

# 杉木种子涩籽地理分布的空间插值法研究

吴承祯 洪伟\* 林思祖

(福建农林大学森林生态研究所, 福建福州 350002)

**摘要:**应用距离反比插值法进行杉木种子涩籽率的空间插值,并提出一种改进的距离反比插值法(MID),其包含了距离反比法(ID)和距离平方反比法(IDS)。根据福建省25个杉木种子产地的涩籽率观察值,建立了杉木种子涩籽率的改进距离反比插值模型。改进距离反比插值法经交叉验证,其平均误差及平均误差平方的平方根均较距离反比法和距离平方反比法更小,具有较高的精度,平均精度达88.91%,可应用于杉木种子涩籽的空间内插。这不仅为杉木种子涩籽的空间插值提供了一种新方法,而且为地理信息系统(Geographic information system, GIS)辅助的杉木种子涩籽区域分布与流行趋势研究提供了理论依据,为福建省杉木种子园的合理布局与优化奠定基础。

**关键词:**空间插值;距离反比法;杉木;涩籽;地理分布

中图分类号:S722.11

文献标识码:A

文章编号:1005-3395(2004)04-0318-05

## Studies on Interpolation Method for Estimating the Geographical Distribution of Sterile Seeds of *Cunninghamia lanceolata*

WU Cheng-zhen HONG Wei\* LIN Si-zu

(Institute of Forest Ecology, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

**Abstract:** The inverse distance method (ID) was used to interpolate the sterile seed rate of *Cunninghamia lanceolata* distributed in various sites in Fujian Province in order to establish rational seed production bases. A modified inverse distance method (MID) is presented which includes ID and inverse distance squares method (IDS). Based on sterile seeds collected from 25 sites, models of MID, ID and IDS were built. The results of cross-validation for the three methods showed that mean error and root-mean-square interpolation error in MID were lowest, and the average accuracy being 88.91%. MID could be useful in the study of epidemic trend of *C. lanceolata* sterile seeds and its distribution, as well as for establishing seed production base in optimum habitats.

**Key words:** Spatial interpolation; Inverse distance method; *Cunninghamia lanceolata*; Sterile seeds; Geographical distribution

杉木(*Cunninghamia lanceolata*)为我国特有树种,也是我国南方主要用材造林树种,栽培于南方16个省市、自治区,每年需要大量杉木苗木。但由于杉木种子涩籽比例较大(30%-60%),严重影响种子播种品质,造成种子短缺、苗木不足等后果。涩籽是种子不孕性的一种形式,是种子发芽率低的根本原因。研究杉木涩籽的形成原因、预测与防治方法不仅是杉木种子园合理布局的基础,也是杉木种子生产中急需解决的问题。俞新妥<sup>[1]</sup>、杜宏彬等<sup>[2-4]</sup>对杉

木种子涩籽形成的原因进行了初步探讨,认为不利的生态因素包括气象、病虫害等。而建立杉木种子涩籽地理流行规律模型是种子涩籽预测与防治的重要途径,是探讨涩籽在地理学方面的发生、发展和分布趋势规律的重要手段,具有涩籽流行病学研究的理论意义。为此,林思祖等<sup>[5]</sup>采用趋势面分析法探讨杉木种子涩籽地理分布的数学模型,但由于经济和人的原因,收集的涩籽观察值有限,在许多栅格杉木种子涩籽率的大小无法获得,这在一定程

