doi:10.11707/j.1001-7488.20140109

WinNC2:因子交配设计遗传分析软件*

童春发[1] 蒋安纳[2] 杨立委[1] 施季森[1]

(1. 南京林业大学 林木遗传与生物技术省部共建教育部重点实验室 南京 210037;
2. 南京林业大学信息科学技术学院 南京 210037)

摘 要: 因子交配设计在林木遗传育种研究中已经得到广泛应用,然而有关该设计的遗传参数和统计量计算还存在不少问题,这在数据不平衡情况下尤为突出。WinNC2 是一个新开发的用于计算因子交配设计遗传模型参数的 Windows 软件,不但可以处理平衡数据,而且还可以处理不平衡数据。针对 3 种不同的因子交配设计遗传统计模型,该软件使用线性约束模型理论计算一般配合力和特殊配合力以及其他固定效应,使用方差分量模型计算遗传方差和协方差进而估算遗传力。对于每一个参数估计,WinNC2 还给出其标准误和显著性检验统计量;对于多性状数据,WinNC2 可以计算父本、母本以及父母本交互水平上 2 个性状的遗传相关系数及其标准误。WinNC2 为林木遗传育种工作者分析因子交配设计试验数据提供了一个界面友好、易于使用的计算工具。

关键词: 因子交配设计;配合力;方差分量;遗传力

中图分类号: S722.3; Q348 **文献标识码**: A **文章编号**: 1001-7488(2014)01-0055-08

WinNC2: A New Software for Genetic Analysis of Factorial Mating Design

Tong Chunfa[1] Jiang Anna[2] Yang Liwei[1] Shi Jisen[1]

(1. *Key Laboratory of Forest Genetics & Biotechnology of Ministry of Education Nanjing Forestry University Nanjing* 210037;
2. *College of Information Science & Technology, Nanjing Forestry University Nanjing* 210037)

Abstract: Although a factorial mating design has been extensively applied in forest breeding programs, there exist some problems with the design in calculating the estimates of genetic parameters and some statistics, especially when the experimental data are unbalanced. A newly developed Windows-based software (WinNC2), for calculating model parameters of the genetic factorial mating design, can not only deal with balanced data but also with unbalanced data. For three different factorial mating designs, the software estimates the general and special combining abilities of parents and other fixed effects with the restricted linear model, and calculates the genetic variance components and further the heritabilities with the random effect model. In the mean time, WinNC2 can provide standard errors and statistic tests for each parameter estimate. Furthermore, WinNC2 can also present the genetic correlation coefficient between any two traits at each level of paternal, maternal, and their interaction effects. In summary, WinNC2 is a user-friendly computing tool for forest breeders to analyze genetic models of factorial mating designs.

Key words: factorial mating design; combining ability; variance components; heritability

因子交配设计(factorial mating design)是 Comstock 和 Robinson(1948; 1952)提出的 3 种北卡罗林纳(North Carolina)遗传交配设计之一,又称为 NCⅡ交配设计、AB 设计、析因设计或测交设计(沈熙环, 1990)。这种交配设计的特点是选择 1 组父本和 1 组母本,每个父本与每个母本交配,可产生数量众多的杂交组合群体。由于该试验设计不但能提供亲本一般配合力的估计,而且还能给出特殊配合力的估计,同时也能给出有关遗传方差分量以及遗传力的估计,因此它在动植物遗传育种研究中有着极其广泛的应用。林木上,国内外先后在杉木(*Cunninghamia lanceolata*)、松树(*Pinus*)、桉树(*Eucalyptus*)等树种的遗传变异研究中应用了因子交配试验设计(Barnes *et al.*, 1978; Blada, 1992; 郑

收稿日期: 2013-06-06; 修回日期: 2013-07-20。

基金项目: 国家自然科学基金面上项目(31070601); 江苏省高校优势学科建设工程资助项目(PAPD)。

* 施季森为通讯作者。

仁华，2003；Volker *et al.*，2008；蒋焱等，2008；金国庆等，2008；陆钊华，2009；杨萍等，2011）。

应用因子交配设计估算遗传参数需要使用统计模型和相应的计算软件。统计模型主要是线性模型，分为固定效应模型和随机效应模型。利用线性固定效应模型可以估计出一般配合力和特殊配合力并进行相关因素的显著性检验（Wu *et al.*，2000；Tong *et al.*，2012；童春发等，2013），而估计遗传方差分量进而估算遗传力等重要遗传参数则要使用随机效应模型（童春发等，2010；2013）。也就是说，对于每一个遗传交配设计，在分析遗传参数时需要用到 2 种有着本质区别的线性统计分析模型。然而，经典的林木数量遗传学著作（Wright，1976；Namkoong，1979；Williams *et al.*，2002；续九如，2006）在这方面没有很好的论述，使得林木育种工作者在分析林木遗传设计数据时感到诸多不便和困惑，这在不平衡数据条件下或模型含有较多因素情况下尤其突出。在软件应用方面，大多数研究者使用著名的 SAS 分析系统来分析林木遗传交配设计模型；但是 SAS 是通用的统计分析软件，不是针对特定的遗传分析模型，因此在计算林木遗传参数中存在一些不便和不足之处（童春发等，2013）。为了方便分析林木遗传参数并提供更多的统计量，近年来编写了一些 Windows 软件用于林木半同胞子代、巢式设计和双列杂交设计遗传模型分析（童春发等，2010；2013；Tong *et al.*，2012），这些软件的主要特点在于易于使用、能处理不平衡数据以及能计算出比较全面的遗传参数。此外，国内开发的 ForState（唐守正等，2009）和 DPS（唐启义，2010）比较大型的数据分析系统虽然涉及到遗传交配设计模型分析，但是由于其理论方法来源于经典的数量遗传学文献资料，因此它们分析遗传模型的功能存在一定的局限性，如不能提供方差分量估计的标准误和假设检验统计量以及不平衡数据条件下遗传力估计及其标准误等。尽管在某些遗传模型计算软件开发方面已经取得一定的进展，但是对于因子交配设计统计分析目前还没有易于使用并能给出比较全面的遗传参数的计算软件。

