Neural network-based modelling for forest biomass assessment

Subrata Nandy, Rajpal Singh, Surajit Ghosh, Taibanganba Watham, Satya Prakash Singh Kushwaha, Arumugam Senthil Kumar & Vinay Kumar Dadhwal

Published online: 04 Aug 2017.

Submit your article to this journal

Article views: 500

View related articles

View Crossmark data

Citing articles: 25 View citing articles

Taylor & Francis
Taylor & Francis Group

Check for updates

Neural network-based modelling for forest biomass assessment

Subrata Nandy [iD][a], Rajpal Singh[b], Surajit Ghosh [iD][a], Taibanganba Watham[a], Satya Prakash Singh Kushwaha[a], Arumugam Senthil Kumar[a] and Vinay Kumar Dadhwal[c]

[a]Indian Institute of Remote Sensing, ISRO, Dehradun 248001, India; [b]Directorate of Forest Education, Ministry of Environment, Forests and Climate Change, Government of India, Dehradun 248006, India; [c]Indian Institute of Space Science and Technology, Dept. of Space, Govt. of India, Thiruvananthapuram 695 547, India

ABSTRACT

Forest biomass is an important parameter for assessing the status of forest ecosystems. In the present study, forest biomass was assessed by integrating remotely-sensed satellite data and field inventory data using an artificial neural network (ANN) technique in Barkot forest, Uttarakhand, India. Spectral and texture variables were derived from Resourcesat-1 (RS1) LISS-III (Linear Imaging Self-Scanning Sensor) data of April 24, 2013. ANN was used for finding the relation of spectral and texture variables to field-measured biomass. The top 10 variables, namely shortwave infrared (SWIR) band reflectance, near infrared (NIR) band reflectance, normalized difference vegetation index (NDVI), difference vegetation index (DVI), green band contrast, green band variance, SWIR band contrast, NIR band dissimilarity, SWIR band second angular moment, and red band mean, were selected for generating a multiple linear regression model to predict the biomass. The predicted biomass showed a good relationship ($R^2 = 0.75$ and root mean square error (RMSE) = 85.32 Mg ha^{-1}) with field-measured biomass. The model was validated yielding $R^2 = 0.74$ and RMSE = 93.41 Mg ha^{-1}. The results showed that RS1 LISS-III satellite data have good capability to estimate forest biomass, and the ANN technique can be used to enhance the scope of biomass estimation with a minimum number of spectral and texture variables.
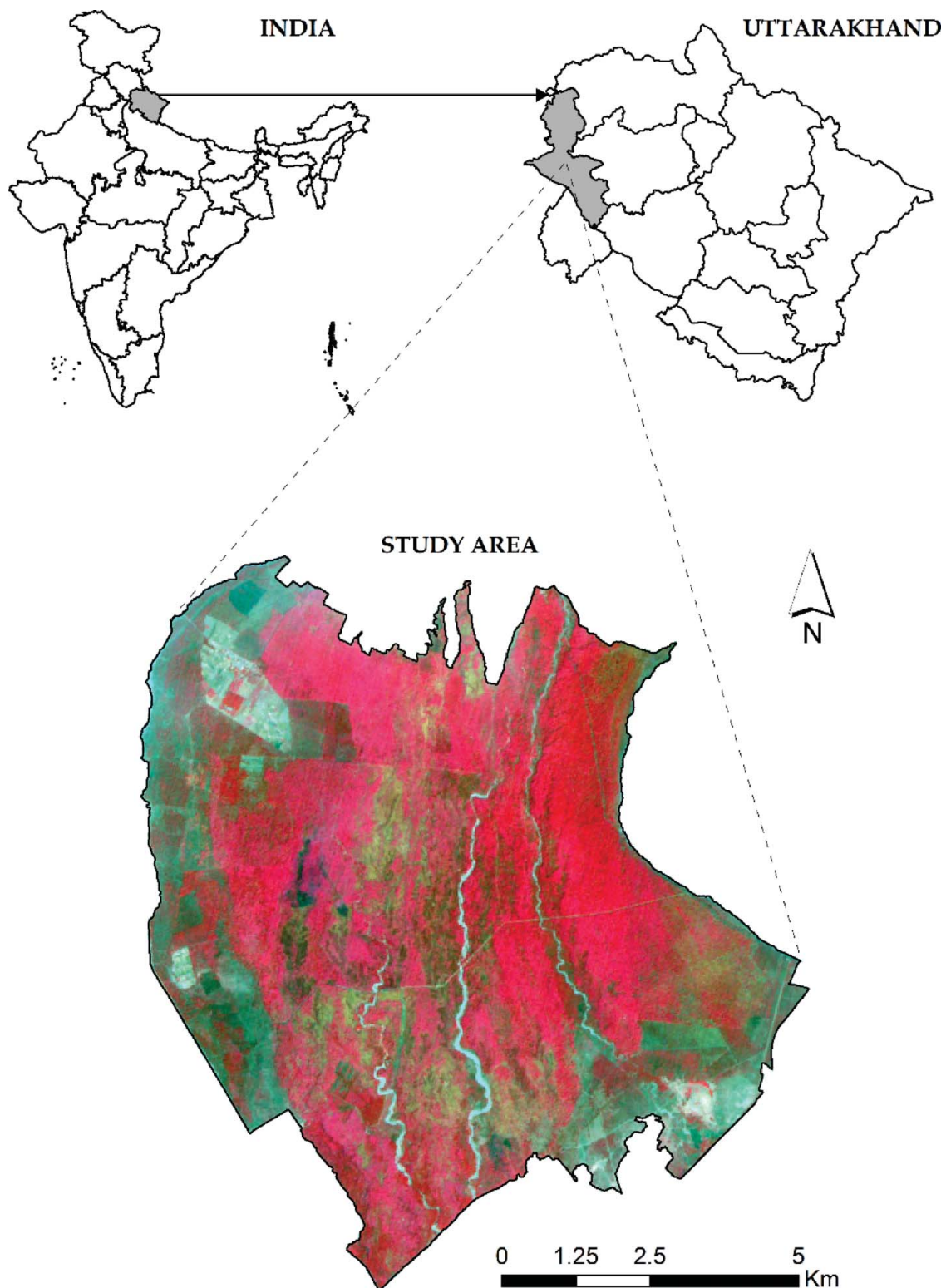
Introduction

Forest acts as an important carbon reservoir, in addition to being a source of livelihood to billions of people across the globe [1]. The International Union of Forest Research Organizations (IUFRO) recognized forest biomass as one of the most significant areas of priority [2], which makes it an important forest attribute to explore. The forest biomass governs the potential carbon emission that could be released to the atmosphere due to deforestation, and change of regional biomass is associated with changes in the climate and ecosystem [3]. Accounting for baseline carbon stocks in forests has been noted as essential for the implementation of climate change policies [4], and as necessary for increasing the accuracy of global carbon cycle models [5]. Forest biomass is thus an important forest characteristic to study for providing valuable inputs to conservation of forests and mitigation of adverse effects of climate change.

The rates of carbon emission are considered the largest source of uncertainty in climate change scenarios due to the difficulty in spatial estimation of carbon stocks and dynamics. Traditional inventory links ground measurements and tree parameters to tree biomass using species-specific allometric relationships [6], and further extrapolates these measurements to entire forest areas [7]. Apart from the traditional methods for biomass estimation, the modern remote sensing-based methods have widely been used due to their comprehensive spatial and temporal coverage, and cost- and time effectiveness. Optical satellite data in conjunction with the field inventory data have been widely used for estimating forest biomass [8–14]. Past studies have produced significant advances in forest biomass estimation including the application of different sensor data and the development of advanced techniques such as regression analysis, geostatistical analysis, neural network, and process-based ecosystem models [3,15,16]. The present study aimed to assess forest biomass by integrating spectral and textural variables of optical satellite and field inventory data using an artificial neural network (ANN) technique.

