Analysis of the Relationship Between Chinese Fir Forest Productivity and Environmental and Stand Factors Using Structural Equation Modeling: Postprint

Authors: Huang Xingzhao, Xu Chonghua, Xu Jun, Tao Xiao, Xu Xiaoniu

Date: 2017-04-18T00:00:00+00:00

Abstract

By collecting 644 productivity data entries from 155 Chinese fir forest studies and utilizing structural equation modeling, we analyzed the relationships between Chinese fir forest net primary productivity and mean annual precipitation, mean annual temperature, stand density, and stand age. The results indicated that Chinese fir forest net productivity exhibited significant positive correlations with mean annual precipitation and mean annual temperature, with correlation coefficients of 0.63 and 0.378, respectively. Chinese fir forest net productivity showed significant negative correlations with stand age and stand density, with correlation coefficients of -0.332 and -0.408, respectively. The structural equation model effectively elucidated the relationships between Chinese fir net primary productivity and environmental and stand factors. Mean annual precipitation, mean annual temperature, stand age, and stand density all significantly influenced Chinese fir forest net productivity, with total path coefficients of 0.398 ($P<0.01$), 0.746 ($P<0.01$), -0.321 ($P<0.01$), and -0.738 ($P<0.01$), respectively. Mean annual temperature and stand age not only directly affected Chinese fir forest net productivity, but also indirectly influenced stand net productivity by affecting mean annual precipitation and stand density. The direct path coefficients for mean annual temperature and stand age were 0.494 ($P<0.01$) and -0.700 ($P<0.01$), respectively; the indirect path coefficients were 0.252 ($P<0.05$) and 0.379 ($P<0.05$), respectively. Employed as a large-scale analytical approach for net primary productivity, the structural equation model revealed that 62% of the variation in Chinese fir forest net primary productivity was explained by mean annual precipitation, mean annual temperature, stand age, and stand density.

Full Text

Using Structural Equation Modeling to Analyze the Relationship Between Environmental and Stand Factors and Net Primary Productivity in *Cunninghamia lanceolata* Forests

HUANG Xingzhao[1,2], **XU Chonghua**[1], **XU Jun**[1,2], **TAO Xiao**[1], **XU Xiaoniu**[1,2,*] [1]School of Forestry & Landscape Architecture, Anhui Agricultural University, Hefei 230036, China [2]Collaborative Innovation Center of Agroforestry Industry in Dabie Mountain Area, Hefei 230036, China

Abstract

We collected 644 data points from 155 published studies on net primary productivity (NPP) measurements of *Cunninghamia lanceolata* forests to analyze the relationships between NPP and environmental and stand factors using structural equation modeling (SEM). Environmental factors included mean annual precipitation (MAP) and mean annual temperature (MAT), while stand factors included stand age and density. Correlation analysis revealed that NPP was significantly positively correlated with both MAP ($r = 0.630$) and MAT ($r = 0.378$), and significantly negatively correlated with both stand age ($r = -0.332$) and density ($r = -0.408$). Each variable approximated a normal distribution after natural logarithmic transformation.

The SEM effectively parsed the relationships between *Cunninghamia lanceolata* NPP and environmental and stand factors. MAP, MAT, stand age, and density all significantly influenced NPP, with total path coefficients of 0.398 ($P < 0.01$), 0.746 ($P < 0.01$), -0.321 ($P < 0.01$), and -0.738 ($P < 0.01$), respectively. Notably, MAT and stand age affected NPP both directly and indirectly, as MAT directly influenced MAP and stand age directly influenced stand density. The direct path coefficients for MAT and stand age were 0.494 ($P < 0.01$) and -0.700 ($P < 0.01$), respectively, while their indirect path coefficients were 0.252 ($P < 0.05$) and 0.379 ($P < 0.05$), respectively. The SEM analysis indicated that MAP and MAT were the strongest positive drivers of NPP, whereas stand age and density were the strongest negative drivers. The model explained 62% of the variation in NPP. We conclude that SEM is an appropriate approach for large-scale analysis of NPP, as understanding NPP patterns requires accurate assessment of ecosystem functioning in relation to environmental and stand factors.

