



林元震, 张卫华, 程玲, 等. 基于空间效应与竞争效应的林木遗传分析模型[J]. 华南农业大学学报, 2017, 38(5): 74-80.

基于空间效应与竞争效应的林木遗传分析模型

林元震¹, 张卫华², 程玲¹, 张心菲¹, 张鑫鑫^{1,2}

(1 华南农业大学 林学与风景园林学院/广东省森林植物种质创新与利用重点实验室, 广东 广州 510642;
2 广东省林业科学研究院, 广东 广州 510520)

摘要:【目的】建立空间效应与竞争效应的分析模型, 以提高林木遗传评估的准确性。【方法】利用 R 软件及其程序包 breedR 模拟数据, 结合实测数据, 采用 XFA1 结构拟合加性效应、近邻竞争效应和 AR1 结构拟合空间效应, 建立随机区组设计模型(RCBM)、空间模型(SM)、空间与测量误差模型(SUM)和空间与竞争模型(SCM), 运行 ASReml 估算遗传参数, 进行模型比较。【结果】对于模拟数据, 估计的参数结果均显示 SCM 是最优模型, 其大大降低了随机误差方差, 随机误差方差分别由 7.56(RCBM)、5.72(SUM)降低到 3.13(SCM), 分别降低了 58.6%、45.3%, 并估算到四周近邻竞争方差; SCM 模型估算的单株狭义遗传力在 0.40 左右, 高于 RCBM(0.24)和 SUM 模型(0.30); 设置参数的不同初始值, SCM 估计的参数结果均较为稳定; 对于实测数据, 估算结果与模拟结果比较一致。【结论】SCM 模型是新的单株混合模型, 可用于林木遗传分析。

关键词: 空间效应; 竞争效应; 单株混合模型; REML; 遗传分析
中图分类号: S722.5 文献标识码: A 文章编号: 1001-411X(2017)05-0074-07

Genetic analysis model of forest based on space and competition effects

LIN Yuanzhen¹, ZHANG Weihua², CHENG Ling¹, ZHANG Xinfei¹, ZHANG Xinxin^{1,2}

(1 College of Forestry and Landscape Architecture, South China Agricultural University/Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm, Guangzhou 510642, China;
2 Guangdong Academy of Forestry, Guangzhou 510520, China)

Abstract: 【Objective】To improve the accuracy in genetic analysis of forest establishing an analysis model based on space and competition effects. 【Method】The data was simulated by R software and its package breedR. The additive effect and neighbor competition effect were fitted using XFA1 structure and the spatial effect was fitted using AR1 structure for both simulated and measured data. Four models (randomized block design model, RCBM; spatial model, SM; spatial with measured error model, SUM; spatial and competition model, SCM) were established and analyzed using ASReml to estimate genetic parameters for comparison. 【Result】The estimated results showed that SCM was the best model for the simulation data. SCM greatly reduced the random error variance from 7.56 (RCBM) and 5.72 (SUM) to 3.13 (SCM), decreased by 58.6% and 45.3%, respectively. SCM could estimate the genetic variance of competition from surrounding neighbors. The individual heritability assessed by SCM was around 0.40, higher than those of RCBM (0.24) and SUM (0.30). SCM obtained stable estimated results under different settings of initial values for parameters. Furthermore, for the measured data, the estimated results were consistent with the simulation data. 【Conclusion】SCM is a new individual-tree mixed model and could be used for genetic analysis of forest.

Key words: spatial effect; competition effect; individual-tree mixed model; REML; genetic analysis

林木遗传改良通常需在田间试验基础上, 估算参试样本(种源、家系、单株或无性系)目标性状的遗传参数, 并运用一定的评价方法筛选出具有优异目标性状的种源、家系或无性系^[1]。目前, 林木遗传评估基本采用混合线性模型, 即根据观测表型值和设计的关联矩阵, 利用最佳线性无偏估计(Best linear unbiased estimation, BLUE)方法获得固定效应估计值, 利用最佳线性无偏预测(Best linear unbiased prediction, BLUP)方法获得随机效应预测值。而育种值(Breeding value)即属随机效应预测值, 且 BLUP 值依赖于混合线性模型的关联设计矩阵。然而, 对林木田间试验而言, 个体间存在竞争, 且基因型与环境互作较为显著, 即空间效应明显。如竞争效应和空间效应没有被考虑到混合线性模型中, 遗传评估则会产生偏差, 其中空间正相关易导致加性遗传方差偏大, 中等竞争效应则抑制偏差^[2]。因而, 在林木遗传评估模型中, 应纳入竞争效应和空间效应。

空间效应(或称环境异质性)可以是空间连续的, 其表现为土壤和小气候效应的相似模式, 也可以是不连续的, 表现为育林措施和测量方法的不同效应, 此外还可以是随机的, 其表现为微环境的异质性^[3]。针对空间效应的空间分析(Spatial analysis)已被广泛用于农作物品种田间试验研究。与传统试验设计相比, 空间分析可减少试验误差并提高农作物品种评估的准确性。空间分析方法通常有趋势面分析法和空间自相关分析法等。其中, 趋势面分析法是采用空间坐标的多项式函数来模拟环境的变异^[4], 而空间自相关分析法则通过行和列的空间自相关分析剔除地点内的空间相关误差来提高遗传试验分析的精确性^[5]。

