Evaluating the productivity of ancient Pu'er tea trees (*Camellia sinensis* var. *assamica*): a multivariate modeling approach

Shuqiao Zhang[1,2*], Wendou Liu[1], Xinmeng Cheng[1], Zizhi Wang[1], Fengjun Yuan[3], Wengui Wu[1] and Shengxi Liao[1*]

Abstract

Background: The demand for productive economic plant resources is increasing with the continued growth of the human population. Ancient Pu'er tea trees [*Camellia sinensis* var. *assamica* (J. W. Mast.) Kitam.] are an important ecological resource with high economic value and large interests. The study intends to explore and evaluate critical drivers affecting the species' productivity, then builds formulas and indexes to make predicting the productivity of such valuable plant resources possible and applicable.

Results: Our analysis identified the ideal values of the seven most important environmental variables and their relative contribution (shown in parentheses) to the distribution of ancient Pu'er tea trees: annual precipitation, ca. 1245 mm (28.73%); min temperature of coldest month, ca. 4.2 °C (18.25%); precipitation of driest quarter, ca. 47.5 mm (14.45%); isothermality, 49.9% to 50.4% (14.11%); precipitation seasonality, ca. 89.2 (6.77%); temperature seasonality, ca. 391 (4.46%); and solar radiation, 12,250 to 13,250 kJ m^{-2} day^{-1} (3.28%). Productivity was indicated by the total value (viz. fresh leaf harvested multiplied by unit price) of each tree. Environmental suitability, tree growth, and management positively affected productivity; regression weights were 0.325, 0.982, and 0.075, respectively. The degree of productivity was classified as follows: > 0.8, "highly productive"; 0.5–0.8, "productive"; 0.3–0.5, "poorly productive"; and < 0.3, "unproductive". Overall, 53% of the samples were categorized as "poorly productive" or "unproductive"; thus, the management of these regions require attention.

Conclusions: This model improves the accuracy of the predictions of ancient Pu'er tea tree productivity and will aid future analyses of distribution shifts under climate change, as well as the identification of areas suitable for Pu'er tea tree plantations. Our modeling framework provides insights that facilitate the interpretation of abstract concepts and could be applied to other economically valuable plant resources.

Keywords: Economical plant resources, Evaluation index, Maximum information entropy, Productivity prediction, Structural equation modeling, Sustainable development

Background

The demand for economically productive plant resources is increasing as the human population continues to grow [1]. Although there have always been trade-offs between economic growth and nature conservation [2], an increase in agricultural productivity (yield) of 60–120% in 2030 relative to 2005 is needed to meet projected increases in demand [3]. Many studies have characterized

*Correspondence: Shuqiao.zhang@anu.edu.au; cafliao@163.com

[1] Institute of Highland Forest Science, Chinese Academy of Forestry, Kunming 650216, Yunnan, China
[2] Fenner School of Environment & Society, College of Science, Australian National University, Canberra, ACT 2601, Australia
Full list of author information is available at the end of the article

Zhang *et al. Plant Methods*   (2022) 18:95

Page 2 of 13

the relationships of plant productivity with species richness [4] and climate [5], the effects of crown attributes and stand structure on tree productivity [6], and the effects of management policies on plant productivity [7]. Most of these studies have used integrative modeling approaches to explore the relationships among variables [4, 8], but few have evaluated correlations using multisourced factors (e.g., environment, plant attributes, and management) at the individual level. In addition, the difficulty of interpreting the significance of the outputs of these modeling analyses often impedes the ability to extract practical insights that could be applied to improve productivity.

Tea is consumed by over three billion people across 160 countries, making it one of the world's most popular beverages [9, 10]. Its management as a cash crop plays an important role in rural poverty reduction [2, 11] and economic growth. Teas are known to provide benefits to human health and well-being, such as decreasing the risk of cancer [12], cancer recurrence [13], and cardiovascular and nervous system diseases [14]. Rogers et al. [15] demonstrated that an active ingredient in tea, theanine, can have beneficial effects for treating chronic diseases, such as raising blood pressure. The taste of tea can also improve mood and focus and alleviate depression and dementia [16].

China has a long history of tea plantation and culture [17]. Most tea gardens are located in southern China. The Chinese Academy of Agriculture Science classified China into several tea areas based on ecological conditions and geography: South China, Southwest, Jiangnan, and Jiangbei tea areas. Since the economic reform of 1978, the area of tea gardens has increased and reached 0.19 million hectares in 2017, with a total tea yield of 2.46 million tons [18]. Production of the six Chinese traditional tea brands, green tea, yellow tea, white tea, oolong tea, black tea, and dark tea, has grown steadily over the past 40 years, and the tea-based products in the market have diversified. According to Chen et al. [18], China's tea exports (0.37 million tons in volume) account for 15–20% of the world's total, and domestic per capita consumption is 1.42 kg per biennium. Yunnan Province, an important region in the Southwest tea area, harbors 15.37% of the tea plantation area and 16% of the total tea output in 2017, making it one of the key areas of strategic development based on tea growing and processing.

Pu'er region, Yunnan Province, is well-known for its high-quality black and dark tea products and unique tea manufacturing process [19], making it an ideal area to study the effect of environmental conditions and human activities on tea production. In Pu'er, many old tea trees found in natural areas are protected. Some trees in tea gardens regrow from the original ancient trunks.

