Universidade de São Paulo
Escola Superior de Agricultura "Luiz de Queiroz"

Progresso genético simultâneo: um exemplo de aplicação no melhoramento do tabaco

Fernando Henrique Ribeiro Barrozo Toledo

Tese apresentada para obtenção do título de Doutor em Ciências. Área de concentração: Genética e Melhoramento de Plantas

Piracicaba
2014

Fernando Henrique Ribeiro Barrozo Toledo
Engenheiro Agrônomo

Progresso genético simultâneo: um exemplo de aplicação no melhoramento do tabaco

Orientador:
Prof. Dr. **ROLAND VENCOVSKY**

Tese apresentada para obtenção do título de Doutor em Ciências. Área de concentração: Genética e Melhoramento de Plantas

Piracicaba
2014

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Toledo, Fernando Henrique Ribeiro Barrozo
Progresso genético simultâneo: um exemplo de aplicação no melhoramento do tabaco / Fernando Henrique Ribeiro Barrozo Toledo. - - Piracicaba, 2014.
74 p: il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2014.

1. Melhoramento vegetal 2. Genética quantitativa 3. Ganho genético 4. Índice de seleção I. Título

CDD 633.71
T649p

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte -O autor"

À memória dos meus avós Toledo, Ademar, Tiba e Nenê e à minha tia Nizia;
e à nova geração da família, minha sobrinha Maria Carolina
DEDICO

AGRADECIMENTOS

Desde 2010 tenho o privilégio do convívio diário com o Professor Roland Vencovsky, uma sumidade da biometria e exemplo de caráter, ao qual devo meus sinceros agradecimentos pela oportunidade na orientação, pela confiança depositada e pelo inestimável suporte intelectual. Estendo meus votos de gratidão ao Professor e amigo Magno Antonio Patto Ramalho, pelos aconselhamentos e pela lição de abnegação e humildade. Não posso deixar de mencionar minhas desculpas à senhora Maria Olávia Vencovsky e senhora Maria Helena Ramalho pelo tempo que meus orientadores despenderam na minha orientação e renunciaram do convívio familiar.

À Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo (ESALQ/USP) e ao programa de pós-graduação em Genética e Melhoramento de Plantas (PPGGMP) pela oportunidade concedida e por prover estrutura intelectual e física fundamental a realização deste trabalho.

À Companhia Souza Cruz por toda a disponibilidade no planejamento, execução e obtenção dos dados que resultam nesse trabalho, em nome do Dr. Carlos Eduardo Pulcinelli, Dr. Adriano Teodoro Bruzi, Dr. Francislei Raposo, Lúcia Scuciato, Ivan Agotani e Gilson Silveira.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão de financiamento e bolsa de estudo no Brasil. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), nas pessoas de Kristien Roepke e Hannah Leite pela atenção dispendida no financiamento e bolsa de estágio de doutoramento no exterior (Ciências sem Fronteiras, processo n° 0497/13–4).

Ao *Centro Internacional de Mejoramiento de Maíz y Trigo* (CIMMYT) pelo valoroso período sabático, em especial ao Dr. José "Pancho" Crossa, a Senhora Enriqueta Sainz e ao amigo Francelino Rodrigues por toda atenção, ensinamentos e acolhida *en México*.

À Bruna Mendes de Oliveira, Guilherme Barbosa Abreu, Ricardo Augusto Ferreira e Thalita Fernanda Abbruzzini pela enriquecedora, rigorosa e dedicada revisão deste documento.

Aos meus pais Maria Amelia e Anselmo Luiz, e ao meu irmão Luiz Felipe, meus primeiros e mais importantes professores; e aos meus familiares e parentes em reconhecimento ao afeto e apoio incondicional.

A Thalita Fernanda Abbruzzini pelo amor, companheirismo e compreensão, estimulando-me a evoluir e persistir em meus ideais.

Aos professores de ontem e hoje, Juarez Gabardo, Vander Melo, Valmiqui Lima, Magno Ramalho, João Cândido, João Bosco, José Airton, César Brasil, Dr^{a}. Ângela Abreu, Julio Bueno, Daniel Ferreira, José Branco, Giancarlo Xavier, Gerhard Bandel, Isaias Geraldi, Claúdio Lopes, Antônio Augusto Garcia, Gabriel Margarido, José Baldin, Décio Barbin, Gerson Mourão, Clarice Demétrio e César de Lima pelos ensinamentos transmitidos; e aos funcionários do Departamento de Genética Cândida Vanderléia de Oliveira "Léia", Maídia Thomaziello, Salete Gaziola, Antonio Gorga "Berdan", Valdir Próspero, Carlos Macedônio, Carmo de Campos, Benedito Soares "Pimpo", Fernando Leopoldino, Sr. Antônio Rasera, Marcos Fernandes e Rogério Marim pela prestatividade e convívio amistoso.

Aos colegas de laboratório e amigos pós-graduandos Bruna Mendes de Oliveira, Bruno Marco de Lima, Evandro Wagner Tambarussi, Júlio Cesar Mistro, Marcela Mendes, Melina Andrade, Mônica Ferreira, Adriana Gianotto, Carina Anoni, João Ricardo Rosa e Marcelo Mollinari em retribuição as enriquecedoras trocas de conhecimento. Aos amigos de Piracicaba Eloana, Diana, Mariana, Carolina, Suzana, Letícia "Eva", Vinícius, Izabela, Leonardo "Gaucho", José "Mano" Pöppl, Rafael Battisti, Maurício Cherubin, Renes Pinheiro e Matheus Tonini pelos ótimos momentos de descontração.

Aos amigos de Lavras com os quais compartilhei importantes momentos de aprendizado Guilherme Barbosa Abreu, Ricardo Augusto Ferreira, Izabel "Bebel" Figueiredo, Paulo Eduardo Prado, Pedro Henrique Santos, Emiliano Costa, Gustavo Salgado, Matheus Mendes e Walmes Zeviani; e aos amigos de Curitiba *Flopitos*, *Menestréis* e Agrônomos, as quais nomeio: André "Jaca" Neves, Guilherme "Polenta", "Marceleza", Guilherme "Pirula", Miron, Nelson, Brunetti, Vítor "Vj", José Fernando "Afinho", Luís "Fred", Bruno "Butina", Carlos "Casico", Diego, Renan, Vítor "Xina", Gustavo Oliveira, Betina, Faisal, Jonas e Gabriel pela amizade, mesmo que distante.

