

Article

Not peer-reviewed version

ReUse: REgressive Unet for Carbon Storage and above Ground Biomass Estimation

[Antonio Elia Pascarella](#) , [Giovanni Giacco](#) , [Mattia Rigioli](#) , [Stefano Marrone](#) ^{*} , [Carlo Sansone](#)

Posted Date: 15 February 2023

doi: [10.20944/preprints202212.0158.v2](https://doi.org/10.20944/preprints202212.0158.v2)

Keywords: U-Net; Carbon Storage; Above-Ground Biomass; Remote Sensing; Deep Learning; CNN; Sentinel-2; ESA CCI Biomass Project



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ReUse: REgressive UNet for Carbon Storage and Above-Ground Biomass Estimation

Antonio Elia Pascarella ¹, Giovanni Giacco ^{1,2}, Mattia Rigioli ², Stefano Marrone ^{1,*} and Carlo Sansone ¹

¹ Department of Electrical Engineering and Information Technology (DIETI), University of Naples Federico II, Via Claudio 21, 80125 Naples, Italy; antonioelia.pascarella@unina.it (A.E.P.); giovanni.giacco@unina.it (G.G.); carlo.sansone@unina.it (C.S.)

² Latitudo 40, Via Emanuele Gianturco 31/c, 80146 Naples, Italy; mattia.rigioli@latitudo40.com

* Correspondence: stefano.marrone@unina.it

Abstract: United Nations Framework Convention on Climate Change (UNFCCC) has recently established the Reducing Emissions from Deforestation and forest Degradation (REDD+) program that requires countries to report their carbon emissions and sink estimates through national greenhouse gas inventories (NGHGI). Thus, developing automatic systems capable of estimating the carbon absorbed by forests without in-situ observation becomes essential. To support this critical need, in this work, we introduce ReUse, a simple but effective deep-learning approach to estimate the carbon absorbed by forest areas based on remote sensing. The proposed method's novelty is in using the public above-ground biomass (AGB) data from the European Space Agency's Climate Change Initiative Biomass project as ground truth to estimate the carbon sequestration capacity of any portion of land on Earth using Sentinel-2 images and a pixel-wise Regressive UNet. The approach has been compared to two literature proposals using a private dataset and human-engineered features. The results show a more remarkable generalization ability of the proposed approach, with a decrease in Mean Absolute Error and Root Mean Squared Error over the runner-up, respectively of 16.9 and 14.3 in the area of Vietnam, 4.7 and 5.1 in the area of Myanmar, 8.0 and 1.4 in the area of Central Europe. As a case study, we also report an analysis made for the Astroni area, a World Wildlife Fund (WWF) natural reserve struck by a large fire, producing predictions consistent with values found by experts in the field after in-situ investigations. These results further support the use of such an approach for the early detection of AGB variations in urban and rural areas.

Keywords: U-Net; carbon storage; above-ground biomass; remote sensing; deep learning; CNN; Sentinel-2; ESA CCI biomass project

1. Introduction

Accurate assessment of forest above-ground biomass (AGB), which in this work is defined as the mass expressed as oven-dry weight of the woody parts (stem, bark, branches, and twigs) of all living trees excluding stump and roots, and related carbon stock is essential for the sustainable management of forests. Recently, the United Nations Framework Convention for Climate Change (UNFCCC) established the Reducing Emissions from Deforestation and forest Degradation (REDD+) that requires countries to report their carbon emissions and sink estimates through national greenhouse gas inventories (NGHGI) [1,2]. Furthermore, Sustainable Development Goal 15 aims to “protect, restore and promote the sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, halt and reverse land degradation and halt biodiversity loss” [3]. Subsequently, it is paramount to conduct an explicit investigation into the methods and procedures for quantifying carbon sinks. Generally, the above-ground dry biomass holds about 50% of carbon; as such, a friction factor of 0.5 is commonly used for converting AGB into carbon concentration [4,5].

There are two basic approaches to obtaining biomass estimation: traditional field-based and remote sensing (RS) methods. There is no doubt that traditional methods are more accurate [6],

still, they are also way more time-consuming, laborious, challenging to implement in inaccessible areas, and destructive in nature [7]. A more and more investigated solution is to rely on images acquired by satellites and/or drones. Among all, the Sentinel-2 satellite system, launched on 23 June 2015 by the European Space Agency (ESA) and equipped with a multi-spectral instrument (MSI), presents a high potential for applications in land management, agricultural industry (food security), forestry (AGB) disaster control, and humanitarian relief operations [8]. Sentinel-2 is a polar-orbiting satellite system comprised of two satellites, each carrying an MSI characterized by a 290-km swath width, offering a multi-purpose design of 13 spectral bands traversing from visible and near-infrared (NIR) wavelengths to shortwave infra-red wavelengths at refined (10, 20 m) and coarse (60 m) spatial resolution. Furthermore, the presence of four bands within the red-edge region, centred at 705 (band 5), 740 (band 6), 783 (band 7), and 865 nm (band 8a), gives the satellite system the potential for mapping various vegetation characteristics [9].

Most of the approaches so far developed make use of classical machine learning models (such as SVM, Random Forest, etc.) to estimate above-ground biomass and the related carbon stock, leveraging expert-made features to be extracted from Sentinel-2 images [10–14]. More recently, some works [15,16] are investigating the use of Convolutional Neural Network (CNN) to estimate AGB using the commercial satellite Worldview-2 and visible spectrum images captured by an unmanned aerial vehicle. In all the reported examples, the AGB target variable is always collected by field measurements (i.e., they rely on data collected by experts physically located on the target terrain). This characteristic strongly limits the usability of AI-based AGB automatic estimators, as their use in a real context, such as the continual analysis of the impact of natural disasters as well as monitoring the effectiveness of environmental policies, requires the reliable acquisition of the ground truth on wide and heterogeneous landscapes.

