Catena 152 (2017) 182–189

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena

Monitoring soil carbon pool in the Hyrcanian coastal plain forest of Iran: Artificial neural network application in comparison with developing traditional models

Ali Asghar Vahedi

Research Institute of Forests and Rangelands, Agricultural Research Education and Extension Organization (AREEO), P.O. Box 13185-116, Tehran, Iran

ARTICLE INFO

Article history:
Received 22 April 2016
Received in revised form 17 December 2016
Accepted 15 January 2017
Available online 20 January 2017

Keywords:
Artificial intelligence system
Nour Forest Park
Plant species diversity
Regression models
Soil carbon stock

ABSTRACT

Iran is categorized under Low Forest Cover Countries and the Nour Forest Park is the largest coastal plain forest in the country. As forest soils contain higher carbon than plant biomass and play a key role in the global carbon cycle, this study aimed to assess highly accurately usable models for monitoring soil organic carbon (SOC) based on plant- species diversity in the forest. For measuring plant-species diversity indices and taking soil samples at two depths (0–20 and 20–40 cm), a total of 75 plots based on a randomized complete block (RCB) design which included three naturally different stand types were outlined in the forest. Traditional models based on multiple linear regression and curve estimation regression analyses, and artificial neural network (ANN) type of multi-layer perception on the basis of back-propagation training algorithm were used for predicting the SOC. According to the Pearson's correlation, the traditional functions and ANN models including *Abundance* and evenness (J') of herbs layer, and dominance (D) of trees layer, *Abundance* and$J'$of herbs layer in turns were developed for predicting the SOC in the soil top and subsoil layer. The findings showed that the power-law models based on validation parameters among the traditional models were the best predictors for SOC in the different soil layers. Designing the topology of each model in the ANN on the basis of training algorithm and testing data set indicated that the best architectures in the optimum models in turns were composed of one hidden layer with 15 neurons including tangent sigmoid function (A_3) for predicting the SOC in the soil organic layer, and two hidden layers with 35 neurons in each layer including the same function (A_6) for predicting the subsoil (mineral) layer. Also, the optimum models in the ANN including A_3 ($R^2 = 0.87, RMSE\% = 8.17$) and A_6 ($R^2 = 0.93, RMSE\% = 3.73$), in comparison with the traditional models, predicted the SOC with a higher accuracy. Since the optimum models in the ANN predicted the response values with the highest accuracy and without major concerns for statistical issues and uncertainty, the ANN technique is suggested for monitoring SOC in the forest.

1. Introduction

Forest ecosystems globally store >70% of all soil organic carbon (SOC) which plays a key role in mitigating climate change (Jandl et al., 2007). Soil carbon storage of forests can be a significant pool in the global carbon cycle and influences soil nutrients, biological processes, water retention and forest productivity (Zhang et al., 2006; Nave et al., 2010; Yang et al., 2016). As minor changes in the amount of SOC influence CO_2 concentrations in the atmosphere, monitoring and a better understanding of the SOC stock is essential for soil resource management and for analyzing the carbon cycle and potential responses of soils to global environmental change (Yang et al., 2016). Because soil sampling with a high density is time consuming and expensive for measuring SOC in large scales and there also is a high variability of SOC in different soil depths in forest ecosystems, the most accurate protocol with a high certainty is needed for predicting and monitoring SOC stock. That means the SOC can be modeled as a function of biological variables which are strongly correlated with the SOC stock.

Regression analyses as traditional methods are commonly used for predicting response variables in biological studies. However, there are major concerns for using the traditional models because of complex relationships between biologically dependent and independent factors. In addition to complexity of functional forms between biological responses and explanatory variables, some statistical issues such as collinearity among the predictors, reliability of parameters estimate and introduction of dubious models are major concerns for using the traditional models (Vahedi, 2016). Therefore, artificial neural network (ANN) based-system with flexible approach was used to be compared with traditional method in the presented work to reexamine the accuracy and certainty of the SOC stock prediction. It is expected that the ANN has considerable efficiency for taking into consideration conflicting relationships of responses to explanatory variables (Moghadassi et al.,

E-mail address: ali.vahedi60@gmail.com.

http://dx.doi.org/10.1016/j.catena.2017.01.022

2010). This technique is one of the subsets of artificial intelligence system, and is the most commonly modeling technique having highly interconnected structure like human's brain systems that emulate the operations and connectivity of biological neurons (Ozçelik et al., 2010; Tiryaki and Aydin, 2014; Vahedi, 2016). Moreover, the ANN technique application is appropriate for improving local, regional and global responses estimate in complex ecosystems such as forests.

