INTELLIGENT AUTOMATION & SOFT COMPUTING, 2017
http://dx.doi.org/10.1080/10798587.2017.1296660

Forest Above Ground Biomass Estimation from Remotely Sensed Imagery in the Mount Tai Area Using the RBF ANN Algorithm

Liang Wang[a,b], Jiping Liu[a,b], Shenghua Xu[b], Jinjin Dong[c] and Yi Yang[d]

[a]School of Resource and Environmental Science, Wuhan University, Wuhan Hubei, China; [b]Research Center of Government GIS, Chinese Academy of Surveying and Mapping, Beijing, China; [c]Shandong Building Materials Institute of Exploration and Survey, Shandong General Brigade of China National Geological Exploration Center of Building Materials Industry, Jinan, China; [d]School of Geomatics and Marine Information, Huaihai Institute of Technology, Lianyungang Jiangsu, China

ABSTRACT

Forest biomass is a significant indicator for substance accumulation and forest succession, and can provide valuable information for forest management and scientific planning. Accurate estimations of forest biomass at a fine resolution are important for a better understanding of the forest productivity and carbon cycling dynamics. In this study, considering the low efficiency and accuracy of the existing biomass estimation models for remote sensing data, Landsat 8 OLI imagery and field data cooperated with the radial basis function artificial neural network (RBF ANN) approach is used to estimate the forest Above Ground Biomass (AGB) in the Mount Tai area, Shandong Province of East China. The experimental results show that the RBF model produces a relatively accurate biomass estimation compared with multivariate linear regression (MLR), k-Nearest Neighbor (KNN), and backpropagation artificial neural network (BP ANN) models.

KEYWORDS

Remote-sensing-based estimation; AGB; RBF ANN; MIV

1. Introduction

Forests are primary components of terrestrial ecosystems. They are extremely rich in biodiversity and provide important ecosystem services, such as food, fiber, and water regulation. In addition to its role in reducing greenhouse gas emissions, reducing emissions from de-forestation in the developing countries provides an opportunity to value and safeguard these services. Mount Tai is a mountain of historical and cultural significance located at north of the city of Tai'an, in Shandong province, China. Since forest biomass is an important indicator for evaluating an ecosystem's productivity and is the basis for analyzing the substance circulation in the terrestrial ecosystem, an accurate estimation of forest AGB in the Mount Tai is the foundation of the terrestrial ecosystem carbon cycle and dynamic carbon analysis, which plays an important role in global carbon cycle research on various terrestrial ecosystems (Karin, Uwe, Viktor, & Florian, 2012; Vaglio Laurin et al., 2014).

The traditional methods for calculating biomass rely on considerable amounts of *in situ* measurements, which are inaccurate, time-consuming, costly, laborious and poorly distributed in large areas (Vaglio Laurin et al., 2016). With the rapid development of remote sensing techniques, remote-sensing-based methods have been widely used to predict the AGB over large areas; these methods are macro-scale, dynamic, fast, economical, and convenient for estimating large-scale forest AGB and long-term dynamic changes.

Over the past two decades, a large number of researchers have contributed to the study of the application of optical remotely sensed data (Ene et al., 2016). Based on the relative forest area within a MODIS pixel, Kumar et al. used four regression functions, i.e., linear, logarithmic, exponential and power functions to find the best model for estimating the AGB in northern Haryana, India (Kumar, Gupta, Singh, Patil, & Dhadhwal, 2011). Karlson et al. assessed the utility of Landsat 8 OLI imagery for mapping tree canopy cover and AGB in a Sudano-Sahelian woodland landscape (Martin et al., 2015). Dube and Mutanga examined the utility of the Landsat 8 multispectral sensor for estimating the AGB of forests in uMgeni catchment, KwaZulu Natal, South Africa as part of large scale monitoring to understand forest contribution to the regional carbon cycle (Dube & Mutanga, 2015). However, these models cannot explain the relevant physical mechanisms or the relationships between the parameters; thus, they may poorly predict the variable under study when the values are beyond the saturation point of the canopy reflectance (Bénié, Kaboré, Goïta, & Courel, 2005; Lumbres & Lee, 2014). Combining the satellite-derived, climatic, and topographic predictor variables with the Mexican National Forest Inventory (NFI) data, Aguirre-Salado et al. compared four variations in the distance metrics of the KNN for the spatially explicit estimation of the AGB in an area of the Mexican northern border within the intertropical zone (Aguirre-Salado et al., 2014). Seo et al. used the KNN method to estimate the AGB of a logged tropical forest in Sabah, Malaysia. The field data, the digital number and the normalized difference vegetation index (NDVI) from the Landsat TM-5 data were used to determine the optimum horizontal reference area and the number of reference sample plots (k) (Seo, Phua, Ong, Choi, & Lee, 2014). The advantage of KNN method is that no assumptions are made about the nature of the relationship between AGB and remote sensing (Seo et al., 2014). However, the KNN method requires large