According to Lu et al. [9], there is no official definition of ancient tea trees, but natural trees over 100 years of age are often considered ancient. Hence, we define wild trees over 100 years or trees regrown from trunks over 100 years of age as ancient in this study. The main ancient tea species include Dali tea [*Camellia taliensis* (W. W. Sm.) Melch.], tea [*C. sinensis* (L.) Kuntze], Pu'er tea [*C. sinensis* var. *assamica* (J. W. Mast.) Kitam.], and white tea (*C. sinensis* var. *pubilimba* T. L. Ming). Besides values of ecosystem services, ancient teas also have high economic values for their non-timber products. The rich aroma and pleasant taste of these warm dark teas add substantial value to their products [20]. Although the prices of final products depend on fermentation processes, technological infrastructure, and the time required for tea harvesting [21], our field research has indicated that the price of raw teas increases with tree age. Increasing the production of tea from ancient trees could enhance the economic value of these trees, and an analysis of the factors associated with increased productivity is the first step toward achieving this goal.
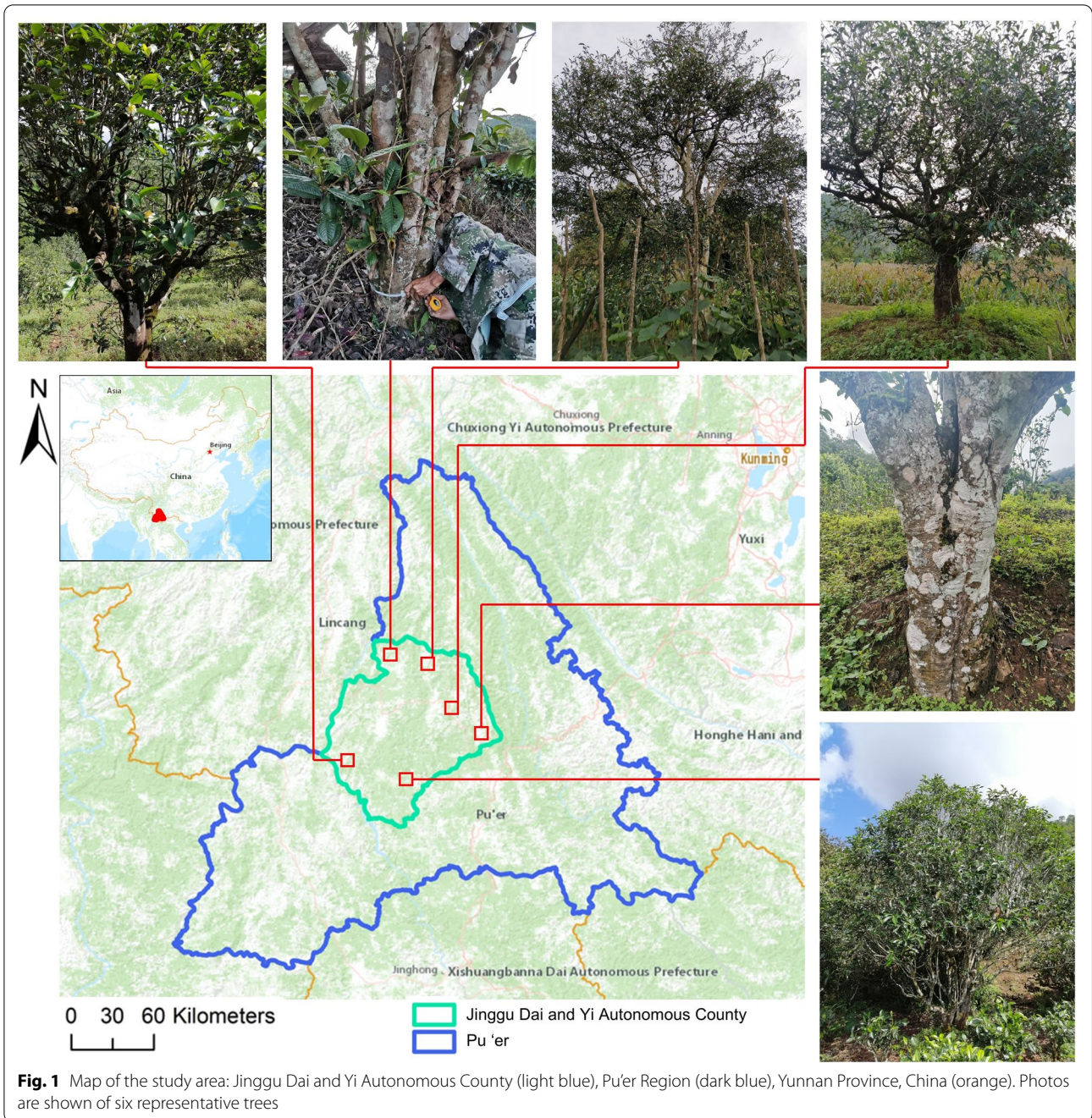
Here, this study evaluates the productivity of an economically valuable plant resource. To improve the current understanding of interactions between plant productivity and impact factors [1, 3], we detect the critical drivers' contribution to ancient tea trees' productivity and build formulas and indexes to help interpretation. We aim to build models that facilitate the objective interpretation of "productivity" and could be applied to other economic plants. The methodology also could be referred to as studying other plant traits and vague concepts.

Materials and methods
Study area
Yunnan Province is located in the southern extension of the Qinghai–Tibet Plateau and Yunnan–Guizhou Plateau (from 21° 8′ 32″ to 29° 15′ 8″ N and from 97° 31′ 39″ to 106° 11′ 47″ E). Mountainous regions comprise 84% of the total area of the province, and the plateau and hilly regions account for 10% [22]. Natural plant resources are abundant in Yunnan Province, and the soils are suitable for the growth of various plant resources, as this region (average altitude of 2000 m) is characterized by a unique plateau monsoon climate formed by the South Bengal high-pressure airflow.

The study area (Fig. 1) is located in southwest Yunnan Province (from 22° 49′ to 23° 52′ N and from 100° 02′ to 101° 07′ E) and covers 777,700 ha. Jinggu Dai and Yi Autonomous County (abbreviated as Jinggu) features a diverse topographic landscape that includes mountainous areas, plateaus, basins, and valleys. The altitude ranges from 600 to 2920 m. The Lancang River flows from northeast to southwest, and there are a total of

Fig. 1 Map of the study area: Jinggu Dai and Yi Autonomous County (light blue), Pu'er Region (dark blue), Yunnan Province, China (orange). Photos are shown of six representative trees

94 rivers with a total length of 1863.54 km and annual average total runoff of 4.7 billion m³. Jinggu receives abundant rainfall and has distinctive seasons due to the influence of the southwest monsoon. Because Ailao Mountain blocks cold air from the north in winter and monsoon-related precipitation limits increases in temperature in the summer, Jinggu has stable temperatures year-round, but temperature noticeably varies with elevation. The annual average temperature is 17.7 to 22.3 °C [23]. The hottest month (June) features an average temperature of 21.7 to 24.6 °C, whereas the coldest month (January) has an average temperature of 11.4 to 13 °C. The daily and annual temperature differences are 13.6 °C and 11.6 °C on average, respectively. The average annual precipitation is over 1200 mm, and approximately 87% of the precipitation falls during the rainy season (May to October). The average relative humidity

Zhang *et al. Plant Methods*   (2022) 18:95

Page 4 of 13

is 78%, the average number of annual sunshine hours is 2056.3 h, and southern winds are predominant.

The unique geographical and climate conditions of this region affect the distribution of ancient tea trees. Soil types in the study area are diverse and include red soil, lateritic red soil, yellow soil, brown soil, purple soil, alluvial soil, and paddy soil. Alkali-resistant tea trees grow better on red soil, lateritic red soil, and yellow soil, and these soil types are present across the study area in both valleys and hilly regions. This high diversity of soils has also contributed to the abundance of forest resources. There are 596,000 ha of forestry land, accounting for 79.2% of this region [23]. The dominant forest types are seasonal rain forest, middle mountain moist evergreen broad-leaved forest, Yunnan pine forest, and coniferous and broad-leaved mixed forest. Many ancient tea trees are present in these forest types, and some are protected or cultivated in tea gardens.

Data

Environmental data collected comprised 28 rasterized parameters (Table 1). The bioclimate variables (bio1–19), wind speed, water vapor pressure, and solar radiation were obtained from WorldClim version 2.1 at a spatial resolution of 30 arc-seconds (1 km^2). The accuracy of these data was evaluated by global correlation coefficients (between estimated and observed values). All temperature parameters had coefficients over 0.99; solar radiation and vapor pressure had coefficients over 0.95; precipitation variables had coefficients of 0.86; and wind speed had a coefficient of 0.76 [24]. We calculated the average values from the 1970 to the 2000 datasets (released in January 2020). We used public soil texture data from the Resource and Environment Science and Data Center [25], which comprised percentages of sand, silt, and clay. We extracted slope and aspect values from the Advanced Land Observing Satellite (ALOS) 12.5 m Digital Elevation Model (DEM) [26]. To make aspect

Table 1 Details of the variables, their value ranges, percent contributions to the Maxent model, and permutation importance