在林木研究上, Costa 等^[6]利用空间自相关分析法分析了辐射松 *Pinus radiata* 等 3 个树种的子代和无性系试验林, 发现空间自相关分析可更有效拟合试验林的空间变异, 显著降低了区组和小区的方差, 且极大提高了遗传评估的准确度与遗传增益。另外, Dutkowski 等^[7-8]在更大数据量的树种与试验林上进一步采用空间分析, 也发现空间分析可显著提高种源、家系、亲本和无性系遗传效应估算的准确性。但林木是多年生植物, 除空间差异外, 往往还存在个体竞争^[9]。虽然 Cappa 等^[10]采用遗传干预模型(Genotypic interference model, GIM)评估了林木遗传加性效应和竞争加性效应, 但是没有考虑空间

效应, 因此模型比较受限。值得注意的是, Stringer 等^[11]提出空间效应与竞争效应的联合模型进行糖蔗品种试验的遗传分析, 模型采用等根二阶自回归法(Equal-roots second-order autoregressive, EAR2)拟合环境的空间效应, 而遗传的竞争效应则采用处理干预模型(Treatment interference model, TIM)。此外, Cappa 等^[12]提出拟合空间效应与竞争效应的林木单株混合模型, 并采用贝叶斯法估计遗传参数。上述模型均在一定程度上提高了林木遗传评估的准确性, 但模型涵盖的遗传竞争效应均为单一值, 而实际上林木个体周边的遗传竞争效应可能不同, 因此有必要提出新的单株混合模型, 同时拟合不同的遗传竞争效应和异质的空间效应。基于此, 本研究根据现代线性统计理论, 构建了异质空间效应和不同竞争效应的联合模型, 并通过数据模拟的方式, 结合实际测定数据进行验证, 同时运用 REML 法估算遗传参数, 对不同模型进行比较, 进而探讨异质空间效应和不同竞争效应的分析模型在林木遗传测定中的有效性, 以期为林木精确的遗传评估和后续可靠的良种选育提供理论参考和技术支撑。

1 方法

1.1 统计模型

单株混合模型为: $y = Xb + Z_g u_g + e$ 。式中: y 为个体性状表型值构成的向量; b 为固定效应构成的向量; u 为随机遗传效应构成的向量, $u'_g = (u'_{gd}, u'_{gen}, u'_{gwn}, u'_{gsn}, u'_{gnn})$, u'_{gd} 、 u'_{gen} 、 u'_{gwn} 、 u'_{gsn} 和 u'_{gnn} 分别为直接遗传效应、东向近邻效应、西向近邻效应、南向近邻效应和北向近邻效应; e 为剩余误差向量; X 为固定效应的关联矩阵; Z_g 为随机效应的关联矩阵。

遗传效应的方差-协方差矩阵为:

$$\text{var}(u_g) = \begin{bmatrix} \delta^2_{gd} & \delta_{gden} & \delta_{gdwn} & \delta_{gdsn} & \delta_{gdnn} \\ \delta_{gden} & \delta^2_{gen} & \delta_{gewn} & \delta_{gesn} & \delta_{genn} \\ \delta_{gdwn} & \delta_{gewn} & \delta^2_{gwn} & \delta_{gwsn} & \delta_{gwnn} \\ \delta_{gdsn} & \delta_{gesn} & \delta_{gwsn} & \delta^2_{gsn} & \delta_{gsnn} \\ \delta_{gdnn} & \delta_{genn} & \delta_{gwnn} & \delta_{gsnn} & \delta^2_{gnn} \end{bmatrix} \otimes I_m,$$

其中, 对角线上依次为直接遗传效应、东向近邻效应、西向近邻效应、南向近邻效应和北向近邻效应的方差分量, 非对角线为上述两两效应之间的协方差分量。 var 代表方差, δ^2 和 δ 分别代表方差和协方差, I_m 代表单位矩阵。

为减少计算量和促进模型收敛, 可通过因子分析法^[1]将遗传效应的方差-协方差矩阵改写为如下

矩阵:

$$\text{var}(\boldsymbol{u}) = (\boldsymbol{\Gamma} \times \boldsymbol{\Gamma}' + \boldsymbol{\Psi}) \otimes \boldsymbol{A},$$

其中, $\boldsymbol{\Gamma}$ 是因子载荷矩阵, $\boldsymbol{\Psi}$ 是特殊方差矩阵, $\boldsymbol{A}$ 是加性遗传相关矩阵。