CONTACT Shenghua Xu xushh@casm.ac.cn

numbers of widely distributed field plots that cover the complete range of values for the AGB, forest structure and remote sensing signals in the study area (Liu, Shen, Zhao, & Xu, 2013; Skowronski, Clark, Gallagher, Birdsey, & Hom, 2014). Based on the combined use of the Landsat Thematic Mapper (TM) and field measurements, linear regression, partial least squares (PLS) regression and BP ANN have been used to estimate the aboveground carbon (AGC) stock of Moso bamboo in Anji, Zhejiang Province, China. The results indicated that the Erf-BP model performed the best and the linear regression model performed the poorest (Xu et al., 2011). The textural measures were derived using wavelet analysis and Grey Level Co-occurrence Matrix methods, which were coupled with multispectral data to provide inputs to artificial neural networks trained under four different scenarios and validated using biomass measured from field plots. Stümer et al. utilized *in situ* national forest inventory data and satellite remote sensing data (Landsat 7 ETM+) to apply self-organizing map algorithms (neural networks) for the spatially explicit estimation of forest carbon stocks in a test region in Thuringia (Germany) (Stümer, Kenter, & Köhl, 2010). The strongest relationships between the predicted biomass and the measured biomass from the field survey were obtained with a neural network that was specifically developed for the site. Artificial neural networks are general-purpose computing tools that can solve complex non-linear problems and that can obtain highly accurate forest biomass.

Neural networks consist of a large class of different architectures. Multi-Layer Perceptron (MLP) and RBF are two of the most widely used neural network architectures for regression problems, such as biomass estimation. The RBF neural network structure is simple, adaptive, and less adjustable in terms of parameters, and it produces an irrelevant correlation between the output and initialized weights (Aguirre-Salado et al., 2014; Seo, Park, Yim, & Lee, 2012); thus, a remote-sensing-based estimation of AGB using the RBF ANN algorithm is proposed. This study presents a cross validation method to determine the optimal parameter value of the spread, to construct the optimal RBF neural network model and to combine the algorithm of the MIV for the selection of the variables.

The objective of this study, therefore, is to link biomass and field inventory data to Landsat 8 OLI data and to analyze their relationships and the potential of Landsat 8 OLI data for biomass estimation. The goal is to develop improved biomass estimation models, evaluating spectral information, textural information and terrain factors.

2. Materials

2.1. Study Area

The study area is located in Mount Tai in the center of Shandong Province. The area is approximately 24200 *ha*, extending from 36°11′N, 116°50′E to 36°31′N,117°12′E (Figure 1). Mount Tai has been listed on the World Cultural and Natural Heritage site list since 1985. The land cover is fragmented with small tourist areas, and the main areas are occupied by forests. Mount Tai is a heterogeneous region with complex topography, where the elevation ranges from 180 to 1530 *m*. The climatic factors are highly variable throughout the year, with the annual mean precipitation in the range of 750–1130 *mm*. These physical conditions are conducive for forest growth. Mount Tai is typical of the broad-leaved forest community in the warm temperate zone of China. Forest cover composes 81% and vegetation

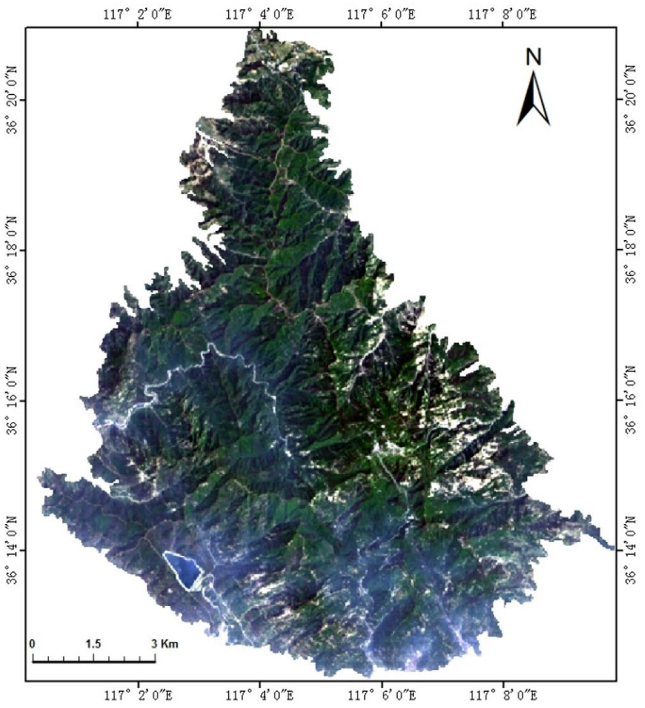

Figure 1. Study Area.

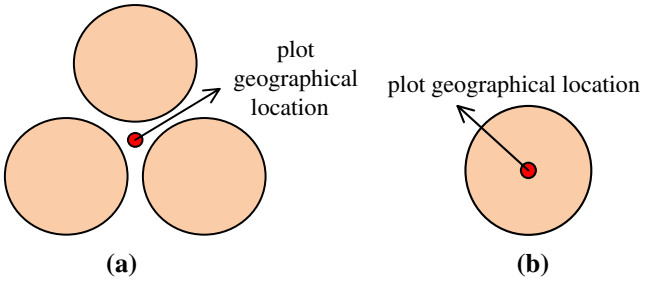

Figure 2. Sample Plots: (a) Clustered Circular Plot; (b) Circular Plot.

cover composes 90% of the total area. The dominant species include pine, oriental arborvitae, oak, Chinese scholar tree, poplar, and willow.