To cope with this need, in this work we propose to use the open data released by the ESA Climate Change Initiative (CCI) BIOMASS project as ground truth for AGB. This, together with the images acquired by the Sentinel-2 system allows, unlike field measurements, the acquisition of AGB on a global scale. To the best of our knowledge, this is the first time that this data combination has been used for AGB estimation. The rest of the paper is as follows: Section 2 reports a detailed review of the approaches so far proposed for AGB estimation from satellite images; Section 3 details the proposed approach, as well as the considered dataset, competition and experimental setup; Section 4 reports the obtained results and a case study analysis performed on a WWF natural reserve struck by a severe wildfire; Finally, Section 5 draw some conclusions.

2. Related Works

Literature on remote sensing for vegetation detection is wide. However, Above-Ground Biomass detection is a different and more crucial task, especially when needed for detailed estimation of other indicators, such as the carbon sequestration ability of a portion of land. Thus, in this section, we focus only on works aimed at estimating above-ground biomass and carbon storage from Sentinel-2 images.

In [13], the authors attempted to examine the prospects of Sentinel-2 spectral data for quantifying the carbon stock in a reforested urban landscape, using Random Forest and, as predictors, 10 Sentinel bands plus 15 spectral indices that summarise the spectral content without taking into account in any way the spatial correlations typical of an image. Similarly, in [10] the authors applied Sentinel-2 satellite images combined with field-measured biomass using Random Forest (RF) to estimate above-ground biomass in Yok Don National Park, Vietnam. A total of 132 spectral and texture variables were extracted from the Sentinel-2 images; the grey level co-occurrence matrix (GLCM) method [17] was used to compute the texture variables.

In [12], Sentinel-2 performance was evaluated for a buffer zone community forest in Parsa National Park, Nepal, using field-based AGB as a dependent variable, as well as spectral band values and spectral-derived vegetation indices as independent variables in the Random Forest algorithm; in this study, no features were extracted from the spatial dimensions, but indicators were only extracted

from the spectral dimension of the input tensor. Spectral bands, vegetation indices (VIs) and texture variables derived from processed Sentinel-2 data and topographic parameters have been used in [14] to statistical link with field-based AGB by implementing random forest and stochastic gradient boosting (SGB) algorithms. The grey level co-occurrence matrix method [17] and wavelet decomposition were applied using the first principal component of the Sentinel-2 multispectral tensor.

In [18], to estimate the AGB from remotely sensed data, parametric and non-parametric methods, including Multiple Regression (MR), k-Nearest Neighbour (kNN), Random Forest, the multi-layer perceptron, which performed best among the various methods, were applied to a single Sentinel-2 image using spectral bands and derived indices. Similarly, in [19] the authors explored the capability of spectral and texture features from the Sentinel-2 Multispectral Instrument (MSI) for modelling grassland AGB using random forest (RF) and extreme gradient boosting (XGBoost) algorithms in Shengjin Lake wetland (a Ramsar site), showing that the RF and XGBoost models had a robust and efficient performance and that the introduction of eight grey-level co-occurrence matrix (GLCM) textures moderately improved the accuracy of modelling AGB. The texture is also the core of [20], in which texture metrics were derived based on different working window sizes (3×3 , 5×5 , 7×7 , and 9×9), and the results were compared with those obtained using raw traditional bands (Band 2, 3, 4, 8, 11 and 12), raw traditional and red edge bands (Band 5, 6, 7 and 8A), and red edge bands only; the use of texture with a 7×7 window size and vegetation indices (VIs) yielded higher biomass estimates.

In [21] the authors proposed an innovative and dynamic architecture based on the generative neural network that extracts target-oriented generative features for forest prediction AGB using satellite data. The architecture exploits its generative capacity to produce variables in a latent space to predict AGB, exploiting only the spectral dimension and not the spatial correlations of the images. In the same year, in [22] the authors present an automated machine learning (AutoML) framework for modelling, evaluating, and stacking multiple base models for AGB prediction. This work incorporates a hyperparameter optimization procedure for automatically extracting targeted features from multitemporal Sentinel-2 data that minimizes human bias. Also, in this context, the automatic feature extraction took into account only the spectral dimension. Finally, a recent work [23] mainly discusses three non-parametric models: the artificial neural network (ANN) and the random forests, and in particular, explores the quantile regression neural network (QRNN), using spectral index and texture features as variables.

These works use machine learning techniques combined with intensive feature extraction, with some focusing only on the spectral dimension and others involving both the spatial and spectral dimensions. Besides the used features, some works [21,22] also exploited generative networks and AutoML pipelines to minimize the human bias in the feature extraction phase. Moreover, as for other image processing related domains, researchers [15,16] are also working on the use of Convolutional Neural Networks (CNNs), designed to produce numerical values for AGB prediction (one for each input image), using input from commercial satellites such as Worldview-2 or visible spectrum images captured by an unmanned aerial vehicle instead of Sentinel-2 open data.

As far as we know, no deep learning approach relying on fully convolutional architecture (such as UNet) trained on Sentinel-2 multi-spectral images has been proposed so far to predict AGB and carbon storage. In this work, we cope with this lack with the aim of introducing a new approach that embeds feature extraction within the network to produce not a simple numerical value for the whole input image (as in the case of CNNs) but a mask of numerical values associating the AGB value to each pixel in the input image. This combines a very fine-grained estimation and wide computational advantages, together with a greater generalization ability, especially in the case of learning over large and variegated geographic areas. Moreover, all the reported works, including recent ones [24] designed to provide an in-depth analysis of forests and individual trees' carbon storage estimation, use AGB field measurements as the target variable and/or require input data for which field surveys are necessary. Instead, in our work we rely on the ESA public data for AGB. The aim is to show that the combined use of a Regressive UNet network with public data (Sentinel-2 and ESA AGB) can help monitor carbon

content in forest areas at a global scale in close a real-time manner (i.e., as soon as a new image is available from the Sentinel-2 system).

3. Materials and Methods

In this paper, we introduce a Regressive UNet trained on the public above-ground biomass data from the European Space Agency's Climate Change Initiative Biomass project as ground truth to estimate the carbon sequestration capacity of any portion of land on Earth using Sentinel-2 images, comparing its performance against two literature proposals [10,14] on their respective study areas. Section 3.1 introduces the proposed approach, describing ideas and motivations. Section 3.2 details the experimental setup as well as the considered competitors. Finally, Section 3.3 describes the considered dataset, focusing on data acquisition and pre-processing.

