Postprint of Health Assessment of Plantation Forests in Nur-Sultan, Capital of Kazakhstan

Authors: Yan Jinsheng

Date: 2021-09-14T00:00:00+00:00

Abstract

By analyzing different plantations in Nur-Sultan, screening comprehensive health evaluation indicators, and establishing a comprehensive health evaluation model, this study provides a theoretical foundation for the comprehensive health evaluation of plantations in Nur-Sultan and its surrounding areas. The study investigated 25 plantation sample plots and 2 natural forest sample plots in Nur-Sultan, selecting a total of 14 indicators including understory vegetation Shannon-Wiener index (X1), Pielou index (X2), Simpson index (X3), stand spatial comprehensive structure index (X4), soil organic matter (X5), total nitrogen (X6), total phosphorus (X7), pH (X8), mass water content (X9), average tree height (X10), average DBH (X11), average height under branch (X12), average crown width (X13), and understory regeneration (X14). Using multivariate statistical analysis methods such as factor analysis, cluster analysis, discriminant analysis, and stepwise regression analysis, a comprehensive evaluation of the health status of plantations in Nur-Sultan was conducted. Through factor analysis, the 14 individual indicators were transformed into 4 mutually independent comprehensive indicators with contribution rates of 30.482%, 24.374%, 19.711%, and 8.646%, respectively, representing 83.212% of the total information. The comprehensive health score was obtained by combining the factor score coefficient matrix with the weight of each factor. Cluster analysis was performed on the comprehensive health scores, dividing the selected sample plots into 5 categories: excellent health ( ), good health ( ), moderate health ( ), sub-health ( ), and unhealthy ( ). Discriminant analysis was used to verify the effectiveness of the cluster analysis, with self-validation and cross-validation accuracies of 100% and 85.185%, respectively. Stepwise regression analysis was employed to establish the optimal mathematical model for plantation health evaluation in Nur-Sultan: H= 0 + 0.293X13 +0.186X5 + 0.079X3 + 0.100X2 + 0.038X7 (R2=0.987), screening out 5 indicators for determining plantation health status: average crown width, soil organic matter, Simpson index, Pielou index, and soil total phosphorus. Average crown width,

soil organic matter, Simpson index, Pielou index, and soil total phosphorus can serve as indicators for determining the health status of plantations in Nur-Sultan. These 5 indicators can be measured under the same conditions to calculate the comprehensive health evaluation value and predict their health status.

Full Text

Health Assessment of Plantations in Nursultan, Capital of Kazakhstan

YAN Jinsheng[1,2,3], **WANG Yongdong**[1,3], **LOU Boyuan**[1,2,3], **Akida Askar**[1,2,3], **XU Xinwen**[1,3]

[1]Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, Xinjiang 830011, China

[2]University of Chinese Academy of Sciences, Beijing 101408, China

[3]National Engineering Technology Research Center for Desert Oasis Ecological Construction, Urumqi, Xinjiang 830011, China

Abstract

This study analyzes different plantations in Nursultan to screen comprehensive health evaluation indicators and establish a health assessment model, providing a theoretical foundation for health evaluation of plantations in Nursultan and surrounding areas. We investigated 25 plantation plots and 2 natural forest plots in Nursultan, evaluating 14 indicators: Shannon-Wiener index (X_1), Pielou index (X_2), Simpson index (X_3), stand spatial structure comprehensive index (X_4), soil organic matter (X_5), total nitrogen (X_6), total phosphorus (X_7), soil pH (X_8), mass water content (X_9), mean tree height (X_{10}), mean diameter at breast height (X_{11}), mean height under branches (X_{12}), mean canopy width (X_{13}), and understory regeneration (X_{14}). Using multivariate statistical methods including factor analysis, cluster analysis, discriminant analysis, and stepwise regression, we conducted a comprehensive health assessment of Nursultan plantations.

