

Original Research Paper

Vol. 19

Open Access Journal

Machine Learning Model for Revealing the Characteristics of Soil Nutrients and Aboveground Biomass of Northeast Forest, China

di https://doi.org/10.46488/NEPT.2020.v19i02.005

Chunyan Wu*, Lifeng Pang**, Jun Jiang***†, Miaoying An*** and Yuanjun Yang****

*Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China

**Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China

***Research Center of Forest Management Engineering of State Forestry and Grassland Administration,

Beijing Forestry University, Beijing 100083, China

****College of Water Resources and Hydraulic Engineering, Yunnan Agricultural University, Kunming, China †Corresponding author: Jun Jiang; jiang@bjfu.edu.cn

Nat. Env. & Poll. Tech. Website: www.neptjournal.com

Received: 29-07-2019 Accepted: 08-10-2019

Key Words: Mt. Changbai; Soil properties; Species diversity; Aboveground biomass; Climate change

ABSTRACT

Declining soil quality and climate change may affect species diversity and forest biomass productivity in many temperate regions in the future. Our research objective is to reveal the characteristics of soil nutrients and biomass of forests in Northeast China with climate change. The purpose of this study was to determine the soil physical and chemical properties of mature broad-leaved forest in the cold temperate zone of Mt. Changbai, Jilin Province, by measuring pH, NH^{4 +}, organic matter (%), C/N, available phosphorus, alkali-hydrolysable N, rapidly available K, and Cr etc., analysing species diversity characteristics, and estimating aboveground biomass (AGB) of tree species with machine learning models. The results showed that with the increase of soil depth, the soil physical and chemical parameters have a decreasing trend; with the increase of soil depth, the soil nutrient content decreased; the main tree species were the Acer barbinerve (6937), Carpinus cordata BI. (6682) and Acer mandshuricum Maxim. (5447) etc. The total difference (SOR) showed a similar trend in the four directions and central point; the reference sample size at central point, north, west, south and east direction was 903, 954, 971, 1005 and 1016, respectively; GRNN model was the relatively best model among these models for modelling the aboveground biomass of the trees. Therefore, the diversity of tree species in north-eastern forests was affected by soil nutrients, climate change also has a significant impact on the aboveground biomass of northeast forests, which provides a theoretical basis for the management of northeast forests about soil physical and chemical properties and species diversity.

INTRODUCTION

Global climate change influences the distribution (Wang et al. 2017) and structure of forests (Peñuelas et al. 2016), the stability and balance of ecosystems (Hautier et al. 2015, Kilpeläinen et al. 2017), ecosystem restoration (Deng et al. 2017a, Deng et al. 2017b), the functional composition of trees, and biological accumulation in forests through changes in biodiversity (Van der Sande et al. 2017). The tree mortality induced by global drought and warming reveals the damage to forests caused by climate change (Kennedy et al. 2018). Forest biomass was identified as an important variable by the global climate observing system (GCOS) (Fu et al. 2017) to improve our knowledge of the forest ecosystems (Le Toan et al. 2011). Information about the impact of climate on soil physical and chemical properties can provide valuable insights to promote the natural regeneration of forests, respond to sustainability, and adapt to climate change (Kennedy et al. 2018). Therefore, understanding how the characteristic of soil physical and chemical properties of the forest has become particularly important.

The researchers have determined that soil nutrients play an important role in the soil fertility and environmental condition of plant growth and development (Delgado-Baquerizo et al. 2013, Camenzind et al. 2018). While, seed germination, seedling establishment, growth and reproduction during tree recruitment (Kulmatiski et al. 2008, Collin et al. 2018, Wang et al. 2018), the composition, development, spatial patterns, and geographical distribution of species (Schulz & Glaser 2012, Lu et al. 2017), community biomass and succession, and ecosystem stability are mainly affected by soil nutrients (Andrew et al. 2014, Laliberté et al. 2017). Also, increasing soil nutrient content will increase tree species richness (De Deyn et al. 2004, Cline et al. 2018). Most researches of soil nutrition have focused on the effects of alkaline nitrogen, available phosphorus, available potassium, and water content on plant growth and distribution (Mackay et al. 2017, Netzer

et al. 2017). They have determined that soil nutrient availability of nitrogen, phosphorus, and potassium strongly affect the ecosystem community structure (Sardans et al. 2017, Waldrop et al. 2017). Soil pH value, soluble ion, calcium carbonate content (Carter 1987), different condition of KNO₃ or NH₄NO₃ and soil water content have an impact on germination characteristics, leaf development, photosynthetic pigment content, and seedling photosynthetic efficiency. All of them have a significant correlation with species diversity of the forest (Lal et al. 2017).

Transpiration and reduced rainfall caused by global warming may threaten the vitality of trees, soil nutrient, species diversity and forest biomass. Process-based models have been used to simulate the impacts of climate change on forest biomass (Yousefpour et al. 2018, Kang et al. 2017). However, all of these studies only have different plant characteristics in different climatic regions or populations of a certain place. Although the model method has been developed and created with some new ideas, it is still relatively traditional as a whole (Stegen et al. 2011, Aherne et al. 2012). Besides, climate change is very complex. It is particularly important to explore a fast and efficient model for estimating and predicting precise biomass.

The use of ML models in agricultural industry and forestry has become more popular in classification and discrimination is one of the most important challenges in solving specific non-linear problems in dynamic change systems in the process of model prediction (Soares & Araújo 2016). Artificial neural network (ANN) and support vector machine (SVM) can solve this problem and have been widely used (Ottoy et al. 2017), which can represent the potential organic matter and complex process of climate affecting forest biomass. With the rapid development of machine learning, many advanced data processing technologies have included the combination method of data processing (GMDH), generalized regression neural network (GRNN), adaptive neuro-fuzzy inference system (ANFIS) (Dou & Yang 2017). These methods have been used to identify shell-free rice fungal populations (Chang et al. 2017), predict and map soil organic carbon stocks in landscapes (Were et al. 2015), and accurately predict biomass heating values (Akkaya 2016).

