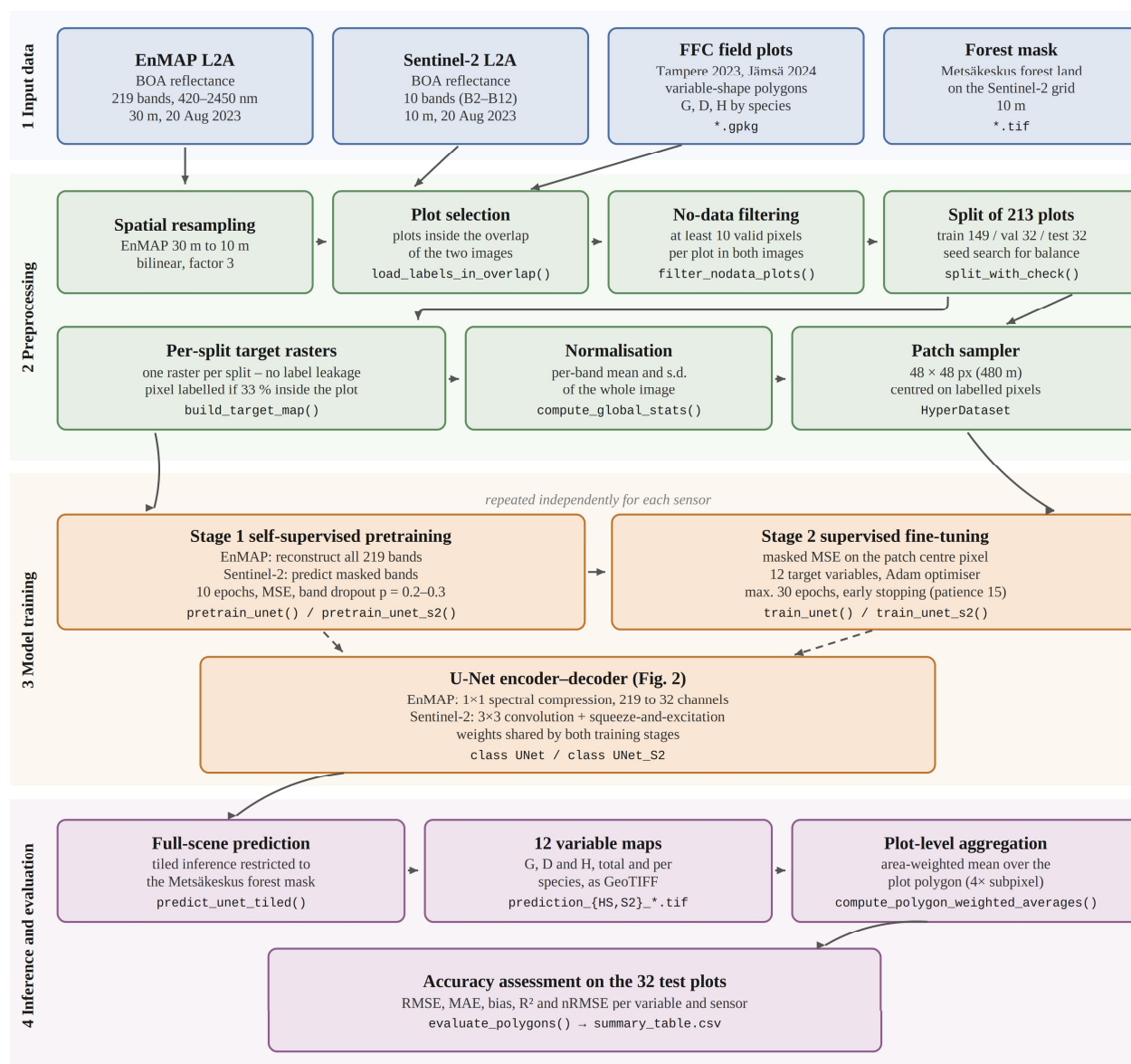


Figure 3. Processing chain. Boxes give the processing step, its key parameters and the name of the corresponding function or class in the analysis code. The chain is identical for the two sensors apart from the sensor-specific network stem and the self-supervised pretraining task.