2.2. Field Inventory

The field campaign was performed from May 13 to May 18, 2013. This period was characterized by high biomass productivity. Random sampling was adopted in this study. The field data were collected at 48 temporary sample plots, which included 6 clustered circular plots (Figure 2a) and 42 circular plots (Figure 2b). These plots covered all of the major forest types in Mount Tai. The clustered circular plot was sampled via 3 sub-plots, where each sub-plot area was as large as 200 m^2. The circular plot had an area of 450 m^2.

The geographical locations of the plots were determined by non-differential GPS with a positional error of < 15*m* under a canopy. In each sample plot, the diameter of the individual living forest trees was measured (at breast height, 1.30*m* above ground, DBH), along with the tree height; each tree species was identified. The trees with DBH values below 5 *cm* were not included in the survey. The tree height was measured with a laser range finder (OPTI-LOGIC 400LH, Opti-Logic Corporation); for trees lower than 3 *m*, a leveling rod was

used. The DBH was measured at a height of 1.3 *m* using a tape measure. The latitude, longitude, and elevation of the forest site were measured by a GPS device (HOLUX M-241, Holux Technology Inc.).

2.3. Remote Sensing Data and DEM

During the field inventory time, a Landsat 8 OLI image (WRS-2, Path 122/Row 35) captured the entire study area on May 21, 2013. The Landsat 8 satellite payload consists of two scientific instruments; the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The OLI images consist of nine spectral bands with a spatial resolution of 30 *m* for Bands 1 to 7 and 9. New band 1 (ultra-blue) is useful for coastal and aerosol studies. New band 9 is useful for cirrus cloud detection. The resolution of Band 8 (panchromatic) is 15 *m*. The spectral ranges of the nine bands are 433–453 *nm* (B1-coastal aerosol), 450–515 *nm* (B2-blue), 525–600 *nm* (B3-green), 630–680 *nm* (B4-red), 845–885 *nm* (B5-near infrared), 1560–1660 *nm* (B6-SWIR 1), 2100–2300 *nm* (B7-SWIR 2), 500–680 *nm* (B8-panchromatic), and 1360–1390 *nm* (B9-cirrus).

DEM data covering the entire study area is from the 1:50,000 national fundamental geographic data base in China, which is based on 1985 National Height Datum on CGCS2000 ellipsoid.

3. Methodology

The proposed method is based on the RBF ANN model. First, feature extraction is adopted to extract the textural features and vegetation indices from the Landsat 8 OLI data. Second, to select the non-linear variables rapidly and accurately, the method of selecting biomass estimation model parameters for the neural network's mean impact value (MIV) is implemented based on the ideology of the MIV. Third, the RBF neural network model is constructed, and the radial based distribution function "spread" is determined by the cross-validation method. The simulation is applied to the study area of Mount Tai to estimate the local forest biomass.

3.1. Overview of the Method

The flow of the proposed method is shown in Figure 3. First, image pre-processing and feature extraction are adopted to extract the texture features and the vegetation indices from the Landsat 8 OLI data, and the terrain factors, such as the slope, aspect and height, are analyzed from the DEM. Second, for rapid and accurate selection of non-linear variables, the selection method for the biomass-estimation model parameters for the neural network's MIV is based on the ideology and advantages of the MIV, such as the training capabilities, the rapid convergence, and the adaptive fault tolerance of the neural network. Third, the RBF neural network model is constructed with a good approximation of the data. The RBF neural network is a three-layer network that is identified by the nodes of each layer. The hidden layer transfer function is a Gaussian function, and the output layer transfer function is a linear function; after repeated testing, the radial-based distribution function "spread" is determined by the cross-validation