本文根据现代线性模型理论，针对 3 种因子交配设计遗传模型开发了用于计算多种遗传参数和统计量的 Windows 应用软件 WinNC2。3 种因子交配设计分别为多地点多区组设计、单地点多区组设计和单地点不含区组设计。对于每一种因子交配设计，该软件使用约束线性模型理论计算一般配合力和特殊配合力以及其他固定效应，使用随机效应模型计算遗传方差分量进而计算遗传力，无论对平衡数据还是不平衡数据均可以进行计算分析。对于每一个参数估计，WinNC2 还可以给出它们的标准误和显著性检验统计量。另外，WinNC2 还可以计算父本、母本以及父母本交互水平上 2 个性状的遗传相关系数及其标准误。

1 统计模型

由于试验目的和条件的不同，因子交配设计统计分析模型也有多种形式。WinNC2 主要关注 3 种因子交配设计统计分析模型，即多地点多区组设计、单地点多区组设计和单地点不含区组设计统计分析模型。其他因子交配设计可用同样的线性统计模型理论方法来对遗传参数进行估计和计算。

1.1 多地点多区组设计统计模型

设有 m 个父本、f 个母本两两交配产生 mf 个全同胞家系，然后布置在 s 个地点进行子代试验，每个地点设置 b 个区组，每一区组若干个单株，每一个单株的数量性状值可用线性模型表示为：

$$y_{ijklt} = \mu + S_i + B_{ij} + M_k + F_l + \mathrm{MF}_{kl} + \mathrm{MS}_{ik} + \mathrm{FS}_{il} + \mathrm{MFS}_{ikl} + \mathrm{MFBS}_{ijkl} + e_{ijklt}。 \quad (1)$$

式中：y_{ijklt} 为第 i 个地点第 j 个区组中第 k 个父本与第 l 个母本交配子代的第 t 个个体数量性状值；μ 为总平均值；S_i 为第 i 个地点的效应（$i = 1, \cdots, s$）；B_{ij} 为第 i 个地点内第 j 个区组的效应（$j = 1, \cdots, b$）；M_k 为第 k 个父本的效应（$k = 1, \cdots, m$）；F_l 为第 l 个母本的效应（$l = 1, \cdots, f$）；MF_{kl} 为第 k 个父本与第 l 个母本的交互效应；MS_{ik} 为第 i 个地点与第 k 个父本的交互效应；FS_{il} 为第 i 个地点与第 l 个母本的交互效应；MFS_{ikl} 为第 i 个地点、第 k 个父本和第 l 个母本的交互效应；MFBS_{ijkl} 为第 i 个地点内第 j 个区组与第 k 个父本和第 l 母本组合的交互效应；e_{ijklt} 为随机误差效应（$t = 1, \cdots, n_{ijkl}$），一般假定它服从均值为 0、方差为 σ_e^2 的正态分布。该模型在 Wright（1976）和续九如（1989；2006）中均有描述。

对于模型（1），为了计算亲本的一般配合力和特殊配合力，将除随机效应 e_{ijklt} 外所有的效应都看作是固定效应，此时模型（1）被称为固定效应模型。根据约束线性模型理论可以估计配合力 M_k，F_l，MF_{kl} 和其他固定效应，还可以对这些效应进行统计检验（Tong *et al.*，2012；童春发等，2013）。这里固定效应应满足如下约束条件：

$$\begin{cases}\sum_i S_i = 0;\\ \sum_j B_{ij} = 0(i = 1,\cdots,s);\\ \sum_k M_k = 0;\\ \sum_l F_l = 0;\\ \sum_k \mathrm{MF}_{kl} = 0(l = 1,\cdots,f);\\ \sum_l \mathrm{MF}_{kl} = 0(k = 1,\cdots,m);\\ \sum_k \mathrm{MS}_{ik} = 0(i = 1,\cdots,s);\\ \sum_i \mathrm{MS}_{ik} = 0(k = 1,\cdots,m);\\ \sum_l \mathrm{FS}_{il} = 0(i = 1,\cdots,s);\\ \sum_i \mathrm{FS}_{il} = 0(l = 1,\cdots,f);\\ \sum_i \mathrm{MFS}_{ikl} = 0(k = 1,\cdots,m;\ l = 1,\cdots,f);\\ \sum_k \mathrm{MFS}_{ikl} = 0(i = 1,\cdots,s;\ l = 1,\cdots,f);\\ \sum_l \mathrm{MFS}_{ikl} = 0(i = 1,\cdots,s;\ k = 1,\cdots,m);\\ \sum_{k,l} \mathrm{MFBS}_{ijkl} = 0(i = 1,\cdots,s;\ j = 1,\cdots,b);\\ \sum_j \mathrm{MFBS}_{ijkl} = 0(i = 1,\cdots,s;\ k = 1,\cdots,m;\\ \qquad l = 1,\cdots,f)。\end{cases} \tag{2}$$