Factor analysis converted 14 individual indicators into 4 independent comprehensive indicators with contribution rates of 30.482%, 24.374%, 19.711%, and 8.646%, respectively, representing 83.212% of the total information. Health composite scores were calculated using factor score coefficient matrices and factor weights. Cluster analysis of these scores classified sample plots into five categories: high-quality health (7.407%), good health (11.111%), moderate health (44.444%), sub-health (29.630%), and unhealthy (7.407%). Discriminant analysis validated the clustering results with self-verification and cross-validation accuracies of 100% and 85.185%, respectively. Stepwise regression established the optimal mathematical model for plantation health assessment: $H = 0 + 0.293X_{13} + 0.186X_5 + 0.079X_3 + 0.100X_2 + 0.038X_7$ ($R^2 = 0.987$). Five key

indicators were selected: mean canopy width, soil organic matter, Simpson index, Pielou index, and soil total phosphorus. These indicators can be measured under similar conditions to calculate comprehensive health evaluation values and predict plantation health status.

Keywords: plantations; factor analysis; cluster analysis; discriminant analysis; stepwise regression analysis; health assessment

1. Introduction

Forests constitute a major component of many terrestrial ecosystems, providing not only important economic products including timber, pulp, fiber, and food, but also diverse ecosystem services. Forest health is crucial for maintaining services upon which humans depend and for effectively protecting forest organisms. Plantations currently cover a significant portion of global forest area and may account for an increasing share by the end of this century. However, plantations often feature single-species cultivation and intensive management, making monoculture stands more vulnerable to health problems. Although existing forests have adapted to certain degrees of disturbance, all forests now face new pressures, necessitating the development of systems to assess forest health conditions. Plantation health assessment forms the foundation for plantation construction and sustainable development, providing references for plantation management to ensure continuous service delivery.

Numerous studies on forest health assessment have been reported, though definitions of forest health vary among researchers, as do evaluation indicators and methods. No standardized evaluation model has yet been established. Some scholars argue that healthy forests must be structurally sustainable and meet ecological or economic management objectives without compromising sustainability. Others contend that healthy forests maintain ecosystem stability with strong self-regulation capacity while fully and continuously delivering ecological, social, and economic benefits. Forest health is also associated with biodiversity, ecosystem services, and forest economic value, requiring consideration of multiple factors including temperature, precipitation, CO_2 concentration, nitrogen deposition, air pollutants, pathogens, and pests when assessing current and future forest health.

Recent studies have employed various methods for forest health evaluation. Some used neural networks to assess forest health in Hunan Province using structural integrity, functional stability, system vitality, and sustainability indicators. Others applied cloud models to evaluate forest health around Dongting Lake using structural, vitality, sustainability, and resistance indicators. Despite these efforts, current research faces several challenges: evaluation indicators are diverse and overly complex; forest value target achievement receives insufficient consideration; and assessments focus primarily on static health conditions with limited attention to dynamic trends.

To address these issues, this study employs multivariate statistical methods to evaluate Nursultan plantations using five categories of indicators: tree growth, stand spatial structure, understory plant diversity, soil physicochemical properties, and regeneration. We aim to establish a health assessment mathematical model for Nursultan plantations, identify key indicators for health determination, and provide a theoretical foundation for predicting plantation health status and sustainable management.

1.1 Study Area Overview

The study area is located in and around Nursultan, the capital of Kazakhstan. The region has a typical continental climate, with Nursultan known as the world' s second-coldest capital. Winters are long, cold, and snowy, while summers are hot and dry. The winter season extends for 5-6 months with a snow cover period of 130-140 days. The average annual temperature is 3.2°C, with July averaging 20.8°C and January averaging -14.2°C. Annual sunshine duration is 2,200 hours. The average annual precipitation is 299.55 mm, with summer accounting for the largest share (59.74% of annual precipitation) occurring during the growing season from April to September. The multi-year average potential evapotranspiration is 721.23 mm, and the multi-year average dryness index is 2.8. The average wind speed is 3.4 $\mathrm{m \cdot s^{-1}}$, with minimum speeds in July and maximum in March (average 3.9 $\mathrm{m \cdot s^{-1}}$). Prevailing winds are southerly and southwesterly in autumn-winter, shifting to northeasterly in summer.