在常规的空间分析^[7]中, 剩余误差(R)可分解为空间自相关误差方差(δ_{ξ}^2)和随机测量误差方差(δ_{η}^2), 因此剩余误差的方差协方差矩阵为:

$$\boldsymbol{R} = \delta_{\xi}^2 \left(\sum_c (\rho_c) \otimes \sum_r (\rho_r) \right) + \delta_{\eta}^2 \boldsymbol{I},$$

其中, $\sum_c (\rho_c)$ 、 $\sum_r (\rho_r)$ 分别是列、行自回归的相关矩阵, ρ_c 、 ρ_r 分别是列、行自相关值, $\boldsymbol{I}$ 是单位矩阵。

$\sum(\rho)$ 自回归的相关矩阵, 具体如下:

$$\sum(\rho) = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \vdots \\ \rho^2 & \rho & 1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \cdots & \cdots & \cdots & 1 \end{bmatrix}.$$

1.2 数据模拟方法

采用 R 软件^[1]和 R 程序包 `breedR`^[13] 进行数据模拟。为简化数据模拟, 参考下述松树数据, 以该松树胸径均值(13)和估算的加性方差(δ_a^2 为 2.46)、剩余残差(δ_e^2 为 3.00)作为模拟初始值, 4 次重复, 5 株行式小区, 10 个父本, 10 个母本, 共 30 行、40 列。为简化不同模型的比较, 设置加性遗传效应与竞争遗传效应相关值[简称竞争相关值(C_r)]为-0.6, 空间行、列自相关值分别为 0.6、-0.6, 比较模型包括随机完全区组设计模型(RCBM)、空间模型(SM)、空间与测量误差模型(SUM)和空间与竞争模型(SCM)。为探究遗传竞争效应和环境空间效应的影响, 设置竞争相关值(C_r =-0.3, -0.6, -0.9)以及空间行相关值(ρ_r =0.3, 0.6, 0.9)、列自相关值(ρ_c =-0.3, -0.6, -0.9), 通过正交设计^[1], 产生 9 组数据。

1.3 实测数据

数据来源于文献[1]中的例 6.4.4.2, 该数据有 46 个全同胞家系, 采用完全随机区组设计, 设 7 个区组, 5 株行式小区, 试验林共 35 行 40 列, 测定性状是 10 年生胸径。

1.4 模型分析方法

所有模型分析均采用 ASReml 软件 V4.0^[14], 除总体均值作为固定因子外, 所有其他试验因子均作为随机因子。空间相关的残差采用行列 AR1 自相关结构, 遗传竞争效应的方差采用单因子 XFA1 结构。拟合模型的原始数据图和残差图采用 R 包 `AAfun`^[15] 生成, 空间变异图 `variogram` 由 ASReml 运行生成, 所有估计的参数由 ASReml 运行获得。

单株狭义遗传力(h_i^2)的计算公式如下:

$$h_i^2 = \delta_a^2 / (\delta_a^2 + \delta_e^2),$$

式中, δ_a^2 为加性方差, δ_e^2 为随机误差方差(剩余残差)。

AIC 值的计算公式如下:

$$\text{AIC} = -2 \times (\ln L - P),$$

式中, AIC 值是赤池信息量准则, $\ln L$ 是模型迭代收敛似然值, P 是模型中随机参数的数量。

2 结果与分析

2.1 不同模型比较

对模拟数据设置随机区组设计模型(RCBM)、空间模型(SM)、空间与测量误差模型(SUM)和空间与竞争模型(SCM)4 种模型, 模型的拟合结果(表 1)显示: 模型 RCBM 的参数最少(4 个), 收敛似然值($\ln L$)最小(-1 927), AIC 值最大(3 863), 剩余残差(7.56)也最大, 所估算的单株狭义遗传力也最小(0.24±0.08); 模型 SM 的参数为 6 个, 收敛似然值($\ln L$)稍微增加, 为-1 904, AIC 值下降至 3 820, 空间

表 1 模拟数据不同模型估计参数的比较¹⁾

Tab. 1 Comparison of estimated parameters from different models for simulated data

模型	参数/ 个	lnL	AIC	δ_{rep}^2	δ_{plot}^2	δ_e^2	δ_{η}^2	δ_{ξ}^2	ρ_r	ρ_c	δ_a^2	δ_{gen}^2	δ_{wn}^2	δ_{gsn}^2	δ_{gmn}^2	h_i^2
RCBM	4	-1 927	3 863	0.06 ^{0.85}	0.10 ^{1.03}	7.56 ^{13.52}					2.37 ^{2.67}					0.24±0.08
SM	6	-1 904	3 820	0.08 ^{0.81}	0.07 ^{0.80}			7.04 ^{8.92}	0.25 ^{8.92}	-0.07 ^{5.87}	3.44 ^{2.52}					...
SUM	7	-1 897	3 808	0.07 ^{0.89}	0.08 ^{0.84}		5.72 ^{9.02}	1.83 ^{4.21}	0.61 ^{5.84}	-0.70 ^{8.00}	2.43 ^{2.70}					0.30±0.10
SCM	10	-1 853	3 734				3.13 ^{4.74}	2.20 ^{4.75}	0.53 ^{5.24}	-0.70 ^{9.63}	2.37 ^{2.81}	0.59 ^{2.22}	0.34 ^{1.60}	0.68 ^{2.39}	0.74 ^{2.54}	0.43±0.13