ANN has been considered an important non-parametric algorithm for forest biophysical parameter estimation. In contrast to conventional parametric approaches, ANN does not require any assumptions about statistical distribution of the data [17,18]. It is more successful at solving non-linear problems of the parametric models. The network commonly consists of one input layer, one or more hidden layers, and one output layer. In ANN a relatively large

CONTACT Subrata Nandy ✉ subrato.nandy@gmail.com

INDIA

UTTARAKHAND

STUDY AREA

N

0   1.25   2.5   5
                         Km

Figure 1. Location of the study area.

amount of data is required for training and learning. During the training phase, the neural network learn about the patterns present in the training data, and, considering the learned pattern, the network constructs rules that can be applied for testing the data set. Multilayer perceptron (MLP) neuronal network is the most widely used neural network model [19]. MLP is a predictive model for a dependent variable based on the values of predictor variables. The feed-forward MLP is used in this study because of its desirable computational and approximation capabilities [20]. In this study MLP was applied to optimize the number of spectral and textural variables derived from Linear Imaging Self-Scanning Sensor-III (LISS-III) data based on their contribution toward addressing the field-measured biomass. These selected variables were further used as inputs in a multiple linear regression model to predict forest biomass.
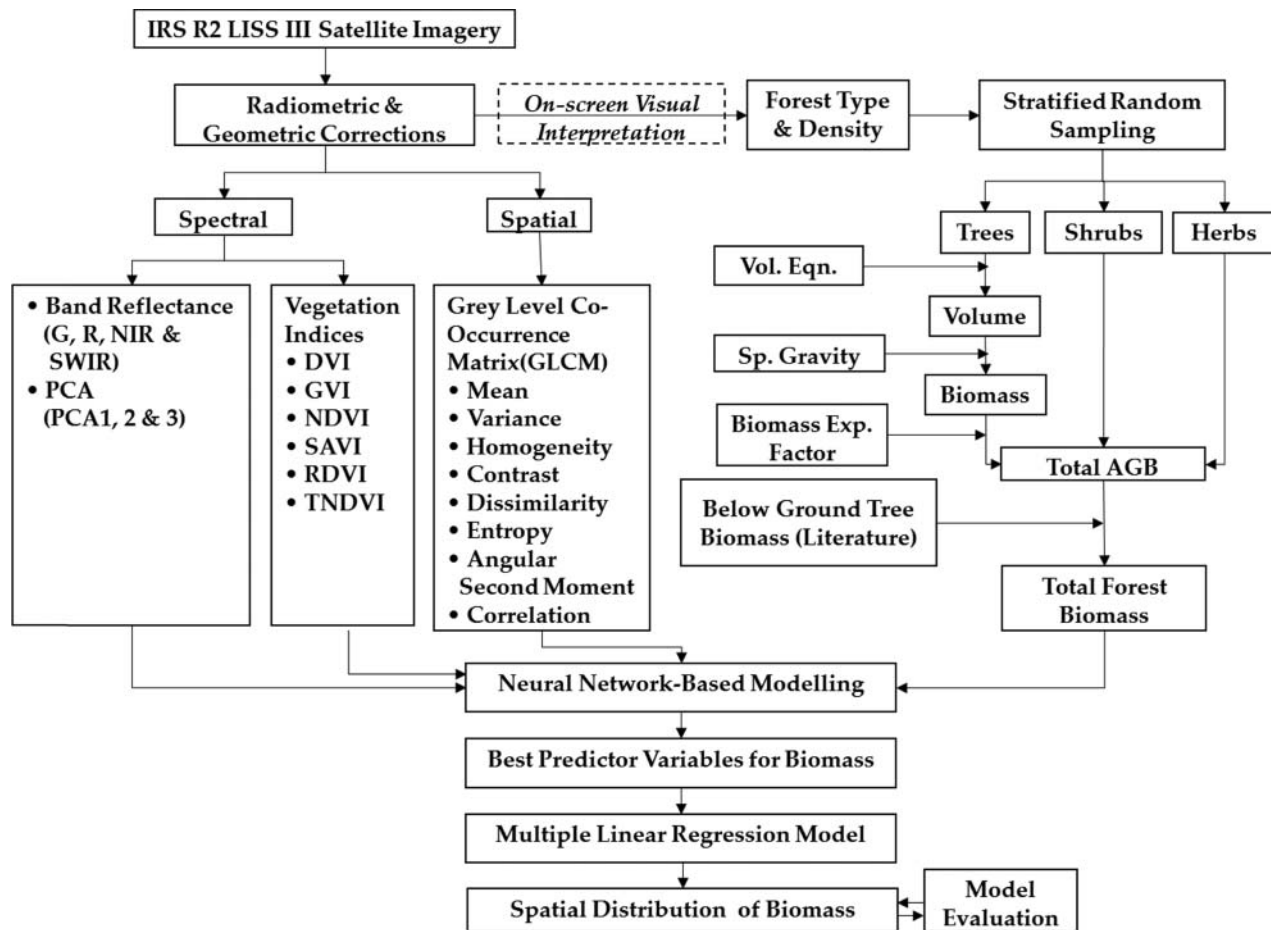
Figure 2. The methodology. [G: green; R: red; NIR: near infrared; SWIR: shortwave infrared; PCA: principal component analysis; DVI: difference vegetation index; GVI: green vegetation index; NDVI: normalized difference vegetation index; SAVI: soil adjusted vegetation index; RDVI: re-normalized difference vegetation index; TNDVI: transformed normalized difference vegetation index; Vol. Eqn.: volumetric equations; Sp.: specific; Exp.: expansion; AGB: aboveground biomass]

Materials and methods

Study area

The area selected for the present study is Barkot forest (30°03′52″–30°10′43″N, 78°09′49″–78°17′09″E) in Uttarakhand, India (Figure 1). The forest of the study area can be broadly classified as Tropical Moist Deciduous forest [21] dominated by sal (*Shorea robusta*) and its associates. Major tree species present here are sal, kamala tree (*Mallotus philippensis*), chamror (*Ehretia laevis*), golden rain tree (*Cassia fistula*) and plantations of teak (*Tectona grandis*). The area has a predominantly tropical to sub-tropical humid climate with temperature ranging from 2 to 41 °C. The average annual rainfall is typically 2300 mm. The abundance of forest cover, accessibility and data availability were the motivating factors for the choice of the study area.

Methodology

Satellite data processing and variable extraction

For the present study, Resourcesat-1 LISS-III satellite data of April 24, 2013, was used. LISS-III provides multispectral images in four bands (green, red, near infrared and shortwave infrared) with 23.5-m spatial resolution. Figure 2 illustrates the methodology in detail. The digital numbers of all the bands of LISS-III data were converted to reflectance image [101]. The image was geometrically referenced with Universal Transverse Mercator (UTM) 44 N projection and World Geodetic System (WGS) 84 datum. A subset image covering the study area was extracted from the whole scene.

Spectral variables including band reflectance, principal components (PCs) and vegetation indices, namely normalized difference vegetation index (NDVI) [22], renormalized difference vegetation index (RDVI) [23], difference vegetation index (DVI) [24], green vegetation index (GVI) [25], soil-adjusted vegetation index (SAVI) [26] and transformed normalized difference vegetation index (TNDVI) [27], of the study area were extracted from the reflectance image. However, the NDVI is saturated in high biomass and it is sensitive to a number of perturbing factors, such as atmospheric effects, cloud, soil effects, anisotropic effects, etc. These factors may lead an erroneous estimation. To overcome these drawbacks, various derivatives of and alternative indices to NDVI have been developed. Considering the vegetation condition of the study area and number of bands of the LISS-III
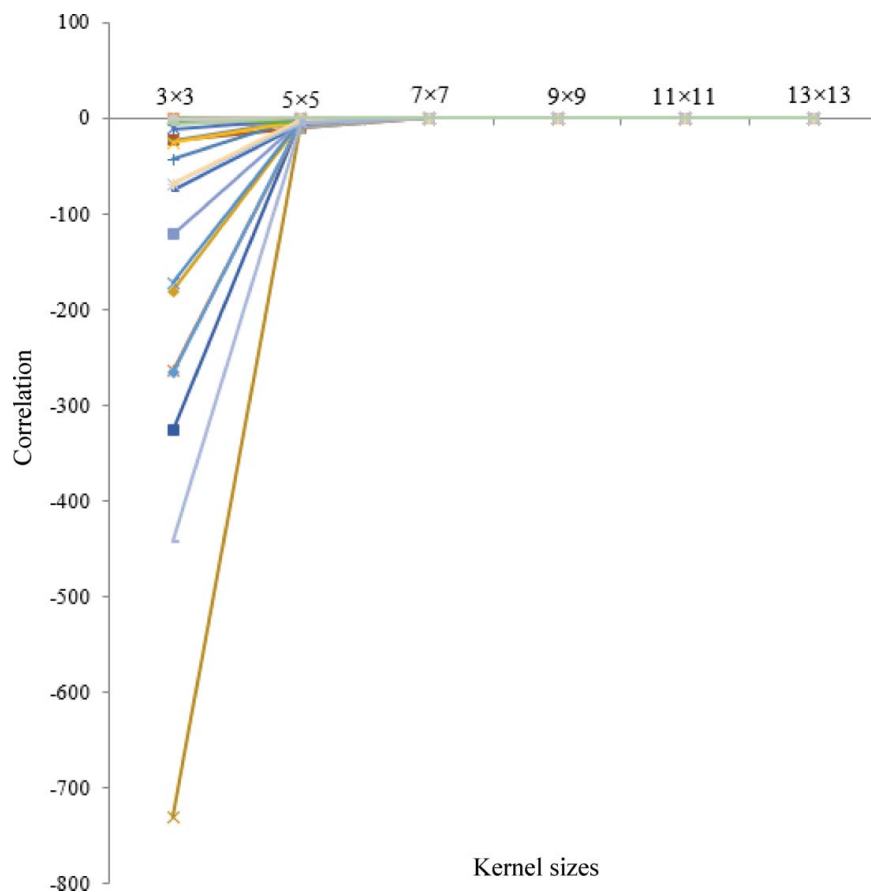
Figure 3. Values of correlation images of different kernel sizes at the field points.

