

文章编号: 1001-1498(2001)03-0271-07

马尾松管胞长度分布规律及 计算机随机模拟

侯祝强, 姜笑梅, 殷亚方

(中国林业科学研究院 木材工业研究所, 北京 100091)

摘要: 根据 20 年生人工林马尾松早晚材管胞长度测量结果, 运用随机变量正态性检验的偏斜系数和峰态系数方法, 系统地分析了 6 株马尾松标准试材测量样品不同组合的偏斜系数和峰态系数。结果表明, 马尾松木材的管胞长度是服从正态的随机变量, 3 株试材测量样品就可揭示出其正态分布特性。另外, 利用计算机直接随机抽样的方法对马尾松管胞长度进行了模拟, 模拟结果的平均值具有较高的精度, 由此可以获得一个关于研究木材细胞分子解剖结构参数的新方法。

关键词: 马尾松; 管胞长度; 随机抽样模拟

中图分类号: S781.1 **文献标识码:** A

马尾松(*Pinus massoniana* Lamb.) 是我国南方大面积栽种的针叶材树种, 它在工业生产特别是在造纸和人造纤维的生产中有着广泛的应用。马尾松作为一个较为重要的人工林树种, 曾经对其解剖性质展开了一系列研究工作。王章荣等^[1]研究了马尾松木材的密度、管胞长度和晚材率等在林分间和林分内个体的变异情况; 周志春等^[2]分析了马尾松幼林材密度和管胞长度在种源内和种源间遗传变异及地理环境对其的影响; 姜笑梅^[3,4]对马尾松幼林材和成熟材的多种解剖结构参数进行了系统全面的研究。上述有关的研究工作, 为马尾松材性的选育和木材的加工利用提供了科学的参考和依据。

不论是在同一树种内还是在不同的树种之间, 木材的解剖结构参数变异都很大。如何针对木材解剖性质随机多变的特点, 在当前计算机应用及其技术不断普及和发展的条件下, 寻求木材解剖性质研究工作的新方法, 这是值得认真考虑的问题。本文将从分析讨论马尾松管胞长度的分布规律出发, 利用计算机随机抽样的方法, 对马尾松管胞长度的变化特性进行模拟, 以探索关于树木性状与木材解剖性质研究的新方法。

1 材料与方法

试材采集地点是广西壮族自治区凭祥市中国林科院热带林业实验中心大青山林场, 在生长了 20 a 的马尾松人工林中, 选择 6 棵正常生长的树, 从距地面 2.3 m 处向上各截取 1.5 m 长的树段, 计 6 根, 各试材树干(湿材)大头直径 19~24 cm, 小头直径 18~22 cm。试材自树干截锯之

收稿日期: 2000-01-08

基金项目: 国家攀登计划预选项目子课题——人工林木材性质特点与规律研究的部分内容

作者简介: 侯祝强(1956-), 男, 河北南宫人, 副教授, 博士。

后,马上在两端锯口处涂上防腐药物以避免霉变。试材自采集地点运回后,立即解锯成用作加工测量试件的毛坯。用于管胞长度测量的各个测量样品,均是从6棵试材下端的10 cm处开始,按50 cm相等间隔在3个不同的高度、距髓心第15~16个年轮处截取,共计18个样品,按常规方法分早、晚材离析、测量。管胞长度的测量结果是根据18个样品,分早晚材各随机选取50根管胞进行测量所得。即每一棵试材随机选取测量早晚材各150个管胞长度,6棵试材共随机选取测量了早材和晚材各900个管胞长度。

2 马尾松管胞长度分布规律的确定

姜笑梅根据马尾松人工林和天然林解剖性质的研究结果指出^[3],其幼林材和成熟材管胞长度分布频率呈正态分布。本文视马尾松管胞长度为一系列随机变量的集合,应用国家标准所推荐使用的随机变量正态性检验的“偏斜系数和峰态系数法”^[5],根据各株试材的样品以及由不同数目试材的样品组合,逐一对照马尾松管胞长度的测量结果的偏斜系数和峰态系数进行分析计算,以确定马尾松管胞长度的分布规律。