1) $\ln L$ 为迭代收敛似然值, AIC 为赤池信息量准则, δ_{rep}^2 为区组方差, δ_{plot}^2 为小区方差, δ_e^2 为剩余残差, δ_{η}^2 、 δ_{ξ}^2 分别为随机测量误差方差和空间自相关误差方差, ρ_r 、 ρ_c 分别为行、列自相关值, δ_a^2 为加性方差, δ_{gen}^2 、 δ_{wn}^2 、 δ_{gsn}^2 、 δ_{gmn}^2 分别为东向、西向、南向和北向近邻遗传竞争方差, h_i^2 为单株狭义遗传力; 表中的上标值是 t 检验统计量, 为估计值与标准误的比值; ... 表示单株遗传力因没有随机测量误差而不能计算。

自相关误差为 7.04, 加性方差为 3.44, 比 RCBM 提高了 45%, 但估计的行列相关值远低于模拟设置值, 行自相关值为 0.25, 列自相关值为-0.07; 模型 SUM 的参数中, 随机测量误差方差为 5.72, 空间自相关误差方差为 1.83, 行列相关值与模拟设置值一致, 加性方差为 2.43, 估计的单株狭义遗传力为 0.30 ± 0.10 ; 模型 SCM 参数最多(10 个), 似然值($\ln L$)最大(-1 853), 但 AIC 值(3 734)和随机测量误差方差(3.13)均最小, 估计的单株狭义遗传力最高, 为 0.43 ± 0.13 。此外, 模型 SCM 估计的 4 个近邻效应方差在 0.34~0.74, 并且均显著。根据似然值 $\ln L$ 最大和 AIC 值最小原则, 可知 SCM 为最佳模型。

残差图显示, RCBM 的残差图(图 1C)与原始数据图(图 1A)非常相似, 表明 RCBM 模型未能很好

地拟合空间异质性。虽然 SM 模型比 RCBM 有所改进, 但也只拟合了残差小的空间部分(图 1D)。SUM 和 SCM 模型, 图形比较相似, 有明显的行列空间趋势, 但 SCM 模型更能消除部分残差高的区域(图 1E、1F)。从空间变异图可以看出, SM 模型的图形比较平坦(图 2B), 表明捕获的空间变异很小, 而 SUM 和 SCM 模型的图形比较相似(图 2C、2D), 在行方向上存在明显的空间变异趋势, 在列方向的突起尖峰表示列相关为负值, 存在环境的竞争效应, 表明 SUM 和 SCM 模型捕获空间变异的效果不错。此外, 残差值与近邻表型均值的散点图(图 2A)显示两者存在明显的负相关($r=-0.30$), 表明存在遗传竞争效应。上述结果与表 1 的结果相一致。