imagery, RDVI, DVI, GVI, SAVI and TNDVI were used to minimize the saturation issue and other perturbing factors related to NDVI. The gray level co-occurrence matrix (GLCM) method of texture analysis [28] was used to derive different texture variables (e.g. correlation, contrast, dissimilarity, entropy, homogeneity, mean, second angular moment and variance) for individual bands of the LISS-III image. First, the correlation variable was tested for different kernel sizes (3 × 3, 5 × 5, 7 × 7, 9 × 9, 11 × 11 and 13 × 13 pixels) to determine the optimum kernel size for further analysis. It was observed that there was great variation in correlation variable for 3 × 3 kernel size which almost converged at 5 × 5 kernel size (Figure 3) and became almost constant at 7 × 7 and larger kernel sizes. Optimum kernel size was determined when the correlation variable became stable. Hence, 5 × 5 kernel was taken as optimum kernel size for extraction of all the texture variables. The spectral and textural information derived from the satellite image were considered independent variables. A total of 13 spectral and 32 texture variables were derived (Table S1).

Sample design and field data collection
The false color composite (FCC) of the study area (Figure 1) was taken to the field to relate the image features with the vegetation types/land uses and the forest canopy density; this information was used later for

the stratification of different vegetation types/land uses and canopy density categories through on-screen visual image interpretation at 1:50,000 scale. Three canopy density classes, namely very dense (>70%), moderately dense (40–70%) and open (10–40%), were delineated. Stratified random sampling was applied to lay out sample plots in different strata for field data collection. A pilot survey was conducted in the study area
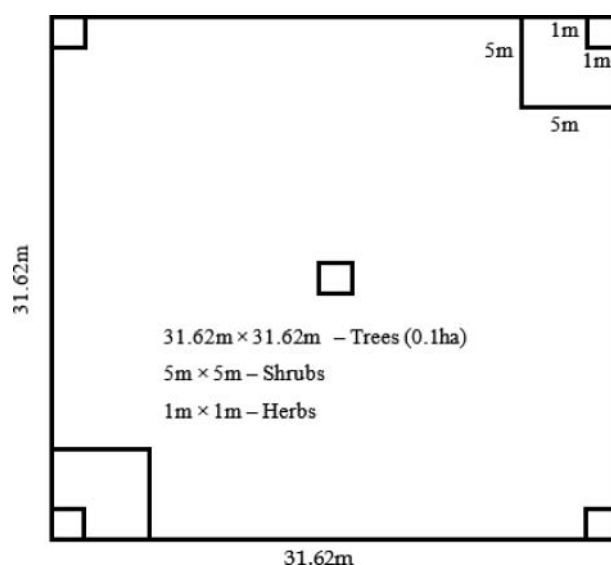

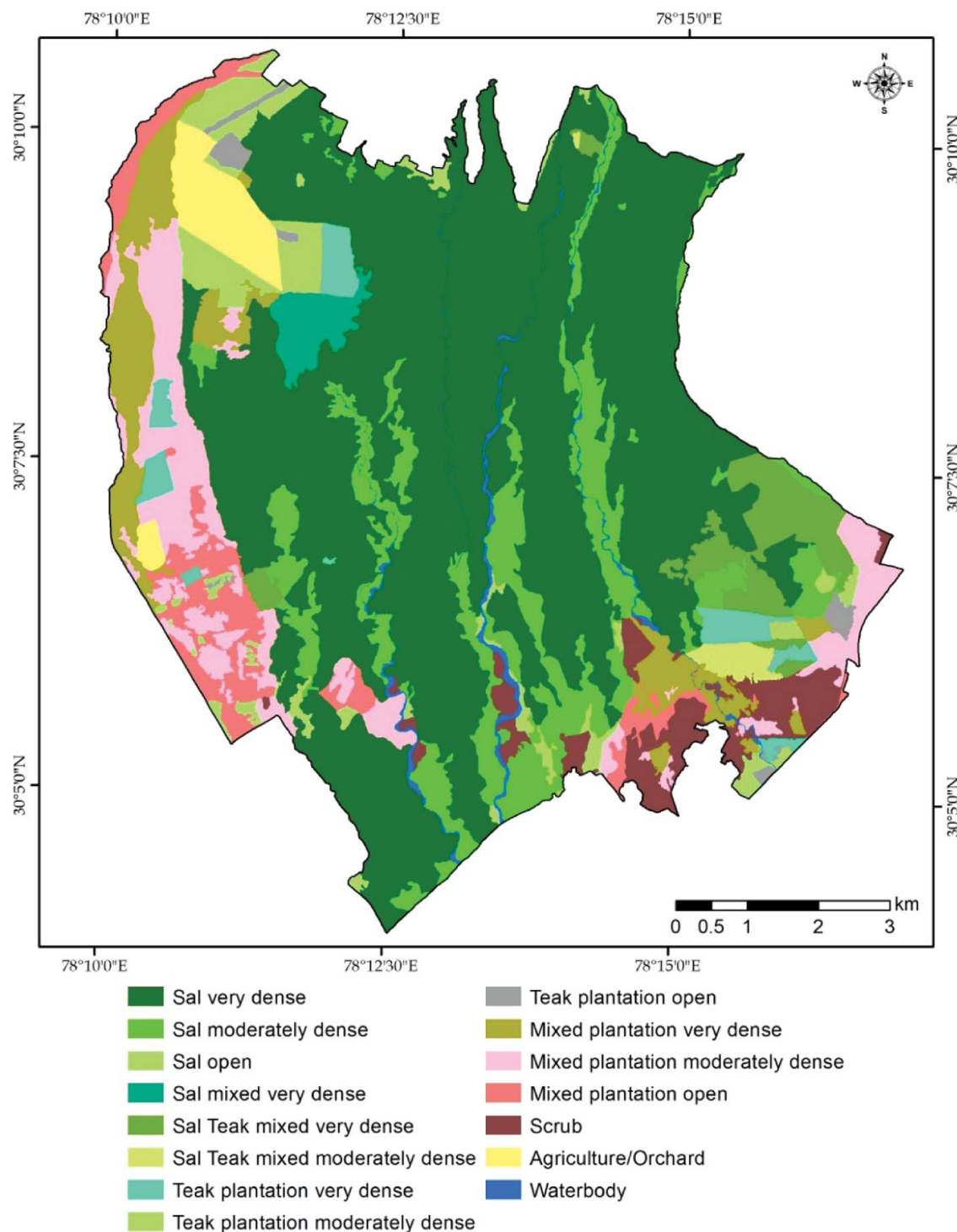**Figure 4.** The field sampling plot design.

Figure 5. Forest type-density map.

Table 1. Area under different forest type-density classes.