Após percorrer essa importante etapa, posso compreender a importância de expressar gratidão. Por isso, peço licença para agradecer a todos aqueles, que de alguma forma colaboraram, apoiaram, incentivaram ou torceram para essa realização. Essa conquista é também de vocês, meu honesto Muito Obrigado.

"Eis o homem de Platão."
Diógenes de Sínope

SUMÁRIO

RESUMO 11
ABSTRACT 13
LISTA DE FIGURAS 15
LISTA DE TABELAS 17
1 INTRODUÇÃO 19
2 DESENVOLVIMENTO 21
2.1 Revisão Bibliográfica 21
2.1.1 Metodologia de estimação do Progresso Genético 21
2.1.2 Estimativas do Progresso Genético 25
2.1.3 Metodologias de Índices de Seleção 30
2.2 Material e Métodos 38
2.2.1 Progressos Genéticos Univariados 38
2.2.2 Índice de Seleção com Diferenciais de Seleção Restritos 41
2.2.3 Exemplo de Aplicação 45
2.3 Resultados 46
2.3.1 Progressos Genéticos Univariados 46
2.3.2 Progresso Genético Simultâneo 51
2.4 Discussão 58
2.4.1 Progressos Genéticos Univariados 58
2.4.2 Progresso Genético Simultâneo 62
3 CONSIDERAÇÕES FINAIS 67
REFERÊNCIAS 69

RESUMO

Progresso genético simultâneo: um exemplo de aplicação no melhoramento do tabaco

É inegável a importância do melhoramento genético. Na literatura reporta-se que na ausência do melhoramento a produtividade dos cultivos agrícolas poderia ter diminuído. Deste modo, é fundamental avaliar o progresso genético para otimizar a eficiência da seleção e avaliar as técnicas empregadas. O progresso genético foi estudado em várias culturas no Brasil e no mundo. Apesar de feitas interpretações do relacionamento entre caracteres, não foram encontrados na literatura métodos ou estimativas relacionadas ao progresso simultâneo quando mais de um caráter sofre seleção. Como os melhoristas se atém a seleção de mais de um caráter, ao estimar o progresso genético dever-se-ia considerar esse aspecto. Nesse sentido, vislumbrou-se a possibilidade de se reunir a informação de mais de um caráter, tal qual um índice em que restrições sobre o original de Smith representam os progressos individuais estimados. Isto posto, o objetivo deste trabalho foi estimar os progressos genéticos no melhoramento do tabaco para os caracteres de interesse e apresentar uma alternativa interpretativa de reunir os progressos individuais na estimativa do progresso simultâneo. Mensurou-se o progresso genético dos caracteres agronômicos e morfológicos ao longo de 40 anos de melhoramento de tabaco do grupo varietal *Virginia* pela avaliação de nove linhagens oriundas de quatro fases distintas no delineamento casualizado em blocos com quatro repetições em quatro locais representativos da cultura. Com base nas estimativas e no desenvolvimento teórico do índice com restrições apresentado, reuniram-se as estimativas referentes aos progressos individuais como uma forma de representar o progresso genético simultâneo. A análise de regressão mostrou ganhos significativos por era de 275,29$kg\,ha^{-1}$ para a produtividade de folhas curadas, o que representa um progresso de 10% da média do ciclo inicial. A altura de plantas, número de folhas expandidas, bem como a largura e comprimento das folhas também mostraram incrementos significativos. Devido à interação genótipos por ambientes, os progressos estimados na média dos quatro locais não são bons preditores dos progressos por local. O índice de qualidade geral e a remuneração do produtor não apresentaram os progressos esperados em razão de essas avaliações terem sido realizadas apenas nas fases finais da seleção. A estimativa do progresso simultâneo aferiu um incremento de 6% por era no cômputo dos caracteres avaliados e reflete o sucesso na seleção e melhoramento simultâneo desses caracteres. O índice desenvolvido não necessita arbitrar pesos econômicos, atende às restrições, e mostrou-se eficaz em reunir os progressos genéticos dos caracteres além de poder ser empregado com o propósito da seleção em níveis não independentes de seleção.

Palavras-chave: Melhoramento vegetal; Genética quantitativa; Ganho genético; Índice de seleção

ABSTRACT

Simultaneous genetic progress: long-term results in flue-cured tobacco

The importance of plant breeding is unquestionable. In the literature, it is reported that in the absence of crop improvement crop yield could have decreased. Therefore, it is crucial to evaluate the genetic progress in order to optimize the efficiency of the selection procedures and evaluate the employed techniques. Data on genetic progress were studied throughout the world as well as in Brazil. Besides the understanding of the underlying relationships between traits, there are no current methods or estimates regarding simultaneous progress, i.e., whenever more than one trait undergo selection. As the breeders generally carry out selection considering more than one trait, when estimating the genetic progress, this particular issue should be considered. In this context, it has become possible to summarize information from more than one trait, as an index, based on the classic Smith index with constraints which represent the estimates of individual traits progress. Thus, the objective of the present work was to estimate genetic progress in a tobacco breeding program and to present an interpretative and promising alternative to address the individual progress and hence estimate the simultaneous progress. The genetic progress of agronomic and morphological traits was estimated over a period of 40 years in which selection was carried out for obtaining superior inbred lines of the Virginia variety group. Nine inbred lines, derived from four distinct eras, were evaluated in randomized complete block designs with four replicates in four Brazilian representative sites of the tobacco crop. Based on both the results and theoretical developments of the constrained index, the estimates of individual progress were combined in order to represent the simultaneous genetic progress. The regression analysis showed a significant gain per era of $275.29 kg\, ha^{-1}$ for cured leaf weight, corresponding to a progress of 10% relative to the estimated intercept. Traits such as plant height, expanded leaves, as well as width and length of leaves also showed a significant increase. Given the genotype by environment interaction, the estimated marginal progress of the four sites was not a good predictor of the progress at each site. On the other hand, the leaf quality index and the corresponding commercial value did not show the aimed trend due to the fact that both evaluations were performed at the final trials. The simultaneous progress increased the set of traits by 6% per era relative to the intercept. This reflects the sucess of selection and simultaneous improvement of traits. The proposed index does not require attributing economic weights, satisfies the restrictions, is effective in summarizing the genetic progress of multipletraits and can be applied considering multitrait selection through non independent culling levels.