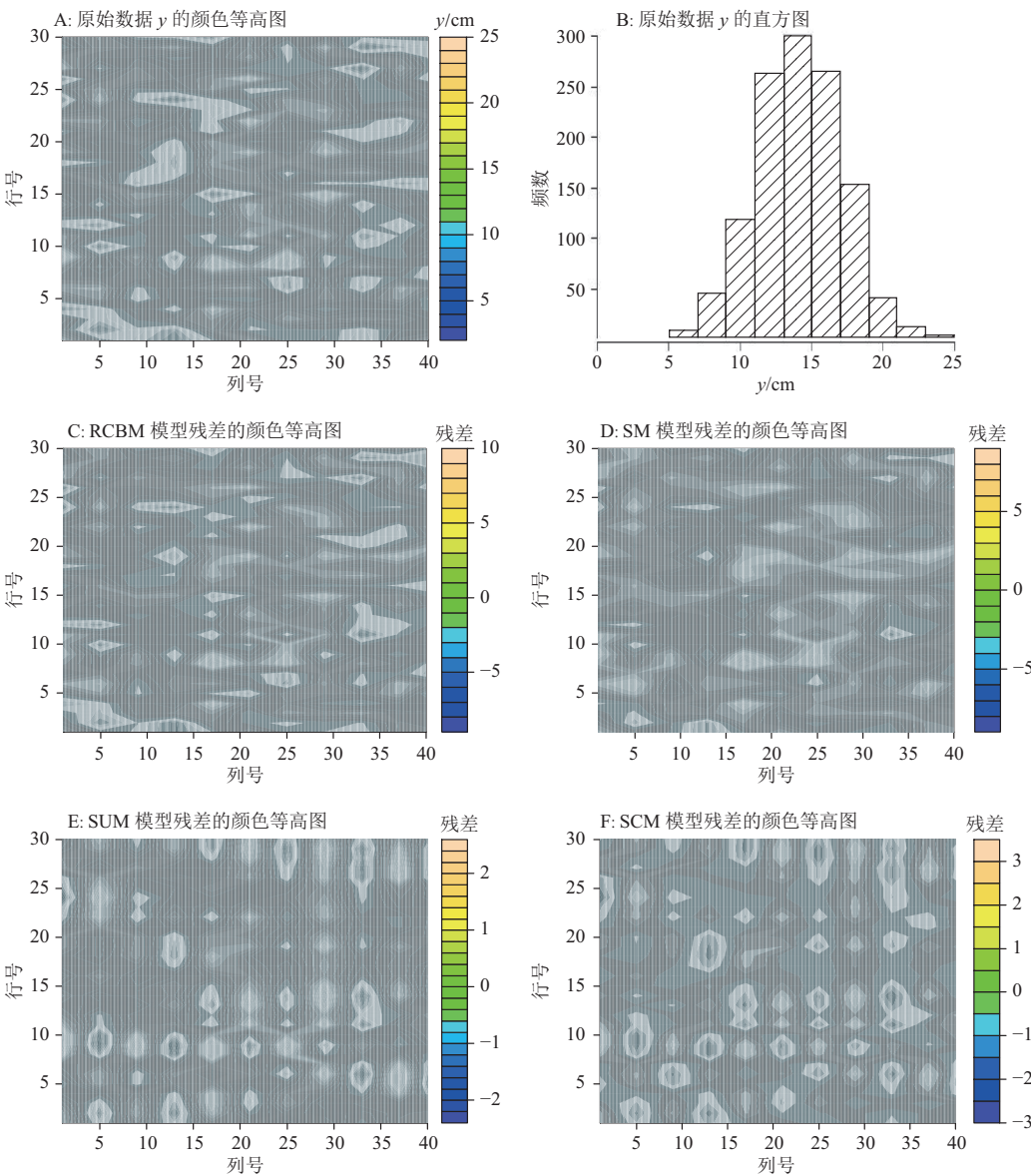


图 1 不同模型的剩余残差图

Fig. 1 Residual plot of different model

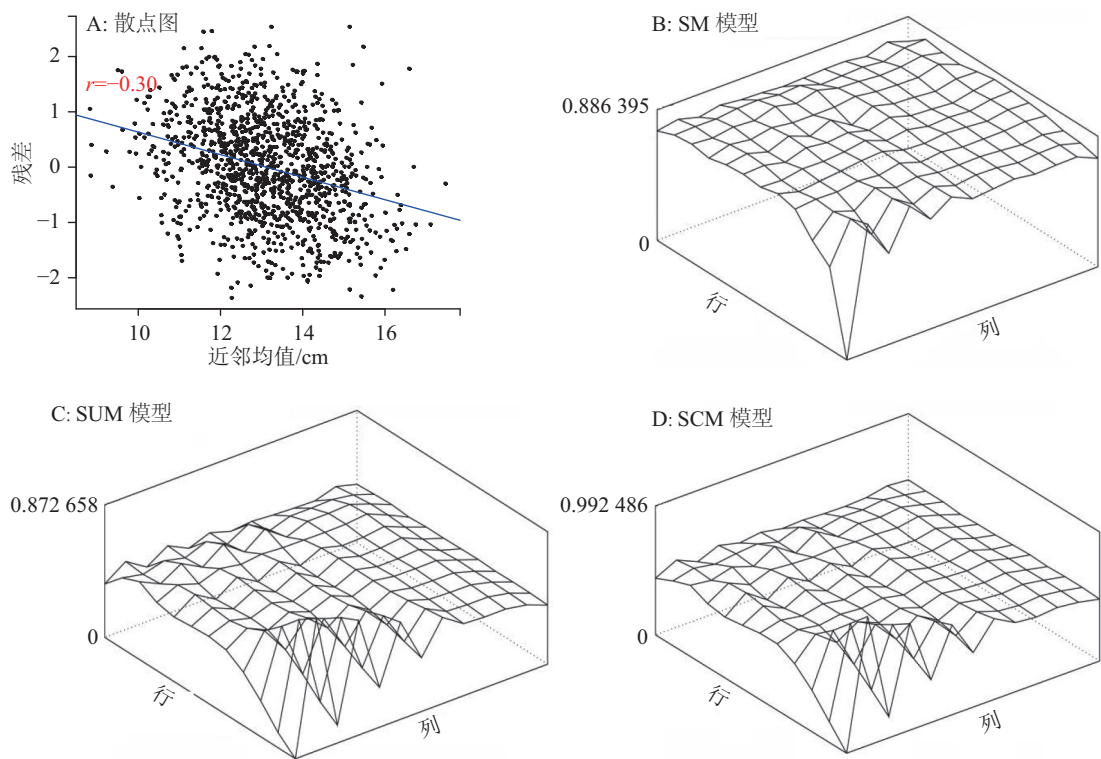


图 2 残差与近邻均值的散点图(A)和不同模型的空间变异图(B、C、D)

