Pak. j. sci. ind. res. Ser. A: phys. sci. 2021 **64A**(2) 119-130

Integration of Sentinel-2 Derived Spectral Indices and *In-situ* Forest Inventory to Predict Forest Biomass

Areeba-Binte-Imran, Samia Ahmed, Waqar Ahmed, Muhammad Zia-ur-Rehman, Arif Iqbal, Naveed Ahmad* and Irfan Ullah

Department of Forestry and Range Management, PMAS Arid Agriculture University, Rawalpindi-46300, Pakistan

(received April 8, 2019; revised June 14, 2020; accepted July 7, 2020)

Abstract. Forest biomass estimation is the central part of sustainable forest management to assess carbon stocks and carbon emissions from forest ecosystem. Sentinel-2 is state-of-art sensor with refined spatial and recurrent temporal resolution data. The present study explored the potential of Sentinel-2 derived vegetation indices for above ground biomass prediction using four regression models (linear, exponential, power and logarithmic). Sentinel-2 indices includes Global environmental monitoring index, transformed normalized difference vegetation index, normalized difference water index, normalized difference infrared index and red-edge normalized difference vegetation index. The performances of Sentinel-2 indices were assessed by simple single variable (index) based regression for GEMI, TNDVI, NDII, NDWI and RENDVI versus AGB values. Further, stepwise linear regression was also developed in which used all indices entered into stepwise selection and the best index was selected in the final model. Results showed that linear model of all indices performance best compared to the rest three models and R^2 values 0.12, 0.39, 0.46, 0.44 and 0.37 for Global environmental monitoring index, transformed normalized. Vegetation index, normalized difference water index, infrared index and red-edge vegetation index, respectively. Normalized difference water index was considered the best index among five computed indices in simple linear as well as in stepwise linear regression, whereas rest of the indices were removed because they were not significant under the stepwise criteria. Further, the accuracy of normalized difference water index model was determined by root mean square error and final prediction model has 28.27 t/ha error for both simple linear and stepwise linear regression. Therefore, normalized difference water index was selected for biomass mapping and resultant biomass showed up to 339 t/ha in the study area. The resultant biomass map also showed consistency with global datasets which include global forest canopy height and global forest tree cover change maps. The study suggest that Sentinel-2 product has great potential to estimate above ground biomass with accuracy and can be used for large scale mapping in combination with national forest inventory for carbon emission accounting.

Keywords: normalized difference water index (NDWI), transformed normalized difference vegetation index (TNDVI), normalized difference infrared index (NDII), red-edge normalized difference vegetation index (RENDVI)

Introduction

Forest biomass can be estimated by two basic approaches, the first one is *in-situ* forest inventory method and the second the proxy estimation through remote sensing (RS) techniques. The former one (forest inventory) is no doubt has more accuracy (Lu *et al*., 2016) but these methods have many limitations such as they are time consuming, require labour, implementation become difficult for large areas and have spatio-temporal constrains (Lu *et al*., 2016; Henry *et al*., 2011). Traditional inventories are also destructive in nature for developing proper allometric equations. On the other hand, remote sensing techniques are favoured because, it can minimize the aforesaid limitations of *in-situ* forest inventory (Shi, 2010). Remote sensing data can provide wall-to-wall coverage of large forest areas in both spatial and temporal context. Although remote sensing provide spectral information from space; large distances from the earth (forest resource) thus it provide only proxy estimation. That's why remote sensing methods required *in-situ* forest inventory data for reliable estimation. Biomass prediction and mapping can only be accurate and precise when field attributes are combined with spectrally-derived information (indices) from sensor reflectance (bands) (Pandit *et al*., 2018). Thus integration of remote sensing and field inventory can be improve the accuracy and biomass can estimated for large areas at various scales and time. Numerous studies are used such integration of various remote sensing data (Landsat, LiDAR, hyperspectral) and field sample plots data for

*Author for correspondence;
E-mail: naveedahmad795@gmail.com

above ground biomass (AGB) estimation (Ali *et al*., 2018; Imran and Ahmed, 2018; Shen *et al*., 2016; Dube *et al*., 2014; Rana *et al*., 2014; Chen *et al*., 2009). Among remote sensing data (Landsat, LiDAR, Radar, hyperspectral) the image cost, information volume, data redundancy, storage and data processing costs limit data selection and use for AGB estimation (Plaza *et al*., 2009). Mathieu *et al*. (2013) discussed these limitations of using various remote sensing technologies and compared it with other medium-resolution data. Therefore, forest managers have shifted towards cost effective and freely available open source broadband remote sensing data including Land sat, MODIS, ASTER and other recently launched Sentinel-2 (Sibanda *et al*., 2015). European Space Agency (ESA) has launched Sentinel-2 (multi-spectral instrument, MSI) sensor, which provide improved spatial resolution and spectral bands combination and high temporal frequency compared to Land sat sensors (Drusch *et al*., 2012). Sentinel-2 sensor comprised of two satellites (Sentinel-2A and Sentinel-2B), offering 13 spectral band combinations starting from visible followed by near-infrared (NIR) to short wave infra-red (SWIR) and these are available with improved resolution of (10, 20 m) and 60 m respectively (Wittke *et al*., 2019). Moreover, the addition of three red-edge bands which appear immediately after Red band, provide the Sentinel-2 product high potential for mapping various vegetation characteristics. These Red-edge bands offer great improvement for mapping various vegetation characteristics (Cao *et al*., 2016). Large swath width offered by Sentinel-2 allow forest managers to quantify AGB from local to regional scale and carbon accounting can be timely updated (Majasalmi and Rautiainen, 2016; Frampton *et al*., 2013). Such band combination of Sentinel-2 make it analogous to some marketable satellite data, such as world view-2, thus Sentinel-2 product has high potential for applications in forest resource (AGB and carbon stocks estimation), land management, food security and precision agriculture, disaster management and humanitarian relief operations (Chang and Shoshany, 2016). Forest attributes and growth parameters can be efficiently quantified by using Band 5, Band 6 and Band 7 of Sentinel-2 product. Ramoelo *et al*. (2015) successfully estimated grass nutrients by exploring the potential of Band 5, Band 6 and Band 7 of Sentinel-2. Similarly, comparison of Land sat-8 and Sentinel-2 products are studied by Forkuor *et al*. (2018) and land use land cover change was studied in Burkino Faso and reported that Sentinel-2 image has produced better results than open source Land sat 8 sensor. Chen *et al*. (2018) collected AGB data from *in-situ* forestry (n=56) in China and modeled AGB with applying regression models and machine learning techniques. Nuthammachot *et al*. (2018) assessed the potential of various vegetation indices computed from Sentinel-2 indices and estimated AGB. The study used forest inventory data of forty five sample plots and seven Sentinel-2 image based vegetation indices. Chen *et al*. (2019) studied AGB for Bamboo plantations using Sentinel-2 image in Zhejiang Province, China and reported that Red-edge bands and NIR bands in Sentinel-2 have improved identification of key AGB variables. Navarro *et al*. (2019) integrated Unmanned Aerial Vehicle (UAV), Sentinel-1, and Sentinel-2 data for AGB estimation in Mangroves plantation using 95 ground sample plots. The study used support vector regression models and demonstrated that combination of UAV-based estimates with Sentinel-1 and Sentinel-2 has provided accurate and cost-effective AGB estimates. Chrysafis *et al*. (2019) assessed growing stock volume of Mediterranean forest using the spectral and spatial information of Sentinel-2. The study developed regression models for growing stock volume using spectral bands and other derived spectral and spatial features from Sentinel-2 images. Astola *et al*. (2019) discussed comparison of Landsat-8 and Sentinel-2 for estimating different forest biophysical variables which included stem diameter and volume, height of the tree and tree cross-sectional area, and specie-specific components of pine, spruce and broadleaved associates. The study concluded that Sentinel-2 performed better than Landsat-8 because of its finer spatial resolution and Red-edge bands presence. Ali *et al*. (2018) also used Sentinel-2 image for biomass estimation in Sub-tropical Scrub forests in Pakistan and compared it with Land sat-8. The study concluded that Sentinel-2 has better potential for biomass estimation as compared to Landsat-8 images. The present study integrated vegetation indices derived from Sentinel-2 image (including Red-edge band) and field biomass AGB data for biomass estimation. The study was conducted in sub-tropical chirpine forest in Shinkiarii, Pakistan in 2018.