3.3 Two-stage training

Training proceeded in two stages. In the first, self-supervised stage, the network was trained for ten epochs on unlabelled patches drawn from the whole image, with the reconstruction head attached. For EnMAP, the task was to reconstruct the full 219-band spectrum from a spectrally degraded input, and for Sentinel-2 to predict randomly masked bands from the remaining ones (He et al., 2022; Wang et al., 2022). Bands were dropped with probability 0.2–0.3. The purpose is the same in both cases: to force the encoder to learn the internal structure of the spectrum before any field data are used, which matters because only 149 labelled plots are available.

In the second stage, the reconstruction head was replaced by the regression head, and the whole network was fine-tuned on the labelled patches for at most 30 epochs with the Adam optimiser, using early stopping on the validation RMSE with a patience of 15 epochs. The loss was the mean squared error of the twelve outputs at the centre pixel of the patch only, masked for targets

that were missing — a plot with no spruce, for instance, contributes to the total and to the pine and broadleaved outputs but not to the spruce outputs. Restricting the loss to the centre pixel keeps the spatial alignment between prediction and reference exact while still allowing the network to use the full 480 m context.

3.4 Inference and plot-level aggregation

The trained networks were applied to the full scene in a tiled, fully convolutional manner, producing one georeferenced raster per variable at 10 m (or 30 m for the native-resolution control). Plot-level predictions were then obtained as area-weighted averages of the predicted pixels over each plot polygon, the weight of a pixel being the fraction of its area inside the polygon, evaluated on the same four-times refined grid used for the target rasters. This matches the support of the prediction to the support of the field measurement.

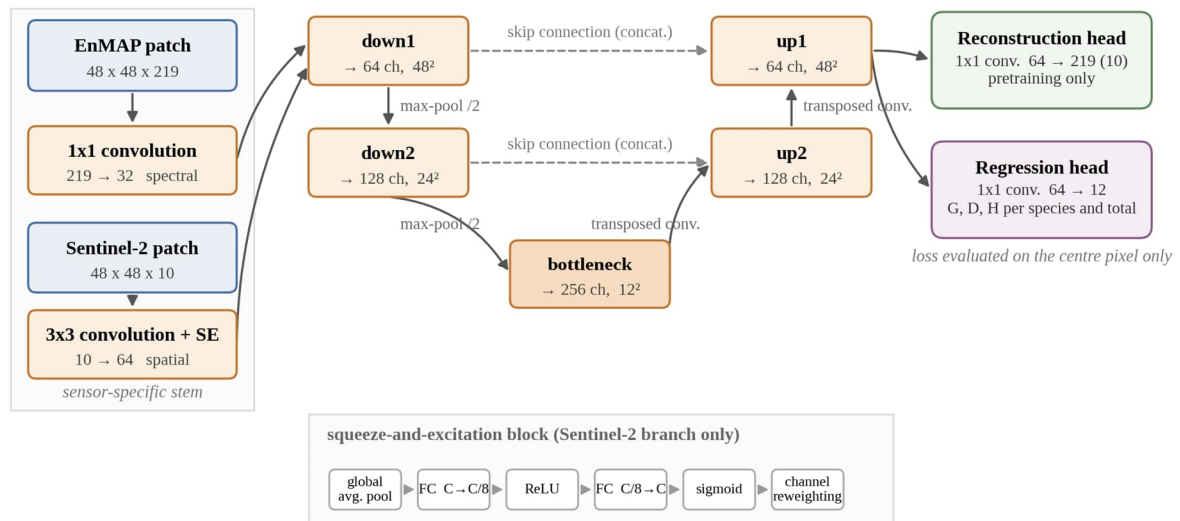


Figure 4. The two U-Net variants. Encoder, bottleneck and decoder are identical; the sensor-specific stem compresses the 219 EnMAP bands spectrally with a 1×1 convolution, whereas the Sentinel-2 branch uses a spatial convolution with channel attention. During supervised training the loss is evaluated on the centre pixel of the patch only.