如果要计算遗传力等遗传参数，那么首先要计算遗传方差分量，此时要将模型(1)中的效应 M_k，F_l，MF_{kl}，MS_{ik}，FS_{il}，MFS_{ikl} 和 MFBS_{ijkl} 都看作是随机效应，并假定这些随机效应的均值为 0，方差分别为 σ^2_{M}，σ^2_{F}，σ^2_{MF}，σ^2_{MS}，σ^2_{FS}，σ^2_{MFS} 和 σ^2_{MFBS}。根据方差分析法(Henderson，1953)可以估计这些方差分量，进一步还可以计算出其标准误以及

$$h^2_{\mathrm{f}} = \frac{\sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}}}{\frac{1}{n_{\mathrm{M\times F}}}\sum_{k,l}\mathrm{var}(\bar{y}_{\cdot\cdot kl\cdot})} = \frac{\sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}}}{\sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}} + k_1\sigma^2_{\mathrm{MS}} + k_1\sigma^2_{\mathrm{FS}} + k_1\sigma^2_{\mathrm{MFS}} + k_2\sigma^2_{\mathrm{MFBS}} + k_3\sigma^2_{\mathrm{e}}}。 \tag{3}$$

式中：$n_{\mathrm{M\times F}}$表示家系的个数；$k_1 = \frac{1}{n_{\mathrm{M\times F}}}\sum_{k,l}\frac{1}{n^2_{\cdot\cdot kl}}\sum_{i=1}^{s} n^2_{i\cdot kl}$；

$k_2 = \frac{1}{n_{\mathrm{M\times F}}}\sum_{k,l}\frac{1}{n^2_{\cdot\cdot kl}}\sum_{i=1}^{s}\sum_{j=1}^{b_i} n^2_{ijkl}$；$k_3 = \frac{1}{n_{\mathrm{M\times F}}}\sum_{k,l}\frac{1}{n_{\cdot\cdot kl}}$。

1.2　单地点多区组设计统计模型

设有 m 个父本、f 个母本两两交配产生 mf 个全同胞家系，然后在某个地点设置 b 个区组进行子代试验，每一区组若干个单株，则单株的数量性状值可用线性模型表示为：

$$y_{ijkl} = \mu + B_i + M_j + F_k + \mathrm{MF}_{jk} + e_{ijkl}。 \tag{4}$$

式中：y_{ijkl} 为第 i 个区组中第 j 个父本与第 k 个母本交配子代的第 l 个个体数量性状值；μ 为总平均值；B_i 为第 i 个区组的效应($i = 1,\cdots,b$)；M_j 为第 j 个父本的效应($j = 1,\cdots,m$)；F_k 为第 k 个母本的效应($k = 1,\cdots,f$)；MF_{jk} 为第 j 个父本与第 k 个母本的交互效应；e_{ijkl} 为随机误差效应($l = 1,\cdots,n_{ijk}$)，并假定它服从均值为 0、方差为 σ^2_{e} 的正态分布。该模型可见于贝克尔(1987)。

显著性检验统计量，具体方法步骤可参考童春发等(2010；2013)。

遗传力是林木遗传分析中的一个重要指标，根据因子交配设计可以计算单株遗传力和家系遗传力。对于单株遗传力，无论数据是平衡还是不平衡，其计算表达式都是一样的。然而，对于家系遗传力，文献中只给出了平衡数据条件下的结果(Wright，1976；续九如，2006)。这里给出不平衡设计条件下家系遗传力的计算表达式。

记 n_{ijkl} 为第 i 个地点第 j 个区组中第 k 个父本与第 l 母本交配子代的个数，那么家系 (k,l) 性状的均值为：

$$\begin{aligned}\bar{y}_{\cdot\cdot kl\cdot} = {} & C + M_k + F_l + \mathrm{MF}_{kl} + \\ & \frac{1}{n_{\cdot\cdot kl}}\sum_{i=1}^{s} n_{i\cdot kl}\,\mathrm{MS}_{ik} + \frac{1}{n_{\cdot\cdot kl}}\sum_{i=1}^{s} n_{i\cdot kl}\,\mathrm{FS}_{il} + \\ & \frac{1}{n_{\cdot\cdot kl}}\sum_{i=1}^{s} n_{i\cdot kl}\,\mathrm{MFS}_{ikl} + \\ & \frac{1}{n_{\cdot\cdot kl}}\sum_{i=1}^{s}\sum_{j=1}^{b_i} n_{ijkl}\,\mathrm{MFBS}_{ijkl} + \frac{1}{n_{\cdot\cdot kl}}\sum_{i=1}^{s}\sum_{j=1}^{b}\sum_{m=1}^{n_{ijkl}} e_{ijklt}。\end{aligned}$$