Materials and Methods

Study area. The Shinkiarii forests situated in Mansehra district of Khyber Pukhtunkhwa province of Pakistan and are situated between 34°-15' to 34°-38' N and 72°-20' E. The Shinkiarii forests surrounded by three ranges i.e. lower Siran, Mansehra, and Hilkot ranges (range is

a territorial unit of forest). The highest elevation of study area reaches 7579 feet (2297 meters) at Siri above mean sea level, where as the lowest elevation falls down to 1835' feet at Kalinger area. The topography of forest covered area is mostly composed of moderate to steep slopes, while the non-forest area is mostly moderate to flatter in its nature. The total area covered by Shinkiarii forests are 3408 hectares which comprised of various compartments and the area has great potential for tourists due to its hygienic climate and fascinating green belt hillside. A number of different species of Wildlife i.e. monkeys; jackals and birds like partridges etc. are more common. The main forest species are *Pinus roxburghii* (chir) with scattered associates of *Quercus incana* (oak) and other broadleaved like *Alnus nitida* (spach), *Morus alba* (mulberry), *Poplus ciliata* (Poplar) and *Robinia pseudoacacia*. In the some areas of community forests there is abundant pole crop of *Pinus roxburghii*, whereas mature trees have severely been depleted because of anthropogenic pressure. The tract is mountainous; therefore, the climate is sub-tropical. Generally the summers are severe at foot fills, while higher up these are mild and pleasant. The entire tract is within the reach of summer Monsoons, which generally start early in July and continue up to mid September, when nearly half of the annual precipitation is received. Another wet spell is experienced in the winter months. i.e. from December to February and during this season snowfall is frequent particularly above 5000 feet.

Forest inventory. For the present research random sampling strategy was used and basic statistical principles (of coverage and representation) were observed. The area including of pure Chirpine forest (*Pinus roxburghii*) with small percentage of *Quercus incana* in depressions. Forest inventory comprises 20 samples plots were taken randomly laid out within forests. The size of each sample plot was 0.1 ha ($1/10^{th}$ of ha) having 17.84 meters radius as per departmental protocols (Ali *et al*., 2018; Imran and Ahmed, 2018; Nizami, 2012). Different field equipment's were used during data collection sampling which include diameter Calliper, Haga Altimeter, ranging rods, measuring tape and Sunnto Compass were used. Geographical position of each sampling was recorded by Garmin Global Positioning System (Garmin GPS 30X). Diameter of each tree in every plot was measured at 1.3 meters (4.5 feet) which is conventional measuring point in forestry. Haga Altimeter was used for height of the tree while considering various situations (such as leening tree, sloping ground, invisible top etc.). The general eq. 1 used by Haga Altimter, (Ali *et al*., 2018) is as follows.

$$H = (TanA_1 + TanA_2) \times d \quad \text{..... (1)}$$

where:

$TanA_1$ is the angle of tree top; $TanA_2$ is the angle of tree base, "d" is distance from the observer to tree.

Further, volume was calculated from plot wise diameter and height data. Tree basal area was calculated at 1.3 m above the ground which is actually the cross-sectional area at DBH point. Tree basal area is measured in square meters which was simply determined using area of a circle given in eq. 2.

$$\text{Tree basal area } (m^2) = (DBH/200)^2 \times 3.142 \quad \text{..... (2)}$$

where:

DBH is the diameter at breast height (cm) and 3.142 is π.

The tree basal area was converted into stem volume (m^3) by using height of tree in meter and tree form factor. The volume (m^3) was simply obtained by multiplying basal area with height and form factor. Form factor was conical with value 0.33 (one-third of cylindrical shape) which was obtained from local volume table entitled "local volume table of sub-tropical Chirpine forests Shinkiarii". The above ground biomass was obtained by using volume (m^3), basic wood density (BWD) and biomass expansion factor (BEF) using the formula given in eq. 3. Both BWD and BEF were obtained from available literature (Ali *et al*., 2018). Biomass expansion factor is actually conversion factor

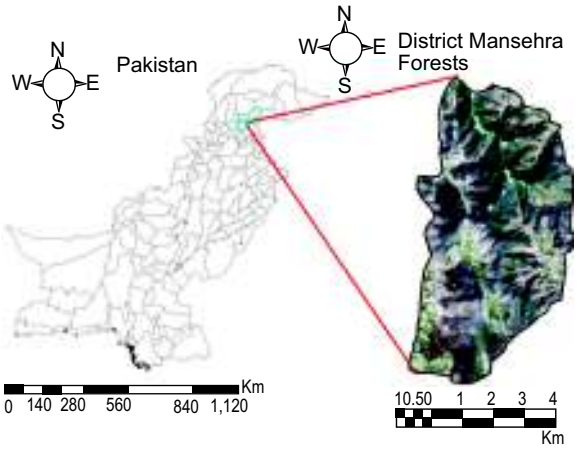

Fig. 1. Geographic location of the study area in district Mansehra, Pakistan.

which expand stem volume (derived from DHB and height of stem) to tree volume (extended lateral branches of the whole tree). BEF was actually used to expand volume or biomass estimates of merchantable parts (of stem or main trunk) to non-merchantable parts (branches of the tree). BEF value of 1.5 was a generic value derived as per FAO guidelines for sub-tropical cChirpine forests.