Keywords: *Cunninghamia lanceolata* forests; net primary productivity; structural equation model; path coefficient

Introduction

Forest productivity, representing the production capacity of plant communities under natural environmental conditions, constitutes a core research focus in forest ecosystem science. As the foundation for studying nutrient cycling and energy flow in forest ecosystems, forest productivity data not only reflect community structure but also embody specific forest-environment relationships. Investigating the relationships between forest productivity and environmental and stand structure factors is crucial for understanding how forest productivity responds to changes in these factors and for providing a scientific basis for forest management and utilization.

Forest productivity integrates comprehensive information from both environmental and stand factors. Environmental factors are associated with geographic location, altitude, and site conditions, among which temperature and precipitation are essential for forest productivity development and represent the primary drivers of spatiotemporal productivity patterns. Numerous studies have explored relationships between forest productivity and temperature/precipitation across different regions, establishing predictive models for various forest types. Most current research examines relationships between environmental factors and productivity, or between stand factors and productivity, but few have integrated both factor types simultaneously. Stand density and age are primary factors affecting productivity, which changes with variations in these characteristics.

Cunninghamia lanceolata is an important timber species in southern China, ranking first in plantation area and ninth in total forest area according to the Eighth National Forest Inventory. Despite extensive research on *C. lanceolata* biomass and productivity, most studies have employed linear models to analyze relationships with temperature, precipitation, or stand age, ignoring indirect effects among factors. No studies have reported comprehensive analyses linking productivity simultaneously with both environmental and stand characteristics. Structural equation modeling (SEM) offers a robust method for establishing, estimating, and testing causal relationships, incorporating both observable and latent variables. SEM can replace multiple regression and covariance analysis to clearly dissect the effects of multiple factors and their interrelationships. This study employs SEM to elucidate the effects of environmental and stand factors on *C. lanceolata* productivity, providing a theoretical basis for sustainable management and long-term productivity maintenance.

---

1. Data Collection and Processing

We conducted a comprehensive literature search using the CNKI database, VIP Chinese Journal Database, and ScienceDirect to identify studies on *Cunninghamia lanceolata* forest productivity. This yielded a dataset comprising 644 data points from 155 publications, covering sample plots primarily in Anhui, Jiangxi, Guangxi, Guizhou, and Sichuan provinces [Figure 1: see original paper]. The dataset included geographic coordinates (latitude/longitude), mean annual precipitation (MAP), mean annual temperature (MAT), altitude (ALT), stand age, diameter at breast height (DBH), tree height (H), and net primary productivity (NPP). We excluded data from natural or planted stands younger than 5 years. For multiple studies from the same location with identical stand age and density, we averaged the productivity values. Outliers were evaluated and retained or removed based on expert consultation.

NPP was defined as the net primary productivity of *C. lanceolata* forests—the organic matter produced through photosynthesis per unit time and area after autotrophic respiration, expressed in Mg ha ^{1} yr ^{1}. Missing climate data were supplemented using the China Meteorological Data Service Center (http://cdc.cma.gov.cn), with MAP and MAT estimated from sample plot coordinates.

2. Data Analysis

We employed structural equation modeling (SEM) using the **lavaan** package in R to parse relationships between NPP and MAP, MAT, stand age, and density. SEM is a statistical method based on analyzing correlation or covariance matrices among variables, assuming causal relationships among latent variables that are typically linear combinations of observed variables. By verifying correlations among observed variables, SEM estimates path coefficients to statistically test whether the hypothesized model appropriately represents the studied process.

The general SEM formulation is:

$$\eta = \beta\eta + \Gamma\xi + \zeta$$

$$\Lambda_X\xi + \delta$$

$$\Lambda_Y\eta + \varepsilon$$

where η represents endogenous latent variables, ξ represents exogenous latent variables, β and Γ are coefficient matrices (path coefficients), Λ_X and Λ_Y are factor loading matrices linking observed and latent variables, and ζ, δ, and ε are residual vectors.

Prior to SEM analysis, we applied natural logarithmic transformation to MAP, MAT, stand density, and NPP to normalize distributions and meet model assumptions. Statistical analyses and visualizations were performed using R software.

---

1. Distribution Characteristics of NPP Relative to Environmental Factors

The spatial distribution of 644 sample plots spanned southern China' s major *Cunninghamia lanceolata* regions [Figure 1: see original paper]. Scatter plots revealed no clear trend between NPP and MAP, while NPP showed an increasing trend with MAT when temperature was below 16°C. Frequency distributions indicated that NPP samples were normally distributed across MAP and MAT gradients, concentrating in the ranges of 1400–1900 mm for precipitation and 16–20°C for temperature [Figure 2: see original paper].