Keywords: Plant breeding; Quantitative genetics; Genetic gain; Selection Index

LISTA DE FIGURAS

Figura 1 – Representação do progresso genético das médias de um caráter hipotético em diferentes "eras" de melhoramento sob seleção recorrente 21

Figura 2 – Representação de dois caracteres hipotéticos x_1 e x_2 em três diferentes métodos de seleção: (A) seleção em *tandem*; (B) níveis independentes de eliminação; e (C) índice de seleção ($I = \beta_1 x_1 + \beta_2 x_2$) 31

Figura 3 – Progressos genéticos na média dos quatro locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*), número de folhas expandidas (*nfe*), circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*), Brasil 2011/2012 49

Figura 4 – Progressos genéticos por locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*) e número de folhas expandidas (*nfe*), Brasil 2011/2012 51

Figura 5 – Progressos genéticos por locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os caracteres circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*), Brasil 2011/2012 52

Figura 6 – Progressos genéticos na média dos quatro locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os "caracteres índice" índice de qualidade geral (*iqg*), remuneração ao produtor em *US$* por hectare (*$ha*) e para o índice construído (índice), Brasil 2011/2012 56

Figura 7 – Progressos genéticos por locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os "caracteres índice" índice de qualidade geral (*iqg*), remuneração ao produtor em *US$* por hectare (*$ha*) e para o índice construído (índice), Brasil 2011/2012 57

Figura 8 – Correspondência entre as estimativas da herdabilidade antes (h_0^2) e após (h_s^2) a seleção em função da herdabilidade e intensidade de seleção, com e sem redução da variância residual 65

LISTA DE TABELAS

Tabela 1 – Resumo de estimativas do progresso genético publicadas para algumas culturas agrícolas no Brasil e no Mundo 29

Tabela 2 – Esquema da análise de variância, fontes de variação, graus de liberdade e respectivas esperanças dos quadrados médios com desdobramentos complementares ao estudo do progresso genético 41

Tabela 3 – Resultado do testes da homegeneidade de variâncias (K), normalidade dos resíduos (W) e aditividade dos efeitos (F); coeficiente de variação (CV) e acurácia seletiva (Ac) para os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*), número de folhas expandidas (*nfe*), circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*), Brasil 2011/2012 46

Tabela 4 – Análise de variância e desdobramentos complementares ao estudo do progresso genético para os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*) e número de folhas expandidas (*nfe*), Brasil 2011/2012..... 47

Tabela 5 – Análise de variância e desdobramentos complementares ao estudo do progresso genético para os caracteres circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*), Brasil 2011/2012..... 48

Tabela 6 – Correlações fenotípicas na média dos quatro locais, entre os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*), número de folhas expandidas (*nfe*), circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*) 53

Tabela 7 – Resumo das estimativas relativas ao índice reunindo os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*), número de folhas expandidas (*nfe*), circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*) 54

Tabela 8 – Análise de variância e desdobramentos complementares ao estudo do progresso genético para o índice de qualidade geral (*iqg*), para a remuneração paga pelo tabaco em *US$* por hectare (*$ha*) e para o índice construído (índice), Brasil 2011/2012 55

1 INTRODUÇÃO

A importância do melhoramento genético é inquestionável. A partir da segunda metade do século passado realizaram-se alguns trabalhos com o objetivo de mensurar a contribuição relativa do progresso genético frente ao incremento em produtividade de algumas culturas. Reanalisando os dados de avaliação final de cultivares no Reino Unido, ao longo de 58 anos, observou-se que para o trigo e a cevada a importância do melhoramento é tão expressiva que, na sua ausência, a produtividade dessas culturas poderia ter diminuído (MACKAY et al., 2011).

Dada a relevância do melhoramento, é fundamental que se avalie o sucesso obtido, fruto da seleção, e que auxilia no processo de decisão: minimizando perdas, maximizando a eficiência e avaliando as técnicas empregadas ou a serem implementadas. Todavia, os caracteres com que os melhoristas trabalham possuem controle poligênico, assim é virtualmente impossível fixar todos os alelos em um único ciclo seletivo. Portanto, ao longo de ciclos sucessivos de seleção, a estimativa do progresso genético é uma forma de avaliar o acúmulo dos alelos favoráveis nos genótipos com o transcorrer dos ciclos seletivos.

Na literatura reportam-se apenas estimativas de progresso genético dos caracteres de forma independente (univariado). Não foram encontrados relatos em que se obtiveram estimativas do progresso simultâneo dos caracteres de importância agronômica, apesar de em muitos trabalhos serem feitas elocubrações quanto à relação dos progressos de caracteres sob seleção (UNDERDAHL et al., 2008; BRESEGHELLO et al., 2011).

Melhoristas, todavia, quase nunca se atém à seleção de apenas um caráter. Portanto, pela natureza das avaliações praticadas nos programas de melhoramento, indubtavelmente multivariadas, para estimar o progresso genético, dever-se-ia considerar esse aspecto. Deste modo, a estimativa do progresso simultâneo dos caracteres vai de encontro às expectativas dos melhoristas, trazendo um aporte interpretativo interessante. Nesse escopo, descortinou-se a possibilidade de estimar o progresso genético no cômputo dos caracteres sob seleção, que decorre diretamente da possibilidade de se reunir a informação de mais de um caráter, tal qual um índice.