Above ground biomass (Kg)=V×BWD×BEF ... (3)

where:

V is the volume in m^3; BWD is the Basic Wood Density in Kilograms per cubic meter; BEF Biomass expansion factor (1.5 for this study).

The volume per hectare values were obtained by multiplying, it with plot expansion factor, (PEF) which is 10 (because plot size 0.1 ha). This conversion factor has been globally used by (Ali *et al*., 2018; Nizami, 2012). The AGB (t/ha) values were then integrated with spectral values of Sentinel-2 for regression model development.

Sentinel-2 image acquisition and processing. The Sentinel-2A Level-1C (L1C) product (Single tile) acquired was downloaded from The Copernicus open access hub. The Sentinel-2 images was acquired for (path: 150 and row: 36) with minimum cloud cover and with nearest possible date close to the field survey. The Sentinel-2 product was an ortho-corrected image in UTM/WGS 84, 43N projection and having top-of-atmosphere (TOA) reflectance values (per-pixel radiometric values). The Sentinel-2 image was transformed from sensor radiance to surface reflectance by applying the bottom-of-atmospheric (BOA) correction and dark object subtraction (DOS) method using Sentinel application toolbox (SNAP) software (version 5.0). The SNAP 5.0 used third party plugin named as Sen2Cor for the preprocessing of Sentinel-2 image. The atmospheric correction was completed through Sen2Cor plugin which was installed under the umbrella of Anaconda Python library in SNAP 5.0. Pre-processing also removed the atmospheric scattering and dark pixels in spectral bands during image processing. The Sentinel-2 image has thirteen spectral bands with three types spatial resolutions (10, 20, and 60 m) extended from visible to infrared wavelengths (Ali *et al*., 2018). In this research, Band 4 (Red), Band 5 (Red-edge), Band 8 (NIR), Band 12 (SWIR) and Band 13 (SWIR) were used for vegetation analysis. Moreover, all bands with lower resolution (Band 6, Band 12, 13) were resampled to ten m resolution using SNAP 5.0. As original pixel size of Sentinel-2 bands was 10 m, whereas field plot size was 0.1 hectar therefore, it was necessary that first field sample plots and pixel size should be brought to same size. Mean filter (window of 3 by 3) was applied to all bands (of spectral indices) to make it equivalent to field plot size. Further, the image subset was created to minimize the time of computation. Based on previous literature (Ali *et al*., 2018; Adan, 2017), vegetation indices were selected for their performance in forest bio-physical studies. Five vegetation indices (VIs) were selected and computed in SNAP 5.0 using vegetation radiometric indices tool. These VIs include broadband indices, canopy water indices and Red-edge band index. The broadband VIs were Global Environmental Monitoring Index (GEMI) and Transformed normalized difference vegetation index (TNDVI). Canopy water contents VIs includes normalized difference infrared index (NDII) and normalized difference water index (NDWI). Whereas Red-edge (RE) band index was RE normalized difference vegetation index (RENDVI) was also computed to study the impact of RE band (band 6). The details formulas, bands used and their sources of all VIs are given in the Table 1.

Table 1. Sentinel-2 derived vegetation indices

Indices	Landsat-8	Reference study
Transformed normalized difference vegetation	sqrt((NIR–R/NIR+R)+0.5)	Nouri *et al*., 2018
~~index (TNDVI) Global environment monitoring~~	~~ε(1-0.25 ε) (RED-0.125) / (1-RED)~~	~~Schultz *et al*., 2016~~
index (GEMI)	where: $\varepsilon = (2(NIR^2 - RED^2) + 1.5NIR + 0.5 *RED)/(NIR+RED+0.5)$	
Normalized difference water index (NDWI)	ρNIR - ρB11/ NIR + ρB11	Adan, 2017
Normalized difference infrared index (NDII)	ρNIR - ρB12/ NIR + ρB12	Hunt *el al*., 2012
Red-edge normalized difference vegetation index (RENDVI)	(NIR-RE)/(NIR+RE)	Adan, 2017

Statistical analysis. The relationships of field data and performance of VIs were assessed through regression analysis. The dependent variable was AGB (t/ha) and independent variables were Sentinel-2 image spectral indices. First, simple single variable (index) based regression were developed for each of the computed indices (GEMI, TNDVI, NDII, NDWI and RENDVI) versus AGB values and four regression models (linear, power, exponential and logarithmic) were evaluated. Secondly, stepwise linear regression which used all indices (as input variables) entered into stepwise selection and as a result the best index was selected in the final model. These models were developed in satistical program for social survey (SPSS) software. The model selection and performance were based on coefficient of determination (eq. 4) and level of significance (P-value). Accuracy assessment was determined by the root mean square error (RMSE), BIAS and mean absolute percentage error (MAPE) given in eq. 5, 6, 7, respectively. The statistical relationship was consi-dered significant if it's P-value is less than 0.05.

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}} \quad \text{..... (4)}$$

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(AGB1 - AGB2)} \quad \text{..... (5)}$$

$$BIAS = \frac{1}{n}\Sigma_{i=1}^{n}(AGB1 - AGB2) \quad \text{..... (6)}$$

$$MAPE = \frac{100\%}{n}\Sigma_{i=1}^{n}(AGB1 - AGB2 \div AGB1) \quad \text{..... (7)}$$

where:

r = coefficient of correlation; n = number of inventory plots; Σxy = the sum of the products of AGB and VI; Σx = total of VI values; Σy = total of AGB values; Σx^2 = total of squared VI values; Σy^2 = total of squared AGB values; AGB1 was the AGB value from field inventory; AGB2 was the predicted value of biomass by Sentinel-2 VI.