In this study, we used forest tree species data from permanent sample plots to study soil nutrient characteristics and species diversity characteristics and used machine learning-based data processing technology to estimate biomass of trees from forest comprehensive survey data to assess tree species response to climate and soil properties. Specifically, we ask the following questions: (1) what was the soil nutrient characteristics of Mt. Changbai? (2) what was the tree species diversity characteristics? (3) how was the relationship between climate factors and forest biomass?

MATERIALS AND METHODS

Learning Area

The study was conducted in Northeast China along a latitudinal gradient from 42.35° N to 43.57°N. The annual mean temperature in this region ranges between -2.1°C and 2.6°C, and the annual precipitation is between 510 mm and 810 mm. The broad-leaved Korean pine forest, a climax forest of the temperate zone, is characterized by the complex structure and abundant biodiversity. The study sites covered most of the latitudinal range of the broad-leaved Korean pine forest. The plot was near a mature broad-leaved Korean pine forest located in the cold temperate zone of Mt. Changbai, Jilin Province (Fig. 1). The broad-leaved Korean pine forest zone in Mt. Changbai is a typical zonal forest with the greatest number of animal and plant species and the most abundant vegetation on this mountain. It mainly grows at an altitude of 720-1100 m in the areas with a mild climate, heavy rainfall, and dark-brown forest soil. The main tree species include Abiesholophylla, Betulaplatyphylla, Carpinuscordata, Fraxinusmandshurica, Maackiaamurensis, Pinuskoraiensis, Populusussuriensis, Quercusmongolica, Tiliaamurensis, Ulmus japonica, and Ulmuslaciniata (Hu et al. 2014).

MATERIALS AND METHODS

Sampling plot setting: At the study sites described above, fixed sampling plots of 40 km² in size were established in 2015. Both the plots were composed of *Pinus koraiensis* and broad-leaved species such as *Acer* spp., *Fraxinus mandschurica*, and *Tilia* spp. They were divided into 1050 continuous sampling plots of 20 m × 20 m in size, and each of them was further divided into 5 sub-sampling plots of 5 m × 5 m in size in the east, south, west, north direction, and central point of the plot. In each plot, the name of an individual species, diameter at breast height (DBH), tree height, under-branch height, and crown width (east-west crown length, north-south crown length) were identified, measured and mapped for each plant with DBH > 1 cm. The overview of sampling plots is given in Table 1.

Soil data survey: Soil samples at three depths (0-10, 11-20, and 21-30 cm) were randomly collected from three random points using a stainless steel auger at each site. The soil was weighed to 200 g and returned to the laboratory before removing the residual artefacts. The samples were stored under dry, well-ventilated conditions to allow them to dry naturally before they were sieved through a 200-micron mesh, and then



Fig. 1: The sampling plots of the mature broad-leaved forest of Changbai Mountain in Jilin Province.

they were sent to the laboratory for physical and chemical analysis (Pingree & DeLuca 2018).

Climate data: Climate data in tropical China was obtained from the China Meteorological Data Network http://data. cma.cn/data/index/0b9164954813c573.html of the Meteorological Data Center of China Meteorological Administration. Influenced by global climate change, extreme weather (typhoons, tsunami natural disasters) is entering in a more frequent and more normalized trend, which is prone to extreme temperatures and sudden changes in rainfall. Therefore, five bioclimatic variables, average maximum temperature, average minimum temperature, average temperature, relative humidity and precipitation were chosen as independent variables to analyse the changes of biomass (Petus et al. 2018).

Biomass data: In this study, the diameter at breast height (DBH), height (H) and crown width of trees were investigated. One standard tree was selected from each plot for biomass determination. We cut down the tree and measured the fresh weights of the stems, branches, leaves and fruits to determine the water content in the field. The samples were brought to the laboratory and oven-dried at 75° C to a constant weight. The biomass of stems, branches, leaves and fruits was calculated using the dry weight to fresh

weight ratio, and the aboveground biomass (AGB) of the tree was obtained by summing all biomass values (Fu et al. 2017). The integral biomass dataset was randomly divided into three groups: training, validation and testing datasets. After several calculations, we found that 735 (70% of 1050 samples) trees for model training, 157 (15%) trees for model validation, and 157 (15%) trees for model testing produced the best results.

Models applications: Using climate variables estimate AGB of trees through machine learning modelling. In this study, the five machine learning algorithms include artificial neural network (ANN), support vector machine (SVM), generalized regression neural network (GRNN), adaptive neuro-fuzzy inference system (ANFIS) and the combination method of data processing (GMDH) (Dou & Yang 2017).

Data Analysis

Beta diversity analysis: Beta diversity measures changes in species composition between locations or communities (Baselga & Orme 2012). Baselga (2012) developed a unified framework for assessing beta diversity that uses the Sørensen or Jaccard indices and their turnover and nesting components to calculate the total differences. We use the Sørensen index:

Table 1: A summary of permanent forest plots in Mt.Changbai.

Forest type	Average elevation (m)	Total basal area (m ² km ⁻²)	Dominant species	Primary vegetation	
Primary Pinuskoraiensis-Tiliaamurensis mixed forests	779	25.36	Pinuskoraiensis Fraxinusmandshurica Tiliaamurensis Quercusmongolica	Broad-leaved Pinuskoraiensis forests	

Sørensen dissimilarity

$$\beta_{sor} = \frac{b+c}{2a+b+c} \qquad \dots (1)$$

turnover/replacement

$$\beta_{sim} = \frac{\min(b,c)}{a + \min(b,c)} \qquad \dots (2)$$

And nestedness/ richness difference

$$\beta_{sne} = \frac{|b-c|}{2a+b+c} \times \frac{a}{a+\min(b,c)} \qquad \dots (3)$$

Where, β_{sor} is the Sørensen dissimilarity, β_{sim} is the Simpson dissimilarity, β_{size} is the nestedness or richness difference, a is the number of species shared between two cells, b is the number of unique species at the poorest sites, and c is the number of unique species at the richest site.

Rarefaction and extrapolation curves with Hill numbers: Rarefaction and the extrapolation of Hill numbers are used for sampling and estimation in species diversity studies (Chao et al. 2014a). This criterion quantifies and assesses changes in biodiversity, allowing for the differential weighting of rare and abundant species, which is similar to a diversity index, but its meaning is easier to understand. Chao et al. (2014b) applied a uniform approach to sample and individual-based data to estimate the first three Hill numbers for characterizing the diversity of a species assemblage: species richness (q = 0), Shannon entropy index (Shannon diversity, q = 1) and the anti-Simpson concentration (Simpson diversity, q = 2). These proposed estimators are accurate for both sparse and short-range extrapolation.