The study area represents a typical combination of grassland and plantations, with plantations primarily consisting of mixed forests. Representative plantation species include Scots pine (*Pinus sylvestris*), Russian olive (*Elaeagnus angustifolia*), boxelder maple (*Acer negundo*), white birch (*Betula platyphylla*), Siberian elm (*Ulmus pumila*), raspberry (*Rubus*), blackcurrant (*Ribes nigrum*), and red-barked dogwood (*Cornus alba*). Before afforestation, the area was a vast grassland.

1.2 Data Sources and Processing

To scientifically assess plantation health conditions, we selected 27 sample plots (25 plantations and 2 natural forests) following comprehensive, representative, and typical principles based on plantation growth and distribution, remote sensing imagery, and forest compartment divisions from the Kazakhstan Forestry Committee. Field surveys were conducted in July-August 2019 using plot sampling methods.

Plot sizes were flexibly set according to actual site conditions while adhering to these principles: (1) total plot area ≥ 600 m^2; (2) minimum plot specification ≥ 10 m $\times$ 10 m; (3) each plot contained ≥ 3 herbaceous subplots ≥ 1 m $\times$ 1 m; (4) each subplot was replicated 3 times. Survey content included: (1) basic plot information (coordinates, elevation, species composition, age group classification); (2) tree measurements for every individual (species, location, height, DBH, height under branches, crown width) plus understory regeneration status; (3) stand spatial structure characteristics using structured forest management survey protocols to eliminate systematic edge effects; (4) herbaceous layer surveys using the French-Swiss school method to record frequency, coverage, height, dominance, and aggregation; (5) soil samples collected via soil pits and augers at 0-300 cm depth for moisture and 0-60 cm for other properties, with 3 replicates per plot. Soil physicochemical properties were measured in the laboratory.

Data processing was performed in Excel 2010 and SPSS 26.0.

1.3.1 Evaluation Index System Construction

Based on four value dimensions—ecological, living, aesthetic, and cultural—we selected 14 representative, quantifiable indicators following scientific, objective, systematic, operational, and applicable principles:

(1) Tree growth indicators: Mean DBH, mean tree height, mean canopy width, and mean height under branches, calculated as plot averages.

(2) Understory plant diversity indicators: Shannon-Wiener index, Pielou index, and Simpson index calculated following Fang et al.' s methods.

(3) Soil physicochemical property indicators: Soil organic matter, total nitrogen, total phosphorus, pH, and mass water content, with water content averaged across soil depths.

(4) Understory regeneration indicator: Quantified according to continuous forest inventory technical regulations.

(5) Stand spatial structure comprehensive index: Calculated as:

$$\text{Stand spatial structure index} = \sqrt{(MA-0.25)^2+(UA-0.25)^2+(WA-0.5)^2+(K-1)^2}$$

where MA is stand average mingling, UA is stand average size ratio, WA is stand average angular scale, and K is mean stand crowding. MA, UA, and WA calculations follow structured forest management protocols, while K follows Hui et al.' s methods.

Tree growth reflects plantation development patterns and management effectiveness, representing ecological, living, and aesthetic values. Understory diversity forms the foundation of ecosystem integrity, representing ecological value while enhancing living and aesthetic values. Soil properties, as essential ecological factors, influence material cycling and energy flow. Natural regeneration relates to sustainable management goals and future value realization. Stand spatial structure reflects interspecific patterns, crucial for stability and management, representing ecological, living, and cultural values.

1.3.2 Analysis Methods

This study integrated four multivariate statistical methods: factor analysis, cluster analysis, discriminant analysis, and stepwise regression.

Factor Analysis: Indicators were classified as positive (higher values = better health), negative (lower values = better health), or bidirectional (optimal range). Negative indicators were transformed to their opposites. All data were Z-score standardized (bidirectional indicators also standardized). Factor analysis reduced dimensionality, with KMO and Bartlett' s tests verifying suitability. Principal components were extracted based on eigenvalues > 1. The health composite score formula was:

$$H_j = \sum_{m=1}^{4} W_m \times \sum_{i=1}^{14} V_{mi} \times Z_{ij}$$

where H_j is the health score for plot j, W_m is the weight of factor m, V_{mi} is the score coefficient of indicator i for factor m, and Z_{ij} is the standardized value of indicator i for plot j.