Results and Discussion

Computation of vegetation indices. The vegetation indices (VIs) were calculated using optical thematic land processing tool in SNAP 5.0 software. All five indices and computation process were explained in methodology. These indices include global environmental monitoring index (GEMI), transformed normalized difference vegetation index (TNDVI), normalized difference water index (NDWI), normalized difference infrared index (NDII) and Red-edge normalized difference vegetation index (RENDVI). The broadband indices (GEMI and TNDVI) and canopy water indices (NDWI and NDII) are sensitive to higher biomass density and have saturation issue, whereas narrow band index (RENDVI) have overcome this saturation issue in higher biomass density. Lu *et al*. (2016) discussed that accuracy of biomass estimation may have various uncertainties due to saturation problems of VIs. Adan, (2017) also reported that narrowband indices have greater efficiency to differentiate higher biomass values, while on the other hand broadband indices have saturation issues. Frampton *et al*. (2013) demonstrated that Sentinel-2 has comparatively higher resolution and the presence of Red-edge bands was very useful for AGB estimation, forest monitoring and bio-physical parameters estimation.

Single predictor regression models (Sentinel-2 indices)

Global environmental monitoring index (GEMI). GEMI was regressed against AGB and four different regression models were developed including linear, power, exponential and logarithmic model (Table 2). According to Table 2, GEMI has shown much less correlation compared to other indices. The linear model and logarithmic models have explained 12% of data variations while, the rest of 88% data variations were not explained by GEMI models. Similarly, the lower correlations were observed in other non-linear models (power and exponential) with the R^2 of 0.113. Moreover, all the models of GEMI have not performance significantly as shown in Table 2. The GEMI is broadband index using Band 4 (Red) and Band 5 (NIR) and have saturation issue in higher density area as shown in Fig. 2. Overall, the performance of GEMI was not good and hence did not as act best model for biomass estimation. Wijaya *et al*. (2010) evaluated various vegetation indices for estimation of stem volume and above ground biomass using moderate resolution data. The study reported that GEMI performed better having correlation (R^2) of 0.368 and 0.313 with stem volume and AGB, respectively. Lu *et al*. (2004) used GEMI for various forest attributes estimation and reported that GEMI has good correlation (R^2 = 0.48) with AGB. Gupta *et al*. (2018) demonstrated that mangroves

vegetation (coastal) can be best modeled by using two VIs (Modified Soil Adjusted Vegetation Index (MSAVI) and GEMI). The study also reported that monitoring was best by using these two indices for Tropical Littoral forests.

Transformed normalized difference vegetation index (TNDVI). Four different regression models were developed including linear, power, exponential and logarithmic model for TNDVI (Table 2). According to Table 2, linear model of TNDVI has showed the highest correlation (R^2= 0.392) followed by logarithmic model with coefficient of correlation of 0.386, whereas power and exponential models have lower coefficient of correlation of 0.289 and 0.294 respectively. Compared to GEMI, TNDVI has good relationship with AGB data as linear model of TNDVI has explained 39 % of the data variations. All regression models of TNDVI have showed significant relationships with AGB as shown in Table 2. However, TNDVI have less correlation as compared to canopy water indices (NDWI and NDII). Therefore, based on performance (R^2 values), TNDVI was not selected as best model for AGB estimation but nevertheless it may considered as vegetation index for AGB estimation (Fig. 3). Nouri *et al*. (2018) used TNDVI and other indices to estimates various forest attributes (density of forest, canopy closure and basal area) and reported correlation of 0.90 for TNDVI and NDVI. Ahmad (2012) also used TNDVI for quantitatively evaluation of dense and sparse vegetation using Landsat ETM+ in Cholistan Desert. Chen *et al*. (2018) has estimated AGB in Changbai Mountains, China and used various predictive for biomass estimation. The study estimated AGB by combined used of Sentinel-1, Sentinel-2 and Digital Elevation Model (DEM) and showed that TNDVI was one of the important vegetation index derived from Sentinel-2 index. Alrababah *et al*. (2011) estimated forest parameters (crown percentage and above ground biomass) in Meditarranean forest in Jordan using Landsat data. The study reported that TNDVI (computed from Landsat image) has significant high correlation of 0.8 and was selected to map volume and crown cover.

Normalized difference water index (NDWI). Four different regression models were developed including linear, power, exponential and logarithmic models when Sentinel-2 computed NDWI was regressed against AGB (Table 2). The spatial distribution of biomass has been shown in Fig. 4. As shown in Table 2, NDWI has shown highest correlation as compared to other indices (GEMI, TNDVI, NDII and RENDVI). The linear model of NDWI was considered best for biomass mapping because it has highly significant relation with AGB (P value=0.05) with correlation (R^2 =0.469). The rest of models also performed well; logarithmic model has

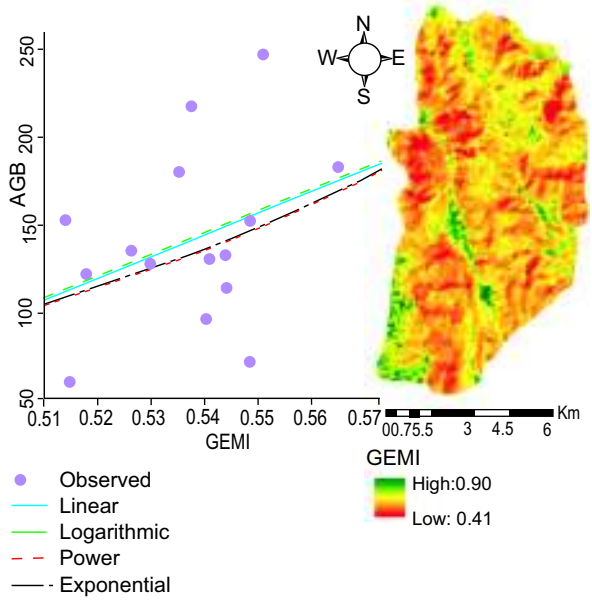

Fig. 2. Global environmental monitoring index (GEMI).