3.1. Regressive UNet

In the following study, we introduce ReUse, a UNet network trained to carry out a pixel-wise regression task to map Sentinel-2 images into AGB rasters. The UNet was developed by [25] for biomedical semantic segmentation. In the original proposal, the architecture contains two paths. The first is a contraction path (also known as the encoder) designed to capture the context in the image. Several structures are possible, but usually, they all involve sequences of convolutional and max pooling layers. The second path is a symmetric expanding path (also known as the decoder) designed to produce a pixel-wise prediction using transposed convolutions. These two paths are connected by some skipping connections, designed to improve the localization ability of the network by combining the high-resolution features from the contracting path with the corresponding one in the expanding path. A final convolution layer can then be used to learn and assemble a more precise output based on this information.

The main difference we introduce to the original UNet architecture is that the network has been trained not for performing semantic segmentation but to produce a pixel-wise regression map. This has been obtained by omitting the softmax in the last layer, thus forcing the optimiser to minimise the loss function based on the actual values predicted by the network, directly comparing them with the AGB ground truth (Figure 1). To the best of our knowledge, this is the first time such an approach has been used to estimate AGB. The main advantage is that the proposed variant is able to extract both spatial and spectral features from the satellite multi-spectral images using an end-to-end paradigm. In particular, a patch-wise approach has been used [26] dividing each Sentinel-2 input and AGB-raster into non-overlapping patches of 16*16 pixels.

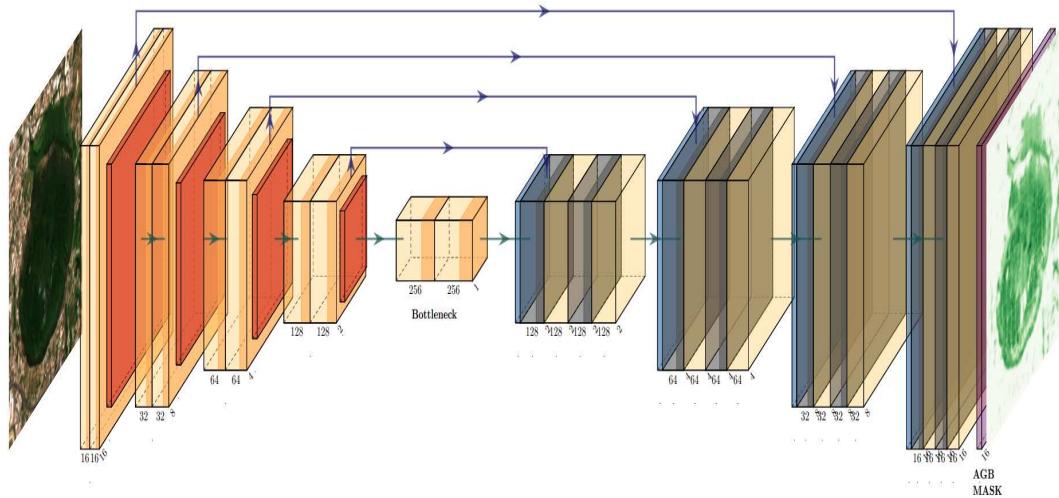


Figure 1. ReUse architecture for pixel-wise regression. The input is the Sentinel-2 image with dimensions (patch size, patch size, number of channels); the output is the AGB image with dimensions (patch size, patch size, 1).

3.2. Experimental Setup

One of the problems with the evaluation of remote-sensing applications is the lack of a standard experimental protocol. Indeed, given the high spatial and temporal variability associated with images captured by satellites' sensors, selecting different Earth zones and acquisition times can result in different performance, making any comparison with other literature approaches less fair. To cope with this, we perform different evaluation on different Earth zones, comparing ReUse against two literature proposals [10,14] both leveraging a Random Forest [27] classifier trained on spatial and spectral features in a end-to-end paradigm. It is worth noting that in [14] the authors also use topographical parameters, such as altitude as a variable. However, as that data is not always available, in this work only Sentinel-2 data were used as input.

Comparisons between ReUse and the [10,14] approaches were made on study areas used by the latter, located in the Central Highlands of Vietnam, and the Yinmar Forest (YM) located in the northern and central-eastern part of Myanmar, respectively. In [10], a total of 132 spectral and texture variables were extracted from the Sentinel-2 images; the grey level co-occurrence matrix (GLCM) method [17] was used to compute mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation. In [14], principal component analysis (PCA) [28] was used to eliminate correlated information in satellite images and simultaneously reduce their dimensionality. The first principal component (PC1) was used for texture extraction. When extracting textural features, the grey level co-occurrence matrix method [17] was used to compute mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation, and wavelet decomposition was also applied considering their usefulness for the representation of relevant features [29]. The wavelet analysis produces four essential components: the approximation image, horizontal detail, vertical detail, and diagonal detail images. The latter three are regarded as helpful textural measures. In particular, the Coiflet discrete wavelet function was chosen. Thus, based on the first principal component, a three-level decomposition strategy was implemented to generate nine detailed images as independent textural variables for AGB modeling. Finally, two types of textures derived from GLCM-based and wavelet analysis were included in the AGB modeling in combination with 11 spectral indices.

The proposed ReUse architecture has instead been trained using two different setups: one relying only on the raw Sentinel-2 bands, suitably normalized; the other leveraging raw band together with textural and spectral indices as in [10,14] respectively. Focusing on the latter setup, concerning the texture variables, the GLCM method was adopted to compute mean, variance, homogeneity, contrast,

dissimilarity, entropy, second moment, and correlation to which the nine detail images obtained by applying the Coiflet wavelets have been added by adopting a three-level strategy. The texture variables were computed using the first principal component of the raw Sentinel bands. In this work, a 5×5 kernel was used to construct the GLCM-based features. Indeed, considering that training was conducted at a spatial resolution of 100 meters (because this is the spatial resolution of ESA's AGB data), such a kernel involves an area of 500 meters by 500 meters.