式中：C 为固定效应代数和。因此家系均值的方差为：

$$\begin{aligned}\mathrm{var}(\bar{y}_{\cdot\cdot kl\cdot}) = {} & \sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}} + \\ & \frac{1}{n^2_{\cdot\cdot kl}}\sum_{i=1}^{s} n^2_{i\cdot kl}\sigma^2_{\mathrm{MS}} + \frac{1}{n^2_{\cdot\cdot kl}}\sum_{i=1}^{s} n^2_{i\cdot kl}\sigma^2_{\mathrm{FS}} + \\ & \frac{1}{n^2_{\cdot\cdot kl}}\sum_{i=1}^{s} n^2_{i\cdot kl}\sigma^2_{\mathrm{MFS}} + \\ & \frac{1}{n^2_{\cdot\cdot kl}}\sum_{i=1}^{s}\sum_{j=1}^{b_i} n^2_{ijkl}\sigma^2_{\mathrm{MFBS}} + \frac{1}{n_{\cdot\cdot kl}}\sigma^2_{\mathrm{e}}。\end{aligned}$$

由于数据不平衡性会导致各家系均值的方差不一致，因此取家系均值方差的平均值作为计算家系遗传力的背景，于是不平衡数据条件下家系遗传力可表示为：

同样地，根据模型(4)来估计亲本的一般配合力和特殊配合力，首先需要将效应 B_i，M_j，F_k 和 MF_{jk} 看作固定效应，并施加如下的约束条件：

$$\begin{cases} \sum_i B_i = 0; \\ \sum_j M_j = 0; \\ \sum_k F_k = 0; \\ \sum_j \mathrm{MF}_{jk} = 0(k = 1, \cdots, f); \\ \sum_k \mathrm{MF}_{jk} = 0(j = 1, \cdots, m)。 \end{cases} \tag{5}$$

然后根据线性约束模型理论来估计这些固定效应。

若要估计方差分量，则在模型(4)中将 M_j，F_k 和 MF_{jk} 看作随机效应，并假定它们的均值为0，方差分别为 σ^2_{M}，σ^2_{F} 和 σ^2_{MF}。利用方差分析法可以对这些方差分量进行估计，并可以计算出估计的标准误以及检验这些参数显著性的统计量，进一步还可以计算出遗传力。

记 n_{ijk} 为第 i 个区组中第 j 个父本与第 k 个母本交配子代的个数，那么家系 (j,k) 性状的均值为：

$$\bar{y}_{\cdot jk\cdot} = C + M_j + F_k + \mathrm{MF}_{jk} + \frac{1}{n_{\cdot jk}} \sum_{i=1}^{b} \sum_{l=1}^{n_{ijk}} e_{ijkl},$$

式中：C 为固定效应代数和。家系均值的方差为：

$$\mathrm{var}(\bar{y}_{\cdot jk\cdot}) = \sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}} + \frac{1}{n_{\cdot jk}} \sigma^2_{\mathrm{e}}。$$

因此，根据模型(4)计算不平衡数据条件下家系遗传力的表达式为：

$$h^2_{\mathrm{f}} = \frac{\sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}}}{\sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}} + k_4 \sigma^2_{\mathrm{e}}}。 \tag{6}$$

式中：$k_4 = \dfrac{1}{n_{\mathrm{M\times F}}} \sum_{j,k} \dfrac{1}{n_{\cdot jk}}$。

1.3 单地点不含区组设计统计模型

设有 m 个父本、f 个母本两两交配产生 mf 个全同胞家系，每个家系种植若干个单株，则单株的数量性状值可用线性模型表示为：

$$y_{ijk} = \mu + M_i + F_j + \mathrm{MF}_{ij} + e_{ijk}。 \tag{7}$$

式中：y_{ijk} 为第 i 个父本与第 j 个母本交配子代的第 k 个个体数量性状值；μ 为总平均值；M_i 为第 i 个父本的效应($i = 1, \cdots, m$)；F_j 为第 j 个母本的效应($j = 1, \cdots, f$)；MF_{ij} 为第 i 个父本与第 j 个母本的交互效应；e_{ijk} 为随机误差效应($k = 1, \cdots, n_{ij}$)，并假定它服从均值为0、方差为 σ^2_{e} 的正态分布。该模型比较简单，在 Namkoong (1979) 和贝克尔(1987)中均有提到。

在模型(7)中，若将 M_j，F_k 和 MF_{jk} 看作固定效应，并施加如下的约束条件：

$$\begin{cases} \sum_i M_i = 0; \\ \sum_j F_j = 0; \\ \sum_i \mathrm{MF}_{ij} = 0(j = 1, \cdots, f); \\ \sum_j \mathrm{MF}_{ij} = 0(i = 1, \cdots, m)。 \end{cases} \tag{8}$$

则按约束线性模型参数估计方法可以计算出一般配合力和特殊配合力。若将 M_j，F_k 和 MF_{jk} 看作随机效应，并假定它们的均值为0，方差分别为 σ^2_{M}，σ^2_{F} 和 σ^2_{MF}，则使用方差分析法可对这3个方差分量进行估计。