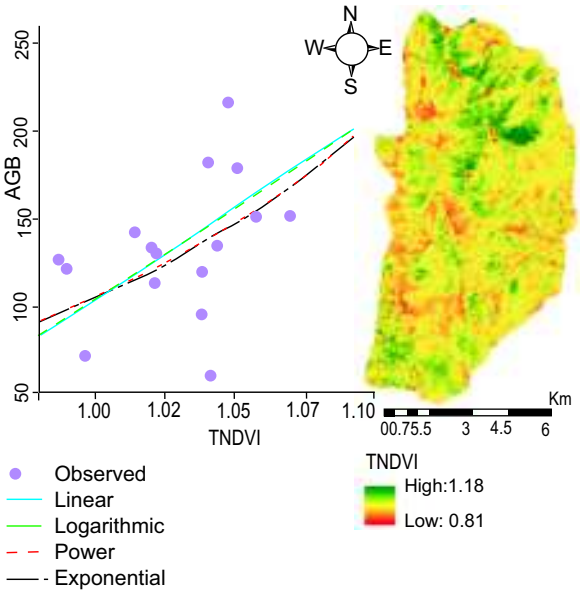

Fig. 3. Transformed normalized difference vegetation Index (TNDVI).

correlation of 0.351 with P-value of 0.20 followed by exponential model with correlation of 0.328 (P-value =0.26). The lowest performance was shown by power model which explained only 25% of the data variation (R^2=0.25) with significance value (P-value = 0.55). Therefore, among the four regression models, only power model was insignificant and become unsuitable for biomass estimation. Adan (2017) reported that NDWI has better performance as compared to NDII and NDWI has explained about 31% of the variation (R^2= 0.311) in linear regression model. Similarly, in non-linear relationship, NDWI has also slightly better performance than NDII with R^2 of 0.30.

Normalized difference infrared index (NDII). NDII was computed from Sentinel-2 image using Band 13 (SWIR) and Band 4 (Red) and the resultant values at sample plots locations were regressed with field data (Fig. 5). Four different types of regression models were developed including one linear and three non-linear relationships (power, exponential and logarithmic) as shown in Table 2. As depicted in Table 2, linear model of NDII shown the highest correlation with AGB having R^2 value 0.44 and the level of significance was also best (0.007). Thus the linear model explained 44% of data variation while 56% of the AGB data was remained unexplained. Regarding to non-linear models, the

Table 2. Summary of regression models of vegetation indices

	Global environmental monitoring index (GEMI)				
	Model summary			Parameter estimates	
Equation	R square	F	Sig.	Constant	b1
Linear	0.124	1.846	0.197	-523.216	1.237E3
Logarithmic	0.124	1.840	0.198	553.038	662.226
Power	0.113	1.655	0.221	2.558E3	4.759
Exponential	0.113	1.653	0.221	1.129	8.873
	Transformed normalized difference vegetation index (TNDVI)				
Linear	0.392	8.398	0.012	-950.373	1.055E3
Logarithmic	0.386	8.172	0.013	104.819	1.083E3
Power	0.289	5.292	0.039	104.571	7.059
Exponential	0.294	5.406	0.037	0.108	6.873
	Normalized difference water index (NDWI)				
Linear	0.469	11.473	0.005	69.653	711.333
Logarithmic	0.351	7.039	0.020	273.278	54.518
Power	0.255	4.456	0.055	309.516	0.350
Exponential	0.328	6.336	0.026	84.492	4.478
	Normalized difference infrared index (NDII)				
Linear	0.441	10.249	0.007	-26.281	542.541
Logarithmic	0.404	8.827	0.011	333.345	160.945
Power	0.296	5.457	0.036	456.836	1.036
Exponential	0.318	6.050	0.029	45.448	3.468
	Red-edge normalized difference vegetation index (RENDVI)				
Linear	0.378	7.905	0.015	-128.166	738.708
Logarithmic	0.354	7.130	0.019	404.853	259.748
Power	0.252	4.380	0.057	707.605	1.650
Exponential	0.267	4.728	0.049	24.132	4.672

where:

'F' means F-statistic named in honor of Ronald Fisher; 'b1' is the slope of regression line; 'R' is coefficient of correlation; 'Sig' is value of significance; Dependent variable was AGB and independent variables were GEMI, TNDVI, NDWI, NDII and RENDVI. Best Index among five indices was NDWI.

logarithmic model perform well with R^2 of 0.404 and P-value 0.11; followed by the exponential model with R^2 value of 0.318 (P-value=0.29). The least correlation was shown by power model with correlation (R^2=0.29) which explained only 29% of data variation. However, the power model of NDII has performed significantly with AGB as compared to power model of NDWI which was not significant. Adan, (2017) has reported performance of NDII was less than NDWI with correlation (R^2 = 0.22) in linear relationship, while it has bit better correlation (R^2 = 0.24) in non-linear relationship. Hunt *et al*. (2012) explored that NDII has better efficiency to express canopy water contents and can be used for vegetation monitoring.

Red-edge band normalized difference vegetation index (RENDVI). RENDVI was computed from Sentinel-2 image by using Red-Edge Band (Band 6) and NIR band (Band 8) as shown in Fig. 6. Sentinel-2 has three Red-Edge bands (Band 5, Band 6 and Band 7) which are in addition to Red Band and can be used for vegetation biophysical parameters. Four different regression models were developed for RENDVI including linear, power, exponential and logarithmic model (Table 2). According to Table 2, RENDVI has shown better as compared to broadband indices (GEMI and TNDVI) because RENDVI used narrowband combination. The linear model and logarithmic models have explained 37% of data variations followed by exponential model which explained 35% of AGB variation. Similarly, the lower correlations were observed in other non-linear models (power and exponential) with the R^2 of 0.25 and 0.26. Adnan, (2017) has reported bit higher correlation for RENDVI (R^2=0.59) in linear relationship while in non-linear regression model that correlation was increased

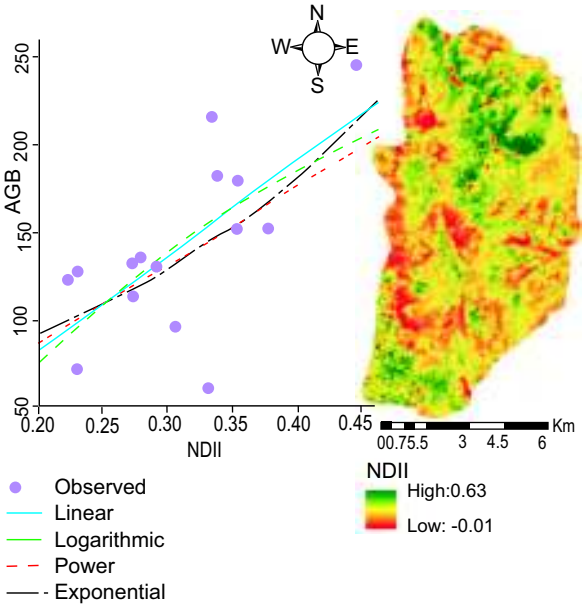

Fig. 5. Normalized difference infrared index (NDII).