Parameter	Full name or description (*)	Value range	Percent contribution	Permutation importance
bio1	Annual mean temperature (℃)	10.29–23.05	0.17	0.33
bio2	Mean diurnal range [mean of monthly (max temp—min temp)] (℃)	9.25–12.95	0.47	1.21
bio3	Isothermality (BIO2/BIO7) (× 100) (%)	47.02–53.25	14.11	6.04
bio4	Temperature seasonality (standard deviation × 100)	307.35–448.04	4.45	5.92
bio5	Max temperature of warmest month (℃)	16.4–36.3	0.01	0.01
bio6	Min temperature of coldest month (℃)	− 2.2 to 11.3	18.25	21.14
bio7	Temperature annual range (BIO5–BIO6) (℃)	18.6–25	0.36	1.71
bio8	Mean temperature of wettest quarter (℃)	14.55–26.45	0.06	0.00
bio9	Mean temperature of driest quarter (℃)	4.87–19.35	0.37	1.15
bio10	Mean temperature of warmest quarter (℃)	14.55–26.45	0.07	0.01
bio11	Mean temperature of coldest quarter (℃)	4.87–17.73	0.34	0.92
bio12	Annual precipitation (mm)	916–1616	29.73	26.75
bio13	Precipitation of wettest month (mm)	173–342	0.12	0.04
bio14	Precipitation of driest month (mm)	5–22	0.04	0.40
bio15	Precipitation seasonality (coefficient of variation)	74.48–90.55	6.77	4.57
bio16	Precipitation of wettest quarter (mm)	486–911	0.01	0.02
bio17	Precipitation of driest quarter (mm)	28–73	14.45	11.97
bio18	Precipitation of warmest quarter (mm)	486–911	0.33	0.65
bio19	Precipitation of coldest quarter (mm)	31–79	1.67	9.78
wind	Wind speed (m s^{-1})	0.7–2.5	0.90	0.55
vapr	Water vapor pressure (kPa)	0.62–1.51	0.21	0.08
srad	Solar radiation (kJ m^{-2} day^{-1})	10,858–14,288	3.28	3.99
sand	*Particle size from 0.05 to 2 mm	22–61	0.13	0.17
silt	*Particle size from 0.002 to 0.05 mm	17–49	0.84	0.96
clay	*Particle size less than 0.002 mm	14–48	0.06	0.07
slope	*Extract from DEM (°)	0–70.49	1.59	0.86
Sin_aspect	*Aspect (east to west) = sin((π/180) × aspect (degree))	− 1 to 1	0.53	0.25
Cos_aspect	*Aspect (north to south) = cos((π/180) × aspect (degree))	− 1 to 1	0.67	0.46

* indicates explanations or essential information of parameters

Zhang *et al. Plant Methods* (2022) 18:95

Page 5 of 13

values more mathematically reasonable, we used sine and cosine functions to constrain the values from −1 to 1, which represent east–west and north–south degrees, respectively (Table 1).

We measured, monitored, and collected field data from individual ancient tea trees to build a field dataset of ancient tea trees for the past decade. All coordinates of known trees were recorded in Jinggu, which met the "unbiased data" assumption of using Maxent. We obtained comprehensive data from 1282 individuals for the structural equation modeling. Specifically, we measured tree age (years), diameter at breast height (DBH, cm), ground diameter (cm), tree height (m), crown width (m), and height under branch (m). The crown width was the average of the north–south and east–west widths. We also subjectively scored growth vigor (strong=3, medium=2, weak=1), harvest intensity (strong=1, medium=0.5, week=0), and protection (strong=1, medium=0.5, weak=0.1, none=0). Specifically, trees showing strong morphological growth trends (e.g., crown width and tree height) were considered to show "strong" growth vigor. Harvest intensity was considered "strong" when all new leaves and branches were periodically harvested; by contrast, harvest intensity was considered "weak" when most leaves were had not been removed. Comprehensive protection was considered "strong" if an enclosure was present around the trees, monitoring efforts were in place, and weeds were regularly cleared. These indicators were assessed and recorded in the field. In addition, we summarized each tea tree's total value based on fieldwork and market research in 2019. The annual production (kg) of green leaf was recorded for each tree by our research group and the local government. Although the final products have different prices due to different processing technology and places it sells, the raw tea from the primary harvesting has nearly the same prices in Yunnan, and the price is related to tree ages [23]. Hence, we used the raw tea's unit price (RMB/kg) in the study area to calculate the total value by multiplying output (kg).

Analyses
Here, we evaluated the productivity of tea production from Pu'er ancient tea trees using a structural equation modeling approach. We defined "productivity" as the value generated by an individual tree and used the 1-year output value (viz. fresh leaf harvested multiplied by the unit price) as a direct indicator of productivity. We then developed four models to analyze correlations. The environmental suitability model comprised 28 rasterized parameters. The tree growth model comprised all tree attributes and growth vigor. The management model comprised harvest intensity and protection as variables.

The productivity model comprised total value as an indicator.

We used ArcGIS Desktop 10.8 to standardize all environmental raster data into the same band, cell size, pixel type, pixel depth, coordinate system, and spatial datum. We then used Maxent 3.4.4 to calculate contributions from each environmental parameter [27]. Maxent's maximum information entropy-based machine learning method has been used to filter significant parameters and reduce the number of dimensions [28, 29]. Specifically, we input the geographic coordinates (latitude and longitude) of all known ancient tea trees in Jinggu and 28 rasterized parameters into the software and used the jackknife procedure to measure variable importance. We set the random test percentage as 75, replicates as 10 with cross-validation running type, and the output format as logistic [30]. After running the model, we used the average of 10 iterations to weight the importance of each parameter. We also used the coordinates of 1282 trees to extract values from the 28 rasterized environmental parameters to prepare the dataset for the productivity model using the "Extract Multi values to Points" tool in ArcGIS.

We used JMP 15.2.0 [31] to characterize the distribution of points and detect bivariate relationships between each field parameter and total value. We selected the seven most important environmental variables based on their relative contributions to build the initial integrative model. The seven tree-related and two management-related variables were also comprised. Since the number of parameters is associated with model fitness and redundancy, we tried to filter the most significant variables for the final model by their contributions and rerun the initial model.

We used IBM SPSS AMOS 24.0 to build the integrative model, a graphic-based software to visualize correlations and regression weights among variables [32]. The structural equation modeling method applied in the software has been widely used in recent ecological modeling studies [4, 8]. We built the environment suitability, tree growth, and management models and evaluated their relationships with the productivity model. The initial regressions were calculated based on the following assumptions: environmental suitability affects tree growth, management, and productivity; tree growth affects productivity; and management affects tree growth and productivity (Fig. 4). In addition, we found that fluctuations in total value were correlated with adjustments to the harvest intensity and protection implemented by managers from our fieldwork. Hence, we assumed that the residual errors of harvest intensity and protection were correlated with the residual error of the total value. The final model was

Zhang *et al. Plant Methods*   (2022) 18:95

Page 6 of 13

standardized by maximum likelihood discrepancy, and unbiased covariances were used as the input matrix. Using the coefficients, we derived mathematical equations to categorize the productivity of our samples into four classes based on the calculated values.