For all the considered approaches (including ours and competitors), the images have been rescaled to a spatial resolution of 100 meters to match the ESA CCI Biomass Project AGB data. All the experiments have been run using an 8-fold cross-validation strategy, where each fold contains only patches associated with the same zone on Earth while different folds refer to different portion of Earth (thus ensuring that there is no data leaks between different folds). For each iteration, 6 of the eight folds were used for training, 1 for validation, and 1 for testing. The validation set was used to optimize the number of epochs for ReUse with an early stopping procedure [30] and the number of trees for the two competitors using two possible values: 250 or 500. In this study, concerning the early stopping procedure, training is stopped when the monitored validation loss has stopped improving after 35 epochs; the maximum number of epochs is set at 500. The optimizer used is Adam [31] with the default parameters. For the learning rate, following the approach introduced in [32], if no improvement in validation loss is seen for 25 epochs, the learning rate is reduced by a factor of 0.2. The Mean Absolute Error is used as a loss function for ReUse. The trees in the Random Forest are maximally grown, and the number of variables that each tree can choose at each split is equal to the square root of the number of features, as suggested by [33]. For reproducibility purposes, the source code is available at our GitHub repository¹.

3.3. Image Acquisition and Pre-Processing

Starting from the Global Dataset of above-ground biomass of the year 2018 version 3 of the ESA CCI BIOMASS Project, three study areas were downloaded and used separately to compare ReUse with the competitors [10,14]. The file “N20E100” contains the AGB of the study area of [10] in Vietnam, while the file “N30E90” contains the AGB of the study area of Yinmar (YM) forest of [14] in Myanmar. It is again emphasized that [10,14] use field AGB data, while in this work, public AGB data provided by ESA are used. The third study area cut from the “N60E00” file in central Europe was used to compare the approaches in a western area and for the Astroni use case (Section 4.1). Table 1 shows the names of the downloaded AGB files in GeoTIFF format and the clipped areas of interest in Well-known text (WKT), a text markup language for representing vector geometry objects. From the WKTs, it was possible to download the corresponding Sentinel-2 L2A satellite multispectral images acquired during available cloud-free days. The dates of the downloaded Sentinel-2 images of Vietnam, Myanmar, and Europe are three April, seven March, and twenty-seven July 2018, respectively.

¹ <https://github.com/priamus-lab/ReUse>

Table 1. List of the file names downloaded from the ESA Biomass Climate Change Initiative of the Global above-ground forest biomass for 2018, v3 in Geotiff format containing AGB rasters. In each GeoTiff file, a study area was cut out and reported in Well-known text (WKT) format.

AGB file name	Area of interest represented in WKT
N60E00	Polygon ((6.41116296045328227 50.733179027500789, 7.47810311940281469 50.733179027500789, 7.47810311940281469 51.57252282701965385, 6.41116296045328227 51.57252282701965385, 6.41116296045328227 50.733179027500789))
N20E100	Polygon ((107.10766723771612874 12.51511413035953346, 107.83207927588165376 12.51511413035953346, 107.83207927588165376 13.26250192346686596, 107.10766723771612874 13.26250192346686596, 107.10766723771612874 12.51511413035953346))
N30E90	Polygon ((96.00143548690360262 22.97297258442770485, 96.49905818628494103 22.97297258442770485, 96.49905818628494103 23.42872411816529876, 96.00143548690360262 23.42872411816529876, 96.00143548690360262 22.97297258442770485))

The Sentinel-2 satellite system acquires images with 13 spectral channels at variable spatial resolutions of 10, 20 and 60 meters. This satellite system covers the red-edge region (i.e., b5, 6, 7, and 8A), strategically positioned in the electromagnetic spectrum with unique band settings critical for vegetation modelling [34]. For ReUse, bands 1, 9, and 10 were eliminated due to their coarse spatial resolution, resulting in the use of only 10-bands images. The values in the retrieved rasters are digital numbers (DN) that must be transformed into reflectance by dividing them by the quantification value. The quantification value in the Sentinel-2 product metadata is equal to 10000 [35]. The infrastructure provided by the company Latitudo 40 was used to download and prepare the Sentinel-2 data described above. Concerning the AGB data, the dataset [36] comprises estimates of forest above-ground biomass for 2010 [37], 2017, and 2018. They are derived from a combination of Earth observation data, depending on the year, from the Copernicus Sentinel-1 mission, Envisat's ASAR instrument, and JAXA's Advanced Land Observing Satellite (ALOS-1 and ALOS-2), along with additional information from Earth observation sources. The data has been produced as part of the European Space Agency's Climate Change Initiative program by the Biomass CCI team. The mapping is at 100 m grid spacing with a target relative error of less than 20% where AGB exceeds 50 Mg/ha.

4. Results

As described in the previous section, the proposed ReUse approach (both versions) has been compared against two machine learning approaches [10,14] using Random Forest and leveraging both spatial and spectral dimensions during the feature-engineering phase. The study areas used are the area of Vietnam and the area of Myanmar, containing the areas of the articles mentioned above, respectively, to compare ReUse with competitors in the areas in which these methodologies were initially designed. Moreover, we also consider an area in Central Europe to test the models on the territorial characteristics of a western country. Eight-fold cross-validation was used to estimate the error of the models. Table 2 presents the experiments for Vietnam, Myanmar, and Central Europe, highlighting that ReUse performs better than its competitors in terms of MAE, RMSE, and R^2 .

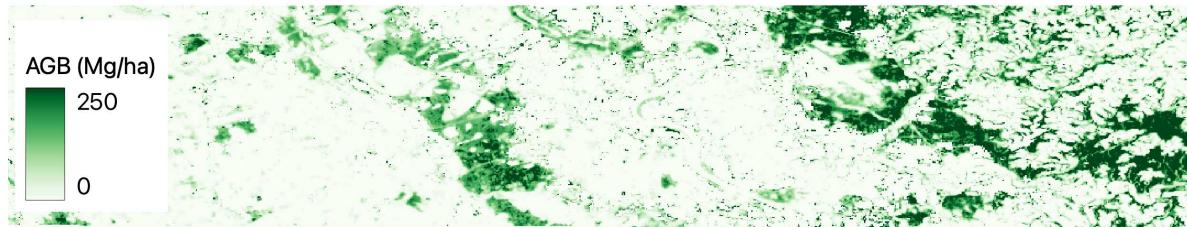


Figure 2. Base-map of the test area in Central Europe.