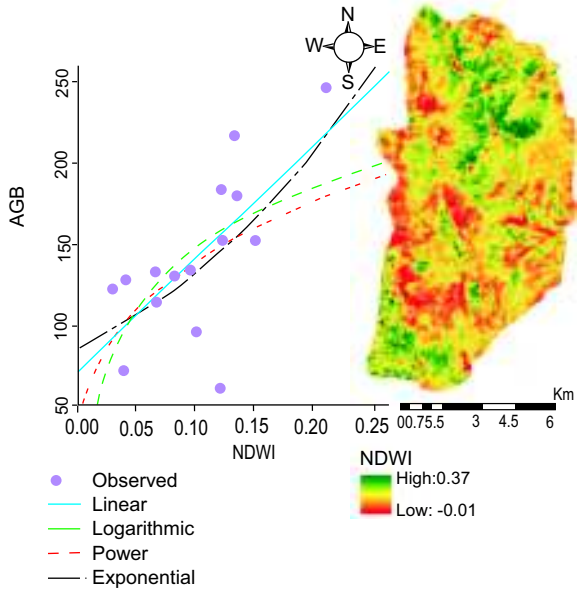

Fig. 4. Normalized difference water index (NDWI).

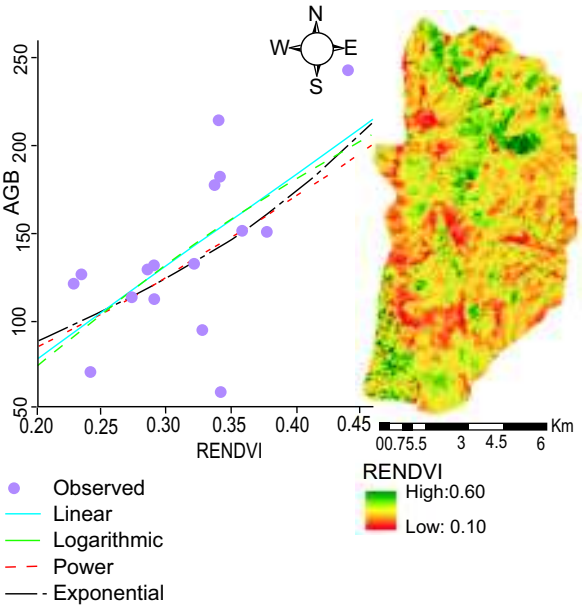

Fig. 6. Red-edge normalized difference vegetation index (RENDVI).

to 0.61. Thus explaining 61 percent of data variation and was considered best compared to broadband indices. Slonecker *et al*. (2009) demonstrated that narrowband (Red-edge) was more sensitive to leaf internal structure, properties and chlorophyll contents. Zhao *et al*. (2007) and Chen *et al*. (2007) have comparatively assessed the performance of NDVI, Simple ratio (SR) and Red-edge indices for biophysical parameters estimation. They reported that Red-edge band inclusion in vegetation indices have significantly increased the correlation (R^2 values).

Stepwise regression model (Sentinel-2 indices). Stepwise linear regression was used for best index selection for biomass mapping while considering all explanatory variables (GEMI, TNDVI, NDWI, NDII and RENDVI). The stepwise linear regression uses probability for selection in prediction model by considering the criteria, variable is selected when the significance is less than 0.50 and it is removed when it is greater than 0.10. The summary of stepwise correlation between AGB (dependent variable) and five indices (GEMI, TNDVI, NDWI, NDII, RENDVI) has been shown in Table 3. Similar to simple linear regression, NDWI was also selected in stepwise process whereas the other four indices were removed because they are not significant. The overall model explained 46% of the data variations and the rest of 54% data remained unexplained (R^2 was 0.46 and adjusted R^2 was 0.42 with 38.07 standard error). The correlation matrix depicted strong correlation between AGB and Sentinel-2 vegetation indices. The coefficient of correlation (R^2) for GEMI, TNDVI, NDWI, NDII and RENDVI were 0.35, 0.62, 0.68, 0.66 and 0.61 respectively. Previously, Imran and Ahmed, (2018) demonstrated stepwise linear regression for Landsat-8 spectral indices (NDVI, SAVI, PVI, DVI and ARVI) and AGB (t/ha) for the district Manshera in Pakistan. The study reported that stepwise linear regression was a robust method for best index selection for AGB estimation and reported coefficient of correlation (R^2 =0.63 and Adjusted R^2 was 0.60).

Accuracy assessment and biomass mapping. Normalized difference water index (NDWI) was selected for biomass estimation as it was considered best spectral index in single predictor regression model and stepwise linear regression model. The accuracy of linear model of NDWI and AGB was assessed by root mean square Error (RMSE), RMSE%, BIAS, BIAS% and MAPE%. As per selection condition, the RMSE and RMSE% of NDWI was equal to 28.27 (t/ha) and 28%, respectively. Similarly, BIAS, BIAS% and MAPE% were 20.44, 13.82% and 19.82% respectively. The resultant biomass

Table 3. Stepwise linear regression (AGB versus indices)

	Correlations					
	AGB	RENDVI	GEMI	TNDVI	NDWI	NDII
AGB	1.000	0.615	0.353	0.626	0.685	0.664
RENDVI	0.615	1.000	0.157	0.991	0.981	0.976
GEMI	0.353	0.157	1.000	0.124	0.145	0.167
TNDVI	0.626	0.991	0.124	1.000	0.970	0.965
NDWI	0.685	0.981	0.145	0.970	1.000	0.987
NDII	0.664	0.976	0.167	0.965	0.987	1.000

Variables selection			Model summary			
Entered	Removed	Significance	R	R Square	Adjusted R Square	Std. error of the estimate
NDWI		0.005	.685[a]	.469	.428	38.071
	RENDVI	0.152				
	GEMI	0.218				
	TNDVI	0.460				
	NDII	0.724				

Model coefficients			
Unstandardized coefficients			
B	Std. Error	t	Sig.
69.653	23.340	2.984	0.011
711.333	210.009	3.387	0.005

where:

'R' means correlation; 'R Square' is coefficient of correlation; 'B' is beta value which represent slope of regression line; 't' is the value of t-statistic; 'Sig' is value of significance; Model equation: AGB = 69.653+711.333*NDWI. Dependent variable: AGB; Stepwise method (variable selected <= .050 and removal >= .100).