Cluster Analysis: System clustering classified plots into five health levels: high-quality, good, moderate, sub-health, and unhealthy.

Discriminant Analysis: Fisher' s discriminant functions validated clustering results.

Stepwise Regression: Established a predictive health assessment model.

---

2. Results and Analysis

Given the complexity of plantation ecosystems, individual indicators exhibit varying correlations, creating information overlap and indicating that plantation health is a multifaceted concept. Indicators also contribute differently to health status, making direct evaluation inaccurate. Multivariate statistical methods address these limitations.

2.1 Screening and Analysis of Plantation Health Evaluation Factors

Factor analysis of 14 indicators was conducted. KMO (0.715) and Bartlett' s test ($P < 0.05$) confirmed data suitability. Using eigenvalues > 1 and scree plot analysis [Figure 1: see original paper], four factors were extracted with rotated variance contributions of 30.482%, 24.374%, 19.711%, and 8.646%, cumulatively explaining 83.212% of variance . Factor weights were 36.632%, 29.291%, 23.688%, and 10.390%, respectively.

The rotated component matrix showed: - **Factor F_1** correlated highly with tree height, DBH, height under branches, and canopy width, representing tree growth. - **Factor F_2** correlated with Shannon-Wiener, Pielou, Simpson indices and stand spatial structure index, representing understory diversity and spatial structure. - **Factor F_3** correlated with soil total nitrogen, total phosphorus, pH, and moisture, representing soil properties. - **Factor F_4** correlated highly with understory regeneration.

Factor score equations were:

$$F_1 = -0.050X_1 + 0.315X_{10} + 0.233X_{11} + 0.043X_{12} + 0.231X_{13}$$
$$F_2 = 0.319X_1 + 0.298X_2 + 0.330X_3 + 0.318X_4 + 0.021X_{14}$$
$$F_3 = 0.169X_5 + 0.076X_6 + 0.180X_7 + 0.063X_8 + 0.111X_9$$
$$F_4 = -0.060X_5 + 0.004X_6 + 0.785X_{14}$$

Composite health scores were calculated for each plot . The highest score belonged to a natural birch forest (plot 2), indicating optimal health with complex vertical structure, high species richness, abundant regeneration, and tallest trees. It also had the highest soil organic matter, total nitrogen, and total phosphorus. The lowest score was a young Scots pine plantation (plot 25) lacking herbaceous understory and regeneration.

2.2 Plantation Health Level Classification

Based on composite health scores, cluster analysis classified plots into five categories : - **High-quality health (I)**: 2 plots (7.407%), H $\geq$ 0.773 - **Good health (II)**: 3 plots (11.111%), 0.267 $\leq$ H $<$ 0.773 - **Moderate health (III)**: 12 plots (44.444%), -0.107 $\leq$ H $<$ 0.267 - **Sub-health (IV)**: 8 plots (29.630%), -0.400 $\leq$ H $<$ -0.107 - **Unhealthy (V)**: 2 plots (7.407%), H $<$ -0.400

Overall, 62.963% of plantations were in good or better condition.

Discriminant analysis validated these results. Fisher' s linear discriminant functions were:

$$Y_1 = 1.433F_1 + 0.773F_2 + 0.127F_3 + 0.398F_4$$
$$Y_2 = 0.751F_1 + 1.837F_2 + 1.091F_3 + 1.684F_4$$
$$Y_3 = -0.100F_1 + 0.806F_2 + 1.433F_3 + 0.398F_4$$
$$Y_4 = -0.110F_1 - 0.100F_2 + 0.773F_3 + 1.091F_4$$

The first discriminant function strongly distinguished classes (R = 0.864) [Figure 2: see original paper]. Validation showed 100% self-verification and 85.185% cross-validation accuracy, with misclassifications only between adjacent health levels, confirming the clustering validity.

2.3 Plantation Health Prediction

To simplify fieldwork while maintaining accuracy, stepwise regression identified key indicators. Using health composite scores as the dependent variable and standardized indicator values as independent variables, the optimal model was:

$$H = 0 + 0.293X_{13} + 0.186X_5 + 0.079X_3 + 0.100X_2 + 0.038X_7 \quad (R^2 = 0.987, F = 381.471, P = 0.000)$$

Five indicators were selected: mean canopy width (X_{13}), soil organic matter (X_5), Simpson index (X_3), Pielou index (X_2), and soil total phosphorus (X_7).