Sl. no.	Forest type-density class	Area (ha)
1	Sal very dense	4493.22
2	Sal moderately dense	943.68
3	Sal open	95.87
4	Sal mixed very dense	109.07
5	Sal teak mixed very dense	290.26
6	Sal teak mixed moderately dense	50.14
7	Teak plantation very dense	168.27
8	Teak plantation moderately dense	302.19
9	Teak plantation open	55.45
10	Mixed plantation very dense	459.42
11	Mixed plantation moderately dense	552.25
12	Mixed plantation open	340.30
13	Scrub	291.09

to calculate the required number of sampling plots, using the following equation [29]:

$$N = \frac{t^2 \times (CV)^2}{(SE\%)^2} \qquad (1)$$

where N is the number of sample plots, t is the statistical value at 95% significance level, CV is the coefficient of variation and SE% is the standard error percentage. In this study, CV of diameter at breast height (*dbh*) of trees was considered.

Table 2. Mean total biomass in different strata.

Sl. no.	Forest type-density	Mean tree AGB (Mg ha^{-1})	Mean BGB (Mg ha^{-1})	Mean shrub, herbs, litter (Mg ha^{-1})	Mean total biomass (Mg ha^{-1})
1	Sal very dense	519.42	153.42	9.91	682.75
2	Sal moderately dense	365.67	107.80	14.48	487.95
3	Sal open	286.10	74.73	5.96	366.79
4	Sal mixed very dense	208.46	58.90	6.58	273.95
5	Sal teak mixed very dense	618.52	180.85	8.68	808.05
6	Sal teak mixed moderately dense	514.81	137.34	2.86	655.02
7	Teak plantation very dense	429.01	111.54	10.92	551.47
8	Teak plantation moderately dense	242.28	62.99	2.84	308.11
9	Teak plantation open	207.64	53.99	10.10	271.73
10	Mixed plantation very dense	289.19	76.56	6.57	377.87
11	Mixed plantation moderately dense	294.44	75.19	13.50	377.57
12	Mixed plantation open	118.69	30.86	10.89	160.44
13	Scrub	104.41	27.15	5.64	137.20

AGB: aboveground biomass; BGB: belowground biomass.
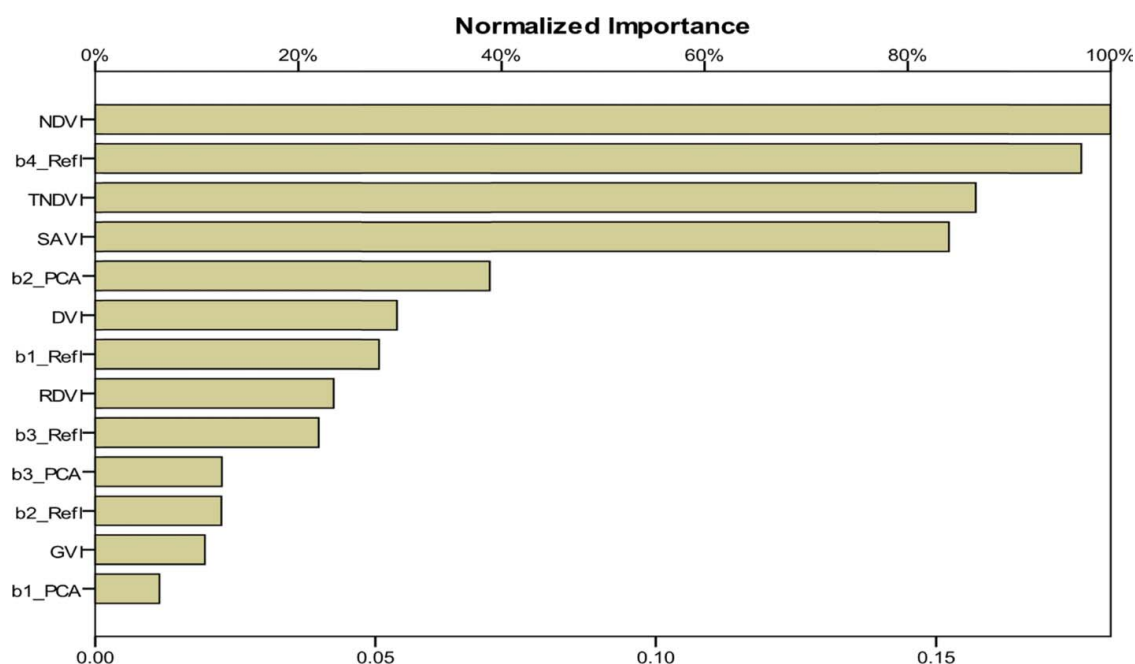
Figure 6. Normalized importance of spectral variables calculated using an artificial neural network (ANN). [NDVI: normalized difference vegetation index; b4_Refl: Band 4 reflectance; TNDVI: transformed normalized difference vegetation index; SAVI: soil adjusted vegetation index; b2_PCA: second principal component; DVI: difference vegetation index; b1_Refl: Band 1 reflectance; RDVI: re-normalized difference vegetation index; b3_Refl: Band 3 reflectance; b3_PCA: third principal component; b2_Refl: Band 2 reflectance; GVI: green vegetation index; b1_PCA: first principal component]

The total number of sample plots required to be laid in different strata was found to be 36. Of these, 70% (26) of the plots were randomly selected as training data for the analyses and the remaining 30% (10) plots were used for accuracy assessment. The total number of sample plots was proportionately distributed in each of the strata as:

$$n_h = \frac{N_h}{N} \times n \qquad (2)$$

where n_h = number of samples in h stratum, N_h = size of h stratum, N = total population size and n = total number of samples.

Field inventory for trees was carried out in square plots of size 0.1 ha (31.62 m × 31.62 m) in all 36 sample plots laid down randomly in different strata. At each sample plot, species name, dbh of all trees (1.37 m), height and crown density were noted. Other details such as geographic coordinates, location, slope, aspect and evidence of disturbances were also recorded. Two sample plots of 5 m × 5 m for shrubs at two opposite corners, four sample plots of 1 m × 1 m for herbs at the four corners and one sample plot of 1 m × 1 m at the center for litter were also laid in each sample plot (0.1 ha; Figure 4). The total number of individuals for each shrub species was counted and a specimen sample from each species was collected including roots. Fresh weight was noted with the use of an electronic balance. A representative part of the sample was kept in a hot-air oven at 80°C for drying to constant weight, and the