3.5 Accuracy assessment

Accuracy was assessed on the 32 test plots, which were used neither for fitting nor for early stopping. With $\hat{y}_i$ being the predicted and y_i the measured value for plot i and n the number of plots,

$$\text{RMSE} = \left[(1/n) \sum (\hat{y}_i - y_i)^2 \right]^{1/2}, \quad (1)$$

$$\text{MAE} = (1/n) \sum |\hat{y}_i - y_i|, \quad (2)$$

$$\text{bias} = (1/n) \sum (\hat{y}_i - y_i), \quad (3)$$

$$\text{nRMSE} = \text{RMSE} / s_y, \quad (4)$$

where s_y is the standard deviation of the reference values over the test plots. The coefficient of determination R^2 was computed in the standard one-minus-residual-to-total-variance form, so that negative values are possible when the model performs worse than the mean of the test set.

3.6 Use of Artificial Intelligence (AI)

The code was originally written by the authors and was subsequently restructured, cleaned and debugged with the assistance of Claude (Anthropic) used as an AI coding agent. That work included a systematic review of the coherence of the approach including how the training, validation and test sets are separated throughout the processing chain. Claude identified that a training patch centred near a test plot could contain that plot's labelled pixels and contribute to the loss; this was corrected to the per-split construction described in Section 3.1, and all results reported here were produced with the corrected implementation on a fully independent test data set. The authors reviewed the changes. Claude has also been used to generate the flowcharts and figures used in this proceedings article. Responsibility for the correctness of the analysis rests with the authors.

Variable	Sensor	RMSE	MAE	bias	R^2	nRMSE
G ($\text{m}^2 \text{ha}^{-1}$)	EnMAP	5.50	3.58	-2.22	0.650	0.57
	Sentinel-2	5.31	4.11	-2.62	0.674	0.55
D (cm)	EnMAP	3.19	2.35	+0.71	0.803	0.40
	Sentinel-2	3.24	2.59	-0.47	0.797	0.41
H (m)	EnMAP	2.75	1.98	-0.24	0.792	0.43
	Sentinel-2	2.79	2.27	-1.31	0.787	0.44

Table 2. Accuracy of the total growing stock variables on the 32 test plots. G is basal area, D basal-area-weighted mean diameter and H basal-area-weighted mean height with the test-set means of $16.3 \text{ m}^2 \text{ha}^{-1}$, 13.9 cm and 12.5 m , respectively

4. RESULTS

4.1 Total growing stock

Test-set accuracies for the three total-stand variables are given in Table 2. The two sensors performed almost identically. EnMAP gave marginally higher R^2 for mean diameter (0.803 vs. 0.797) and mean height (0.792 vs. 0.787) and lower R^2 for basal area (0.650 vs. 0.674). The differences in R^2 are all below 0.03 and are well within the sampling variability of a 32-plot test set.

The mean absolute error was slightly more consistent: it was lower for EnMAP for all three variables — by 13 % for basal area (3.58 vs. $4.11 \text{ m}^2 \text{ha}^{-1}$), by 9 % for mean diameter (2.35 vs. 2.59 cm) and by 13 % for mean height (1.98 vs. 2.27 m). Since RMSE was essentially equal, this means that the EnMAP errors are more strongly concentrated in a few plots while the Sentinel-2 errors are spread more evenly. Both sensors underestimated basal area on average (bias -2.2 and $-2.6 \text{ m}^2 \text{ha}^{-1}$), which is the familiar saturation of optical estimators in dense stands.

Relative to the variability of the test set (nRMSE, Eq. 4), mean diameter and mean height were retrieved to about 0.40 and 0.43 of the reference standard deviation. Basal area was only retrieved with nRMSE = 0.56, making it the hardest of the three variables for both sensors.

4.2 Species-specific variables

The same network also predicted the three variables separately for Scots pine, Norway spruce and broadleaved trees (Table 3).

Species	Variable	RMSE		R^2	
		EnMAP	S-2	EnMAP	S-2
Scots pine	<i>G</i>	5.46	5.95	0.51	0.41
	<i>D</i>	5.53	5.27	0.65	0.68
	<i>H</i>	4.64	4.08	0.64	0.72
Norway spruce	<i>G</i>	4.52	5.34	0.26	−0.04
	<i>D</i>	4.32	4.24	0.53	0.55
	<i>H</i>	3.14	3.22	0.69	0.67
Broadleaved	<i>G</i>	4.20	3.57	0.47	0.62
	<i>D</i>	6.95	5.75	0.01	0.33
	<i>H</i>	6.90	5.57	0.04	0.38

Table 3. Species-specific accuracy on the test plots for EnMAP and Sentinel-2 (S-2). Units as in Table 2.

Accuracy dropped sharply and was no longer consistent between the sensors. Pine and spruce variables were retrieved with moderate success (R^2 between 0.25 and 0.72), with EnMAP better for spruce basal area (0.253 vs. −0.041) and Sentinel-2 better for pine height (0.719 vs. 0.636). For broadleaved trees, the EnMAP model essentially failed for diameter and height (R^2 of 0.014 and 0.041 against 0.326 and 0.376 for Sentinel-2).

The species-specific results should be read as a limitation of the reference data rather than as a property of the sensors. Broadleaved trees form a minor component of most plots – their mean basal area over the test set was 3.2 m² ha^{−1} against 16.3 m² ha^{−1} for the total – so the species-specific targets are dominated by small values with a long tail, and the number of plots carrying usable species-level information is far smaller than 149. We report the species-specific figures for completeness and do not draw sensor comparisons from them.

4.3 Spatial context and EnMAP sampling

The size of the image patch given to the network sets the spatial context available for each prediction. Table 4 gives the test-set R^2 of models trained with patches from 16 to 128 pixels on the bilinearly resampled EnMAP data and the corresponding Sentinel-2 data. Accuracy rose steeply from 160 m to about 480 m of context and then declined again: the smallest patches do not give the network enough neighbourhood to work with, whereas the largest ones dilute the signal of the plot itself and reduce the effective number of independent training samples. A patch of 48 × 48 pixels was used for the main results.

The lower block of Table 4 gives the control experiment in which EnMAP was used at its native 30 m sampling, with the patch size in pixels reduced so that the ground extent of the context matched that of the Sentinel-2 patch. At a matched extent of 480 m, the native-resolution EnMAP model reached R^2 of 0.68 for basal area, 0.72 for mean diameter and 0.75 for mean height. Basal area was thus retrieved better than from the resampled data, whereas mean diameter and mean height were retrieved slightly worse. Spatial resampling of the EnMAP cube is therefore convenient but not essential, and none of the conclusions depend on it.

Note that the runs in Table 4 predicted only the three total-stand variables and used a shorter fine-tuning schedule than the main model, so their values are not directly comparable with Table 2.

5. DISCUSSION

The key result is that over the set of 213 boreal forest plots which were covered by a single EnMAP scene, a nine-fold coarser ground sampling did not cost measurable accuracy in the retrieval

Context	Patch	<i>G</i>		<i>D</i>		<i>H</i>	
		EnMAP	S-2	EnMAP	S-2	EnMAP	S-2
EnMAP resampled to 10 m							
160 m	16 px	−0.03	−0.15	0.62	0.15	0.54	0.14
320 m	32 px	0.46	0.55	0.81	0.63	0.78	0.61
480 m	48 px	0.59	0.48	0.80	0.75	0.79	0.79
640 m	64 px	0.59	0.50	0.78	0.74	0.78	0.72
1280 m	128 px	0.33	0.54	0.62	0.70	0.60	0.72
EnMAP at native 30 m							
240 m	8 px	0.55	0.46	0.62	0.74	0.66	0.73
480 m	16 px	0.68	0.58	0.72	0.73	0.75	0.73
960 m	32 px	0.50	0.46	0.54	0.72	0.63	0.72

Table 4. Test-set R^2 as a function of the ground extent of the training patch. The patch column gives the EnMAP patch size in pixels; the Sentinel-2 (S-2) patch always covered the same ground extent at 10 m. Rows in bold are the final configurations used for the main results.

of basal area, mean diameter or mean height. The loss of spatial detail by the hyperspectral sensor is compensated for by the additional spectral bands. At the same time, the differences we observed are small compared with the uncertainty of the comparison itself. With 32 test plots, a difference of 0.03 in R^2 is not statistically meaningful, and the range of forest conditions represented in a single image over one study area is limited.