根据随机变量正态性检验的偏斜系数和峰态系数法^[5],若总体的偏斜系数和峰态系数都接近于零,则从实用的角度来说,可以认为总体服从正态分布。反之,当总体的偏斜系数和峰态系数有一个不接近零时,则总体不服从正态分布。此外,当总体的偏斜系数和峰态系数不知道时,可以通过样本来进行估计。若设随机变量 X_1, \dots, X_n 是总体的一个样本,则样本的偏斜系数 m_1 和峰态系数 m_2 分别为

$$m_1 = B_3 / B_2^{3/2} \quad (1)$$

$$m_2 = B_4 / B_2^2 - 3 \quad (2)$$

(1)式和(2)式中的 B_2 是样本的方差(二阶中心距), B_3 和 B_4 分别是样本的三阶距和四阶距,即

$$B_2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (3)$$

$$B_3 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3 \quad (4)$$

$$B_4 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4 \quad (5)$$

(3)式、(4)式和(5)式中的 n 是样本的数目, \bar{X} 是样本的平均值。

根据本文关于马尾松试材样品管胞长度的测量结果,利用(1)式、(2)式、(3)式、(4)式和(5)式所得到的偏斜系数和峰态系数,如表1~3所示。表1为从6株试材中随机抽取单株试材组成的测量样品组,早材和晚材测量样品数各为150,早晚材测量样品总数为300;表2为从6株试材中随机抽取3株试材组成测量样品组,早材和晚材测量样品数各为450,早晚材测量样品总数为900;表3为从6株试材中随机抽取5株试材组成测量样品组,早材和晚材测量样品数各为750,早晚材测量样品总数为1500;全部6株试材组成的测量样品组合的偏斜系数和峰态系数,早材和晚材测量样品数各为900,早晚材测量样品总数为1800;且早材的 m_1 和 m_2 为-0.263和-0.343,晚材的 m_1 和 m_2 为0.333和-0.293,早晚材的 m_1 和 m_2 为-0.004和-0.302。

表 1 单株试材测量样品组合的偏斜系数(m_1)和峰态系数(m_2)

样品组别		1	2	3	4	5	6
早 材	m_1	0.297	0.305	0.196	0.008	0.912	0.271
	m_2	-0.192	-0.233	-0.416	-0.713	1.146	-0.350
晚 材	m_1	0.298	0.266	0.289	0.229	0.764	0.248
	m_2	-0.348	-0.465	0.702	-0.644	0.045	-0.618
早晚材	m_1	0.201	0.324	0.115	0.027	0.768	0.244
	m_2	-0.266	-0.271	0.037	-0.475	0.476	-0.441

表 2 3 株试材测量样品组合的偏斜系数(m_1)和峰态系数(m_2)

样品组别		1	2	3	4	5	6
早 材	m_1	0.231	0.534	0.292	0.169	0.331	0.090
	m_2	-0.511	0.452	-0.257	-0.305	-0.150	-0.479
晚 材	m_1	0.399	0.424	0.175	0.220	0.399	0.449
	m_2	0.021	-0.427	-0.654	-0.563	-0.274	0.080
早晚材	m_1	0.253	0.189	0.431	0.253	0.247	0.091
	m_2	-0.044	-0.408	-0.148	-0.044	-0.316	-0.495

表 3 5 株试材测量样品组合的的偏斜系数(m_1)和峰态系数(m_2)

样品组别		1	2	3	4	5	6
早 材	m_1	0.188	0.389	0.314	0.180	0.225	0.292
	m_2	-0.375	-0.248	-0.186	-0.418	-0.380	-0.335
晚 材	m_1	0.304	0.351	0.383	0.234	0.225	0.247
	m_2	-0.304	-0.306	-0.255	-0.437	-0.380	-0.426
早晚材	m_1	0.211	0.349	0.210	0.312	0.192	0.079
	m_2	-0.260	-0.242	-0.355	-0.178	-0.420	-0.375

根据表 1 的结果, 单株试材的早材测量样品中第 4 组的 m_2 、第 5 组的 m_1 与 m_2 , 晚材测量样品中第 3 组的 m_2 、第 5 组的 m_1 , 早晚材测量样品中第 5 组的 m_1 的绝对值均大于 0.7。根据文献[5]提出的一个实用标准, 当 m_1 和 m_2 的绝对值大于 0.7 时即可判定其偏离零较远, 故可推知单株马尾松管胞长度测量结果的分布与正态分布有差异。根据表 2 和表 3 的结果, 3 株试材的早材、晚材、早晚材测量样品组合中 m_1 和 m_2 的绝对值均不超过 0.7, 5 株试材的早材、晚材、早晚材测量样品组合中 m_1 和 m_2 的绝对值均不超过 0.5, 这表明选取 3 株试材测量管胞长度就已基本能反映出马尾松早材、晚材及整株管胞长度正态分布的特性, 而使用 5 株标准试材则可较好地反映出马尾松早材、晚材及整株的管胞长度正态分布的特性。同时, 根据 6 株试材测量样品的 m_1 和 m_2 , 可知其能更好地反映出马尾松管胞长度服从正态分布的特性。测量样品越多则越能反映出马尾松管胞长度所服从的分布规律, 样品的偏斜系数和峰态系数随测量试材数目增大而减小, 这一结果证实了马尾松管胞长度变化服从正态分布的客观事实。