Sparse and extrapolated curves and Hill numbers were used to compare the plant diversity patterns (Chao et al. 2014a). We estimated plant diversity (species richness and Shannon and Simpson diversity) as the mean of 200 replicates with a 95% confidence interval (Chao et al. 2014b).

Univariate analysis of variance (ANOVA) was used to test significant differences in soil physical and chemical parameters at different position points. All statistical analyses were implemented using R version 3.5.1 (R Development Core Team 2017) and GS+7.0. Beta diversity was analysed by betapart package, and sparse and extrapolated curves were compiled by iNEXT package. Draw numbers and process data using ggplot2 packages (R Development Core Team, 2017).

Data for Model Training and Validation

The input variables for all models are climate variables and diameter at breast-high. The five different ML modelling techniques presented in the previous section have been developed and compared. To achieve accurate and reasonable comparisons of different AGB estimates, all AGB values of trees were modelled using the same input variables (average minimum temperature, average maximum temperature, average temperature, average relative humidity, precipitation) and prediction methods. These input variables and corresponding output variables are normalized to a range between 0 and 1 before training each application model. The toolboxes in the MATLAB software (version 8.2, The MathWorks, Inc., Natick, MA, USA), including the Neural Network Toolbox 8.1 and the Fuzzy Logic Toolbox 2.2.18, were utilized for all models (Li et al. 2014).

Model Evaluation

The data processing performance of the developed model was evaluated based on several statistical data indices, including the coefficient of determination (R^2), Nash-Sutcliffe efficiency (E_{NS}), root mean squared error (RMSE), and mean absolute error (MAE) (Fu et al. 2017). The above statistical indicators are described as follows:

$$R^{2} = \left[\frac{\sum_{i=1}^{N} (Y_{a,c} - \overline{Y_{a}})(Y_{b,c} - \overline{Y_{b}})}{\sqrt{\sum_{i=1}^{N} (Y_{a,c} - \overline{Y_{a}})^{2} \sum_{i=1}^{N} (Y_{b,c} - \overline{Y_{b}})^{2}}}\right]^{2} \dots (6)$$
$$RMSE = \sqrt{\sum_{i=1}^{N} (Y_{a,c} - Y_{b,c})^{2} / N} \dots (7)$$

$$E_{NS} = 1 - \sum_{i=1}^{N} \left(Y_{a,c} - Y_{b,c} \right)^2 / \sum_{i=1}^{N} \left(Y_{a,c} - Y_b \right)^2$$
(8)

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |Y_{a,c} - Y_{b,c}| \qquad \dots (9)$$

Where, Y_a and Y_b denote the observed and modelled values, respectively; \overline{Y}_a and \overline{Y}_b are the means of the observed and modelled values, respectively; and N is the number of observed values.

RESULTS

Physical and Chemical Properties of Soil

The physical and chemical parameters of soils at Mt. Changbai are different in different soil depths (Table 2). Soil physical and chemical parameters show a decreasing trend with the increase of soil depth. The content of 0-10 cm soil depth was the highest, and 21-30 cm soil depth was the lowest in the three types of soil depth. The change ranges of pH value were relatively smaller than that of other indicators with the increase of the soil depth, the soil of Mt. Changbai was acidic and neutral in general. The content of NH⁴⁺ ranged from 0.4 to 0.08. The highest C/N ratio was 13.6. The content of available phosphorus was almost 10 mg kg⁻¹ in the 0-10 cm soil layer, and that of the alkali-hydrolysable N was almost 5 mg kg⁻¹ in the 21-30 cm soil layer, the rapidly available K was only 2.5 mg kg⁻¹ in the 0-10 cm soil layer.

Soil Nutrient Content

Soil nutrients also vary significantly with different locations and soil depths (Table 3). Soil nutrient content decreases as soil depth increases. In the 0-10 cm soil layer of Mt. Changbai, the contents of Cr, Mn, Co, Ni, Cu, Zn and Pb are 45.710, 746, 12.612, 26.531, 13.452, 130.638, 28.711, respectively, while the content of that was decreased sharply in 11-20 cm and the 21-30 cm soil layer.

Species Composition of the Trees

We identified 48 species of trees in the 49684 trees, belonging to 32 genera and 20 families in the Mt. Changbai. Among them, the number of the *Acer barbinerve* (6937), *Carpinus cordata* Bl. (6682), *Acer mandshuricum* Maxim. (5447), *Corylus mandshurica* Maxim. (5387), *Acer mono* Maxim. (4113), *Syringa reticulata* (Blume) H. Hara var. *amurensis* (Rupr.) J. S. Pringle (3026), *Ulmus laciniata* (Trautv.) Mayr. (2813), *Acer ukurunduense* Trautv. et Mey. (2058), *Tilia amurensis* Rupr. (1963), *Acer tegmentosum* Maxim. (1818), *Pinus koraiensis* Sieb. et Zucc. (1339), and *Sorbus alnifolia* (Sieb. et Zucc.) K. Koch (1065) was exceeded 1000 (Fig. 2). The 13 species, such as *Padus racemosa* (Linn.) Gilib., *Juglans mandshurica*, and *Quercus mongolica* Fisch. ex Ledeb, were more than 100 but less than 1000. The 23 species, such as *Betula platyphylla* Suk. and *Fraxinus rhynchophylla* Hance were less than 100.

Changes in the species composition of Mt. Changbai was estimated by beta diversity analysis (Fig. 3). The total difference (SOR) showed a similar trend in the four directions and central point: the greatest difference was the north direction, but with similar values at south direction and west direction, and the lowest difference was on the east direction. The turnover difference (SIM) showed a little different from the SOR in the four directions and central point: the lowest differences were in the south direction. The nesting difference (SNE) was significantly different in the east direction, with the central point being the largest followed by the south direction and the west direction, which were in the middle, and the north direction, which was the smallest.