Applying this model to original data, 85.185% of predicted scores fell within correct health intervals. Misclassifications occurred mainly between high-quality and good health categories, likely due to small sample sizes, but this has minimal practical impact as management prioritizes sub-health and unhealthy stands.

Comparative analysis revealed: - **Tree growth**: High-quality plots had significantly greater height and DBH than sub-health plots. - **Understory diversity**: High-quality plots showed 51.770% higher Shannon-Wiener and 35.974% higher Pielou indices than sub-health plots. - **Soil properties**: High-quality plots had 45.947% higher organic matter, 123.917% higher total nitrogen, and 65.093% higher total phosphorus than sub-health plots. Moisture content was slightly lower in high-quality plots, likely due to greater water consumption by older trees. - **Regeneration**: High-quality plots demonstrated substantially higher regeneration than sub-health plots.

Age-related patterns showed young plantations (mostly moderate/sub-health) have disadvantages in growth compared to older stands. Mature plantations (good/high-quality health) exhibited improved understory diversity, regeneration, and soil properties. Two unhealthy plots lacked both herbaceous understory and tree regeneration, resulting in very low soil nutrients, possibly due to anthropogenic disturbance.

---

3. Discussion

Plantation health is a complex concept interpreted differently by scholars. Plantations provide social, economic, and ecological services, with health status determining service quality and sustainable management viability. Different indicator selection and evaluation methods yield varying results and emphasis.

As complex systems, plantations exhibit multivariate data with interrelated variables. This study combined multiple multivariate statistical methods to address current limitations: inconsistent indicators, insufficient consideration of forest value targets, and static rather than dynamic assessments. We focused on objective indicator selection aligned with health concepts, quantitative weight assignment, result validation, and predictive capability.

From four value perspectives (ecological, living, aesthetic, cultural), we selected 14 easily measurable indicators reflecting different value targets. We incorporated future development trends by evaluating different age classes together to predict future health trajectories under similar site conditions. Factor analysis objectively weighted indicators, avoiding subjective expert scoring limitations. Discriminant analysis validated clustering reliability. Stepwise regression simplified the indicator set to five key metrics—mean canopy width, soil organic matter, Simpson index, Pielou index, and soil total phosphorus—practical for field application.

The results demonstrate that multivariate statistical methods provide a scientific approach for Nursultan plantation health assessment. Current overall health is good, with optimistic future prospects. The five selected indicators can effectively evaluate and predict health status for other plantations in the region.

---

4. Conclusion

This study classified 27 Nursultan sample plots into five health categories: high-quality (7.407%), good (11.111%), moderate (44.444%), sub-health (29.630%), and unhealthy (7.407%). Overall, 62.963% of plantations were in good or better condition. The optimal health assessment model ($R^2 = 0.987$) identified five key indicators: mean canopy width, soil organic matter, Simpson index, Pielou index, and soil total phosphorus. These indicators enable calculation of comprehensive health scores to predict plantation health status. The methodology addresses current research limitations and provides practical guidance for plantation management in Nursultan and similar regions.

References

[1] Boyd I L, Freer Smith P H, Gilligan C A, et al. The consequence of tree pests and diseases for ecosystem services[J]. Science, 2013, 342(6160): 823-830.

[2] Andrew, Sugden, Julia, et al. Forest health. Forest health in a changing world. Introduction[J]. Science, 2015, 349(6250), 800-801.

[3] Wingfield M J, Brockerhoff E, Wingfield B D, et al. Planted forest health: The need for a global strategy[J]. Science, 2015, 349(6250): 832-836.

[4] Trumbore S, Brando P, Hartmann H. Forest health and global change[J]. Science, 2015, 349(6250): 814.

[5] Wang Qiuyan, Chen Pengfei, Li Xuedong, et al. Review of forest health assessment methods[J]. Journal of Nanjing Forestry University (Natural Sciences Edition), 2018, 42(2): 177-183.

