Acta Ecologica Sinica 36 (2016) 62–71

Contents lists available at ScienceDirect

Acta Ecologica Sinica

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

Multi-scale regional forest carbon density estimation based on regression and sequential Gaussian co-simulation

Gaoyun Shen [a], Maozhen Zhang [a,b,*]

[a] School of Environmental & Resource Sciences, Zhenjiang A&F University Lin'an, Zhejiang, 311300, China
[b] Zhejiang Provincial Key Laboratory of Carbon Cycling in Forest Ecosystems and Carbon Sequestration, China

ARTICLE INFO

Article history:
Received 7 August 2014
Received in revised form 20 April 2015
Accepted 28 June 2015

Keywords:
Above-ground forest carbon storage
Forest inventory
Carbon density
Carbon distribution
Sequential Gaussian co-simulation
Multi-scale

ABSTRACT

By applying nonlinear regression of a unary cubic equation and sequential Gaussian co-simulation to Forest Inventory (plot) data in Xianju county, Zhejiang, from 2008, and Landsat TM image data collected in the same region in 2007, this research estimated the above-ground forest carbon density and its distributions at 30 m × 30 m and 270 m × 270 m resolutions, and analyzed the results comparatively. The results showed that the above-ground forest carbon density of Xianju county was continuously distributed, and was surrounded by high carbon density forestland, and the majority of the intermediate region was filled with low carbon density non-forestland. Using the random sampling method, the total carbon estimate is 5,289,437.11 Mg. At 30 m × 30 m resolution, with nonlinear regression of a unary cubic equation, the total carbon is 5,246,749.81 Mg, and the R^2 of the model is 0.1353. At the same scale, with sequential Gaussian co-simulation, the total carbon is 5,692,875.69 Mg, and the R^2 of the model is 0.6203. Compared with the results in the 270 m × 270 m resolution, the former total carbon amount is larger, the range of distribution is wider, and the model's precision is higher. Comparing the two methods, the results estimated by the sequential Gaussian co-simulation are better than those of the unary cubic nonlinear regression. The result of sequential Gaussian co-simulation, which considers the spatial distribution of carbon density, is closer to that estimated from plot data, and better represents the continuous change of the carbon distribution.

Forest ecosystem is the main part of the terrestrial ecosystem, which includes lots of terrestrial ecosystem organic carbon pool. As the important component of the global climate system, forests play a very essential role on terrestrial ecosystem carbon cycle, and therefore, take effect in the balance of the regulation of the global carbon and in slowing the rise of the greenhouse gas concentrations in the atmosphere [1]. The regional forest carbon distribution which is key to forest ecosystem management and relevant decision-making has been attracted into the study of forest ecology and other related fields [2]. Although there are lots of studies on vegetation carbon stock and density of forest ecosystem [3–5], most studies focused on the overall forest carbon stock estimation [4–7], the studies about forest carbon distribution are not enough [8]. Gaining regional forest carbon stock and its distribution timely and accurately is of great significance to evaluating the ability of forest C source and sink [9].

Forest carbon distribution estimation is a spatial estimation problem. Traditional statistical methods provide the estimation of overall eigenvalues, and the results of random sampling are only the estimation of total, variance and the sampling precision calculated by variance, which are unable to obtain the spatial carbon distribution. But the accurate data of regional forest carbon distribution is significant to forest resource management, regional carbon source/sink evaluation and carbon trading. Geostatistics based on spatial autocorrelation theory provides a method to discuss spatial distribution about these characteristics so that we can solve the basic problems of the spatial estimation theoretically [10]. In recent years, many comparative studies between Kriging interpolation (Kriging) and sequential Gaussian simulation (sequential Gaussian co-simulation, SGCS) are conducted in the soil science and ecology fields [11–14]. It is showed that the estimated value of forest carbon is a regionalized variable, but the Kriging has its strong smooth effect as well as the method of ordinary Kriging which is a kind of Kriging interpolation methods. However, the spatial stochastic simulation method of geostatistics can overcome the defects of Kriging method [15–18]. It views the data as a whole to restore the overall spatial structure so that it can better pursuit the reality simulation of the spatial distribution. At the same time, stochastic simulation can be directly used in the study of spatial uncertainty, which is what Kriging method lacks.

In addition to the use of geostatistics theory, the regional forest carbon storage can also be estimated by regression method. Taking advantage of regional remotely sensed data and plot data to build regression model can also estimate the regional distribution of forest carbon stock. But, at present in terms of the spatial distribution of forest carbon,

* Corresponding author at: School of Environmental & Resource Sciences, Zhenjiang A&F University Lin'an, Zhejiang, 311300, China.
E-mail address: zhangmaozhen@163.com (M. Zhang).

http://dx.doi.org/10.1016/j.chnaes.2016.01.002
1872-2032/

there are not enough comparative studies between regression and sequential Gaussian co-simulation methods at different resolutions. Based on Forest Inventory (plot) data in Xianju county, this research estimated the above-ground forest carbon density and its distribution at 30 m × 30 m and 270 m × 270 m resolutions with nonlinear regression of a unary cubic equation and sequential Gaussian co-simulation respectively. The results were compared with the random sampling estimates from the plot data, and the two methods were evaluated.