收稿日期:2003-09-25 接受日期:2004-02-23

基金项目:中国博士后科研基金;福建省科技厅重大项目(2001F007、2001Z025);福建省自然科学基金(B0110026)资助

\* 通讯作者 Corresponding author

度上影响了模型的可靠性和精度。在实际工作中,利用邻近区域的杉木种子涩籽量的观察资料,通过空间插值生成所需要的栅格区域的杉木种子涩籽量,就成为一种有效的解决方法。

空间插值方法有距离权重法、多项式插值法、克里格法、样条插值法等<sup>[4-9]</sup>。在这些方法中,距离权重法最为简便;多项式值的物理学意义不是很明确,容易得出一些难以解释的值;样条插值是对一些限定的点值,通过控制估计方差,利用一些特征节点,用多项式拟合的方法产生平滑的插值曲线,多用于观察要素的时间序列插值;克里格方法产生于地质采矿学中的品位估计,以能提供最佳线性无偏估计而逐渐运用于需要空间插值的诸多领域<sup>[10,11]</sup>。然而,并非所有这些方法均适用于同一研究背景的最佳插值。选用一种相对适合而且便于运用的方法,对杉木种子涩籽的空间插值是非常重要的。在开展地理信息系统(Geographic information system, GIS)辅助的杉木种子涩籽区域分布与流行

趋势研究中,首要的任务就是栅格尺度的杉木种子涩籽场的生成。为此,笔者<sup>[12]</sup>曾采用地统计学方法探讨了杉木种子涩籽分布规律的内插,收到理想效果。本文将探讨距离反比法在杉木种子涩籽插值研究中的应用,并提出一种改进的距离反比插值法。

## 1 资料和方法

### 1.1 资料来源

调查区域福建省地处我国东南沿海,位于东经115°50′-120°43′,北纬23°33′-28°19′之间,东西宽约540 km,南北宽约550 km,属杉木适生区范围。收集福建省25个杉木种子产地(县、市)的杉木种批,供试种子取量250-500 g,每一产地种批随机抽取并解剖500粒种子。以50粒种子为一重复,计算每个重复的涩籽数,并求该产地平均涩籽流行百分率(简称涩籽率),将收集的不同产地供试种批的经纬度及涩籽率列于表1<sup>[9]</sup>。

表1 基础数据及插值估计结果

Table 1 Basic data of seed collection and estimated results of interpolation

产地 Sites	东经 Longitude (E)	北纬 Latitude (N)	涩籽率 Sterile seeds (%)	理论值 Theoretical value (%)		
				距离反比法 Inverse distance method	距离平方反比法 Inverse distance squares method	改进距离反比法* Modified inverse distance method
平和 Pinghe	117°19′	24°22′	42.9	40.47	35.03	33.90
南靖 Nanjing	117°22′	24°31′	33.4	43.82	42.99	42.89
华安 Huaan	117°32′	25°00′	42.8	43.13	41.40	40.48
安溪 Anxi	118°11′	25°04′	43.7	44.42	44.22	44.03
龙岩 Longyan	117°02′	25°06′	40.8	43.35	42.27	41.88
连城 Liancheng	116°45′	25°43′	44.2	43.06	41.40	40.38
长汀 Changting	116°22′	25°51′	37.2	44.16	43.91	43.94
永泰 Yongtai	118°56′	25°52′	64.5	45.30	48.68	51.17
闽侯 Minhou	119°09′	26°09′	56.4	46.70	51.70	55.22
尤溪 Youxi	118°09′	26°10′	44.9	44.36	43.79	43.30
三明 Sanming	117°37′	26°16′	41.4	44.16	43.46	42.93
明溪 Mingxi	117°09′	26°24′	44.8	44.03	43.93	43.93
沙县 Shaxian	117°48′	26°24′	41.2	44.28	43.73	43.35
古田 Gutian	118°44′	26°35′	42.8	45.61	46.38	46.72
南平 Nanping	118°10′	26°39′	41.2	44.22	43.36	42.86
将乐 Jiangle	117°28′	26°44′	45.2	44.48	45.15	45.71
顺昌 Shunchang	117°48′	26°48′	48.2	43.60	42.96	42.77
建宁 Jianning	116°51′	26°50′	44.2	44.61	45.81	46.76
霞浦 Xiapu	119°59′	26°53′	50.8	46.22	49.24	50.93
泰宁 Taining	117°10′	26°54′	49.2	43.99	44.21	44.42
屏南 Pingnan	118°59′	26°55′	48.6	44.62	44.15	43.68
柘荣 Zherong	119°54′	27°15′	54.0	44.81	45.27	45.59
建阳 Jianyang	118°07′	27°20′	42.2	43.34	40.10	37.68
建瓯 Jian'ou	118°19′	27°03′	32.2	44.93	44.12	43.56
寿宁 Shouning	119°25′	27°32′	35.6	46.81	49.23	50.61

\* 改进距离法中的  $\alpha=2.63$ 。  $\alpha$  in the modified inverse distance method is 2.63.