2. Distribution Characteristics of NPP Relative to Stand Factors

Analysis of NPP against stand factors demonstrated a declining trend with stand age. The relationship with stand density was unimodal, initially increasing then decreasing [Figure 3: see original paper]. While NPP samples showed normal distribution across stand density, the distribution by stand age was less clearly normal, with most samples concentrated in the 5–30 year age range.

3. Correlations Between NPP and Environmental/Stand Factors

Following natural logarithmic transformation of MAP, MAT, stand density, and NPP, correlation analysis revealed significant positive relationships between NPP and both MAP ($r = 0.630$) and MAT ($r = 0.378$), and significant negative relationships with stand age ($r = -0.332$) and density ($r = -0.406$). Stand density was significantly positively correlated with MAT ($r = 0.489$) and negatively correlated with stand age ($r = -0.467$). Stand age showed no significant correlation with either MAT or MAP .

4. Structural Equation Model of NPP with Environmental and Stand Factors

Based on log-transformed data, SEM analysis using **lavaan** revealed that MAP, MAT, stand age, and density all directly affected NPP, with total path coefficients of 0.398 ($P < 0.01$), 0.746 ($P < 0.01$), -0.321 ($P < 0.01$), and -0.738 ($P < 0.01$), respectively. The path coefficient from MAP to MAT was 0.633 ($P < 0.01$), while that from stand age to density was -0.514 ($P < 0.01$) [Figure 4: see original paper].

Crucially, both MAT and stand age influenced NPP indirectly as well as directly. MAT directly affected MAP, while stand age directly influenced stand density. The direct path coefficients for MAT and stand age on NPP were 0.494 ($P < 0.01$) and -0.700 ($P < 0.01$), respectively. Their indirect path coefficients, operating through MAP and density, were 0.252 ($P < 0.05$) and 0.379 ($P < 0.05$), respectively . The SEM explained 62% of variation in NPP, with MAP and MAT identified as the strongest positive drivers, and stand age and density as the strongest negative drivers.

---

3. Conclusion and Discussion

Based on 644 carefully screened *Cunninghamia lanceolata* productivity records, we applied natural logarithmic transformation to MAP, MAT, stand density, and NPP before analyzing relationships among these variables. NPP showed significant positive correlations with MAP and MAT, and significant negative correlations with stand age and density—findings consistent with most previous research. The spatial gradient of decreasing precipitation from west to east and temperature from south to north across the study region directly influences regional NPP patterns.

We selected SEM over traditional linear models for two primary reasons. First, while NPP samples approximated normal distributions for MAP, MAT, and stand density, the distribution by stand age was less normal. Direct analysis using untransformed stand age would risk heteroscedasticity. Log transformation of variables before SEM construction ensured robust analysis. Second, separate models for each factor would ignore significant indirect effects, as correlations existed not only between NPP and individual factors but also among the factors themselves.

SEM effectively disentangled the complex relationships between NPP and environmental and stand factors, quantifying both direct effects (MAP: 0.398; MAT: 0.746; age: -0.321; density: -0.738) and indirect effects. MAT and stand age affected NPP both directly (0.494 and -0.700, respectively) and indirectly through their influence on MAP and stand density (0.252 and 0.379, respectively). The model explained 62% of NPP variation, indicating that additional important factors remain unaccounted for. Future research should incorporate soil physicochemical properties and site conditions, which have been identified as important drivers of *C. lanceolata* productivity in previous studies, to further refine our understanding of productivity variation.

References

[1] Bloom AJ, Chapin III FS, Mooney HA. Resource limitation in plants-an economic analogy. *Annual Review of Ecology and Systematics*, 1985, 16(1): 363-392.

[2] Cornelissen JHC, Castro-Díez P, Carnelli AL. Variation in relative growth rate among woody species // Lambers H, Poorter H, van Vuuren MMI, eds. *Inherent Variation in Plant Growth: Physiological Mechanisms and Ecological Consequences*. Leiden, the Netherlands: Backhuys, 1998: 363-392.

[3] Brown S. Measuring carbon in forests: current status and future challenges. *Environmental Pollution*, 2002, 116(3): 363-372.

[4] Poorter H, Sack L. Pitfalls and possibilities in the analysis of biomass allocation patterns in plants. *Frontiers in Plant Science*, 2012, 3: 259.