Diversity Pattern of the Trees

The results showed that the reference sample size (i.e., the number of individual plants) at the central point, north, west, south and east direction was 903, 954, 971, 1005 and 1016, respectively, with Hill numbers q = 0, 1 and 2 curves (Fig. 4). The sample size was close to the complete recovery phase with a base coverage rate of 1.0. The reference sample size was extrapolated to 1746. The corresponding observed species richness (q = 0), Shannon diversity (q = 1) and Simpson diversity (q = 2) showed a consistent trend: gradually increased at the west and south directions but remaining stable

Table 2: Soil physical and chemical parameters (mean ± standard error) for each experimental site.

Forest type	layer	рН	NH ⁴⁺	Organic matter (%)	NO ³⁻	C/N	Available phos- phorus (mg kg ⁻¹)	Alkali-hydro- lysable N (mg kg ⁻¹)	Rapidly available K (mg kg ⁻¹)
Changbai	0-10cm	6.08±0.57a	0.371±0.121a	2.471±0.231a	25.301±2.748b	13.6b	10.514±2.215b	92.813±19.824a	56.763±11.434b
Mountain	11-20cm	5.62±0.46b	0.107±0.083b	1.057±0.115b	12.745±1.115b	10.4b	2.451±0.968b	17.836±8.610b	9.632±2.618b
	21-30cm	5.53±0.12bc	0.084±0.021bc	0.685±0.057c	7.163±0.476bc	9.7	1.424±0.043bc	5.725±3.764c	2.451±0.068bc

Note: the compartments with different letters represent the statistical differences of each parameter.

Table 3: Soil nutrient concentration (mg g^{-1} dry soil) at each experimental site (mean \pm standard error).

Forest type	Layer	$Cr (mg g^{-1})$	$Mn \ (mg \ g^{-1})$	Co (mg g^{-1})	Ni (mg g ⁻¹)	$Cu (mg g^{-1})$	$Zn \ (mg \ g^{-1})$	$Pb (mg g^{-1})$
Changbai	0-10cm	45.710a	746a	12.612a	26.531a	13.452a	130.638b	28.711a
Mountain	11-20cm	19.235b	324b	3.752b	11.673b	5.632b	86.452b	17.134b
	21-30cm	6.107b	105bc	1.146bc	4.650c	2.187c	24.951c	4.135c

Note: Compartments with different letters indicate statistical differences in each soil nutrient.





Note: BH: Betula platyphylla Suk.; BNQ: Acer mandshuricum Maxim.; BMDX: Syringa reticulata (Blume) H. Hara var. amurensis (Rupr.) J. S. Pringle; CXL: Salix koreensis Anderss.; CWM: Euonymus phellomana Loes.; CL: Padus racemosa (Linn.) Gilib.; CS: Abies nephrolepis (Trautv.ex Maxim) Mazim.; CY: Ulmus pumila 1.; CWJ: Acanthopanax senticosus (Rupr. Maxim.) Harms; CMQ: Acer barbinerve; DGY: Ulmus macrocarpa Hance; DBSMH: Philadelphus schrenkii Rupr.; DBSL: Rhamnus schneideri Levl.; DBSS: Deutzia parviflora Bge. var. amurensis Regel; FH: Betula costata Trautv.; HH: Betula davurica Pall.; HYT: Cerasus maximowiczii (Rupr.) Kom.; HS: Pinus koraiensis Sieb. et Zucc.; HTQ: Juglans mandshurica; HKQ: Acer ukurunduense Trautv. et Mey.; HQ: Sorbus pohuashanensis; HQL: Fraxinus rhynchophylla Hance; HH1: Maackia amurensis Rupr. et Maxim.; HB: Phellodendron amurense Rupr.; HHRD: Lonicera chrysantha Turcz.; JGM: Sambucus williamsii Hance; JYRD: Lonicera maackii (Rupr.) Maxim.; KD: Tilia mandshurica Rup et Maxim.; SMQ: Acer triflorum; QIY: Carpinus cordata BI: QKQ: Acer tegmentosum Maxim.; SMQ: Acer mono Maxim.; SS: Abies holophylla Maxim; SDZ: Malus baccata (L.) Borkh.; SPT: Vitis amurensis Rupr.; SY: Populus koreana Rehd.; CBRD: Lonicera ruprechtiana Regel; ZD: Tilia amurensis Rupr.; a: amount of species less than 100; b: number of species more than 1000.



Fig. 3: The beta diversity analysis of the density of the difference in species composition of trees on the forest in study sites.

Note: total dissimilarity (SOR) is divided into conversion difference (SIM) and nesting (SNE) components. Density estimation is performed by using a kernel smoothing method.

at the central point, diversity peaks at the east direction, and then decreased at the north direction.

Modelling Aboveground Biomass

Performance metrics, including R^2 , E_{NS} , RMSE and MAE, were used to evaluate the application model in predicting AGB efficiency. The larger the R^2 and E_{NS} values and the smaller RMSE and MAE values, the higher the model efficiency. The results of the estimation were used to predict the training of five data processing methods (ANN, GRNN, ANFIS, SVM and GMDH). The validation and testing times were summarized in Table 4. The table shows the R² and E_{NS} size order of model calculation results during training: ANFIS> GRNN> SVM> ANN> GMDH, RMSE and MAE values were in the same order as R² and E_{NS} ; and verification and testing (The R² and E_{NS} order of the model calculation results in the process of validation and testing: GRNN> SVM> ANN> ANFIS> GMDH, RMSE and MAE were in



Fig. 4: Individual- and coverage-based rarefaction and extrapolation curves, based on the Hill numbers (q = 0, 1, 2) for the vegetation on Changbai Mountain, Jilin province of China.

Note: the 95% confidence intervals (shaded area) were obtained by bootstrapping (200 copies). Reference samples are denoted by solid dots; the numbers in parentheses are the respective sample size and observed Hill number of each reference sample. In this study, we used the first three Hill numbers to characterize the diversity of a species assemblage based on the individual-based data: species richness (q = 0), the exponential of the Shannon entropy (Shannon diversity, q = 1), and the inverse Simpson concentration (Simpson diversity, q = 2). The proposed estimators are accurate for both rarefaction (solid line) and short-range extrapolation (dashed line, up to double the reference sample size).

the same order as R^2 and E_{NS}).