Índices são conhecidos em genética quantitativa e no melhoramento. Fundamentalmente, os índices se propõem a agrupar múltiplas informações, tal que a inferência se baseie em um complexo de variáveis, reunindo vários atributos de importância para o melhoramento (CRUZ; REGAZZI; CARNEIRO, 2001, capítulo 6). Por premissa, no trabalho seminal de Smith, o índice é uma função linear que congrega os caracteres maximizando os ganhos genéticos, que estabelece uma combinação ótima dos vários caracteres e se constitui como um caráter adicional e artificial.

Para o emprego do índice *natural*, faz-se necessário arbitrar pesos econômicos de natureza empírica. Pesek e Baker (1969), na tentativa de mitigar a influência destes pesos econômicos, substitui os referidos, por ganhos desejados pelo melhorista. Deste modo, a importância que o melhorista atribui aos caracteres reflete-se por meio da intensidade de

seleção imposta. Lin (2005b) aprofunda-se em extensa teoria relacionada aos índices restritos, primeiramente descritos por Kempthorne e Nordskog (1959). A inovação de Lin (2005b) é a possibilidade de atribuir múltiplas restrições de diferentes naturezas, inclusive sobre ganhos desejados. O princípio envolvido é estritamente matemático, podendo ser considerado como um caso especial do índice *natural*.

Tendo a estimativa do progresso genético como objetivo, adaptou-se a teoria de índices de seleção que Lin apresentou, no desenvolvimento de um índice em que as restrições são diferenciais de seleção. Das estimativas dos progressos genéticos individuais é possível estimar os diferenciais de seleção "realizados" e esses são agrupados a fim de reunir as informações dos progressos genéticos dos vários caracteres e constituir um "supercaráter" que expresse o progresso genético obtido no melhoramento de múltiplos caracteres, ou progresso simultâneo.

Diante do exposto, o objetivo da presente tese é estimar os progressos genéticos obtidos ao longo de 40 anos de melhoramento genético dos caracteres de interesse do tabaco do grupo varietal *Virginia* no Brasil e apresentar uma alternativa interpretativa, viável, promissora e formal de reunir as informações dos progressos individuais na estimativa do progresso simultâneo.

2 DESENVOLVIMENTO

2.1 Revisão Bibliográfica

2.1.1 Metodologia de estimação do Progresso Genético

Ao falar em progresso genético, deve-se primeiro fazer a distinção entre o progresso esperado, baseado no coeficiente de herdabilidade e no diferencial de seleção, e o progresso obtido, relativo ao ganho em performances dos genótipos no transcorrer dos ciclos recorrentes de seleção (FALCONER; MACKAY, 1996, capítulo 11).

De forma geral, os caracteres com que os melhoristas trabalham possuem controle poligênico. Deste modo, é virtualmente impossível agregar todos os alelos favoráveis, em um dado genótipo em um único ciclo seletivo (RAMALHO et al., 2012, capítulo 10). Ao longo dos ciclos de melhoramento a estimativa do progresso genético é uma forma de se observar o acúmulo desses alelos favoráveis nos genótipos com o transcorrer dos ciclos, por meio da mudança da média fenotípica (FALCONER; MACKAY, 1996, capítulo 11).

A figura 1 representa o progresso genético da média de um caráter hipotético ao longo dos ciclos de seleção recorrente. Observa-se a alteração na média na expressão do caráter ao longo dos ciclos e a variância que propicia a seleção dos melhores indivíduos dentro de cada ciclo, como consequência da recombinação dos selecionados (RAMALHO et al., 2012, capítulo 10).

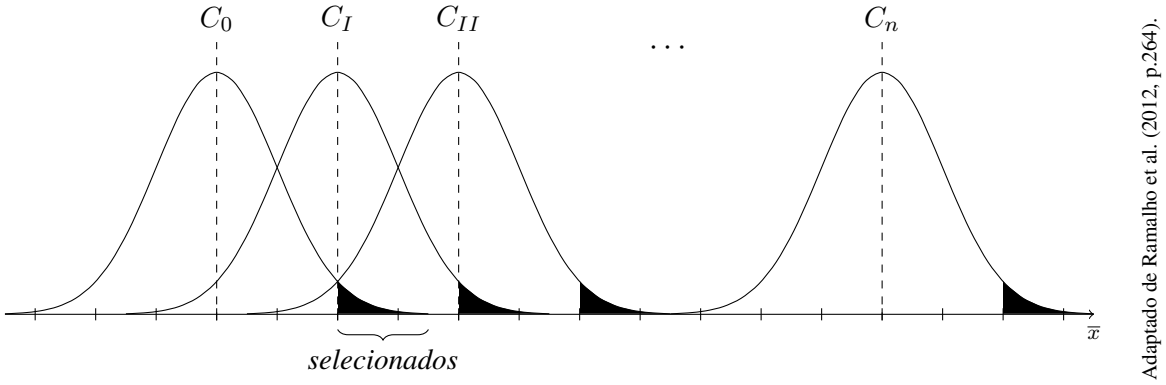

Figura 1 – Representação do progresso genético das médias de um caráter hipotético em diferentes “eras” de melhoramento sob seleção recorrente

Uma forma de se avaliar a eficiência de um programa de melhoramento é por meio do número de cultivares recomendadas e da respectiva área ocupada com tais cultivares. Entretanto, para Abreu et al. (1994) esse procedimento é falho, pois não afere somente a eficiência do processo de melhoramento, mas também da difusão de cultivares, o que, via de regra, independe da atividade do melhorista. Portanto, estimativas formais do progresso genético obtido são úteis em mensurar os ganhos, minimizar as perdas e maximizar a eficiência do processo. Tais estimativas auxiliam no processo de decisão e na avaliação das técnicas empregadas ou a serem implementadas (SOARES; RAMALHO; SOUSA, 1994)

Vencovsky et al. (1986) foram os precursores no desenvolvimento de metodologias com o intuito de estimar o progresso genético do melhoramento por meio indireto, ou seja, utilizando dados retrospectivos dos ensaios de avaliação de cultivares ao longo dos anos. O fundamento do método consiste em assumir um modelo linear, em que $\overline{Y}_i$ é a média geral de um ensaio, para um dado local i.