记 n_{ij} 为第 i 个父本与第 j 个母本交配子代的个数，那么家系 (i,j) 性状的均值为：

$$\bar{y}_{ij\cdot} = \mu + M_i + F_j + \mathrm{MF}_{ij} + \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} e_{ijk};$$

家系均值的方差为：

$$\mathrm{var}(\bar{y}_{ij\cdot}) = \sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}} + \frac{1}{n_{ij}} \sigma^2_{\mathrm{e}}。$$

因此，根据模型(7)计算不平衡数据条件下家系遗传力的表达式为：

$$h^2_{\mathrm{f}} = \frac{\sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}}}{\sigma^2_{\mathrm{M}} + \sigma^2_{\mathrm{F}} + \sigma^2_{\mathrm{MF}} + k_5 \sigma^2_{\mathrm{e}}}, \tag{9}$$

式中：$k_5 = \dfrac{1}{n_{\mathrm{M\times F}}} \sum_{j,k} \dfrac{1}{n_{ij}}$。

2 结果

针对因子交配设计遗传统计分析模型(1)，(4)和(7)，用 VC + + 2008 编写了 Windows 计算软件 WinNC2。该软件分别根据固定效应和随机效应模型理论给出亲本配合力估计和遗传方差分量估计，进一步还给出遗传力和遗传相关系数的估计，无论对平衡数据还是不平衡数据都能进行处理计算。WinNC2 界面友好，使用方便，可以自由下载使用，其网址为：http：// fgbio. njfu. edu. cn/tong / WinNC2/WinNC2. htm。

图1显示了 WinNC2 分析3种因子交配设计的 Excel 输入数据格式。图1a 中第1列至第4列分别为地点、区组、父本和母本的代码，第5列起为性状1，2，3的值等，这里只列出了2个性状值。显然图1b 和图1c 是从图1a 简化而来的，其每列意义如表头所示。使用者可点击 WinNC2 的 File 菜单打开某个输入数据，然后选择合适的因子交配设计分析模型计

算,计算结果会在窗口中显示,并在输入数据所在的目录里找到计算结果文件 result. txt。

图 2 给出了 WinNC2 对 1 个性状使用固定效应模型计算模型(1)的部分结果。对于每个性状,WinNC2 首先对模型本身和各个因素进行检验,然后给出亲本的一般配合力和特殊配合力以及其他固定效应的估计、标准误、显著性检验 t 统计量和 p 值,接着给出父本配合力两两之间、母本配合力两两之间和父母本特殊配合力两两之间是否有显著差异的 t 检验统计量。从图 2 的前部分可以看出,对于性状 1 模型本身和各个因素经检验都是极显著的。从图 2 的参数估计部分可以看出,母本 3 和 4 的一般配合力以及父本 5 和母本 4 的特殊配合力都是极显著的,而父本 1 和母本 1 的特殊配合力则不显著。图 2 的后部分表明,父本 3,4 和 5 一般配合力两两之间差异不显著,而其他情况则是显著的,母本一般配合力两两之间差异都是显著的,特殊配合力两两之间差异有些显著而其他情况则不显著。

Site	Block	Male	Female	Trait1	Trait2
1	1	1	1	6.3645	21.687
1	1	1	1	8.1177	21.2014
1	1	1	1	7.9913	21.0049
1	1	1	1	7.16	20.6614
1	1	1	1	9.1057	19.4439
1	1	1	1	9.4658	21.4799
1	1	1	2	6.8373	19.7776
1	1	1	2	7.0785	22.042
1	1	1	2	7.1525	22.4098
1	1	1	2	7.1679	21.6285
1	1	1	2	7.3534	22.1432
1	1	1	2	6.8711	22.8269
1	1	1	2	9.4161	20.3676
1	1	1	2	5.54	20.9573
1	1	1	2	7.7597	21.4602
1	1	1	2	6.6312	21.2424
1	1	1	3	5.9122	17.368
1	1	1	3	5.978	18.4213
1	1	1	3	5.5674	21.6301
1	1	1	3	5.701	18.0363
1	1	1	3	3.6209	20.3385

a

Block	Male	Female	Trait1	Trait2
1	1	1	9.0764	19.0997
1	1	1	9.31	19.0145
1	1	1	6.8235	18.7858
1	1	1	8.7053	17.726
1	1	2	7.7272	17.9362
1	1	2	7.2259	16.8178
1	1	2	7.53	17.6716
1	1	2	7.9392	17.4364
1	1	3	9.4825	19.0707
1	1	3	8.8653	19.5513
1	1	3	7.1515	20.8569
1	1	3	9.4943	19.1388
1	1	4	7.8581	19.3564
1	1	4	6.9542	17.3791
1	1	4	5.7817	18.5775
1	1	4	7.8524	17.2762
1	1	5	7.8879	17.9703
1	1	5	8.4406	19.8737
1	1	5	7.3686	18.8877
1	1	5	9.9252	19.1194
1	1	6	7.7087	19.4982