Iran is categorized under Low Forest Cover Countries (LFCC), and only 7.6% of its land is covered by forest ecosystems (IUFRO, 2004). The temperate and ancient Hycarian forests located at the northern slopes of Alborz Mountains are one of the most important ecoregions in Iran. These forests cover an area of almost 2 million ha extending in the southern Caspian Sea from −20 to 2500 m a.s.l (IUFRO, 2004). There are only small remnants of lowland forests in the southern coastal plain of the Caspian Sea remaining. Majority of the forests were deforested based on human's activities, urbanization, land-use changes and agricultural purposes. Therefore, the remnants of plain forests in the Hyrcanian region have significantly environmental values. The current study was conducted in one of the most prominently important plain forests, the so-called Nour Forest Park, located in Nour County in the north of Iran. Nour Forest Park is one of the largest natural plain patches in the southern coastal plain of the Caspian Sea that still remains. The forest includes mixed broad-leaved tree species and considerable plant diversity, and is protectively managed for its valuably ecological attributes. There was no study found to date that assessed SOC stocks in the mentioned forest for better understanding of their contribution for ecosystem sustainability and mitigating climate change. As actual variation in SOC storage is now a prominent item for climate change negotiations and for optimal management in forest ecosystems, accurate prediction models with high certainty are needed for monitoring SOC variability. Having major role for ecosystem services such as hydrological cycle, pollution reduction and carbon sequestration, the SOC in the Nour Forest Park should be monitored by an accurate and reliable model for improving optimal management processes in the forest. This study aimed to introduce commonly basic regression functions and ANN models for deciding which technique based on measurably environmental variables could be preferably applied for monitoring the SOC stock with the highest accuracy and certainty in the forest ecosystem.

2. Material and methods

2.1. Study area

This study was conducted in the Nour Forest Park (36° 32°–36° 36° N and 52° 08°–53° 02° E) which covers 3600 ha of the Nour county in Mazandaran Province, in northern Iran. The forest is extended in the southern Caspian Sea from −20 to 30 m a.s.l. The forest area is generally flat, and has no particularly physiographic units. Surface soils are alluvial and deep and with clay loam texture in the forest. Furthermore, the forest soils are developed from the same parent material in the whole forest. The forest is protectively managed by cooperation of municipality and Forestry Department of Nour. Some parts of the forest are determined for recreational and tourist visit. The current study was carried out in the natural stands unavailable to visitors and tourists. In fact, the natural stands have been protected by the natural resources organization in the country with no forestry operations or plantations for >40 years. The natural stands of Alder-Ironwood (AI: *Alnus glatinosa* L. – *Parrotia caspica* CA May.), Maple-Ironwood (MI: *Acer velutinum* Bioss - *Parrotia caspica* CA May.) and Elm-Hornbeam (EH: *Ulmus glabra* Huds – *Carpinus betulus* L.) are commonly broadly mixed-species stands in the forest. These natural stands are broadly distributed in different parts of the forest. Meteorological parameters directly measured between 1985 and 2014 at Nowshahr Meteorological Station, the nearest station to the forest, indicated that mean annual precipitation and temperature of this studied area were 1293.5 mm and 16.1 °C, respectively.

2.2. Field data

The three natural stands of AI, MI and EH were selected for taking soil samples based on a randomized complete block (RCB) design. There were five replications for each stand type and a total of 15 replications in the forest, and five plots were randomly allocated within each replication. The number of blocks is the number of replications in the RCB design. Therefore, twenty five plots represented each stand type and a total of 75 plots were laid out at the studied forest. The plots' area was 400 m^2 based on the commonly used protocol for measuring trees biophysical attributes and plant diversity in the Hyrcanian forests of the country (Mesdaghi, 2006). Moreover, a micro-plot with area of 1 m^2 was nested in the center of each plot for measuring ground vegetation coverage (Mesdaghi, 2006). There was no shrubs layer in the plots. Hence, the measurements were just carried out in two layers of trees and herbs in the stands. The soils were dug up to 40 cm depth in center and corners of each plot, with soil samples taken from a depth of 0–20 cm and of 20–40 cm. The soil samples of each plot were stored per layer, mixed accordingly and transported to the laboratory. In general, a total of 150 mixed soil samples (75 samples from the depth of 0–20 cm and 75 samples from the depth of 20–40 cm) were used for the statistical modeling.

2.3. Analysis

2.3.1. Estimation of SOC stock

SOC concentrations were measured in the laboratory by the wet chemical oxidation of organic carbon, referred to as the Walkley-Black method (Walkley and Black, 1934). The SOC concentrations were originally reported in percentage mass unit. The SOC stock, i.e., mass of carbon per unit area for a given depth, were calculated using the following formula:

$$SOC_i = C_i\% \times B_i \times BD_i \tag{1}$$

where SOC_i is the SOC stock ($g \cdot cm^{-2}$), $C\%$ is the SOC concentration (%), B is depth interval thickness (cm), and BD is bulk density ($g \cdot cm^{-3}$). The SOC was converted from $g \cdot cm^{-2}$ to $t \cdot ha^{-1}$.

2.3.2. Plant diversity indexes

The plant diversity indexes such as Shannon's diversity index (H'), Species richness (S), Evenness of Pielou (J') and Species dominance (D) (Barnes et al., 1998; Mesdaghi, 2006) represented plant diversity of the stands in the current study. The function of each index is as following:

$$H' = -\sum_{i=1}^{s} P_i Ln(P_i) \tag{2}$$

$$J' = \frac{H'}{H'_{\max}} = \frac{-\sum_{i=1}^{s} P_i Ln(P_i)}{Ln(s)} \tag{3}$$

$$D = \sum_{i=1}^{s} (P_i)^2 \tag{4}$$

where P is the relative abundance of species i, and S is the total number of species. Furthermore, direct measures of total abundance of plant species (trees and herbaceous species coverage) in each plot was introduced as one of the plant diversity indexes in the forest.

Pearson correlation analyses were used to explore relationships between the plant diversity indexes and the response values within the forest.

2.3.3. Regression analyses approach

For regression analyses, two protocols consisting of multiple linear regression (MLR) and curve estimation regression (CER) analysis were