3 马尾松管胞长度的计算机抽样模拟

3.1 正态分布随机数的计算机抽样

一株马尾松试材包含着数目巨大的管胞,各个管胞的长度都不相同而且是随机变化的,对于这样数目巨大的随机变量系统,通常只能进行抽样测量并依据抽样测量结果对试材管胞长度进行统计描述。随着科学的进步与发展,越来越要求我们对于研究对象进行更为具体的量化描述,而计算机随机模拟则是处理这类随机变量系统或随机现象的一个有效的方法和工具^[6],计算机随机模拟(又称为蒙特卡罗方法)在包括木材科学在内的许多科学研究领域中已有广泛的应用^[5-8]。本节将在管胞长度服从正态分布的前提下,根据马尾松管胞长度的测量结果,运用样本参数的平均值和方差估算总体的平均值和方差,以构成管胞长度的正态分布密度函数。然后,按照管胞长度的分布密度函数利用计算机进行直接随机抽样,其所得的随机数即代表着马尾松管胞长度的值,从而实现对于马尾松管胞长度的计算机抽样模拟。

当标准正态分布的随机数 $U \sim N(0, 1)$ 已知时,那么由正态分布样本的平均值 μ 和方差 σ^2 有 $X = \sigma U + \mu$, 并且 $X \sim N(\mu, \sigma^2)$ ^[7]。本文关于标准正态分布随机变量的计算机抽样,是基于均匀分布随机变量的抽样方法^[7],所使用的均匀分布随机数发生器则为大部分统计软件包中常用的素数模积式发生器;同时,本文采用抽样效率较高的“极坐标”抽样法实现标准正态分布的随机数的抽样。用于计算机直接抽样的随机数发生器,使用 Turbo C 语言编写在 32 位微机上实现,其抽样序列的周期不小于 2.147×10^9 ^[7]。

3.2 马尾松管胞长度的抽样模拟结果

在 3 株测量样品组、5 株测量样品组、6 株测量样品组中任意选取早材、晚材和早晚材管胞长度的各一组平均值和标准差,使用正态分布随机数发生器进行管胞长度的计算机随机抽样模拟。表 4 所列出的各测量样品组平均值与标准差,作为估计值用作随机抽样的总体正态分布密度函数的参数,表 8 所列为计算机抽样模拟的平均值及其与测量结果平均值的比较。由于模拟数据数目太大无法逐一列出,表 5~7 分别为早材、晚材、早晚材管胞长度模拟结果的分布频率,为了比较也列出了相应测量结果的分布频率。

表 4 随机抽样的样品组管胞长度的平均值与标准差

样品组	项目	3 株试材测量样品组	5 株试材测量样品组	6 株试材测量样品组
早材管胞	平均值/ μm	4 277.40	4 461.80	4 383.90
	标准差	785.51	815.78	816.98
晚材管胞	平均值/ μm	4 678.70	4 874.60	4 721.10
	标准差	737.55	783.01	855.31
早晚材管胞	平均值/ μm	4 484.90	4 672.30	4 581.50
	标准差	772.07	814.68	836.63

将模拟结果与测量结果作一比较可知,在远离样品组管胞长度平均值的区域,模拟结果与测量结果间有一定的差异;而在较接近样品平均值的区域,模拟结果与测量结果间差异不大。另外,各个样品组模拟结果的平均值与测量结果的平均值非常接近,相对误差都不大于 0.55%,这表明马尾松管胞长度的计算机抽样模拟结果,具有足够的精度,完全可以满足实际应用的需要。这也表明本文利用测量样品的有关参数,使用计算机随机抽样的方法模拟马尾松管胞长度是可行的。

表5 早材管胞长度测量值和模拟值的分布频率

管胞长度/ μm	3株试材测量样品组		5株试材测量样品组		6株试材测量样品组	
	测量结果	抽样结果	测量结果	抽样结果	测量结果	抽样结果
$X > 7\ 000$	0	0.002	0.001	0.001	0.001	0.001
$6\ 500 < X < 7\ 000$	0.009	0.004	0.020	0.007	0.019	0.006
$6\ 000 < X < 6\ 500$	0.049	0.031	0.073	0.024	0.062	0.017
$5\ 500 < X < 6\ 000$	0.060	0.100	0.096	0.051	0.083	0.049
$5\ 000 < X < 5\ 500$	0.176	0.207	0.201	0.145	0.176	0.114
$4\ 500 < X < 5\ 000$	0.256	0.251	0.248	0.209	0.226	0.197
$4\ 000 < X < 4\ 500$	0.253	0.218	0.223	0.213	0.214	0.198
$3\ 500 < X < 4\ 000$	0.149	0.120	0.105	0.188	0.140	0.208
$3\ 000 < X < 3\ 500$	0.047	0.060	0.031	0.096	0.066	0.122
$2\ 500 < X < 3\ 000$	0.002	0.007	0.001	0.065	0.012	0.089
$X < 2\ 500$	0	0	0	0	0.001	0