Fig. 2 Scatter plot of residuals and neighbor(A) means as well as spatial variogram of different models(B, C, D)

2.2 SCM 模型拟合结果的稳定性分析

为了解不同初始参数值条件下,SCM 模型拟合的效果,根据竞争相关值、行列自相关值设计了 9 组数据,结果(表 2)表明,所有数据拟合的加性方差为 2.25~2.52,随机测量误差方差为 2.90~3.96,空间自相关误差方差为 1.31~2.43,除了列相关值($\rho_c=-0.3$)拟合得一般外,其他均与初始值接近。各近邻效应方差多数在 0.45~0.80,基本上都显著;此外,估

计的单株狭义遗传力差异不大,均在 0.40 左右。可见,SCM 模型对于空间效应和竞争效应的拟合效果稳定性较好。

2.3 真实数据的模型拟合结果比较

真实数据拟合的结果(表 3)显示,常规的 RCBM 模型,估计的区组方差为 0.19,加性方差为 2.46,剩余残差为 2.96;仅有空间行列自相关的模型 SM,加性方差被高估为 7.58;加上区组因子后的

表 2 不同参数初始值的 SCM 模型估计结果¹⁾

Tab. 2 Estimated results of SCM with different initial values for parameters

参数初始值	ρ_r	ρ_c	δ_η^2	δ_ξ^2	δ_a^2	δ_{gen}^2	δ_{gwn}^2	δ_{gsn}^2	δ_{gnn}^2	h_i^2
$C_r-0.3/\rho_r0.3/\rho_c-0.3$	0.36 ^{2.24}	-0.67 ^{5.45}	3.96 ^{5.20}	1.31 ^{2.77}	2.47 ^{2.72}	0.74 ^{2.43}	0.24 ^{1.32}	0.56 ^{2.01}	0.65 ^{2.43}	0.38±0.12
$C_r-0.3/\rho_r0.6/\rho_c-0.9$	0.58 ^{6.00}	-0.91 ^{33.47}	3.28 ^{5.61}	2.39 ^{5.05}	2.45 ^{2.75}	0.65 ^{2.36}	0.33 ^{1.71}	0.49 ^{2.32}	0.76 ^{2.58}	0.43±0.13
$C_r-0.3/\rho_r0.9/\rho_c-0.6$	0.91 ^{25.41}	-0.73 ^{8.56}	3.24 ^{5.37}	1.80 ^{3.95}	2.52 ^{2.77}	0.91 ^{2.60}	0.47 ^{2.03}	0.52 ^{2.35}	0.83 ^{2.83}	0.44±0.13
$C_r-0.6/\rho_r0.3/\rho_c-0.9$	0.40 ^{3.35}	-0.91 ^{31.91}	3.35 ^{6.09}	2.37 ^{5.08}	2.25 ^{2.89}	0.53 ^{2.28}	0.41 ^{1.89}	0.60 ^{2.63}	0.62 ^{2.44}	0.40±0.12
$C_r-0.6/\rho_r0.6/\rho_c-0.6$	0.53 ^{5.24}	-0.70 ^{9.63}	3.13 ^{4.74}	2.20 ^{4.75}	2.36 ^{2.81}	0.59 ^{2.22}	0.34 ^{1.60}	0.68 ^{2.39}	0.74 ^{2.54}	0.43±0.13
$C_r-0.6/\rho_r0.9/\rho_c-0.3$	0.94 ^{40.34}	0.01 ^{0.04}	3.21 ^{4.77}	2.21 ^{3.69}	2.27 ^{3.29}	0.45 ^{2.12}	0.45 ^{1.88}	0.80 ^{2.61}	0.51 ^{2.18}	0.41±0.11
$C_r-0.9/\rho_r0.3/\rho_c-0.6$	0.38 ^{3.04}	-0.68 ^{7.71}	3.96 ^{5.20}	1.31 ^{2.77}	2.47 ^{2.72}	0.74 ^{2.43}	0.24 ^{1.32}	0.56 ^{2.01}	0.65 ^{2.43}	0.44±0.08
$C_r-0.9/\rho_r0.6/\rho_c-0.6$	0.54 ^{5.36}	-0.66 ^{8.38}	2.90 ^{4.56}	2.41 ^{4.51}	2.35 ^{3.24}	0.45 ^{2.10}	0.43 ^{1.87}	0.80 ^{2.58}	0.59 ^{2.33}	0.43±0.13
$C_r-0.9/\rho_r0.9/\rho_c-0.9$	0.91 ^{20.45}	-0.92 ^{24.47}	3.24 ^{6.82}	2.43 ^{3.05}	2.38 ^{3.64}	0.51 ^{2.41}	0.57 ^{2.48}	0.65 ^{2.77}	0.53 ^{2.51}	0.42±0.10