The study here has several clear shortcomings. The evaluation rests on a single date and a single site. A proper test requires spatially continuous acquisitions across several sites and phenological stages. Second, only 213 plots were available, which limits both the accuracy attainable by a convolutional model and the precision of the comparison. The self-supervised pretraining stage was introduced to mitigate this, and quantifying its contribution would be a useful ablation. Third, the species-specific results (Section 4.2) show that the reference data did not carry enough species-level information to support the twelve-variable formulation which would be needed in a digital twin. Nevertheless, they serve as indication of the capability of future hyperspectral mapping missions over forests where Sentinel-2 has proven its value (e.g., Astola et al., 2019, Fang et al., 2023).

These shortcomings will be mitigated by the increasing availability of hyperspectral data. Since the acquisition of the imagery, EnMAP has switched to always recording at least three consecutive tiles. In cloudy areas such as the boreal zone, making best use of cloudless conditions is a necessity for programmable imagers such as EnMAP or PRISMA (Cogliati et al., 2021) to acquire data which can be matched to the routine field data — e.g., environmental or agricultural — measured regardless of weather.

For a forest digital twin, the practical question is not which sensor is marginally better, but whether imaging spectroscopy can serve as a baseline information source. The results here suggest that it can: the structural variables that initialise and constrain a forest growth model are retrieved from EnMAP as accurately as from Sentinel-2, and they are obtained from data that additionally carry information on canopy biochemistry (Farmonov et al., 2025) and biodiversity (Torresani et al., 2025). A twin fed by CHIME or SBG would obtain the structural state and the functional state of the forest from a single data stream, making data assimilation more robust.

6. CONCLUSIONS

Spaceborne hyperspectral imagery contains the same information as the Sentinel-2 MultiSpectral Imager, which has 9 times better spatial resolution. Using a near-simultaneous EnMAP and Sentinel-2 image pair, identical reference data and a common U-Net framework, basal area, mean diameter and mean height were retrieved with equal accuracy from the two sensors (R^2 of 0.65, 0.80 and 0.79, respectively, for EnMAP against 0.67, 0.80 and 0.79 for Sentinel-2), with mean absolute errors being 9–13 % lower for EnMAP. The result held whether EnMAP was resampled to the Sentinel-2 grid or used at its native 30 m sampling.

This result supports a conclusion that spaceborne imaging spectroscopy can therefore serve as a baseline information source for key forest structural variables in a forest digital twin, with the additional, so far unexploited potential of mapping canopy biochemistry, biodiversity and vegetation functioning from the same data.

ACKNOWLEDGEMENTS

This work has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101187992 (TerraDT) and the European Space Agency (Forest DTC project). Code and data are available from the authors upon request subject to the restrictions of the hyperspectral and forest data licenses.

REFERENCES

- Astola, H., Häme, T., Sirro, L., et al., 2019. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. *Remote Sens. Environ.*, 223, 257–273. doi.org/10.1016/j.rse.2019.01.019.
- Astola, H., Seitsonen, L., Halme, E., et al., 2021. Deep neural networks with transfer learning for forest variable estimation using Sentinel-2 imagery in boreal forest. *Remote Sens.*, 13(12), 2392. doi.org/10.3390/rs13122392.
- Bauer, P., Stevens, B., Hazeleger, W., 2021. A digital twin of Earth for the green transition. *Nat. Clim. Change*, 11(2), 80–83. doi.org/10.1038/s41558-021-00986-y.
- Cawse-Nicholson, K., Townsend, P. A., Schimel, D., et al., 2021. NASA's surface biology and geology designated observable: A perspective on surface imaging algorithms. *Remote Sens. Environ.*, 257, 112349. doi.org/10.1016/j.rse.2021.112349.
- Chabrilat, S., Foerster, S., Segl, K., et al., 2024. The EnMAP spaceborne imaging spectroscopy mission: Initial scientific results two years after launch. *Remote Sens. Environ.*, 315, 114379. doi.org/10.1016/j.rse.2024.114379.
- Cogliati, S., Sarti, F., Chiarantini, L., et al., 2021. The PRISMA imaging spectroscopy mission: overview and first performance analysis. *Remote Sens. Environ.*, 262, 112499. doi.org/10.1016/j.rse.2021.112499.
- Drusch, M., Del Bello, U., Carlier, S., et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.*, 120, 25–36. doi.org/10.1016/j.rse.2011.11.026.
- Fang, G., Xu, H., Yang, S.-I., et al., 2023. Synergistic use of Sentinel-1, Sentinel-2, and Landsat 8 in predicting forest variables. *Ecol. Indic.*, 151, 110296. doi.org/10.1016/j.ecolind.2023.110296.
- Farmonov, N., Walden, S., Martinée, E., et al., 2025. Optimizing hybrid models for forest leaf and canopy trait mapping from EnMAP hyperspectral data with limited field samples. *Sci. Remote Sens.*, 12, 100253. doi.org/10.1016/j.srs.2025.100253.
- Finnish Forest Centre, 2026. Open forest and nature information. www.metsakeskus.fi/en/open-forest-and-nature-information (26 August 2026).
- Guanter, L., Kaufmann, H., Segl, K., et al., 2015. The EnMAP spaceborne imaging spectroscopy mission for Earth observation. *Remote Sens.*, 7(7), 8830–8857. doi.org/10.3390/rs70708830.
- Hari, P., Kulmala, M., 2005. Station for Measuring Ecosystem–Atmosphere Relations (SMEAR II). *Boreal Environ. Res.*, 10(5), 315–322.
- He, K., Chen, X., Xie, S., et al., 2022. Masked autoencoders are scalable vision learners. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, 15979–15988. doi.org/10.1109/CVPR52688.2022.01553.
- Hill, J., Buddenbaum, H., Townsend, P. A., 2019. Imaging spectroscopy of forest ecosystems: Perspectives for the use of space-borne hyperspectral Earth observation systems. *Surv. Geophys.*, 40(3), 553–588. doi.org/10.1007/s10712-019-09514-2.
- Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-excitation networks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, 7132–7141. doi.org/10.1109/CVPR.2018.00745.
- Juola, J., Hovi, A., Rautiainen, M., 2024. Comparison of contemporaneous Sentinel-2 and EnMAP data for vegetation index-based estimation of leaf area index and canopy closure of a boreal forest. *Eur. J. Remote Sens.*, 57(1), 2432975. doi.org/10.1080/22797254.2024.2432975.
- Lang, N., Jetz, W., Schindler, K., Wegner, J. D., 2023. A high-resolution canopy height model of the Earth. *Nat. Ecol. Evol.*, 7(11), 1778–1789. doi.org/10.1038/s41559-023-02206-6.
- Minunno, F., Peltoniemi, M., Launiainen, S., et al., 2016. Calibration and validation of a semi-empirical flux ecosystem model for coniferous forests in the boreal region. *Ecol. Modell.*, 341, 37–52. doi.org/10.1016/j.ecolmodel.2016.09.020.
- Minunno, F., Peltoniemi, M., Härkönen, S., et al., 2019. Bayesian calibration of a carbon balance model PREBAS using data from permanent growth experiments and national forest inventory. *For. Ecol. Manage.*, 440, 208–257. doi.org/10.1016/j.foreco.2019.02.041.
- Möttus, M., Dees, M., Astola, H., et al., 2021. A methodology for implementing a digital twin of the Earth's forests to match the requirements of different user groups. *GI Forum*, 9(1), 130–136. doi.org/10.1553/giscience2021_01_s130.
- Rast, M., Nieke, J., Adams, J., Isola, C., Gascon, F., 2021. Copernicus Hyperspectral Imaging Mission for the Environment (CHIME). In: 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Brussels, 108–111. doi.org/10.1109/IGARSS47720.2021.9553319.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science 9351, Springer, Cham, 234–241. doi.org/10.1007/978-3-319-24574-4_28.
- Torresani, M., Rossi, C., Mina, M., et al., 2025. Spectral heterogeneity from the spaceborne imaging spectrometer EnMAP reveals biodiversity patterns in forest ecosystems. *Int. J. Appl. Earth Obs. Geoinf.*, 144, 104902. doi.org/10.1016/j.jag.2025.104902.
- Wang, Y., Albrecht, C. M., Braham, N. A. A., Mou, L., Zhu, X. X., 2022. Self-supervised learning in remote sensing: A review. *IEEE Geosci. Remote Sens. Mag.*, 10(4), 213–247. doi.org/10.1109/MGRS.2022.3198244.