表6 晚材管胞长度测量值和模拟值的分布频率

管胞长度/ μm	3株试材测量样品组		5株试材测量样品组		6株试材测量样品组	
	测量结果	抽样结果	测量结果	抽样结果	测量结果	抽样结果
$X > 7\ 000$	0	0	0.001	0.003	0.001	0.002
$6\ 500 < X < 7\ 000$	0.006	0.001	0.013	0.010	0.013	0.009
$6\ 000 < X < 6\ 500$	0.030	0.006	0.047	0.035	0.041	0.031
$5\ 500 < X < 6\ 000$	0.052	0.021	0.087	0.106	0.079	0.089
$5\ 000 < X < 5\ 500$	0.137	0.247	0.168	0.195	0.152	0.176
$4\ 500 < X < 5\ 000$	0.234	0.227	0.232	0.224	0.212	0.222
$4\ 000 < X < 4\ 500$	0.254	0.217	0.226	0.210	0.228	0.215
$3\ 500 < X < 4\ 000$	0.177	0.164	0.148	0.127	0.171	0.146
$3\ 000 < X < 3\ 500$	0.090	0.077	0.067	0.069	0.088	0.076
$2\ 500 < X < 3\ 000$	0.019	0.032	0.011	0.021	0.015	0.034
$X < 2\ 500$	0.001	0.009	0.001	0	0.001	0

表7 早晚材管胞长度测量值和模拟值的分布频率

管胞长度/ μm	3株试材测量样品组		5株试材测量样品组		6株试材测量样品组	
	测量结果	抽样结果	测量结果	抽样结果	测量结果	抽样结果
$X > 7\ 000$	0	0	0.001	0.003	0.001	0.002
$6\ 500 < X < 7\ 000$	0.006	0.001	0.013	0.010	0.013	0.009
$6\ 000 < X < 6\ 500$	0.030	0.006	0.047	0.035	0.041	0.031
$5\ 500 < X < 6\ 000$	0.052	0.021	0.087	0.106	0.079	0.089
$5\ 000 < X < 5\ 500$	0.137	0.247	0.168	0.195	0.152	0.176
$4\ 500 < X < 5\ 000$	0.234	0.227	0.232	0.224	0.212	0.222
$4\ 000 < X < 4\ 500$	0.254	0.217	0.226	0.210	0.228	0.215
$3\ 500 < X < 4\ 000$	0.177	0.164	0.148	0.127	0.171	0.146
$3\ 000 < X < 3\ 500$	0.090	0.077	0.067	0.069	0.088	0.076
$2\ 500 < X < 3\ 000$	0.019	0.032	0.011	0.021	0.015	0.034
$X < 2\ 500$	0.001	0.009	0.001	0	0.001	0

表 8 管胞长度模拟结果和测量结果平均值

样品组	早材测量样品			晚材测量样品			早晚材测量样品		
	3 株	5 株	6 株	3 株	5 株	6 株	3 株	5 株	6 株
测量值/ μm	4 277	4 462	4 384	4 679	4 875	4 721	4 485	4 672	4 582
模拟值/ μm	4 291	4 439	4 367	4 691	4 851	4 704	4 469	4 662	4 566
误差/ $\%$	0.31	0.52	0.38	0.27	0.48	0.37	0.35	0.22	0.35

4 小结与讨论

本文对 6 株马尾松管胞长度不同的测量样品组合所进行的偏斜系数和峰态系数的分析表明, 马尾松管胞长度是一系列服从正态分布的随机变量。选取 3 株以上的标准木试材就可较好地揭示其正态分布特性, 选取 5 株标准木则能更好地描述其正态分布特性。根据马尾松管胞长度抽样测量的样品平均值和标准差, 使用计算机进行直接随机抽样的方法可以对马尾松管胞长度进行模拟, 且模拟所得的平均值具有极好的精度, 可以实际用于科研与生产中。要指出的是, 除马尾松管胞长度服从正态分布之外, 以往的一些研究工作已经证明一些针叶材细胞分子其它的几何参数^[9, 10], 如管胞的径向直径和弦向直径、纹孔和纹孔膜的大小也服从正态分布。