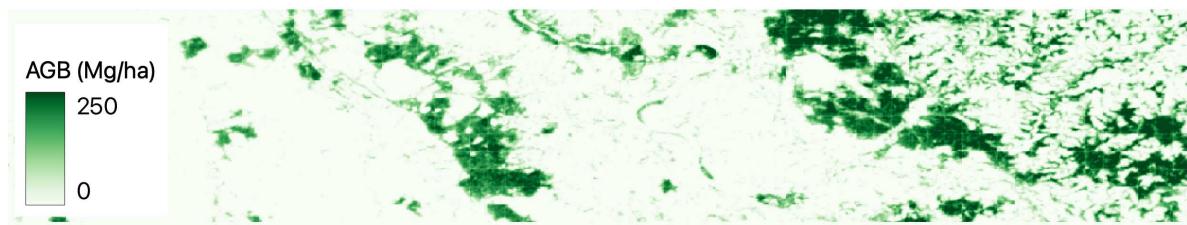


Figure 3. ReUse's AGB predictions with raw bands in Central Europe.

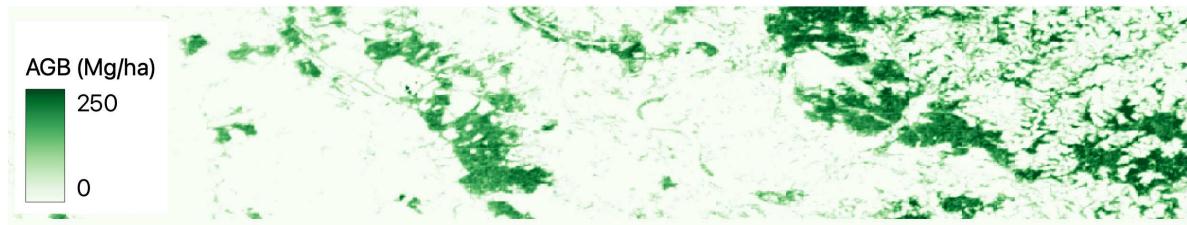


Figure 4. ReUse's AGB predictions with raw bands and feature extraction in Central Europe.



Figure 5. Predictions of AGB using the machine learning approach [10] in Central Europe.

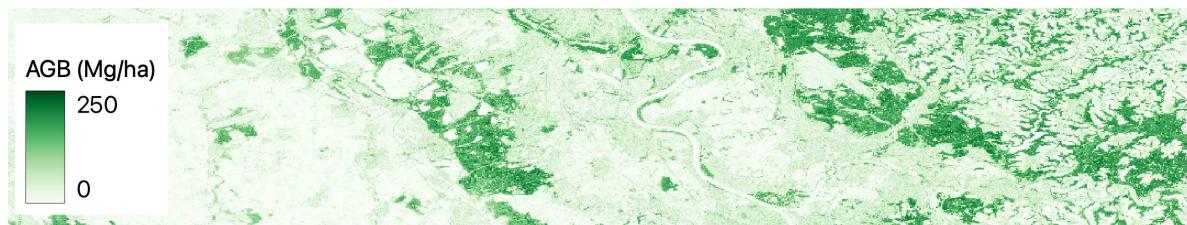


Figure 6. Predictions of AGB using the machine learning approach [14] in Central Europe.

Figures 2 to 6 show a base map of the test area in Central Europe and the AGB predictions of the different approaches. It can be seen that the city areas where there is no greenery are appropriately set to 0, and the tree areas are all highlighted. In particular, our solutions (Figure 3 and 4) show more marked differences between green and non-green areas than the classical machine learning approaches (Figure 5 and 6). Furthermore, the experiments show that ReUse with feature extraction, 42 spectral indexes as in [10] in conjunction with texture variables obtained with GLCM method and

wavelet analysis as in [14], does not significantly improve compared to ReUse with raw bands. This suggests that such a deep approach, in the contest of AGB prediction, can avoid the feature engineering phase. It is worth noting that errors reported for the competitors are higher than those reported in the corresponding papers. This is because the ground truth used in this paper comes from ESA, while the original works used field measurements, not released for reproducibility purposes.

Table 2. The results of the experiments performed on Vietnam and Myanmar study areas of [10,14] and Central Europe study area. At each iteration of the eight-fold cross-validation, six folds are used for training, one for validation, and one for testing. The averages and standard deviations of the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 metrics are calculated on the test set. In bold, the best value for each metric, per study area.

area	model	MAE	RMSE	(R^2)
Vietnam	ReUse with raw bands	42.0 ± 6.6	57.7 ± 7.3	0.4 ± 0.2
	ReUse with feature extraction	44.4 ± 6.0	59.5 ± 4.7	0.4 ± 0.2
	Competitor 1 [10]	60.1 ± 8.3	73.0 ± 9.4	0.2 ± 0.2
	Competitor 2 [14]	58.9 ± 8.6	72.0 ± 9.7	0.2 ± 0.2
Myanmar	ReUse with raw bands	10.8 ± 2.0	15.0 ± 2.4	0.7 ± 0.1
	ReUse with feature extraction	10.7 ± 2.2	14.9 ± 2.6	0.7 ± 0.1
	Competitor 1 [10]	15.7 ± 1.9	20.2 ± 2.3	0.4 ± 0.1
	Competitor 2 [14]	15.5 ± 1.5	20.1 ± 1.8	0.4 ± 0.1
Europe	ReUse with raw bands	24.5 ± 3.3	46.6 ± 5.2	0.6 ± 0.1
	ReUse with feature extraction	24.1 ± 3.4	46.9 ± 4.2	0.6 ± 0.1
	Competitor 1 [10]	32.5 ± 3.1	48.0 ± 4.4	0.5 ± 0.5
	Competitor 2 [14]	34.8 ± 3.1	51.1 ± 3.9	0.5 ± 0.5

4.1. Case Study: Astroni Nature Reserve

A case study for the Astroni nature reserve in southern Italy was presented to show how ReUse, Sentinel-2, and ESA's AGB public data can help estimate CO₂ in forest areas and monitor deforestation downstream of events such as fires. ReUse is adopted with raw Sentinel-2 bands alone without using other extracted features that would not bring decisive benefits, as demonstrated in the previous section.