The AGB using the data model to compare measurements and predictions is shown in Fig. 5 as a scatter plot during the testing. As shown in Fig. 5, the fitted line of GRNN model (slope = 0.937) and SVM model (slope = 0.931) was closer to the ideal line (1:1 line). For the GMDH model, the fitted line (slope = 0.862) was more inconsistent with the ideal-fit line than the other models (ANN model (slope = 0.919), ANFIS model (slope = 0.902)). Besides, none of the estimated values of the GMDH model matches the minimum value around the corresponding observations. Although the applied models have significant AGB prediction performance differences (Table 4), these models provide a similar overall dispersion estimate (Fig. 5). Therefore, the GRNN model was the relatively best model among these models for modelling the AGB of the trees.

DISCUSSION

Soil texture has many effects on forest tree species. Our study shows that there are significant differences in soil physical and chemical properties between different geographical locations and soil depths. This is similar to previous studies by experts. Different authors have demonstrated that soil nutrients are the main factors affecting the composition and structure of forest communities (Valentín-Vargas et al. 2014). Besides, soil texture and topographic factors may change solar incidence angle, slope change, soil organic carbon content, temperature and humidity conditions (Brabcová et



Fig. 5: The AGB measured and predicted by models for the testing period.

Note: (a) ANFIS model; (b) GMDH model; (c) ANN model; (d) SVM model; and (e) GRNN model. AGB is above-ground biomass of trees; ANFIS refer to adaptive neuro-fuzzy inference systems; GMDH refers to group method of data handling; ANN refers to artificial neural network; SVM refers to support vector machine; GRNN refers to generalized regression neural networks.

Model	Training				Validation				Testing			
	R^2	E _{NS}	RMSE	MAE	\mathbb{R}^2	E _{NS}	RMSE	MAE	R^2	E _{NS}	RMSE	MAE
GMDH	0.8909	0.7912	0.9917	0.6733	0.8142	0.6612	0.9990	0.9564	0.8087	0.6661	0.9293	0.8553
GRNN	0.9273	0.8776	0.6361	0.4022	0.8748	0.7684	0.7512	0.4661	0.8792	0.7858	0.6727	0.4199
ANFIS	0.9331	0.8820	0.5484	0.3580	0.8478	0.7574	0.9130	0.9017	0.8528	0.7720	0.8066	0.7516
SVM	0.9150	0.8773	0.8277	0.4956	0.8747	0.7676	0.9039	0.6198	0.8765	0.7774	0.7742	0.5387
ANN	0.8993	0.8674	0.9127	0.6655	0.8735	0.7656	0.9086	0.8295	0.8757	0.7754	0.7987	0.7217

Table 4: Comparisons of models for AGB for the training, validation and testing periods.

Note: GMDH refer to group method of data handling; GRNN refer to generalized regression neural networks; ANFIS refer to adaptive neuro-fuzzy inference systems; SVM refer to support vector machine; ANN refer to artificial neural network; R^2 is coefficient of determination; E_{NS} is nash-sutcliffe efficiency; RMSE is mean squared error; MAE is mean absolute error.

al. 2018). The advanced succession of plant communities can significantly change soil water content, soil compactness, improve soil aeration and water holding capacity, and promote the development of soil fertility. The improvement of soil physical properties is also conducive to the process of community species replacement and succession. Soil properties of successive communities were influenced by slope and parent material of sample plots (Zhang et al. 1997). The composition of tree species was also affected by soil properties, such as the C/N ratio, soil organic carbon (Jiang 2016). Different tree species have different soil quality (Yang et al. 2008). The longer the forest age, the higher the soil water content. It shows that tree species are the main factors affecting soil quality. Tree species significantly improve the soil total nitrogen and soil organic matter content (Jiang 2016). Soil texture of parent rock materials and topography is almost the same in each sample plot in the study area, and the differences related to soil properties may be related to the characteristics of vegetation on the ground.

The trend of soil nutrient storage is obvious, and a large quantity of soil nutrients are stored in the surface layer (0-10 cm). Also, there was no significant difference in soil nutrient status between Mt. Changbai, and there was no significant correlation between soil organic matter and nutrients. This result is similar to previous studies, in which different soil nutrient concentrations decrease with increasing soil depth (Berger et al. 2015). In addition, the accumulation of NH_4^+ , soil organic matter, NO³⁻, C/N, available phosphorus and Cr, Mn, Co, Ni, Cu, Zn, Pb were significantly different in different soil depths. The nutrient accumulation in the surface soil layer led to the increase of primary productivity of community plants, while the nutrient in the lower soil layer was significantly less. Vegetation, defoliation, wood debris and dead roots are closely related to soil microorganisms and organic matter.

Due to the low level of soil nutrients and the decline

of plant species diversity (Deng et al. 2018), soil nutrient dynamics should be taken into account in any study on the changes of plant species diversity after regional vegetation restoration. Previous studies have shown that changes in species diversity and spatial distribution of most tree species after vegetation restoration are closely related to combinations of multiple factors, including soil nutrients (Pingree & DeLuca 2018). Also, the ratio of soil nutrient supply is a factor that determines the change of plant species diversity in forest communities (Deng et al. 2018). However, functional diversity of nutrient strategies may lead to higher plant species diversity in barren soils (Pingree & DeLuca 2018). Species distribution of tree communities can respond to individual soil nutrients and seasonal drought (Deng et al. 2018). Soil nutrients also play an important role in the nutrient cycle of forest ecosystems and have a great impact on plant species, age and density. The response of soil nutrients to vegetation activities mainly depends on vegetation types, species diversity and restoration activities (Ye et al. 2019). Therefore, soil nutrient supply has an important impact on some important aspects of plant community structure, and soil nutrient has an important impact on plant diversity.

The effects of species diversity on productivity are affected by factors such as resource supply rate (Wang et al. 2008). Total biomass and species richness were negatively correlated. Diversity tended to increase with the increase of soil nutrients (Wang et al. 2008). The topsoil nutrient content of the community was higher than that of deep soil, and the Simpson diversity index of community species was only related to soil potassium (Tan et al. 2017). Therefore, there was a significant correlation between soil nutrients and richness index and diversity index in species diversity.