试材所包含的木材细胞分子数目巨大, 同时其大小形状又是无规多变的, 寻求木材细胞分子几何参数的定量描述方法, 具有理论和实际的应用意义。要指出的是, 正因为木材细胞分子数目巨大且其形状无规多变, 从而可将其视为一个服从某一统计分布规律的随机变量系统。在 32 位微机上进行直接抽样模拟, 其不重复的抽样样本可达 20 多亿个, 若以马尾松木材管胞数目来计的话, 对应着平均胸径 30 cm、长 12 m 左右的木材所包含的管胞, 这样的抽样模拟范围完全能够满足实际应用的需要。本文所得到的模拟结果表明, 在确定木材分子的统计分布规律后运用计算机进行直接抽样模拟, 是定量描述木材细胞解剖结构参数的一个有效的方法和途径。

参考文献:

- [1] 王章荣, 陈天华, 周志春, 等. 马尾松木材性状在林分间和林分内个体间的变异[J]. 南京林业大学学报, 1988, (2): 38~42.
- [2] 周志春, 金国庆, 秦国峰. 马尾松幼林材密度、管胞长度的地理遗传变异及形状相关[J]. 林业科学研究, 1990, 3(4): 393~397.
- [3] 姜笑梅. 人工林和天然林马尾松幼林材与成熟材解剖性质比较研究[J]. 世界林业研究, 1995, 8(专集): 1~9.
- [4] 姜笑梅. 人工林和天然林马尾松木材解剖性质差异及在径向上的变异的比较研究[J]. 世界林业研究, 1995, 8(专集): 10~19.
- [5] 方开泰, 许建伦. 统计分布[M]. 北京: 科学出版社, 1987. 321~332.
- [6] 裴鹿成. 蒙特卡罗方法及其应用[M]. 长沙: 国防科技大学出版社, 1993. 1~25.
- [7] 高惠璇. 统计计算[M]. 北京: 北京大学出版社, 1995. 98~101, 146~148.
- [8] Taylor S E, Triche M H, Bender D A, et al. Monte Carlo simulation methods for engineering wood system [J]. Forest Prod J, 1995, 45(7/8): 43~50.
- [9] Sirvio P, Karenlampi P. Pits as natural irregularities in softwood fibers [J]. Wood and Fiber Science, 1998, 30(1): 27~39.
- [10] Vysotskaya, Vaganov E A. Components of the variability of radial cell size in tree rings of conifers [J]. IAWA Bulletin n. s., 1989, 10(4): 417~428.

The Probability Distribution and Sampling Simulation of the Tracheid Length in Masson Pine

HOU Zhu-qiang, JIANG Xiao-mei, YIN Ya-fang

(Research Institute of Wood Industry, CAF, Beijing 100091, China)

Abstract: The tracheid length in six tested logs from 20-year-old plantation-grown tree of masson pine (*Pinus massoniana* Lamb.) were studied. The skew factors and kurtosis factors were evaluated to test the probability distribution towards the tracheid length from the measurements of different sample combinations. Results showed that the length of tracheid in masson pine can be regarded as the random variable and is normally distributed. The property of normal distribution could be revealed just with three woods tested. The computer simulation of tracheid length was made by means of random sampling. There was great agreement between the means of tracheid length from the simulation and the ones from measurement and the relative error was below 0.55%.

Key words: masson pine; tracheid length; sampling simulation

中国林科院惠刚盈研究员荣获台湾“刘业经教授奖励基金”

“刘业经教授奖励基金”每年奖励在林业中作出重要贡献的科技人员。在第五届的获奖名单中,中国林科院惠刚盈研究员榜上有名,是获奖者中最年轻的一位。

惠刚盈研究员 1961 年 2 月出生于陕西富平。1983 年毕业于西北林学院,同年分配到中国林科院从事科研工作。1996 年在德意志学术交流中心的资助下赴德国留学,在不到两年的时间里,以优异的成绩获得了德国哥廷根大学林学博士学位。2000 年荣获第 21 届国际林联 (IUFRO) 杰出博士研究奖,成为国际上获此殊荣的七位杰出青年科学家之一。

惠刚盈研究员在林业科学研究方面成绩显著。在国内外著名期刊上发表了 30 多篇科学论文,在国外出版专著 3 部;完善和发展了林分生长与收获模型理论,提出了人工林经营模型的构建方法;在天然林林分空间结构研究方面取得重大突破,提出描述林分空间结构的新指标和方法,特别是首次成功解决了不用测距就可获得林木在地域上的分布格局这一世界性难题。