[6] Teale S, Castello J. Forest health: An Integrated Perspective[M]. Cambridge: Cambridge University Press, 2011: 3-16.

[7] Gao Zhiliang, Yu Xinxiao, Chen Guoliang, et al. Forest health assessment in Badaling forest farm of Beijing[J]. Forest Resources Management, 2008, 37(4): 77-82.

[8] Loehle Craig, Idso Craig, Bently Wigley T. Physiological and ecological factors influencing recent trends in United States forest health responses to climate change[J]. Forest Ecology and Management, 2016, 363: 179-189.

[9] Liu Xiaonong, Song Yabin, Xing Yuanjun. Forest health assessment of Xinhua county based on SOM neural network[J]. Journal of Central South University of Forestry and Technology, 2017, 37(4): 21-26.

[10] Li Xianliang, Zhang Gui, Li Jianjun. Assessment of forest health around Dongting lake based on entropy weight cloud model[J]. Journal of Central South University of Forestry and Technology, 2020, 40(11): 119-128.

[11] Investigation, Planning and Design Institute of State Forestry and Grassland Administration. Technical Regulations for Continuous Forest Inventory (GB/T 38590-2020)[S]. National Forestry and Grassland Administration, 2020.

[12] Research Institute of Forestry Chinese Academy of Forestry. Survey of Data Technical Specification of Structure based Forest Management (LY/T 2811-2017)[S]. The State Forestry Administration of the People' s Republic of China, 2017.

[13] Hui Gangying, Hu Yanbo, Liu Ruhong. Methods of analyzing stand spatial dominance in forest observational studies[J]. Journal of Temperate Forestry Research, 2019, 2(1): 1-6, 12.

[14] Zhou Hongmin, Hui Gangying, Zhao Zhonghua, et al. Treatment methods of plot boundary trees in spatial forest structure analysis[J]. Scientia Silvae Sinicae, 2009, 45(2): 1-5.

[15] Wang Fang. Health Assessment and Management Model of Larch Plantation in Jingouling Forest Farm[D]. Beijing: Beijing Forestry University, 2012.

[16] Hui Gangying, Zhang Lianjin, Hu Yanbo, et al. Stand crowding degree and its application[J]. Journal of Beijing Forestry University, 2016, 38(10): 1-6.

[17] Xu Laixian, Yao Lan, Guo Qiuju, et al. Forest health assessment of *Pinus massoniana* natural secondary forest in Lizhong basin in southwestern Hubei[J]. Journal of Southwest Forestry University (Natural Sciences), 2021, 41(3): 69-77.

[18] Xu Junli. Study on Urban Forest Health Evaluation Based on Community Structure, Regeneration Capacity and Soil Quality in Shanghai[D]. Shanghai: East China Normal University, 2018.

[19] Ma Zhilin, Chen Lihua, Yu Xianwei, et al. Construct and analysis for healthy assessment mode of forest ecological system[J]. Journal of Inner Mongolia Agricultural University (Natural Science Edition), 2009, 30(3): 264-270.

[20] Chen Gao, Dai Limin, Ji Lanzhu, et al. Assessing forest ecosystem health I. model, method, and index system[J]. Chinese Journal of Applied Ecology, 2004, 15(10): 1743-1749.

[21] Gu Xinxin, Si Jianhua. Health evaluation of *Pinus tabuliformis* Carr. plantation in Xining City based on analytic hierarchy process[J]. Journal of Qinghai University, 2020, 38(3): 34-43.

[22] Zhao Yongjun, Xie Yangsheng, Wang Jianjun, et al. Health assessment of *Pinus massoniana* plantation on multivariate statistical analysis: A case study of Guangxi tropical forestry experimental center[J]. Journal of Central South University of Forestry and Technology, 2019, 39(7): 100-107.

[23] Fang Jingyun, Wang Xiangping, Shen Zehao, et al. Methods and protocols for plant community inventory[J]. Biodiversity Science, 2009, 17(6): 533-548.

[24] Zheng Xueliang, Chen Lihua, Li Hongyang, et al. Health assessment of Liaodong shelterbelt system based on water conservation[J]. Science of Soil and Water Conservation, 2020, 18(2): 102-110.

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