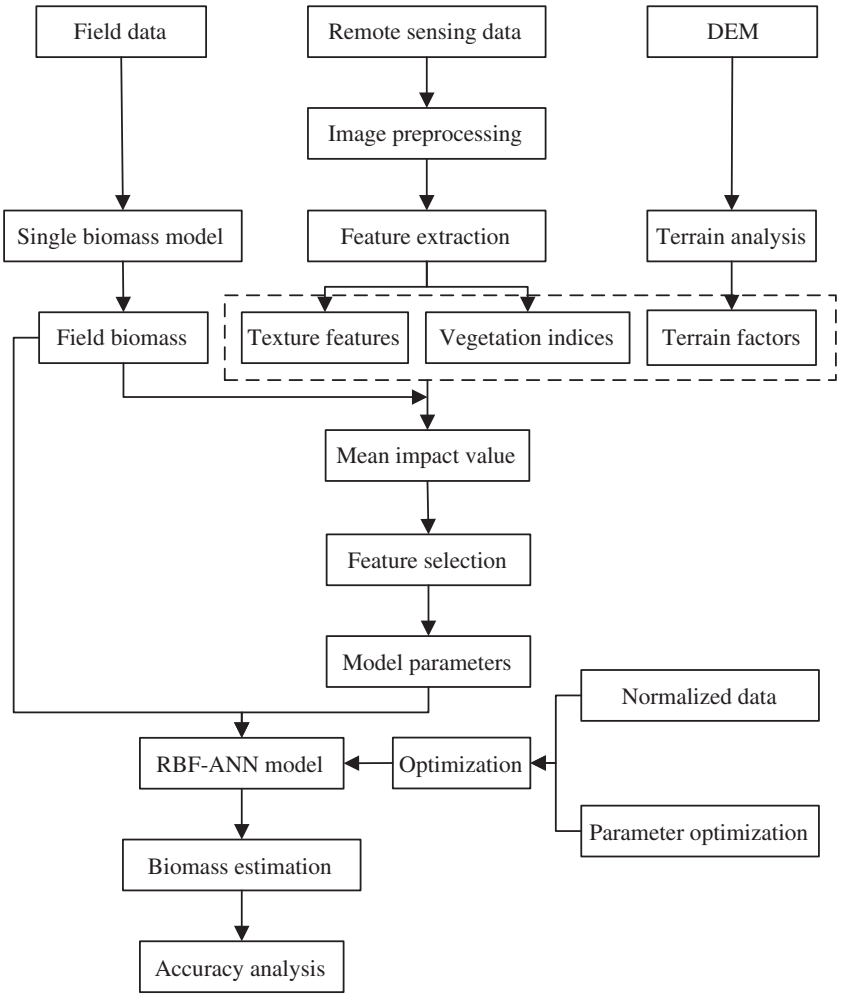

Figure 3. The Flow of the Proposed Method.

method. The RBF neural network is trained with the known field samples, and the simulation is applied to the study area to estimate the forest biomass.

3.2. Field Biomass Calculation

The calculation of the biomass of the individual trees is commonly based on the species, DBH and height. In our study, the forest-field AGB is calculated from equations based on the destructive techniques proposed by Fernández-Manso et al. (Fernández-Manso, Fernández-Manso, & Quintano, 2014), as follows:

$$W_{\text{field}} = e^a d^b c \tag{1}$$

$$c = e^{s^2/2} \tag{2}$$

where W_{field} is the AGB of the individual trees, a and b are the regression parameters, d is the diameter at breast height, s is the standard error of the estimation, and c is the correction factor, which enabled the calculation of W_{field} in all of the plots, e is the Euler number. Within each of the 48 plots, the AGB is calculated for each individual tree and is then averaged to a plot value. Subsequently, the AGB results are converted to tons per hectare (t/ha).

3.3. Satellite Image Preprocessing

The image pre-processing includes all of the common steps of multispectral remote sensing data treatment, such as radiometric, geometric and atmospheric corrections and conversion of digital numbers to reflectance values (Cutler, Boyd, Foody, & Vetrivel, 2012). For Landsat 8, a full computation of the reflectivity in each band is performed. As an initial step, the radiance values of Landsat 8 are computed from the DN values. Generally, the Landsat 8 OLI data are inappropriate for quantitative analyses and comparisons; therefore, a particular form of radiometric calibration or normalization is required. Consequently, an atmospheric correction is performed by the FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes) algorithm. Subsequently, the image is geometrically transformed and registered to the UTM projection to facilitate the linkage with the ground data. Based on the C-correction algorithm (Fan, Koukal, & Weisberg, 2014), the topographic correction is tested with the DEM, which is arranged as a regular grid of elevation points with 10 m spacing.

3.4. Selection of Variables

The biomass model parameters can be generally divided into the following four categories: Single band information, vegetation indices, texture features and terrain factors. There are eight individual bands [coastal aerosol, blue, green, red (R), near-infrared (NIR), cirrus, and two shortwave infrared (SWIR) bands]. The vegetation indices and the textural features are calculated from the individual bands as independent variables. The vegetation indices include the normalized difference vegetation index (NDVI), the difference vegetation index (DVI), the ratio vegetation index (RVI), the soil and atmospherically resistant vegetation index (SARVI), the transformed soil atmospherically resistant vegetation index (TSARVI), the multi vegetation index (MVI), and the perpendicular vegetation index (PVI). The entropy, contrast, homogeneity, means, correlation, variance, dissimilarity, correlation and second moment are the primary textural-feature statistical measures. All the spectral information is obtained by a single pixel format. In this study, the terrain factors primarily include the slope, aspect and height. The calculation of the terrain factors is based on the altitude values (grid cells) and the altitude value of the direct neighbors (typically, the 8 neighboring cells).