Normalized Importance
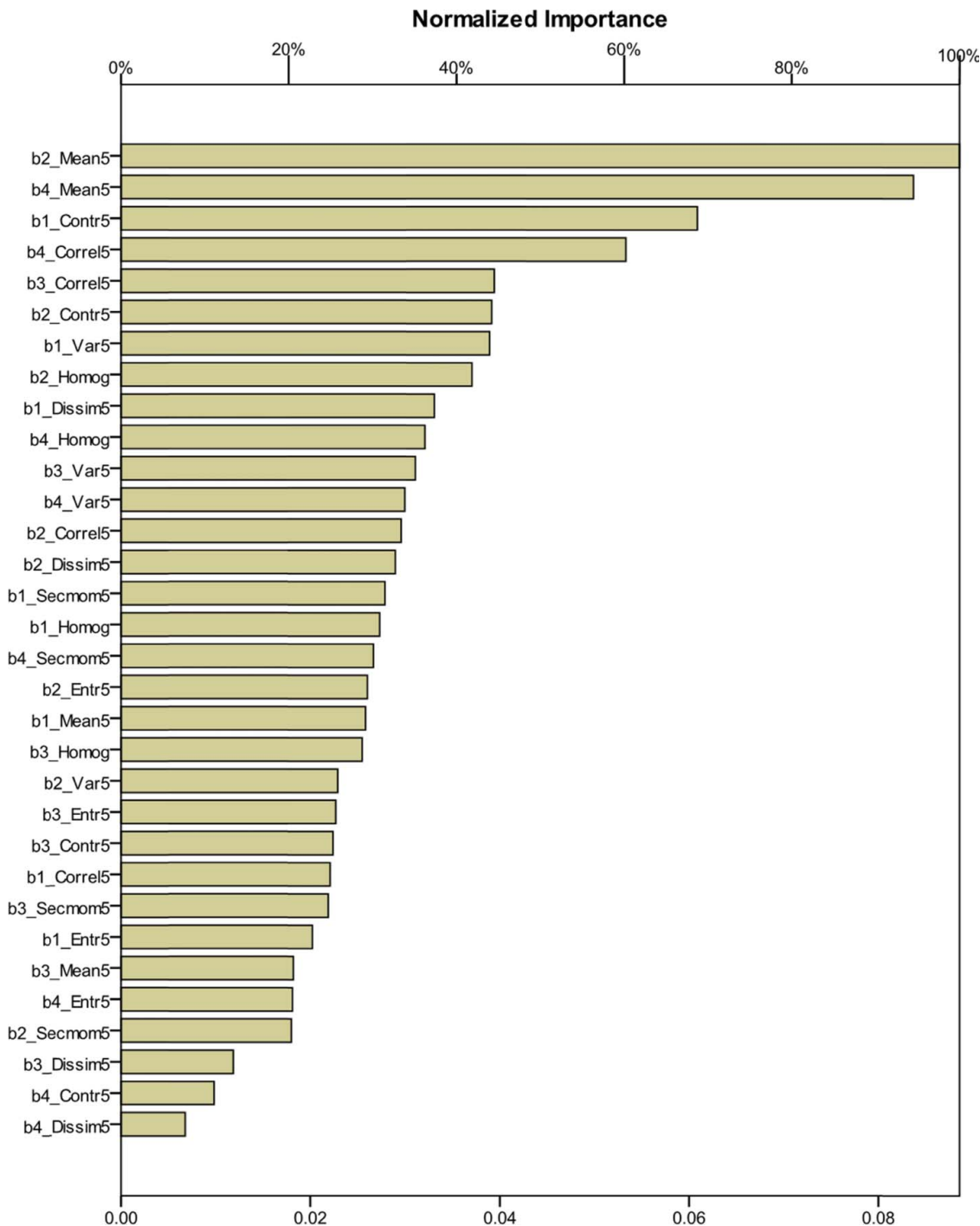
Figure 7. Normalized importance of texture variables calculated using an artificial neural network (ANN). [b2_Mean5: Band 2 mean; b4_Mean5: Band 4 mean; b1_Contr5: Band 1 contrast; b4_Correl5: Band 4 correlation; b3_Correl5: Band 3 correlation; b2_Contr5: Band 2 contrast; b1_Var5: Band 1 variance; b2_Homog: Band 2 homogeneity; b1_Dissim5: Band 1 dissimilarity; b4_Homog: Band 4 homogeneity; b3_Var5: Band 3 variance; b4_Var5: Band 4 variance; b2_Correl5: Band 2 correlation; b2_Dissim5: Band 2 dissimilarity; b1_Secmom5: Band 1 angular second moment; b1_Homog: Band 1 homogeneity; b4_Secmom5: Band 4 angular second moment; b2_Entr5: Band 2 entropy; b1_Mean5: Band 1 mean; b3_Homog: Band 3 homogeneity; b2_Var5: Band 2 variance; b3_Entr5: Band 3 entropy; b3_Contr5: Band 3 contrast; b1_Correl5: Band 1 correlation; b3_Secmom5: Band 3 angular second moment; b1_Entr5: Band 1 entropy; b3_Mean5: Band 3 mean; b4_Entr5: Band 4 entropy; b2_Secmom5: Band 2 angular second moment; b3_Dissim5: Band 3 dissimilarity; b4_Contr5: Band 4 contrast; b4_Dissim5: Band 4 dissimilarity]

oven-dry weight for shrub species was noted. Finally, total shrub biomass was worked out for each plot. A similar process was used for herb biomass estimation. Litter from the 1 m × 1 m sample plot was collected and fresh weight was measured. A representative part

of the litter was oven dried to obtain the litter biomass per plot.

The volume of individual trees of the sample plot was calculated by using *dbh* value in species-specific volumetric equations developed by the Forest

Table 3. The variables used in multiple linear regression and the corresponding R^2 value.

Sl. no.	Variables	R^2
Spectral variables		
1.	Band4	0.389
2.	Band4, NDVI	0.494
3.	Band4, Band3, NDVI	0.511
4.	Band4, Band3, NDVI, DVI	0.511
Texture variables		
1.	contrast1	0.052
2.	contrast1, variance1	0.055
3.	contrast1, variance1, contrast4	0.056
4.	contrast1, variance1, contrast4, dissimilarity3	0.305
5.	contrast1, variance1, contrast4, dissimilarity3, angular second moment4	0.683
6.	Contrast1, variance1, contrast4, dissimilarity3, angular second moment4, mean2	0.684
Spectral and texture variables combined		
1.	Band4, contrast1	0.416
2.	Band4, NDVI, contrast1, variance 1	0.591
3.	Band4, Band3, NDVI, contrast1, variance1, contrast4	0.621
4.	Band4, Band3, NDVI, DVI, contrast1, variance1, contrast4, dissimilarity3	0.693
5.	Band4, Band3, NDVI, DVI, contrast1, variance1, contrast4, dissimilarity3, angular second moment4	0.702
6.	Band4, Band3, NDVI, DVI, contrast1, variance1, contrast4, dissimilarity3, angular second moment, mean2	0.745
7.	Band4, Band3, NDVI, DVI, contrast1, variance1, contrast4, dissimilarity3, angular second moment, mean2, dissimilarity4	0.746

NDVI: normalized difference vegetation index; DVI: difference vegetation index.

Survey of India [30] for the same locality. The aboveground biomass (AGB) was calculated by multiplying species-wise tree volume by specific gravity [31], and further multiplying by biomass expansion factor [32]. For calculating belowground biomass (BGB), a root: shoot ratio of 0.30 was used for *Shorea robusta*, and 0.26 was used for other species [33]. Total biomass of the sample plot was worked out by adding AGB, BGB, shrub biomass, herb biomass and litter biomass. The biomass was converted into biomass per ha for each sample plot.

ANN-based modelling and multiple linear regression for predicting biomass

The MLP neural network algorithm was used in the present study. Biomass was considered the dependent variable and all the spectral and textural variables were considered independent variables. The ANN algorithm was used to optimize the independent variables with respect to the dependent variable. The ANN model was run in three stages, first using only spectral variables with biomass, second using only textural variables with biomass, and third by combining all of the spectral and textural variables with biomass. The model was run repeatedly for the selected independent variables for different activation functions and numbers of hidden layers until a satisfactory coefficient of determination (R^2) value was obtained. The model gave a ranking of the independent variables for predicting biomass. The ranking is essentially based on the contribution of each variable to predict the biomass. The ANN model was used to optimize the number of variables by retrieving the non-linear relationship among the variables. A multiple linear regression (MLR) was run between biomass and the top-ranked spectral and textural variables separately, and then both combined together, by sequentially increasing the number of variables in order of ranking until the R^2 value was stabilized. The MLR

model was selected as MLR is easy to implement, and the ANN regression model and MLR produce similar kinds of results with a limited number of variables.

Next, an empirical equation was established for biomass prediction using the coefficients for selected variables after reaching optimum R^2 and root mean square error (RMSE) value with a minimum set of independent variables. Finally, the accuracy assessment was done by comparing predicted biomass with the field-estimated biomass using data from 42 sample plots.

Results and discussion

The forest type-density map prepared for the study area is shown in Figure 5. There are altogether 13 forest and three non-forest classes. Forests covers 93% of the study area (Table 1), of which the 'sal very dense' class covered 53.11% of the area. The area statistics of different forest type-density classes were useful for the distribution of field sample plots for biomass assessment. From the field inventory data it was observed that the mean biomass ranged from 137.20 (scrub) to 808.05 Mg ha^{-1} (sal teak mixed very dense), indicating the presence of high forest biomass in the study area (Table 2).

The relationship between predicted and field-estimated biomass showed an R^2 value of 0.61. Independent variable importance analysis was done as it is crucial to determine the importance of the explanatory variables in estimating the target variable. Importance of variables can be normalized by subtracting the minimum value, and the subtracted value was then divided by the range of the importance of variables, and finally scaled up from 0 to 100. The normalized importance of different spectral variables is given in Figure 6. The top six variables were NDVI, SWIR band reflectance, TNDVI, SAVI, PCA2 and DVI. Next, ANN modelling was done involving all 32 texture variables, with one hidden layer.

Normalized Importance
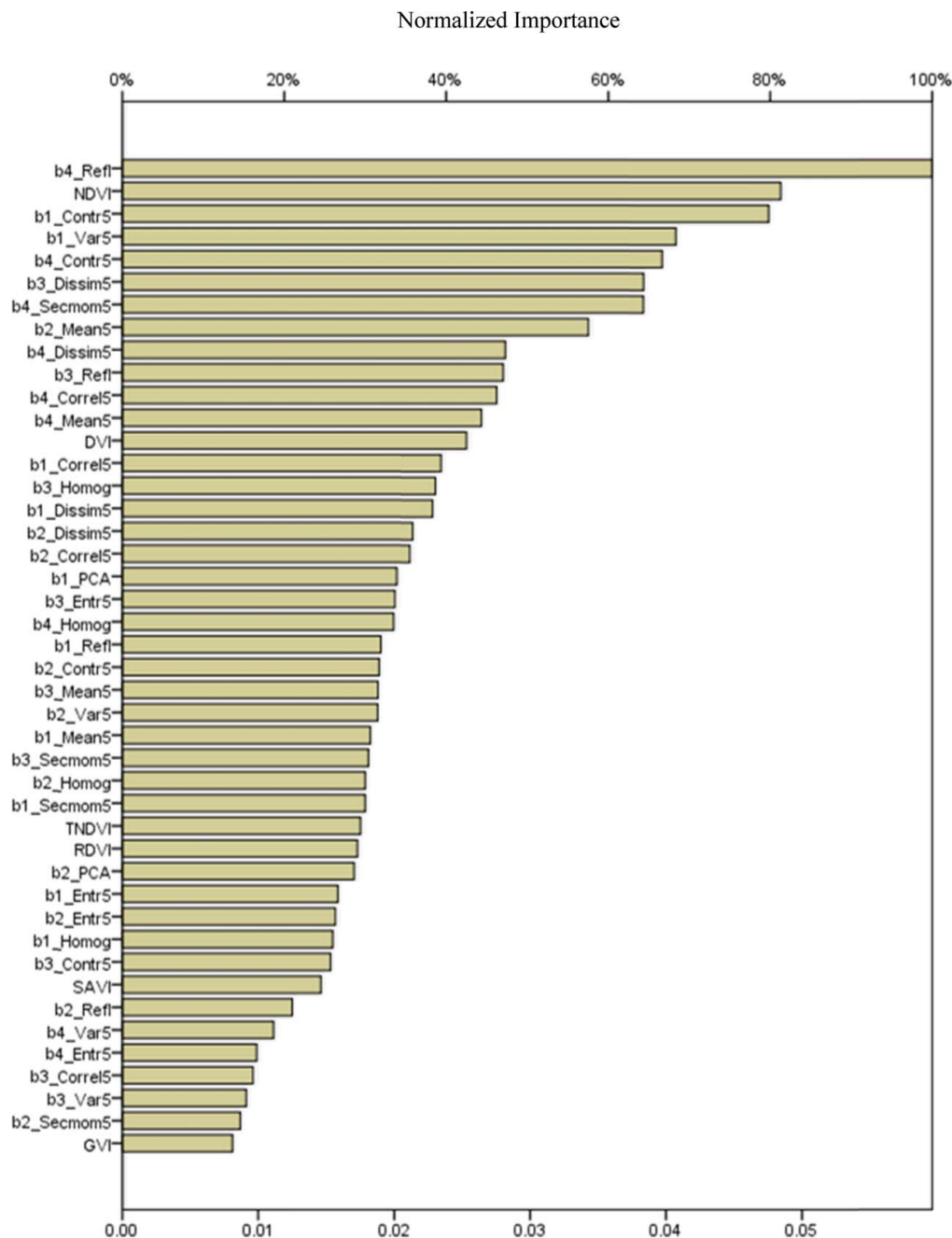
Figure 8. Normalized importance of spectral and texture variables calculated using an artificial neural network (ANN). [b4_Refl: Band 4 reflectance; NDVI: normalized difference vegetation index; b1_Contr5: Band 1 contrast; b1_Var5: Band 1 variance; b4_Contr5: Band 4 contrast; b3_Dissim5: Band 3 dissimilarity; b4_Secmom5: Band 4 angular second moment; b2_Mean5: Band 2 mean; b4_Dissim5: Band 4 dissimilarity; b3_Refl: Band 3 reflectance; b4_Correl5: Band 4 correlation; b4_Mean5: Band 4 mean; DVI: difference vegetation index; b1_Correl5: Band 1 correlation; b3_Homog: Band 3 homogeneity; b1_Dissim5: Band 1 dissimilarity; b2_Dissim5: Band 2 dissimilarity; b2_Correl5: Band 2 correlation; b1_PCA: first principal component; b3_Entr5: Band 3 entropy; b4_Homog: Band 4 homogeneity; b1_Refl: Band 1 reflectance; b2_Contr5: Band 2 contrast; b3_Mean5: Band 3 mean; b2_Var5: Band 2 variance; b1_Mean5: Band 1 mean; b3_Secmom5: Band 3 angular second moment; b2_Homog: Band 2 homogeneity; b1_Secmom5: Band 1 angular second moment; TNDVI: transformed normalized difference vegetation index; RDVI: re-normalized difference vegetation index; b2_PCA: second principal component; b1_Entr5: Band 1 entropy; b2_Entr5: Band 2 entropy; b1_Homog: Band 1 homogeneity; b3_Contr5: Band 3 contrast; SAVI: soil adjusted vegetation index; b2_Refl: Band 2 reflectance; b4_Var5: Band 4 variance; b4_Entr5: Band 4 entropy; b3_Correl5: Band 3 correlation; b3_Var5: Band 3 variance; b2_Secmom5: Band 2 angular second moment; GVI: green vegetation index]

The relationship between predicted and field-estimated biomass had an R^2 value of 0.87. The texture variable based on GLCM increases the accuracy of the classification and regression-based model [34,35]. Texture variables contain more information about the image structure which cannot be detected by spectral
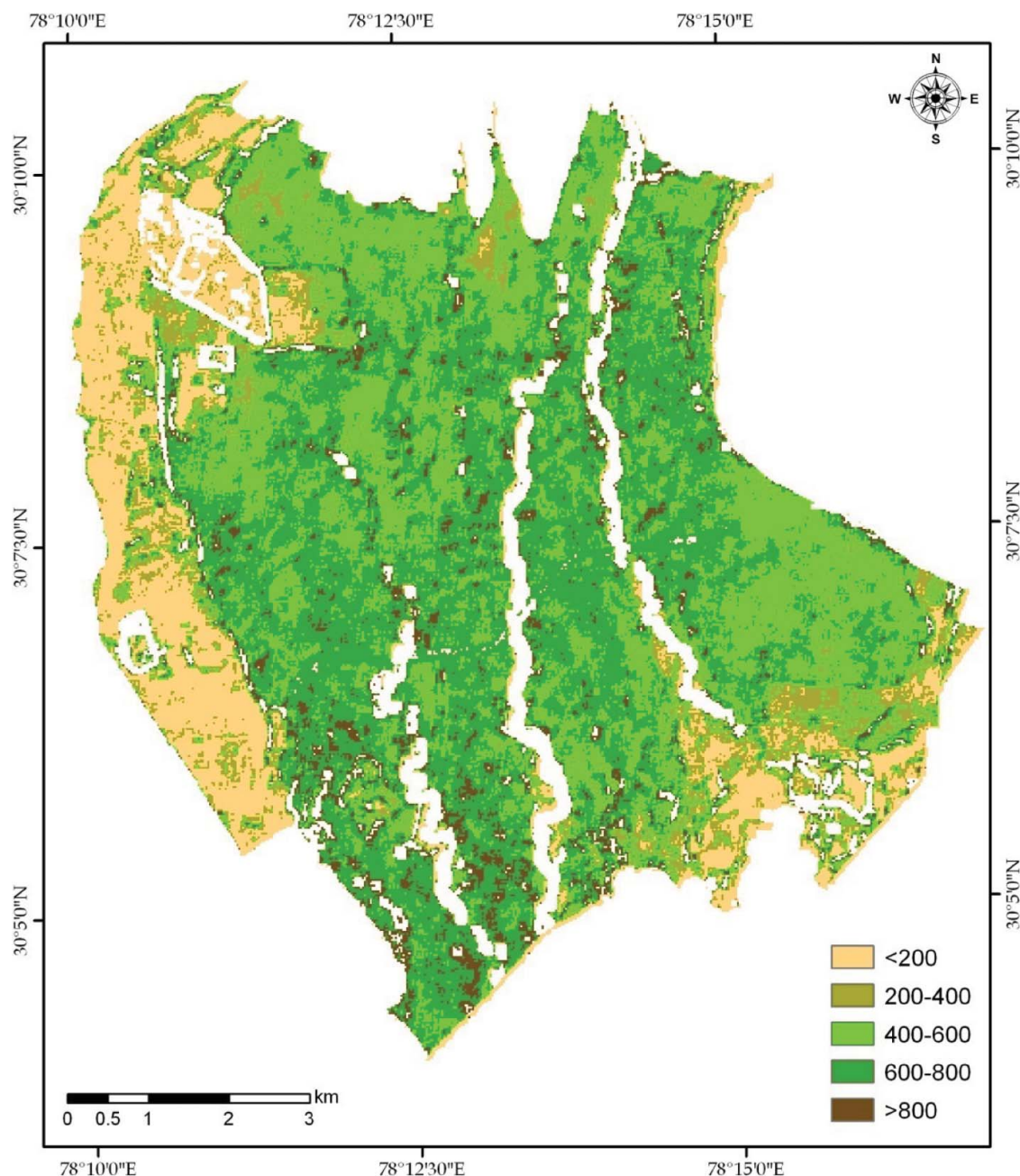
Figure 9. Spatial distribution of biomass (Mg ha^{-1}).