Results

Environmental suitability model

The average omission rate over the replicate runs was close to the predicted omission rate per the definition of the cumulative threshold. The area under the receiver operating characteristic curve (AUC) value was 0.938, and the standard deviation was 0.007. The AUC was significantly high for all data partitions, suggesting excellent reliability and fitness of the model [33]. Table 1 shows value ranges, estimates of the relative contributions of the environmental parameters to the Maxent model (all p-values less than 0.01) and parameters' permutation importance. The seven most-contributed variables were used to build the productivity model and plotted with the logistic output (Fig. 2). These variables are (contribution in parentheses): annual precipitation (29.73%), min temperature of the coldest month (18.25%), precipitation of the driest quarter (14.45%), Isothermality (14.11%), precipitation seasonality (6.77%), temperature seasonality (4.45%), and solar radiation (3.28%). Their curves suggested extreme points that maximize the logistic output, which indicates the effect on environmental suitability. The value 0.5 is a threshold to extract suitable value ranges of parameters, representing suitable environmental conditions of ancient Pu'er tea trees.
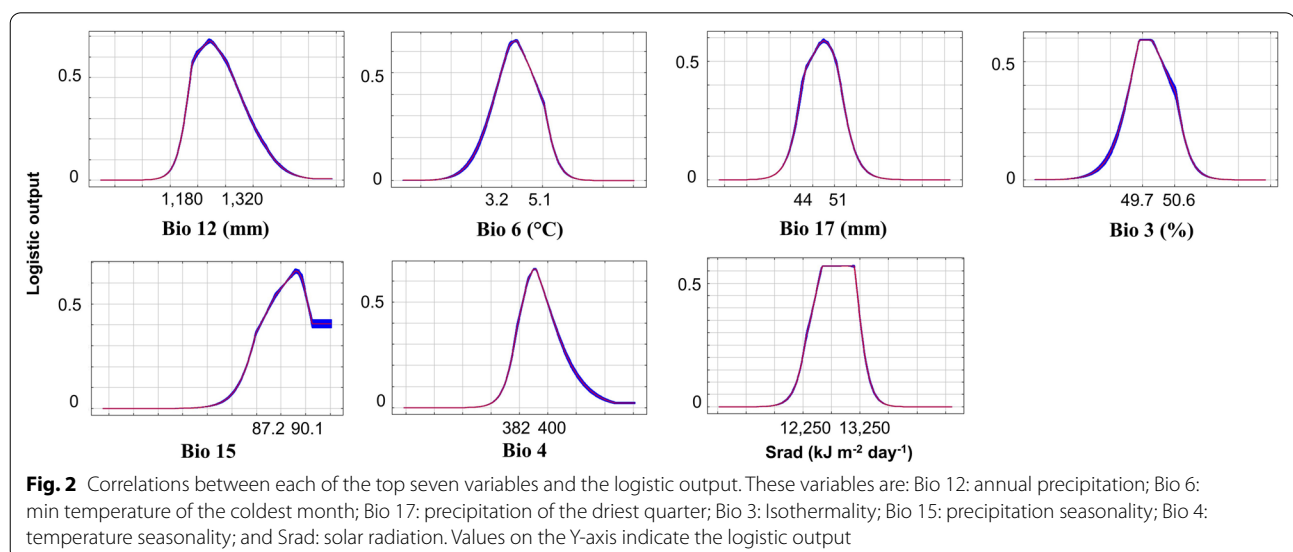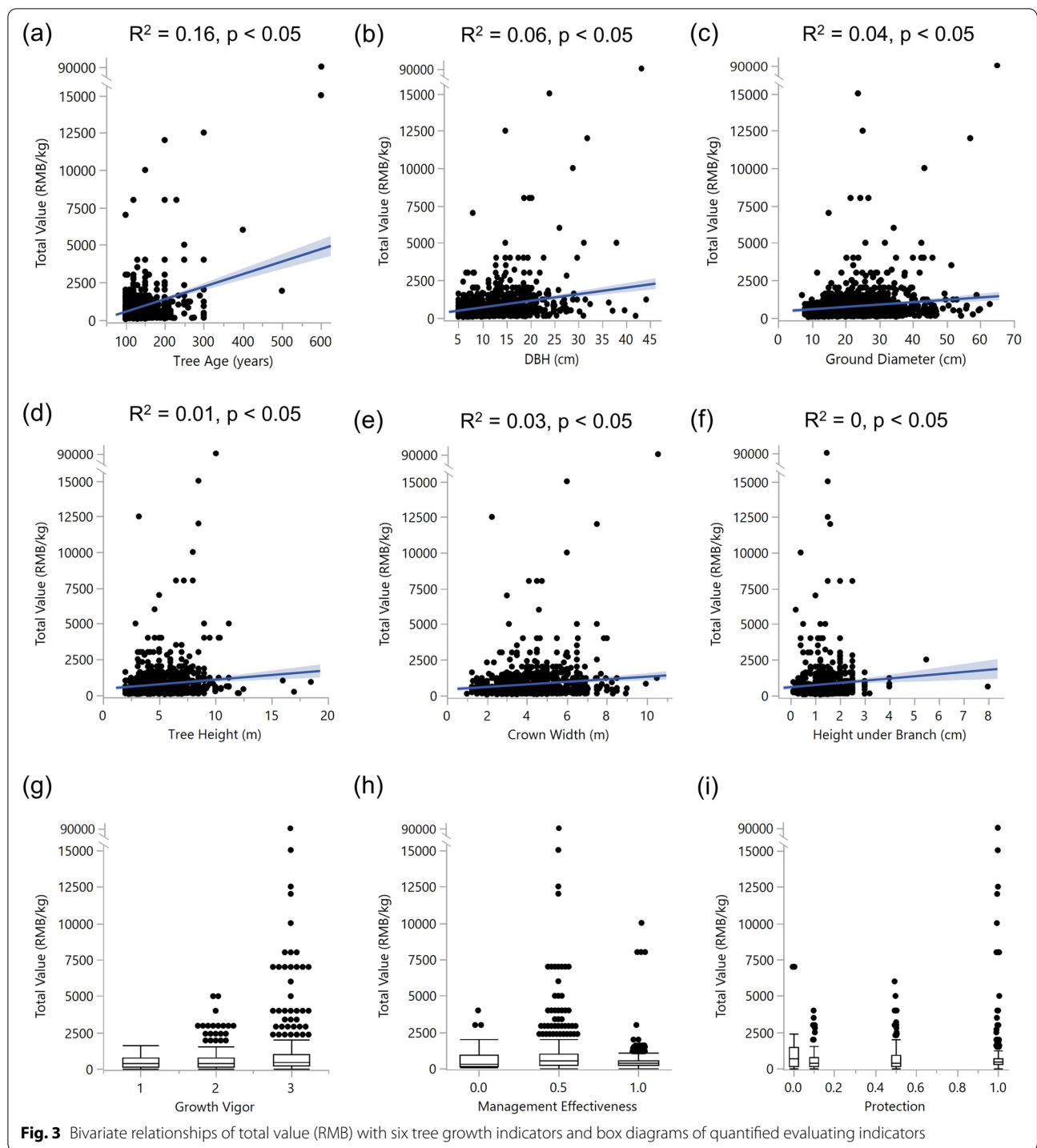
Tree growth and management models

The bivariate relationships between the total value and tree growth and management-related variables are shown in Fig. 3. We did not observe any clear patterns in these bivariate models but some slightly positive correlations. The results suggest that each bivariate model lack explanatory power with a small R squared. Therefore, integrative pathways are required to explain the relationship between the total value and related variables.