The selection of the variables plays an important role in the remote-sensing-based estimation of the AGB. Dombi et al. proposed an effective method, called MIV that reflects the changes in the weight matrix of the neural network as an evaluation of the various independent variables that influence the importance of the size of the dependent variable indicators. The MIV method can improve the accuracy of the proposed model (Dombi, Nandi, Saxe, Ledgerwood, & Lucas, 1995). The biomass model parameters of the field plots are selected as the initial training set. First, the training set is trained with the ANN model. Second, every explanatory variable is increased and decreased by a ratio to constitute two new training sets. Then, these new training sets and the trained neutral network are used to obtain the simulation results. The difference in the simulation values of the two training sets and their average values, which are the MIV values, are calculated. Finally, the variables are sorted based on their absolute values for the variable selection. The obvious variables of the neural network modeling results are selected as the network input parameters, and the other variables are removed. The selected biomass model parameters are shown in Table 1.

3.5. RBF ANN Algorithm

In 1985, radial-basis functions were introduced by Powell to solve the real multivariate interpolation problem (Musavi, Ahmed, Chan, Faris, & Hummels, 1992). Broomhead and Lowe were the first to use the radial-basis functions in the design of neural networks in 1988 (Webb & Lowe, 1990). The RBF neural network is based on supervised learning, which is designed as a solution to the approximation problem in multidimensional spaces. The RBF neural network is good at modeling nonlinear data, it can be trained in one stage rather than using an iterative process (as in MLP), and it can learn the given application quickly (Seo et al., 2014; Yilmaz & Kaynar, 2011). The output of the network is a linear combination of radial basis functions of the inputs and the neuron parameters. Radial basis function networks have many uses, including function approximation, time series prediction, egression, classification, and system control.

In this study, the Matlab 8.3 (2014) software is used in the analyses of a three-layer feed-forward neural network that consists of an input layer (7 neurons), one hidden layer with a non-linear RBF activation function (12 neurons for RBF) and a linear output layer (Figure 4). The seven neurons in the input layer are the variance in the NIR band, the variance in

Table 1. The MIV values of the selected variables.

No.	Model parameters	Absolute MIV value
1	Variance in the NIR band	2.93
2	Variance in the SWIR-2 band	1.60
3	Contrast in the NIR band	1.25
4	Homogeneity in the blue band	1.20
5	Entropy in the SWIR-1 band	1.14
6	Slope	1.10
7	Homogeneity in the SWIR-1 band	0.76

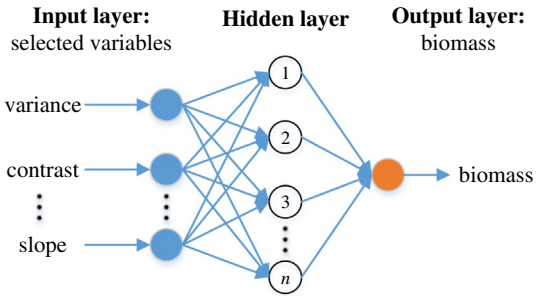

Figure 4. Architecture of an RBF ANN.

the SWIR-2 band, the contrast in the NIR band, the homogeneity in the blue band, the entropy in the SWIR-1 band, the slope, and the homogeneity of the SWIR-1 band. The neuron numbers in the hidden layers are selected from a series of trial runs of the networks with 1 neuron to 25 neurons to obtain the neuron number in the network with minimal error. The radius (also called the spread) of the radial basis function is selected by the cross validation method from a series of trial runs of the networks, which output values of 0.8 to 1.3 to obtain the spread in the network with minimal error.

This study uses training samples from the plot vegetation indices, textural features, and terrain factors as the input and the plot biomass as the output, using the newrb function in the MATLAB artificial neural network toolbox to train the network with the samples.

The radial basis network is iteratively created by the newrb function. Initially, the number of hidden-layer neurons is zero. Then, the hidden layer increases by one neuron in each iteration. The network simulates and determines the input sample vector corresponding to the maximum output error. Next, a neuron is added to the hidden layer, and the input vector is set as the weight vector. Finally, the weights of the linear layer are modified until the sum of the squared error is less than the error threshold or until a maximum number of neurons are reached. The newrb function calls the variables as follows:

$$net = newrb(P, T, goal, spread, MN, DF) \tag{3}$$

The default parameters are shown in Table 2.

Table 2. Default newrb parameter settings.

Parameter name	Parameter value	Description
P	7 × 48 matrix of plot parameters	Input vectors
T	1 × 48 matrix of plot biomass data	Target vectors
goal	0.0001	Mean squared error goal
spread	1	Spread of the radial basis functions
MN	7	Maximum number of neurons
DF	1	Number of neurons to add between displays

4. Results and Discussion

To evaluate the performance of the multivariate linear regression (MLR), KNN, and BP ANN models in predicting the AGB in Mount Tai, experiments and comparisons are performed as follows:

4.1. Validation and Assessment of the Prediction Techniques

To compare the different AGB estimation approaches, the following two aspects of the estimation accuracy are measured to control the model performance of the prediction capacity; the root-mean-square error (RMSE) and the mean absolute percentage error (MAPE).

The RMSE is a frequently used measure of the differences between the values estimated by a model and the observed values (Yilmaz & Kaynar, 2011). Generally, the RMSE represents the sample standard deviation of the differences between the estimated and observed values. The RMSE is defined according to the following equation:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2} \tag{4}$$

where n is the total number of the sample plots, y_i is the measured biomass at the ith plot, and y_i' is the predicted biomass at the ith plot.