[5] Enquist BJ, Niklas KJ. Global allocation rules for patterns of biomass partitioning in seed plants. *Science*, 2002, 295(5559): 1517-1520.

[6] Enquist BJ, Allen AP, Brown JH, Gillooly JF, Kerkhoff AJ, Niklas KJ, Price CA, West GB. Biological scaling: does the exception prove the rule? *Nature*, 2007, 445(7127): E9-E10.

[7] Read J, Stokes A. Plant biomechanics in an ecological context. *American Journal of Botany*, 2006, 93(10): 1546-1565.

[8] [Chinese reference removed for consistency]

[9] [Chinese reference removed for consistency]

[10] [Chinese reference removed for consistency]

[11] [Chinese reference removed for consistency]

[12] Cramer W, Kicklighter DW, Bondeau A, Iii BM, Churkina G, Nemry B, Ruimy A, Schloss AL. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Global Change Biology*, 1999, 5(S1): 1-15.

[13] Prince SD, Goward SN. Global primary production: a remote sensing approach. *Journal of Biogeography*, 1995, 22(4/5): 815-835.

[14] [Chinese reference removed for consistency]

[15] [Chinese reference removed for consistency]

[16] [Chinese reference removed for consistency]

[17] State Forestry Administration. *China Forest Resources Report (2009-2013)*. China Forestry Publishing House, 2014: 10-16.

[18] Luo YJ, Wang XK, Zhang XQ, Booth TH, Lu F. Root:shoot ratios across China' s forests: forest type and climatic effects. *Forest Ecology and Management*, 2012, 269: 19-25.

[19] [Chinese reference removed for consistency]

[20] Reich PB, Luo YJ, Bradford JB, Poorter H, Perry CH, Oleksyn J. Temperature drives global patterns in forest biomass distribution in leaves, stems, and roots. *Proceedings of the National Academy of Sciences of the United States of America*, 2014, 111(38): 13721-13726.

[21] [Chinese reference removed for consistency]

[22] Poorter H, Jagodzinski AM, Ruiz-Peinado R, Kuyah S, Luo YJ, Oleksyn J, Usoltsv VA, Buckley TN, Reich PB, Sack L. How does biomass distribution change with size and differ among species? An analysis for 1200 plant species from five continents. *New Phytologist*, 2015, 208(3): 736-749.

[23] Olive RL. *Satisfaction: A Behavioral Perspective on the Consumer*. New York: McGraw Hill, 1997.

[24] Woodruff RB, Gardial SF. *Know Your Customer: New Approaches to Understanding Customer Value and Satisfaction*. Cambridge, Massachusetts: Blackwell Business, 1996.

[25] [Chinese reference removed for consistency]

[26] Kang MY, Dai C, Ji WY, Jiang Y, Yuan ZY, Chen HYH. Biomass and its allocation in relation to temperature, precipitation, and soil nutrients in Inner Mongolia Grasslands, China. *PLoS One*, 2013, 8(7): e69561.

[27] R Core Team. R: a language and environment for statistical computing, version 3.1.2. Lavaan, China: R Foundation for Statistical Computing. 2015, 11, 7. URL https://cran.r-project.org/web/packages/lavaan/index.html.

[28] [Chinese reference removed for consistency]

[29] [Chinese reference removed for consistency]

[30] [Chinese reference removed for consistency]

[31] [Chinese reference removed for consistency]

[32] Enquist BJ, Allen AP, Brown JH, Gillooly JF, Kerkhoff AJ, Niklas KJ, Price CA, West GB. Biological scaling: does the exception prove the rule? *Nature*, 2007, 445(7127): E9-E10.

[33] [Chinese reference removed for consistency]

[34] Clough Y. A generalized approach to modeling and estimating indirect effects in ecology. *Ecology*, 2012, 93(8): 1809-1815.

[35] Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 1996, 15(4): 361-387.

[36] Coupland RT. *Grassland Ecosystems of the World: Analysis of Grasslands and Their Uses: International Biological Programme Synthesis Series*. Cambridge: Cambridge University Press, 1979, 41-45.

[37] Lamb EG, Mengersen KL, Stewart KJ, Attanayake U, Siciliano SD. Spatially explicit structural equation modeling. *Ecology*, 2014, 95(9): 2434-2442.

Note: Figure translations are in progress. See original paper for figures.