Paper ID: 769
Theme 1. Grassland resources
Sub-theme 1.1. Dynamics of grassland resources – global database

Estimating above ground tree biomass of semi-arid Bundelkhand region using satellite data, regression modelling and ANN technique

Dibyendu Deb*, J.P. Singh, R.S. Chaurasia
ICAR-Indian Grassland and Fodder Research Institute, Jhansi-284003, India
*Corresponding author e-mail: debiasri@gmail.com

Keywords: ANN technique, Bundelkhand region, Regression modelling, Satellite data, Tree biomass

Introduction

Remotely sensed satellite imagery can be used to classify and monitor vegetation dynamics (Tucker, 1979). The Normalized Difference Vegetation Index (NDVI) computed from satellite data is a good measure of photosynthetic activity at landscape scales, and can be used to estimate vegetation biomass and Net Primary Production (NPP) (Tucker, 1979; Myneni *et al*., 1995, Nemani *et al*., 2003). As in the present environmental condition the climate change has adversely affected the ecosystem and the forest cover, NDVI has an important role to play to track and quantify the change taking place in plant ecosystem process (Myneni *et al*., 1995, Nemani *et al*., 2003). Biomass estimation using NDVI is easy to implement, harvesting of trees is not required and also effectively reduce the time and cost required in case of any other estimation process. In this study regression models and ANN (Artificial Neural Network) model were tried to simulate and predict total biomass production from different districts of Bundelkhand region. NDVI values collected from remote sensing images at particular season of the year to characterize above ground biomass of the study area. The performance of the ANN model was compared with several other commonly used linear and nonlinear models and validation was done based on the model's stability.

Materials and Methods

In the present study IRS P6 LISS III (Row 98, Path 53) data dated February 2012 was selected. After the geometric and radiometric correction and histogram enhancement the image was classified using Erdas Imagine Professional s/w ver. 9.3. Finally information on above ground biomass production was generated in ArcGIS ArcInfo Workstation ver10.1. To eliminate the effect of bare soil or non-vegetation canopy surface and atmospheric contamination data with NDVI>0.1 were computed. Those pixels which have no vegetation covers were eliminated computing percent vegetation map with a resolution 36 m, using vegetation classification map. The image was then used for extraction of spectral band values from different sample sites representing the Central Highland physiographic zone of Madhya Pradesh and Uttar Pradesh.

For fitting regression models, the first step was to graphically explore the relation between different vegetative indices and the total biomass per plot obtained as effect and response variables respectively and then different linear and nonlinear models were tried. Eighty percent of the data were selected by random sampling and allocated for model building and the rest 20% for the validation set.
In this study power/allometric and exponential models were tried to find out their applicability after log transformation of the data. The ANN model was also used to simulate and estimate the biomass production of trees grown in different districts. The R software (version 3.1.3) was used to build the model. The feed forward multi-layer perceptron model (MLP) (Bishop, 1995) neural network was built for learning which are well applicable when modelling functional relationships. It calculates the function as defined by Gunther and Fritsch, 2010. The best type of model was chosen based on some reliability statistics including root mean square error (RMSE) and Akaike information criterion (AIC).

Results and Discussion

In order to detect a correlation between the predictor variable (NDVI) and the response variable (biomass) was used the Pearson's product-moment correlation test. This procedure tests the null hypothesis that the correlation between the variables is zero against the alternative hypothesis that this correlation is greater than zero.

Table 1: Results of hypothesis test about correlation between NDVI and Biomass

t-value	Degrees of freedom	p-value	Estimate	CI 95%
13.4532	65	$2.2e^{-16}$	0.72	0.6778 -0.8103

According to the p-value shown in table 1 (lower than 0.01), reject the null hypothesis. The conclusion is that the true correlation is not equal to 0. The correlation estimate is 0.7247, showing potential for fitting the regression model.
Amongst nonlinear models, allometric and exponential models were applied. The fit plots of these two models are given in figure 1. In terms of AIC values we can say that exponential model is better than Allometric model to explain the relationship between NDVI and Biomass production, as the model with the smallest AIC value is the best.

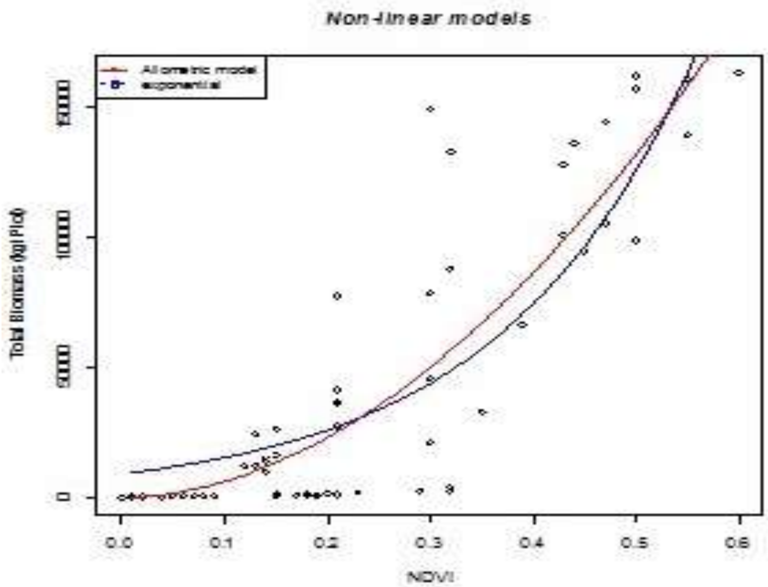

Fig 1: Allometric and exponential models applied taking NDVI as effect variable and biomass (kg/plot) as response variable.

In case of modelling with neural network, here 5 hidden neurons were used this was determined on the basis of needed complexity. Result matrix of the training process showed that the training process required 67 steps until all absolute partial derivatives of the error function were smaller than 0.01 (the default threshold). The estimated weights ranged from -1.3 to 0.88. Here the intercepts for first hidden layer ranges from -0.57 to 2.19. For example the intercepts for the first hidden layer are -0.64, -0.59, -0.61, -0.57 and -1.81. The weight for NDVI was 3.27 in the first layer. Here we can see that the AIC value for this modelling approach is the lowest among all the models tried here. The following diagram (Fig 2) is for visualizing the training process:

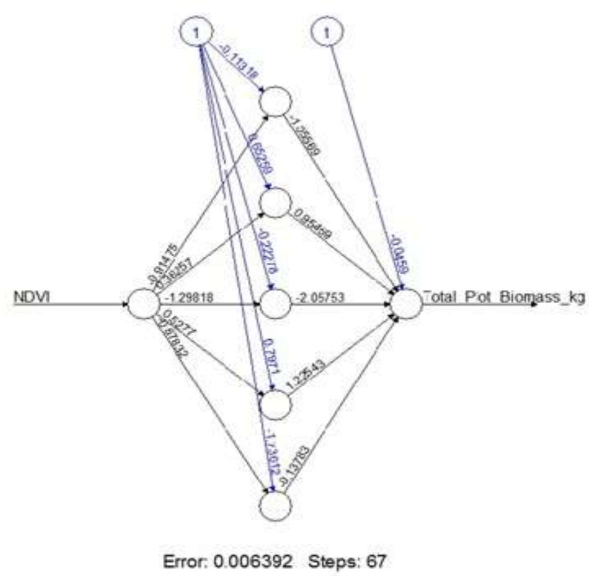

Fig. 2: Trained neural network model using NDVI as independent variable including trained synaptic weights

Conclusion

In the present study it was tried to explore the usefulness of NDVI values collected from satellite data and neural network modelling combined to estimate aboveground biomass. It was shown that artificial neural network can efficiently capture the complex relationship exists between biomass and vegetative indices. Although extrapolating the results obtained here using any data set which does not cover the whole time space can be misleading, but from the present study it can be concluded that NDVI can be a useful tool to estimate biomass production provided rigorous statistical procedures were followed.

References

Bishop, C. 1995. Neural networks for pattern recognition. Ox-ford University Press, New York.

Gunther, F. and S. Fritsch. 2010. neuralnet: Training of Neural Networks, *The R Journal* 2/1:pp.30-38.

Myneni, R.B., F.G. Hall P. J. Sellers, and A. L. Marshak. 1995. The interpretation of spectral vegetation indexes. *IEEE Transactions on Geoscience and Remote Sensing* 33: 481–486.

Nemani, R. R., K. Eeling, C. D. Hashimoto, H. Jolly, W.M. Piper. S. C. Tucker, C. J. Myneni, R. B. and S.W. Running. 2003. Climate driven increases in global terrestrial net primary production from 1982–1999. *Science* 300: 1560–1563

Tucker, C. J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8: 127–150.