Desse modo, tem-se na forma escalar $\overline{Y}_i = \mu + a_i + \overline{g}_i + \overline{\varepsilon}_i$, com μ sendo a constante inerente a todas as observações, representando um ponto de referência (média geral); a_i é o efeito de ambiente, comum a todos os tratamentos sob avaliação; $\overline{g}_i$ é o potencial produtivo médio de todos os tratamentos presentes no experimento, e $\overline{\varepsilon}_i$ é o termo devido ao erro de medida, que inclui o efeito ambiental contido na média além da média das interações com os tratamentos e o ambiente i. Da mesma forma, os autores definem o mesmo modelo para um ano subsequente i' ($\overline{Y}_{i'} = \mu + a_{i'} + \overline{g}_{i'} + \overline{\varepsilon}_{i'}$).

Para a estimativa do progresso genético, o interesse é estimar a diferença $\overline{Y}_{i'} - \overline{Y}_i$ denominada diferença total, observada em um dado ano i', relativo ao anterior i (VENCOVSKY et al., 1986). Os autores comentam, entretanto, que este contraste confunde diferenças relativas a genótipos e a ambientes. Desse modo, ao se isolar os tratamentos que são comuns no biênio, indicado pelo índice c, calcula-se as médias destes como:

$$\begin{aligned}
\overline{Y}_{i_c} &= \mu + a_i + \overline{g}_{ii'_c} + \overline{\varepsilon}_{i_c} \\
\overline{Y}_{i'_c} &= \mu + a_{i'} + \overline{g}_{ii'_c} + \overline{\varepsilon}_{i'_c} \,.
\end{aligned}$$

A diferença genética (Δ_G) é definida como: $\Delta_{G_{i'i}} = (\overline{Y}_{i'} - \overline{Y}_i) - (\overline{Y}_{i'_c} - \overline{Y}_{i_c})$. Assumindo aleatórios os efeitos ambientais e os devidos às interações entre tratamentos e anos e os efeitos genotípicos fixos, a esperança matemática desta diferença resulta em

$$E[\Delta_{G_{i'i}}] = (g_{i'} - g_i)\,.$$

Portanto, $\Delta_{G_{i'i}}$, pode ser tomado como um estimador das alterações na média, ocorrida entre os anos i e i', que se atribui às alterações genotípicas ou incremento no potencial produtivo médio dos tratamentos que foram substituídos (VENCOVSKY et al., 1986).

Para cada biênio é possível obter tais estimativas ($\Delta_{G_{i'i}}, \Delta_{G_{i''i'}}, \ldots, \Delta_{G_{n,n-1}}$). O somatório de tais quantidades é tomado como o progresso genético acumulado; o acréscimo médio por ano é a razão entre o progresso acumulado e o número de anos de avaliação e o progresso relativo pode ser estimado como a razão do progresso estimado pela média do ciclo inicial (VENCOVSKY et al., 1986).

Todo o desenvolvimento proposto por Vencovsky et al. considera o fato que entre os anos subsequentes, alguns tratamentos são mantidos e outros substituídos por novos tratamentos. Esses "novos" tratamentos, por hipótese, devem ser iguais ou superiores aos substituídos da avaliação. Outro ponto necessário à acurácia da estimativa é a manutenção de um número apreciável de tratamentos no biênio.

Pelo método de Vencovsky et al., as diferenças genéticas são estimadas de forma independente entre os biênios. Todavia, Rodrigues (1990) mostra que existe uma variância entre os potenciais produtivos dos tratamentos nos biênios. Essa variância é função do quadrado médio do resíduo σ_ε^2, do número de tratamentos comuns $C_{ii'}$ e do número de tratamentos em cada ano do biênio, respectivamente T_i e $T_{i'}$. Tal variância é definida como:

$$V(\overline{g}_{ii'}) = \sigma_\varepsilon^2 \left(\frac{2}{C_{ii'}} - \frac{1}{T_i} - \frac{1}{T_{i'}} \right) .$$

E, por sua vez, a covariância entre os efeitos genéticos entre três anos consecutivos de ensaios é

$$Cov(\overline{g}_{ii'}, \overline{g}_{i'i''}) = -\sigma_\varepsilon^2 \left(\frac{C_{ii'i''}}{C_{ii'}C_{i'i''}} - \frac{1}{T_i} \right) ,$$

em que $C_{ii'i''}$ representa o número de cultivares comuns nos três anos consecutivos (RODRIGUES, 1990).

A partir dessas variâncias e covariâncias pode-se constituir uma matriz $\boldsymbol{V}$, com as variâncias e covariâncias entre os efeitos g's. Desse modo, por meio de um modelo linear generalizado é possível estimar os efeitos g's levando-se em consideração a correlação dos potenciais produtivos ao longo dos anos de análise e com isso incrementar a acurácia da estimativa do progesso genético, pois o resultado desse ajuste está sujeito ao teste de F-Snedecor entre a variância da média dos g's e a variância dos desvios entre tais valores (RODRIGUES, 1990).

Fernandes e Frazon (1997) aperfeiçoaram o método de Vencovsky et al. considerando que a estimativa do progresso genético precisa levar em conta os efeitos ambientais, enquanto que no original as melhorias do ambiente são consideradas como parte do ganho total.

Estendendo o modelo pioneiro, Fernandes e Frazon (1997) definem: $\overline{Y}_i = \mu + a_i + \overline{g}_i + (ag)_i + \overline{\varepsilon}_i$, em que $(ag)_i$ é a interação genótipos por anos e os demais termos já foram definidos. Para o ano subsequente (i') o modelo é análogo ($\overline{Y}_i = \mu + a_{i'} + \overline{g}_{i'} + (ag)_{i'} + \overline{\varepsilon}_{i'}$). Consideram ainda, os termos: $\overline{Y}_{ii'_n}$ como a média dos tratamentos presentes apenas no ano i e, por sua vez, $\overline{Y}_{i'i_m}$, referindo-se apenas à média dos tratamentos presentes no ano i'. O contraste entre ambos $\overline{Y}_{i'i_m} - \overline{Y}_{ii'_n} = (a_{i'} - a_i) + (g_{i'} - g_i) + (\varepsilon_{i'} - \varepsilon_i)$ é estabelecido como outra forma de estimar a diferença total, como uma alternativa ao contraste estabelecido por Vencovsky et al..

Portanto, a diferença entre ambos os métodos é a natureza do contraste definido, ao invés de utilizar o contraste original, o modelo de Fernandes e Frazon (1997) utiliza apenas as diferenças entre os tratamentos que não são comuns de um ano para outro, ou seja, o contraste $\overline{Y}_{i'i_m} - \overline{Y}_{ii'_n}$. A justificativa é que assim, a estimativa do progresso incide apenas entre os tratamentos substituídos (FERNANDES; FRAZON, 1997).

Desse modo, espera-se a obtenção de estimativas das diferenças totais de maior magnitude do que pelo método original; consequentemente a estimativa do progresso genético, pelo método modificado de Fernandes e Frazon é também maior. Os autores comentam que a

estimativa de Vencovsky et al. subestima o progresso genético, pois por esse método, o progresso estimado está "diluído", sendo portanto menor, já que entre os tratamentos comuns não há progresso genético.

Quanto maior o número de tratamento comuns em um biênio, menores as estimativas do progresso genético. O aperfeiçoamento proposto está mais em concordância com o conceito clássico do progresso genético, já que os tratamentos comuns não são incluídos na estimativa do progresso, refletindo portanto o esforço do melhorista em substituir os tratamentos por outros com maior potencial produtivo (FERNANDES; FRAZON, 1997).

Outro método empregado para a estimativa do progresso genético de modo indireto, foi proposto inicialmente por Breseghello (1995). O método consiste em agrupar os tratamentos genéticos avaliados no conjunto completo das avaliações do período por ano de lançamento e a partir disso ajustar o seguinte modelo estatístico:

$$Y_{ijkl} = \mu + a_i + t(a)_{ij} + g_k + l(g)_{kl} + \varepsilon_{ijkl}\,, \tag{1}$$

em que: Y_{ijkl} é uma observação do experimento, expressa na média do tratamento l no experimento j dentro do ano i; μ é a constante inerente a todas observações; a_i é o efeito de ano; $t(a)_{ij}$ é o efeito de experimento dentro de ano; g_k é o efeito de grupos de tratamentos, ou seja, o efeito genético "cronológico"; $l(g)_{kl}$ é o efeito de tratamentos dentro de grupos; e ε_{ijkl} é o termo do erro de medida, efeito residual inerente a todas observações, por pressuposição, assumido $\varepsilon_{ijkl} \cap N(0, \sigma^2_\varepsilon)$.

O ajuste desse modelo resulta nas estimativas dos efeitos de grupos de experimentos (vetor $\boldsymbol{G}$) e a matriz denominada $\boldsymbol{V}$ das variâncias e covariâncias desses respectivos efeitos. Assim, o progresso genético é estimado como a solução $\hat{\boldsymbol{\beta}}$ computado por meio de um modelo linear generalizado: $\hat{\boldsymbol{\beta}} = (\boldsymbol{X'V}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{X'V}^{-1}\boldsymbol{G})$. Sendo $\boldsymbol{X}$ uma matriz com uma coluna de 1's e outra coluna com os respectivos anos de lançamento dos tratamentos (BRESEGHELLO, 1995). O método permite que o coeficiente angular dessa regressão (estimativa do progresso genético), seja testada pelo teste t. E da mesma forma que no método pioneiro de Vencovsky et al. (1986) possibilita expressar a estimativa do progresso em termos relativos à média do início do processo de melhoramento, ou o coeficiente linear do ajuste (BRESEGHELLO, 1995).

O advento dos modelos mistos, descortinou a possibilidade de obter estimativas por meio da teoria de máxima verossimilhança restrita, que possui propriedades que resultam em estimadores mais fidedignos e comportam dados desbalanceados, sem contudo, prejudicar a acurácia (BERNARDO, 2010, capítulo 11). Nesse contexto, o método proposto por Breseghello (1995) foi também utilizado por meio de modelos mistos (CHIORATO et al., 2010; BRESEGHELLO et al., 2011).

Assim, a partir do modelo (eq. 1), assumindo fixos os efeitos de grupos de tratamentos e aleatórios os efeitos de tratamento dentro de grupos, ano, e experimento dentro de ano, constitui-se matricialmente o seguinte modelo misto: $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\lambda} + \boldsymbol{\varepsilon}$. Em que $\boldsymbol{y}$ é o vetor das observações; $\boldsymbol{X}$ e $\boldsymbol{Z}$ são as matrizes de incidência dos efeitos fixos e aleatórios,

respectivamente; $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$, são os vetores de solução para os respectivos efeitos; e $\boldsymbol{\varepsilon}$ é o vetor de resíduos com a mesma pressuposição do modelo original. Seguindo a teoria dos modelos mistos, a variância dos dados é tida como $V(\boldsymbol{y}) = \boldsymbol{V} \sim \boldsymbol{ZGZ}' + \boldsymbol{R}$, em que $\boldsymbol{G}$ é a matriz de variâncias e covariâncias dos efeitos aleatórios do modelo e $\boldsymbol{R}$, a matriz de variâncias e covariâncias dos resíduos, como já dito, assumida como sendo $\boldsymbol{R} = \boldsymbol{I}\sigma_{\varepsilon}^2$ (MRODE; THOMPSON, 2005, capítulo 3).

No caso específico em que um tratamento, ou grupo de tratamentos estiverem nos experimentos durante todo o período, Abreu et al. (1994) utilizaram a regressão linear, entre anos como variável independente e a produtividade média como dependente restrita apenas aos tratamentos comuns como uma estimativa da alteração ambiental, enquanto que o coeficiente dessa regressão utilizando os demais tratamentos mede a alteração fenotípica (genética mais ambiental). Desse modo, a diferença entre ambas as estimativas, pode ser tomada como uma medida do progresso genético. Os resultados obtidos que comparam esse método ao original de Vencovsky et al. mostram concordância entre as estimativas (ABREU et al., 1994).