### 1.2 方法

距离反比插值法实质上是以插值点与样点间的距离为权重的一种加权平均法,其权重赋予是离插值点越近的样点赋予的估值权重越大。当权重由距离反比给出,称为距离反比法(Inverse distance method, ID);由距离平方的反比给出,则称为距离平方反比法(Inverse distance squares method, IDS),两者计算公式如下<sup>[11]</sup>:

$$Z = \sum_{i=1}^n \frac{Z_i}{d_i} / \sum_{i=1}^n \frac{1}{d_i} \tag{1}$$

$$Z = \sum_{i=1}^n \frac{Z_i}{d_i^2} / \sum_{i=1}^n \frac{1}{d_i^2} \tag{2}$$

式中  $Z$  为所要估计的栅格点的杉木种子涩籽率;  $Z_i$  为杉木种子涩籽在第  $i$  个产地的观察值;  $d_i$  为插值点到第  $i$  产地的距离;  $n$  为用于插值的杉木种子涩籽产地的样点数目。

在(1)、(2)式中,权重是以距离反比和距离平方反比的形式出现,是否存在更好的权重赋予方式值得进一步的探讨。或者说,权重以距离反比和距离平方反比的方式并不一定是最优的内插法,故本文提出一个改进的距离反比方法(Modified inverse distance method, MID),权重以距离的  $\alpha$  次方成反比,具体插值计算公式为:

$$Z = \sum_{i=1}^n \frac{Z_i}{d_i^\alpha} / \sum_{i=1}^n \frac{1}{d_i^\alpha} \tag{3}$$

式中  $\alpha$  为待定参数,表明在进行杉木种子涩籽率插值计算时,权重为距离  $\alpha$  次方反比给出,其它变量意义同前。由(1)-(3)式不难得出,当  $\alpha=1$  时,改进插值法(3)即为距离反比法(1)式;当  $\alpha=2$  时,改进插值法(3)式即为距离平方反比法(2)式,因此,改进距离反比插值法是距离反比法和距离平方反比法的拓展,距离反比法和距离平方反比法仅为其的两个特例。 $\alpha$  值的确定与被内插对象的特性有关,可根据研究对象的样本观察数据进行估计。在对  $\alpha$  进行估计时,可以一定的评价标准为目标函数,如残差平方和最小为目标函数,采用遗传算法、改进单纯形法和三次设计等最优估计。这里采用遗传算法对  $\alpha$  值进行估计,遗传算法的具体计算方法与过程参阅文献[13]和[14]。

### 1.3 检验标准

采用交叉验证法验证插值的效果<sup>[12]</sup>,即首先假定每一产区的杉木种子涩籽率未知,都用其周围产

区的观察值进行估算,然后计算所有产地实际值与估算值的误差,以此评判估值方法的优劣。采用平均误差(Mean error, ME)、平均绝对误差(Mean absolute error, MAE)、插值平均误差平方的平方根(Root mean squared interpolation error, RMSIE)作为评估插值方法的插值效果的标准<sup>[12]</sup>。平均误差总体反映估计误差的大小,平均绝对误差可以估量估计值可能的误差范围,插值平均误差平方的平方根可以反映利用样点数据的估值灵敏度和极值效应。其中插值平均误差平方的平方根的计算式为:

$$RMSIE = \sqrt{\frac{\sum_{i=1}^n (Z_i - \hat{Z}_i)^2}{n}} \tag{4}$$

式中  $Z_i$  为第  $i$  个产地的杉木种子涩籽率的实际观察值;  $\hat{Z}_i$  为其估计值;  $n$  为用于检验的产地数目。

## 2 结果和分析

对模型(3)中的  $\alpha$  值的求解采用遗传算法,即以调查获得的 25 个杉木产地的种子涩籽率观察值为基础资料,以空间内插交叉验证残差平方和最小为目标函数,编制遗传算法求解  $\alpha$  值的计算机程序进行求解。通过遗传算法的复制、交换与变异等遗传操作,优化得到  $\alpha$  的取值为 2.63,相应的目标函数值为 974.33。所以得到杉木种子涩籽率的空间插值计算公式为:

$$Z = \sum_{i=1}^n \frac{Z_i}{d_i^{2.63}} / \sum_{i=1}^n \frac{1}{d_i^{2.63}} \tag{5}$$

由此还可以发现,当以残差平方和最小为目标函数时,改进的距离反比插值法的距离的次数  $\alpha$  为 2.63,既不等于 1 也不等于 2,因此,改进方法较距离反比法和距离平方反比法具有更大的灵活性和适用性。

根据距离反比插值法的 2 个计算公式及笔者提出的改进模型,对福建省 25 个杉木种子产地的种子涩籽率进行空间内插的交叉验证(表 1)。交叉验证结果表明(表 2),杉木种子涩籽率插值的 ME 绝对值大小排序为  $IDS > ID > MID$ ,其值分别为 5.91%、3.92%、3.71%;MAE 的排序为  $MID > ID > IDS$ ,但相差较小,依次为 11.094%、11.014%、10.997%;RMSIE 的排序为  $ID > IDS > MID$ ,其值为 6.72%、6.32%、6.24%。MID 法获得了最低的 ME 和 RMSIE,且其 MAE 值与其它两种方法非常接近,因此可以认为

本文提出的改进模型较距离反比法和距离平方反比法更合理、更科学,可以推广应用于有关气象要素、地理要素及生态学要素的空间内插。

表 2 3 种插值方法的交叉验证误差

Table 2 Cross-validation errors resulted from the three interpolation methods

方法 Method	ME (%)	MAE (%)	RMSIE (%)
Inverse distance	3.92	11.014	6.72
Inverse distance squares	5.91	10.997	6.32
Modified inverse distance	3.71	11.094	6.24

ME: Mean error; MAE: Mean absolute error; RMSIE: Root-mean-squared interpolation error.

由于改进的距离反比插值法较距离反比法和距离平方反比法空间插值结果更好,其平均相对误差为 11.094%,平均估计精度为 88.906%,说明用改进的距离反比法进行栅格点的杉木种子涩籽率的

空间插值估计是可行的。因此,可以利用所建立的杉木种子涩籽空间插值改进模型(5)式进行不同栅格点的预测预报,从而进行 GIS 辅助的杉木种子涩籽区域分布与流行趋势研究。

### 3 结语

涩籽是种子不孕性的一种形式,涩籽量的高低直接影响种子的品质和种子的发芽率,因此涩籽量的空间分布规律研究在杉木种子园布局建设中具有重要意义。开展 GIS 辅助的杉木种子涩籽区域分布与流行趋势研究是探讨杉木种子涩籽空间分布规律的重要措施,但往往受调查产地数量的制约,需要采用一定的空间插值法进行未知栅格点杉木种子涩籽率的内插,本项研究为开展这一领域研究工作提出了新思路。当研究区域较大时,直接采用经、纬度值进行距离计算将存在较大的距离误差,