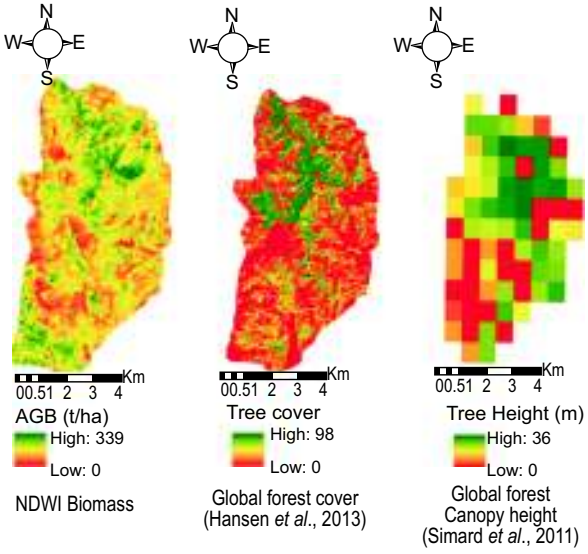

Fig. 7(abc). (a) NDWI biomass map; (b) Global forest tree cover map and (c) Global forest canopy height map.

map was generated in ArcGIS 10.3 software using Raster Calculator tool. NDWI based map AGB map has been shown in Fig. 7(a), which shows that AGB range from minimum (0) to maximum (339 t/ha). Mostly AGB was present in northern upper part of the study area and the resultant biomass was also compared to global forest maps. Global forest tree cover map (GFTCM) showed percent tree cover and range from 0 to 98% forest cover in the study area Fig. 7(b). GFTCM also showed higher forest density towards northern part and some forest patches in the southern side. Similarly, Global LiDAR 1 km forest canopy height showed height of forest canopy and values range from 0 in the south to higher (i.e 36 meters) in the northern parts of the study area Fig. 7(c). The present study AGB spatial distribution map was in consistent with both global vegetation maps which showed that AGB map was developed correctly.

Conclusion

This study investigated the forest above ground biomass prediction and modeling by utilizing Sentinel-2 vegetation indices in District Manshera, Khyber Pukhtunkhwa province, Pakistan. Sentinel-2 is the state-of-art sensor with refined spatial and frequent temporal resolution which makes it more useful for AGB estimation as compared to broadband multispectral sensors such as Landsat-8 (Ali et al., 2018); however Landsat-8 bands may perform better low resolution bands of Sentinel-2 (such as 60 m). Sentinel-2 data (Copernicus Open Acces Hub) provides wider accessibility and additional spectral bands when used in integration with robust commercial datasets such as world view-3 and Sentinel-1 or Synthetic Aperture Radar (SAR) can enhance AGB estimation reliability.

This study concludes the Sentinel-2 spectral indices have good relationship with field AGB plot wise data and can effectively predicted AGB of the Mansehra forests. The best index NDWI has shown an R^2 value 0.46 and an RMSE value of 28.27 t/ha and the remaining four indices (GEMI, TNDVI, NDII and RENDVI) performance were also satisfactory and can be used for biomass mapping over the study area, the inclusion of Red-edge band index also improve the correlation and can be used in determination of biophysical parameters and other forest/tree attributes such as crown, crown cover, volume and basal area. The present study suggests further in-depth work in the future using Sentinel-2 in national forest inventory to monitor temporal changes in AGB and carbon stocks. Sentinel-2 has high temporal resolution which is useful for reporting and measuring carbon emissions in future for implementation of forest conservation projects. Additionally, the present study also suggest applying the Red-edge bands available in Sentinel-2 to various vegetation parameters can further enhance AGB estimation. Therefore, Sentinel-2 sensor can be used as a decisive tool to assess and monitor AGB at local and regional scale. This state-of-art sensor should be further utilized and explored in studying the carbon accounting measurements and vegetation properties in Pakistan's forests and can provide best alternative data when use of commercial data is not available. It is, therefore, more detailed studies are needed to be performed to develop local, regional and national maps of different forest ecosystems.

Conflict of Interest. The authors declare no conflict of interest.

References

Adan, M.S. 2017. Integrating Sentinel-2A derived indices and terrestrial laser scanner to estimate above ground biomass/carbon in Ayer Hitam tropical forest, Malaysia, *Master Thesis*, pp. 30-50, University of Twente, The Netherlands.

Ahmad, F. 2012. Spectral vegetation indices performance evaluated for Cholistan desert. *Journal of Geo-*

graphy and Regional Planning, **5:** 165-172.

Ali, A., Ullah, S., Bushra, S., Ahmad, N., Ali, A., Khan, M.A. 2018. Quantifying forest carbon stocks by integrating satellite images and forest inventory data. *Austrian Journal of Forest Science*, **135:** 93-117.

Alrababah, M.A., Alhamad, M.N., Bataineh, A.L., Bataineh, M.M., Suwaileh, A.F. 2011. Estimating east mediterranean forest parameters using Landsat ETM. *International Journal of Remote Sensing*, **32:** 1561-1574.

Astola, H., Häme, T., Sirro, L., Molinier, M., Kilpi, J. 2019. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. *Remote Sensing of Environment*, **223:** 257-273.

Cao, Q., Miao, Y., Shen, J., Yu, W., Yuan, F., Cheng, S., Liu, F. 2016. Improving in-season estimation of rice yield potential and responsiveness to topdressing nitrogen application with crop circle active crop canopy sensor. *Precision Agriculture*, **17:** 136-154.

Chang, J., Shoshany, M. 2016. Mediterranean shrub lands biomass estimation using Sentinel-1 and Sentinel-2. In 2016 *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5300-5303.

Chen, J.C., Yang, C.M., Wu, S.T., Chung, Y.L., Charles, A.L., Chen, C.T. 2007. Leaf chlorophyll content and surface spectral reflectance of tree species along a terrain gradient in Taiwan's Kenting National Park. *Student*, **48:** 71-77.

Chen, J., Gu, S., Shen, M., Tang, Y., Matsushita, B. 2009. Estimating above-ground biomass of grassland having a high canopy cover: an exploratory analysis of in situ hyperspectral data. *International Journal of Remote Sensing*, **30:** 6497-6517.

Chen, L., Ren, C., Zhang, B., Wang, Z., Xi, Y. 2018. Estimation of forest above-ground biomass by geographically weighted regression and machine learning with Sentinel imagery. *Forests*, **9:** 582.