Soares, Ramalho e Sousa (1994) compararam os métodos de Vencovsky et al. e Rodrigues e observaram que existe uma diferença significativa entre as estimativas dos progressos para os anos extremos. Sendo o método de Vencovsky et al. mais eficaz em detectar tais alterações, enquanto que o de Rodrigues tende a atenuar os sobressaltos nos anos intermediários pelos anos anteriores (SOARES; RAMALHO; SOUSA, 1994). A sugestão para tais situações seria o emprego do método da regressão de erros absolutos. Esse procedimento é uma eficaz alternativa ao método dos quadrados mínimos em se tratando de observações discrepantes e/ou com valores extremos (SOARES; RAMALHO; SOUSA, 1994).

Nos diversos estudos relativos à estimativa do progresso genético das culturas agrícolas, dois modos distintos têm sido usados. Segundo Vencovsky et al. (1986), os métodos indiretos para a estimativa do progresso genético são indubtavelmente passíveis de críticas. A opção mais indicada seria obter a estimativa de modo direto, ou seja, por meio de ensaios específicos, em diferentes locais e anos, e quando oportuno, incluindo outras variáveis de interesse, como práticas de manejo, fazendo uso de sementes renovadas dos principais cultivares lançadas no período histórico.

Em síntese, à parte as especificidades de cada metodologia, a estimativa do progresso genético é obtida como o coeficiente angular da regressão das médias dos genótipos de cada período (FALCONER; MACKAY, 1996, capitulo 12).

2.1.2 Estimativas do Progresso Genético

Pode-se dizer que o melhoramento genético de plantas começou com a domesticação das espécies e consequentemente com a origem da agricultura, um marco das civilizações (RAMALHO et al., 2012, capítulo 1). Diversos trabalhos foram realizados, sobretudo após a segunda guerra mundial, com o intuito de quantificar a efetividade do melhoramento genético por meio do progresso genético das principais culturas agrícolas, o que evidenciou expressivos

sucessos nesse intento (MILES; PANDEY, 2004; TRACY; TIEFENTHALER; SCHABER, 2004). Vencovsky e Ramalho (2006) apresentam estimativas do progresso genético em importantes culturas agrícolas no Brasil. Esses incrementos de produtividade, são, as vezes, até superiores aos relatados ao redor do globo. A seguir são comentados alguns estudos realizados, no Brasil e no mundo, que quantificam o progresso genético da produtividade de grãos de importantes culturas agrícolas e o relacionamento com outros caracteres importantes na visão dos melhoristas.

Em estudo de duplo propósito, Soares, Ramalho e Sousa (1994) estimaram o progresso genético em 10 anos de melhoramento do arroz comparando diferentes metodologias. O incremento médio anual da produtividade de grãos pelo método de Vencovsky et al. (1986) foi de 82,50$kg\,ha^{-1}$, enquanto que ao empregar a regressão de erros absolutos a estimativa do progresso foi de 62,9$kg\,ha^{-1}$, tais estimativas, respectivamente, representam progressos genéticos de 2% e 1% ao ano respectivamente. Apesar de diferentes, em termos absolutos, as estimativas foram próximas. A regressão de erros absolutos torna-se recomendável com o objetivo de atenuar os valores aberrantes, todavia, o método de Vencovsky et al. possui aplicação mais geral e flexível (SOARES; RAMALHO; SOUSA, 1994).

Abreu et al. (1994), em trabalho semelhante, estimaram o progresso genético do feijoeiro durante 20 anos de melhoramento no estado de Minas Gerais e compara diferentes métodos de estimação. De forma geral, os diferentes métodos, assim como em Soares, Ramalho e Sousa (1994), resultaram em estimativas equivalentes. Ao propôr um novo método de estimação, os autores reportam um incremento médio anual da produtividade de 42,62$kg\,ha^{-1}$. Essa quantidade, tem significado fenotípico, de modo que, dada uma melhoria do ambiente de 28,10$kg\,ha^{-1}$, o progresso foi estimado por diferença, como 14,52$kg\,ha^{-1}$, que representa um progresso da produtividade de 2% ao ano (ABREU et al., 1994).

Mais recentemente, em um estudo com o feijoeiro, Chiorato et al. (2010) avaliaram cerca de 20 anos de melhoramento, destacando duas distintas fases. Em um primeiro momento constatou-se um progresso genético de 1% ao ano, enquanto que na fase posterior, tal estimativa não foi significativamente diferente de zero. O observado no segundo período segue o contexto da seleção que fora praticada, priorizando cultivares com elevada qualidade tecnológica de grãos em detrimento à produtividade. Tal estagnação dos incrementos genéticos, todavia, representa um alerta em relação às estratégias empregadas apesar de o rendimento médio de grãos dos cultivares da segunda fase ter sido, em média, uma tonelada maior que os do período prévio (CHIORATO et al., 2010).

Ao distinguir entre grupos de maturação de 20 anos de melhoramento de arroz Soares et al. (1999) relacionam a estimativa dos progressos aos diferentes ambientes de produção e grupos de maturação. Os progressos genéticos dos grupos precoce e tardio, foram respectivamente de 1% e 3%, que em números absolutos representam, na mesma ordem 23,82$kg\,ha^{-1}$ e 46,43$kg\,ha^{-1}$ ao ano. A estimativa do progresso do grupo tardio foi mais que o dobro da estimativa do grupo precoce. A razão para tal discrepância é o fato de que, efetivamente, os cultivares desse

grupo possuirem elevada resistência à ocorrência de doenças. Assim, o progresso genético em produtividade dos cultivares do grupo precoce foi atenuado devido ao escape da resistência nas linhagens com essa maturação (SOARES et al., 1999).

Também com a cultura do arroz, em um período de 50 anos, Souza et al. (2007) estimaram, para os cultivares de maturação tardia, um progresso na produtividade de grãos de 27,6$kg\,ha^{-1}$, que equivale a 2% ao ano. O grupo de maturação precoce apresentou um progresso genético da produtividade de grãos de 14,6$kg\,ha^{-1}$ ($^1/_2$% ao ano), um acréscimo médio da duração do ciclo vegetativo de dez dias e uma redução da altura média das plantas de 21cm. Além do incremento em produtividade de grãos, os cultivares apresentaram progressos tanto na duração do ciclo, como da altura da planta em resposta à seleção praticada, que almejou cultivares com ciclo em torno de 85 a 90 dias e baixo índice de acamamento (SOUZA et al., 2007).