variables [34]. In the present study, inclusion of texture variables also improved the value of R^2 from 0.61 to 0.87. Zheng *et al.* [15] also found that the texture features were highly correlated with forest growing stock. Figure 7 shows the normalized importance of different textural variables. The top six variables were red band mean, SWIR band mean, green band contrast, SWIR band correlation, NIR band correlation and red band contrast. Then, ANN modelling was done involving all spectral and textural variables, with one hidden layer. The relationship between predicted and field biomass had an R^2 value of 0.92. Using 32 variables, we obtained $R^2 = 0.87$, and including another 13 variables (total 45 variables) increased the value of R^2 to 0.92 with RMSE

of 58.30 Mg ha^{-1}. This occurs as more variables can capture or detect more features which increases the accuracy. Figure 8 shows the normalized importance of different spectral and textural variables. The top six variables were SWIR band reflectance, NDVI, green band contrast, green band variance, SWIR band contrast and NIR band dissimilarity.

MLR was carried out with top spectral and texture variables for predicting biomass. Table 3 shows the variables involved in MLR, and the corresponding R^2 values. It can be seen that the R^2 stabilizes at 0.75 with RMSE of 85.32 Mg ha^{-1} for a combination of four spectral and six textural variables. The MLR equation (Equation 3) is given by:
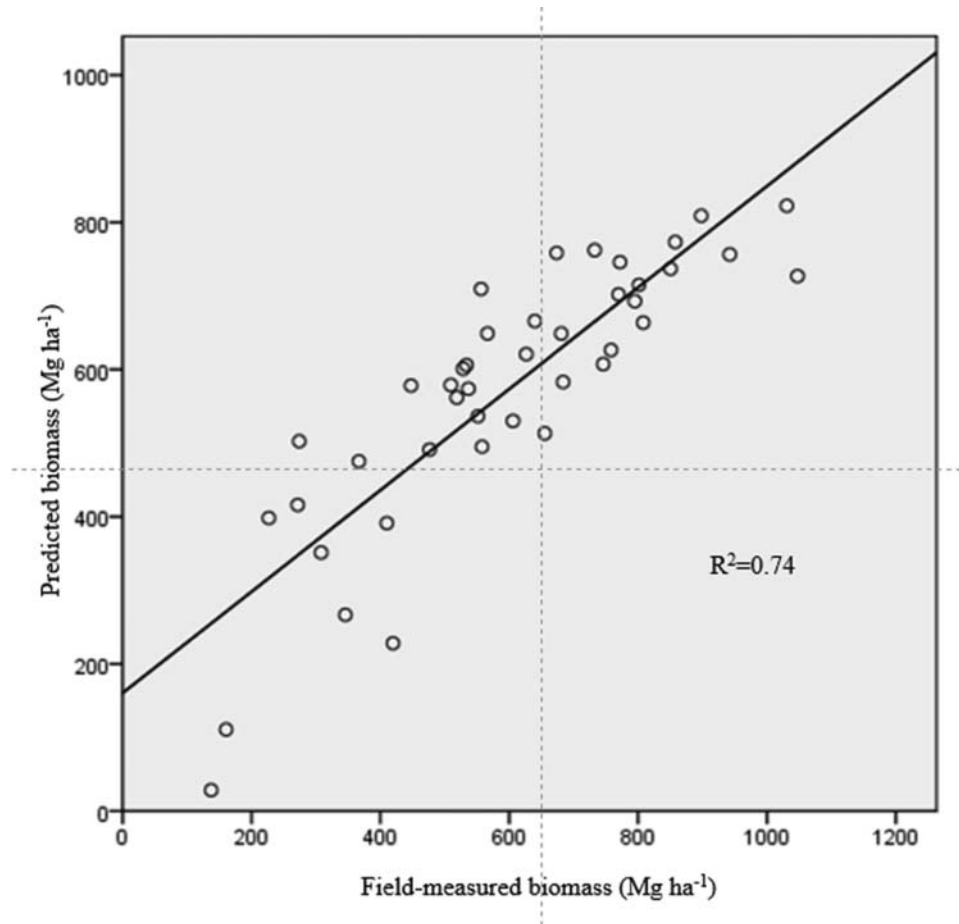
Figure 10. Field measured vs. predicted biomass.

Biomass $= 1811.890 - 80333.101 \times$ SWIR_Refl
$\qquad - 141.412 \times$ Green_Contr5 $- 374.411 \times$ NDVI
$\qquad + 450.507 \times$ Green_Var5 $+ 202726.766$
$\qquad \times$ NIR_Refl $+ 4.334 \times$ SWIR_Contr5 $- 22.879$
$\qquad \times$ DVI $+ 85.090 \times$ NIR_Dissim5 $- 284.773$
$\qquad \times$ SWIR_Secmom5 $- 138.614 \times$ Red_Mean5

$$(3)$$

By applying the above equation the spatial distribution of biomass in the study area was generated (Figure 9). The accuracy assessment was done by comparing predicted biomass with the field-estimated biomass. An R^2 value of 0.74 (Figure 10) was achieved with an RMSE of 93.41 Mg ha^{-1}.

It was observed that SWIR band was consistently present in the top variables. Yadav and Nandy [16] also found a stronger relationship between plot biomass and SWIR band than with red band in Timli, Uttarkhand, India (same forest type and near our study area). NIR band is extremely important for ecological study as vegetation reflects maximum energy in the NIR region. But NIR band does not provide any information about the soil under the vegetation. SWIR band has strong contrasts with soil. SWIR sensors can distinguish the moisture content of vegetation and soil. This happens as SWIR band positively correlates to vegetation

[16] and has a strong negative relationship with soil moisture [36]. We used LISS-III data for the month of April, and this period experiences yellowing and/or leaf fall (more soil is exposed) in the study area. Therefore, SWIR band captured the vegetation signal properly and was present in the top variables in the model.

The results obtained in this study are fairly comparable with those of a similar study on oak forest in Xiuyan in southeastern Liaoning, China, conducted using Landsat Thematic Mapper (TM) data [15]. They used a total of 53 spectral and texture variables in that study to predict the growing stock volume (GSV). With spectral variables, among the six bands, TM-1 had the highest correlation ($R^2 = 0.48$) with GSV and the lowest standard error of the estimate (3.10 m^3). After analyzing the correlation between variables and measured forest GSV, variables DVI, PVI and (TM-2)/(TM-4) were found to be significantly correlated with forest GSV. The correlation with texture variables showed that green band mean, red band mean, and NIR band mean were correlated with forest GSV ($R^2 > 0.45$). In an optimal multiple stepwise regression (MSR), green band mean, NIR band skewness, and PCA 3 were included and an R^2 of 0.80 was achieved, while in the accuracy assessment, an R^2 of 0.70 was achieved with RMSE of 2.3 m^3. Kelsey and Neff [37] evaluated regional

biomass maps generated using physical variables, spectral vegetation indices and image texture of Landsat TM imagery in the San Juan National Forest in southwest Colorado. The map with the highest prediction quality was created using three texture metrics calculated from Landsat Band 2 on a 3 × 3 window and an offset of [0,1]: entropy, mean and correlation; and one physical variable: slope. The correlation of predicted versus observed biomass values for texture-based biomass map is r = 0.86 and RMSE = 45.6 Mg ha^{-1}. These results are in agreement with our results. In another study, in-situ data collected with the Relasphone were successfully combined with optical satellite images to produce growing stock volume and aboveground biomass maps in Mexico [38]. Green and SWIR bands were found to be suitable in predicting aboveground biomass. In our study also, features derived from SWIR band were consistently present in the top variables, as this band highly correlates to vegetation.