b

Male	Female	Trait1	Trait2
1	1	10.5089	21.6292
1	1	11.2584	20.4409
1	1	10.9494	20.4145
1	1	9.746	19.8537
1	2	10.4071	20.6039
1	2	11.981	20.4355
1	2	10.7801	19.116
1	2	9.6909	20.4343
1	3	10.0842	21.982
1	3	9.6085	20.4305
1	3	8.6097	20.0356
1	3	8.5946	21.5001
1	4	7.1339	19.1993
1	4	10.2891	22.1373
1	4	10.2568	20.3355
1	4	9.7053	19.3981
1	5	9.1174	21.0095
1	5	9.8992	20.4402
1	5	9.4585	19.7231
1	5	7.5197	19.2943
1	6	9.2378	19.5511

c

图 1　WinNC2 用于分析(a) 模型(1)、(b) 模型(4) 和 (c) 模型(7)的 Excel 输入数据格式

Fig. 1　The Excel format of input data for (a) model (1), (b) model (4) and (c) model (7) in WinNC2

图 3 是 WinNC2 对 2 个性状使用随机效应模型计算模型(1)的结果。对于每个性状,WinNC2 均给出每个方差分量的估计、标准误、F 检验统计量和 p 值,同时还计算出单株遗传力和家系遗传力以及它们的标准误。对于 2 个以上的性状,软件还进一步计算出两两性状间在父本效应、母本效应和父母本交互效应水平上的遗传相关系数。

3　讨论

因子交配设计在林木遗传育种研究中是一个重要的交配试验设计,它能提供父母本及其大量子代的遗传信息,因此在遗传分析上能够获得比较全面的遗传参数,如双亲的一般配合力和特殊配合力、各个因子上的遗传方差分量等。然而,在实际分析因子交配设计试验数据时,不少研究者会遇到许多困难:一是没有意识到在用因子交配设计进行遗传分析时会用到 2 类线性模型,即固定效应模型和随机效应模型,使得即使有通用的软件(如 SAS 等)也不能得到很好的应用;二是实际获取的数据往往会在不同的水平上是不平衡的,使得经典的方差分析法在稍微复杂的模型上难以实现。本文开发的 Windows 应用软件 WinNC2 能让使用者很方便地分析计算 3 种类型的因子交配设计遗传参数,包括一般配合力和特殊配合力、遗传方差分量、单株遗传力和家系遗传力以及性状间的遗传相关系数,同时还提供遗传参数估计的标准误和假设检验的统计量以对遗传参数估计作进一步评估。

同常用的统计软件相比,WinNC2 简单易用,使用者无须按照遗传模型编写代码,只要打开数据文件然后选择适当的因子交配设计模型就能计算出相应的遗传参数。SAS 是著名的统计分析系统,但它不是针对特定遗传分析模型。如果用 SAS 分析因子交配设计模型的配合力,就需要使用者熟悉 SAS 编程并将固定效应模型的约束条件考虑进去(Wu *et al.*, 2000; Tong *et al.*, 2012; 童春发等, 2013),这让很多不熟悉统计模型的研究者无所适从。SAS 的

Analysis of the Genetic Model for NC Design II:
yijklm = u + Si + SBij + Mk + Fl + MFkl + SMik + SFil + SMFikl + SBMFijkl + eijklm

* Result for the fixed model. Wed May 22 14:11:52 2013 *

Testing the model for trait 1:

Source	DF	Sum of Squares	Mean Squares	F Value	Pr>F
Model	359	15181.5105	42.2883	38.2113	0.0000
Error	3209	3551.3891	1.1067		
Total	3568	18732.8996			

Testing factors for trait 1:

Source	DF	Sum of Squares	Mean Squares	F Value	Pr>F
Site	2	11766.3099	5883.1550	5315.9605	0.0000
Block/Site	15	1425.1880	95.0125	85.8524	0.0000
Male	4	293.2679	73.3170	66.2485	0.0000
Female	3	602.2161	200.7387	181.3855	0.0000
Male x Female	12	165.0350	13.7529	12.4270	0.0000
.....					

Parameter estimates for trait 1:

Parameter	Estimate	Std Error	T for H0: Parameter=0	Pr>|T|
u	9.8618	0.0177	558.6541	0.0000
S1	-2.4103	0.0250	-96.5342	0.0000
.....				
F3	-0.6198	0.0306	-20.2692	0.0000
F4	0.2113	0.0306	6.9063	0.0000
MF(1,1)	-0.0014	0.0615	-0.0229	0.9817
.....				
MF(5,4)	0.1839	0.0618	2.9769	0.0029
.....				
SBMF(3,6,5,4)	-0.4319	0.2961	-1.4586	0.1448

T values of test for difference between two males for trait 1 (1):

	M1	M2	M3	M4
M2	-2.2860*			
M3	8.8261**	11.0790**		
M4	9.4517**	11.7090**	0.5989	
M5	9.5436**	11.7949**	0.7134	0.1165

*p<0.05, **p<0.01

T values of test for difference between two females for trait 1 (1):

	F1	F2	F3
F2	-11.6005**		
F3	-22.2678**	-10.7261**	
F4	-5.6472**	5.9473**	16.6376**

*p<0.05, **p<0.01

T values of test for difference between the interactive effects of males and females for trait 1:

	MF(1,1)	MF(1,2)	MF(1,3)	...	MF(5,2)	MF(5,3)
MF(1,2)	2.1050*					
MF(1,3)	-3.0100**	-5.1536**				
MF(1,4)	0.9612	-1.1491	3.9962**			
.....						
MF(5,3)	0.6569	-1.8887	3.6968**	...	3.2842**	
MF(5,4)	2.2226*	-0.3037	5.8571**	...	4.5542**	1.3032

*p<0.05, **p<0.01

图 2 WinNC2 分析固定效应模型计算结果

Fig. 2 Output of analyzing fixed effects model with WinNC2

VARCOMP 等过程可以用于计算因子交配设计模型的方差分量，但是不能直接计算出遗传力和遗传相关系数，并且如果选择使用方差分析（ANOVA）法，那么 SAS 还不能给出方差分量估计的标准误。最近，ASReml 软件（Gilmour *et al.*, 2009）在育种领域里应用较多，但是此软件也是一个通用软件，需要使用者熟悉特定遗传模型并编写代码以实现相应的计算。

遗传力在林木遗传育种学中是一个重要的遗传参数，利用因子交配设计可以计算单株遗传力和家系遗传力（Wright, 1976; 续九如, 1989; 2006），甚至还可以计算父本家系遗传力和母本家系遗传力

```
*****************************************************************************
*                                                                           *
*  Result for the random effect model.                  Wed May 22 14:11:52 2013 *
*                                                                           *
*****************************************************************************

Estimates of variance components and heritabilites for trait 1:
=============================================================================
   Parameter            Estimate           Std Error        F For H0:        Pr > F
                                                          Parameter=0
-----------------------------------------------------------------------------
   Sigma_M               0.0728             0.0742          2.4619          0.1295
   Sigma_F               0.1716             0.1856          2.1084          0.2005
   Sigma_MF              0.0671             0.0327          4.5907          0.0007
   Sigma_MS              0.0333             0.0215          3.5039          0.0081
   Sigma_FS              0.1135             0.0699         13.4434          0.0000
   Sigma_MFS             0.0099             0.0111          1.3284          0.1434
   Sigma_MFBS            0.0533             0.0141          1.4772          0.0000
   Sigma_E               1.1067             0.0276

   Hi                    0.3002             0.2116
   Hf                    0.8352             0.1096
=============================================================================

Estimates of variance components and heritabilites for trait 2:
=============================================================================
   Parameter            Estimate           Std Error        F For H0:        Pr > F
                                                          Parameter=0
-----------------------------------------------------------------------------
   Sigma_M               0.1492             0.1314          5.6078          0.0189
   Sigma_F               0.6494             0.5536          8.7136          0.0132
   Sigma_MF              0.0339             0.0277          1.2647          0.3000
   Sigma_MS              0.0578             0.0409          2.8413          0.0226
   Sigma_FS              0.0465             0.0378          2.9335          0.0272
   Sigma_MFS             0.0589             0.0265          2.5614          0.0001
   Sigma_MFBS            0.0639             0.0166          1.4917          0.0000
   Sigma_E               1.2877             0.0321

   Hi                    0.6804             0.3212
   Hf                    0.9273             0.0507
=============================================================================

Genetic correlation matrix of male effects with std. in brackets
==========================================
    Trait            1             2
------------------------------------------
       1          1.0000        0.3412
                 (0.0000)      (0.5749)

       2          0.3412        1.0000
                 (0.5749)      (0.0000)
==========================================

Genetic correlation matrix of female effects with std. in brackets
==========================================
    Trait            1             2
------------------------------------------
       1          1.0000        0.8718
                 (0.0000)      (0.2071)

       2          0.8718        1.0000
                 (0.2071)      (0.0000)
==========================================

Genetic correlation matrix of male x female effects with std. in brackets
==========================================
    Trait            1             2
------------------------------------------
       1          1.0000        0.9946
                 (0.0000)      (0.2200)

       2          0.9946        1.0000
                 (0.2200)      (0.0000)
==========================================
```

图 3　WinNC2 分析随机效应模型计算结果

Fig. 3　Output of analyzing random effects model with WinNC2

（Holland *et al.*，2003；续九如，1989；2006）。文献中有关家系遗传力的计算公式一般都是针对平衡试验设计的，但是没有说明是如何推导出来的。本文针对 3 种因子交配设计推导出了不平衡数据条件下家系遗传力的计算公式，这些公式可以简化成平衡数据条件下遗传力的计算公式。例如，在平衡设计条件下，公式（3）可以简化成 Wright（1976）和续九如（1989；2006）中全同胞家系遗传力计算公式。但是令人遗憾的是，在 Wright（1976，p247）中这一公式分母中项 σ^2_{MF} 多了因子 1/FM，而在续九如（2006）公式（4－28）中分子分母项 σ^2_M 和 σ^2_F 的系数都存在问题。只有续九如（1989）公式（23）同本文

