

Estimation of genotypic correlation and heritability of apricot traits, using restricted maximum likelihood in repeated measures data

Ahmad Ismaili, Farhad Karami, Omidali Akbarpour, and Abdolhossein Rezaei Nejad

Abstract: In estimation of genetic parameters in perennial tree species on the basis of analysis of variance (ANOVA), heterogeneity of years and genotype \times environment interaction for data sets during the juvenility to maturity life period is ignored. Therefore, a linear mixed model based on restricted maximum likelihood (REML) approximation for modeling of covariance structure of longitudinal data can improve our ability to analyze repeated measures data. In the present research, a modeling of variance-covariance structure by mixed model based on the REML approach has been used for characteristics of 26 apricot genotypes recorded during three years. Fitting unstructured covariance (UN) models for all traits indicated a great heterogeneity of variances among repeated years and the trends of response variables in the genotypes (except for RWC) was due to imperfect correlation of subjects measured in different years. Based on the same structure, positive correlations were estimated among fruit set, potassium content, and yield of pistil in repetitive years, and most traits showed high heritability estimation. To our knowledge, this is the first report in plant that genotypic correlation and heritability and their standard errors are estimated in a repeated measures data over years using REML approximation.

Key words: longitudinal data, unstructured covariance, mixed model, biplot, genotype \times year interaction, *Prunus armeniaca*.

Résumé : Lorsqu'on estime les paramètres génétiques des essences d'arbre pérennes par analyse de la variance (ANOVA), on ne tient pas compte de l'âge hétérogène des arbres ni des interactions entre leur génotype et l'environnement dans les jeux de données prélevées durant leur jeunesse, jusqu'à maturité. Un modèle linéaire mixte reposant sur l'approximation du maximum de vraisemblance restreint (REML) visant à reproduire la structure de la covariance des données longitudinales pourrait nous aider à mieux analyser les mesures répétitives. Les auteurs de cette étude ont modélisé la variance-covariance avec un modèle mixte s'appuyant sur le REML afin de vérifier les caractères de 26 génotypes d'abricot observés pendant trois ans. L'ajustement des modèles non structurés de covariance pour l'ensemble des caractères révèle la forte hétérogénéité de la variance les années où les relevés sont répétés, mais aussi que les tendances suivies par les variables-réaction du génotype (sauf la concentration d'eau relative) résultent d'une corrélation imparfaite entre les sujets évalués différentes années. Recourant à la même structure, les auteurs ont évalué les corrélations positives entre la nouaison, la teneur en potassium et le rendement des pistils lors des années répétitives et constaté que la plupart des caractères présentent une grande héritabilité. Pour autant qu'on le sache, il s'agit du premier rapport sur les plantes où l'on estime la corrélation génotypique, l'héritabilité et leurs écarts-types pour des mesures prises à répétition plusieurs années à partir d'une approximation du REML. [Traduit par la Rédaction]

Mots-clés : données longitudinales, covariance non structurée, modèle mixte, diagramme de double projection, interaction génotype \times année, *Prunus armeniaca*.

Received 22 August 2015. Accepted 2 December 2015.

A. Ismaili and O. Akbarpour. Department of Agronomy and Plant breeding, Faculty of Agriculture, Lorestan University, PO Box 465, Korramabad, Iran.

F. Karami. Agricultural and Natural Resources Research Center of Kurdistan, Gerizeh, PO Box 714, Sanandaj, Iran.

A. Rezaei Nejad. Department of Horticultural Sciences, Faculty of Agriculture, Lorestan University, PO Box 465, Korramabad, Iran.

Corresponding author: Ahmad Ismaili (email: ismaili.a@lu.ac.ir).

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from [RightsLink](https://www.rightslink.com).

Introduction

Tree breeders face several challenges when studying quantitative traits due to the ontogeny of these traits (Kozłowski et al. 1991). However, the challenges of quantitative traits for trees are derived from the fact that breeding of trees acquired more time to find out the nature of plants and their interaction with environment during the juvenility to maturity life period. Soil erosion, soil water-holding capacity, soil physical characteristics, competition for light, nutrients, water and other environmental factors often vary across repeated years leading to the genotype \times environment interaction.