In statistics, MAPE values, which are a measure of the accuracy of a method for constructing a fitted time series, are used for comparing the prediction performances of the models (Yilmaz & Kaynar, 2011). This value typically expresses accuracy as a percentage and is defined by the following formula:

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_i'}{y_i}\right| \tag{5}$$

Lower RMSE values and lower MAPE values indicate less residual variance. An excellent biomass estimation model should achieve a low RMSE and a low MAPE value.

We perform a leave-one-out cross-validation to obtain RMSE and MAPE values, where in each trial we left one participant's data out for training the method, and tested the resulting method on the left out participant.

4.2. Comparison of Different Values of the Spread

One of the most important parameters of the newrb function is the spread, which is the distribution density of the RBF function. The spread can greatly affect the performance of the network. A larger spread indicates a smoother function approximation. An excessively large spread indicates that many neurons are required to fit a fast-changing function. An excessively small spread indicates that many neurons are required to fit a smooth function, and the network might not generalize well. Therefore, the network must find the best value of the spread to reach the ideal accuracy during the network design. The spread is set to 0.8, 0.9, 1.0, 1.1, 1.2 and 1.3.

The network mean square error is 0.0001, the maximum number of neurons is 7, and the number of neurons to add between the displays is 1. As shown in Figure 5, a ratio shows the relative sizes of the training samples and the total samples. Figures 5a and 5b show that the red dotted line with a spread of 1.1 has the best performance, and the black solid line with a spread of 0.8 has the largest error. Therefore, this study chooses the optimal spread value of 1.1, which is used in the comparison of the four algorithms. Overall, as the ratio of the training

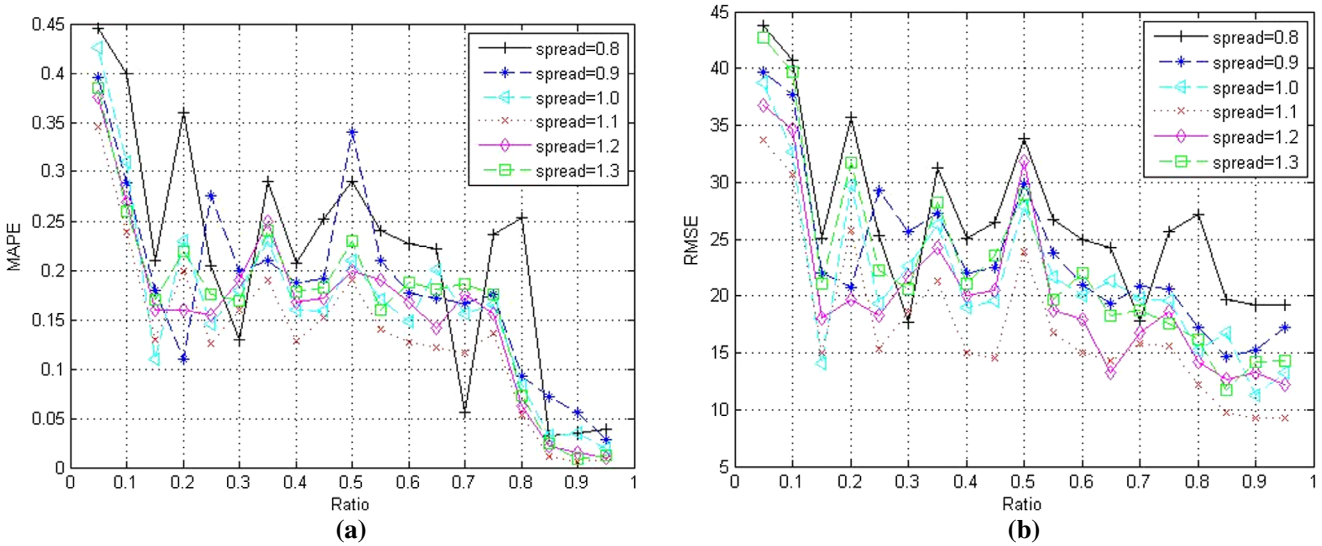

Figure 5. Error Curves with Different Spread Values: (a) MAPE; (b) RMSE.

Table 3. Accuracy of the four different methods.

Methods	MAPE (%)	R^2	RMSE
MLR	16.5	0.905	19.77
KNN	21.4	0.882	22.64
BP ANN	11.2	0.928	18.22
RBF ANN	9.5	0.936	9.87

data increases, there is a general downward trend in the MAPE and RMSE values.

4.3. Comparison of the Four Algorithms

In the following analyses, the optimal value for k is 8, and the optimal spread is 1.1. The BP ANN network parameters of the learning rate and the momentum are set to 0.01 and 0.95, respectively. The variable learning rate with the momentum (trainbfg) as the networks training function and tansig as an activation (transfer) function for all of the layers are used.