Productivity model

The initial productivity model is shown in Fig. 4. We optimized the model based on variable analysis and their percent contributions to improve model fit. We kept annual precipitation and min temperature of coldest month for the environmental suitability model because they had the highest percent contributions. We left ground diameter and crown width for the tree growth model because they had substantial weights in the initial model. We standardized the units of the parameters by setting residual errors and the magnitudes of some variables to one to make the estimates comparable. The performance of the final model (Fig. 5) was acceptable with the following parameters: chi-square (χ^2), 34.088; degrees of freedom, eight; goodness of fit index (GFI), 0.993; adjusted GFI, 0.974; comparative fit index (CFI), 0.989; and Akaike Information Criterion (AIC), 74.091.
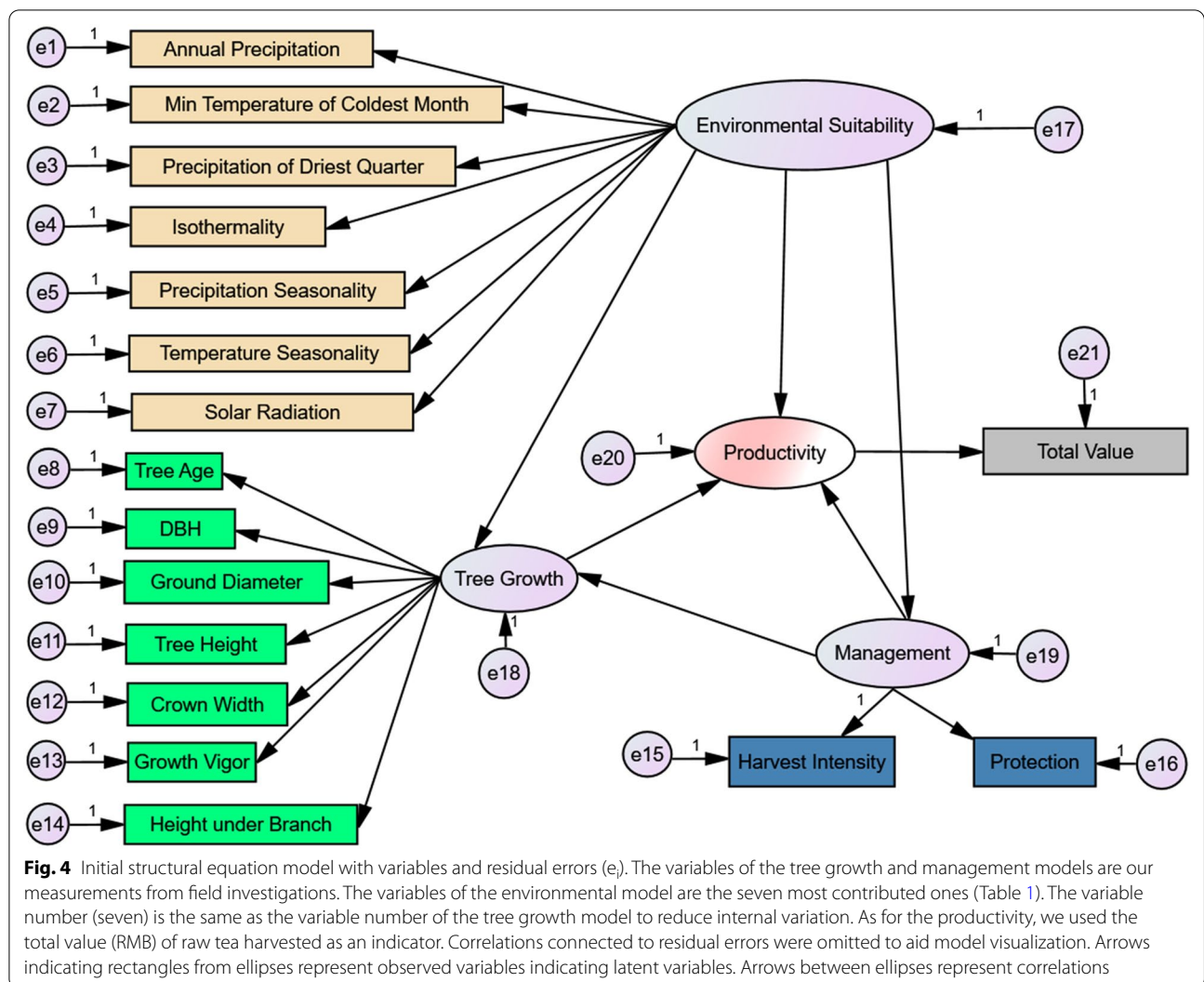
The results suggest that environmental suitability, tree growth, and management positively contribute to the productivity of ancient tea trees. Because the units were standardized, we used the coefficients (standardized regression weights) to indicate the importance of the parameters: path values indicate variables' influence

Fig. 2 Correlations between each of the top seven variables and the logistic output. These variables are: Bio 12: annual precipitation; Bio 6: min temperature of the coldest month; Bio 17: precipitation of the driest quarter; Bio 3: Isothermality; Bio 15: precipitation seasonality; Bio 4: temperature seasonality; and Srad: solar radiation. Values on the Y-axis indicate the logistic output

Fig. 3 Bivariate relationships of total value (RMB) with six tree growth indicators and box diagrams of quantified evaluating indicators

on their arrow-pointed models. For example, environmental suitability had a positive but weak effect (coefficient of 0.068) on tree growth, and the effect of environmental suitability was weaker than the effect of management (coefficient of 0.325) on tree growth.

The values of the coefficients suggested that the effect of tree growth on productivity was more than three times the magnitude of the effect of environmental suitability on productivity. The effect of management was very weak. Likewise, the contribution of ground diameter
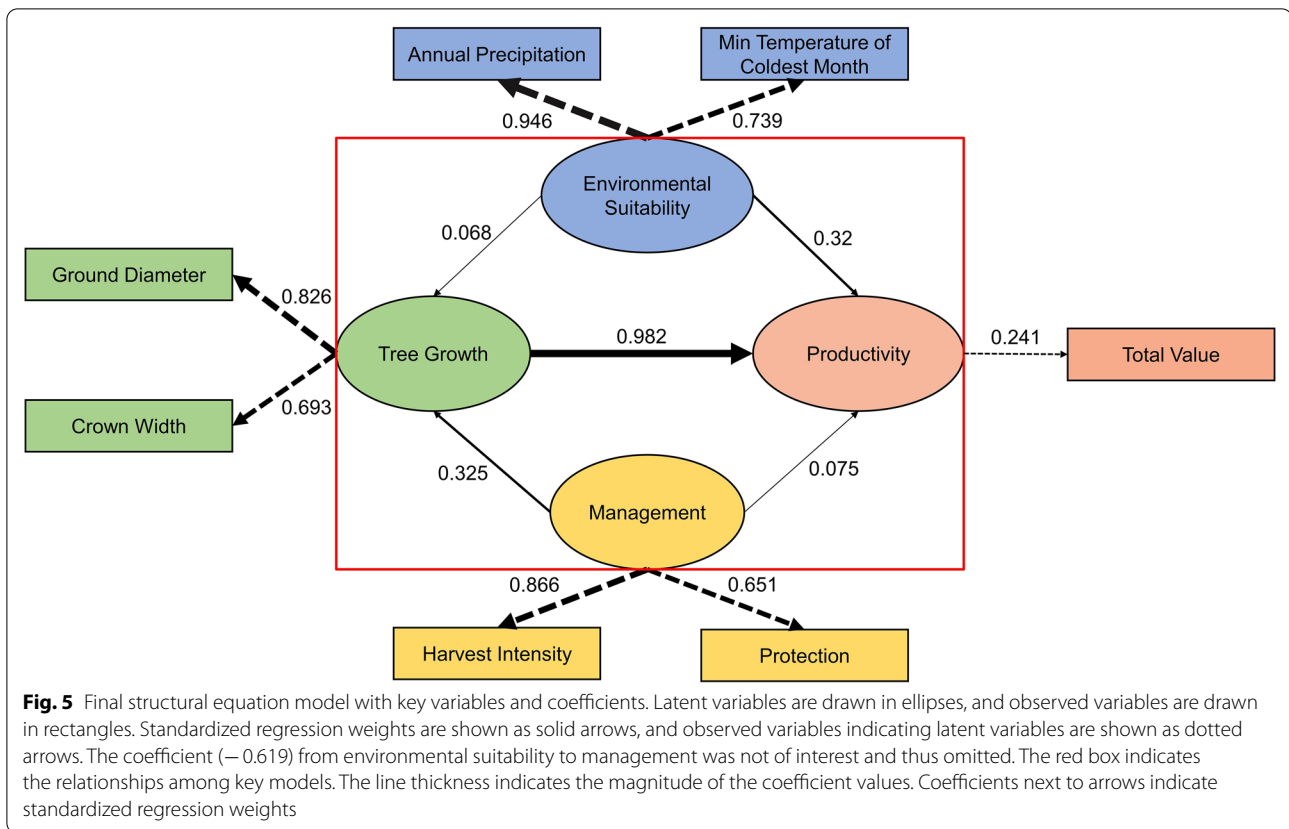
Fig. 4 Initial structural equation model with variables and residual errors (e$_i$). The variables of the tree growth and management models are our measurements from field investigations. The variables of the environmental model are the seven most contributed ones (Table 1). The variable number (seven) is the same as the variable number of the tree growth model to reduce internal variation. As for the productivity, we used the total value (RMB) of raw tea harvested as an indicator. Correlations connected to residual errors were omitted to aid model visualization. Arrows indicating rectangles from ellipses represent observed variables indicating latent variables. Arrows between ellipses represent correlations

to tree growth was greater than that of crown width, and the effect of harvest intensity on management was greater than the magnitude of the effect of protection on management. The coefficient between total value and productivity was only 0.241, suggesting that there might be other observable parameters (such as residual errors in the model not shown in the Fig. 5) that are better indicators of productivity.