The area of central Europe contained in the file 'N60E00' was chosen to monitor the Astroni reserve to train ReUse because it is certainly an area with characteristics closer to those of southern Italy than the other two datasets containing the areas of Vietnam and Myanmar. In [38], research was conducted within the Astroni Crater World Wildlife Fund (WWF) Reserve in the volcanic area of the Campi Flegrei in the urban area of Naples, Italy. The Reserve (247 hectares) lies within the caldera of an extinct volcano with a maximum altitude of 255 m above sea level and an elliptical shape (2 × 1.6 km). The inner part of the crater has a deep depression containing a lake, where a minimum altitude of 9 m above sea level is reached. In the central part of the crater, near the largest lake, two other small lakes and three hills rise to the bottom at 45, 74, and 82 m above sea level, respectively. Throughout the crater, [38] focused on the area of holm oak forest (127 ha) and the area of mixed forest (104 ha) to define the two main ecosystems and found that the total carbon stocks of the phytomass of these two ecosystems were 22173 ± 7054 tonnes using sampling from April to October 2016.

In order to make inferences, the Sentinel-2 images downloaded on 31 May 2017 were upscaled to a spatial resolution of 10 meters, assuming that the spatial correlations learned from the network at 100 meters are also reproducible at 10 meters; this improves the resolution of the predictions compared to the resolution of the AGB raster of the ESA CCI Biomass project which is at 100 meters. The predictions of Carbon stocks were made by creating the predicted raster of AGB ten by ten non-overlapping patches, corresponding to one hectare each. Then, for each patch, the average AGB value expressed in tonnes per hectare was taken, which, when multiplied by 1 hectare, which is the extent of the patch, yields a value in tonnes. These values were summed over all patches, and the final result multiplied

by 0.5 [4,5] to obtain the value of absorbed carbon in tonnes. From this procedure, the estimate of the Carbon stock for Astroni on 31 May 2017 was 18748 tons, in line with the forecast of [38], which is 22173 \pm 7054 tons for phytomass in the year 2016. Please note that the latter includes the roots of the plants, which is not included in our estimate, but which can be considered around 30% of above-ground biomass for temperate oak forest [39]. Thus, considering only the $0.3*18748 + 18748$ estimates, we obtain a final value of 24372 tons which is within the range of Astroni's estimate of 22173 \pm 7054 tons. In order to confirm the fact that the trained network recognizes a decrease in AGB downstream of a fire, the AGB raster of Astroni on 24 August 2017 downstream of the fire on July 2017 is shown in Figure 7 compared with the AGB raster of the same area before the fire. The estimated above-ground biomass carbon stock on 24 August 2017 for the nature reserve downstream of the fire on July 2017 is 10104 tonnes, confirming the above.

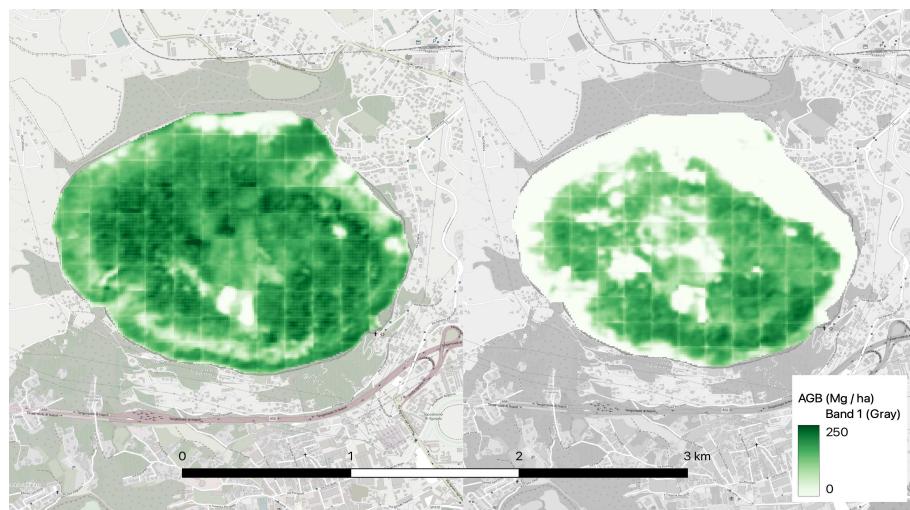


Figure 7. On the left is the predicted above-ground biomass raster of the Astroni nature reserve before the July 2017 fire; on the right is the predicted above-ground biomass raster after a major fire event for the same area.

It is stressed that the work done is based entirely on open data, and the fact that a prediction based on public data is in agreement with ranges given by a ground truth gives worth to a tool that can be obtained without field measurements and that could be useful for monitoring carbon stocks in forest areas. The prediction after the fire cannot be verified in any way; however, the fact that the prediction before the fire is in line with the ground truth and that the forecast after this event shows a decrease due to the fire is encouraging.

5. Discussion with Conclusions

Sustainable Development Goal 15 aims to 'protect, restore and promote the sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, halt and reverse land degradation and halt biodiversity loss' [3]. Furthermore, REDD+ projects aim to reduce Greenhouse Gases (GHG) concentrations in the atmosphere and contribute to climate change mitigation through various activities, including carbon stock enhancement [40]. Developing systems that can estimate the carbon absorbed by forests globally and monitor losses associated with deforestation phenomena such as fires on time is essential. As explained in Section 1, systems that rely on field measurements are the most reliable. However, this work aimed to show how it is possible to estimate the carbon absorbed by forests and nature reserves such as Astroni using open AGB data from the ESA CCI Biomass Project in conjunction with Sentinel-2 images. To the best of our knowledge, this is the first time such a study has been done by totally relying on open data only, without field measurements and using a deep approach based on a UNet architecture.