The results showed that the variables related to temperature, humidity, and precipitation are of great significance in explaining the changes in AGB. Increasing temperatures and new patterns of precipitation due to climate change result in changes in the morphology, physiology, growth, reproduction, and biology of trees, all of which affect AGB (Stegen et al. 2011). In addition, enhancing the use efficiency of water and light resources could increase forest productivity (Van der Sande et al. 2017). Additionally, water deficiency has the most severe effect on the process of biomass formation (Van der Sande et al. 2017). Climate change is leading to increasing global temperatures and droughts in many areas, which have changed the nutrients in vegetation, AGB and belowground biomass (Peñuelas et al. 2016). Rainfall is highly correlated with the increase in tree stem diameter, which affects the accumulation of AGB, while rainfall together with temperature affects the AGB (Ogaya & Peñuelas 2007).

The AGB values of trees were estimated by the model, which showed that the average value for trees was 90% (Table 4). Therefore, these ML techniques are capable of describing the relationships between climate change and the AGB of trees. These results are similar to previous studies that found that ML-based data processing techniques are effective at modelling and predicting forest biomass (Were et al. 2015, Akkaya 2016). The best model was the ANN, which predicted the AGB with higher accuracy than the allometric equation. The ANFIS-based model is an effective technique for obtaining high-precision predictions of biomass with high heating value (Akkaya 2016). Furthermore, ANN and ANFIS can be used to establish a dynamic behavioural model (Were et al. 2015). Additionally, the models we proposed have achieved higher accuracy in testing compared with the above results. In summary, our current proposed models have great potential for estimating AGB. These models can be considered alternative tools for the traditional AGB models across regional or global scales for different vegetation types.

According to the performance of the various AGB models in this study, no model consistently outperformed the others for all AGB estimates. The GMDH model consistently produced relatively worse simulation results compared with the other models. This result may be partly due to the inherent limitations of the GMDH model, such as the choice of input parameters, multicollinearity and overfitting (Dou & Yang 2017). In addition, the effects of random errors in AGB observations on different time scales (e.g., daily, monthly, and yearly) should be taken into consideration when assessing the differences in predictions (Dou & Yang 2017). Systematic errors can also increase the uncertainty of the AGB estimates. Moreover, AGB is strongly influenced by the complex interactions between photosynthesis and photosynthetic efficiency (Huntingford et al. 2008). The selection of effective driving variables as model inputs could be used to further improve the predictive power of the model, especially for the AGB, and this is of significant importance for future work.

CONCLUSION

Our results show that the diversity of tree species in north-eastern forests is affected by soil nutrients. Also, climate change has a significant impact on the aboveground biomass of forests in northeast China. It shows that the high quality soil in the forest environment is a long-term process. Besides, it is difficult to interpret the field data of forest soil nutrient storage, especially when several factors affect the nutrient cycle. The soil type of Mt. Changbai is dark brown soil, and the climate type is warm and humid. There is a certain relationship between soil nutrient level and climatic conditions: in a hot and humid climate, surface vegetation is abundant, soil humus sources are more natural and fertile. In low temperature or arid climate, surface vegetation is difficult to grow, and soil humus sources are less natural than soil nutrients.

ACKNOWLEDGEMENTS

This study was conducted under the financial supports of The National Key Research and Development Program of China (2017YFC0504106), and the National Natural Science Foundation of China, China (31901306).

REFERENCES

- Aherne, J., Posch, M., Forsius, M., Lehtonen, A. and Härkönen, K. 2012. Impacts of forest biomass removal on soil nutrient status under climate change: A catchment-based modelling study for Finland. Biogeochemistry, 107(1-3): 471-488.
- Akkaya, E. 2016. ANFIS based prediction model for biomass heating value using proximate analysis components. Fuel, 80: 687-693.
- Andrew, K., Anderson-Smith, A., Beard, K.H., Doucette-Riise, S., Mazzacavallo, M., Nolan, N.E., Anderson-Smith, A., Mazzacavallo, M., Ramirez, R.A. and Stevens J.R. 2014. Most soil trophic guilds increase plant growth. A meta-analytical review. Oikos, 123: 1409-1419.
- Baselga, A. and Orme, C. 2012. Betapart: An R package for the study of beta diversity. Methods Ecol. Evol., 3: 808-812.
- Baselga, A. 2012. The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. Glob. Ecol. Biogeogr., 21: 1223-1232.
- Berger, T.W., Duboc, O., Djukic, I., Tatzber, M., Gerbazek, M.H. and Zehetner, F. 2015. Decomposition of beech (*Fagus sylvatica*) and pine (*Pinus nigra*) litter along an alpine elevation gradient: Decay and nutrient release. Geoderma, 251-252: 92-104.
- Brabcová, V., Tursová, M. and Baldrian, P. 2018. Nutrient content affects the turnover of fungal biomass in forest topsoil and the composition of associated microbial communities. Soil Biol. Biochem., 118: 187-198.
- Camenzind, T., Hättenschwiler, S., Treseder, K.K., Lehmann, A. and Rillig, M.C. 2018. Nutrient limitation of soil microbial processes in tropical forests. Ecol. Monogr., 88: 4-21.
- Carter, M.R. 1987. Physical properties of some Prince Edward island soils in relation to their tillage requirement and suitability for direct drilling. Can. J. Soil Sci., 67: 473-487.
- Chang, N.B., Bai, K.X. and Chen, C.F. 2017. Integrating multisensor satellite data merging and image reconstruction in support of machine learning for better water quality management. J. Environ. Manag., 201: 227-240.