Conclusions

Forest biomass is one of the most important indicators in the field for forest resource assessment and monitoring. The present study highlighted the usefulness of Indian Remote Sensing (IRS) LISS-III satellite data in combination with field inventory data for assessment of the spatial distribution of forest biomass. The study also emphasized the utility of neural network-based modelling for forest biomass assessment with a minimum number of spectral and textural variables of LISS-III data. ANN was used to identify the important spectral and texture variables for forest biomass prediction. A combination of the top 10 spectral and texture variables showed a strong relationship with forest biomass. The empirical equation developed using the top 10 variables was effectively used to assess the spatial distribution of forest biomass in the study area.

The present study area is almost flat and, hence, topographic variables were not considered. But one can use a combination of spectral, textural and topographical variables to predict forest biomass using the same approach in a hilly/mountainous area. Since spatial distribution of forest biomass is important for planning and management of forest, carbon accounting, global change monitoring and forest productivity modelling, there is a need for reliable methods of assessment and monitoring of forest biomass. As forests play an important role in the carbon cycle, variation in biomass quantity can be a good indicator of climatic change.

Acknowledgements

Disclosure statement

ORCID

Subrata Nandy http://orcid.org/0000-0003-4127-4035
Surajit Ghosh http://orcid.org/0000-0002-3928-2135

References

1. Anaya JA, Chuvieco E, Palacios-Orueta A. Aboveground biomass assessment in Colombia: a remote sensing approach. *For. Ecol. Manag.* 257(4), 1237–1246 (2009).
2. Zhao M, Zhou GS. Estimation of biomass and net primary productivity of major planted forests in China based on forest inventory data. *For. Ecol. Manag.* 207(3), 295–313 (2005).
3. Lu D, Chen Q, Wang G *et al*. Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. *Int. J. For. Res.* 21(2), 486–499 (2012).
4. Rosenqvist Å, Milne A, Lucas R, Imhoff M, Dobson C. A review of remote sensing technology in support of the Kyoto Protocol. *Environ. Sci. Policy.* 6(5), 441–455 (2003).
5. Goodale CL, Apps MJ, Birdsey RA *et al*. Forest carbon sinks in the Northern Hemisphere. *Ecol. Appl.* 12(3), 891–899 (2002).
6. Jenkins JC, Chojnacky DC, Heath LS, Birdsey RA. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1), 12–35 (2003).
7. Hall RJ, Skakun RS, Arsenault EJ, Case BS. Modeling forest stand structure attributes using Landsat ETM+ data: application to mapping of aboveground biomass and stand volume. *For. Ecol. Manag.* 225(1), 378–390 (2006).
8. Lu D, Mausel P, Brondízio E, Moran E. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. *For. Ecol. Manag.* 198(1), 149–167 (2004).
9. Lu D. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *Int. J. Remote. Sens.* 26 (12), 2509–2525 (2005).
10. Muukkonen P, Heiskanen J. Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: a possibility to verify carbon inventories. *Remote. Sens. Environ.* 107(4), 617–624 (2007).
11. Heyojoo BP, Nandy S. Estimation of above-ground phytomass and carbon in tree resources outside the forest (TROF): a geo-spatial approach. *BankoJanakari* 24(1), 34–40 (2014).
12. Kushwaha SPS, Nandy S, Gupta M. Growing stock and woody biomass assessment in Asola-Bhatti Wildlife Sanctuary, Delhi, India. *Environ. Monitor. Assess.* 186(9), 5911–5920 (2014).
13. Manna S, Nandy S, Chanda A, Akhand A, Hazra S, Dadhwal VK. Estimating aboveground biomass in Avicennia marina plantation in Indian Sundarbans using high-resolution satellite data. *J. Appl. Remote. Sens.* 8(1), 083638 (2014).

14. Lu D. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote. Sens.* 27(7), 1297–1328 (2006).

15. Zheng S, Cao C, Dang Y *et al*. Retrieval of forest growing stock volume by two different methods using Landsat TM images. *Int. J. Remote. Sens.* 35(1), 29–43 (2014).

16. Yadav BK, Nandy S. Mapping aboveground woody biomass using forest inventory, remote sensing and geostatistical techniques. *Environ. Monitoring Assess.* 187(5), 1–12 (2015).

17. Atkinson PM, Tatnall ARL. Introduction neural networks in remote sensing. *Int. J. Remote. Sens.* 18(4), 699–709 (1997).

18. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* 3(Mar), 1157–1182 (2003).

19. Mather P, Tso B. *Classification Methods for Remotely Sensed Data* (2nd Edition). CRC Press, New York (2009).

20. Cybenko G. Approximation by super positions of a sigmoidal function. *Math. Control. Signal. Syst.* 2(4), 303–314 (1989).

21. Champion HG, Seth SK. *A Revised Survey of the Forest Types of India*. Manager of Publications, Government of India, New Delhi (1968).

22. Rouse JW, Haas RH, Deering DW, *et al*. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. NASA/GSFC Type III, Final Report. NASA, Greenbelt, MD (1974)

23. Roujean JL, Breon FM. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* 51(3), 375–384 (1995).

24. Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150 (1979).

25. Eckert S. Improved forest biomass and carbon estimations using texture measures from WorldView-2satellite data. *Remote Sens*. 4(4), 810–829 (2012).

26. Huete AR. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25(3), 295–309 (1988).

27. Deering DW, Rouse JW, Haas RH, Schell JA. Measuring forage production of grazing units from Landsat MSS data. *Proceedings of the 10th International Symposium on Remote Sensing of Environment*. ERIM, Ann Arbor, Michigan, vol. 2, 1169–1178, 6–10 October (1975).

28. Haralick RM. Statistical and structural approaches to texture. *Proc. IEEE* 67(5), 786–804 (1979).

29. Chacko VJ. *A Manual on Sampling Techniques for Forest Surveys*. Manager of Publications, Government of India, New Delhi (1965).

30. FSI. *Volume Equations for Forests of India, Nepal and Bhutan*. Forest Survey of India, Ministry of Environment and Forests, Government of India, Dehradun (1996).

31. FRI. *Indian Woods: Their Identification, Properties and Uses*, Vol. I-VI (Revised Edition). Forest Research Institute, Indian Council of Forestry Research and Education, Ministry of Environment and Forests, Government of India, Dehradun (2002).

32. Haripriya GS. Estimates of biomass in Indian forests. *Biomass Bioenerg*. 19(4), 245–258(2000).

33. Negi JDS. *Biological productivity and cycling of nutrients in managed and man-made ecosystems [Ph.D. Thesis]*. Garhwal University, Srinagar, India (1984).

34. Salas EAL, Boykin KG, Valdez R. Multispectral and texture feature application in image-object analysis of summer vegetation in Eastern Tajikistan Pamirs. *Remote Sens*. 8 (1), 78 (2016).

35. Dhanda P, Nandy S, Kushwaha SPS, Ghosh S, Krishna Murthy YVN, Dadhwal VK. Optimising spaceborne LiDAR and very high resolution optical sensor parameters for biomass estimation at ICESat/GLAS footprint level using regression algorithms. *Prog. Phys. Geogr.* 41(3), 247–267, (2017).

36. Brown D, Jorgenson M, Kielland K, Verbyla DL, Prakash A, Koch, JC. Landscape effects of wildfire on permafrost distribution in interior Alaska derived from remote sensing. *Remote. Sens.* 8(8), 654 (2016).

37. Kelsey KC, Neff JC. Estimates of aboveground biomass from texture analysis of Landsat imagery. *Remote Sens*. 6 (7), 6407–6422 (2014).

38. Molinier M, Lopez-Sanchez CA, Toivanen T *et al*. Relasphone—mobile and participative in situ forest biomass measurements supporting satellite image mapping. *Remote Sens*. 8(10), 869 (2016).

Website

101. http://calval.cr.usgs.gov/documents/IRSP6.pdf