The proposed ReUse architecture is based on a pixel-wise regressive UNet, able to generate a pixel mask of AGB predictions with computational advantages, particularly when monitoring large areas. This is a great advantage over classical machine learning algorithms that require feature extraction work to derive indices that capture both spectral and spatial information content [10,14] or over solutions based on convolutional neural network approaches to estimate AGB using the commercial Worldview-2 satellite and visible spectrum images captured by an unmanned aerial vehicle [15,16] which produce a single value as a prediction of AGB. The computational advantage of UNet over simple CNNs lies in the fact that with CNNs, for each input pixel, its neighbourhood and associated bands are exploited to produce a single prediction of AGB, whereas with UNet, a patch of input pixels is associated with a patch of output pixels (e.g., for 16×16 pixels patches as in our case, 16×16 inferences with a simple CNN would be equivalent to a single inference with our UNet architecture).

ReUse, both when using raw bands only and raw bands in combination with spectral indices and texture features, showed better cross-validation performance in terms of MAE, RMSE, and R^2 against the two considered competitors [10,14] on three different areas (Table 2). It should also be noted that using a ReUse version leveraging raw bands together with some features does not lead to substantial advantages in terms of result accuracy compared to ReUse with raw bands alone, thus demonstrating that this approach can dispense with the feature engineering phase and work directly with the Sentinel-2 raw bands appropriately normalized. When trained on the central European zone, the network's predictions were also validated on Astroni, a WWF reserve located close to the metropolitan area of Naples in southern Italy. The predicted Carbon storage of above-ground biomass of Astroni, prior to the July 2017 fire, is 18748 tonnes in line with the estimates reported in [38]. In addition, an estimate of carbon storage of above-ground biomass of Astroni was also made on 24 August 2017 after a major fire event, resulting in 10104 tonnes. This last forecast cannot be verified in any way; however, the fact that the first prediction is in line with the ground truth and the second forecast shows a decrease due to the fire is encouraging.

In conclusion, the combined use of Sentinel-2 data and ESA's AGB data with a UNet approach could be suitable for estimating the carbon absorbed in urban and rural areas and help monitor deforestation events without field measurements. To this aim, we have released the code to be used as a monitoring tool for deforestation. Future research directions could concern the use of multi-temporal Sentinel-2 data in which spectral images of several dates are examined, the joint use of Sentinel-2 and Sentinel-1 as input for ReUse, the application of super-resolution to improve the estimations of AGB and carbon stock in particular in urban areas. We will also explore other case studies to further verify how well the AGB data of ESA's Climate Change Initiative Biomass project is suitable as ground truth for AI-based models for forest monitoring.

Author Contributions: Conceptualization, A.E.P., G.G., S.M. and C.S.; methodology, A.E.P., G.G. and S.M.; software, A.E.P.; validation, A.E.P., G.G., M.R., S.M. and C.S.; formal analysis, A.E.P., G.G. and S.M.; investigation, A.E.P., G.G., M.R., S.M. and C.S.; resources, C.S.; data curation, A.E.P., G.G., M.R. and S.M.; writing—original draft preparation, A.E.P., G.G. and S.M.; writing—review and editing, A.E.P., G.G., S.M. and C.S.; visualization, A.E.P., G.G. and S.M.; supervision, S.M. and C.S.; project administration, S.M. and C.S. . All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Sentinel-2 L2A data are available at <https://scihub.copernicus.eu/> and ESA's AGB data are available at <https://climate.esa.int/en/projects/biomass/>.