1) C_r 为竞争相关值, ρ_r 、 ρ_c 分别为行、列自相关值, δ_η^2 、 δ_ξ^2 分别为随机测量误差方差和空间自相关误差方差, δ_a^2 为加性方差, δ_{gen}^2 、 δ_{gwn}^2 、 δ_{gsn}^2 、 δ_{gnn}^2 分别为东向、西向、南向和北向近邻遗传竞争方差, h_i^2 为单株狭义遗传力;表中的上标值是t检验统计量,为估计值与标准误的比值。

表 3 真实数据不同模型估计参数的比较¹⁾

Tab. 3 Comparison of estimated parameters from different models for real data

模型	参数/ 个	lnL	AIC	δ^2_{rep}	δ^2_{plot}	δ^2_e	δ^2_{η}	δ^2_{ξ}	ρ_r	ρ_c	δ^2_a	δ^2_{gen}	δ^2_{gwn}	δ^2_{gsn}	δ^2_{gnn}	h^2_i
RCBM	4	-1 605	3 219	0.19 ^{1.54}	0.00	2.96 ^{8.04}					2.46 ^{3.82}					0.45±0.09
SM	4	-1 578	3 165					0.81 ^{6.84}	0.98 ^{39.29}	0.98 ^{42.75}	7.58 ^{24.29}					...
SM1	6	-1 600	3 212	0.19 ^{1.49}	0.00			2.84 ^{6.84}	0.11 ^{8.92}	0.13 ^{2.69}	2.74 ^{3.69}					...
SUM	7	-1 569	3 153	0.01 ^{0.14}	0.00		2.80 ^{8.37}	0.80 ^{1.26}	0.97 ^{38.63}	0.98 ^{39.76}	2.21 ^{3.81}					0.44±0.10
SCM	10	-1 559	3 138				2.41 ^{6.43}	0.72 ^{2.96}	0.91 ^{20.33}	0.93 ^{26.38}	2.25 ^{3.75}	0.13 ^{0.75}	0.09 ^{1.13}	0.01 ^{0.32}	0.05 ^{1.00}	0.48±0.10

1) lnL 为迭代收敛似然值, AIC 为赤池信息量准则, δ_{rep}^2 为区组方差, δ_{plot}^2 为小区方差, δ_e^2 为剩余残差, δ_{η}^2 、 δ_{ξ}^2 分别为测量随机误差方差和空间自相关误差方差, ρ_r 、 ρ_c 分别为行、列自相关值, δ_a^2 为加性方差, δ_{gen}^2 、 δ_{gwn}^2 、 δ_{gsn}^2 、 δ_{gnn}^2 分别为东向、西向、南向和北向近邻遗传竞争方差, h_i^2 为单株狭义遗传力; 表中的上标值是 t 检验统计量, 为估计值与标准误的比值; ... 表示单株遗传力因没有随机测量误差而不能计算。

空间模型 SM1, 这种偏差被修正, 加性方差为 2.74, 空间自相关误差方差为 2.84, 但行、列自相关值仅为 0.11、0.13; “空间+测量” 误差模型 SUM, 有较好的拟合结果, 加性方差为 2.21, 随机测量误差方差为 2.80, 空间自相关误差方差为 0.80; 空间与竞争模型 SCM 为最优模型, 加性方差为 2.25, 随机测量误差方差和空间自相关误差方差均比 SUM 有所降低, 分别为 2.41 和 0.72, 所估算的近邻竞争方差在 0.01~0.13, 但没达到显著水平。从似然值 lnL 和 AIC 值可知, SCM 模型仍是最优。

3 讨论与结论

本研究利用 R 软件及其程序包 breedR 模拟数据, 结合实测数据, 采用 XFA1 结构拟合加性效应、近邻竞争效应和 AR1 结构拟合空间效应, 建立了 RCBM、SM、SUM 和 SCM 4 种模型并进行比较, 发现 SCM 是最优模型, 该模型可大大降低随机误差方差, 并估算到四周近邻竞争方差, 设置参数的不同初始值, SCM 估计的参数结果均较为稳定。

对林木研究而言, 试验林面积一般较大, 采用常规试验设计难以避免环境异质性, 传统的分析模型因忽略环境的异质性而导致遗传参数和育种值估算的偏差, 因此 Dutkowski 等^[8]提出林木遗传试验数据应尽可能采用空间分析。对于环境的空间变异, 一般使用 3 次 B 样条函数(Cubic B spline)或行列自回归拟合。考虑到残差结构的灵活性, 本文采用行列自回归法进行空间分析。从模拟数据的似然值 lnL 和 AIC 值结果可以看出, 空间分析模型比常规分析模型拟合的结果更好, 但是没有随机误差方差的空间模型, 随机误差被纳入空间相关误差, 从而导致空间相关误差方差被高估, 因此, 也造成空