With the same sample-plot data-set, the biomass estimation model is run using the MLR, KNN, BP ANN and RBF ANN approaches. The RMSE, MAPE and coefficient of correlation (R^2) values are shown in Table 3. For the estimation of the forest biomass, the MAPE values of the estimated values are less than 16.5%, the RMSE values of the estimated values are less than 19.77, and the R^2 values of the estimated values are more than 0.882. Compared with the MLR, KNN and BP ANN approaches, the RBF ANN method has more applicable advantages: the MAPE is 9.5%, the RMSE is 9.87, and the R^2 is 0.936. Therefore, the RBF ANN method has significantly higher accuracy.

The above comparison of the RMSE, MAPE and R^2 values in predicting the AGB demonstrates that the prediction performance of the RBF ANN model is higher than that of BP ANN, MLR and KNN. Because the KNN model and the MLR model require field plot data distributions to address particular substitution (e.g., linearity, normality, equal variance, or independence), remote sensing information is insufficiently used; as a result, these models are less effective. In the BP ANN method, the weights are adjusted in the steepest descent direction (negative of the gradient). However, BP ANN has a slow learning convergent velocity, and it may be trapped within local minima. In addition, the performance of the BP ANN depends on the learning rate parameter and the complexity of the problem. The RBF network retains training speeds that are comparable to the BP networks, with the advantages of requiring a smaller number of units to cover high-dimensional input spaces and producing high approximation accuracy. In addition, the RBF networks have several advantages over the BP networks, including a faster learning rate and resistance to problems, such as paralysis and local minima. Therefore, the RBF ANN model exhibits more reliable predictions compared with the BP ANN model.

5. Conclusions

This study proposes a remote sensing-based estimation method for AGB in the Mount Tai area using the RBF ANN algorithm. In this study, the performance of the RBF ANN algorithm is largely influenced by the spread parameter. The RBF ANN with two layers and a spread of 1.1 is the best model for the biomass estimation by a test trial based on MAPE and RMSE. The experimental results clearly indicate that the RBF model possesses a relatively accurate biomass estimation compared with MLR, KNN and BP ANN. The results also show that the designed model addresses the characteristics of biomass estimation and accurately reflects the conditions of the AGB in Mount Tai. The combination of remote sensing and forest field-inventory data is a practical and effective technique for estimating the AGB distribution required for regional carbon stock assessments.

The limitations of the RBF model are that it is more sensitive to dimensionality and that it has greater difficulties when the number of units is large. Further studies are required to ascertain whether a mixed sigmoid–RBF network can effectively extrapolate to unknown biomass regions without losing the predictive accuracy within the known field biomass regions. Additional research is required to estimate forest biomass based on very-high-resolution satellite data.

Disclosure statement

No potential conflict of interest was reported by the authors

Funding

This research was funded by the national key research and development program [grant number 2016YFC0803101], National High Technology Research and Development Program of China (863 Program) under grant number 2013AA122003, the key laboratory of watershed ecology and geographical environment monitoring, National Administration of Surveying, Mapping and Geoinformation under [grant No. WE2016005], and the Basic Research Fund of CASM.

Notes on contributors

Liang Wang received a B.S. degree in Geography from Nanjing University, China, in June 1985 and M.S. degrees in in Cartography and GIS from Wuhan University, China, in June 2006,. His research interest includes e-government GIS, spatial analysis, and emergency geographic information services.

Jiping Liu received a B.S. degree in Department of Cartography from Wuhan Technical University of Surveying and Mapping, China, in June 1989, an M.S. degree in Computer Aided Cartography from Wuhan Technical University of Surveying and Mapping, China, in June 1992, and a Ph.D. degree in Cartography and GIS from the PLA Information Engineering University, China, in June 2004. His recent research topics include emergency geographic information service, geospatial big data analysis, E-government geographic information service, online geographic information monitoring and spatial decision-making.

Shenghua Xu received a B.S. degree in Information Engineering from Wuhan University, China, in June 2002 and a Ph.D. degree in State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing from Wuhan University, China, in July 2007. His research interests include artificial intelligent, remote sensing image processing and spatial analysis.

Jinjin Dong received a B.S. degree in College of Information Science and Engineering from Shandong Agricultural University, China, in June 2011 and an M.S. degree in Cartography and Geographic Information Engineering from Shandong Agricultural University, China, in July 2014. Her research interests include artificial intelligent and remote sensing image processing.

Yi Yang received a B.S. degree in College of Surveying and Geo-Informatics from Tongji University, China, in June 2012 and a Ph.D. degree in School of Resource and Environmental Sciences from Wuhan University, China, in December 2016. His research interests include artificial intelligent and spatial data mining.

References

Aguirre-Salado, C.A., Treviño-Garza, E.J., Aguirre-Calderón, O.A., Jiménez-Pérez, J., González-Tagle, M.A., Valdéz-Lazalde, J.R., & Miranda-Aragón, L. (2014). Mapping aboveground biomass by integrating geospatial and forest inventory data through a k-nearest neighbor strategy in North Central Mexico. *Journal of Arid Land, 6*, 80–96.