We quantified productivity based on the structural equation model (Fig. 6). The equation was formulated using a linear function by normalizing the coefficients 0.32 (from the environmental suitability model), 0.982 (from the tree growth model), and 0.075 (from the management model) to 0.24, 0.71, and 0.05 (sum is 1), respectively. Likewise, the coefficients of the submodels were determined by the coefficients of the observed variables and standardized by dividing by the maximum values. Specifically, the coefficients 0.54 and 0.46 were

determined by the regression weights of 0.826 and 0.693 in the tree growth model, respectively. The denominators 71.5 and 11.5 were the maximum values of the ground diameter and crown width, respectively. The purpose of the division was to standardize their units and make their magnitudes comparable. The coefficients 0.57 and 0.43 were determined by the regression weights of 0.866 and 0.651, respectively. We used quadratic functions to formulate the environmental suitability model (y$_1$) because they are the most straightforward pathways for representing the relationships with extrema. The extreme values of parameters (representing the most suitable conditions) and percent contribution (shown in parentheses) to the distribution of ancient Pu'er tea trees are annual precipitation, ca. 1245 mm (28.73%); min temperature of coldest month, ca. 4.2 °C (18.25%); precipitation of driest quarter, ca. 47.5 mm (14.45%); isothermality, 49.9% to 50.4% (14.11%); precipitation seasonality, ca. 89.2 (6.77%);

Zhang *et al. Plant Methods*   (2022) 18:95

Page 9 of 13

Fig. 5 Final structural equation model with key variables and coefficients. Latent variables are drawn in ellipses, and observed variables are drawn in rectangles. Standardized regression weights are shown as solid arrows, and observed variables indicating latent variables are shown as dotted arrows. The coefficient (− 0.619) from environmental suitability to management was not of interest and thus omitted. The red box indicates the relationships among key models. The line thickness indicates the magnitude of the coefficient values. Coefficients next to arrows indicate standardized regression weights

Fig. 6 Equations for quantifying productivity (y). Submodel (y_1, y_2, and y_3) weights were determined by coefficients (Fig. 5). The environmental suitability model (y_1) was formed by quadratic functions based on Fig. 2, and linear models were used for the productivity (y), tree growth (y_2), and management (y_3) models

temperature seasonality, ca. 391 (4.46%); and solar radiation, 12,250 to 13,250 kJ m^{-2} day^{-1} (3.28%). The quadratic functions were fit based on these values and Fig. 2. Likewise, their domain definitions were extracted from the function images in Fig. 2; when the logistic output was 0, the environment was not suitable for ancient tea trees, and the factor's contribution was 0.

After formulating the model for productivity, we applied the formulas to our samples and set criteria to classify productivity based on the calculations of the equations (Fig. 6), and example photographs of the criteria taken in the field were selected for visual assistance. The criteria were as follows: > 0.8, "highly productive"; 0.5–0.8, "productive"; 0.3–0.5, "poorly productive"; and

Zhang *et al. Plant Methods*   (2022) 18:95

Page 10 of 13

< 0.3, "unproductive" (Fig. 7). The calculations suggested that 7 samples were highly productive, 594 samples were productive, 667 samples were poorly productive, and 14 samples were unproductive. Overall, 53% of the samples were categorized as "poorly productive" or "unproductive"; thus, these regions' management requires attention for conservation and productivity improvement.

Discussion

A comprehensive dataset was obtained in our study. For the environmental suitability model, we comprised nine factors in addition to the commonly used "bio1–bio19" [24] to better reflect environmental factors. The wind speed, solar radiation, and water vapor pressure data were averaged annually instead of quarterly or monthly because ancienttrees have long growth cycles [34]. However, we did not consider whether the extremes (e.g., wind in spring and solar radiation in autumn) limit the distribution of tea trees [35]. Future phytogeographical studies are needed to test this possibility. Although the sprouts and leaves are usually harvested in the spring and autumn, the branches and trunks are slow-growing [36]. Data from the WorldClim database are averaged and interpolated to rasterize the data on maps, which may not be accurate enough for modeling data from individual trees. Our approach could be refined by establishing monitoring points near these ancient tea trees and collecting data regularly, especially in light of the recent impacts of climate change on trees [2].

Among topographic indexes, we used slope and aspect but excluded altitude because we suspected that the effect of altitude on the total value of ancient tea trees would be indirect [37]; we expected that environmental factors such as temperature and precipitation would have direct effects on the total value of ancient tea trees. We used sine and cosine functions to constrain the aspects to the east–west and north–south dimensions (formulas shown in Table 1). Transformed aspect values show better performance in mathematical calculations and theoretically have greater explanatory power. For soils, we used texture because the data were percentages and were easy to calculate. Although this indicator can reflect the soil conditions for tea trees, we did not quantify other important factors, such as types, thickness, and nutrients. These indicators could be used in future studies to better evaluate soil conditions.

The environmental suitability model is a point-based (i.e., individual-based) model. More area-based research is needed to analyze specialized niches, as vegetation structure and landscape components are important in area studies [38]. In future niche analyses, we suggest classifying ancient tea trees into areas consisting of different tea gardens or land types. As Maxent is a coordinate-based method and has been widely used in previous distribution studies because of its simplicity [39], there were studies used random forest or deep learning models to evaluate the contributions of environmental factors [40, 41]. The accuracy and efficiency of the machine learning-based merits comparison in terms of variables' contribution [42].

The evaluating indicators used in our dataset comprise growth vigor, harvest intensity, and protection. These variables were evaluated in the fieldwork and scored based on our experience. This approach [43] can achieve its intended purpose, but the accuracy and theoretical significance of the characteristics obtained remain questionable. Poor performance of these indicators might be why the management model had a minimal coefficient on

Fig. 7 The classification index with representative trees shown. The classification criteria were calculation-based, and experts' opinions were referred to for assistance

Zhang *et al. Plant Methods*   (2022) 18:95

Page 11 of 13

productivity. Evaluation of factors such as wildness [44], human activities, and the vegetation community [45] may also provide useful management insights. Although the total value is a straightforward measure for evaluating productivity, other significant indicators could be used in future studies, such as biomass above ground, net primary production, and the dry weight of harvested leaves [46].

The initial structural equation model is comprehensive but shows low fitness and performance because of the high number of variables. The initial model was built using the maximum likelihood technique, and the output parameters were as follows: chi-square (χ^2), 5987.978; degrees of freedom, 114; all fitness indexes less than 0.7; and AIC, 6065.978. These fitness parameters suggested that the model was unacceptable [32]. Thus, we retained the top two most important variables for the environmental suitability model. For the tree growth model, we omitted the other five parameters for the following reasons aside from their lower regression weights [23]: (1) tree age: numbers are reported by locals and are often rough estimates (more robust ways to measure it are required in the future, such as isotope tracing methods [47] and dendrochronology methods [48]); (2) DBH: most trunks bifurcate below breast height; (3) tree height: some managers artificially limit heights to promote the growth of well-developed canopies [23]; (4) growth vigor: it is quantified in only three classes and thus cannot clearly separate individuals; and (5) height under branch: it determines intervals for intercropping and thus does not strongly reflect tree growth conditions [49].