1. Study area

Xianju county (120°17′16″-120°55′ 51″E, 28°28′14″-28°48′59″N) is located in the southeast of Zhejiang province, east–west length 63.6 km, north–south width 57.3 km. The total area is 2013.18 km^2 which is made up of 80.6% of hilly mountainous area and 11.1% of plain. The county is typical subtropical monsoon climate, warm and humid, four distinctive seasons, and the average annual precipitation is 1444 mm and the sunshine hours are 1786 h.

The forestry land area is 1.64583×10^5 hm^2 including 1.53369×10^5 hm^2 of forestland, 6.289×10^5 hm^2 of scrubland, 2.579×10^3 hm^2 of young afforested land and 1.889×10^3 hm^2 of non-forest area. The forest coverage rate is 77.9% and the total forest volume amounts to 5.555×10^6 m^3 (2008).

2. Data sources and research methods

2.1. Data sources

2.1.1. The remote sensing image data

The remotely sensed image data is a Landsat TM image of 2007 (October) in the study area. Because the sixth band of Landsat TM sensor is an infrared band which mainly receives heat radiation information, it will not be used for the forest carbon density estimation. This study adopts the six bands, from first band to fifth band and seventh band, the spatial resolution is 30 m × 30 m. Firstly, the image should be processed with geometric correction and radiation correction. On the basis of better quality image, we should extract these six bands' pixel values from corresponding location plots for calculating band ratio in different combinations. After correlation analysis in the SPSS20.0, the band conbination (band5/(band5+band7)) which has highest correlation with the forest carbon in sample plots was selected to participate in the regression estimation and sequential Gaussian co-simulation.

2.1.2. The above-ground sampling survey data

The study adopts Forest Inventory (plot) data in Xianju county, Zhejiang in 2008.With the range of the whole county, the survey sets square plots in the spacing of 2 km × 3 km with the systematical sampling method. Each plot area is 0.08 hm^2 (28.28 m × 28.28 m) and the number of plots is 302, of which forest land is 251, accounting for 83.11%. The carbon density calculating from plot data was used to build simulated model.

2.2. Data preprocessing

2.2.1. Resampling

Based on the min resolution of MODIS image (250 m × 250 m), the image from Landsat TM was resampled from 30 m × 30 m to 270 m × 270 m. Both images are used respectively to estimate carbon density in multi-scale, so that the estimated results of forest carbon density can be analyzed in different resolutions.

2.2.2. Plot forest carbon stock

According to the sample tree tally data of the 302 effective plots and the existing tree spcies specified biomass model, the above–ground biomass of each tree was calculated. The biomass of each plot results from summing up the biomass of each tree within the plot. Referencing the average carbon/biomass conversion coefficient 0.5, the forest total biomass was converted to carbon stock.The above-ground forest carbon stock be converted into carbon density according to the plots area [19].

2.2.3. Normal distribution test of sample data

Through the frequency distribution histogram of the forest carbon density in study area, we can know whether the data distribution was normal, which is the premise to use sequential Gaussian co-simulation to estimate the spatial distribution of forest carbon density. Only the data approximately obeys the normal distribution, can the simulation results be more effective. Using the histogram drawing function of SPSS20.0 for testing the normal distribution of forest carbon density in study area, we found that the training plots data didn't meet the normal distribution (see Fig. 1). However, after the cube root calculation, the rest of the plots carbon density except the 0 value tended to be the normal distribution that roughly satisfied the requirement of sequential Gaussian co-simulation (see Fig. 2) [20]. Therefore, we use the cube root transformed data in the calculation of variation function.

2.3. Method of forest carbon distribution estimation

2.3.1. Nonlinear regression of a unary cubic equation

The nonlinear regression is a method to estimate the relationship between dependent variable and independent variable because of the assumption that there is a nonlinear regularity between these two variables. Generally speaking, the nonlinear regression method can reflect the relationship among objective phenomenon more correctly than the linear regression method. Nonlinear regression of a unary cubic equation has the characteristic of nonlinearity and the advantage of the low complexity.

The relationship between forest carbon density and band ratio of TM4/(TM5 + TM7) was cubic. Building the model: $y = ax^3 + bx^2 + cx + d$, d, in this model, y is carbon density, x is the value of band ratio of TM4/(TM5 + TM7), a, b, c and d are solve-for parameters.

The images of 30 m × 30 m and 270 m × 270 m resolutions were fitted by the plots data respectively. The model of nonlinear regression of a unary cubic equation at 30 m × 30 m resolution was:

$$y = -27.570x^3 + 30.449x^2 + 32.340x - 10.805.$$

The model of nonlinear regression of a unary cubic equation at 270 m × 270 m resolution was:

$$y = -51.628x^3 + 86.704x^2 - 24.764x + 8.928.$$

By using these regression models and the band4, band5 and band7 from Landsat TM remotely sensed image, the forest carbon density distribution was estimated, and the distribution maps on corresponding resolution was exported.

2.3.2. Sequential Gaussian co-simulation

Based on the geostatistical variation (variogram), plots data and remotely sensed image of Landsat TM, and the method of sequential Gaussian co-simulation, the forest carbon density for all pixels of the study area were estimated through the stochastic simulation, which could export the forest carbon density estimation map.

The method of sequential Gaussian co-simulation assumed that the estimated carbon density of each pixel was calculated from a random function of the random variable $Z(u)$. The implementation or estimate is obtained from random sampling from the conditional cumulative distribution function of the sample data and the existing estimates. And then assumed the distribution is normal and the remote sensing image data in estimated position was used to determine the conditional distribution [9].

The following basic steps involved in the process of sequential Gaussian co-simulation (SGS) were made: