# Bayesian estimation of a surface to account for a spatial trend using a semiparametric mixed model

Eduardo P. Cappa[12*] and Rodolfo J. C. Cantet[13]

[1] Department of Animal Production, University of Buenos Aires,
Avenida San Martín 4453, C1417DSQ Buenos Aires, Argentina.
[2] Doctoral fellow, "Fondo para la Investigación Científica y Tecnológica" (FONCyT), Argentina.
[3] "Consejo Nacional de Investigaciones Científicas y Técnicas" (CONICET), Argentina.

* Corresponding author
Eduardo Pablo Cappa
Departamento de Producción Animal, Facultad de Agronomía
Universidad de Buenos Aires
Av San Martín 4453
C1417DSQ Buenos Aires
Argentina

ph: +54 11 4524 8000 ext. 8192
fax: +54 11 4514 8735 or 8737
e-mail: ecappa@mail.agro.uba.ar

# Abstract

Unaccounted spatial variability leads to bias in estimating genetic parameters and predicting breeding values from forest genetic trials. Previous attempts to account for continuous spatial variation employed spatial coordinates in the direction of the rows (or columns). In this research, we use an individual tree mixed model and the tensor product of B-spline bases with a proper covariance structure for the knot effects to account for spatial variability. Dispersion parameters were estimated using Bayesian techniques via the Gibbs sampling. The procedure is illustrated with data from a progeny trial of *E. globulus*. Four different models were used in the sequel. The first model included block effects and the three other models included a surface on a grid of either 8 × 8, 12 × 12, or 18 × 18 knots. The three models with B-splines displayed a sizeable lower value of the Deviance Information Criterion than the model with blocks. Also, the mixed models fitting a surface displayed a consistent reduction in the posterior mean of $\sigma_e^2$, an increase in the posterior means of $\sigma_A^2$ and $h^2_{DBH}$, and an increase of 66 % (for parents) or 60% (for offspring) in the accuracy of breeding values.

# Introduction

Forest genetic trials are prone to a high degree of environmental heterogeneity as compared to other cultivated plants (Libby and Cockerham, 1980): trees are large living creatures and occupy more space than most cultivated plant species. Moreover, trees are often planted in places with heterogeneous levels of fertility, humidity, soil depth, or slope. Although spatial heterogeneity is a nuisance effect in forest genetic evaluation where the main goal is the prediction of breeding values, ignoring such a source can lead to biases in the estimation of genetic parameters and the prediction of individual additive genetic effects (breeding values, Magnussen 1993, 1994). To account for environmental gradients, tree breeders have devised forest trials using randomized complete blocks or incomplete block designs. However, setting fixed limits for the blocks makes it difficult to account for continuously varying environmental factors. Additionally, establishing *a priori* a design that properly account for all sources of environmental heterogeneity may be a hopeless task as "environmental variation is never known prior to establishment" (Fu et al. 1999a). Alternatively, the spatial variation can be accounted for *a posteriori* within the model of evaluation. In these so called 'spatial models', variability has two main sources: the *local trend*, or small-scale variation, and the *global trend* or large-scale variation across a spatial gradient. The two sources are observable in forest genetic trials: either component alone or in combination with each other (e.g., Fu et al. 1999b; Costa e Silva et al. 2001; Dutkowsky et al. 2002; Hamann 2002; Dutkowsky et al. 2006).

Models that account for continuous spatial variation include spatial coordinates expressed as either classification variables or covariables. The latter are non-stochastic functions such as polynomials (Federer 1998) or smoothing splines (Verbyla et al. 1999).

Costa e Silva et al. (2001) and Dutkowsky et al. (2002) considered the global trend in one dimension, either row-wise or column-wise, after adjusting first order autoregressive (AR(1)) and separable covariance structures (Gilmour et al. 1997). Costa e Silva et al. (2001) proposed the use of a classification variable for columns. Also, Dutkowsky et al. (2002) modeled global variation with linear models of fixed effects that included spatial coordinates in one dimension, fitted as quadratic polynomials or cubic smoothing splines (Verbyla et al. 1999). In the latter case, the resulting variogram was not stationary, so that patterns of unaccounted variability were still present in the residuals, most probably associated with rows by columns interactions (Dutkowsky et al. 2002, p. 2205). Therefore, the analysis of forest genetic trials where continuous spatial variation may develop in two dimensions, using classification variables or covariables only in the direction of the rows (or columns), may not completely account for the spatial variability. Thomson and El-Kassaby (1988) fitted sixth order degree polynomials in two dimensions by least-squares to compare different provenances of Douglas-fir. The use of polynomials for the analysis in two dimensions (*trend analysis*) of forest genetic data can also be found in the work of Liu and Burkhart (1994) and Saenz-Romero et al. (2001). However, the fit of polynomials suffer from several drawbacks (Green and Silverman 1994, p. 2). First of all, the fit is global and not local, which means that: 1) the method is not capable of accounting for local variations present in the data; 2) few influential observations exert a large influence in the resulting fit; 3) the fit in the extremes is usually poor. Another serious drawback with polynomials is its numerical instability as the order of the polynomial increases.

Splines are a more efficient approach to the use of polynomials. They are segmented polynomial functions that are locally fitted such that the resulting function is differentiable at the joints of the segments (*knots*), up to the order of fit. Splines are able to capture most

89 sinuosities present in the data and do not suffer from numerical instability. Eilers and Marx
90 (1996) introduced *penalized* splines in one dimension using B-splines with equally spaced
91 knots, and a linear model approach with a roughness penalty consisting on the differences
92 among the parameters, i.e. the effects of the knots. T. Speed (see Robinson 1991) first
93 pointed out the connection between splines and mixed models, a subject further expanded
94 by Ruppert et al. (2003) and Wand (2003). Cantet et al. (2005) approached P-splines in one
95 dimension using proper covariance structures rather matrices of differences, in an animal
96 breeding context. Eilers and Marx (2003) extended their methodology to estimate a surface
97 along two dimensions, using the tensor product of B-splines. The goal of the present
98 research is to show how to fit a surface using the tensor product of B-spline bases, to
99 account for continuous spatial variation in an individual tree mixed model for forest genetic
100 evaluation. To do that, we superimpose a covariance structure for the knot effects in a two-
101 dimensional grid. As in some recent contributions to forest breeding (Soria et al. 1998;
102 Gwaze and Woolliams 2001; Zeng et al. 2004; Cappa and Cantet 2006a; Waldmann and
103 Ericsson 2006), we employed the Bayesian approach *via* Gibbs sampling to make
104 inferences in all dispersion parameters of the model. Developments are illustrated by means
105 of a progeny trial data on diameter at breast height in *Eucalyptus globulus* ssp. *globulus*.
106 The resulting estimates of all dispersion parameters for mixed models that include the fitted
107 surface are finally compared with corresponding estimates from the classical model
108 including blocks.
109

# Methods

## Two-dimensional tensor product of B-splines

We first briefly introduce penalized splines (P-splines) in one dimension as suggested by Eilers and Marx (1996). Then, we take the approach of Eilers and Marx (2003) and Green and Silverman (1994) and extend P–splines to two dimensions using the tensorial product of P–splines.

Eilers and Marx (1996) advocated using B-splines with equally spaced knots to obtain penalized splines. B-splines are local basis functions, consisting of polynomial segments of degree $d$, in general quadratic or cubic, that have $d - 1$ continuous derivatives at the joining points, or knots. A B-spline of degree $d$ is positive on a domain spanned by $d + 2$ knots and is zero elsewhere. All in all, $d + 1$ B-spline coefficients are nonzero. Eilers and Marx (1996) introduced a penalty that affects first or second differences of B-spline coefficients. The penalty controls the degree of smoothness while fitting the function. Let $\boldsymbol{y}$ and $\boldsymbol{x}$ be vectors of length $n$ containing the observed and explanatory variables, respectively, and let $\boldsymbol{s}(\boldsymbol{x})$ be a spline function written as:

$$\boldsymbol{s}(\boldsymbol{x}) = \sum_{i=1}^{k} \boldsymbol{B}_i(x)\,\boldsymbol{b}_i \qquad [1]$$

where $\boldsymbol{B}_i = \left(\boldsymbol{B}_1(x), \boldsymbol{B}_2(x), ..., \boldsymbol{B}_k(x)\right)'$ is a column vector with B-spline bases (De Boor, 1993), and $\boldsymbol{b}_i = \left(\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_k\right)'$ denotes the vector of spline coefficients in one dimension. In matrix form, expression [1] can be written as $\boldsymbol{Bb}$, being $\boldsymbol{B}$ the $n \times k$ matrix that contains the $\boldsymbol{B}_i$'s, and $\boldsymbol{b}$ is the parametric vector ($k \times 1$) containing the $\boldsymbol{b}_i$'s to form $\boldsymbol{s}(\boldsymbol{x})$. The functional [1] is generally fitted by least-squares with an additive penalty. Eilers and Marx

(1996) observed that the penalized estimator of $\boldsymbol{b}$ is the solution of the following system of equations:

$$\left(\boldsymbol{B}'\boldsymbol{B} + \lambda\, \boldsymbol{D}_d'\boldsymbol{D}_d\right)\hat{\boldsymbol{b}} = \boldsymbol{B}'\boldsymbol{y} \qquad [2]$$

where the positive scalar $\lambda$ controls the amount of smoothing and $\boldsymbol{D_d}$ is the matrix of differences of order $d$. For $d = 1$ and $d = 2$, we respectively have:

$$\boldsymbol{D}_1 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}; \qquad \boldsymbol{D}_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{bmatrix} \qquad [3]$$

Ruppert et al. (2003) and Wand (2003) discussed the connection between P-splines and mixed models (Henderson, 1984). The smoothing parameter $\lambda$ is seen as the ratio of the error variance to the variance of the B-spline coefficients $\boldsymbol{b}_i$. Moreover, $\boldsymbol{D}'\boldsymbol{D}$ is interpreted as a g-inverse of the covariance matrix of the B-spline coefficients (Cantet et al. 2005), and acts as a singular penalization matrix.

Tensor products of B-splines allow a natural extension of one dimensional P-spline smoothing to two dimensions by means of the Kronecker product of single structures. A more rigorous approach can be found in Green and Silverman (1994, p. 155-159). The tensor product of two univariate B-splines along the rows ($r$) and columns ($c$) is defined as the $r \times c$ rectangle in $\Re^2$ such that $\boldsymbol{T}_{kl}(r,c) = \boldsymbol{Br}_k(r)\,\boldsymbol{Bc}_l(c)$, where $\boldsymbol{Br}_k(r)$ and $\boldsymbol{Bc}_l(c)$ are B-spline bases for the row ($k$ = 1, 2, …, $nx_r$) and column ($l$ = 1, 2, …, $nx_c$) knot effects, respectively. If row and column knots are chosen to be equally spaced, the $r \times c$ space can be divided in small rectangular panels such that $[r_k, r_{k+6}] \times [c_l, c_{l+6}]$. Let $\boldsymbol{S} = [\boldsymbol{\gamma}_{kl}]$ be the $nx_r \times nx_c$ matrix containing the coefficients from the tensor product of B-splines that have to be

estimated. Then, for a given set of knots the surface $\left(\alpha\left(r,c\right)\right)$ can be approximated using the following matrix expression

$$\mathrm{vec}\left\{\alpha\left(r,c\right)\right\} = \boldsymbol{B}\,\boldsymbol{b} \qquad [4]$$

where $\boldsymbol{B}$ has dimension $\boldsymbol{n}\times\left(nx_r \times nx_c\right)$ and is equal to $\boldsymbol{B} = \left(\boldsymbol{B}_r \otimes 1'_{nx_c}\right)\odot\left(1'_{nx_r} \otimes \boldsymbol{B}_c\right)$. The notation vec stands for the operator that results from stacking the columns of a matrix into a vector, and the symbols $\otimes$ and $\odot$ indicate the Kronecker and Hadamard products of matrices, respectively (Harville, 1997).

In analogy to what they had done for one dimension (Eilers and Marx, 1996), Eilers and Mark (2003) and Marx and Eilers (2005) proposed a two-dimensional penalized estimation of a surface. Let $\lambda_r$ and $\lambda_c$ be the parameters controlling the degree of smoothness for rows and columns, respectively, whereas $\boldsymbol{D_r}$ and $\boldsymbol{D_c}$ are the respective difference matrices [3]. Then, the solution for $\hat{\boldsymbol{b}}$ is obtained by solving the equations

$$\left(\boldsymbol{B}'\boldsymbol{B} + \lambda_r\left(\boldsymbol{I}_{\boldsymbol{nxr}} \otimes \boldsymbol{D}_r'\boldsymbol{D}_r\right)_r + \lambda_c\left(\boldsymbol{D}_c'\boldsymbol{D}_c \otimes \boldsymbol{I}_{\boldsymbol{nxc}}\right)\right)\hat{\boldsymbol{b}} = \boldsymbol{B}'\boldsymbol{y} \qquad [5]$$

The expression above is similar to the system in one dimension where $\boldsymbol{B}$ is replaced by $\boldsymbol{Br}$ or $\boldsymbol{Bc}$, and $\lambda\ \boldsymbol{D}'\boldsymbol{D}$ is replaced by $\lambda_r\left(\boldsymbol{I}_{\boldsymbol{nxr}} \otimes \boldsymbol{D}_r'\boldsymbol{D}_r\right)_r + \lambda_c\left(\boldsymbol{D}_c'\boldsymbol{D}_c \otimes \boldsymbol{I}_{\boldsymbol{nxc}}\right)$. In the next section, we show how to fit data in two dimensions using the tensor product of B-splines by means of a mixed linear model.

**Mixed model representation of a two-dimensional tensor product of B-splines**

In forest genetic trials trees are usually arranged in regular grids arrayed in rows and columns. In order to position any tree, let *r* and *c* be the row and column coordinates,

172 respectively, measured in meters or degrees. Let $\boldsymbol{Y}$ be an $n_r$ (number of rows) × $n_c$ (number
173 of columns) containing the observations for a trait (such as height, or diameter). Consider
174 also the vector $\boldsymbol{y}$ such that $\boldsymbol{y}$ = vec($\boldsymbol{Y}$), so that data are ordered by column within row. Then,
175 an individual tree mixed model with a smoothed surface to account for environmental
176 heterogeneity is equal to

177 $$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{B}\,\boldsymbol{b} + \boldsymbol{Z}\,\boldsymbol{a} + \boldsymbol{e} \qquad [6]$$

178 In [6], $\boldsymbol{\beta}$ is a $p \times 1$ vector of fixed effects associated to $\boldsymbol{y}$ by the incidence matrix $\boldsymbol{X}$ ($n \times p$)
179 such that $r[\boldsymbol{X}] = p$. In case $r[\boldsymbol{X}] < p$, it is always possible to find a reparametrization that
180 turns $\boldsymbol{X}$ into a matrix of full-column rank (Christensen, 1987). The random $q \times 1$ vector $\boldsymbol{a}$
181 contains the breeding values, and is related to $\boldsymbol{y}$ by the incidence matrix $\boldsymbol{Z}$ (of order $n \times q$).
182 The expectation of $\boldsymbol{a}$ is $\boldsymbol{0}$ and the covariance matrix is $\boldsymbol{A}\ \sigma^2_{\mathrm{A}}$ where $\boldsymbol{A}$ is the additive
183 relationship matrix (Henderson, 1984) among trees, and $\sigma^2_{\mathrm{A}}$ is the additive genetic
184 variance. The distribution of the random vector $\boldsymbol{b}$ containing the coefficients of the tensor
185 product of B-splines is such that $\boldsymbol{b} \sim N\left(\boldsymbol{0}, \boldsymbol{U}\sigma^2_b\right)$. The scalar $\sigma^2_b$ is the variance of the
186 coefficients for rows and columns and $\boldsymbol{U}$ is the covariance structure in two-dimensions.
187 Finally, random error terms are included in the $n \times 1$vector $\boldsymbol{e}$, which is distributed as
188 $\boldsymbol{e} \sim N\left(\boldsymbol{0}, \boldsymbol{I}\sigma^2_e\right)$ and $\sigma^2_e$ is the error variance.

189 The covariance structure $\boldsymbol{U}$ plays an important role in model [6]. The matrix should
190 reflect the correlation decay among B-spline knots that are further apart, either row or
191 column-wise. A possible choice for $\boldsymbol{U}$ is $\boldsymbol{\Sigma}_r \otimes \boldsymbol{\Sigma}_c$, a Kronecker product of matrices for the
192 rows ($\boldsymbol{\Sigma}_r$) and for the columns ($\boldsymbol{\Sigma}_c$). If $\boldsymbol{U}$ is a linear covariance structure (Anderson, 1973),
193 the estimation process is simplified and there is only one parameter to estimate: $\sigma^2_b$. Then,

194 estimation can be performed with simpler methods and algorithms, i.e. REML-EM or
195 Gibbs sampling. The challenge is to find a $\boldsymbol{U}$ that is informative enough among the
196 correlation decay among knot effects, at the same time that does not depend on extra
197 parameters. In this regard, we will set $\boldsymbol{\Sigma}_r$ and $\boldsymbol{\Sigma}_c$ to be equal to the one-dimensional
198 covariance structure originally proposed by Green and Silverman (1994, p. 13) and then
199 used by Durban et al. (2001) to fit a fertility trend. In this tridiagonal matrix, correlations
200 are non-zero for neighbor knots, and are 0 otherwise. More explicitly, if $\varsigma_{ij}$ is element *ij* of
201 any of the matrices $\boldsymbol{\Sigma}_r$ or $\boldsymbol{\Sigma}_c$, diagonals are $\varsigma_{ii} = 4/6$, whereas off-diagonals are either
202 $\varsigma_{i+1,i} = \varsigma_{i,i+1} = 1/6$ or $\varsigma_{ij} = 0$ for $|i - j| \geq 2$, $i = j = 1, 2, \ldots, nx_r$ or $nx_c$. Thus, besides being
203 positive definite, $\boldsymbol{U} = \boldsymbol{\Sigma}_r \otimes \boldsymbol{\Sigma}_c$ is strictly diagonally dominant as $|\varsigma_{ii}| > \sum_{j \neq i} |\varsigma_{ij}|$ for every *i*.
204 To exemplify, suppose $nx_r = nx_c = 4$, then

205
$$\boldsymbol{\Sigma}_r = \boldsymbol{\Sigma}_c = \frac{1}{6}\begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

206 and $\boldsymbol{U} = \boldsymbol{\Sigma}_r \otimes \boldsymbol{\Sigma}_c$ is equal to

207
$$\frac{1}{6}\begin{bmatrix}
16 & 4 & 0 & 0 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
4 & 16 & 4 & 0 & 1 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 4 & 16 & 4 & 0 & 1 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 4 & 16 & 0 & 0 & 1 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
4 & 1 & 0 & 0 & 16 & 4 & 0 & 0 & 4 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 4 & 1 & 0 & 4 & 16 & 4 & 0 & 1 & 4 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 4 & 1 & 0 & 4 & 16 & 4 & 0 & 1 & 4 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 4 & 0 & 0 & 4 & 16 & 0 & 0 & 1 & 4 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 4 & 1 & 0 & 0 & 16 & 4 & 0 & 0 & 4 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 4 & 1 & 0 & 4 & 16 & 4 & 0 & 1 & 4 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 4 & 1 & 0 & 4 & 16 & 4 & 0 & 1 & 4 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 & 0 & 4 & 16 & 0 & 0 & 1 & 4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 1 & 0 & 0 & 16 & 4 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 1 & 0 & 4 & 16 & 4 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 1 & 0 & 4 & 16 & 4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 & 0 & 4 & 16
\end{bmatrix}$$

208 In this example non-zero elements of $\boldsymbol{U}$ are correlations between neighbor knots.
209 Take for example, the second knot (row 2 of $\boldsymbol{U}$) having as proximal neighbors the knots 1,
210 3 and 6, and as diagonal neighbors the knots 5 and 7. Notice that correlations with
211 neighbors in proximal positions are stronger (4/6) than with neighbors located diagonally
212 (1/3). Implicit is the assumption that the spacing between both columns and rows is equal.
213 There other structures that allow modeling a gradual decay in correlation as knots are
214 separated further in the direction of the rows or of the columns, such as those proposed by
215 Hyndman et al. (2005) or Cantet et al. (2005). Finally, given the random effects in [6], the
216 covariance matrix $\boldsymbol{y}$ (say $\boldsymbol{V}$) is as follows:

217
$$\boldsymbol{V} = \boldsymbol{Z}\,\boldsymbol{A}\,\boldsymbol{Z}'\sigma^2_{\mathrm{A}} + \boldsymbol{B}\,\boldsymbol{U}\,\boldsymbol{B}'\sigma^2_b + \boldsymbol{I}_n\,\sigma^2_e \qquad [7]$$

218 and mixed model equations (Henderson, 1984) for [6] are

219
$$\begin{bmatrix}
\boldsymbol{X}'\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{B} & \boldsymbol{X}'\boldsymbol{Z} \\
\boldsymbol{B}'\boldsymbol{X} & \boldsymbol{B}'\boldsymbol{B}+\boldsymbol{U}^{-1}\lambda & \boldsymbol{B}'\boldsymbol{Z} \\
\boldsymbol{Z}'\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{B} & \boldsymbol{Z}'\boldsymbol{Z}+\boldsymbol{A}^{-1}\alpha
\end{bmatrix}
\begin{bmatrix}
\hat{\beta} \\ \hat{\boldsymbol{b}} \\ \hat{\boldsymbol{a}}
\end{bmatrix}
=
\begin{bmatrix}
\boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{B}'\boldsymbol{y} \\ \boldsymbol{Z}'\boldsymbol{y}
\end{bmatrix} \qquad [8]$$

where $\lambda = \sigma_e^2/\sigma_b^2$ and $\alpha = \sigma_e^2/\sigma_A^2$. Notice that in the Bayesian view of the mixed linear model (Sorensen and Gianola, 2002) the likelihood of the data is proportional to

$$p(\boldsymbol{y}|\boldsymbol{\beta},\boldsymbol{a},\boldsymbol{b}) \propto (\sigma_e^2)^{-\frac{1}{2}} \exp\left[\frac{1}{2\sigma_e^2}(\boldsymbol{y}-\boldsymbol{X\beta}-\boldsymbol{Za}-\boldsymbol{Bb})'(\boldsymbol{y}-\boldsymbol{X\beta}-\boldsymbol{Za}-\boldsymbol{Bb})\right] \quad [9]$$

**Bayesian estimation**

The Bayesian approach *via* Gibbs sampling was used to estimate the parameters in model [6] (Sorensen and Gianola, 2002). We now specify the prior distributions, as well as the joint and marginal conditional posterior densities.

**Specification of prior distributions**: Conjugate prior densities were chosen for all parameters. To reflect a prior state of uncertainty for the fixed effects and to keep a proper posterior distribution (Hobert and Casella 1996), we set $\boldsymbol{\beta} \sim N_p(\boldsymbol{0}, \boldsymbol{K})$ and $\boldsymbol{K}$ is a diagonal matrix with very large elements ($k_{ii} > 10^8$). Therefore, this prior density is proportional to:

$$p(\boldsymbol{\beta}|\boldsymbol{K}) \propto |\prod_{i=1}^{p} k_{ii}|^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\sum_{i=1}^{p}\frac{\boldsymbol{\beta}_i^2}{k_{ii}}\right\} \quad [10]$$

The vector of the tensor product of B-spline coefficients $\boldsymbol{b}$ is distributed *a priori* as $\boldsymbol{b} \sim N_b(\boldsymbol{0}, \boldsymbol{U}\,\sigma_b^2)$, so that:

$$p(\boldsymbol{b}|\sigma_b^2) \propto (\sigma_b^2)^{-\frac{nx*nx}{2}} \exp\left\{-\frac{\boldsymbol{b}'\boldsymbol{U}^{-1}\boldsymbol{b}}{2\sigma_b^2}\right\} \quad [11]$$

The prior density for the vector of breeding values is $\boldsymbol{a} \sim N_q(\boldsymbol{0}, \boldsymbol{G}_0 \otimes \boldsymbol{A})$ (see (13.38) in Sorensen and Gianola, 2002, p. 578), so that:

$$p(\boldsymbol{a}|\sigma_A^2) \propto (\sigma_A^2)^{-\frac{q}{2}} \exp\left\{-\frac{\boldsymbol{a}'\boldsymbol{A}^{-1}\boldsymbol{a}}{2\sigma_A^2}\right\} \quad [12]$$

Following Sorensen and Gianola (2002), we chose to use independent scaled inverted chi-square densities as prior distributions for the variance components $\sigma^2_b$, $\sigma^2_A$ and $\sigma^2_e$:

$$p\left(\sigma^2_b \mid \upsilon_b, \delta^2_b\right) \propto \left(\sigma^2_b\right)^{-\left(\frac{\upsilon_b}{2}+1\right)} \exp\left\{-\frac{\upsilon_b \delta^2_b}{2\sigma^2_b}\right\} \quad [13]$$

$$p\left(\sigma^2_A \mid \upsilon_A, \delta^2_A\right) \propto \left(\sigma^2_A\right)^{-\left(\frac{\upsilon_A}{2}+1\right)} \exp\left\{-\frac{\upsilon_A \delta^2_A}{2\sigma^2_A}\right\} \quad [14]$$

$$p\left(\sigma^2_e \mid \upsilon_e, \delta^2_e\right) \propto \left(\sigma^2_e\right)^{-\left(\frac{\upsilon_e}{2}+1\right)} \exp\left\{-\frac{\upsilon_e \delta^2_e}{2\sigma^2_e}\right\} \quad [15]$$

Parameters in the densities [13], [14], and [15], are the hypervariances $\delta^2_b$, $\delta^2_A$ and $\delta^2_e$, and the degrees of freedom $\upsilon_b$, $\upsilon_A$ and $\upsilon_e$, respectively.

**Joint and conditional posterior densities**: By multiplying [9] with [10], [11], [12], [13], [14], and [15], the joint posterior density of all parameters is proportional to:

$$p\left(\boldsymbol{\beta}, \boldsymbol{a}, \boldsymbol{b}, \sigma^2_b, \sigma^2_A, \sigma^2_e \mid \boldsymbol{y}, \upsilon_b, \upsilon_A, \upsilon_e, \delta^2_b, \delta^2_A, \delta^2_e\right) \propto$$

$$\left(\sigma^2_e\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2_e}(\boldsymbol{y}-\boldsymbol{X\beta}-\boldsymbol{Bb}-\boldsymbol{Za})'(\boldsymbol{y}-\boldsymbol{X\beta}-\boldsymbol{Bb}-\boldsymbol{Za})\right] \exp\left\{-\frac{1}{2}\sum_{i=1}^{p}\frac{\boldsymbol{\beta}_i^2}{k_{ii}}\right\}$$

$$\left(\sigma^2_b\right)^{-\frac{nx*nx}{2}} \exp\left\{-\frac{\boldsymbol{b}'\boldsymbol{U}^{-1}\boldsymbol{b}}{2\sigma^2_b}\right\} \left(\sigma^2_A\right)^{-\frac{q}{2}} \exp\left\{-\frac{\boldsymbol{a}'\boldsymbol{A}^{-1}\boldsymbol{a}}{2\sigma^2_A}\right\} \left(\sigma^2_b\right)^{-\left(\frac{\upsilon_b}{2}+1\right)} \exp\left\{-\frac{\upsilon_b \delta^2_b}{2\sigma^2_b}\right\}$$

$$\left(\sigma^2_A\right)^{-\left(\frac{\upsilon_A}{2}+1\right)} \exp\left\{-\frac{\upsilon_A \delta^2_A}{2\sigma^2_A}\right\} \left(\sigma^2_e\right)^{-\left(\frac{\upsilon_e}{2}+1\right)} \exp\left\{-\frac{\upsilon_e \delta^2_e}{2\sigma^2_e}\right\} \quad [16]$$

Inference on any parameter by means of the Gibbs sampler requires conditional posterior densities in close form. The joint conditional density of $\boldsymbol{\beta}$ and $\boldsymbol{b}$, and $\boldsymbol{a}$ is equal to

$$\begin{bmatrix}\boldsymbol{\beta}\\ \boldsymbol{b}\\ \boldsymbol{a}\end{bmatrix} \Big| \; \boldsymbol{y}, \sigma^2_{\mathrm{A}}, \sigma^2_b, \sigma^2_e \sim N\left(\begin{bmatrix}\boldsymbol{\beta}\\ \boldsymbol{b}\\ \boldsymbol{a}\end{bmatrix}, \begin{bmatrix}\boldsymbol{X}'\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{B} & \boldsymbol{X}'\boldsymbol{Z}\\ \boldsymbol{B}'\boldsymbol{X} & \boldsymbol{B}'\boldsymbol{B}+\boldsymbol{U}^{-1}\lambda & \boldsymbol{B}'\boldsymbol{Z}\\ \boldsymbol{Z}'\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{B} & \boldsymbol{Z}'\boldsymbol{Z}+\boldsymbol{A}^{-1}\alpha\end{bmatrix}^{-1}\right) \qquad [17]$$

Vectors $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{a}}$ are the solutions to equations [8]. The conditional posterior distribution of $\sigma^2_{\mathrm{A}}$ is scaled inverted chi-square

$$p\left(\sigma^2_{\mathrm{A}} \,|\, \boldsymbol{\beta}, \boldsymbol{b}, \boldsymbol{a}, \sigma^2_b, \sigma^2_e, \boldsymbol{y}\right) \propto \mathrm{Inv}\text{-}\chi^2\left(\tilde{\upsilon}_{\mathrm{A}}, \tilde{\delta}^2_{\mathrm{A}}\right) \qquad [18]$$

with parameters $\tilde{\upsilon}_{\mathrm{A}} = q + \upsilon_{\mathrm{A}}$ and $\tilde{\delta}^2_{\mathrm{A}} = \left(\boldsymbol{a}'\boldsymbol{A}^{-1}\boldsymbol{a} + \upsilon_{\mathrm{A}}\delta^2_{\mathrm{A}}\right)/\tilde{\upsilon}_{\mathrm{A}}$. Also, for $\sigma^2_b$ we have

$$p\left(\sigma^2_b \,|\, \boldsymbol{\beta}, \boldsymbol{b}, \boldsymbol{a}, \sigma^2_{\mathrm{A}}, \sigma^2_e, \boldsymbol{y}\right) \propto \mathrm{Inv}\text{-}\chi^2\left(\tilde{\upsilon}_b, \tilde{\delta}^2_b\right) \qquad [19]$$

with $\tilde{\upsilon}_b = nx * nx + \upsilon_b$ and $\tilde{\delta}^2_b = \left(\boldsymbol{b}'\boldsymbol{U}^{-1}\boldsymbol{b} + \upsilon_b\delta^2_b\right)/\tilde{\upsilon}_b$. Finally, the error variance has the following conditional posterior

$$p\left(\sigma^2_e \,|\, \boldsymbol{\beta}, \boldsymbol{a}, \boldsymbol{b}, \sigma^2_b, \sigma^2_{\mathrm{A}}, \boldsymbol{y}\right) \propto \left(\sigma^2_e\right)^{-\left(\frac{n+\upsilon_e+2}{2}+1\right)} \exp\left\{-\frac{\tilde{\upsilon}_e\tilde{\delta}^2_e}{2\sigma^2_e}\right\} \qquad [20]$$

with $\tilde{\upsilon}_e = n + \upsilon_e$ degrees of freedom and scale parameter $\tilde{\delta}^2_e = \left(\boldsymbol{e}'\boldsymbol{e} + \upsilon_e\delta^2_e\right)/\tilde{\upsilon}_e$. At any iteration of the Gibbs algorithm, we first sampled from distribution [17], then from [20], then from [18], and finally from [19], to start the process back again. A program was written in FORTRAN to perform all calculations.

# A working example: Analysis of an *E. globulus* progeny trial

**Data**

A *Eucalyptus globulus* ssp. *globulus* progeny trial was used in the study. The data were collected at Licenciado Matienzo (lat. 37º 59' 578" S long. 59º 00' 107" W), in the

272 southeastern part of Buenos Aires province, Argentina, where *E. globulus* has traditionally
273 being planted (Lopez et al. 2001). The soil was a fine Petrocalcic Paleudoll, with
274 subsurficial petrocalcic horizon (locally known as "tosca") at variable depth. There were
275 1080 trees from seventy two seed lots: 36 open pollinated families from 8 native stand sites
276 in Australia, 30 open pollinated families and 6 bulk collections from land race from
277 Argentina, Portugal, Spain and Chile (Lopez et al. 2001). After including all known genetic
278 relationships, a total of 1148 individual trees were used in the pedigree file. The trait was
279 diameter at breast height (1.3 m, DBH), measured when trees were 6 year-old in cm. Trees
280 were planted in single-tree plots on a rectangular grid of 32 rows and 36 columns (93 m ×
281 105 m) arrayed in squares of 3 by 3 meters, with 15 replicates per family. Then, rows have
282 coordinates $r_i$, $i = 1, 2, \ldots, R = 32$ and columns coordinates $c_j$, $j = 1, 2, \ldots, C = 36$. For the
283 purpose of model fitting, row ($r$) and column ($c$) spatial coordinates were expressed in
284 meters and the origin was taken to be the north corner. The first tree ($r = 1$, $c = 1$) was set to
285 coordinates (0, 0), so that $R$ was equal to 93 m and $C$ to 105 m.
286
287 **Models of analysis**

288 Four individual additive tree models were evaluated. All models included a fixed
289 effect of genetic group to account for the means of the different origins of parents, random
290 additive genetic effects (breeding values), and random errors. Model 1 also included fixed
291 block effects. In the other three models (2, 3, and 4), a surface was fitted using the tensor
292 products of cubic B-splines. These models differ in the number of knots: $8 \times 8$, $12 \times 12$ and
293 $18 \times 18$, for models 2, 3 and 4, respectively. The coefficients for the cubic B-splines in $\boldsymbol{B}$
294 were calculated using the recursive algorithm of De Boor (1993), and the order of the

295 resulting matrix was $n \times (nx_r \times nx_c)$. Accordingly, the vector $\boldsymbol{b}$ was of order $(nx_r \times nx_c) \times 1$,

296 and the covariance structure $\boldsymbol{U}$ of order $(nx_r \times nx_c) \times (nx_r \times nx_c)$. The Deviance Information

297 Criterion (DIC, Spiegelhalter et al. 2002) was employed to compare the fit from different

298 models. The model with the smallest value of DIC should be favored, as this indicates a

299 better fit and a lower degree of model complexity. Numerical details for the calculus of

300 DIC in individual tree models were given by Cappa and Cantet (2006a).

301 Further model comparison was provided by the accuracy of prediction of breeding

302 values, which was computed using the following expression:

303 $$\boldsymbol{r} = \sqrt{\frac{1 - \mathrm{PEV}}{\sigma^2_{\mathrm{A}}}}$$

304 The acronym PEV stands for 'prediction error variance' (Henderson, 1984) of predicted

305 breeding values using the "Best linear unbiased predictors" (BLUPs) of parent and

306 offspring. The PEV is calculated as the diagonal elements of the inverse of the coefficient

307 matrix from the mixed model equations (Henderson, 1984) in [8]. The required variance

308 components to set up the mixed model equations were those estimated from the Bayesian

309 analysis. Spearman correlations were also estimated to compare predicted breeding values

310 from different models.

311

312 **Spatial analysis of residuals**

313 In order to identify spatial patterns in the data, we examined the spatial distribution

314 and the variogram of the residuals as suggested by Gilmour et al. (1997), using a model

315 with fixed genetic groups and random breeding values. The distribution of the DBH

316 residuals is displayed in Figure 1. The color intensity represents the magnitude of the

317 residuals: the darker the dot, the larger the residual value. Additionally, residuals were
318 plotted against row and column position, to detect dissimilar patterns in any row (across
319 columns, Figure 2a), or in any column (across rows, Figure 2b). To exemplify, only rows 1,
320 16, and 32, and columns 1, 16 and 32, are displayed. Notice the different residual patterns
321 across rows or columns, which indicate the presence of interaction between row and
322 column position and the need for a two-dimensional smoothing. This effect is also observed
323 in the sample variogram displayed in Figure 3, where there is a consistent increase in the
324 semivariance as the displacements in the row and column directions increase. Note the
325 steeper slope row-wise (on the left side of the figure), as compared to the column-wise
326 slope (on the right side of the figure).

327 [Insert **Figure 1** about here]

328 [Insert **Figure 2** about here]

329 [Insert **Figure 3** about here]

330

331 **Computational details and posterior inference**

332 The values of the hypervariances $\delta^2_{\mathrm{A}}$, $\delta^2_b$ and $\delta^2_e$ were calculated using a single trait
333 Gibbs sampler from the same data set. The degrees of belief were set to 10 (i.e. $n_A = \nu_k =$
334 10) to reflect a relatively high degree of uncertainty. The deviance information criterion
335 (DIC) was computed for each model using the output from the Gibbs sampling. At the end
336 of each iteration, heritability of DBH was calculated as $h^2_{\mathrm{DBH}} = \tilde{\sigma}^2_{\mathrm{A}} \big/ \left( \tilde{\sigma}^2_{\mathrm{A}} + \tilde{\sigma}^2_e \right)$ where $\tilde{\sigma}^2_{\mathrm{A}}$ and
337 $\tilde{\sigma}^2_e$ are the values of the additive and error variance sampled at a given iteration.

A single Gibbs chain of 1 010 000 iterations was sampled, and the first 10 000 iterates were discarded due to *burn-in*. Autocorrelations were calculated with "*Bayesian Output Análisis*" (BOA version 1.0.1, Smith 2003) for all lags from 1 to 50. ). To evaluate the impact of autocorrelations in the variability of the samples, the 'effective sample size' (ESS) proposed by R. Neal (Kass et al. 1998) was calculated for each parameter as:

$$\mathrm{ESS} = \frac{1000\,000}{1 + 2\sum_{i=1}^{50} \rho(i)}$$

where $\rho(i)$ is the autocorrelation measured at lag *i*. Marginal posterior densities for all parameters were estimated by the Gaussian kernel method (Silverman 1986; chapter 2):

$$f(\theta) = \frac{1}{1\,000\,000\,h} \sum_{i=1}^{1000000} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\theta_i}{h}\right)^2\right] \qquad [21]$$

In (16), $f(\theta)$ is the estimated posterior density, $\theta_i$ ($i$ =1,..., 1 000 000) is a sampled value and *h* is the window width estimated by unbiased cross-validation. Mean, mode, median, standard deviation (SD), and 95% high posterior density interval (95% HPD), were then calculated with BOA for all parameters from the individual marginal posteriors using the free-software *R* (http://www.r-project.org/).

# Results

The values of DIC for models 1 to 4 were 3152.66, 2868.64, 2833.46, and 2835.12, respectively. Note that all models that included a tensor product of B-splines had a smaller DIC (i.e. better fits) than model 1 with block effects. Model 3 (12 × 12 knots) showed the smallest DIC, closely followed by model 4 (18 × 18 knots). The presence of spatial effects

could be observed in Figure 4, which displays the estimates of the block effects for model 1, or the estimated surface for models 2 to 4. There seems to be similarities in the locations of the 'hills' and 'valleys' in all four graphs. The fit for model 1 is expectedly abrupt as block effects are parameters for a categorical variable. On the other hand, the estimated surfaces with models 2 to 4 show that the degree of smoothness increases with the increase in the number of knots from 8 to 18.

[Insert **Figure 4** about here]

Posterior statistics for $\sigma^2_A$, $\sigma^2_b$, $\sigma^2_e$ and $h^2_{DBH}$ are shown in Table 1. Posterior means, medians and modes of the variance components and $h^2_{DBH}$ were similar except for $\sigma^2_A$ from models 2 and 3 and $\sigma^2_e$ from model 1, where the modes were smaller than means and medians. Estimates of $\sigma^2_A$ and $\sigma^2_e$ were similar in models 2 to 4, and this resulted in similar posteriors means of $h^2_{DBH}$: 0.237, 0.261, and 0.256 for the models with 8, 12 and 18 knots, respectively. Conversely, the estimated posterior mean of $h^2_{DBH}$ from the model with blocks was sensibly smaller (0.08). Also, the estimate of $\sigma^2_b$ from model 2 (17.35) was smaller than the estimated values from models 3 (22.31) and 4 (21.76). The estimates of $\sigma^2_e$ from models 2 to 4 were about half the magnitude of the parameter estimate for model 1. This is due to the spatial variation not being completely accounted for by the blocking procedure in model 1. None of the 95% HPD for $\sigma^2_A$, $\sigma^2_b$, $\sigma^2_e$ and $h^2_{DBH}$ included 0, suggesting that no parameter is equal to zero. The standard errors indicate that all estimates were quite precise, though large numbers of samples were drawn to attain reasonable ESS (last column in Table 1).

[Insert **Table 1** about here]

The average accuracy of prediction of breeding values, calculated from model 3 (the one with the smallest DIC) was higher for parents (0.61) and progeny (0.54), than corresponding values (0.40 and 0.32) calculated from model 1. Thus, fitting a surface using B-splines resulted in a gain in accuracy of 66 % for parents and 60% for offspring, a result which is due to the larger value of $h^2_{DBH}$ estimated in the model with B-splines. The Spearman correlation between predicted breeding values from models 1 and 3 was equal to 0.97 for parents and 0.94 for offspring, indicating that some re-ranking took place between the individuals with the least information, i.e. the progenies.

# Discussion

Unaccounted spatial variability in forest genetic trials leads to bias in estimating genetic parameters and predicting breeding values (Magnussen 1993, 1994), so that accuracy of selection decreases, thus reducing genetic gain. In the current research, we showed how to fit a two-dimensional surface using the tensor product of B-splines bases by means of a mixed model, in the spirit of Eilers and Marx (1996, 2003). P-splines in two dimensions have also been obtained by a Bayesian approach, as shown by Lang and Brezger (2004). These authors regarded the difference matrices [3] as a first or a second order random walk, respectively. Our approach is different from theirs in the replacing of the singular matrix of the differences [3] by a proper variance-covariance matrix of the random coefficients for the knot effects in two dimensions. In doing so, we extend the tensor product of B-spline bases to an individual tree mixed model to account for continuous spatial variability. Thus, the model incorporates a surface that is smoothed in the direction of both columns and rows. Gilmour at al. (1997) modeled the large scale

403 variation in one dimension of agricultural trials by fitting either polynomials or a cubic
404 smoothing spline. However, in forest genetic trials where trees are planted in squares or
405 rectangles, a large portion of the continuous spatial variation is usually present in the two
406 dimensions. Moreover, it is extremely rare that continuous spatial variability is found only
407 in the direction of the rows or of the columns, and some sort of interaction between rows
408 and columns has to be considered in order to account for such variability (Federer, 1998).
409 Although there exist several statistical methods of smoothing to capture non linearity of the
410 variation in one dimension, methods in two dimensions are less abundant. For such a
411 purpose, Federer (1998) proposed fitting interactions between polynomials for rows and
412 columns. However, polynomials do a poor job when fitting observations in the extremes.
413 Moreover, small changes in the data produce a dramatic effect in the estimated values of
414 the coefficients, and this is specially so for polynomials of higher degree. Additionally, the
415 degree of the polynomial should be selected, which in turn introduces the issue of model
416 selection. Instead, we propose estimating a smoothed surface using penalized splines. The
417 approach is flexible as B-spline functions are locally sensitive to the data and are
418 numerically well conditioned. The variance $\sigma_b^2$ was used to smooth the effects of both rows
419 and columns. In the approach of Eilers and Marx (2003) and Lang and Brezger (2004),
420 different variances for rows and columns were used. Lang and Brezger (2004) went further
421 and used a locally adaptive estimator of the dispersion parameters. In future research, we
422 may consider smoothing rows and columns with different dispersion parameters, although
423 it is not clear to us that this approach may be more advantageous than ours regarding the
424 quality of the fit, i.e. the value of the DIC.

425 The P-splines methodology of Eilers and Marx (1996, 2003) consists of using cubic
426 B-splines with equally spaced knots. In this approach, the crucial parameter is the penalty
427 or smoothing factor λ (see [2] and [5]), and the number of knots in the spline is not vital to
428 the fit as long as there are "sufficiently" many (Eilers and Marx, 1996; Cantet et al. 2005).
429 In the mixed model approach to P-splines, λ is the ratio $\sigma_e^2/\sigma_b^2$ (Cantet et al. 2005) in [8].
430 Looking at Table 1 one may infer that the magnitude of $\sigma_b^2$ (the denominator of λ) was
431 sensitive to the number of knots, as compared to the other variance components. It is
432 known that the fit of very few knots produces bias, which rapidly decreases as the number
433 of knots increases (Ruppert 2002). Cantet et al. (2005) found almost equal values of the
434 modified Akaike Information criterion for models with 20, 40, 60, 80, or 120 equally
435 spaced knots. However, Restricted Maximum Likelihood estimators for the variance
436 components did not converge for certain models with 120 knots. For those situations with
437 120 knots where convergence was attained, there were some inconsistencies in the fit for
438 intervals where no data was recorded. In the current research, increasing the number of
439 knots from 8 to 18 produced a smoother surface (Figure 4). Although the model with 12 ×
440 12 knots displayed the smallest DIC, the difference in DIC between the models with 12 ×
441 12 and 18 × 18 knots was minor. This was also true for the estimates of $h^2_{\mathrm{DBH}}$ obtained
442 from both models: a difference in the third decimal place. In the mixed model approach to
443 P-splines, the covariance structure of the knot coefficients replaces any of the singular
444 matrices of differences [3]. In the current research, the tridiagonal matrix proposed by
445 Durban et al. (2001) is used to model the correlations between the knots for columns and
446 for rows. The formulation is simpler than the dense correlation structures used by Hyndman
447 et al. (2005) and Cantet et al. (2005), where there is complete dependence among all knot

448 effects. In order to check the impact of the covariance matrix on the fit, we adjusted three
449 models with 12 × 12 knots differing only in the covariance matrix of knot effects, and run
450 30 000 Gibbs samples. The values of DIC obtained were 2882.33, 2871.58, and 2850.97,
451 for the structures used by Cantet et al. (2005), Hyndman et al. (2005), and Durban et al
452 (2001), respectively, which supports the use of the latter structure for the current data set.
453 There are several examples of the use of B-spline functions in one dimension when
454 analyzing breeding data. Thus, animal breeders used splines to model functional breeding
455 values (White et al. 1999; Bohmanova et al. 2005) or the effects of management unit and
456 time (Cantet et al. 2005). In forest genetic breeding, Cornillon et al. (2003) used B-splines
457 to model time functional breeding values of clones in Eucalyptus using a fixed effects
458 model. Magnussen and Yanchuk (1994) fitted spline functions to observed data so as to
459 estimate the individual heights at non-recorded times from Douglas-fir trees. The resulting
460 data was then used to predict breeding values at non-recorded ages and genetic dispersion
461 parameters. The fit of a smoothed surface to the progeny trial in *E. globulus* with tensor
462 product of B-splines instead of the '*a priori*' block design, consistently increased the
463 posterior means of $\sigma^2_A$ and of $h^2_{DBH}$ (Table 2). The results agree with those of Zas (2006)
464 that accounted for spatial variability using Kriging, and are different from those of
465 Dutkowsky et al. (2002). In the latter case, inconsistent estimates of $\sigma^2_A$ were obtained after
466 adjusting an AR(1) × AR(1) covariance structure to the residuals of the model. In our data,
467 the spatial models produced an increase in precision for the estimation of $\sigma^2_e$, which can be
468 noticed by the much lower standard deviations and the narrower values for the 95% high
469 posterior probability density intervals, when compared to the estimate from the model with
470 blocks (Table 1). Moreover, accuracy of breeding values from parents and offspring

calculated with the spatial models were higher than corresponding values estimated from the model with block effects. This result agrees with those of Costa e Silva et al. (2001) for tree height and Zas (2006) for tree diameter. Our results suggest that analysis of data displaying large scale continuous spatial variation, such as the one induced by a petrocalcic layer at variable depth, could hardly be accounted for by blocking techniques.

In the current research, we modeled spatial variability that is continuous and permanent along a site, using an individual tree mixed model with a smoothed surface. In forest genetic evaluation, the spatial variation at the microsite level has been modeled with nearest neighbor techniques (Magnussen 1990; Costa e Silva et al. 2001; Dutkowski et al. 2002, 2006) or with Kriging (Hamann et al. 2002; Zas 2006). Nevertheless, interplant competition may be another source for small scale spatial variation which affects the correlation between neighbors (Magnussen 1994). The mixed model [6] does not account for genetic competition among trees, and this can bias the estimation of $\sigma_A^2$ (Cappa and Cantet 2006b). However, the trees used in the analysis were 6 yr-old, so that competition was weak or absent. For those situations where trees are measured at an age where competition effects are sizeable, it would be desirable to fit simultaneously continuous spatial variation and genetic effects of competition.

## Acknowledgements

This research was supported by grants of Agencia Nacional de Ciencia y Tecnología (FONCyT PICT 08-09502); Universidad de Buenos Aires (UBACyT G018, 2004-2007), and CONICET (PIP 5338, 2005-2006) of Argentina. The authors would like to thank to

493 INTA-Soporcel (Instituto Nacional de Tecnología Agropecuaria – Sociedad Portuguesa de
494 Papel S.A., Portugal) for allowing us to use their data.
495

# References

496 Anderson, T.W. 1973 . Asymptotically efficient estimation of covariance matrices with
497 linear structure. Ann. Statist. **1**: 135-141.

498 Bohmanova, J., Misztal, I., Bertrand, J.K. 2005. Studies on multiple trait and random
499 regression models for genetic evaluation of beef cattle for growth. J. Anim. Sci. **83**: 62-
500 67.

501 Cantet R.J.C., Birchmeier A.N., Canaza Cayo A.W., and Fiorett, C. 2005. Semiparametric
502 animal models via penalized splines as alternatives to models with contemporary
503 groups. J. Anim. Sci. **83**: 2482–2494.

504 Cappa, E.P., and Cantet, R.J.C. 2006a. Bayesian inference for normal multiple-trait
505 individual-tree models with missing records via full conjugate Gibbs. Can. J. For. Res.
506 **36**: 1276-1285.

507 Cappa, E.P., and Cantet, R.J.C. 2006b. Direct and competition additive effects in tree
508 breeding: Bayesian estimation from an individual tree mixed model. Silvae Genetica
509 (submitted).

510 Christensen, R. 1987. Plane Answer to Complex Questions. The Theory of Linear Models,
511 New York: Springer Verlag.

512 Costa e Silva, J., Dutkowski, G.W., Gilmour, A.R. 2001. Analysis of early tree height in
513 forest genetic trials is enhanced by including a spatially correlated residual. Can. J. For.
514 Res. **31**: 1887-1893.

515 Cornillon, P.A., Saint-Andre, L., Bouvet, J.M., Vigneron, P., Saya, A., and Gouma R.
516 2003. Using B-splines for growth curve classification: applications to selection of
517 eucalypt clones. Forest Ecology and Management, **176**: 75-85.
518 De Boor, C. 1993. B(asic)-spline basics. Fundamental Developments of Computer-Aided
519 Geometric Modeling. L. Piegl, ed. Academic Press, San Diego, CA.
520 Durban, M., Currie I., and Kempton, R. 2001. Adjusting for fertility and competition in
521 variety trials. J. Agric. Sci. (Camb.) **136**: 129–140.
522 Dutkowski, G.W., Costa e Silva, J., Gilmour, A.R., and Lopez, G.A. 2002. Spatial analysis
523 methods for forest genetic trials. Can. J. For. Res. **32**: 2201-2214.
524 Dutkowski, G.W., Costa e Silva J., Gilmour A.R., Wellendorf H., Aguiar A. 2006. Spatial
525 analysis enhances modeling of a wide variety of traits in forest genetic trials. Can. J.
526 For. Res. **36**: 1851-1870.
527 Eilers, P.H.C., and Marx, B.D. 1996. Flexible smoothing with B-splines and penalties (with
528 comments and rejoinder). Stat. Sci. **11**: 89–121.
529 Eilers, P.H.C., and Marx, B.D. 2003. Multivariate calibration with temperature interaction
530 using two-dimensional penalized signal regression. Chemometr. Intell. Lab. Syst. **66**:
531 159-174.
532 Federer, W.T. 1998. Recovery of interblock, intergradient, and intervarietal information in
533 incomplete block and lattice rectangle designed experiments. Biometrics, **54**: 471-481.
534 Fu, Y.B., Yanchuk, A.D., and Namkoong, G. 1999a. Incomplete block designs for genetic
535 testing: some practical considerations. Can. J. For. Res. **29**: 1871-1878.
536 Fu, Y.B., Yanchuk, A.D., and Namkoong, G. 1999b. Spatial patterns of tree height
537 variations in a series of Douglas-fir progeny trials: implications for genetic testing. Can.
538 J. For. Res. **29**: 714-723.

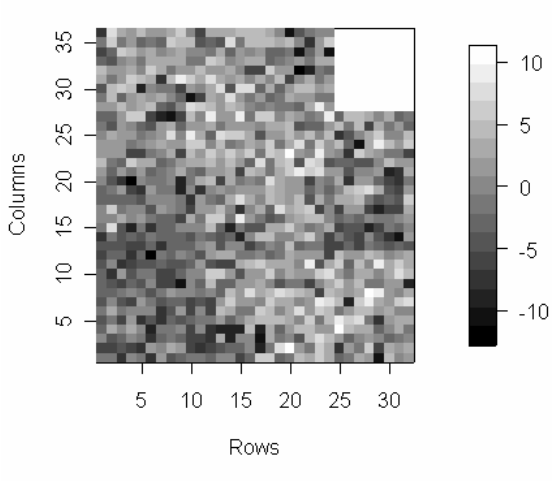
539 Gilmour A.R., Cullis B.R., and Verbyla A.P. 1997. Accounting for Natural and Extraneous
540 Variation in the Analysis of Field Experiments. Journal of Agricultural, Biological, and
541 Environmental Statistics, **2**: 269-293.
542 Green, P. J., and Silverman, B. W. 1994. Nonparametric Regression and Generalized
543 Linear Model. Chapman & Hall, London, UK.
544 Gwaze, D.P., and Woolliams, J.A. 2001. Making decisions about the optimal selection
545 environment using Gibbs sampling. Theor. Appl. Genet. **103**: 63–69.
546 Hamann, A., Koshy, M., and Namkoong, G. 2002. Improving precision of breeding values
547 by removing spatially autocorrelated variation in forestry field experiments. Silvae
548 Genetica, **51**: 210-215.
549 Harville, D.A. 1997. Matrix algebra from a statistician's perspective. Springer-Verlag. New
550 York.
551 Henderson, C.R. 1984. Applications of Linear Models in Animal Breeding. Canada,
552 University of Guelph, Guelph, Ont.
553 Hobert, J.P., and Casella, G. 1996. The effects of improper priors on Gibbs sampling in
554 hierarchical linear models, J. Amer. Statist. **91**: 1461-1473.
555 Hyndman, R.J., King, M.L., Pitrun, I., and Billah, B. 2005. Local lineal forecasts using
556 cubic smoothing splines. Aust. N. Z. J. Stat. **47**: 87–99.
557 Kass, R.E., Carlin, B.P., Gelman, A., and Neal, R.M. 1998. Markov chain Monte Carlo in
558 practice: a roundtable discussion. Amer. Stat. **52**: 93-100.
559 Lang, S., Brezger, A. 2004. Bayesian P-Splines. Journal of Computational & Graphical
560 Statistics, **13**: 183-212.
561 Libby, W.J., Cockerham, C.C. 1980. Random non-contiguous plots in interlocking field
562 layouts. Silvae Genetica, **29**: 183-190.

563 Liu, J., and Burkhart, H.E. 1994. Spatial characteristics of diameter and total height in
564 juvenile loblolly pine (*Pinus taeda* L.) plantations. For. Sci. **40**: 774-786.
565 Lopez, G.A., Potts, B.M., Dutkowski, G.W., Rodriguez Traverso, J.M. 2001. Quantitative
566 genetics of Eucalyptus globulus: Affinities of land race and native stand localities.
567 Silvae Genetica, **50**: 244-252.
568 Magnussen, S. 1990. Application and comparison of spatial models in analysing tree-
569 genetics field trials. Can. J. For. Res. **20**: 536-546.
570 Magnussen, S. 1993. Bias in genetic variance estimates due to spatial autocorrelation.
571 Theor. Appl. Genet. **86**: 349-355.
572 Magnussen, S. 1994. A method to adjust simultaneously for statial microsite and
573 competition effects. Can. J. For. Res. **24**: 985-995.
574 Magnussen, S., and Yanchuk, A.D. 1994. Time trends of predicted breeding values in
575 selected crosses of coastal Douglas-fir in British Columbia: a methodological study.
576 For. Sci. **40**: 663-685.
577 Mark, B.D., and Eilers, P.H. 2005. Multidimensional penalized signal regression.
578 Technometrics, **47**: 13-22.
579 Robinson, G.K. 1991. That BLUP is a good thing: The estimation of random effects.
580 Statistical Science, **6**: 15-32.
581 Ruppert, D. 2002. Selecting the number of knots for penalized splines. Journal of
582 Computational and Graphical Statistics, **11**: 735–757.
583 Ruppert, D., Wand, M.P., and Carroll, R.J. 2003. Semiparametric Regression. Cambridge
584 Univ. Press, Cambridge, UK.

585 Saenz-Romero, C., Nordheim, E.V., Guries, R.P., and Crump, P.M. 2001. A Case Study of
586 a Provenance/Progeny Test Using Trend Analysis with Correlated Errors and SAS
587 PROC MIXED. Silvae Genetica, **50**: 127 – 135.
588 Silverman, B. 1986. Density Estimation for Statistics and Data Analysis, Chapman and
589 Hall, London.
590 Smith, B.J. 2003. Bayesian Output Analysis Program (BOA) version 1.0 user's manual.
591 Available from http://www.public-health.uiowa.edu/boa/Home.html.
592 Soria F., Basurco, F., Toval, G., Silió, L., Rodriguez, M.C., and Toro, M. 1998. An
593 application of Bayesian techniques to the genetic evaluation of growth traits in
594 Eucalyptus globulus. Can. J. For. Res. **28**: 12861294.
595 Sorensen, D., and Gianola, D. 2002. Likelihood, Bayesian, and MCMC Methods in
596 Quantitative Genetics. Springer-Verlag, New York.
597 Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Van der Linde, A. 2002. Bayesian
598 measures of model complexity and fit (with discussion). Journal of the Royal Statistical
599 Society Series B, **64**: 583-639.
600 Thomson, A.J. and El-Kassaby Y.A. 1988. Trend surface analysis of a Douglas-fir
601 provenace-progeny transfer test. Can. J. For. Res. **18**: 515-520.
602 Verbyla, A.P., Cullis, B.R., Kenward, M.G. and Welham, S.J. 1999. The analysis of
603 designed experiments and longitudinal data by using smoothing splines (with
604 discussion). Applied Statistics, **48**: 69-311.
605 Waldmann, P., and Ericsson, T. 2006. Comparison of REML and Gibbs sampling estimates
606 of multi-trait genetic parameters in Scots pine. Theor. Appl. Genet. **112**: 1441-1451.
607 Wand, M.P. 2003. Smoothing and mixed models. Comput. Stat. **18**: 223–249.

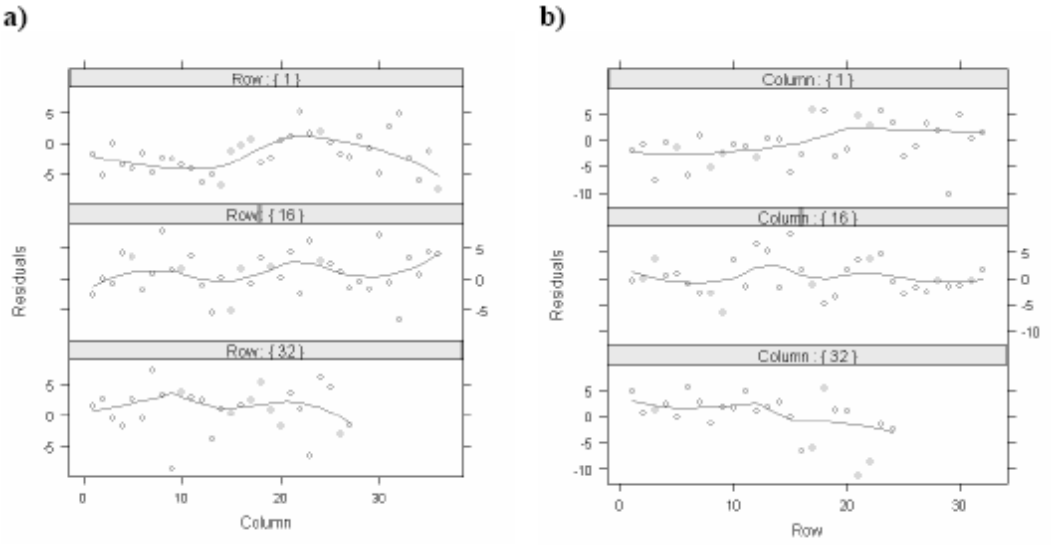
608 Zas, R. 2006. Iterative kriging for removing spatial autocorrelation in analysis of forest
609 genetic trials. Tree Genetics and Genomes, **2**: 177-185.
610 Zeng, W., Ghosh, S.K., Li, B. 2004. A Blocking Gibbs Sampling Method to Detect Major
611 Genes Affecting a Quantitative Trait for Diallel Mating Design. Genetical Research **84**:
612 1-12.
613 White, I.M.S., Thompson, R., Brotherstone, S. 1999. Genetic and environmental smoothing
614 of lactation curves with cubic splines. J. Dairy Sci. **82**: 632–638.

615 **Figure 1:** Spatial patterns of the residuals of tree DBH. The colors of the dots represent the
616 magnitude of the residuals: the blacker the dot, the bigger the residual.



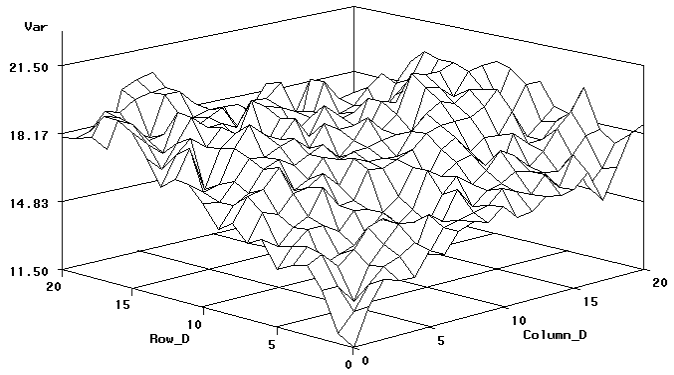

617

618 **Figure 2:** Plot of the residuals after fitting provenance and additive genetic effects: **a**)
619 number of column for different rows and **b**) number of rows for different columns.



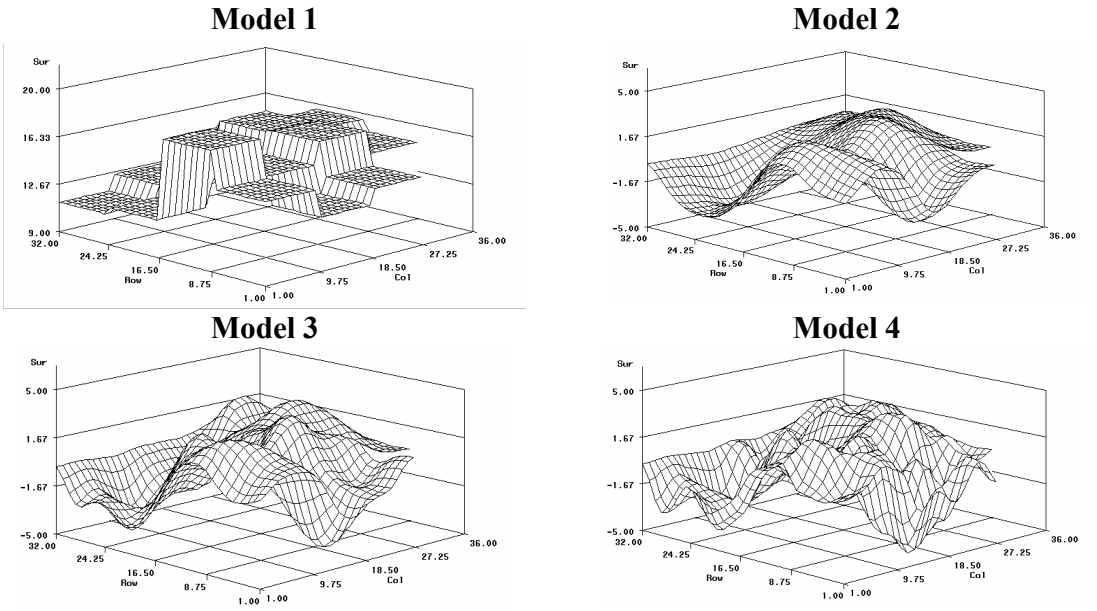

620

621 **Figure 3:** Sample variogram showing the interaction between rows and columns.




622

623 **Figure 4:** Plot of the estimates of block effects (Model 1) and the surfaces from the fitting
624 of tensor product B-splines with either 8 (Model 2), 12 (Model 3), or 18 (Model 4) knots.




625

626 **Table 1**: Posterior statistics for the additive genetic variance ($\sigma^2_A$), the variance of the B-
627 spline coefficients ($\sigma^2_b$), the error variance ($\sigma^2_e$), and the heritability of DBH ($h^2_{DBH}$).

| Model[a] | Parm.[b] | Mean | Median | Mode | SD[c] | 95% HPD[d] | ESS[e] |
|---|---|---|---|---|---|---|---|
| **1** | $\sigma^2_A$ | 1.835 | 1.801 | 1.609 | 0.37149 | 1.291 – 2.503 | 24119 |
| | $\sigma^2_e$ | 23.043 | 20.144 | 14.070 | 8.69251 | 15.182 – 40.520 | 87274 |
| | $h^2_{DBH}$ | 0.080 | 0.079 | 0.084 | 0.02520 | 0.040 – 0.123 | 43572 |
| **2** | $\sigma^2_A$ | 3.596 | 3.480 | 2.642 | 0.98973 | 2.191 - 5.381 | 16181 |
| | $\sigma^2_b$ | 17.351 | 16.558 | 16.875 | 5.17173 | 10.457 - 26.887 | 169158 |
| | $\sigma^2_e$ | 11.156 | 11.191 | 10.476 | 1.01469 | 9.432 - 12.760 | 24207 |
| | $h^2_{DBH}$ | 0.243 | 0.237 | 0.259 | 0.06401 | 0.151 - 0.358 | 16254 |
| **3** | $\sigma^2_A$ | 3.754 | 3.643 | 2.933 | 1.00390 | 2.310 – 5.573 | 16474 |
| | $\sigma^2_b$ | 22.317 | 21.649 | 23.716 | 5.47972 | 14.682 – 32.132 | 109973 |
| | $\sigma^2_e$ | 10.275 | 10.301 | 9.900 | 1.01309 | 8.558 - 11.871 | 23568 |
| | $h^2_{DBH}$ | 0.267 | 0.261 | 0.244 | 0.06872 | 0.167 - 0.389 | 16519 |
| **4** | $\sigma^2_A$ | 3.661 | 3.558 | 3.439 | 0.98475 | 2.254 - 5.458 | 16526 |
| | $\sigma^2_b$ | 21.758 | 21.409 | 18.998 | 4.17318 | 15.463 - 29.223 | 81522 |
| | $\sigma^2_e$ | 10.312 | 10.339 | 9.683 | 1.00670 | 8.595 - 11.920 | 24305 |
| | $h^2_{DBH}$ | 0.262 | 0.256 | 0.205 | 0.06706 | 0.164 - 0.383 | 16588 |

628 **Note:**
629 [a] **Model 1**: blocks fitted as fixed effects.
630 **Model 2**: P-splines with 8 knots for rows and 8 knots for columns.
631 **Model 3**: P-splines with 12 knots for rows and 12 knots for columns.
632 **Model 4**: P-splines with 18 knots for rows and 18 knots for columns.
633 [b] **Parm.** = Parameter.
634 [c] **SD** = standard deviation.
635 [d] **HPD =** high posterior density interval.
636 [e] **ESS =** effective sample size.