Bénié, G.B., Kaboré, S.S., Goïta, K., & Courel, M.F. (2005). Remote sensing-based spatio-temporal modeling to predict biomass in Sahelian grazing ecosystem. *Ecological Modelling, 184*, 341–354.

Dombi, G.W., Nandi, P., Saxe, J.M., Ledgerwood, A.M., & Lucas, C.E. (1995). Prediction of rib fracture injury outcome by an artificial neural network. *The Journal of Trauma: Injury, Infection, and Critical Care, 39*, 915–921.

Dube, T., & Mutanga, O. (2015). Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing, 101*, 36–46.

Ene, L.T., Næsset, E., Gobakken, T., Mauya, E.W., Bollandsås, O.M., Gregoire, T.G., & Zahabu, E. (2016). Large-scale estimation of aboveground biomass in miombo woodlands using airborne laser scanning and national forest inventory data. *Remote Sensing of Environment, 186*, 626–636.

Fan, Y., Koukal, T., & Weisberg, P. (2014). A sun-crown sensor model and adapted C-correction logic for topographic correction of high resolution forest imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 96*, 94–105.

Fernández-Manso, O., Fernández-Manso, A., & Quintano, C. (2014). Estimation of aboveground biomass in Mediterranean forests by statistical modelling of ASTER fraction images. *International Journal of Applied Earth Observation and Geo information, 31*, 45–56.

Karin, K., Uwe, B., Viktor, B., & Florian, S. (2012). Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan using LiDAR data. *International Journal of Applied Earth Observation and Geoinformation, 18*, 37–48.

Kumar, R., Gupta, S., Singh, S., Patil, P., & Dhadhwal, V. (2011). Spatial distribution of forest biomass using remote sensing and regression models in Northern Haryana, India. *International Journal of Ecology and Environmental Sciences, 37*, 37–47.

Liu, J., Shen, J., Zhao, R., & Xu, S. (2013). Extraction of individual tree crowns from airborne LiDAR data in human settlements. *Mathematical and Computer Modelling, 58*, 524–535.

Lumbres, R.I.C., & Lee, Y.J. (2014). Aboveground biomass mapping of La Trinidad forests in Benguet, Philippines, using landsat thematic mapper data and k-nearest neighbor method. *Forest Science and Technology, 10*, 104–111.

Martin, K., Madelene, O., Heather, R., Josias, S., Boalidioa, T., & Eskil, M. (2015). Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using Landsat 8 and random forest. *Remote Sensing, 7*, 10017–10041.

Cutler, M.E.J., Boyd, D.S., Foody, G.M., & Vetrivel, A. (2012). Estimating tropical forest biomass with a combination of SAR image texture and landsat TM data: An assessment of predictions between regions. *ISPRS Journal of Photogrammetry and Remote Sensing, 70*, 66–77.

Musavi, M.T., Ahmed, W., Chan, K.H., Faris, K.B., & Hummels, D.M. (1992). On the training of radial basis function classifiers. *Neural Networks, 5*, 595–603.

Seo, H., Park, D., Yim, J., & Lee, J. (2012). Assessment of forest biomass using k-Neighbor techniques-a case study in the research forest at Kangwon National University. *Journal of Korean Forestry Society, 101*, 547–557.

Seo, H., Phua, M., Ong, R., Choi, B., & Lee, J. (2014). Determining aboveground biomass of a forest reserve in Malaysian Borneo using k-nearest neighbour method. *Journal of Tropical Forest Science, 26*, 58–68.

Skowronski, N.S., Clark, K.L., Gallagher, M., Birdsey, R.A., & Hom, J.L. (2014). Airborne laser scanner-assisted estimation of aboveground biomass change in a temperate oak–pine forest. *Remote Sensing of Environment, 151*, 166–174.

Stümer, W., Kenter, B., & Köhl, M. (2010). Spatial interpolation of *in situ* data by self-organizing map algorithms (neural networks) for the assessment of carbon stocks in European forests. *Forest Ecology and Management, 260*, 287–293.

Vaglio Laurin, G., Chen, Q., Lindsell, J.A., Coomes, D.A., Frate, F.D., Guerriero, L., & Valentini, R. (2014). Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS Journal of Photogrammetry and Remote Sensing, 89*, 49–58.

Vaglio Laurin, G.V., Puletti, N., Chen, Q., Corona, P., Papale, D., & Valentini, R. (2016). Above ground biomass and tree species richness estimation with airborne lidar in tropical Ghana forests. *International Journal of Applied Earth Observation and Geoinformation, 52*, 371–379.

Webb, A.R., & Lowe, D. (1990). The optimised internal representation of multilayer classifier networks performs nonlinear discriminant analysis. *Neural Networks, 3*, 367–375.

Xu, X., Du, H., Zhou, G., Ge, H., Shi, Y., Zhou, Y., … Fan, W. (2011). Estimation of aboveground carbon stock of Moso bamboo (Phyllostachys heterocycla var. pubescens) forest with a landsat thematic mapper image. *International Journal of Remote Sensing, 32*, 1431–1448

Yilmaz, I., & Kaynar, O. (2011). Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. *Expert Systems with Applications, 38*, 5958–5966.