表 3 不同插值方法结果比较

Table 3 Comparison between two interpolation methods

产地 Sites	涩籽率 Sterile seeds (%)	改进距离法 Modified inverse distance method			人工神经网络方法 Artificial neural net method		
		模拟值 Simulated value (%)	绝对误差 Absolute error (%)	相对误差 Relative error (%)	模拟值 Simulated value (%)	绝对误差 Absolute error (%)	相对误差 Relative error (%)
平和 Pinghe	42.9	33.9	9.00	20.98	50.08	7.18	16.74
南靖 Nanjing	33.4	42.89	9.49	28.41	39.98	6.58	19.70
华安 Huaan	42.8	40.48	2.32	5.42	49.52	6.72	15.70
安溪 Anxi	43.7	44.03	0.33	0.76	50.84	7.14	16.34
龙岩 Longyan	40.8	41.88	1.08	2.65	48.33	7.53	18.46
连城 Liancheng	44.2	40.38	3.82	8.64	46.93	2.73	6.18
长汀 Changting	37.2	43.94	6.74	18.12	45.98	8.78	23.60
永泰 Yongtai	64.5	51.17	13.33	20.67	51.13	13.37	20.73
闽侯 Minhou	56.4	55.22	1.18	2.09	51.12	5.28	9.36
尤溪 Youxi	44.9	43.30	1.60	3.56	48.88	3.98	8.86
三明 Sanming	41.4	42.93	1.53	3.70	47.74	6.34	15.31
明溪 Mingxi	44.8	43.93	0.87	1.94	46.78	1.98	4.42
沙县 Shaxian	41.2	43.35	2.15	5.22	47.85	6.65	16.14
古田 Gutian	42.8	46.72	3.92	9.16	49.53	6.73	15.72
南平 Nanping	41.2	42.86	1.66	4.03	48.25	7.05	17.11
将乐 Jiangle	45.2	45.71	0.51	1.13	46.74	1.54	3.41
顺昌 Shunchang	48.2	42.77	5.43	11.27	47.3	0.90	1.87
建宁 Jianning	44.2	46.76	2.56	5.79	45.47	1.27	2.87
霞浦 Xiapu	50.8	50.93	0.13	0.26	51.75	0.95	1.87
泰宁 Taining	49.2	44.42	4.78	9.72	45.98	3.22	6.54
屏南 Pingnan	48.6	43.68	4.92	10.12	49.51	0.91	1.87
柘荣 Zherong	54.0	45.59	8.41	15.57	47.31	5.11	9.46
建阳 Jianyang	42.2	37.68	4.52	10.71	50.94	3.06	7.25
建瓯 Jianou	32.2	43.56	11.36	35.28	37.95	5.75	17.86
寿宁 Shouning	35.6	50.61	15.01	42.16	39.73	4.13	11.60
平均值 Mean	-	-	-	11.09	-	-	11.60

因此需要将其转化为平面距离。此时,须计算每一产区到某一参考点的大圆距离和方位,将极坐标系转换成以该参考点为原点的  $X$ - $Y$  平面坐标系, $X$  轴代表东西向, $Y$  轴代表南北向。具体转换方法参阅文献[6]。

何东进等<sup>[16]</sup>曾采用人工神经网络方法研究杉木种子涩籽地理流行规律,探讨了杉木种子涩籽率的预测与预报方法,建立的 BP 模型模拟效果略差于本文提出的改进距离反比插值法(表 3)。且本文提出的改进的空间插值法操作简单、易于应用,是一种理想的插值法。应用改进的距离反比插值法可对福建省以经纬度  $1' \times 1'$  栅格进行杉木种子涩籽的空间插值,从而生成福建杉木种子涩籽率的空间分布栅格图,得到福建省杉木种子涩籽地理流行区域分布图,为杉木种子园的布局提供了理论依据。

距离反比插值法由于方法简单,在气象要素及自然经济地理变量的空间插值中具有广泛应用前景;本文提出的改进距离反比插值法是距离反比法的拓展,距离反比法仅为改进方法的特例。对于改进距离反比法中的待定参数  $\alpha$ ,其求解必须依赖于一定区域的已知栅格点的观察值,通过以一定的目标函数为收敛准则,可采用遗传算法、改进单纯形法及三次设计等最优化方法求解得到。在对杉木种子涩籽率改进距离反比插值法参数  $\alpha$  的求解时,本文以残差平方和最小为目标函数,收到较好效果。根据研究的需要,还可采用平均误差、平均绝对误差和平均误差平方的平方根等统计学中的评判指标作为目标函数。同时,由于距离反比插值法仅仅依据周围已知栅格点观察值对未知栅格点的数值进行距离加权内插,因此在一定程度上对局部区域的真实的较小值和较大值的插值估计误差较大。这是距离反比插值法的一大缺点,在进行 GIS 辅助的杉木种子涩籽区域分布与流行趋势研究中可采用其它方法加以克服,此内容及空间插值生成的涩籽地理流行区域分布图将另外报道。