In quantifying productivity, we noticed that the curves (Fig. 2) between environmental suitability variables and logistic output have extreme points. This is reasonable given that every plant species has preferences for specific climate factors such as temperature and precipitation [50]. We used quadratic functions with different domain definitions to fit the relationships, given that we sought to minimize the model's redundancy. For this same reason, we used linear functions to formulate the productivity, tree growth, and management models. Simplified functions are efficient for evaluating data from many individuals. Regarding the criteria-based indexes, we categorized ancient tea trees into four levels based on outputs from the formulas (Fig. 6), our field evaluations, and local reports [23].

Bivariate relationships in the tree growth model (Fig. 3) were lacking in explanatory power for various possible reasons. First, the correlations between tree parameters and productivity are indirect. Because we used the total value to indicate productivity, the direct parameters and harvested tea parts (such as sprouts and leaves) are related. Although parameters such as crown width can reflect tree growth conditions and theoretically represent the production of trees, their linear relationships with total values might be weak. Another reason might be that we included many ancient tea trees varying in age in the study area. The tea of some trees over 500 years old may have remarkable prices but low output. Their size-related growth parameters are usually high compared with younger plants. However, young trees having lower growth parameters may have high output and low prices. In this case, their total values are close, but growth differs greatly. The large data size and skewed data distribution may also result in non-significant model fit for some parameters. We did not apply any transformations in terms of specific bivariate relationships because the original data are more logistic in explaining correlations in the final model, and we wanted to ensure that the entire dataset satisfied a multivariate normal distribution.

The significance of our model is that productivity can be determined based on measures from environments, trees, and management evaluations without harvesting the leaves of ancient tea trees. Managers can gain insight and make decisions by classifying trees into categories before harvesting. For example, the term "poorly productive" is more easily interpreted than the term "low net primary production" or "small biomass." "Highly productive" and "productive" trees can be harvested twice a year. "Poorly productive" trees can be harvested once a year. "Unproductive" trees should not be harvested to promote their healthy growth. However, we only used data from one county to build the formula and index. An analysis includes other regions with ancient tea trees distributed will greatly improve the robustness of the evaluation index. Our productivity model of ancient tea trees could also potentially be applied to other economically important tree plants, such as lacquer, fruit, and coffee trees.

The structural equation modeling approach for studying productivity is effective for exploring the significance of unobserved variables [51]. Quantitative modeling can facilitate the interpretation of abstract concepts. The indexes can aid interpretation of the practical significance of model outputs, which can then be used to modify management approaches accordingly. This approach can also be used to clarify other abstract concepts. For example, the heights and growth forms of trees have been quantified and then evaluated in indexed classes such as the Australian National Vegetation Information System [43].

Ancient tea trees are different from other ancient tree resources because of their high economic value and sustainable management models. Unlike other economically important tree species with unsustainable timber products, for example, precious wood [52], tea products are produced periodically. Hence, their conservation

and sustainable production require special attention. A quantitative modeling approach for evaluating productivity represents a first step in the study of ancient tea trees. Several outstanding research questions remain. For example, in addition to their economic and ecological significance, what are some of the political, societal, and cultural implications of ancient tea trees? How generalizable is the productivity model beyond ancient Pu'er tea trees? Are there differences between wild and cultivated ancient tea trees in terms of their productivity? Can Pu'er tea species be cultivated in other regions? Might it be possible to develop technology that allows normal tea products to possess the same tastes and scents of tea derived from ancient Pu'er tea trees? Can the production of ancient teas be increased sustainably? Such questions will require additional research.

In subsequent research, we plan to model the niches and distributions of ancient Pu'er tea species under future scenarios of climate change and globalization. We also plan to explore potentially suitable areas for Pu'er tea tree plantations in regions with tea cultures and plantation interests. Practical management actions need to be implemented in different regions and should be based on current management methodologies.

Conclusion

In this study, we used a structural equation modeling approach to quantify the productivity of tea production from ancient Pu'er tea trees and employed various indexes to permit qualitative evaluation. Overall, the model exhibited acceptable performance and permitted the identification of significant variables to include in the submodels. The final model suggested that environmental suitability, tree growth, and management positively affected productivity; regression weights were 0.32, 0.982, and 0.075, respectively. The index suggested 53% of the samples require management attention. In addition, our environmental suitability submodel revealed the optimal environmental conditions for ancient Pu'er tea trees. Modeling productivity is the first step to achieving the sustainable management and growing of Pu'er tea in potential interested regions. Future studies are needed to analyze specialized niches and distribution patterns under future climate change and the plausibility of establishing plantations in other areas.

The quantitative model and qualitative index provide a more robust approach for quantifying productivity compared with traditional correlation-based approaches. They also contribute to the local managers' interpretation and prediction of the trees' production before harvesting. In addition, the methodology might be applicable beyond ancient Pu'er tea trees (e.g., other economically valuable trees and plant resources requiring conservation attention) and could provide insights that facilitate the interpretation of abstract concepts, but its generality requires further examination.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author details
[1]Institute of Highland Forest Science, Chinese Academy of Forestry, Kunming 650216, Yunnan, China. [2]Fenner School of Environment & Society, College of Science, Australian National University, Canberra, ACT 2601, Australia. [3]Yunnan Institute of Forest Inventory and Planning, Kunming, Yunnan, China.

References

1. Nowicka B, Ciura J, Szymańska R, Kruk J. Improving photosynthesis, plant productivity and abiotic stress tolerance—current trends and future perspectives. J Plant Physiol. 2018;231:415–33.
2. Katila P, Pierce Colfer CJ, de Jong W, Galloway G, Pacheco P, Winkel G. Sustainable development goals: their impacts on forests and people. Cambridge: Cambridge University Press; 2019.
3. Ort DR, Merchant SS, Alric J, Barkand A, Blankenship RE, Bock R, et al. Redesigning photosynthesis to sustainably meet global food and bioenergy demand. Proc Natl Acad Sci. 2015;112(28):8529–36.
4. Grace JB, Anderson TM, Seabloom EW, Borer ET, Adler PB, Harpole WS, et al. Integrative modelling reveals mechanisms linking productivity and plant species richness. Nature. 2016;529(7586):390–3.
5. Van Nuland ME, Ware IM, Schadt CW, Yang Z, Bailey JK, Schweitzer JA. Natural soil microbiome variation affects spring foliar phenology with

Zhang *et al. Plant Methods*   (2022) 18:95

Page 13 of 13

consequences for plant productivity and climate-driven range shifts. New Phytol. 2021;232(2):762–75.

6. Forrester DI, Ammer C, Annighöfer PJ, Barbeito I, Bielak K, Bravo-Oviedo A, et al. Effects of crown architecture and stand structure on light absorption in mixed and monospecific *Fagus sylvatica* and *Pinus sylvestris* forests along a productivity and climate gradient through Europe. J Ecol. 2018;106(2):746–60.

7. Galbraith CS, Nkwenti-zamcho E. The effect of management policies on plant-level productivity: a longitudinal study of three US and Mexican small businesses. J Small Bus Manag. 2005;43(4):418–31.

8. Godlee JL, Ryan CM, Bauman D, Bowers SJ, Carreiras JMB, Chisingui AV, et al. Structural diversity and tree density drives variation in the biodiversity–ecosystem function relationship of woodlands and savannas. New Phytol. 2021;232(2):579–94.

9. Lu L, Chen H, Wang X, Zhao Y, Yao X, Xiong B, et al. Genome-level diversification of eight ancient tea populations in the Guizhou and Yunnan regions identifies candidate genes for core agronomic traits. Hortic Res. 2021. https://doi.org/10.1038/s41438-021-00617-9.

10. Brody H. Tea. Nature. 2019;566(7742):S1.

11. Chen Y, Li M. Evaluation of influencing factors on tea production based on random forest regression and mean impact value. Agric Econ. 2019;65(7):340–7.

12. Wu AH, Yu MC, Tseng CC, Hankin J, Pike MC. Green tea and risk of breast cancer in Asian Americans. Int J Cancer. 2003;106(4):574–9.