- Chao, A., Chiu, C.H. and Jost, L. 2014a. Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. Annu. Rev. Ecol. Evol. Syst., 45: 297-324.
- Chao, A., Gotelli, N.J., Hsieh, T.C., Sander, E.L., Ma, K.H., Colwell, R.K. and Ellison, A.M. 2014b. Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. Ecol. Monogr., 84: 45-67.
- Cline, L.C., Hobbie, S.E., Madritch, M.D., Buyarski, C.R., Tilman, D. and Cavender-Bares, J.M. 2018. Resource availability underlies the plant-fungal diversity relationship in a grassland ecosystem. Ecology, 99: 204-216.
- Collin, A., Messier, C., Kembel, S.W. and Bélanger, N. 2018. Can sugar maple establish into the boreal forest? Insights from seedlings under various canopies in southern Quebec. Ecosphere, 9: 2022-2040.
- De Deyn, G.B., Raaijmakers, C.E. and Van Der Putten, W.H. 2004. Plant community development is affected by nutrients and soil biota. J. Ecol., 92: 824-834.
- Delgado-Baquerizo, M., Maestre, F.T., Gallardo, A., Bowker, M.A., Wallenstein, M.D., Quero, J.L., Ochoa, V., Gozalo, B., García-Gómez, M., Soliveres, S. and García-Palacios, P. 2013. Decoupling of soil nutrient cycles as a function of aridity in global drylands. Nature, 502: 672-676.
- Deng, L., Han, Q.S., Zhang, C., Tang, Z.S. and Shangguan, Z.P. 2017a. Above-ground and below-ground biomass accumulation and carbon sequestration with *Caragana korshinskii* Kom plantation development. Land Degradation & Development, 28: 906-917.
- Deng, L., Kim, D.G., Peng, C.H. and Shangguan, Z.P. 2018. Controls of soil and aggregate-associated organic carbon variations following natural vegetation restoration on the Loess Plateau in China. Land Degrad. Dev., 29: 3974-3984.
- Deng, L., Liu, S., Kim, D.G., Peng, C., Sweeney, S. and Shangguan, Z 2017b. Past and future carbon sequestration benefits of China's grain for green program. Global Environmental Change-Human and Policy Dimensions, 47: 13-20.
- Dou, X.M. and Yang, Y.G. 2017. Modelling and predicting carbon and water fluxes using data-driven techniques in a forest ecosystem. Forests, 8: 498.
- Fu, L., Lei, X., Hu, Z., Zeng, W., Tang, S., Marshall, P., Cao, L., Song, X., Yu, L. and Liang, J. 2017. Integrating regional climate change into allometric equations for estimating tree aboveground biomass of Masson pine in China. Annals of Forest Science, 74: 42-57.
- Hautier, Y., Tilman, D., Isbell, F., Seabloom, E.W., Borer, E.T. and Reich, P.B. 2015. Anthropogenic environmental changes affect ecosystem stability via biodiversity. Science, 6232: 336-340.
- Hu, Y.S., Yao, X.Y. and Liu, Y.H.N. 2014. Stoichiometric traits of plant and soil in different forest succession stages in Changbai Mountain. Chinese J. Applied Ecol., 25(3): 632-638. (In Chinese)
- Huntingford, C., Fisher, R.A., Mercado, L., Booth, B.B., Sitch, S., Harris, P.P., Cox, P.M., Jones, C.D., Betts, R.A., Malhi, Y. and Harris, G.R. 2008. Towards quantifying uncertainty in predictions of Amazon "dieback". Philosophical Transactions of the Royal Society B: Biological Sciences, 363: 1857-1864.
- Jiang, F. 2016. Spatial heterogeneity of soil organic carbon and total nitrogen in deciduous broad-leaved forest and evergreen broad-leaved forest in central subtropical region. Central South Univer. Forestry Sci. Technol., 5: 132-139. (In Chinese)
- Kang, H., Seely, B., Wang, G., Cai, Y., Innes, J., Zheng, D. and Chen P.L. 2017. Simulating the impact of climate change on the growth of Chinese fir plantations in Fujian province, China. New Zealand Journal of Forestry Science, 47(1): 20.
- Kennedy, R.E., Ohmann, J., Gregory, M., Roberts, H., Yang, Z., Bell, D.M., Kane, V.H., Joseph Cohen, M.J., Warren, P., Scott, N., Neeti, L., Tara, H., Sam, K., Jonathan, T. M., David, L.P., James, B.J. and Rupert, S.