Acknowledgments: Special thanks go to © ESA Climate Change Initiative - Project BIOMASS 2021 for placing the AGB data in the public domain. We thank Latitudo 40 for providing the infrastructure to download Sentinel-2 images.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Curiel-Esparza, J.; Gonzalez-Utrillas, N.; Canto-Perello, J.; Martin-Utrillas, M. Integrating climate change criteria in reforestation projects using a hybrid decision-support system. *Environmental Research Letters* **2015**, *10*, 094022.
2. Deo, R.K.; Russell, M.B.; Domke, G.M.; Andersen, H.E.; Cohen, W.B.; Woodall, C.W. Evaluating site-specific and generic spatial models of aboveground forest biomass based on Landsat time-series and LiDAR strip samples in the Eastern USA. *Remote Sensing* **2017**, *9*, 598.
3. United Nations. Sustainable Development Goals. Available online: <https://sdgs.un.org/goals>.
4. Tooqi, E.C. Carbon sequestration: How much can forestry sequester CO₂. *Forestry Research and Engineering: International Journal* **2018**, *2*, 148–150.
5. Birdsey, R.A. *Carbon storage and accumulation in United States forest ecosystems*; Vol. 59, US Department of Agriculture, Forest Service, 1992.
6. Lu, D. The potential and challenge of remote sensing-based biomass estimation. *International journal of remote sensing* **2006**, *27*, 1297–1328.
7. Lu, D.; Chen, Q.; Wang, G.; Liu, L.; Li, G.; Moran, E. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth* **2016**, *9*, 63–105.
8. Aschbacher, J.; Milagro-Pérez, M.P. The European Earth monitoring (GMES) programme: Status and perspectives. *Remote Sensing of Environment* **2012**, *120*, 3–8.
9. Shoko, C.; Mutanga, O. Examining the strength of the newly-launched Sentinel 2 MSI sensor in detecting and discriminating subtle differences between C3 and C4 grass species. *ISPRS journal of photogrammetry and remote sensing* **2017**, *129*, 32–40.
10. Dang, A.T.N.; Nandy, S.; Srinet, R.; Luong, N.V.; Ghosh, S.; Kumar, A.S. Forest aboveground biomass estimation using machine learning regression algorithm in Yok Don National Park, Vietnam. *Ecological Informatics* **2019**, *50*, 24–32.
11. Moradi, F.; Sadeghi, S.M.M.; Heidarlu, H.B.; Deljouei, A.; Boshkar, E.; Borz, S.A. Above-ground biomass estimation in a Mediterranean sparse coppice oak forest using Sentinel-2 data. *Annals of Forest Research* **2022**, *65*, 165–182.
12. Pandit, S.; Tsuyuki, S.; Dube, T. Estimating above-ground biomass in sub-tropical buffer zone community forests, Nepal, using Sentinel 2 data. *Remote Sensing* **2018**, *10*, 601.
13. Mngadi, M.; Odindi, J.; Mutanga, O. The utility of Sentinel-2 spectral data in quantifying above-ground carbon stock in an urban reforested landscape. *Remote Sensing* **2021**, *13*, 4281.
14. Wai, P.; Su, H.; Li, M. Estimating Aboveground Biomass of Two Different Forest Types in Myanmar from Sentinel-2 Data with Machine Learning and Geostatistical Algorithms. *Remote Sensing* **2022**, *14*, 2146.
15. Dong, L.; Du, H.; Han, N.; Li, X.; Zhu, D.; Mao, F.; Zhang, M.; Zheng, J.; Liu, H.; Huang, Z.; others. Application of convolutional neural network on lei bamboo above-ground-biomass (AGB) estimation using Worldview-2. *Remote Sensing* **2020**, *12*, 958.
16. Schreiber, L.V.; Atkinson Amorim, J.G.; Guimarães, L.; Motta Matos, D.; Maciel da Costa, C.; Parraga, A. Above-ground Biomass Wheat Estimation: Deep Learning with UAV-based RGB Images. *Applied Artificial Intelligence* **2022**, *36*, 2055392.
17. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics* **1973**, pp. 610–621.
18. Moradi, F.; Darvishsefat, A.A.; Pourrahmati, M.R.; Deljouei, A.; Borz, S.A. Estimating aboveground biomass in dense Hyrcanian forests by the use of Sentinel-2 data. *Forests* **2022**, *13*, 104.
19. Li, C.; Zhou, L.; Xu, W. Estimating aboveground biomass using Sentinel-2 MSI data and ensemble algorithms for grassland in the Shengjin Lake Wetland, China. *Remote Sensing* **2021**, *13*, 1595.
20. Pandit, S.; Tsuyuki, S.; Dube, T. Exploring the inclusion of Sentinel-2 MSI texture metrics in above-ground biomass estimation in the community forest of Nepal. *Geocarto International* **2020**, *35*, 1832–1849.
21. Naik, P.; Dalponte, M.; Bruzzone, L. Generative Feature Extraction From Sentinel 1 and 2 Data for Prediction of Forest Aboveground Biomass in the Italian Alps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2022**, *15*, 4755–4771.

22. Naik, P.; Dalponte, M.; Bruzzone, L. Automated Machine Learning Driven Stacked Ensemble Modelling for Forest Aboveground Biomass Prediction Using Multitemporal Sentinel-2 Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2022**.
23. Li, L.; Zhou, B.; Liu, Y.; Wu, Y.; Tang, J.; Xu, W.; Wang, L.; Ou, G. Reduction in Uncertainty in Forest Aboveground Biomass Estimation Using Sentinel-2 Images: A Case Study of *Pinus densata* Forests in Shangri-La City, China. *Remote Sensing* **2023**, *15*, 559.
24. Song, P.; Kim, G.; Mayer, A.; He, R.; Tian, G. Assessing the ecosystem services of various types of urban green spaces based on i-Tree Eco. *Sustainability* **2020**, *12*, 1630.
25. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.
26. Fazzini, P.; De Felice Proia, G.; Adamo, M.; Blonda, P.; Petracchini, F.; Forte, L.; Tarantino, C. Sentinel-2 Remote Sensed Image Classification with Patchwise Trained ConvNets for Grassland Habitat Discrimination. *Remote Sensing* **2021**, *13*, 2276.
27. Breiman, L. Random forests. *Machine learning* **2001**, *45*, 5–32.
28. Jolliffe, I.T.; Cadima, J. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2016**, *374*, 20150202.
29. Banskota, A.; Wynne, R.H.; Kayastha, N. Improving within-genus tree species discrimination using the discrete wavelet transform applied to airborne hyperspectral data. *International Journal of Remote Sensing* **2011**, *32*, 3551–3563.
30. Prechelt, L. Early stopping—but when? In *Neural Networks: Tricks of the trade*; Springer, 1998; pp. 55–69.
31. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* **2014**.
32. Keras: ReduceLROnPlateau. Available online: https://keras.io/api/callbacks/reduce_lr_on_plateau/.
33. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An introduction to statistical learning*; Vol. 112, Springer, 2013.
34. Korhonen, L.; Packalen, P.; Rautiainen, M.; others. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. *Remote sensing of environment* **2017**, *195*, 259–274.
35. Sentinel-2 Products Specification Document. Available online: <https://sentinel.esa.int/documents/247904/685211/sentinel-2-products-specification-document>.
36. ESA's Climate Change Initiative Biomass project: Global datasets of forest above-ground biomass. Available online: <https://climate.esa.int/en/projects/biomass/data/>.
37. Santoro, M.; Cartus, O.; Carvalhais, N.; Rozendaal, D.; Avitabile, V.; Araza, A.; De Bruin, S.; Herold, M.; Quegan, S.; Rodríguez-Veiga, P.; others. The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations. *Earth System Science Data* **2021**, *13*, 3927–3950.
38. Innangi, M.; Balestrieri, R.; Danise, T.; d'Alessandro, F.; Fioretto, A. From soil to bird community: A Partial Least Squares approach to investigate a natural wooded area surrounded by urban patchwork (Astroni crater, southern Italy). *Ecological Modelling* **2019**, *394*, 1–10.
39. Mokany, K.; Raison, R.J.; Prokushkin, A.S. Critical analysis of root: Shoot ratios in terrestrial biomes. *Global change biology* **2006**, *12*, 84–96.
40. United Nations. What is REDD+? Available online: <https://unfccc.int/topics/land-use/workstreams/redd/what-is-redd>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.