In general, identification of morphological, phenological, and pomological characteristics and their genetic correlations are the first effort of preliminary breeding processes (Kramer and Twigg 1966; Perez-Gonzales 1992; Badenes et al. 1998; Asma and Ozturk 2005). In addition, the assessment of proportion of phenotypic variation that is due to variation in genetic values, defined as heritability, directly helps breeders to have a good evaluation and selection for their breeding's targets (Falconer 1960). Genetic assessment of perennial tree species usually encompass with the data in which the response variable is observed in sequence on the same subject or experimental unit. These types of data measured repeatedly over time or space are repeated measures data (Littel et al. 2006). The analyzing and estimation of genetic parameters in these data types always have special problems because of seasonal fluctuations as the source of heterogeneity and genotype \times year interaction. For instance, the importance of juvenile–mature correlations and the reliability of predictions of later performance from early assessment can vary with species, environments, and characteristics (Zamudio and Wolfinger 2002). However, improved statistical tools never compensate for precision lost while conducting an experiment using poor design or inappropriate agronomical and experimental practices. But nowadays, applying good statistical models and new technical analyses with increased computing power have led to robust estimation, providing ways to minimize or avoid bias and present valid results (Burgueño et al. 2000).

Plant breeders have traditionally estimated variance and covariance components using the method of moments on the basis of analysis of variance (ANOVA). Experimental designs use ANOVA to estimate experimental error and obtain unbiased results. The moment-based estimators are associated with sum of squares, expected mean squares (EMS), and the solution of EMS equations. Drawbacks of using method of moments to estimate variance and covariance components include ignorance of the estimators' distributional properties when data are unbalanced and the possibility of obtaining estimates outside of parameter bounds (Liu et al. 1997; Holland 2006).

Current linear mixed model methodology not only permits the presence of heterogeneity of variance in the linear model but also allows the researcher to address directly the covariance structure. Providing valid standard errors and efficient statistical tests, and modeling the covariance structure of the data can improve our ability to analyze repeated measures data (Zamudio et al. 2008). Holland (2006) presented outstanding codes using the Proc MIXED in SAS system to implement restricted maximum likelihood (REML) estimation of genotypic and phenotypic correlations and their standard errors concerning widely used experimental designs. Holland et al. (2001) used the REML approach to estimate genotypic and phenotypic correlations and their approximate standard errors for grain oil content and other agronomic traits in oat (*Avena sativa* L.). Zamudio and Wolfinger (2002) used a similar approach to estimate genetic covariances between measurements made at different ages on trees.

Apricot (*Prunus armeniaca* L.) is one of the most important fruit crops cultivated in wide geographical areas. Due to long its life period, nature of longitudinal data, and effect of seasonal fluctuations as the source of heterogeneity and genotype \times year interaction, the early evaluation based on best linear unbiased predictor (BLUP) approach for genetic parameters could improve apricot breeding programs. The objectives of this research were (i) to study the fixed effects of factors in response variables such as morphological and physiological traits in apricot genotypes through modeling and identifying the variance-covariance structure of repeated measure analysis to make inferences of mean traits trends, and (ii) to estimate genotypic correlations, heritability, and their standard errors in apricot genotype traits using the mixed model. To our knowledge, this is the first report in plant that genotypic correlation and heritability are estimated in repeated measures data over years using the REML approximation.

Materials and Methods

This study was conducted at the Agricultural Research Farm of Gerizeh, Sanadaj, Iran based on a randomized complete block design with three replications. The evaluations were done on eight commercial apricot cultivars of 'Ghermez-e-Shahroud', 'Dorosht-e-Malayer', 'Ghorban-e-Maragheh', 'Nasiri, Ordobad', 'Tilton', 'Royal', and 18 local genotypes of 'Ebrahimi', 'Azizi', 'Hashtaloui', 'Biglary', 'Bidaneh Kurdistan', 'Jahangiri', 'Sefid-e-Mezrah', 'Abdollahi', 'Malayer', 'Mezrah', 'Zoudras-e-Nisar', 'Zoudras-e-Gazneh', 'Ghamishlou', 'Khorramta 1', 'Khorramta 2', 'Naran', 'Ghaderi', and 'Noshour'. All genotypes were grafted on apricot seedlings and all trees were five years old when investigation was started in this research. Flowering time, flowering period, severity of frost injury in flowers, fruit set, relative water content (RWC), proline and K⁺ contents in pistil of flower, harvest time, yield, fruit weight, fruit size, and total soluble