间变异趋势被抑制, 使得行、列自相关值被低估, 这从残差图和空间变异图也可得到佐证。然而, 仅有行、列自相关的空间模型会导致加性方差被严重高估^[3], 本文也得到一致的结果。Magnussen^[2] 指出, 环境异质性一般存在于试验林早期, 林木个体周围相似的微环境会导致个体间存在正相关, 但随着林木生长, 个体间的竞争日趋严重, 将会导致个体间存在负相关。因此, 对于多年生的林木试验林, 个体间存在竞争效应是必然的, 这是造成个体致死率的重要因素^[9]。换言之, 林木性状的表型中, 会同时存在空间效应和竞争效应, 忽略任何一个因素, 都可能会造成遗传评估的偏差, 并进一步影响后续的选择工作。对于遗传竞争效应, 拟合的方法有 GIM^[10]、TIM^[11]和 CSM^[12], 但上述研究中遗传竞争效应均为单一值, 忽略了林木个体周边的遗传竞争效应可能不同。本文提出的以加性效应及其四周近邻竞争效应的方差-协方差矩阵作为 G 结构, 替代常规的加性 G 结构, 从而达到拟合近邻竞争效应值不唯一的目的。考虑到降低运算量和模型收敛, 采用因子分析法来拟合 G 结构。模拟数据的拟合结果表明, 新的 SCM 模型, 拟合方差虽然没有增加, 但随机误差方差大大降低了, 分别由 7.56(RCBM)、5.72(SUM)降低到 3.13(SCM), 分别降低了 58.6%、45.3%, 这与同时考虑竞争效应和空间效应的研究模型结果^[12]相一致。此外, 拟合数据的 SCM 模型单株狭义遗传力保持在 0.40 左右, 高于 RCBM(0.24)和 SUM 模型(0.30)。而针对参数不同初始值的模拟数据, SCM 模型的拟合结果都比较稳定, 表明其可用于林木的遗传分析。最后, 利用松树的实测数据验证 SCM 模型, 估算结果与模拟结果比较一致, 笔者认为, 新单株混合模型即 SCM 模型可用于林木遗传分析, 在提高遗传评估的准确性和可靠性方面具有优势。

参考文献:

- [1] 林元震, 陈晓阳. R 与 ASReml-R 统计分析教程[M]. 北京: 中国林业出版社, 2014.
- [2] MAGNUSSEN S. A method to adjust simultaneously for spatial microsite and competition effects[J]. Can J For Res, 1994, 24(5): 985-995.
- [3] DUTKOWSKI G. Improved models for the prediction of breeding values in trees [D]. Australia Tasmania: University of Tasmania, 2005.
- [4] YANG R C, YE T Z, BLADE S F, et al. Efficiency of spatial analyses of field pea variety trials[J]. Crop Sci, 2004, 44(1): 49-55.
- [5] GILMOUR A R, CULLIS B R, VERBYLA A P. Accounting for natural and extraneous variation in the analysis of field experiments[J]. J Agric Biol Environ Stat, 1997, 2(3): 269-293.
- [6] COSTA E SILVA J, DUTKOWSKI G W, GILMOUR A R. Analysis of early tree height in forest genetic trials is enhanced by including a spatially correlated residual[J]. Can J For Res, 2001, 31(11): 1887-1893.
- [7] DUTKOWSKI G W, COSTA E S J, GILMOUR A R, et al. Spatial analysis methods for forest genetic trials[J]. Can J For Res, 2002, 32(12): 2201-2214.
- [8] DUTKOWSKI G W, COSTA E S J, GILMOUR A R, et al. Spatial analysis enhances modeling of a wide variety of traits in forest genetic trials[J]. Can J For Res, 2006, 36(7): 1851-1870.
- [9] LIN Y Z, YANG H X, IVKOVIĆ M, et al. Effect of genotype by spacing interaction on radiata pine genetic parameters for height and diameter growth[J]. For Ecol Manage, 2013, 304: 204-211.
- [10] CAPP A E P, CANTET R J C. Direct and competition additive effects in tree breeding: Bayesian estimation from an individual tree mixed model[J]. Silvae Genet, 2008, 57(2): 45-56.
- [11] STRINGER J K, CULLIS B R, THOMPSON R. Joint modeling of spatial variability and within-row interplot competition to increase the efficiency of plant improvement[J]. J Agric Biol Environ Stat, 2011, 16(2): 269-281.
- [12] CAPP A E P, MUÑOZ F, SANCHEZ L, et al. A novel individual-tree mixed model to account for competition and environmental heterogeneity: A Bayesian approach [J]. Tree Genet Genomes, 2015, 11: 120. doi: 10.1007/s11295-015-0917-3.
- [13] MUÑOZ F, SANCHEZ L. R package version 0.7-16.2014: BreedR: Statistical methods for forest genetic resources analysts [CP/OL]. [2016-12-12]. <https://github.com/famuvie/breedR>.
- [14] GILMOUR A R, GOGEL B J, CULLIS B R, et al. ASReml user guide release 4.0[M]. UKHemel Hempstead: Vsn International Ltd Hemel, 2016.
- [15] LIN Y Z. R package version 2.6.1.2016: AAfun: ASReml-R Added Functions [CP/OL]. [2016-12-12]. <https://github.com/yzhlinscau/AAfun>.

【责任编辑 李晓卉】