13. Inoue M, Tajima K, Mizutani M, Iwata H, Iwase T, Miura S, et al. Regular consumption of green tea and the risk of breast cancer recurrence: follow-up study from the hospital-based epidemiologic research program at Aichi Cancer Center (HERPACC). Japan Cancer Lett. 2001;167(2):175–82.

14. Cabrera C, Artacho R, Gimenez R. Beneficial effects of green tea—a review. J Am Coll Nutr. 2006;25(2):79–99.

15. Rogers PJ, Smith JE, Heatherley SV, Pleydell-Pearce CW. Time for tea: mood, blood pressure and cognitive performance effects of caffeine and theanine administered alone and together. Psychopharmacology. 2007;195(4):569–77.

16. Gilbert N. The science of tea's mood-altering magic. Nature. 2019;566(7742):S8–9.

17. Benn JA. Tea in China: a religious and cultural history. Honolulu: University of Hawaii Press; 2015.

18. Chen F, Hu L, Jiang A. The 40-year development of China tea industry. China Tea. 2019;41(10):1–5.

19. Zhang J. Puer tea: ancient caravans and urban chic. Seattle: University of Washington Press; 2014.

20. Jiang H, Tang Y, Chen L, Wang P, Cai X, Yu F, et al. Survey and analysis of ancient tea plant resources in Yunnan Province, China. J Plant Genet Resour. 2020;21(02):296–307.

21. Feenstra RC, Xu M, Antoniades A. What is the price of tea in China? Goods prices and availability in Chinese cities. Econ J. 2020;130(632):2438–67.

22. Jiang H, Lei B, Guo Y. Suitability analysis of planting Puer tea in Yunnan Province based on natural conditions. Chin J Trop Agric. 2013;33(10):19–23.

23. Liao S, Li L, Lv H, Wu W, Wang Z, Kong W, et al. Jinggu Dai and Yi Autonomous county ancient (wild) tea trees resources. Unpublished: Chinese Academy of Forestry Research Institute of Resource Insects; Yunnan Institute of Forest Inventory and Planning; 2019.

24. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 2017;37(12):4302–15.

25. Center RaESaD. Soil texture spatial distribution data in China resource and environment science and data center; 2021. www.resdc.cn.

26. NASA. Advanced land observing satellite-1 EARTHDATA; 2021. https://search.asf.alaska.edu.

27. Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME. Opening the black box: an open-source release of Maxent. Ecography. 2017;40(7):887–93.

28. Kumar S, Neven LG, Zhu H, Zhang R. Assessing the global risk of establishment of *Cydia pomonella* (Lepidoptera: Tortricidae) using CLIMEX and MaxEnt Niche models. J Econ Entomol. 2015;108(4):1708–19.

29. Sobek-Swant S, Kluza DA, Cuddington K, Lyons DB. Potential distribution of emerald ash borer: what can we learn from ecological niche models using Maxent and GARP? For Ecol Manag. 2012;281:23–31.

30. Von Toussaint U, Preuss R. MaxEnt 2019—proceedings, 2019, MaxEnt 2019 the 39th international workshop on Bayesian inference and maximum entropy methods in science and engineering: MDPI—Multidisciplinary Digital Publishing Institute; 2020.

31. Jones B, Sall J. JMP statistical discovery software. Wiley Interdiscip Rev Comput Stat. 2011;3(3):188–94.

32. Blunch NJ. Introduction to structural equation modelling using IBM SPSS statistics and AMOS. 2nd ed. Los Angeles: SAGE; 2013.

33. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190(3):231–59.

34. Nolan V, Reader T, Gilbert F, Atkinson N. The Ancient Tree Inventory: a summary of the results of a 15 year citizen science project recording ancient, veteran and notable trees across the UK. Biodivers Conserv. 2020;29(11–12):3103–29.

35. Reyer CPO, Leuzinger S, Rammig A, Wolf A, Bartholomeus RP, Bonfante A, et al. A plant's perspective of extremes: terrestrial plant responses to changing climatic variability. Glob Change Biol. 2013;19(1):75–89.

36. Hung P-Y. Tea production, land use politics, and ethnic minorities: struggling over dilemmas in China's Southwest frontier. Houndmills: Palgrave Macmillan; 2015.

37. Gale J. Plants and altitude—revisited. Ann Bot. 2004;94(2):199.

38. Hanson JO, Rhodes JR, Butchart SHM, Buchanan GM, Rondinini C, Ficetola GF, et al. Global conservation of species' niches. Nature (London). 2020;580(7802):232–4.

39. Yackulic CB, Chandler R, Zipkin EF, Royle JA, Nichols JD, Campbell Grant EH, et al. Presence-only modelling using MAXENT: when can we trust the inferences? Methods Ecol Evol. 2013;4(3):236–43.

40. Furbank RT, Silva-Perez V, Evans JR, Condon AG, Estavillo GM, He W, et al. Wheat physiology predictor: predicting physiological traits in wheat from hyperspectral reflectance measurements using deep learning. Plant Methods. 2021;17(1):108.

41. Hill L, Hector A, Hemery G, Smart S, Tanadini M, Brown N. Abundance distributions for tree species in Great Britain: a two-stage approach to modeling abundance using species distribution modeling and random forest. Ecol Evol. 2017;7(4):1043–56.

42. Meyer H, Reudenbach C, Wöllauer S, Nauss T. Importance of spatial predictor variable selection in machine learning applications—moving from data reproduction to spatial prediction. Ecol Model. 2019;411: 108815.

43. Bolton MP, deLacey C, Bossard KBE. Australian vegetation attribute manual: national vegetation information system, version 7.0. Department of the Environment and Energy, Canberra: NVIS Technical Working Group; 2017.

44. Ridder B. The naturalness versus wildness debate: ambiguity, inconsistency, and unattainable objectivity. Restor Ecol. 2007;15(1):8–12.

45. Price JN, Pärtel M, Pillar VD, Chytrý M. Restoration and management of plant communities in applied vegetation science. Appl Veg Sci. 2022;25(1):1–6.

46. Johnson MO, Galbraith D, Gloor M, De Deurwaerder H, Guimberteau M, Rammig A, et al. Variation in stem mortality rates determines patterns of above-ground biomass in Amazonian forests: implications for dynamic global vegetation models. Glob Change Biol. 2016;22(12):3996–4013.

47. Hancock GJ, Murray AS, Brunskill GJ, Argent RM. Ra isotopes in trees: their application to the estimation of heartwood growth rates and tree ages: RA isotopes in trees. Glob Biogeochem Cycles. 2006. https://doi.org/10.1029/2005GB002641.

48. Koch J. Improving age estimates for late Holocene glacial landforms using dendrochronology—some examples from Garibaldi Provincial Park, British Columbia. Quat Geochronol. 2009;4(2):130–9.

49. Shen Y, Xiang Y, Xu G, Ge X, Li Z. Major co-localized QTL for plant height, branch initiation height, stem diameter, and flowering time in an alien introgression derived *Brassica napus* DH population. Front Plant Sci. 2018;9:390.

50. Kfoury N, Scott ER, Orians CM, Ahmed S, Cash SB, Griffin T, et al. Plant-climate interaction effects: changes in the relative distribution and concentration of the volatile tea leaf metabolome in 2014–2016. Front Plant Sci. 2019;10:1518.

51. Grace JB, Anderson TM, Seabloom EW, Borer ET, Adler PB, Harpole WS, et al. Integrative modelling reveals mechanisms linking productivity and plant species richness. Nature (London). 2016;529(7586):390–3.

52. Waeber PO, Schuurman D, Ramamonjisoa B, Langrand M, Barber CV, Innes JL, et al. Uplisting of Malagasy precious woods critical for their survival. Biol Conserv. 2019;235:89–92.

Publisher's Note