公式(3)的简化结果一致。至于不平衡数据条件下因子交配设计的父本家系遗传力和母本家系遗传力计算似乎有点复杂,留待以后作进一步的研究。

方差分量的假设检验问题比较复杂,其统计方法仍然在不断发展研究过程中。对于平衡设计试验数据,一般用 Wald 法对方差分量是否为零进行显著性检验(Wald, 1941; 王松桂等, 2007)。对于不平衡设计试验数据,本文采用 Ofversten(1993)提出的重抽样法(resampling)对方差分量进行显著性检验。由于模型(1)涉及到地点等较多的因素,因此在方差分量假设检验过程中本文多次使用了 Ofversten 重抽样法。当然,Ofversten 方差分量假设检验方法在因子交配设计模型中应用的统计性质需要作进一步评估,可以考虑将置换法(permutation)(Fitzmaurice *et al.*, 2007)应用到方差分量假设检验中去以便同 Ofversten(1993)方法作比较。

参 考 文 献

贝克尔 W A. 1987. 数量遗传学手册. 区靖祥,译. 北京: 科学出版社.

蒋 焱,龚建英,雷布先,等. 2008. 杉木测交系设计子代遗传分析. 江西林业科技,(6):16 – 18.

金国庆,秦国峰,刘伟宏,等. 2008. 马尾松测交系杂交子代生长性状遗传分析. 林业科学, 44(1):70 – 76.

陆钊华. 2009. 尾叶桉种内种间交配遗传分析及 F1 选择研究. 北京: 中国林业科学研究院博士学位论文.

沈熙环. 1990. 林木育种学. 北京: 中国林业出版社.

唐启义. 2010. DPS 数据处理系统. 北京: 科学出版社.

唐守正,郎奎建,李海奎. 2009. 统计和生物数学模型计算(ForStat 教程). 北京: 科学出版社.

童春发,卫 巍,尹 辉,等. 2010. 林木半同胞子代测定遗传模型分析. 林业科学,46(1): 29 – 35.

童春发,杨立伟,蒋安纳,等. 2013. 不平衡巢式设计遗传模型分析. 林业科学,49(1): 1 – 8.

王松桂,史建红,尹素菊,等. 2007. 线性模型引论. 北京: 科学出版社.

续九如. 1989. 树木遗传力估算方法的研究. 内蒙古林学院学报,(1): 32 – 45.

续九如. 2006. 林木数量遗传学. 北京:高等教育出版社.

杨 萍,石志玲,陈晓阳. 2011. 杉木 3 种交配设计试验种子品质的遗传分析. 华南农业大学学报,32(4):67 – 70.

郑仁华. 2003. 不同年龄杉木无性系交配的研究. 林业科学, 39(Sp.1):81 – 86.

Barnes R D, Mullin L J. 1978. Three-year height performance of *Pinus elliottii* Engelm. var. *elliottii* × *P. taeda* L. hybrid families on three sites in Rhodesia. Silvae Genetica, 27:217 – 223.

Blada I. 1992. Analysis of genetic variation in a *Pinus strobus* × *P. griffithii* F1 hybrid population. Silvae Genetica, 41 (4/5): 282 – 289.

Comstock R E, Robinson H F. 1948. The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics, 4 (4): 254 – 266.

Comstock R E, Robinson H F. 1952. Estimation of average dominance of genes// Gowen J W. Heterosis. Ames, IA: Iowa State University Press, 494 – 516.

Fitzmaurice G M, Lipsitz S R, Ibrahim J G. 2007. A note on permutaion tests for variance components in multilevel generalized linear mixed models. Biometrics, 63(3): 942 – 946.

Gilmour A R, Gogel B J, Cullis B R, *et al.* 2009. ASReml User Guide Release 3.0. Hemel Hempstead, UK: VSN International Ltd.

Henderson C R. 1953. Estimation of variance and covariance components. Biometrics, 9(2):226 – 252.

Holland J B, Nyquist W E, Cervantes-Martinez C T. 2003. Estimating and interpreting heritability for plant breeding: An update. Plant Breeding Reviews, 22: 9 – 112.

Namkoong G. 1979. Introduction to quantitative genetics in forestry. U. S. Department of Agriculture Forest Service Technical Bulletin 1588.

Ofversten J. 1993. Exact tests for variance components in unbalanced mixed linear models. Biometrics, 49(1): 45 – 57.

Tong C F, Liu G X, Yang L W, *et al.* 2012. GSCA: new software and algorithms to analyse diallel mating designs based on restricted linear model. Silvae Genet, 61(3): 126 – 132.

Volker P W, Potts B M, Borralho N G M. 2008. Genetic parameters of intra- and inter-specific hybrids of *Eucalyptus globulus* and *E. nitens*. Tree Genetics and Genomes, 4(3): 445 – 460.

Wald A. 1941. On the analysis of variance in case of multiple classifications with unequal class frequencies. The Annals of Mathematical Statistics, 12(3): 346 – 350.

Williams E R, Matheson A C, Harwood C E. 2002. Experimental design and analysis for tree improvement. Collingwood, VIC, Australia: CSIRO Publishing.

Wright J W. 1976. Introduction to forest genetics. New York: Academic Press, Inc.

Wu H X, Matheson A C. 2000. Analysis of half-diallel mating design with missing crosses: Theory and SAS program for testing and estimating GCA and SCA fixed effects. Silvae Genet, 49(3):130 – 137.

(责任编辑 徐 红)