

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS RURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA**

**GANHO GENÉTICO PARA PRODUTIVIDADE DE
GRÃOS DE MILHO NA REGIÃO SUL DO BRASIL**

DISSERTAÇÃO DE MESTRADO

Éder David Borges da Silva

Santa Maria, RS, Brasil.

2015

GANHO GENÉTICO PARA PRODUTIVIDADE DE GRÃOS DE MILHO NA REGIÃO SUL DO BRASIL

Éder David Borges da Silva

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Agronomia, Área de Concentração em Produção Vegetal, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Agronomia.**

Orientador: Prof. Dr. Alberto Cargnelutti Filho

Santa Maria, RS, Brasil.

2015

Ficha catalográfica elaborada através do Programa de Geração Automática da Biblioteca Central da UFSM, com os dados fornecidos pelo(a) autor(a).

Silva, Éder David Borges da
Ganho genético para produtividade de grãos de milho na região sul do Brasil / Éder David Borges da Silva.- 2015.
61 p.; 30cm

Orientador: Alberto Cargnelutti Filho
Dissertação (mestrado) - Universidade Federal de Santa Maria, Centro de Ciências Rurais, Programa de Pós-Graduação em Agronomia, RS, 2015

1. Zea mays L. 2. Modelos mistos 3. Mínimos quadrados
I. Cargnelutti Filho, Alberto II. Título.

© 2015

Todos os direitos autorais reservados a Éder David Borges da Silva. A reprodução de partes ou do todo deste trabalho só poderá ser feita mediante a citação da fonte.

E-mail: ederdb@gmail.com

Universidade Federal de Santa Maria
Centro de Ciências Rurais
Programa de Pós-Graduação em Agronomia

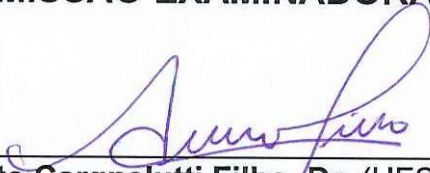
A Comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**GANHO GENÉTICO PARA PRODUTIVIDADE DE GRÃOS DE
MILHO NA REGIÃO SUL DO BRASIL**

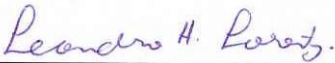
elaborada por
Éder David Borges da Silva

como requisito parcial para obtenção do grau de
Mestre em Agronomia

COMISSÃO EXAMINADORA:



Alberto Cargnelutti Filho, Dr. (UFSM)
(Presidente/Orientador)



Leandro Homrich Lorentz, Dr. (UNIPAMPA)



Lindolfo Storck, Dr. (UFSM)

Santa Maria, 11 de fevereiro de 2015

Aos meus Pais, **Elecir e Deoclides**,
E as minhas irmãs, **Deonete, Jania, Solange e Marcia**,
Dedico este trabalho!

AGRADECIMENTOS

À Universidade Federal de Santa Maria e ao Programa de Pós-Graduação em Agronomia, pela oportunidade de realização deste trabalho.

Aos meus pais, Elecir Azeredo da Silva (*in memoriam*) e Deoclides Borges da Silva.

As minhas irmãs: Deonete, Jania, Solange e Marcia, e minhas sobrinhas Tainá e Fernanda, por todo o apoio.

Ao meu orientador Alberto Cargnelutti Filho, pela dedicação e orientação deste trabalho, e pela compreensão durante este período.

Aos professores do grupo de pesquisa em experimentação, Alessandro Dal'Col Lúcio e Lindolfo Storck, pelo apoio e ajuda na realização deste trabalho.

Aos colegas que conheci durante este período, em especial, Bruna Alves, Cláudia Burin, Fernando Haesbaert, Riteli Mambri, Giovani Facco e Marcos Toebe.

À Dupont do Brasil – Divisão Pioneer Sementes, pela oportunidade e pela cessão dos dados para a realização deste trabalho, em especial na pessoa de Itavor Nummer Filho pelo incentivo neste trabalho e aos colaboradores que conduziram estes ensaios.

Aos demais colegas da Dupont Pioneer, que me apoiaram, em especial Rogério Ferreira, Roger Butzge, Fabricio Passini, Jose Vecchio, Robson de Paula, Ricardo Zottis, Henrique Maldaner e Vanderson Costa, pelo companheirismo.

Aos colegas do Laboratório de estatística e Geoinformação (LEG) da UFPR, em especial Paulo Justiniano Jr, Walmes Veziani, Wagner Bonat, e Silvia Shimakura, Renato Oliveira, que sempre me apoiaram/incentivaram no estudo da estatística.

A todos os amigos, colegas e demais familiares, que contribuíram de forma direta ou indiretamente para a realização desta dissertação.

Meu sincero muito obrigado!

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Agronomia
Universidade Federal de Santa Maria

GANHO GENÉTICO PARA PRODUTIVIDADE DE GRÃOS DE MILHO NA REGIÃO SUL DO BRASIL

AUTOR : ÉDER DAVID BORGES DA SILVA
ORIENTADOR: ALBERTO CARGNELUTTI FILHO
Data e Local da Defesa: Santa Maria, 11 de fevereiro de 2015.

Os objetivos deste trabalho foram calcular o ganho genético de produtividade de grãos de milho na região Sul do Brasil e calcular o ganho genético em duas classes de altitude. As metodologias utilizadas foram a de Vencovsky et al. (1988), baseada em modelos lineares com inferência por mínimos quadrados (ML/MQ) e a de Borges et al. (2009), baseada em modelos mistos com inferência por verossimilhança residual (MM/REML). Foi utilizada uma base de dados com 30.292 observações de produtividade de grãos de milho, resultantes de 135 genótipos avaliados em 2.826 ensaios conduzidos em 13 anos, de 2001 a 2013, sendo o ano referente à data de semeadura do ensaio. Para a estratificação dos ensaios, com o objetivo de formar classes de altitude, utilizou-se como critério da altitude, inferior e superior a 700m. Pela metodologia de ML/MQ o ganho genético médio anual foi de 121 kg ha⁻¹ ano⁻¹ com intervalo de confiança a 95% de [11;232] kg ha⁻¹ ano⁻¹, e pela metodologia de MM/REML foi de 79 kg ha⁻¹ ano⁻¹ com intervalo de confiança a 95% de [70;98] kg ha⁻¹ ano⁻¹, para a região sul do Brasil. Em altitude superior a 700 m, o ganho genético foi de 94 kg ha⁻¹ ano⁻¹ e em altitude inferior a 700 m o ganho genético foi de 74 kg ha⁻¹ ano⁻¹. A metodologia MM/REML proporcionou intervalos de confiança mais precisos e valores de ganho genético inferiores quando comparada à metodologia de ML/MQ. Em todas as situações analisadas o ganho genético foi positivo, e significativo entre os anos de 2001 a 2013.

Palavras-chave: *Zea mays* L. Modelos mistos. Mínimos quadrados.

ABSTRACT

Master Dissertation
Programa de Pós-graduação em Agronomia
Universidade Federal de Santa Maria

GENETIC GAIN FOR CORN GRAIN YIELD IN SOUTHERN BRAZIL

AUTHOR: ÉDER DAVID BORGES DA SILVA

ADVISOR: ALBERTO CARGNELUTTI FILHO

Local and Date: Santa Maria/RS, February, 11, 2015.

The objectives of this study was to calculate the genetic gain corn grain yield in southern Brazil and calculate the genetic gain in two classes of altitude. The methods used were the Vencovsky et al. (1988) based on linear models with inference least squares (ML/MQ) and the Borges et al. (2009), based on mixed models with inference by residual likelihood (MM/REML). We used a database with 30,292 yield observation from corn grain but resulting in 135 genotypes in 2,826 trials conducted in 13 years, from 2001 to 2013, being the year regarding the planting date of the trials. For the stratification of trials, in order to form similar environments, we used as criteria the altitude, down and up to 700m. The methodology of ML/MQ annual mean genetic gain was $121 \text{ kg ha}^{-1} \text{ year}^{-1}$ with a confidence interval to 95% [11;232] $\text{kg ha}^{-1} \text{ year}^{-1}$, and by MM/REML methodology was $79 \text{ kg ha}^{-1} \text{ year}^{-1}$ with a confidence interval to 95% [70; 98] $\text{kg ha}^{-1} \text{ year}^{-1}$ for the entire southern region of Brazil. In regions with altitude of 700 m, the genetic gain was $94 \text{ kg ha}^{-1} \text{ year}^{-1}$ and in regions with altitude less than 700 mo genetic gain was $74 \text{ kg ha}^{-1} \text{ year}^{-1}$.MM/REML methodology provided intervals more accurate and reliable values of genetic gain lower when compared to ML/MQ methodology. In all cases analyzed, the genetic gain was positive and significant between the years 2001-2013.

Keywords: *Zea mays*. *Mixed Models*. *Least Squares*.

LISTA DE TABELAS

- Tabela 1 - Análise descritiva da produtividade de grãos de milho na região Sul do Brasil, e das covariáveis nitrogênio (N), fósforo (P) e potássio (K) e altitude representadas pelo valor médio e o número de locais e número de observações por ano. 29
- Tabela 2 - Comportamento dos genótipos de milho nos ensaios de avanço de produto ao longo dos biênios, sendo C o número de genótipos comuns aos dois anos do biênio, n_{c1} e n_{c2} , sendo respectivamente o número de genótipos presentes no ano 1 e 2, TM a taxa de manutenção dos genótipos, dada por $[C/(n_{c1}+n_{c2}-C)].100$, TS taxa de substituição dos genótipos dada por $[(n_{c2}-C)/n_{c2}].100$, N_1 e N_2 o número de observações no ano 1 e 2 respectivamente. 31
- Tabela 3 - Resultados das análises de variâncias realizadas para estimativa do modelo de Vencovsky et al. (1988), sendo QM_{e1} quadrado médio do erro do ano 1, GL_{e1} grau de liberdade do erro no ano 1, SW_1 valor-p do teste de Shapiro Wilk para normalidade do erro no ano 1, $Bartlett_1$ é o valor-p do teste de Bartlett para homogeneidade do erro do ano 1, QM_{e2} quadrado médio do erro do ano 2, GL_{e2} grau de liberdade do erro no ano 2, SW_2 valor-p do teste de Shapiro Wilk para normalidade do erro no ano 2, $Bartlett_2$ é o valor-p do teste de Bartlett para homogeneidade do erro do ano 2. 32
- Tabela 4 - Resultados gerais do modelos de Vencovsky et al, (1988), sendo Y_1 média geral do ano 1, Y_2 média geral do ano 2, Y_{C1} média geral dos genótipos comuns aos dois anos no ano 1, Y_{C2} média geral dos genótipos comuns aos dois anos no ano 2, G_g ganho genético, V_{Cg} variância do ganho genético, t_{cal} valor calculado do teste t, sob hipótese de ganho diferente de 0, valor-p é valor-p do teste t, $IC_{(95\%)}$ é o intervalo de confiança para o ganho genético. 33
- Tabela 5 - Resultados dos componentes de variâncias, e dos efeitos fixos, do modelo misto ajustado, sendo θ , o parâmetro do modelo, e $\hat{\theta}$, a estimativa obtida para o parâmetro via REML, e σ o desvio padrão da estimativa do parâmetro. 35
- Tabela 6 - Resultado do ganho via modelos mistos, sendo u_1 e u_2 , valor genético do ano 1 e 2 respectivamente, V_{u1} e V_{u2} variância do ganho genético no ano 1 e 2 respectivamente, G_g ganho genético, V_{Gg} variância do ganho genético, valor-p é o valor-p sob o teste da diferença do ganho em relação a 0, $IC_{(95\%)}$ intervalo de confiança do ganho genético com 95% de probabilidade. 35
- Tabela 7 - Dados descritivos de médias de produtividade(x), e número de locais (NL) estratificando em duas classes de altitude superior e inferior a 700m ao longo dos anos de estudo. 46
- Tabela 8 - Componentes de variância dos modelos ajustado para cada condição de estratificação, sendo θ o parâmetro, $\hat{\theta}$ o valor estimado para o parâmetro, σ o desvio padrão do parâmetro, $\% \sigma^2$ percentual total da variância e valor-p para o teste de significância, GL o numero de graus de liberdade aproximados por Satterthwaite, t o valor t calculo, e valor-p sob hipótese de H_0 , onde a estimativa seria igual a 0. 47
- Tabela 9 - Resultado do ganho via modelos mistos, em diferentes classes, sendo u_1 e u_2 , ganho genético do ano 1 e 2 respectivamente, V_{u1} e V_{u2} variância do ganho genético no ano 1 e 2 respectivamente, G_g ganho genético, V_{Gg} variância do ganho genético, valor-p é o valor-p sob o teste da diferença do ganho em relação a 0, IC intervalo de confiança a 95% de probabilidade do ganho genético. 48

LISTA DE FIGURAS

Figura 1 - Croqui de um ensaio de avanço de produto.....	24
Figura 2 - Distribuição espacial dos ensaios de milho de avanço de produto. Cada círculo preto representa um local de realização do ensaio, entre os anos de 2001 a 2013, na região Sul do Brasil	25
Figura 3 - Croqui de um ensaio de avanço de produto.....	41
Figura 4 - Distribuição espacial dos ensaios de milho de avanço de produto. Realizados entre os anos de 2001 a 2013, na região sul do Brasil, pontos cinza forte indicam ensaios em altitude inferior a 700m e pontos cinza claro indicam ensaios em altitude superior a 700m.	42
Figura 5 - Comportamento da produtividade de grãos de milho na região sul do Brasil, em função da altitude que os ensaios foram conduzidos.....	45

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Métodos de inferência estatística.....	13
1.2 Modelos para determinação de ganho genético	17
2 CAPÍTULO 1 - GANHO GENÉTICO PARA A PRODUTIVIDADE DE GRÃOS DE MILHO ESTIMADO POR MÍNIMOS QUADRADOS E VEROSSIMILHANÇA.....	20
2.1 Resumo.....	20
2.2 Introdução	20
2.3 Material e métodos	23
2.3.1 Metodologia dos ensaios de avanço de produto.....	23
2.3.2 Dados	24
2.3.3 Metodologias	25
2.4 Resultados e discussão	28
2.4.1 Análises descritivas	28
2.4.2 Resultados da predição pela metodologia de Vencovsky et al (1988).....	30
2.4.3 Resultados da predição pela metodologia de Borges et al (2009).....	33
2.5 Conclusão.....	37
3 CAPÍTULO 2 – GANHO GENÉTICO PARA A PRODUTIVIDADE DE GRÃOS DE MILHO EM DUAS CLASSES DE ALTITUDE.....	38
3.1 Resumo.....	38
3.2 Introdução	38
3.3 Material e métodos	40
3.3.1 Metodologia dos ensaios de avanço de produto.....	40
3.3.2 Dados	41
3.4 Resultados e discussão	44
3.4.1 Análise exploratória	44
3.4.2 Resultado das predições	46
3.5 Conclusão.....	49
4 DISCUSSÃO	50
REFERÊNCIAS.....	51
ANEXO A - Código em R para a metodologia de Vencovsky et al. (1988).....	56
ANEXO B - Código em R para a metodologia de Borges et al. (2009).	59

1 INTRODUÇÃO

O milho (*Zea mays* L.) é uma planta originária da América Central, onde começou a sua domesticação, processo que durou milhares de anos, pois há relatos de que o milho já era conhecido pelos povos indígenas da América Central em 8.700 a.C. no México, de acordo com Ranere (2009).

O interesse do ser humano pelo milho é devido a sua excelente produção de energia para consumo humano, animal e industrial. Este fato fez o milho se tornar uma das plantas mais cultivadas pela humanidade, tendo, atualmente, uma produção mundial de 1,16 bilhões de toneladas, em uma área de 184 milhões de hectares, com produtividade média de grãos de 5.520 kg ha⁻¹, tendo como principais produtores os Estados Unidos, China, Brasil e Argentina, segundo dados de 2013 (FAO, 2013).

No Brasil, as condições são favoráveis ao cultivo do milho devido ao clima, bem como os intensos trabalhos de melhoramento genético, que buscaram adaptar e selecionar genótipos mais produtivos em condições brasileiras. Em 2013, no Brasil, foram produzidas 79,9 milhões de toneladas de grãos de milho, em área de 15,8 milhões de hectares, com produtividade média de grãos de 5.057 kg ha⁻¹ (CONAB, 2014)

O cultivo de híbridos simples de milho, geneticamente modificados com a incorporação de genes *Bt* (que gera tolerância a algumas espécies de lepidópteros), e mais recentemente os genes RR (que gera tolerância ao herbicida glifosato), tem dominado o mercado brasileiro de sementes de milho, apesar de híbridos triplos, duplos, simples modificados e cultivares de polinização aberta também participarem, em algumas regiões. Na safra 2013/14 existiam no Brasil 317 genótipos registrados disponíveis aos produtores, com predomínio de híbridos simples (56,15%), seguidos de híbridos triplos (18,61%), híbridos duplos (13,56%) e as variedades de polinização aberta (11,68%) segundo Cruz et al. (2014).

A pesquisa da cultura de milho, a nível mundial, está concentrada nas empresas privadas, que detém o processo de melhoramento genético, produção e comercialização de sementes de milho. No Brasil, a reutilização de sementes de milho pelos produtores é baixa, pois 90% da semente necessária é adquirida a cada ano (ABRASEM, 2014).

O foco do melhoramento genético na cultura de milho é a produtividade de grãos, a adaptação geral aos ambientes e a resistência aos fatores abióticos (BORÉM & MIRANDA, 2013) e em segundo plano a resistência às doenças. Diante da importância do ganho de produtividade de grãos na cultura de milho, a estimativa do ganho genético é uma forma de mensurar se os programas de melhoramento estão progredindo e lançando genótipos mais produtivos.

O ganho genético é definido como o incremento na característica melhorada em função da substituição por um novo genótipo. Em diferentes fases, durante o desenvolvimento de um genótipo de milho, é possível calcular o ganho genético, seja para a produtividade de grãos, ou para outras variáveis, como ciclo, altura de plantas e resistência a doenças. Normalmente, nas etapas finais do programa de melhoramento, há maior número de locais de avaliação, quando comparada às etapas preliminares, submetendo os genótipos a condições ambientais diferenciadas. Portanto, o cálculo de ganho genético mais próximo à condição de lançamento de um novo produto é importante, pois os genótipos com superioridade podem estar acessíveis aos produtores com menor tempo, incrementando a produtividade da cultura.

Diversas metodologias foram propostas para calcular o ganho genético, sendo que no Brasil a metodologia de Vencovsky et al. (1988), foi empregada para o cálculo de ganho genético de produtividade de grãos de milho em 1988. Da metodologia originalmente proposta, surgiram algumas modificações conforme Fernandes (1988), Morais & Abbud (1993), Abreu et al. (1994), Breseghello et al. (1998), Fonseca Junior (1997), sendo que todas as metodologias são baseadas em inferência por mínimos quadrados. Borges et al. (2009) propuseram o cálculo de ganho genético para a cultura do arroz em terras altas, com metodologia baseada em modelos mistos e inferência por verossimilhança restrita (REML).

1.1 Métodos de inferência estatística

Dentre os métodos de inferência estatística, os mais conhecidos e utilizados atualmente, são os métodos baseados em mínimos quadrados. Este método proposto por Carl Friedrich Gauss em 1809 (STIGLER, 1981), baseia-se na resolução de um sistema de equações normais, com o intuito de minimizar a soma

de quadrados dos erros, obtendo as melhores estimativas dos parâmetros, dada está condição.

Para modelos de baixa complexidade, ou seja, com dados balanceados, com poucos parâmetros, é assumido que os dados ajustam-se à distribuição normal, este método possui solução analítica, o qual facilita até mesmo a realização de análises sem recursos computacionais. Este fato foi um dos mais importantes que fez com que este método seja popular e muito utilizado. Dado um modelo linear, apenas com efeitos fixos, menos o erro que é sempre aleatório, pode-se escrever um modelo na forma matricial da seguinte forma:

$$Y = X\beta + \varepsilon$$

Onde:

Y: é o vetor de observações da variável resposta.

X: é a matriz de incidência do modelo.

β : é o vetor dos parâmetros do modelo.

ε : é o vetor de erros do modelo, assumindo que possui distribuição normal com média 0 e variância σ^2 .

Neste exemplo de modelo, o principal objetivo é encontrar os valores para β , que minimizem a soma de quadrados de ε .

Reescrevendo o modelo em termos probabilísticos, tem-se que:

$$\hat{Y} \sim N(\hat{\beta}, \hat{\sigma}^2(X'X)^{-1})$$

Ou seja, \hat{Y} tem distribuição normal com média $\hat{\beta}$ e variância $\hat{\sigma}^2(X'X)^{-1}$, como melhor solução para este modelo, tem-se o estimador de mínimos quadrados para $\hat{\beta}$ que é dado por:

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Esta é a forma matricial para as estimações por mínimos quadrados. Outros métodos algébricos para estimação podem ser encontrados em Resende (2002), Resende (2004) e Casella & Berger (2010).

O método de mínimos quadrados apesar das vantagens, tem a limitação de não considerar efeitos fixos e aleatórios de forma direta. Também não é possível assumir uma estrutura de dependência entre os erros, fato que o faz mais limitado se comparado aos métodos de estimação por verossimilhança que podem ser utilizados tanto em modelos de efeitos fixos como mistos.

A inferência por verossimilhança foi proposta por Fischer em 1922, porém, a sua utilização somente foi explorada após o trabalho de Hartley e Rao (1967), que apresentaram uma especificação matricial para inferência em um modelo misto e as derivações da função verossimilhança para várias classes de modelos. Desta forma, segundo Resende (2004), o trabalho de Hartley e Rao, possibilitou os trabalhos desenvolvidos por Edwards (1972), Bardorff-Nielsen (1976), Gomes (1981), Perreira (1997), Lindsey (1999), ambos reconhecidos como pioneiros na utilização da inferência por verossimilhança segundo Resende (2004).

Um modelo misto tem parâmetros de efeitos aleatório e fixo, e pode ser escrito na forma matricial como:

$$Y = X\beta + Zu + \varepsilon$$

Y: é o vetor de observações da variável resposta.

X: é a matriz de incidência do modelo de efeitos fixos.

β : é o vetor dos parâmetros do modelo.

Z: é a matriz de incidência do modelo de efeitos aleatórios.

u: é o vetor de efeitos aleatórios do modelo.

ε : é o vetor de erros do modelo, assumindo que possui distribuição normal com média 0 e variância σ^2 .

Escrevendo de forma probabilística, assume-se que:

$$\hat{Y} \sim N(\hat{\beta}, V)$$

Onde V é: $V = R + ZGZ'$, sendo R a matriz de efeitos dos erros, e G a matriz de covariância entre os efeitos aleatórios.

Ou seja, pode ter efeitos tanto na média como na variância, isso possibilita representar melhor os fenômenos por meio deste modelo devido a sua robustez. Um outro ponto importante dos modelos mistos é que pode-se definir estruturas de correlações associadas à variância, modelando assim de forma temporal, espacial, ou de formas de correção como o caso de informações de parentesco, popularmente chamadas de *pedigree*.

Toda esta robustez nos modelos mistos, faz com que os métodos de inferência sejam mais complexos. Neste contexto, os métodos de inferência via verossimilhança (ML) e verossimilhança restrita/residual (REML), tem-se mostrado eficientes e com custo computacional adequado para tal tarefa (RESENDE, 2004).

O método da verossimilhança baseia-se em maximizar em conjunto as probabilidades de cada observação. Devido a sua complexidade computacional, aos poucos está se difundindo, devido, em grande parte, à popularização da computação, que permitiu a utilização de métodos numéricos em computadores pessoais.

Os métodos de verossimilhança para modelos complexos (mais que dois parâmetros) não possuem solução analítica na maioria dos casos, sendo necessária a utilização de um método numérico para a solução do mesmo. Dentre os métodos, o de Newton-Rapson tem grande destaque (RESENDE, 2004).

A função de verossimilhança para um modelo com distribuição normal é dada de acordo com a função abaixo.

$$\ln L = \prod_{i=1}^n f(y_i; \theta) = -\frac{1}{2} \{N \ln(2\pi) + \ln|V| + (y - X\beta)'V^{-1}(y - X\beta)\}$$

Da mesma forma que o método de mínimos quadrados, o interesse é achar uma solução para os parâmetros, que maximize a densidade de probabilidade em conjunto das observações.

Baseado na verossimilhança, Patterson & Thompson (1971) aprimoraram a verossimilhança eliminando o seu vício, denominando esta nova metodologia como verossimilhança residual ou restrita. O proposto pelos autores foi maximizar somente a porção da verossimilhança que é invariante aos efeitos fixos, ou seja, “separar” a estimação dos efeitos fixos dos aleatórios, além de ser possível também impor a restrição de não negatividade aos parâmetros de variância do modelo (RESENDE, 2004), e sua equação para a função de verossimilhança é dada por:

$$\ln L = \prod_{i=1}^n f(Ky_i; \theta) = -\frac{1}{2} \{[N - r(X)] \ln(2\pi) - \ln|X'X| + \ln|X'V^{-1}X| + \ln|V| + (y - X\beta)'V^{-1}(y - X\beta)\}$$

Onde:

$$K'y \sim N(0, K'VK)$$

Independentemente do método de inferência, o principal objetivo são os cálculos dos efeitos associados a cada fator. Neste contexto, o BLUE (*Best linear unbiased estimator*) e BLUP (*Best linear unbiased predict*), são os métodos de cálculos dos efeitos para cada fator. O BLUE é oriundo do método dos mínimos quadrados, e o BLUP da verossimilhança. As equações para estimação dos mesmos são as seguintes:

$$BLUE(b) = (X'V^{-1}X)^{-1}X'V^{-1}y$$

$$BLUP(u) = GZ'V^{-1}(y - X\hat{\beta})$$

Em ambos os casos, a dificuldade nos cálculos destes efeitos é baseada na matriz V . Esta matriz pode ser uma matriz identidade, quando os efeitos são independentes, ou pode ser uma matriz densa, como exemplo os efeitos aleatórios espacialmente dependentes. Obter uma V^{-1} de uma matriz densa (com mais de 50% dos elementos não nulos) de dimensão grande, gera um grande custo computacional, fato que durante muito tempo inviabilizou o uso de BLUP em análises. Atualmente, dentro dos problemas mais frequentes, com menos de 5.000 observações, quase todo o computador pessoal pode realizar os cálculos.

Mais detalhes sobre os métodos de inferência e BLUP/BLUE, podem ser obtidos em Patterson & Thompson (1971), Robinson (1991), Searle (1992), Searle (1997), Resende (2002), Resende (2004), Casella & Berger (2010).

1.2 Modelos para determinação de ganho genético

Diversos modelos para o cálculo de ganho genético são encontrados na literatura, sendo clássico o trabalho de Vencovsky et al. (1988), que permite determinar o ganho genético, levando em consideração as substituições de genótipos entre os anos de avaliação. O modelo baseia-se na diferença da média dos genótipos nos ensaios no primeiro e segundo ano de avaliação, em contraste com as médias dos genótipos comuns que estão presentes no primeiro e no segundo ano de avaliação. Por contraste entre os genótipos comuns e diferentes entre os dois anos do biênio, o ganho genético é estimado. Este método foi utilizado por Storck et al. (2005) para a avaliação do ganho genético de milho com dados de ensaios conduzidos em Santa Maria, no Estado do Rio Grande do Sul. Nesse estudo, os autores propuseram um teste de hipóteses para verificar a significância do ganho genético. No método, proposto por Vencovsky et al. (1988), os parâmetros são estimados via inferência pelo método de Mínimos Quadrados (MQ), devido ao desbalanceamento, originário, principalmente, da substituição de genótipos e locais. Já que alguns genótipos são substituídos a cada ano, ou por diferentes genótipos em cada local, as estimativas podem estar sendo imprecisas conforme discutido por Resende (2002), Resende (2004) e Resende (2007).

Alternativamente aos métodos de cálculo de ganho genético, que utilizam estimativas baseadas em inferência por MQ, há o método baseado em inferência por verossimilhança, que tem como principal vantagem ser mais flexível, quanto ao desbalanceamento dos experimentos e ter estimativas melhores destes parâmetros em questão, principalmente, dos componentes de variância (RESENDE, 2002), que são importantes para calcular o ganho genético. Além de proporcionar melhores estimativas para os parâmetros, estes métodos possibilitam obter intervalos de confiança mais precisos para os parâmetros (ROBINSON, 1991; RESENDE, 2004;).

O cálculo de ganho genético com base em verossimilhança é mais popularizado em programas de melhoramento animal do que em vegetal, em partes, devido ao fato de que o melhoramento animal tem situações de desbalanceamento mais acentuadas, se comparado com o vegetal, e também pela falta de cultura de adoção de métodos mais elaborados que é comum na área de melhoramento vegetal. No melhoramento vegetal, estes métodos foram empregados em espécies florestais e frutíferas (RESENDE, 2002) e, em culturas anuais, como o arroz (BORGES et al., 2009).

Com o uso de métodos baseados em verossimilhança, é possível assumir estruturas de correlação para os erros, que podem ser correlacionados tanto temporalmente ou espacialmente, bem como em outras estruturas, como em modelos espaço-temporal, como por exemplo matriz de parentesco (matriz que define o grau de associação por parentesco entre os indivíduos) (*pedigree*). Estas matrizes de correlação, entre os parâmetros são utilizadas no melhoramento genético animal, onde são incorporadas estas informações em modelos, a exemplo a matriz de parentesco dos animais (*pedigree*). No caso de melhoramento vegetal, além da possibilidade de incorporação de matriz de parentesco, também é possível incorporar no modelo, a matriz de correlação dos erros tanto temporalmente como espacialmente, conforme utilizado por Bakhsh et al. (2000) e Cullis & Gleeson (1991).

Na quantificação do ganho genético, é importante uma estimativa pontual do ganho genético, bem como seu respectivo intervalo de confiança, possibilitando assim, inferir sob a significância estatística ou não do ganho calculado a certo grau de probabilidade de erro. Informações quanto ao ganho genético de produtividade de grãos na cultura do milho, com vários anos de avaliações, são bastante escassos na literatura. No Brasil, os trabalhos de Vencovsky et al. (1988), Fernandes e Franzon

(1997) e Storck et al. (2005) foram alguns que calcularam o ganho genético em milho.

Este estudo tem como objetivo estimar o ganho genético da produtividade de grãos de milho na região Sul do Brasil, durante o período de 2001 a 2013, verificando se existem diferenças entre estas situações distintas de cultivo de milho, utilizando as metodologias de Vencovsky et al. (1988) e de Borges et al. (2009). Para a determinação do ganho genético, primeiramente será analisada a região Sul com um único ambiente, e depois será feita estratificação deste ambiente em duas classes baseadas na altitude.

2 CAPÍTULO 1 – GANHO GENÉTICO PARA A PRODUTIVIDADE DE GRÃOS DE MILHO ESTIMADO POR MÍNIMOS QUADRADOS E VEROSSIMILHANÇA

2.1 Resumo

O objetivo deste trabalho foi estimar o ganho genético para produtividade de grãos de milho na região Sul do Brasil. Foi utilizado uma base de dados com 30.292 observações de produtividade de grãos de milho, resultantes 135 genótipos avaliados em 2.826 ensaios de avanço de produto do programa de melhoramento de milho, conduzido durante 13 anos (2001-2013). As metodologias utilizadas foram a de Vencovsky et al. (1988), baseada em modelos lineares com inferência por mínimos quadrados (ML/MQ), e a de Borges et al. (2009), baseada em modelos mistos com inferência com verossimilhança residual (MM/REML). Pela metodologia de ML/MQ, o ganho genético médio anual foi de 121 kg ha⁻¹ ano⁻¹ com intervalo de confiança a 95% de [11;232] kg ha⁻¹ ano⁻¹, e pela metodologia de modelos mistos foi de 79 kg ha⁻¹ ano⁻¹ com intervalo de confiança a 95% de [70;98] kg ha⁻¹ ano⁻¹. A metodologia de MM/REML se demonstrou mais flexível que a metodologia de ML/MQ, por considerar todos dos dados disponíveis na estimação do ganho genético, ao contrário da outra metodologia (ML/MQ), que apenas considera os dados dos biênios em questão para a análise do ganho genético. Para futuros estudos de cálculo de ganho genético, a metodologia baseada em MM/REML, com inferência via MM/REML deve ser priorizada. Houve ganho genético significativo a 5% de probabilidade na produtividade de grãos de milho na região Sul do Brasil, pelas duas metodologias.

Palavras-chave: *Zea mays*, Mínimos quadrados, Modelos mistos

2.2 Introdução

O milho (*Zea mays* L.) é uma planta originária da América Central, onde começou a sua domesticação, processo o qual, durou milhares de anos, pois há

relato que o milho já era conhecido pelos povos indígenas da América Central em 8.700 a.C. no México de acordo com Ranere (2009).

O interesse no milho pelo ser humano, é devido a sua excelente capacidade de produção de energia, para alimentação animal e humana, ou utilização industrial. Este fato, fez o milho se tornar uma das plantas mais cultivadas pela humanidade, tendo atualmente uma produção de 1,16 bilhões de toneladas de grãos de milho, em uma área de 184 milhões de hectares, com produtividade média de grãos de 5.520 kg ha⁻¹ tendo como principais produtores os Estados Unidos, China, Brasil e Argentina, segundo dados de 2013 (FAO, 2014).

No Brasil as condições são favoráveis ao cultivo do milho, devido ao clima, bem como os intensos trabalhos de melhoramento genético, que buscaram adaptar e selecionar genótipos mais produtivos em condições brasileiras. Em 2013, no Brasil foram produzidas 79,9 milhões de toneladas, em área de 15,8 milhões de hectares com produtividade média de 5.057 kg ha⁻¹ (CONAB, 2014)

O foco do melhoramento genético na cultura de milho, é a produtividade de grãos, adaptação geral e resistências a fatores abióticos (Borém & Miranda, 2013) e em segundo plano a resistência a doenças. Diante da importância de ganho de produtividade de grãos na cultura de milho, a estimativa do ganho genético é uma forma de medir se os programas de melhoramento, estão progredindo e lançando genótipos mais produtivos.

Em diferentes fases do desenvolvimento de um genótipo de milho, é possível calcular o ganho genético. O ganho pode ser estimado para a produtividade de grãos, ou outras variáveis (ciclo, altura de plantas, resistência a doenças). O cálculo mais próximo a condição de lançamento de um novo produto é importante, pois os genótipos com superioridade podem estar acessíveis aos produtores, com menor tempo, incrementando a produtividade desta cultura.

Diversos modelos para o cálculo de ganho genético são encontrados na literatura, sendo uma referência para o Brasil o trabalho de Vencovsky et al. (1988), que permite determinar o ganho genético, considerando as substituições de genótipos que ocorrem entre os anos de avaliação. O modelo baseia-se na comparação da média dos ensaios no biênio, com a média dos genótipos comuns aos dois anos em cada ano do biênio analisado. Por contraste destes dois grupos, o ganho genético é estimado. Este método foi utilizado por Storck et al. (2005) para a avaliação do ganho genético de milho com dados de ensaios conduzidos em Santa

Maria, no Rio Grande do Sul. Nesse estudo os autores propuseram um teste de hipóteses para avaliar a significância do ganho genético. No método, proposto por Vencovsky et al. (1988), os parâmetros são estimados via inferência pelo método de Mínimos Quadrados, porém devido ao desbalanceamento que ocorre nesta situação, já que alguns genótipos são substituídos a cada ano, ou por genótipos diferentes em cada local, ou pela variação no número de locais, estas estimativas podem estar sendo imprecisas conforme discutido por Resende (2002), Resende (2004) e Resende (2007).

O cálculo de ganho genético baseado em verossimilhança é mais popularizado em programas de melhoramento animal do que em vegetal, em partes, devido ao fato de que o melhoramento animal tem situações de desbalanceamento mais acentuadas se comparado com o vegetal, e também pela falta de cultura de adoção de métodos mais elaborados que é comum na área de melhoramento vegetal. No melhoramento vegetal, estes métodos foram empregados em espécies florestais e frutíferas (Farias Neto & Resende, 2001; Resende, 2002; Farias Neto & Resende 2008; Farias Neto et al., 2008; Oliveira et al., 2008; Costa et al., 2009; Costa et al., 2010; Farias Neto et al., 2011; Viana et al., 2011a; Viana et al., 2011b; Viana et al., 2011c; Atroch et al., 2011) e, em culturas anuais, como o arroz (Borges et al., 2009; Fritsche-Neto et al., 2012).

O ganho genético de produtividade de grãos de milho foi estudado por alguns pesquisadores em outras partes do mundo, como Duvick (2005) que analisando os dados de 1934 a 2004, encontrou ganho genético $65 \text{ a } 75 \text{ kg ha}^{-1} \text{ ano}^{-1}$, utilizando métodos de regressão linear. Eyherabide & Damilano (2001) na Argentina, com dados entre 1979 a 1998, via modelos de regressão linear, calcularam um ganho genético de $83,2 \text{ kg ha}^{-1} \text{ ano}^{-1}$. Ci et al. (2010) calcularam o ganho genético via ANOVA/Regressão linear, na China, de 1970 a 2010, constataram um ganho genético de $94,7 \text{ kg ha}^{-1} \text{ ano}^{-1}$. Na Africa, Apraku et al. (2015), calcularam ganho de $40 \text{ kg ha}^{-1} \text{ ano}^{-1}$ utilizando ANOVA/Regressão linear. No Brasil, Fernandes e Franzon (1997), em estudo com dados de 30 anos de 1963 a 1993, calcularam ganho de 1.798 kg ha^{-1} ao longo dos anos, que resulta em $59,9 \text{ kg ha}^{-1} \text{ ano}^{-1}$, utilizando a metodologia de Vencovsky et al. (1988). No estado do Rio Grande do Sul, Storck et al. (2005), analisando os dados obtidos em um local, ao longo de 4 anos, não encontraram ganhos genéticos significativos quanto a produtividade de grãos, utilizando a metodologia de Vencovsky et al. (1988).

Diante dos resultados disponíveis na literatura, verifica-se que são escassos os estudos de ganhos genéticos com longa série de dados. Nos estudos nem sempre é apresentada a inferência estatística necessária para afirmar com certo grau de probabilidade, se os valores obtidos são estatisticamente diferentes de zero, que possibilita afirmar se o ganho foi positivo ou negativo ao longo do estudo.

O objetivo desse estudo foi estimar o ganho genético da produtividade de grãos de milho na região Sul do Brasil em um período de 13 anos, utilizando as metodologias de Vencovsky et al. (1988), baseadas em inferência por mínimos quadrados, e de Borges et al., (2009), baseada em modelos mistos com inferência por verossimilhança residual.

2.3 Material e métodos

2.3.1 Metodologia dos ensaios de avanço de produto

O programa de melhoramento genético de milho da Dupont do Brasil S/A – Divisão Pioneer Sementes, possui várias fases, incluindo a obtenção de linhagens endogâmicas, ensaios de *topcross*, ensaios de valor de cultivo e uso (VCU) e, por último, os ensaios de avanço de produto. A cada fase do programa, o número de locais, bem como o tamanho das parcelas são alterados, mas de uma maneira geral; a cada fase, o genótipo é submetido a uma quantidade maior de locais de ensaios, e com parcelas maiores.

Neste estudo, o foco foi na zona ambiental sul, que abrange a região sudoeste e leste do Paraná e os Estados de Santa Catarina e do Rio Grande do Sul. Os ensaios, denominados ensaio de avanço de produto, são conduzidos com o objetivo de avaliar os novos genótipos desenvolvidos pelo programa de melhoramento com os principais genótipos disponíveis no mercado. Estes ensaios são desenvolvidos em áreas de cultivo de milho dos produtores, utilizando todas as práticas de manejo que o produtor adota na área.

Esta diversidade de manejo, época de semeadura, tipo de solo e condições meteorológicas, faz com que os genótipos sejam submetidos a diversas condições ambientais, próximas as quais os genótipos vão se encontrar na produção em larga escala, caso sejam comercializados.

Nestes ensaios há dois grupos de genótipos:

- Novos genótipos desenvolvidos pelo programa de melhoramento;
- Principais genótipos cultivados pelos produtores (testemunhas).

Na figura 1, é apresentado um croqui esquemático do ensaio.

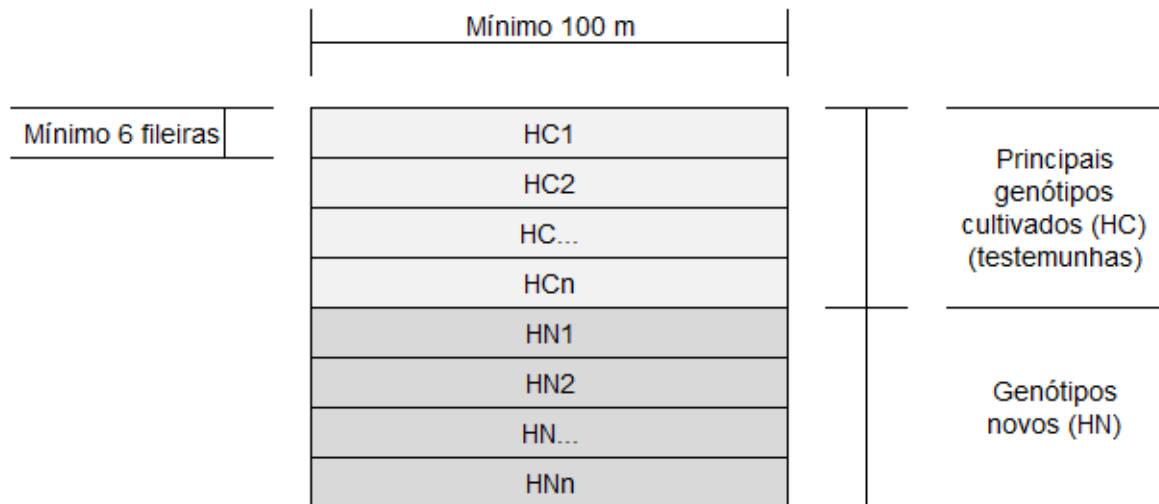


Figura 1 - Croqui de um ensaio de avanço de produto.

Conforme é verificado na Figura 1, o ensaio não possui tamanho de parcela uniforme entre os locais, pois apenas é especificado padrões mínimos. Apenas é definido um número mínimo de fileiras, e um comprimento mínimo da parcela, portanto dependendo das condições do local, as parcelas são de tamanho diferente entre os locais, mas sempre de igual tamanho dentro do local.

2.3.2 Dados

A produtividade de grãos é a principal variável coletada nos ensaios, mas outros dados relacionados ao manejo, também são avaliados e armazenados, o que possibilita utilizar estas variáveis como covariáveis nos modelos, bem como realizar a separação do banco de dados pelas condições específicas, por exemplo, altitude superior a 700 metros.

Na Figura 2 está representada a distribuição espacial dos ensaios durante os anos de 2001 a 2013, onde se verifica que os locais dos ensaios não são os mesmos ao longo dos anos, porém sempre se mantém dentro das principais áreas de produção de milho na região sul.

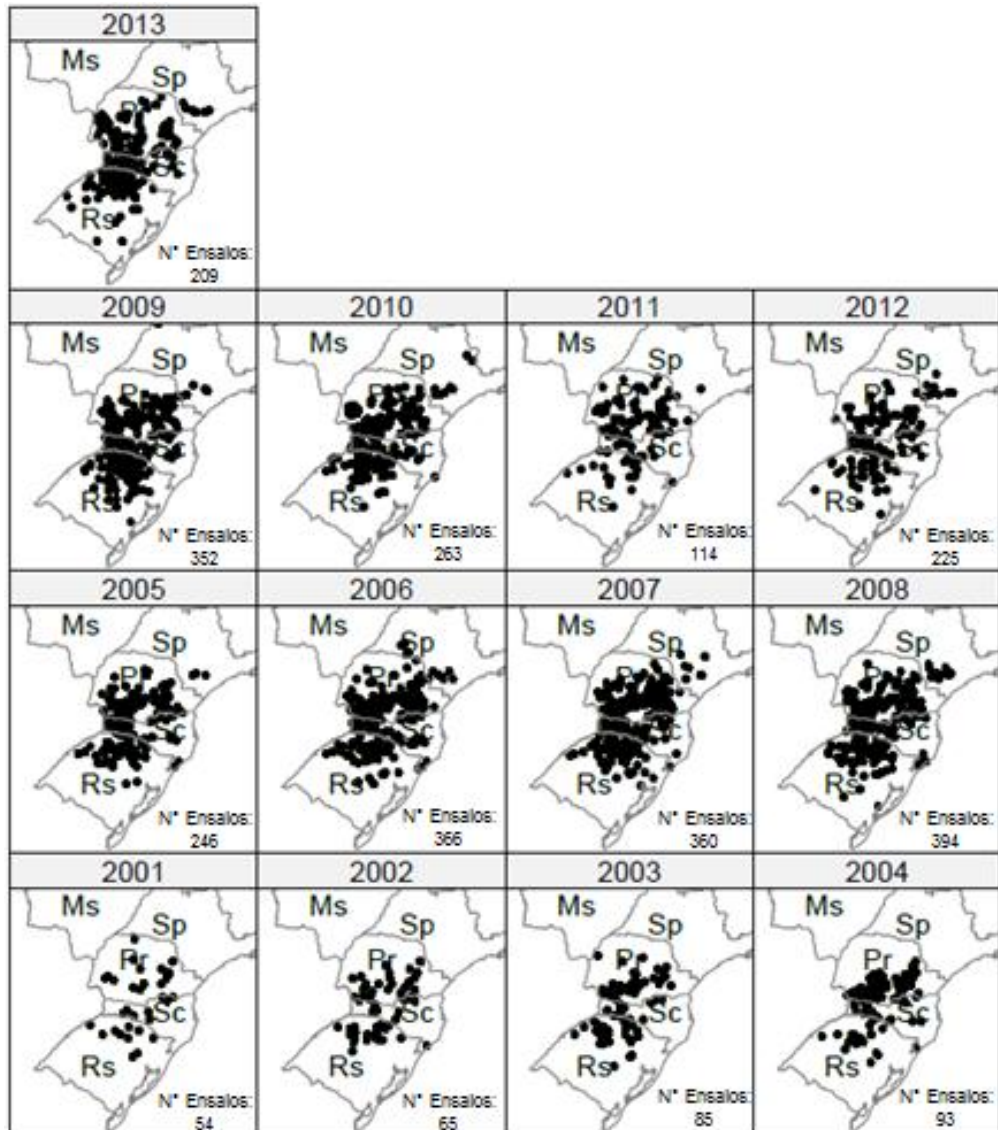


Figura 2 - Distribuição espacial dos ensaios de milho de avanço de produto. Cada círculo preto representa um local de realização do ensaio, entre os anos de 2001 a 2013, na região Sul do Brasil.

O estudo foi realizado com dados de 2.826 locais. Quanto ao número de observações de produtividade de grãos de milho disponíveis, ao longo dos anos esta variou de 701 a 4.176 observações por ano com um total de 30.292. A variável utilizada foi a produtividade de grãos, em kg ha^{-1} , corrigido a 13% de umidade. Como covariáveis, foi utilizada a altitude do local do ensaio, e doses de nitrogênio, fósforo e potássio.

2.3.3 Metodologias

2.3.3.1 Modelo de Vencovsky et al. (1988)

O modelo proposto por Vencovsky et al. (1988) permite calcular o ganho genético comparando dois anos consecutivos. A metodologia deste modelo baseia-se na adoção de um modelo linear para o primeiro ano, conforme descrito abaixo:

$$\bar{Y}_1 = m + a_1 + \bar{g}_1 + \bar{e}_1$$

Onde:

\bar{Y}_1 : média geral do ensaio no ano 1;

m : média geral dos ensaios (dois anos consecutivos);

a_1 : efeito do ambiente no ano 1;

\bar{g}_1 : potencial genético médio de todos os genótipos avaliados no ano 1;

\bar{e}_1 : é o erro, que inclui erro experimental contido na média \bar{Y}_1 e mais a interação dos tratamentos com a ano 1.

O mesmo modelo é adotado para o ano 2:

$$\bar{Y}_2 = m + a_2 + \bar{g}_2 + \bar{e}_2$$

Com estes dois modelos, o interesse está no contraste $\bar{g}_2 - \bar{g}_1$, que é o ganho genético obtido de um ano para o outro.

$$\bar{Y}_2 - \bar{Y}_1 = (a_2 - a_1) + (\bar{g}_2 - \bar{g}_1) + (\bar{e}_2 - \bar{e}_1)$$

Isolando os genótipos comuns:

$$\bar{Y}_{c(1)} = m + a_1 + \bar{g}_{12(c)} + \bar{e}_{1(c)}$$

$$\bar{Y}_{c(2)} = m + a_2 + \bar{g}_{12(c)} + \bar{e}_{2(c)}$$

$$\bar{Y}_{c(2)} - \bar{Y}_{c(1)} = (a_2 - a_1) + (\bar{e}_{2(c)} - \bar{e}_{1(c)})$$

De modo que:

$$\hat{G}_g = (\bar{Y}_2 - \bar{Y}_1) - (\bar{Y}_{c2} - \bar{Y}_{c1}) = (\bar{g}_2 - \bar{g}_1) + (\bar{e}_2 - \bar{e}_1) - [(\bar{e}_{2(c)} - \bar{e}_{1(c)})]$$

Onde:

\hat{G}_g : ganho genético calculado;

\bar{Y}_2 : média geral do ensaio no ano 2;

\bar{Y}_1 : média geral do ensaio no ano 1;

\bar{Y}_{c2} : média geral dos genótipos comuns no ano 2;

\bar{Y}_{c1} : média geral dos genótipos comuns no ano 1.

Neste modelo, o efeito ambiental é calculado pelas genótipos comuns aos dois anos de ensaios, pois toda a variação entre os dois anos de cultivo é devido a efeitos ambientais. Detalhes sobre a demonstração deste modelo podem ser

encontrados em Vencovsky et al. (1988). Para a checagem dos pressupostos do modelo linear, foram realizados os testes de Shapiro-Wilk e Bartlett para normalidade e homogeneidade de variância, respectivamente. Para o teste de hipótese do ganho genético, foi utilizada a metodologia proposta por Storck et al. (2005), e o intervalo de confiança foi realizado baseado na distribuição normal.

2.3.3.2 Modelo de Borges et al. (2009)

Para os estudo de ganho genético via modelos mistos, Borges et al. (2009), propuseram primeiro ajustar um modelo linear misto, e o método de inferência adotada para este modelo foi o de verossimilhança restrita (REML) conforme discutido em Resende (2007). Neste estudo, o modelo adotado, baseado na estrutura dos ensaios, foi:

$$y_{ijklmnop} = \mu + al_i + n_j + p_k + k_l + l_m + a_n + b_o + g_p + e_{ijklmnop}$$

onde:

$y_{ijklmnop}$: Valor da observação referente a altitude i, na dose de nitrogênio j, na dose de fósforo k, na dose de potássio l, no local m, do ano n, da empresa o, do genótipo p

μ : constante geral do modelo;

al_i : efeito da altitude i;

n_j : efeito da dose de nitrogênio j;

p_k : efeito da dose de fósforo k;

k_l : efeito da dose de potássio l;

l_m : efeito do local m;

a_n : efeito do ano n;

b_o : efeito da empresa o;

g_p : efeito do genótipos p ;

$e_{ijklmnop}$: erros referentes a observação na altitude i, na dose de nitrogênio j, na dose de fósforo k, na dose de potássio l, no local m, do ano n, da empresa o, do genótipo p.

Neste modelo, os efeitos das covariáveis al, n, p, k foram considerados efeitos fixos, e os demais aleatórios.

Para o cálculo do ganho genético, depois de ajustado o modelo, foi calculado o valor genético de cada genótipo, somando o efeito individual à média geral, assim o efeito do genótipo foi expresso por: $BLUP = \mu + g$, que por consequência é o BLUP de cada genótipo. Com o BLUP de todos os genótipos, procedeu-se o cálculo do ganho em biênios da seguinte forma subtrai-se os valores genotípicos médios de um determinado ano ao do ano imediatamente anterior por;

$$\hat{G}_g = \frac{\sum_{i=1}^{ng} BLUP_i}{ng_i} - \frac{\sum_{j=1}^{ng} BLUP_{(i-1)}}{ng_{(i-1)}}$$

Onde:

$BLUP_i$: efeito de BLUP dos genótipos no ano i ;

$BLUP_{(i-1)}$: efeito de BLUP dos genótipos no ano $i-1$;

ng_i : número de genótipos no ano i ;

$ng_{(i-1)}$: número de genótipos no ano $i-1$;

O ganho acumulado foi obtido somando-se os ganhos genéticos anuais, ano a ano:

$$\hat{G}_{ga} = \sum_{a=1}^n \left(\frac{\sum_{i=1}^{ng} BLUP_i}{ng_i} - \frac{\sum_{j=1}^{ng} BLUP_{(i-1)}}{ng_{(i-1)}} \right)$$

Para as análises estatísticas foi utilizado o Software R Core Team (2014), sendo os métodos de inferência via verossimilhança restrita implementados com auxílio do pacote *lme4* (BATES; MAECHLER; BOLKER, 2012). Objetivando possuir a distribuição dos parâmetros, posteriormente ao ajuste verossimilhança restrita, o modelo foi ajustado pela metodologia de MCMC (Monte Carlo Markov Chain), a qual permite a obtenção a posteriori dos parâmetros dos modelos, bem como dos efeitos. Os gráficos foram feitos com o pacote *latticeExtra* (DEEPAYAN; ANDREWS, 2012). Para implementar as metodologias, foram programadas funções auxiliares em linguagem R para manipulação de dados e realização das análises.

2.4 Resultados e discussão

2.4.1 Análises descritivas

Na tabela 1, são apresentadas as médias de produtividade de grãos, das covariáveis, e o número de ensaios disponíveis e observações dos dados do estudo.

O número de locais variou entre 54 e 394 ao longo dos anos. Esta variação demonstra a dinâmica com que estes ensaios são conduzidos, pois os mesmos são mudados de local e de número de genótipos ao longo dos anos.

Foi observado, ao longo dos 13 anos, a produtividade média de grãos de milho na região Sul do Brasil, dos ensaios de avanço de produto variou entre 9.383 a 11.331 kg ha⁻¹ com média de 10.233 kg ha⁻¹ que pode se considerada superior a média nacional que foi de 5.057 kg ha⁻¹ no ano de 2013 (CONAB, 2013). Esses dados médios evidenciam que os ensaios foram conduzidos em área de alto investimento, que proporcionou produtividades de grãos próximas à média dos Estados Unidos da América, que foi de 9.969 kg ha⁻¹ no ano de 2013 (FAO, 2013).

Tabela 1 - Análise descritiva da produtividade de grãos de milho na região Sul do Brasil, e das covariáveis nitrogênio (N), fósforo (P) e potássio (K) e altitude representadas pelo valor médio e o número de locais e número de observações por ano.

Ano de semeadura	Número de Locais	Número de observações	Produtividade (kg ha ⁻¹)	N (kg ha ⁻¹)	P (kg ha ⁻¹)	K (kg ha ⁻¹)	Altitude (m)
2001	54	701	9.587	145,0	78,3	84,5	770
2002	65	918	9.961	152,8	88,6	89,9	721
2003	85	1.319	9.962	154,9	97,5	86,0	712
2004	93	1.144	9.636	174,0	98,4	92,3	762
2005	246	2.194	9.312	166,6	93,0	85,9	734
2006	366	4.164	9.894	159,5	89,6	81,4	726
2007	360	4.176	9.956	164,7	89,7	86,0	754
2008	394	3.965	9.383	168,8	88,9	80,7	737
2009	352	3.316	10.857	172,8	93,3	73,1	656
2010	263	3.560	11.196	173,8	97,0	77,0	693
2011	114	1.084	10.399	177,9	95,1	89,2	745
2012	225	1.698	11.277	190,2	106,8	91,9	726
2013	209	2.053	11.331	191,0	105,9	105,9	667
Média	2.826*	30.292*	10.233,5	169,4	93,7	83,5	719

* valor da soma ao longo dos anos

Para o nitrogênio, ao longo dos anos do estudo, a quantidade média utilizada variou entre 145,0 a 191,0 kg ha⁻¹ com média de 169,4 kg ha⁻¹. A menor média da dose de fósforo foi de 78,3 kg ha⁻¹ e a maior de 106,8 kg ha⁻¹, com média de 93,7 kg ha⁻¹, enquanto que para o potássio a variação ficou em média entre 73,1 a 105,9 kg ha⁻¹ com média de 83,5 kg ha⁻¹. Estas três covariáveis, apresentarem tendência de incremento ao longo dos anos. Para a altitude, o valor médio dos ensaios ficou entre 653 a 770 m, com média de 719 metros, porém os ensaios foram conduzidos em locais entre a altitude de 9 a 1314 m.

De forma exploratória, o comportamento dos dados de produtividade de grãos de milho ao longo dos anos observa-se, via análise de regressão linear simples, que aos longo dos anos, a produtividade média aumentou em 179,72 kg ha⁻¹ ano⁻¹. Porém nem todo o incremento é atribuído apenas a ganho genético, pois além da evolução dos novos genótipos, ao longo destes anos a melhoria do manejo que favoreceu a produtividade, como por exemplo, aumento na taxa de semeadura, a redução de espaçamento, e incremento de adubação. Ao analisar o possível incremento na média de produtividade, o importante é separar efeitos genéticos de efeitos ambientais. Segundo Duvick (2005a), nos Estados Unidos como uma análise similar à esta, atribui-se apenas 50-60% devido ao ganho genético, sendo o restante atribuído a melhorias nas praticas de manejo.

De acordo com os percentuais propostos do Duvick (2005a), o ganho genético seria de 89 a 107 kg ha⁻¹, que representaria um ganho percentual sobre a média de 0,99 a 1,2% ao ano. Porém, esta metodologia é simplificada, e resultados mais adequados de ganho serão apresentados com os resultados obtidos pelas duas metodologias que foram utilizadas neste estudo.

2.4.2 Resultados da predição pela metodologia de Vencovsky et al. (1988)

Na Tabela 2, é demonstrado como foi o comportamento da substituição dos genótipos ao longo da série de anos. Observa-se que a taxa de manutenção dos genótipos variou ente 24,59% a 60,87%, isso indicou que em alguns anos os novos genótipos para compor os ensaio não estavam disponíveis como em outros. Quanto a taxa de substituição, teve uma variação entre 15,38% a 66,67% indicando que em alguns anos, a substituição foi bastante elevada de genótipos, indicando maior disponibilidade de novos genótipos para testes.

As diferentes taxas de substituições demonstram a dinâmica que é empregada nos programas de melhoramento. É sempre necessário buscar genótipos com superioridade nas características desejadas, porém em alguns anos, estes genótipos não são encontrados. Por exemplo, em um ano pode-se ter para testes nos ensaios apenas um novo genótipos, em outros anos pode-se ter oito novos genótipos.

Quanto ao número de dados para estimação do ganho genético (N_1 e N_2), pode-se observar que ao longo da série de dados, o mínimo foi de 701 observações de produtividade, valor que é considerado bastante satisfatório para análises estatísticas. Esta variação deve-se a quantidade de ensaios planejados, que podem ser diferentes entre os anos, bem como a disponibilidade de sementes, ou a perda de resultados por fatores abióticos e bióticos, que acontecem com frequência em uma rede de ensaios distribuídas ao longo de uma grande região.

Tabela 2 - Comportamento dos genótipos de milho nos ensaios de avanço de produto ao longo dos biênios, sendo C o número de genótipos comuns aos dois anos do biênio, n_{c1} e n_{c2} , sendo respectivamente o número de genótipos presentes no ano 1 e 2, TM a taxa de manutenção dos genótipos, dada por $[C/(n_{c1}+n_{c2}-C)].100$, TS taxa de substituição dos genótipos dada por $[(n_{c2}-C)/n_{c2}].100$, N_1 e N_2 o número de observações no ano 1 e 2 respectivamente.

Biênio	C	n_{c1}	n_{c2}	TM (%)	TS (%)	N_1	N_2
2001/2002	14	17	20	60,87	30,00	701	918
2002/2003	10	20	18	35,71	44,44	918	1.319
2003/2004	11	18	13	55,00	15,38	1.319	1.144
2004/2005	11	13	33	31,43	66,67	1.144	2.194
2005/2006	27	33	42	56,25	35,71	2.194	4.164
2006/2007	29	42	41	53,70	29,27	4.164	4.176
2007/2008	27	41	41	49,09	34,15	4.176	3.965
2008/2009	15	41	35	24,59	57,14	3.965	3.316
2009/2010	14	35	26	29,79	46,15	3.316	3.560
2010/2011	12	26	15	41,38	20,00	3.560	1.084
2011/2012	11	15	26	36,67	57,69	1.084	1.698
2012/2013	13	26	19	40,62	31,58	1.698	2.053

Para possibilitar o teste de hipótese, conforme proposto por Storck et al. (2005), foi ajustado uma análise de variância para cada ano do biênio. Ao longo dos biênios, apenas em 2011/2012 e 2010/2011 houve homogeneidade das variâncias (valor-p > 0,05), já o pressuposto da normalidade em nenhum dos biênios foi atendido (Tabela 3). Pela característica dos dados que não são balanceados, evidenciam que nestes casos o emprego do cálculo de ganho genético via método de inferência por mínimos quadrados, pode vir a estimar ganhos genéticos sem uma

precisão adequada, pois apresentam violação nos pressupostos dos modelos, conforme discutido por Resende (2004).

Tabela 3 - Resultados das análises de variâncias realizadas para estimativa do modelo de Vencovsky et al. (1988), sendo QM_{e1} quadrado médio do erro do ano 1, GL_{e1} grau de liberdade do erro no ano 1, SW_1 valor-p do teste de Shapiro Wilk para normalidade do erro no ano 1, $Bartlett_1$ é o valor-p do teste de Bartlett para homogeneidade do erro do ano 1, QM_{e2} quadrado médio do erro do ano 2, GL_{e2} grau de liberdade do erro no ano 2, SW_2 valor-p do teste de Shapiro Wilk para normalidade do erro no ano 2, $Bartlett_2$ é o valor-p do teste de Bartlett para homogeneidade do erro do ano 2.

Biênio	QM_{e1}	GL_{e1}	SW_1 (valor-p)	$Bartlett_1$ (valor-p)	QM_{e2}	GL_{e2}	SW_2 (valor-p)	$Bartlett_2$ (valor-p)
2001/2002	833.976	631	< 0,01	0,754	678.829	834	< 0,01	< 0,01
2002/2003	678.829	834	< 0,01	< 0,01	733.393	1.217	< 0,01	< 0,01
2003/2004	733.393	1.217	< 0,01	< 0,01	579.609	1.039	< 0,01	< 0,01
2004/2005	579.609	1.039	< 0,01	< 0,01	585.234	1.916	< 0,01	< 0,01
2005/2006	585.234	1.916	< 0,01	< 0,01	659.747	3.757	< 0,01	< 0,01
2006/2007	659.747	3.757	< 0,01	< 0,01	507.551	3.776	< 0,01	< 0,01
2007/2008	507.551	3.776	< 0,01	< 0,01	633.622	3.531	< 0,01	< 0,01
2008/2009	633.622	3.531	< 0,01	< 0,01	676.007	2.930	< 0,01	< 0,01
2009/2010	676.007	2.930	< 0,01	< 0,01	744.812	3.272	< 0,01	< 0,01
2010/2011	744.812	3.272	< 0,01	< 0,01	669.313	956	< 0,01	0,491
2011/2012	669.313	956	< 0,01	0,491	714.135	1.448	< 0,01	< 0,01
2012/2013	714.135	1.448	< 0,01	< 0,01	792.589	1.826	< 0,01	< 0,01

Analisando os resíduos via gráficos de diagnóstico, os resíduos, não tem um comportamento que pode-se considerar grave a ponto de inviabilizar as análises, mesmo os dados não seguindo a distribuição normal de acordo com o teste de Shapiro-Wilk a 5% de probabilidade de erro, e não atendendo a homogeneidade de variâncias pelo teste de Bartlett a 5% de probabilidade de erro, esta violação, pode estar associada ao número grande de observações. Neste contexto a análise dos resíduos de forma gráfica, conforme discutido por Resende (2007), torna-se satisfatório para continuar com as análises.

Na tabela 4 são apresentados os resultados de ganho genético estimados de acordo com a metodologia proposta por Vencovsky et al. (1988). Observa-se que as produtividades médias entre os anos foram superiores a $10.118 \text{ kg ha}^{-1}$. Analisando o ganho genético estimado ao longo dos biênios, o mesmo apresentou em 2 biênios resultados negativos de ganhos genéticos, sendo os demais positivos, chegando a 354 kg ha^{-1} . Na média geral dos anos o ganho foi de $121 \text{ kg ha}^{-1} \text{ ano}^{-1}$, que comparando com a média dos ensaios corresponde a 1,2% por ano.

Observando a significância dos ganhos genéticos calculados pela metodologia de Vencovsky et al. (1988), que apenas para o biênio de 2001/2002 e para a média geral de ganho, estes foram significativos a 5% de probabilidade.

Tabela 4 - Resultados gerais do modelos de Vencovsky et al. (1988), sendo Y_1 média geral do ano 1, Y_2 média geral do ano 2, Y_{C1} média geral dos genótipos comuns aos dois anos no ano 1, Y_{C2} média geral dos genótipos comuns aos dois anos no ano 2, G_g ganho genético, V_{Gg} variância do ganho genético, t_{Cal} valor calculado do teste t, sob hipótese de ganho diferente de 0, valor-p é valor-p do teste t, $IC_{(95\%)}$ é o intervalo de confiança para o ganho genético.

Biênio	Y_1 (kg ha ⁻¹)	Y_2 (kg ha ⁻¹)	Y_{C1} (kg ha ⁻¹)	Y_{C2} (kg ha ⁻¹)	G_g (kg ha ⁻¹)	V_{Gg}	t_{Cal}	valor-p	$IC_{(95\%)}$ (kg ha ⁻¹)
2001/2002	9.587	9.961	9.709	9.729	354	25.058	2,24	0,013	[44;664]
2002/2003	9.961	9.962	10.069	9.956	114	66.536	0,44	0,329	[-391;619]
2003/2004	9.962	9.636	10.036	9.654	56	34.034	0,31	0,380	[-305;417]
2004/2005	9.636	9.312	9.679	9.539	-184	43.575	-0,88	0,189	[-593;225]
2005/2006	9.312	9.894	9.429	9.866	144	12.667	1,28	0,100	[-76;365]
2006/2007	9.894	9.956	9.975	9.950	87	12.164	0,79	0,215	[-129;303]
2007/2008	9.956	9.383	10.066	9.340	153	14.432	1,28	0,101	[-82;388]
2008/2009	9.383	10.857	9.491	10.797	168	52.539	0,73	0,232	[-281;617]
2009/2010	10.857	11.196	11.166	11.316	188	53.526	0,82	0,207	[-264;642]
2010/2011	11.196	10.399	11.387	10.383	207	44.576	0,98	0,163	[-206;621]
2011/2012	10.399	11.277	10.590	11.516	-47	53.680	-0,21	0,418	[-502;406]
2012/2013	11.277	11.331	11.513	11.349	217	46.719	1,01	0,157	[-205;641]
Geral	10.118	10.264	10.259	10.283	121,7	3191	2,15	0,016	[11,0;232,4]

2.4.3 Resultados da predição pela metodologia de Borges et al. (2009)

Quanto a metodologia proposta por Borges et al. (2009), verifica-se na tabela 5 as informações dos modelos ajustados, onde pode-se verificar o teste de cada componente de variância pela metodologia do teste da razão de verossimilhança. Verificou-se que os componentes de variância foram significativos a 5% de probabilidade, pelo teste de razão de verossimilhança, fato esse, permite manter todos estes componentes de variância no modelo. Nestes casos, onde vários dos componentes são associados a efeitos ambientais, este componentes ajudam a explicar melhor a variação dos dados.

Na tabela 5, são apresentadas as estimativas dos componentes de variância associados a cada componente do modelo ajustado, bem como os efeitos fixos do modelo. Quanto aos componentes de variância, verifica-se que 62,89% da variância dos componentes aleatórios esta associada ao local do experimento (I_m), ou seja, a grande variabilidade deste tipo do ensaio esta nos locais, que carregam junto consigo toda a diferença de manejo, época de semeadura, tipo de solo, regime de hídrico, condições climáticas, manejo de controle de pragas e doenças, sendo que estas variáveis tem relação fisiológica direta com a produtividade de grãos.

Outro fator que obteve o segundo maior componente de variância foi o erro ($e_{ijklmnop}$), com 18,99% demonstrando que é maior que os componentes do genótipo, ano e empresa. Este erro elevado pode estar associado ao tipo de ensaios que possuem poucos genótipos e sem repetições. Este elevado erro, pode ser um

complicador para a seleção de genótipos superiores, se estes não tiveram uma expressão individual significativa.

O componente de variância associado à empresa (b_o), que poderia ser um indicativo de grupos genéticos diferentes entre as empresas, foi de 0,77% da variância total, ou seja, mesmo sendo significativo a 5% de probabilidade, a base genética entre as empresas tem pouca diferença na produtividade de grãos. O componente associado a genótipo (g_p) e ano (a_n), tiveram estimativas bem similares com 9,29% e 8,05% da variância total respectivamente, isso indica, que a variância entre os anos tem uma importância similar aos genótipos testados em cada ano.

Quanto aos efeitos fixos, as adubações com nitrogênio (n_j), fósforo (p_k) e potássio (k_i), bem como a altitude (al_i), capturaram a variação associada as estes quatro componentes, que possuem em nossas condições de cultivo do milho, que é feita em solos que são pobres em fósforo, e com baixo teor de matéria orgânica, que é fonte natural de nitrogênio, estes componentes tem papel importância por ser reconhecidos na literatura as respostas do milho a adubação com estes nutrientes.

Quanto à altitude, a zona ambiental sul de cultivo de milho apresenta grande variabilidade, onde os cultivos, são realizado em zona com altitude próxima a 9, até 1314 m de altitude, esta variação ampla, tem grande associação com as condições climáticas, principalmente temperatura, bem como com os tipos de solos predominantes.

Pelos parâmetros ajustados no modelo, pode-se inferir que na medida que o incremento de 1 m de altitude no local de cultivo, tem uma variação de $1,33 \text{ kg ha}^{-1}$, apesar de parecer pouco, quando comparado com regiões próximas ao nível do mar com regiões de cultivo de maiores altitudes como o segundo planalto paranaense, a diferença de 1000 metros de altitude, que representaria uma diferença de produtividade de aproximadamente 1.333 kg ha^{-1} que é um valor que pode ser considerado representativo em termos de produtividade.

Quanto aos parâmetros associados a adubação, a variação de incremento de 1 kg ha^{-1} de adubo aplicado foi de 5,33 a $6,35 \text{ kg ha}^{-1}$ mostrando ter relação com as praticas de adubação que são utilizadas em lavouras de alto rendimento na produtividade de grãos.

Tabela 5 - Resultados dos componentes de variâncias, e dos efeitos fixos, do modelo misto ajustado, sendo θ , o parâmetro do modelo, e $\hat{\theta}$, a estimativa obtida para o parâmetro via REML, e σ o desvio padrão da estimativa do parâmetro.

Efeitos Fixos					Efeitos Aleatórios				
θ	$\hat{\theta}$	σ	t	valor-p	θ	$\hat{\theta}$	σ	% σ^2	valor-p
μ	6.758,02	230,4	29,3	< 0,001	l	2.274.141	1.508,0	62,89	< 0,001
al	1,33	0,1	10,7	< 0,001	g	335.857	579,5	9,29	< 0,001
n	6,35	0,8	8,0	< 0,001	a	291.213	539,6	8,05	< 0,001
p	6,08	1,2	5,0	< 0,001	b	27.974	167,3	0,77	< 0,001
k	5,33	0,9	5,8	< 0,001	e	686.624	828,6	18,99	< 0,001

μ : Constante geral do modelo, al: efeito da altitude, n: efeito da dose de nitrogênio, p: efeito da doses de fosforo, k: efeito da doses de potássio, l: efeito do local, g: efeito do genótipo, a: efeito do ano, b: efeito da empresa, e: efeito do erro.

Na Tabela 6, é apresentado o ganho genético na produtividade de grãos de milho estimados pela metodologia de Borges et al. (2009), verifica-se que o ganho foi em média ao longo dos anos de 89 kg ha⁻¹ ano⁻¹, com variação entre -230 kg ha⁻¹ a 249 kg ha⁻¹, e com todos os valores estatisticamente significativos a 5% de probabilidade, quando testados sob a hipótese de ganho nulo, verifica-se também, que a amplitude do intervalo de confiança do ganho foi bem menor se comparado aos intervalos obtidos pela metodologia de Vencovsky et al. (1988).

Tabela 6 - Resultado do ganho via modelos mistos, sendo u_1 e u_2 , valor genético do ano 1 e 2 respectivamente, V_{u1} e V_{u2} variância do ganho genético no ano 1 e 2 respectivamente, G_g ganho genético, V_{Gg} variância do ganho genético, valor-p é o valor-p sob o teste da diferença do ganho em relação a 0, IC_(95%) intervalo de confiança do ganho genético com 95% de probabilidade.

Biênio	u_1 (kg ha ⁻¹)	u_2 (kg ha ⁻¹)	V_{u1}	V_{u2}	G_g (kg ha ⁻¹)	V_{Gg}	Valor-p	IC _(95%) (kg ha ⁻¹)
2001/2002	-581	-331	3.298	3.074	249	613	< 0,01	[198;296]
2002/2003	-331	-228	3.074	2.700	103	748	< 0,01	[47;155]
2003/2004	-228	-50	2.700	2.610	177	505	< 0,01	[134;222]
2004/2005	-50	-281	2.610	2.413	-230	736	< 0,01	[-281;-173]
2005/2006	-281	-169	2.413	2.141	111	257	< 0,01	[77;139]
2006/2007	-169	-76	3.684	3.782	92	326	< 0,01	[60;132]
2007/2008	-76	60	3.782	3.407	13	298	< 0,01	[104;171]
2008/2009	60	273	3.407	3.204	212	594	< 0,01	[163;260]
2009/2010	273	431	3.204	3.269	157	592	< 0,01	[108;204]
2010/2011	431	535	3.269	2.938	103	338	< 0,01	[68;139]
2011/2012	535	356	2.938	3.280	-178	407	< 0,01	[-217;-138]
2012/2013	356	496	3.280	3.070	140	978	< 0,01	[73;196]
Geral	-	-	-	-	89	20,2	< 0,01	[80;98]

A maior diferença apresentada pelas duas metodologias são os intervalos de confiança e as médias, onde os intervalos pela metodologia de Borges et al. (2009) foram menores, fato este, deve estar associado a esta metodologia utiliza todos os dados para o cálculo, diferente da metodologia de Vencovsky et al. (1988) que utiliza apenas os dados dos biênios de forma isolada para o cálculo.

Os ganhos obtidos pela metodologia de Borges et al. (2009) são menores se comparados a metodologia de Vencovsky et al. (1988), e estão associados ao efeito de *Shrinkage* (“*Encolhimento*”) que os modelos misto normalmente proporcionam aos valores preditos, conforme discutido por Duarte & Vencovsky (2001).

Quanto a significância dos efeitos, pode-se observar que pela metodologia dos Borges et al. (2009), estes sempre foram significativos a 5% de probabilidade, ao contrário da metodologia de Vencovsky et al. (1988), que como apresenta seu intervalo de confiança de 95% sempre contendo o 0, indicando um ganho não significativo.

Os resultados de ganho genéticos, estimados pelas duas metodologias, são similares aos encontrados por outros pesquisadores, Duvick (2005), realizando o cálculo nos Estados Unidos, no meio oeste, região tradicional de produção de milho, estudando os dados de 1934 a 2004, encontrou ganhos de produtividade de 115 kg ha⁻¹ ano⁻¹, e atribuiu ao ganho genético 65 a 75 kg ha⁻¹ ano⁻¹. Eyherabide & Damilano (2001) analisando ensaios conduzidos na Argentina, entre os anos de 1979 a 1998, utilizando modelos de regressão, calcularam um ganho genético de 83 kg ha⁻¹ ano⁻¹, que representava em termos de percentual 2,9% ao ano. Ci et al. (2010) analisando o ganho genético de milho na China de 1970 a 2010, constataram um ganho genético de 94 kg ha⁻¹ano⁻¹, atribuindo este avanço a melhoria na resistência ao stress e aumento na população de plantas por área, proporcionado pelos novos genótipos.

Apraku et al. (2015), analisando o resultado de ganho genético no oeste e centro da África, em dois anos, chegaram a ganho de 40 kg ha⁻¹ ano⁻¹, que representa um ganho de 1,3% sob suas condições.

Fernandes e Franzon (1997), em estudo no Brasil com dados de 30 anos de 1963 a 1993, via metodologia de Vencovsky et al. (1988) modificada, calcularam um ganho de 1.798 kg ha⁻¹ ao longo dos anos, que resulta em 59 kg ha⁻¹ ano⁻¹.

Com os resultados apresentados, pode-se inferir que a metodologia de Borges et al. (2009) proporcionou uma estimativa de ganho genético mais precisos e mais acurados, se comparado a metodologia de Vencovsky et al. (1988). A metodologia de Borges et al. (2009) possibilitou realizar o cálculo do ganho genético ao longo dos anos, não particionando em biênios como na metodologia de Vencovsky et al. (1988).

Outro aspecto bastante favorável a utilização da metodologia de Borges et al. (1988) se comparado as metodologias de Vencovsky et al. (1988), é a questão da violação dos pressupostos que a metodologia de mínimos quadrados proporciona, pois como neste caso onde os experimentos ao longos dos anos são desbalanceados, a metodologia de Vencovsky et al. (1988) não é tão adequada.

Uma das vantagens da metodologia de Vencovsky et al. (1988) é a simplicidade na realização dos cálculos, e o custo computacional do mesmo que é bastante baixo, se comparado a metodologia de Borges et al. (2009), principalmente quando tem-se uma grande base de dados como esta, com 30.292 observações.

2.5 Conclusão

Houve ganho genético para a produtividade de grãos de milho de $121 \text{ kg ha}^{-1} \text{ ano}^{-1}$ com intervalo de confiança a 95% de $[11;232] \text{ kg ha}^{-1} \text{ ano}^{-1}$ na metodologia de Vencovsky et al. (1988) e de $79 \text{ kg ha}^{-1} \text{ ano}^{-1}$ pela metodologia de Borges et al. (2009), com intervalo de confiança a 95 % de $[70;98] \text{ kg ha}^{-1} \text{ ano}^{-1}$.

3 CAPÍTULO 2 – GANHO GENÉTICO PARA A PRODUTIVIDADE DE GRÃOS DE MILHO EM DUAS CLASSES DE ALTITUDE

3.1 Resumo

O objetivo deste estudo foi estimar o ganho genético da produtividade de grãos de milho em duas classes de altitude na região sul do Brasil. Para a formação das classes, utilizou-se como critérios de altitude inferior e superior a 700m, a base de dados utilizada, é constituída de 30.292 observações de produtividades de grãos de milho, de 135 genótipos, resultantes de 2.826 ensaios, conduzidos em 13 anos (2001 a 2013), resultantes dos ensaios de avanço de produto do programa de melhoramento de milho. Para o cálculo do ganho genéticos foi utilizado a metodologia proposta por Borges et al. (2009), baseada em modelos mistos com inferência via MM/REML. O ganho genético foi diferenciado em cada classe analisada. Verificou-se que em altitude superiores a 700m o ganho foi de 94 kg ha⁻¹ ano⁻¹, já para altitude inferior a 700 m o ganho foi de 74 kg ha⁻¹ano⁻¹. Todas as estimativas sendo estatisticamente significativos a 5% de probabilidade. Pode-se evidenciar que nas classes analisadas obteve-se ganho genético significativo entre os anos de 2001 a 2013.

Palavras-chave: *Zea mays*, Modelos mistos, Ganho genético

3.2 Introdução

O milho (*Zea mays* L.) é uma planta originária da América Central, atualmente é cultivada em todos os continentes. Atualmente uma produção mundial é de 1,16 bilhões de toneladas, em uma área de 184 milhões hectares, com uma produtividade média de 5.520 kg ha⁻¹ tendo como principal produtores os Estados Unidos, China, Brasil e Argentina, segundo dados de 2013 (FAO, 2014). O Brasil tem uma produção de 79,9 milhões de toneladas, em uma área de 15,8 milhões de hectares com produtividade média de 5.057 kg ha⁻¹ segundo dados de 2013 (CONAB, 2014)

A região sul do Brasil, que compreende os estados do Rio Grande do Sul, Santa Catarina e Paraná, é uma importante região produtora de milho sendo a

segunda região que mais produz, ficando atrás apenas para a região Centro Oeste (CONAB, 2013). Diante da importância do milho, o estudo do ganho genético na região sul, bem como dentro da região se faz importante. Ao analisar a região sul, verificou-se que o cultivo de milho é realizado em diferentes condições de altitude, com plantios desde 9 m a até 1314 m.

O foco do melhoramento genético na cultura de milho, é a produtividade de grãos, adaptação geral e resistência a fatores abióticos (Borém & Miranda, 2013) e em segundo plano a resistência a doenças. A estimativa do ganho genético é uma forma de medir se os programas de melhoramento, estão progredindo e lançando genótipos com maior produtividade de grãos.

Em diferentes fases durante o desenvolvimento de um genótipo de milho, é possível calcular o ganho genético, seja para a produtividade de grãos, ou outras variáveis (ciclo, altura de plantas, resistência a doenças). O cálculo de ganho genético, pode ser realizado por diferentes metodologias, dentre elas a metodologia por inferência via verossimilhança residual. Esta metodologia não é até então, muito popularizada em programas de melhoramento vegetal, sendo mais utilizada em programas de melhoramento animal. No melhoramento vegetal, estes métodos foram empregados em melhoramento de espécies florestais e frutíferas (Farias Neto & Resende, 2001; Resende, 2002; Farias Neto & Resende 2008; Farias Neto et al., 2008; Oliveira et al., 2008; Costa et al., 2009; Costa et al., 2010; Farias Neto et al., 2011; Viana et al., 2011a; Viana et al., 2011b; Viana et al., 2011c; Atroch et al., 2011) e, em culturas anuais, como o arroz (Borges et al., 2009; Fritsche-Neto et al., 2012).

Tendo em vista que o programa de melhoramento é realizado para toda a região sul, e um novo genótipo quando é lançado, também é comercializado em toda a região. Buscar verificar se existem diferenças no ganho genético diferentes altitudes se faz importante. Se o ganho não estiver sendo homogêneo dentro da região sul, pode indicar que o programa de melhoramento não está sendo feito de uma maneira ampla para toda a região, deixando áreas sem novos materiais com melhores características agrônômicas, que gerem uma melhor produtividade, dos que os presentes no mercado. Trabalhos que buscam quantificar este ganho em diferentes classes de altitudes são inexistentes para a região sul do Brasil. E os trabalhos de ganho genético no Brasil, como o de Fernandes e Franzon (1997),

exploraram apenas o Brasil de uma forma geral, não uma região, e nem classes dentro desta região.

Este trabalho tem como objetivo calcular do ganho genético da produtividade de grãos de milho na região sul do Brasil, em um período de 13 anos, utilizando a metodologia proposta por Borges (2009), que é baseada em modelos mistos com inferência via verossimilhança residual em duas classes de altitude.

3.3 Material e métodos

3.3.1 Metodologia dos ensaios de avanço de produto.

O programa de melhoramento genético de milho da Dupont do Brasil S/A – Divisão Pioneer Sementes, possui várias fases, incluindo a obtenção de linhagens endogâmicas, ensaios de *topcross*, ensaios de valor de cultivo e uso (VCU) e por último os ensaios de avanço de produto, que são os ensaios de avanço de produto. A cada fase do programa, o número de locais, tamanho das parcelas, são alterados, mas de uma maneira geral, a cada fase, o genótipo é submetido a uma quantidade maior de locais de ensaios, e com parcelas maiores.

Neste estudo o foco foi na zona ambiental sul, que abrange a região sudoeste e leste do Paraná e os Estados de Santa Catarina e do Rio Grande do Sul. Os ensaios, denominados ensaios de avanço de produto, são conduzidos com o objetivo de avaliar os novos genótipos desenvolvidos pelo programa de melhoramento com os principais genótipos disponíveis no mercado. Estes ensaios são desenvolvidos em áreas de cultivo de milho dos produtores, utilizando todas as práticas de manejo que o produtor adota na área.

Esta diversidade de manejo, época de semeadura, tipo de solo e condições meteorológicas, faz com que os genótipos sejam submetidos a diversas condições ambientais, próximas as quais os genótipos vão encontrar na produção em larga escala, caso sejam comercializados.

Nestes ensaios há dois grupos de genótipos:

- Novos genótipos desenvolvidos pelo programa de melhoramento;
- Principais genótipos cultivados pelos produtores (testemunhas).

Na figura 3, é apresentado um croqui esquemático do ensaio.

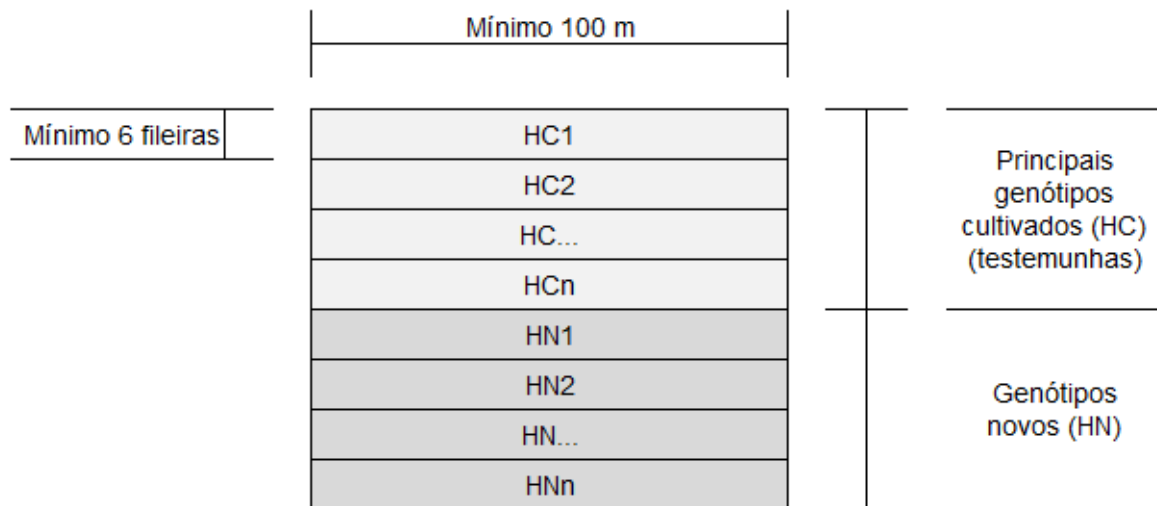


Figura 3 - Croqui de um ensaio de avanço de produto.

Conforme é verificado na figura 3, o ensaio não possui tamanho de parcela uniforme entre os locais, pois apenas é especificado padrões mínimos. Apenas é definido um número mínimo de fileiras, e um comprimento mínimo da parcela, portanto dependendo das condições do local, as parcelas são de tamanho diferente entre os locais, mas sempre de igual tamanho dentro do local.

3.3.2 Dados

A produtividade de grãos, é a principal variável coletada nos ensaios, mas outros dados relacionados ao manejo, também são avaliados e armazenados, o que possibilita utilizar estas variáveis como covariáveis nos modelos, bem como realizar a separação do banco de dados pelas condições específicas, por exemplo, altitude. Na Figura 4, está representada a distribuição espacial dos ensaios durante os anos de 2001 a 2013, onde verifica-se que os locais dos ensaios não são os mesmos ao longo dos anos, porém sempre se mantém dentro das principais áreas de produção de milho na região sul.

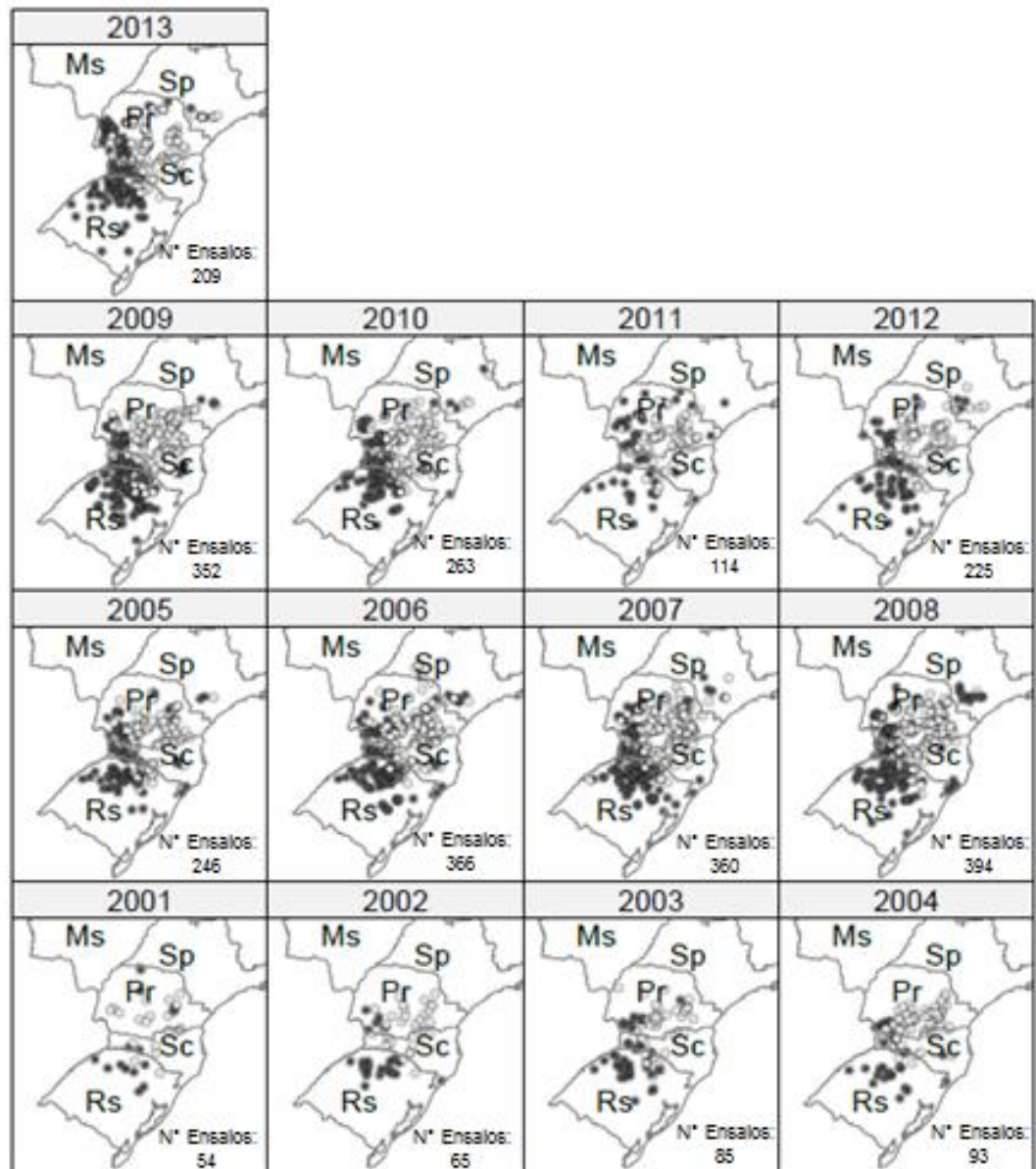


Figura 4 - Distribuição espacial dos ensaios de milho de avanço de produto. Realizados entre os anos de 2001 a 2013, na região sul do Brasil, pontos cinza forte indicam ensaios em altitude inferior a 700m e pontos cinza claro indicam ensaios em altitude superior a 700m.

No total, o estudo está sendo realizado com dados de 2.826 locais, e 135 genótipos, e o número de observações disponíveis, ao longo dos anos esta variou de 701 a 4.176 por ano, observações com um total de 30.292, A variável utilizada foi a produtividade de grãos oriundas destes ensaios, onde esta variável foi padronizada na unidade de quilogramas por hectare, à 13% de umidade. Como covariáveis disponíveis neste banco de dados, pode-se destacar a altitude do local do ensaio, e a adubação por nitrogênio, fósforo e potássio.

Como o objetivo de verificar como foi o comportamento do ganho genético em diferentes classes dentro da zona ambiental sul, foi proposta a divisão por altitude,

onde o valor de 700 metros foi escolhido por ser um valor próximo ao valor mediano desta covariável, e por ser um valor também de referência para o setor.

3.3.2.1 Metodologia

Para os estudo de ganho genético via modelos mistos, Borges et al. (2009), propuseram primeiro ajustar um modelo linear misto, e o método de inferência adotada para este modelo foi o de verossimilhança restrita (REML) conforme discutido em Resende (2007). Neste estudo o modelo adotado, baseado na estrutura dos ensaios, foi:

$$y_{ijklmnop} = \mu + al_i + n_j + p_k + k_l + l_m + a_n + b_o + g_p + e_{ijklmnop}$$

onde:

$y_{ijklmnop}$: Valor da observação referente a altitude i, na dose de nitrogênio j, na dose de fósforo k, na dose de potássio l, no local m, do ano n, da empresa o, do genótipo p

μ : constante geral do modelo;

al_i : efeito da altitude i;

n_j : efeito da dose de nitrogênio j;

p_k : efeito da dose de fósforo k;

k_l : efeito da dose de potássio l;

l_m : efeito do local m;

a_n : efeito do ano n;

b_o : efeito da empresa o;

g_p : efeito do genótipos p ;

$e_{ijklmnop}$: erros referentes a observação na altitude i, na dose de nitrogênio j, na dose de fósforo k, na dose de potássio l, no local m, do ano n, da empresa o, do genótipo p.

Neste modelo os efeitos das covariáveis al, n, p, k foram considerados efeitos fixos, e os demais aleatórios.

Para o cálculo do ganho genético, depois de ajustado o modelo, foi calculado o valor genético de cada genótipo, somando o efeito individual à média geral, assim o efeito do genótipo foi expresso por: $BLUP = \mu + g$, que por consequência é o BLUP de cada genótipo. Com o BLUP de todos os genótipos, procedeu-se o cálculo

do ganho em biênios da seguinte forma subtrai-se os valores genotípicos médios de um determinado ano ao do ano imediatamente anterior por;

$$\hat{G}_g = \frac{\sum_{i=1}^{ng} BLUP_i}{ng_i} - \frac{\sum_{j=1}^{ng} BLUP_{(i-1)}}{ng_{(i-1)}}$$

Onde:

$BLUP_i$: efeito de BLUP dos genótipos no ano i;

$BLUP_{(i-1)}$: efeito de BLUP dos genótipos no ano i-1;

ng_i : número de genótipos no ano i;

$ng_{(i-1)}$: número de genótipos no ano i-1;

O ganho acumulado foi obtido somando-se os ganhos genéticos anuais, ano a ano;

$$\hat{G}_{ga} = \sum_{a=1}^n \left(\frac{\sum_{i=1}^{ng} BLUP_i}{ng_i} - \frac{\sum_{j=1}^{ng} BLUP_{(i-1)}}{ng_{(i-1)}} \right)$$

Para as análises estatísticas, foi utilizado o Software R Core Team (2014), sendo os métodos de inferência via verossimilhança restrita implementado com auxílio do pacote *lme4* (BATES; MAECHLER; BOLKER, 2012). Objetivando possuir a distribuição dos parâmetros, posteriormente ao ajuste verossimilhança restrita, o modelo foi ajustado pela metodologia de MCMC (Monte Carlo Markov Chain). Metodologia a qual, permite a obtenção a posteriori dos parâmetros dos modelos bem como dos efeitos. Os gráficos foram feitos com o pacote *latticeExtra* (DEEPAYAN; ANDREWS, 2012). Para implementar as metodologias, foram programadas função auxiliares em linguagem R para manipulação de dados e realização das análises.

3.4 Resultados e discussão

3.4.1 Análise exploratória

Uma análise exploratória, da produtividade de grão de milho na região sul do Brasil, relacionada a altitude dos ensaios pode ser verificada na Figura 5, onde via um modelo de regressão linear simples, verifica-se que a produtividade aumenta com o acréscimo da altitude, na ordem de 1,27 kg ha⁻¹ m⁻¹. Também verifica-se, produtividades elevadas acima de 12.000 kg ha⁻¹ em quase todas as altitudes,

evidenciando que outras fontes de variações, também podem influenciar a produtividade da cultura de milho.

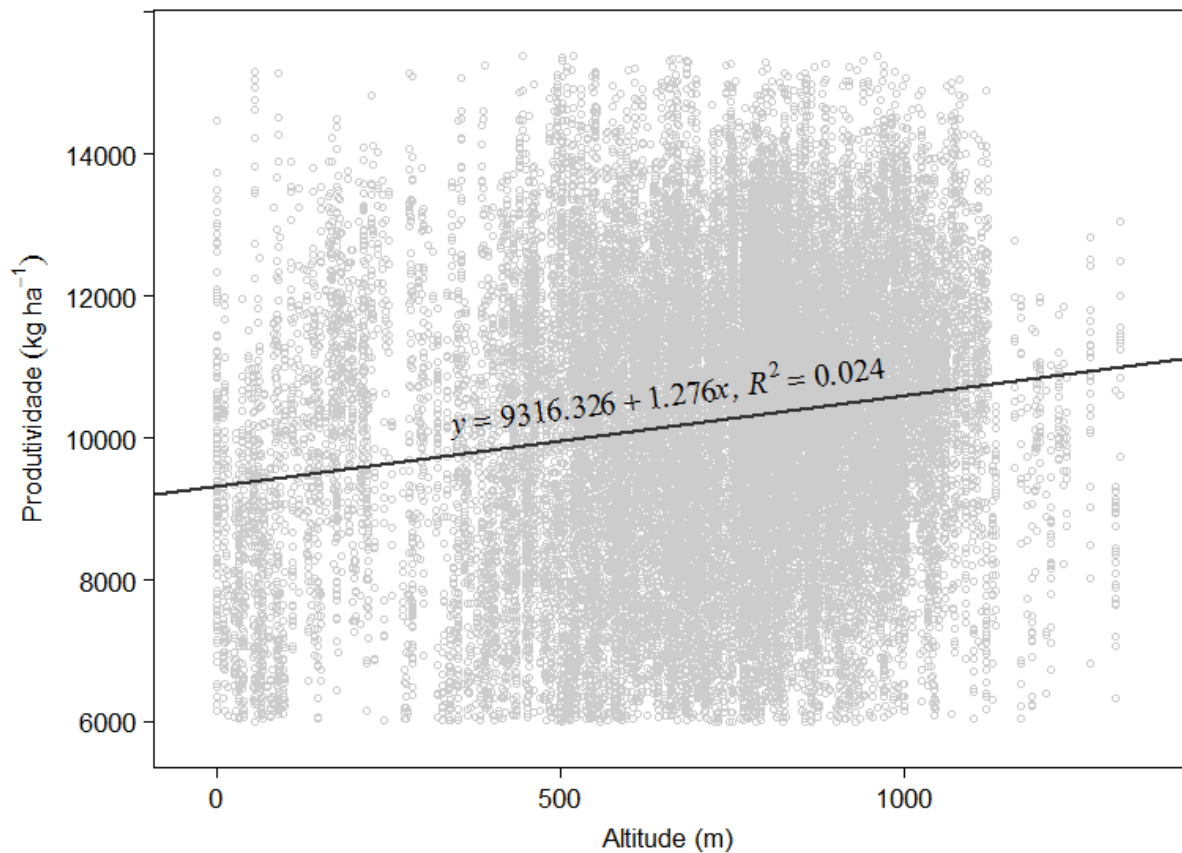


Figura 5 - Comportamento da produtividade de grãos de milho na região Sul do Brasil, em função da altitude que os ensaios foram conduzidos.

A análise da produtividade de grãos em função da altitude também pode ser verificada de forma estratificada, o comportamento da produtividade em altitudes inferiores e superiores a 700 metros, a altitude de 700 m foi escolhida para esta estratificação, pois é um valor próximo a média dos ensaios, e por ser um valor de referência também para o setor, em termos de altitude onde o comportamento dos genótipos tendem a serem distintos.

Quanto observa-se a comportamento da produtividade, em função da altitude, verifica-se que abaixo de 700 m, que a cada metro, a produtividade de grãos também foi incrementada em $1,37 \text{ kg ha}^{-1} \text{ m}^{-1}$ contra $0,99 \text{ kg ha}^{-1} \text{ m}^{-1}$ na classe de altitude superior, esta variação, mesmo aparentemente pequena, representa uma diferença de aproximadamente 37%, demonstrando que na classe inferior a 700 m, os genótipos tem um resposta mais acentuada a altitude de cultivo.

Na Tabela 7, tem-se os dados de produtividade média ao longo dos anos e número de locais em cada um das classes em estudo, demonstrados de uma maneira descritiva. Quanto a produtividade média da cultura ao longo do estudo, verifica-se que esta foi sempre superior a 8.281 kg ha⁻¹ e inferior a 11.869 kg ha⁻¹, valores bem superiores as médias nacionais de produtividade que é de 5.057 kg ha⁻¹ segundo dados de 2013 (CONAB, 2014).

Tabela 7 - Dados descritivos de médias de produtividade(\bar{x}), e número de locais (NL) estratificando em duas classes de altitude superior e inferior a 700m ao longo dos anos de estudo.

Ano	<700 m \bar{x} (NL)	>700 ml \bar{x} (NL)	Total \bar{x} (NL)
2001	8.281(20)	10.157(34)	9.587(54)
2002	9.131(27)	10.498(38)	9.962(65)
2003	9.729(41)	10.181(44)	9.963(85)
2004	9.432(29)	9.714(64)	9.636(93)
2005	8.670(100)	9.687(146)	9.313(246)
2006	9.602(143)	10.072(223)	9.894(366)
2007	9.214(133)	10.368(227)	9.957(360)
2008	9.025(155)	9.601(239)	9.383(394)
2009	10.677(164)	11.011(188)	10.857(352)
2010	11.284(124)	11.120(139)	11.196(263)
2011	9.748(45)	10.777(69)	10.399(114)
2012	10.813(92)	11.584(133)	11.278(225)
2013	10.861(118)	11.869(91)	11.332(209)
Média/Total	9.932(1.191)	10.437(1.635)	10.233(2.826)

3.4.2 Resultado das predições

Para cada classe, foi ajustado um modelo misto com o objetivo de verificar o ganho genético, também foi feito o teste de falta de ajuste do modelo, para testar a significância dos componentes de variância do modelo, via teste de razão de verossimilhança. Na Tabela 8, são apresentados os resultados dos testes de razão de verossimilhança, verifica-se que para todos as classes, a adição dos efeitos aleatórios de ano (a), empresa (b), e genótipo foram significativos a 95% de probabilidade, demonstrando assim a importância destes componentes no modelo.

Também nesta tabela, são apresentados os dados dos parâmetros estimados, pelos modelos, onde foram divididos em efeitos fixos e aleatórios, para todos os modelos todos os parâmetros foram significativos a 5% de probabilidade, demonstrando assim que estes parâmetros são importantes para explicar a produtividades de grãos de milho na região sul do Brasil. Quanto aos componentes

de variância, verifica-se que os locais são os que maior tem participação na variância total, valor este que foi de 60,2% e 63,4% do total da variância, isso indica que mesmo dentro de cada uma das classes estudadas, ainda existe uma variação bastante grande entre os locais dos ensaios. Depois dos locais, o componente mais importante foi o erro, que foi de 16,8% e 20,4%, demonstrando que mesmo tendo experimento distribuído por uma série de locais, o erro ainda foi o segundo maior componente superior ao componente de genótipos que é o mais interessante.

O genótipo teve sua participação na variância total de 6,2% e 10,5 % do total da variância, valor bastante inferior se comparado aos outros componentes, e também quando analisa-se que este é o componente de maior interesse no estudo. O ano teve sua participação na variância total de 8,3% a 12,5%, valores bem similares aos valores dos genótipos. Já o componente das empresas ficou entre 0,6% a 1%, ambos significativos, porém, demonstrando uma importância pequena na variância total, desta forma, pôde-se inferir que entre os germoplasmas de empresas diferentes, os genótipos finais lançados por estes programas de melhoramento não têm uma grande diferença de produtividade entre os mesmos.

Tabela 8 - Componentes de variância dos modelos ajustado para cada condição de estratificação, sendo θ o parâmetro, $\hat{\theta}$ o valor estimado para o parâmetro, σ o desvio padrão do parâmetro, % σ^2 percentual total da variância e valor-p para o teste de significância, GL o número de graus de liberdade aproximados por Satterthwaite, t o valor t calculado, e valor-p sob hipótese de H_0 , onde a estimativa seria igual a 0.

Condição	Efeitos Fixos						Efeitos Aleatórios				
	θ	$\hat{\theta}$	σ	GL	t	valor-p	θ	$\hat{\theta}$	σ	% σ^2	valor-p
Inferior a 700m	μ	6.504,3	321,2	50	29,3	< 0,01	l	2.538.383	1.593,2	63,4	-
	al	1,4	0,3	2.808	10,7	< 0,01	g	249.745	499,7	6,2	< 0,01
	n	6,7	1,3	2.810	8,0	< 0,01	a	500.544	707,5	12,5	< 0,01
	p	7,4	1,9	2.798	5,0	< 0,01	b	38.334	195,8	1,0	< 0,01
	k	5,3	1,6	2.796	5,8	< 0,01	e	673.958	820,9	16,8	< 0,01
		μ	6.838,0	375,3	50	29,3	< 0,01	l	2.016.035	1.419,9	60,2
Superior a 700m	al	1,5	0,3	2.808	10,7	< 0,01	g	350.971	592,4	10,5	< 0,01
	n	6,5	1,0	2.810	8,0	< 0,01	a	276.219	525,6	8,3	< 0,01
	p	4,3	1,6	2.798	5,0	< 0,01	b	20.431	142,9	0,6	< 0,01
	k	5,8	1,1	2.796	5,8	< 0,01	e	683.141	826,5	20,4	< 0,01

μ : Constante geral do modelo, al: efeito da altitude, n: efeito da dose de nitrogênio, p: efeito da doses de fosforo, k: efeito da doses de potássio, l: efeito do local, g: efeito do genótipo, a: efeito do ano, b: efeito da empresa, e: efeito do erro.

Os ganhos genéticos calculados são apresentados na Tabela 9, onde pode-se visualizar cada ambiente, bem como cada ano, e um resumo geral do ambiente ao longo do ano.

Para a altitude inferior a 700m, verifica-se que o ganho genético médio foi de 74 kg ha⁻¹ ano⁻¹, e estatisticamente significativo (valor-p < 0,05). Durante este período o valor de ganho variou de -211,6 kg ha⁻¹ a 188,7 kg ha⁻¹. Em condições de altitude superior a 700m, o ganho ao longo dos anos foi de 94 kg ha⁻¹ ano⁻¹, sendo estatisticamente superior (valor-p < 0,05).

Observando os resultados para as classes, verifica-se que o ganho foi maior em altitudes superiores a 700m, este fato pode estar associado ao melhoramento que foi direcionando e conduzido mais intensamente nestas condições como esta, desta forma as regiões abaixo desta altitude, não ficaram como preferenciais para o melhoramento. Outro aspecto que pode estar relacionado a este ganho maior em altitudes mais elevada, e devido ao fato destes locais possuírem condições ambientais, mais propícias para os genótipos que foram melhorados ao longo destes anos.

Tabela 9 - Resultado do ganho via modelos mistos, em diferentes classes, sendo u_1 e u_2 , ganho genético do ano 1 e 2 respectivamente, V_{u1} e V_{u2} variância do ganho genético no ano 1 e 2 respectivamente, G_g ganho genético, V_{Gg} variância do ganho genético, valor-p é o valor-p sob o teste da diferença do ganho em relação a 0, IC intervalo de confiança a 95% de probabilidade do ganho genético.

Ambiente	Biênio	u_1 (kg ha ⁻¹)	u_2 (kg ha ⁻¹)	V_{u1}	V_{u2}	G_g (kg ha ⁻¹)	V_{Gg}	valor-p	IC _(95%) G_g (kg ha ⁻¹)
Inferior a 700 m	2001/2002	-392	-226	4.885	4.026	166	2.380	< 0,01	[74;264]
	2002/2003	-226	-170	4.026	3.756	56	1.568	0,16	[-24;128]
	2003/2004	-170	-16	3.756	3.682	154	1.495	< 0,01	[78;230]
	2004/2005	-16	-224	3.682	2.966	-208	1.468	< 0,01	[-279;-130]
	2005/2006	-224	-159	2.966	2.689	64	533	< 0,01	[21;112]
	2006/2007	-159	-19	2.689	3.132	140	791	< 0,01	[85;192]
	2007/2008	-19	67	3.132	2.731	87	682	< 0,01	[35;136]
	2008/2009	67	236	2.731	2.785	168	1.590	< 0,01	[-72;245]
	2009/2010	236	335	2.785	3.152	99	1.080	< 0,01	[37;163]
	2010/2011	335	521	3.152	3.748	186	1.832	< 0,01	[99;266]
	2011/2012	521	310	3.748	3.329	-211	1.398	< 0,01	[-287;-142]
	2012/2013	310,1	498	3.329	4.414	188	1.728	< 0,01	[112;273]
	Geral	-	-	-	-	74	48	< 0,01	[61;88]
Superior a 700 m	2001/2002	-645	-368	4.131	3.962	277	641	< 0,01	[220;321]
	2002/2003	-368	-178	3.962	3.554	190	1.348	< 0,01	[116;260]
	2003/2004	-178	-85	3.554	3.316	92	477	< 0,01	[47;133]
	2004/2005	-85	-236	3.316	3.214	-150	604	< 0,01	[-199;-104]
	2005/2006	-236	-149	3.214	3.117	86	463	< 0,01	[45;128]
	2006/2007	-149	-36	3.117	3.200	112	488	< 0,01	[70;155]
	2007/2008	-36	126	3.200	3.132	163	533	< 0,01	[118;207]
	2008/2009	126	282	3.132	3.096	155	858	< 0,01	[-72;211]
	2009/2010	282	407	3.096	3.511	124	945	< 0,01	[68;188]
	2010/2011	407	540	3.511,	3.438	133	584	< 0,01	[85;180]
	2011/2012	540	358	3.438	3.624	-182	614	< 0,01	[-232;-135]
	2012/2013	358	487	3.624	4.312	129	1.134	< 0,01	[59;190]
	Geral	-	-	-	-	94	27	< 0,01	[83;104]

Os resultados de ganhos genéticos são similares aos encontrados por outros pesquisadores ao redor do mundo, como Duvick (2005), nos Estados Unidos, com dados de 1934 a 2004, calculou um ganho de 65 a 75 kg ha⁻¹ ano⁻¹. Eyherabide & Damilano (2001) na Argentina, analisando entre os anos de 1979 a 1998, calculou um ganho de 83,2 kg ha⁻¹.ano⁻¹, Ci et al (2010), na China entre os anos de 1970 a 2010, constataram um ganho genético de 94,7 kg ha⁻¹ ano⁻¹, Apraku et al. (2014), na região oeste e central da África, em dois anos, chegaram a ganho de 40 kg ha⁻¹.ano⁻¹, e Fernandes e Franzon (1997), em estudo no Brasil com dados de 30 anos de 1963 a 1993, calcularam um ganho de 59,9 kg ha⁻¹ ano⁻¹, apesar das similaridades de valores encontrados, cada um destes trabalhos acima citados trabalharam com diferentes metodologias, em não possuíram a mesmas condições de genótipos em ambientes.

Verifica-se que mesmo que um programa de melhoramento tem o foco em um produto para uma determinada área geográfica, dentro da mesma, ocorrem inúmeras variações de respostas dos genótipos, e o posicionamento certo dos melhores genótipos nos melhores locais, pode em muito elevar a produtividade de grãos de milho.

3.5 Conclusão

O ganho genético foi diferenciado em cada classe analisada, verifica-se que em altitude superiores a 700m o ganho foi de 94 kg ha⁻¹ ano⁻¹, já para altitude inferior a 700 m o ganho foi de 74 kg ha⁻¹ ano⁻¹.

Apesar das diferenças apresentadas entre as classes, pode-se afirmar que os programas de melhoramento na região Sul do Brasil para os dados de produtividade de grãos estão tendo um ganho genético estatisticamente significativo a 5% de probabilidade ao longo do período de 2001 a 2013.

4 DISCUSSÃO

Os dois capítulos propostos nesta dissertação buscam explorar melhor as metodologias de cálculos de ganho genético para a produtividade de milho na região Sul do Brasil.

No primeiro capítulo, tem-se uma comparação entre uma metodologia clássica utilizada no Brasil para o cálculo do ganho genético que foi proposta por Vencovsky et al. (1988), e que leva em consideração a inferência via mínimos quadrados. Já a segunda metodologia, a de Borges et al. (2009) é baseada em modelagem via modelos mistos com inferência via verossimilhança restrita (REML), com o intuito de calcular os BLUP para genótipos, esta última metodologia traz uma flexibilidade maior de modelagem, a qual permite adicionar covariáveis que visam melhorar as estimativas realizadas, além propiciar outras informações deste modelo, como por exemplo o efeito das covariáveis sobre a produtividade.

Pelas duas metodologias o ganho genético foi estatisticamente significativo a 5% de probabilidade ao longo dos anos do estudo, demonstrando que os programas de melhoramento estão contribuindo com genótipos melhores, que possuem maior produtividade de grãos de milho.

No segundo capítulo, é utilizada a metodologia de Borges et al. (2009) para cálculo de ganho genético em diferentes classes dentro da região Sul, a fim de verificar se há diferenças em função de diferentes altitudes. Em todas as classes analisadas dentro da região Sul do Brasil, também encontrou-se um ganho genético estatisticamente significativo a 5% para o ganho genético, demonstrando que os avanços trazidos pelos novos genótipos ao longo dos anos do estudo, contribuíram para a produtividade nos diferentes locais em condições de cultivo.

Observou-se maior flexibilidade da metodologia REML/BLUP para a modelagem de dados de melhoramento genético. Portanto, esta dissertação pode contribuir para a difusão e uso desta metodologia, que deve ser sempre preferida se comparada à modelagem via mínimos quadrados em modelos complexos como os apresentados nesta dissertação.

REFERÊNCIAS

- ASSOCIAÇÃO BRASILEIRA DE SEMENTES E MUDAS. Anuário Abrasem 2011. Brasília/DF: ABRASEM. Disponível em: <http://www.abrasem.com.br/wp-content/uploads/2013/09/Anu%C3%A1rio-Abrasem-2014.pdf>. Acesso em: 10 jan. 2015.
- ABREU, A. F. B. et al. Progresso do melhoramento genético do feijoeiro das décadas de setenta e oitenta nas regiões Sul e Alto do Paranaíba em Minas Gerais. **Pesquisa Agropecuária Brasileira**, v. 29, n. 1, p. 105-112, 1994.
- APRAKU, B. et al. Gains in Grain Yield of Early Maize Cultivars Developed During Three Breeding Eras under Multiple Environments. **Crop Science**, v. 55, 2015.
- ATROCH, A. L.; NASCIMENTO FILHO, F. J.; RESENDE, M. D. V.; LOPES, R.; CLEMENT, C. R. Predição de valores genéticos na fase juvenil de progênies de meios irmãos de guaranazeiro. **Revista de Ciências Agrárias**, Belém, v. 54, p. 73-79, 2011.
- BAKSH, A. et al., Spatio-temporal analysis of yield variability for a corn-soybean field in Iowa, **ASAE**, v. 43, n. 31, 2000.
- BARDORFF-NIELSEN, O. Plausibility inference. **Journal of the Royal Statistical Society**, Series B, v. 38, p. 103-131, 1976.
- BATES, D.; MAECHLER, M.; BOLKER, B. **lme4**: Linear mixed-effects models using Eigen and syntax. R package version 0.999999-0. 2012.
- BORÉM, A.; MIRANDA, G. V. **Melhoramento de plantas**. 6. ed. Viçosa: UFV, 2013. 523p.
- BORGES, V. et al. Progresso genético do programa de melhoramento de arroz de terras altas de Minas Gerais utilizando modelos mistos. **Revista Brasileira de Biometria**, v. 27, p. 478-490, 2009.
- _____. Desempenho genotípico de linhagens de arroz de terras altas utilizando metodologia de modelos mistos. **Bragantia**, v. 69, n. 4, p. 833-841, 2010.
- BRESEGHELLO, F.; MORAIS, O. P.; RANGEL, P. H. N. A new method to estimate genetic gain in annual crops. **Genetics and Molecular Biology**, v. 21, n. 4, p. 551-555, 1998.
- CASELLA, G. E.; BERGER, R. L. **Inferência estatística**. Cengage Learning, 2010.
- CI, X. et al. Genetic Contribution to Advanced Yield for Maize Hybrids Released from 1970 to 2000 in China. **Crop science**, v. 51, n. 1, p. 13-20, 2010.
- CONAB. **Acompanhamento da Safra Brasileira de Grãos**. v. 2. Safra 2014/15. n. 3. Terceiro Levantamento. Disponível em: http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14_12_10_08_51_33_boletim_graos_dezembro_2014.pdf. Acesso em: 15 dez. 2014.

- COSTA, R. B.; RESENDE, M. D. V.; ROA, R. A. Melhoramento genético de ervamate nativa do estado de Mato Grosso do Sul. **Bragantia**, São Paulo, v. 68, p. 611-619, 2009.
- _____. GONÇALVES, P. S. Predição de parâmetros e valores genéticos para caracteres de crescimento e produção. **Bragantia**, São Paulo, v. 69, p. 49-56, 2010.
- CRUZ, J. C.; FILHO, A. P.; QUEIROZ, L. R. Quatrocentas e sessenta e sete cultivares de milho estão disponíveis no mercado de sementes do Brasil para a safra 2013/14. **Embrapa**. Disponível em: <<http://www.cnpms.embrapa.br/milho/cultivares/>>. Acesso em: 15 dez. 2014.
- CULLIS, B. R.; GLEESON, A. C. Spatial analysis of field experiments - Na extension to two dimensions. **Biometrics**, v. 47, p. 1449-1460, 1991.
- DEEPAYAN, S.; ANDREWS, F. **latticeExtra**: Extra Graphical Utilities Based on Lattice. R package version 0.6-26, 2013.
- DUARTE, J. B.; VENCOSKY, R. Estimação e predição por modelo linear misto com ênfase na ordenação de médias de tratamentos genéticos. **Scientia Agricola**, Piracicaba, v. 58, n. 1, p. 109-117, 2001.
- DUVICK D.N. Genetic progress In yield of United States maize (*Zea mays* L.). **Maydica**, v. 50. p. 193-202, 2005a.
- _____. The contribution of breeding to yield advances in maize (*Zea mays* L.). **Adv. Agron.** v.86, p. 83–145. 2005b.
- _____. SMITH J.S.C; COOPER M. Long-term selection in a commercial hybrid maize breeding program. **Plant Breeding Review**. v. 24, n. 2, p. 109–150, 2004.
- ECHARTE, L. et al. Grain Yield Determination and Resource Use Efficiency in Maize Hybrids Released. **Agricultural Chemistry**, 2013. 222p.
- EDWARDS, A. W. F. **Likelihood**. Cambridge: Cambridge University Press, 1972. 235p.
- EYHERABIDE, G. H.; DAMILANO, A. L. Comparison of genetic gain for grain yield of maize between the 1980s and 1990s in Argentina, **Maydica**, v. 46, n. 4, p. 277-281. 2001.
- FARIAS NETO, J. T.; RESENDE, M. D. V. Aplicação da metodologia de modelos mistos (REML/BLUP) na estimação de componentes de variância e predição de valores genéticos em pupunheira (*Bactris gasipaes*). **Revista Brasileira de Fruticultura**, Jaboticabal, v. 23, n. 2, p. 320-324, 2001.
- FARIAS NETO, J. T.; RESENDE, M. D. V.; OLIVEIRA, M. S. P.; NOGUEIRA, O. L. Estimativas de parâmetros genéticos e ganho de seleção em progênies de polinização aberta de açaizeiro. **Revista Brasileira de Fruticultura**, v. 30, p. 1051-1056, 2008.

- FARIAS NETO, J. T.; RESENDE, M. D. V.; OLIVEIRA, M. S. P. Seleção simultânea em progênies de açaizeiro irrigado para produção e peso do fruto. **Revista Brasileira de Fruticultura** (Impresso), v. 33, p. 532-539, 2011.
- FERNANDES, J. S. C. Estabilidade ambiental de cultivares de milho na região centro sul do Brasil. 1988. 94f. Dissertação (Mestrado) - Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz", São Paulo, 1988.
- FERNANDES, J. S. C.; FRANZON, J. F. Thirty years of genetic progress in maize (*Zea mays* L.) in a tropical environment. **Maydica**, v. 42, p. 21-27, 1997.
- FONSECA JÚNIOR, N. S. DA. Progresso genético na cultura do feijão no estado do Paraná para o período de 1977 a 1995. Tese (Doutorado em genética e melhoramento de plantas)-**Universidade de São Paulo**, Escola Superior de Agricultura "Luiz de Queiroz", 160p. 1997.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **Value of Agricultural Production**. Disponível em: <<http://faostat.fao.org>>. Acesso em: 17 maio 2013.
- FRITSCHÉ-NETO, R.; RESENDE, M. D. V.; MIRANDA, G. V.; VALE, J. C. Seleção genômica ampla e novos métodos de melhoramento do milho. **Revista Ceres**, v. 59, p. 794-802, 2012.
- GOMES, M. I. Verossimilhança e inferência estatística. In: Colóquio de estatística e investigação operacional, **Actas**, Rio de Janeiro. Universidade federal do Rio De Janeiro, 1981. p. 230-246.
- HARTLEY, H. O.; RAO, J. N. K. Maximum likelihood estimation for the mixed analysis of variance model, **Biometrika**, v. 54, p. 93-108. 1967.
- IBGE. INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA-**Produção agrícola municipal**. SIDRA. Disponível em www.sidra.ibge.gov.br. Acesso: maio 2013.
- LINDSEY, J. K. Some Statistical heresis. **Journal of the Royal Statistical Society**, Series B, v. 48, p. 1-40, 1999.
- MANFIO, C. E.; MOTOIKE, S. Y.; RESENDE, M. D. V.; SANTOS, C. E. M. Avaliação de progênies de macaúba na fase juvenil e estimativas de parâmetros genéticos e diversidade. **Pesquisa Florestal Brasileira** (Online), v. 32, p. 63-68, 2012.
- MORAIS, O. P.; ABBUD, N. S. Subsídios para avaliação do progresso genético dos programas estaduais de melhoramento de arroz no Brasil. **EMBRAPA- CNPAF**, 26p. 1993.
- OLIVEIRA, R.; DAROS, E.; BESPALHOK FILHO, J. C.; ZAMBON, J. L.; IDO, O.; RESENDE, M. D. V.; ZENI NETO, H. Seleção de famílias de cana-de-açúcar via modelos mistos. **Scientia Agraria** (UFPR), v. 9, p. 269-274, 2008.
- PATTERSON, H. D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v. 58, p. 545-554, 1971.

PEREIRA, B. B. **Inferência verossimilhança**. Boletim da associação Brasileira de estatística, v. 38, p. 31-42, 1997.

R CORE TEAM. R. A language and environment for statistical computing. **R Foundation for Statistical Computing**, 2014.

RANERE, A. J. et al. The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. **Proc Natl Acad Sci**, v. 106, p. 5014-5018. 2009.

RESENDE, M. D. V. et al. Melhor predição linear não viciada (BLUP) de valores genéticos no melhoramento de Pinus. **Boletim de Pesquisa Florestal**, v. 32/33, p. 3-22, 1996.

RESENDE, M. D. V. **Genética biométrica e estatística no melhoramento de plantas perenes**. Embrapa Informação tecnológica, 975 p. 2002.

_____. **Métodos estatísticos ótimos na análise de experimentos de campo**. Embrapa Florestas, 65 p. 2004.

_____. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Embrapa Florestas, 561 p. 2007.

ROBINSON, G. K. That BLUP is a Good Thing: The Estimation of Random Effects, **Statistical Science**, v. 6, n. 1, p. 15-32. 1991.

SAMPAIO, P. T. B.; RESENDE, M. D. V.; ARAÚJO, A. J. Estimativas de parâmetros genéticos e métodos de seleção para o melhoramento genético de *Pinus oocarpa* Schiede. **Pesquisa Agropecuária Brasileira**, v. 37, n. 5, p. 625-636, 2002.

SANTOS, V. B. DOS. Avaliação genotípica de linhagens de arroz de terras altas via metodologia de modelos mistos. Tese (doutorado)- **Universidade Federal de Lavras**. 153 p. 2009.

SARKAR, D.; ANDREWS, F. **latticeExtra: Extra Graphical Utilities Based on Lattice**. R package version 0.6-24. 2012.

SEARLE, S. R. Built in restrictions on best linear unbiased predictions (BLUP) of random effects in mixed models. **The American Statistician**, v. 51, n. 1, p. 19-21, 1997.

SEARLE, S. R.; CASELLA, G.; McCULLOCH, C. E. **Variance componentes**. New York; J Wiley, 528 p. 1992

STIGLER, S. M. Gauss and the Invention of Least Squares. **Ann. Statistics**. v. 9, n. 3, p. 465–474. 1981.

STORCK, L.; BISOGNIN, D. A.; CARGNELUTTI FILHO, A. Ganho genético decorrente da substituição anual de cultivares de milho. **Pesquisa Agropecuária Brasileira**, v. 40, n. 9, p. 881-886. 2005.

VENCOVSKY, R. et al. Progresso genético em vinte anos de melhoramento do milho no Brasil. Congresso nacional de Milho e Sorgo. 16. ed., 1986, Belo Horizonte. **Anais...** Sete Lagoas: Embrapa - CNPMS, p. 300-307. 1988.

VIANA, J. M. S.; ALMEIDA, R. V.; FARIA, V. R.; RESENDE, M. D. V.; SILVA, F. F. E. Genetic evaluation of inbred plants based on BLUP of breeding value and general combining ability. **Crop & Pasture Science** (Print), v. 62, p. 515-522, 2011a.

_____. Best linear unbiased prediction and family selection in crop species. **Crop Science**, v. 51, p. 2371-2381, 2011b.

VIANA, J. M. S.; VALENTE, M. S. F.; SCAPIM, C. A.; RESENDE, M. D. V.; SILVA, F. F. E. Genetic evaluation of tropical popcorn inbred lines using BLUP. **Maydica** (Bergamo), v. 56, p. 273-281, 2011c.

ANEXO A – Código em R para a metodologia de Vencovsky et al. (1988)

```
# Função para cálculo de ganho genético pela metodologia de Vencovsky et al. (1988)
# parametros da função:
# dados; Data frame com os dados para o cálculo, conforme exemplo abaixo
## ANO GENOTIPO LOCAL PRODUTIVIDADE
## 2001 G1 L1 8000
## 2001 G1 L2 8135
## 2002 G2 L1 9001
## ... ... ...
## 2013 G135 L2893 10
# ANO: Nome da coluna na base de dados que tem a informação do ano
# GENOTIPO: Nome da coluna na base de dados que tem a informação do genótipo
# LOCAL: Nome da coluna na base de dados que tem a informação do local
# PRODUTIVIDADE: Nome da coluna na base de dados que tem a informação de produtividade
# Diagnos_file: Nome do arquivo para salvar informações de diagnostico do modelo
g_ven <- function(dados,ANO='ANO',GENOTIPO='GENOTIPO',LOCAL='LOCAL',
  PRODUTIVIDADE='PRODUTIVIDADE',diagnos_file='diagnostico'){
  dados <- dados[,c(ANO,GENOTIPO,LOCAL,PRODUTIVIDADE)]
  colnames(dados) <- c('ANO','GENOTIPO','LOCAL','PRODUTIVIDADE')
  res_out <- list(); con <- 1
  pdf(paste(diagnos_file,'.pdf',sep=""))
  for(i in sort(unique(dados$ANO))[-1]){
    dados_s_bi <- dados[dados$ANO %in% c(i-1,i),]
    dados_s_bi_suma <- cast(dados_s_bi,GENOTIPO~ANO,value='PRODUTIVIDADE',fun='length')
    dados_s_bi$STATUS <- NA
    dados_s_bi$STATUS[dados_s_bi$GENOTIPO %in%
      dados_s_bi_suma[dados_s_bi_suma[,2]>0 &
        dados_s_bi_suma[,3] >0,1]] <- 'AMBO'
    dados_s_bi$STATUS[dados_s_bi$GENOTIPO %in%
      dados_s_bi_suma[dados_s_bi_suma[,2]>0 &
        dados_s_bi_suma[,3]==0,1]] <- 'ANO1'
    dados_s_bi$STATUS[dados_s_bi$GENOTIPO %in%
      dados_s_bi_suma[dados_s_bi_suma[,3]>0 &
        dados_s_bi_suma[,2]==0,1]] <- 'ANO2'
    Y1 <- mean(dados_s_bi[dados_s_bi$ANO==(i-1),'PRODUTIVIDADE'])
    Y2 <- mean(dados_s_bi[dados_s_bi$ANO==(i),'PRODUTIVIDADE'])
    N1 <- length(dados_s_bi[dados_s_bi$ANO==(i-1),'PRODUTIVIDADE'])
    N2 <- length(dados_s_bi[dados_s_bi$ANO==(i),'PRODUTIVIDADE'])
    YC1 <- mean(dados_s_bi[dados_s_bi$ANO==(i-1) &
      dados_s_bi$STATUS=='AMBO','PRODUTIVIDADE'])
    YC2 <- mean(dados_s_bi[dados_s_bi$ANO==(i) &
      dados_s_bi$STATUS=='AMBO','PRODUTIVIDADE'])
    Gg <- (Y2-Y1)-(YC2-YC1)
    nc <- length(unique(dados_s_bi[dados_s_bi$STATUS=='AMBO','GENOTIPO']))
    nc1 <- length(unique(dados_s_bi[dados_s_bi$STATUS=='AMBO','GENOTIPO']))+

```



```

length(unique(dados_s_bi[dados_s_bi$STATUS=='ANO1','GENOTIPO']))
nc2 <- length(unique(dados_s_bi[dados_s_bi$STATUS=='AMBO','GENOTIPO']))+
length(unique(dados_s_bi[dados_s_bi$STATUS=='ANO2','GENOTIPO']))
TM <- round(nc/(nc1+nc2-nc)*100,2)
TS <- round(((nc2-nc)/nc2)*100,2)
m1 <- (aov(PRODUTIVIDADE~LOCAL+GENOTIPO,data=dados_s_bi[dados_s_bi$ANO==(i-1),]))
par(mfrow=c(2,2))
plot(m1,main=i-1)
par(mfrow=c(1,1))
hist(m1$res,prob=T,xlab='Resíduos',ylab='Frequência',main='')
lines(density(m1$res),col='blue',lwd=2)
m1_sha <- shapiro.test(m1$res)$p
m1_bar <- bartlett.test(m1$res~m1$model$GENOTIPO)$p.value
m1 <- anova(m1)
QMe1 <- m1[3,'Mean Sq'];DFe1 <- m1[3,'Df']
m2 <- (aov(PRODUTIVIDADE~LOCAL+GENOTIPO,data=dados_s_bi[dados_s_bi$ANO==(i),]))
par(mfrow=c(2,2))
plot(m2,main=i)
par(mfrow=c(1,1))
hist(m2$res,prob=T,xlab='Resíduos',ylab='Frequência',main='')
lines(density(m2$res),col='blue',lwd=2)
m2_sha <- shapiro.test(m2$res)$p
m2_bar <- bartlett.test(m2$res~m2$model$GENOTIPO)$p.value
m2 <- anova(m2)
QMe2 <- m2[3,'Mean Sq'];DFe2 <- m2[3,'Df']
VGg <- QMe1*((1/nc)-(1/nc1))+QMe2*((1/nc)-(1/nc2))
tcal <- Gg/sqrt(VGg)
pvalue <- ifelse(tcal>=0,1-pt(tcal,DFe1+DFe2),pt(tcal,DFe1+DFe2))
res_out[[con]] <- unlist(list(ANO1=i-1,ANO2=i,Gg=Gg,VGg=VGg,
      pvalue=pvalue,tcal=tcal,
      Y1=Y1,Y2=Y2,YC1=YC1,YC2=YC2,
      QMe1=QMe1,DFe1=DFe1,
      m1_sha=m1_sha,m1_bar=m1_bar,
      QMe2=QMe2,DFe2=DFe2,m2_sha=m2_sha,
      m2_bar=m2_bar,
      nc=nc,nc1=nc1,nc2=nc2,TM=TM,TS=TS,
      N1=N1,N2=N2))
con <- con+1
}
dev.off()
res_out <- ldply(res_out)
res_out1 <- res_out[1,]
res_out1[1,] <- NA
res_out1[1,'ANO1'] <- 'GERAL'; res_out1[1,'ANO2'] <- 'GERAL'
res_out1$Gg <- mean(res_out$Gg)

```

```

res_out1$VGg <- sum(res_out$VGg)/(nrow(res_out)^2)
res_out1$tcal <- res_out1$Gg/(sqrt(res_out1$VGg))
DFgeral <- sum(res_out$DFe1)+res_out$DFe2[nrow(res_out)]
res_out1$pvalue <- ifelse(res_out1$tcal>=0,
  1-pt(res_out1$tcal,DFgeral ),
  pt(res_out1$tcal,DFgeral ))
res_out <- rbind(res_out,res_out1)
res_out$IC_hig <- qnorm(c(0.975),res_out$Gg,sqrt(res_out$VGg))
res_out$IC_low<- qnorm(c(0.025),res_out$Gg,sqrt(res_out$VGg))
res_out <- res_out[,c('ANO1','ANO2','Gg','VGg','tcal','pvalue',
  'IC_hig','IC_low',
  'Y1','Y2','YC1','YC2','QMe1','DFe1',
  'm1_sha','m1_bar',
  'QMe2','DFe2','m2_sha','m2_bar','nc','nc1',
  'nc2','TM','TS','N1','N2')]
return(res_out)
}

```

ANEXO B – Código em R para a metodologia de Borges et al. (2009)

```
# Função para cálculo de ganho genético pela metodologia de Borges et al. (2009)
# Parâmetros da função:
# dados; Data frame com os dados para o cálculo, conforme exemplo abaixo
## ANO GENOTIPO LOCAL PRODUTIVIDADE N P K ELEVACAO EMPRESA
## 2001 G1 L1 8000 60 30 100 45 E1
## 2001 G1 L2 8135 70 50 150 1300 E1
## 2002 G2 L1 9001 60 30 100 45 E2
## ... ... ... ... ...
## 2013 G135 L2893 10 200 134 143 123 E10
# modelo: objeto do tipo formula com o modelo especificado, abaixo variáveis deste modelo utilizado:
## ANO: Nome da coluna na base de dados que tem a informação do ano
## GENOTIPO: Nome da coluna na base de dados que tem a informação do genótipo
## LOCAL: Nome da coluna na base de dados que tem a informação do local
## PRODUTIVIDADE: Nome da coluna na base de dados que tem a informação de produtividade
## N: Nome da coluna na base de dados que tem a informação de nitrogênio
## P: Nome da coluna na base de dados que tem a informação de fósforo
## K: Nome da coluna na base de dados que tem a informação de potássio
## ELEVACAO: Nome da coluna na base de dados que tem a informação de elevação
## EMPRESA: Nome da coluna na base de dados que tem a informação da Empresa.
# nmcmc: numero de interações do MCMC.
# diagfile: Nome do arquivo para salvar informações de diagnostico do modelo
g_like <- function(dados,modelo= as.formula(PRODUTIVIDADE~1+ELEVACAO+N+P+K+(1|LOCAL)+
(1|ANO)+(1|EMPRESA)+(1|GENOTIPO)),nmcmc=50000,diagfile='diag.pdf'){
require(lme4);require(coda);require(MCMCglmm);require(lmerTest)
factors <- unlist(strsplit(as.character(modelo),'+')[[3]])
response <- unlist(strsplit(as.character(modelo),'+')[[2]])
con <- 1
for(i in factors[grepl('[]',factors)]){
if(con==1){modeloi <- paste(paste(factors[1:(min(grepl('[]',factors))-1]),collapse='+'),i,sep='+')}
else{ modeloi <- paste(modeloi,i,sep='+')}
eval(parse(text=(paste('model',con,'<-lmer(',gsub(' ','',paste(response,'~',modeloi)),
',data=dados',sep=''))))
con <- con+1
}
cat('###-----\n')
print(eval(parse(text=paste('anova(',paste(paste('model',1:length(grepl('[]',factors)),
sep=''),collapse=',','),sep=''))))
m_reml <- lmer(modelo,data=dados)
cat('###-----\n')
print(summary(m_reml))
print(lmerTest::anova(lmerTest::lmer(modelo,data=dados)))
mc <- MCMCglmm(
fixed=eval(parse(text=paste(response,'~',paste(gsub(' ','',factors[-grepl('[]',factors)]),
collapse='+')))),
```

```

random=eval(parse(text=paste('~',paste(gsub('[ (1)]',",",factors[grep('[]',factors)]),
collapse = '+'))),
data=dados,burnin=nmcmc*0.1,nitt=nmcmc,pr=TRUE,thin=10,verbose=FALSE)
summary(mc)
mc1 <- as.mcmc(cbind(as.matrix(mc$Sol),as.matrix(mc$VCV)))
pdf(diagfile)
print(qqmath(ranef(m_reml, postVar = TRUE),
par.settings= standard.theme("pdf", color=FALSE)))
plot(mc1,ask=FALSE)
dev.off()
umcmc <- as.data.frame((mc$Sol))
#colnames(umcmc) <- unlist(lapply(ranef(m_reml),row.names))
colnames(umcmc) <- gsub('GENOTIPO.',",",colnames(umcmc))
umcmc <- umcmc[,rownames((ranef(m_reml)$GENOTIPO))]
cat('###-----\n')
dadosM <- unique(dados[,c('ANO','GENOTIPO')])
res_out_dis <- res_out <- list()
con <- 1
for(i in sort(unique(dados$ANO))[-1]){
con_mcmc <- as.mcmc(rowMeans(umcmc[,dadosM[dadosM$ANO==(i),'GENOTIPO']])-
rowMeans(umcmc[,dadosM[dadosM$ANO==(i-1),'GENOTIPO']]))
#Ggmcmc <- summary(con_mcmc)$statistics[1]
IC <- HPDinterval(con_mcmc)
Ggmcmc_IC_low <- IC[1]
Ggmcmc_IC_hig <- IC[2]
res_out[[con]] <- unlist(list(ANO1=i-1,ANO2=i,
Gg=mean(con_mcmc),
VGg=var(con_mcmc),
pvalue=round(mcmc.pvalue(as.matrix(con_mcmc)),3),
Ggmcmc_IC_low = Ggmcmc_IC_low,
Ggmcmc_IC_hig = Ggmcmc_IC_hig,
u1=mean(rowMeans(umcmc[,dadosM[dadosM$ANO==(i-1),'GENOTIPO']]),
u2=mean(rowMeans(umcmc[,dadosM[dadosM$ANO==(i),'GENOTIPO']]),
Vu1=var(rowMeans(umcmc[,dadosM[dadosM$ANO==(i-1),'GENOTIPO']]),
Vu2=var(rowMeans(umcmc[,dadosM[dadosM$ANO==(i),'GENOTIPO']]))))
res_out_dis[[con]] <- con_mcmc
con <- con+1
}
res_out <- ldply(res_out)
res_out[nrow(res_out)+1,c('ANO1','ANO2')] <- 'GERAL'

res_out_dis <- ldply(res_out_dis)
con_mcmc <- as.mcmc(colMeans(res_out_dis))
#Ggmcmc <- summary(con_mcmc)$statistics[1]
IC <- HPDinterval(con_mcmc)

```

```
Ggmcmc_IC_low <- IC[1]
Ggmcmc_IC_hig <- IC[2]
res_out[res_out$ANO1=='GERAL','Gg'] <- mean(con_mcmc)
res_out[res_out$ANO1=='GERAL','VGg'] <- var(con_mcmc)
res_out[res_out$ANO1=='GERAL','Ggmcmc_IC_low'] <- Ggmcmc_IC_low
res_out[res_out$ANO1=='GERAL','Ggmcmc_IC_hig'] <- Ggmcmc_IC_hig
res_out[res_out$ANO1=='GERAL','pvalue'] <- round(mcmc_pvalue(as.matrix(con_mcmc)),3)
res_out <- res_out[,c('ANO1','ANO2','Gg','VGg','pvalue','Ggmcmc_IC_low',
  'Ggmcmc_IC_hig','u1','u2','Vu1','Vu2')]
return(res_out)
}
```