Chen, Y., Li, L., Lu, D., Li, D. 2019. Exploring bamboo forest above ground biomass estimation using sentinel-2 data. *Remote Sensing*, **11:** 7-29.

Chrysafis, I., Mallinis, G., Tsakiri, M., Patias, P. 2019. Evaluation of single-date and multi-seasonal spatial and spectral information of Sentinel-2 imagery to assess growing stock volume of a Mediterranean forest. *International Journal of Applied Earth Observation and Geoinformation*, **77:** 1-14.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, **120:** 25-36.

Dube, T., Mutanga, O., Elhadi, A., Ismail, R. 2014. Intra-and-inter species biomass prediction in a plantation forest: testing the utility of high spatial resolution spaceborne multispectral rapideye sensor and advanced machine learning algorithms. *Sensors*, **14:** 15348-15370.

Forkuor, G., Dimobe, K., Serme, I., Tondoh, J.E. 2018. Landsat-8 vs. Sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *GI Science and Remote Sensing*, **55:** 331-354.

Frampton, W.J., Dash, J., Watmough, G., Milton, E.J. 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, **82:** 83-92.

Gupta, K., Mukhopadhyay, A., Giri, S., Chanda, A., Majumdar, S.D., Samanta, S., Mitra, D., Samal, R.N., Pattnaik, A.K., Hazra, S. 2018. An index for discrimination of mangroves from non-mangroves using LANDSAT 8 OLI imagery. *Methods X*, **5:** 1129-1139.

Henry, M., Picard, N., Trotta, C., Manlay, R., Valentini, R., Bernoux, M., Saint André, L. 2011. Estimating tree biomass of sub-Saharan African forests: a review of available allometric equations. *Silva Fennica*, **45:** 477-569.

Hunt, E.R., Wang, L., Qu, J.J., Hao, X. 2012. Remote sensing of fuel moisture content from canopy water indices and normalized dry matter index. *Journal of Applied Remote Sensing*, **6:** 95-119.

Imran, A.B., Ahmed, S. 2018. Potential of Landsat-8 spectral indices to estimate forest biomass. *International Journal of Human Capital in Urban Management*, **3:** 303-314.

Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., Moran, E. 2016. A survey of remote sensing-based above ground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*, **9:** 63-105.

Lu, D., Mausel, P., Brondýzio, E., Moran, E. 2004. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. *Forest Ecology and Management*,

198: 149-167.

Majasalmi, T., Rautiainen, M. 2016. The potential of Sentinel-2 data for estimating biophysical variables in a boreal forest: a simulation study. *Remote Sensing Letters*, **7:** 427-436.

Mathieu, R., Naidoo, L., Cho, M.A., Leblon, B., Main, R., Wessels, K., Asner, G.P., Buckley, J., Van Aardt, J., Erasmus, B.F., Smit, I. P. 2013. Toward structural assessment of semi-arid African savannahs and woodlands: The potential of multitemporal polarimetric RADARSAT-2 fine beam images. *Remote Sensing of Environment*, **138:** 215-231.

Navarro, J., Algeet, N., Fernández-Landa, A., Esteban, J., Rodríguez-Noriega, P., Guillén-Climent, M. 2019. Integration of uav, sentinel-1, and sentinel-2 data for mangrove plantation aboveground biomass monitoring in senegal. *Remote Sensing*, **11:** 77-98.

Nizami, S.M. 2012. The inventory of the carbon stocks in sub-tropical forests of Pakistan for reporting under Kyoto Protocol. *Journal of Forestry Research*, **23:** 377-384.

Nouri, A., Kiani, B., Hakimi, M.H., Mokhtari, M.H. 2018. Estimating oak forest parameters in the western mountains of Iran using satellite-based vegetation indices. *Journal of Forestry Research*, **31:** 541-552.

Nuthammachot, N., Phairuang, W., Wicaksono, P., Sayektiningsih, T. 2018. Estimating aboveground biomass on private forest using sentinel-2 Imagery. *Journal of Sensors*, **20:** 1-11.

Pandit, S., Tsuyuki, S., Dube, T. 2018. Estimating above-ground biomass in sub-tropical buffer zone community forests, Nepal, using Sentinel 2 data. *Remote Sensing*, **10:** 601-627.

Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M. 2009. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, **113:** 110-122.

Ramoelo, A., Cho, M., Mathieu, R., Skidmore, A.K. 2015. Potential of Sentinel-2 spectral configuration to assess rangeland quality. *Journal of Applied Remote Sensing*, **124:** 516-533

Rana, P., Korhonen, L., Gautam, B., Tokola, T. 2014. Effect of field plot location on estimating tropical forest above-ground biomass in Nepal using airborne laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, **94:** 55-62.

Schultz, M., Clevers, J.G., Carter, S., Verbesselt, J., Avitabile, V., Quang, H.V., Herold, M. 2016. Performance of vegetation indices from Landsat time series in deforestation monitoring. *International Journal of Applied Earth Observation and Geoinformation*, **52:** 318-327.

Shen, W., Li, M., Huang, C., Wei, A. 2016. Quantifying live aboveground biomass and forest disturbance of mountainous natural and plantation forests in northern Guangdong, China, based on multi-temporal Landsat, PALSAR and field plot data. *Remote Sensing*, **8:** 595-614.

Shi, L. 2010. Changes of forest in northeast China over the past 25 years: an analysis based on remote sensing technique. In remote sensing of the environment: *The 17th China Conference on Remote Sensing* 2010 Sep. 14.

Sibanda, M., Mutanga, O., Rouget, M. 2015. Examining the potential of sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. *ISPRS Journal of Photogrammetry and Remote Sensing*, **110:** 55-65.

Slonecker, T., Haack, B., Price, S. 2009. Spectroscopic analysis of arsenic uptake in Pteris ferns. *Remote Sensing*, **1:** 644-675.

Wijaya, A., Kusnadi, S., Gloaguen, R., Heilmeier, H. 2010. Improved strategy for estimating stem volume and forest biomass using moderate resolution remote sensing data and GIS. *Journal of Forestry Research*, **21:** 1-12.

Wittke, S., Yu, X., Karjalainen, M., Hyyppä, J., Puttonen, E. 2019. Comparison of two-dimensional multitemporal Sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. *International Journal of Applied Earth Observation and Geoinformation*, **76:** 167-178.

Zhao, D., Huang, L., Li, J., Qi, J. 2007. A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. *ISPRS Journal of Photogrammetry and Remote Sensing*, **62:** 25-33.