### 参考文献

- [1] Yu X T(俞新妥). Primary report on forming of Chinese fir shriveled seed [J]. J Fujian Coll For (福建林学院学报), 1960, (1):1-7. (in Chinese)
- [2] Du H B (杜宏彬). Preliminary study on quality and sterile seed of Chinese fir [J]. Subtrop For Sci Techn (亚热带林业科技), 1980, (4):42-48. (in Chinese)
- [3] He F J(何福基). Study on the cause of formation on Chinese fir sterile seed [J]. Seed (种子), 1986, 20(6):4-6. (in Chinese)
- [4] Lin S Z(林思祖), You S L(游松龄), Lin L L(林如兰). Forecasting models for the epidemic of the Chinese fir sterile seed [J]. J Fujian Coll For (福建林学院学报), 1990, 10(1):72-78. (in Chinese)
- [5] Lin S Z(林思祖), Song S X(宋淑霞). A study on the trend of geographic epidemic of Chinese fir sterile seeds [J]. Acta Phytocool Geobot Sin (植物生态学与地植物生态学学报), 1990, 14(1): 40-45. (in Chinese)
- [6] Ashraf M, Loftis J C, Hubbard K G. Application of geostatistics to evaluate partial station networks [J]. Agri For Meteorol, 1997, 84: 225-271.
- [7] Gottschalk L, Batchvarova E, Gryning S E, et al. Scale aggregation comparison of flux estimates from Nopex [J]. Agri For Meteorol, 1999, 98-99:103-119.
- [8] Nalder I A, Wein R W. Spatial interpolation of climate normals: test of a new method in the Canadian boreal forest [J]. Agri For Meteorol, 1998, 92:211-225.
- [9] Price D T, Mckenney D W, Nalder I A, et al. A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data [J]. Agri For Meteorol, 2000, 101:81-94.
- [10] Li X (李新), Cheng G D (程国栋), Lu L (卢玲). Comparison of spatial interpolation methods [J]. Adv Earth Sci(地球科学进展), 2000, 15(3):260-265. (in Chinese)
- [11] Lin Z H(林忠辉), Mo X G(莫兴国), Li H X(李宏轩), et al. Comparison of three spatial interpolation methods for climate variables in China [J]. Acta Geogr Sin (地理学报), 2002, 57(1) 47-56. (in Chinese)
- [12] Wu C Z(吴承祯), Hong W(洪伟), Lin S Z(林思祖). Study on spatial variability of Chinese fir sterile seeds [J]. Chin J Appl Envir Biol(应用与环境生物学报), 1998, 4(1):15-19. (in Chinese)
- [13] Wu C Z(吴承祯), Hong W(洪伟). An improved method for afforestation planning and design under restricted conditions using the genetic algorithm [J]. Sci Sil Sin (林业科学), 1997, (2):133-141. (in Chinese)
- [14] Wu C Z(吴承祯), Hong W(洪伟). A study on density optimum control during stand management for plantation [J]. J Natl Reso (自然资源学报), 1998, 13(2):175-180. (in Chinese)
- [15] Holdaway M R. Spatial modeling and interpolation of monthly temperature using Kriging [J]. Clim Res, 1996, 24:1835-1845.
- [16] He D J(何东进), Hong W(洪伟), Wu C Z(吴承祯). Study on model of geographic epidemic of Chinese fir sterile seeds [J]. Sil Sin (林业科学), 1999, 35(6):58-62. (in Chinese)