2018. An empirical, integrated forest biomass monitoring system. Environ. Res. Lett., 13: 25004.

- Kilpeläinen, A., Strandman, H., Grönholm, T., Ikonen, V.P., Torssonen, P., Kellomäki, S. and Peltola, H. 2017. Effects of initial age structure of managed Norway spruce forest area on net climate impact of using forest biomass for energy. Bioenergy Research, 10: 499-508.
- Kulmatiski, A., Beard, K.H., Stevens, J.R. and Cobbold, S.M. 2008. Plantsoil feedbacks: A meta-analytical review. Eco. Lett., 11: 980-992.
- Lal, M.K. and Pandey, R. 2017. Interactive effect of elevated (CO₂) on biomass and carbohydrate partitioning under phosphorus stress. Int. J. Curr. Microbiol. App. Sci., 6: 2966-2977.
- Laliberté, B., Wardle, D.A. and Klironomos, J. 2017. Soil fertility shapes belowground food webs across a regional climate gradient. Eco. Lett., 20: 1273-1284.
- Le Toan, T., Quegan, S., Davidson, M.W.J., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H. and Ulander, L. 2011. The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. Remote Sensing of Environment, 115: 2850-2860.
- Li, H., Leng, W., Zhou, Y.B., Chen, F.D., Xiu, Z.L. and Yang, D.Z. 2014. Evaluation models for soil nutrient based on support vector machine and artificial neural networks. Sci. World J., ID:478569.
- Lu, K., Yang, X., Gielen, G., Bolan, N., Ok, Y.S., Niazi, N.K., Xu, S., Yuan, G., Chen, X., Zhang, X. and Liu, D. 2017. Effect of bamboo and rice straw biochars on the mobility and redistribution of heavy metals (Cd, Cu, Pb and Zn) in contaminated soil. J. Environ. Manag., 186: 285-292.
- Mackay, J.E., Cavagnaro, T.R., Stöver, D.S.M., Macdonald, L.M., Grønlund, M. and Jakobsen, I. 2017. A key role for arbuscular mycorrhiza in plant acquisition of P from sewage sludge recycled to soil. Soil Biol. Biochem., 115: 11-20.
- Netzer, F., Schmid, C., Herschbach, C. and Rennenberg, H. 2017. Phosphorus-nutrition of European beech (*Fagus sylvatica* L) during annual growth depends on tree age and P-availability in the soil. Environ. Experi. Bot., 137: 194-207.
- Ogaya, R. and Peñuelas, J. 2007. Tree growth, mortality, and above-ground biomass accumulation in a holm oak forest under a five-year experimental field drought. Plant Ecol., 189: 291-299.
- Ottoy, S., De Vos, B., Sindayihebura, A., Hermy, M. and Van Orshoven, J. 2017. Assessing soil organic carbon stocks under current and potential forest cover using digital soil mapping and spatial generalisation. Ecological Indicators, 77: 139-150.
- Peñuelas, J., Sardans, J., Filella, I., Estiarte, M., Llusià, J., Ogaya, R., Carnicer, J., Bartrons, M., Rivas-Ubach, A., Grau, O. and Peguero, G. 2016. Assessment of the impacts of climate change on Mediterranean terrestrial ecosytems based on data from field experiments and long-term monitored field gradients in Catalonia. Environmental and Experimental Botany, 152: 49-59.
- Petus, C., Devlin, M., da Silva, E.T., Lewis, S., Waterhouse, J., Wenger, A., Bainbridge, Z. and Tracey, D. 2018. Defining wet season water quality target concentrations for ecosystem conservation using empirical light attenuation models: A case study in the Great Barrier Reef (Australia). J. Environ. Manag., 213: 451-466.
- Pingree, M.R.A. and DeLuca, T.H. 2018. The influence of fire history on soil nutrients and vegetation cover in mixed severity fire regime forests of the eastern Olympic Peninsula, Washington, USA. Forest. Ecol. Manag., 422: 95-107.
- R. Development Core Team 2017. A Language and Environment for Statistical Computing Vienna. R Foundation for Statistical Computing.
- Sardans, J., Bartrons, M., Margalef, O., Gargallo-Garriga, A., Janssens, I.A., Ciais, P., Obersteiner, M., Sigurdsson, B.D., Chen, H.Y. and Peñuelas, J. 2017. Plant invasion is associated with higher plant-soil nutrient concentrations in nutrient-poor environments. Glob. Change Biol., 23: 1282-1291.

- Schulz, H. and Glaser, B. 2012. Effects of biochar compared to organic and inorganic fertilizers on soil quality and plant growth in a greenhouse experiment. J. Plant Nutrition Soil Sci., 175: 410-422.
- Soares, S.G. and Araújo, R. 2016. An adaptive ensemble of on-line extreme learning machines with variable forgetting factor for dynamic system prediction. Neurocomputing, 171: 693-707.
- Stegen, J.C., Swenson, N.G., Enquist, B.J., White, E.P., Phillips, O.L., Jørgensen, P.M., Weiser, M.D., Monteagudo Mendoza, A. and Núñez Vargas, P. 2011. Variation in above-ground forest biomass across broad climatic gradients. Global Ecology and Biogeography, 20: 744-754.
- Tan, S.D., Yang, Y.T., Chen, B., Wen, L. and Liu, H.B. 2017. Analysis of correlation between soil nutrients and species diversity in typical communities of Dawei Mountain. J. Zhejiang Agricul. Sci., 5: 887-891. (In Chinese)
- Valentín-Vargas, A., Root, R.A., Neilson, J.W., Chorover, J. and Maier, R.M. 2014. Environmental factors influencing the structural dynamics of soil microbial communities during assisted phytostabilization of acid-generating mine tailings: A Mesocosm Experiment. Sci. Total Environ., 0: 314-324.
- van der Sande, M.T., Peña-Claros, M., Ascarrunz, N., Arets, E.J., Licona, J.C., Toledo, M. and Poorter, L. 2017. Abiotic and biotic drivers of biomass change in a Neotropical forest. J. Ecol., 105: 1223-1234.
- Waldrop, M.P., Holloway, J.M., Smith, D.B., Goldhaber, M.B., Drenovsky, R.E., Scow, K.M., Dick, R., Howard, D., Wylie, B. and Grace, J.B. 2017. The interacting roles of climate, soils, and plant production on soil microbial communities at a continental scale. Ecology, 98: 1957-1967.

Wang, C.T., Long, R.J., Cao, G.M., Wang, Q.L., Jing, Z.C. and Shi, J.J.

2008. The relationship between soil nutrients and diversity productivity of different type grasslands in Alpine Meadow. Chinese J. Soil Sci., 39(1): 1-8. (In Chinese)

- Wang, J., Sun, Z., Hui, D., Yang, L., Wang, F., Liu, N. and Ren, H. 2018. Responses of seedling performance to altered seasonal precipitation in a secondary tropical forest, southern China. For. Ecol. Manag., 410: 27-34.
- Wang, W.J., He, H.S., Thompson, I.I.I.F.R., Fraser, J.S. and Dijak, W.D. 2017. Changes in forest biomass and tree species distribution under climate change in the north-eastern United States. Landscape Ecology, 32: 1399-1413.
- Were, K., Bui, D.T., Dick, Ø.B. and Singh, B.R. 2015. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an afromontane landscape. Ecol. Indic., 52: 394-403.
- Yang, X., Cao, J., Dong, M.X. and Ma, X.J. 2008. Effects of external tree species *Larix japonica* on forest soil quality and bacterial diversity. Chinese J. Applied Ecol., 10: 2109-2116. (In Chinese)
- Ye, L.P., Fang, L.C., Shi, Z.H., Deng, L. and Tan, W.F. 2019. Spatio-temporal dynamics of soil moisture driven by 'Grain for Green' program on the Loess Plateau, China. Agr. Ecosyst. Environ., 269: 204-214.
- Yousefpour, R., Augustynczik, A.L.D., Reyer, C.P.O., Lasch-Born, P., Suckow, F. and Hanewinkel, M. 2018. Realizing mitigation efficiency of European commercial forests by climate smart forestry. Scientific Reports, 8: 1-11.
- Zhang, Q.F., You, W.H. and Song, Y.C. 1997. Influence of plant community succession on soil physical properties in Tiantong Forest Park, Zhejiang Province. J. Plant Resour. and Environ., 6(2): 36-40. (In Chinese)