Os resultados de Breseghello et al. (2011) coincidem com o observado por Souza et al. (2007). Ao longo de 25 anos, estes últimos autores observaram progressos para a produtividade de grãos de arroz, altura da planta e dias para o florescimento. No período, para a produtividade, fases distintas foram observadas: em um primeiro momento, foi obtido pouco progresso em produtividade, avaliando-se muitos genótipos que variavam amplamente para os outros caracteres. Assim, nessa primeira fase houve "indiretamente" a definição de um ideótipo, que propiciou os ganhos posteriores observados.

A cultura do arroz foi também intensamente pesquisada no exterior, Peng et al. (2000) avaliando o progresso genético nas Filipinas desde 1966 estimaram um acréscimo no rendimento médio de grãos de 1% ao ano, que pode não refletir um ganho "real" do potencial produtivo dada a coexistência de fatores bióticos/abióticos aos quais os cultivares recentes devem ser melhor adaptados. Ao relacionar o progresso aos componentes da produção, observaram que em uma primeira fase, o progresso na produtividade esteve atrelado à melhoria no índice de colheita, enquanto que após isso os acréscimos estiveram relacionados à produção de biomassa.

Ao associar o estudo do progresso genético com as melhorias do ambiente Tabien, Samonte e McClung (2008) avaliaram cultivares "elite" de arroz, lançados desde 1944 frente a diferentes doses de Nitrogênio. Marginalmente à aplicação do fertilizante, o progresso genético foi de 23,98$kg\,ha^{-1}$, ou $^1/_2$% da média das cultivares ancestrais. Sob todas as doses, o coeficiente angular positivo demonstrou que o melhoramento resultou em incremento, independente do uso de nitrogênio. Destaca-se que o efeito da dose de N foi significativo e que dobrando-se a dose, praticamente dobrou-se o progresso. Além de mais produtivos, os cultivares atuais são também mais precoces, com plantas menores, têm mais rendimento em arroz beneficiado e são mais responsivas à aplicação do fertilizante (TABIEN; SAMONTE; MCCLUNG, 2008).

Lange e Federizzi (2009) relacionaram a estimativa do progresso genético da soja em quatro diferentes regiões do Rio Grande do Sul. A análise de correlação das médias das cultivares entre regiões permitiu inferir acerca do resultado da seleção praticada em uma região alvo e seu reflexo nas demais regiões. O estudo restringiu-se ao progresso na produtividade de

grãos que, para o grupo de maturação médio variou nas regiões entre 18,5$kg\,ha^{-1}$ e 71,0$kg\,ha^{-1}$ e para o grupo tardio entre 18,1$kg\,ha^{-1}$ e 26,8$kg\,ha^{-1}$. Os autores argumentam que, de sobremaneira, esforços são gastos na seleção de outros caracteres de interesse tais como resistência/tolerância a fatores bióticos/abióticos e características de qualidade. Portanto, recomendam que, para uma avaliação mais ampla do sucesso de um programa, dever-se-iam incluir tais atributos (LANGE; FEDERIZZI, 2009).

Além do incremento médio em produtividade ao longo do melhoramento, Ustun, Allen e English (2001) avaliaram a estabilidade de cultivares de soja ao longo de 15 anos, constatando que os ancestrais, além de menos produtivos, eram menos responsivos à melhoria do ambiente, ou seja, os cultivares modernos produziam mais em todas as condições de cultivo, tanto em ambientes de alto como de baixo potencial produtivo. Progressos simultâneos para o teor de óleo e proteína foram dificultados pela correlação negativa entre ambos os caracteres. Segundo os autores, havendo propósitos específicos, os teores de um caráter podem ser melhorados em detrimento do outro (USTUN; ALLEN; ENGLISH, 2001).

Compilando os resultados de 16 anos de melhoramento da soja na África, Tefera et al. (2010) estimaram um incremento na produtividade de grãos no grupo de maturação médio, aumentando de 1,18$Mg\,ha^{-1}$ para 1,87$Mg\,ha^{-1}$, enquanto que no grupo tardio foi de 1,57$Mg\,ha^{-1}$ para 2,30$Mg\,ha^{-1}$, o que equivale a incrementos anuais da ordem de 2% e $1^1/_2$%, respectivamente. Além da melhoria da produtividade de grãos, houve, para ambos os grupos de maturação, aumento na nodulação natural das plantas, aumento da biomassa e do índice de colheita, com mais vagens por planta e redução na deiscência natural (TEFERA et al., 2010).

Em 40 anos de melhoramento de trigo, Underdahl et al. (2008) estimaram o progresso genético em diferentes locais e anos que variou entre 15,3$kg\,ha^{-1}$ e 60,0$kg\,ha^{-1}$, e na média foi de 30,4$kg\,ha^{-1}$, ou de 1% ao ano. Apesar do resultado obtido para a produtividade de grãos, a altura média de plantas e o número de dias para o florescimento não mostraram alterações significativas na sua expressão, entretanto as alterações na expressão da produtividade foram influenciadas devido ao aperfeiçoamento da tecnologia de produção e melhorias na resistência múltipla dos genótipos aos fatores de estresse (UNDERDAHL et al., 2008).

No Brasil, o acréscimo do potencial produtivo do trigo selecionado para o sistema irrigado foi estimado por Cargnin, Souza e Fronza (2008), como 2,14$Mg\,ha^{-1}$, dos quais, 1,44$Mg\,ha^{-1}$ (67%) são atribuídos ao progresso genético. Essa quantidade representa um incremento anual em produtividade de 48,03$kg\,ha^{-1}$, ou seja, de 2% ao ano. Ao interpretar a estimativa do progresso genético foi possível observar três fases distintas: em um primeiro momento, o progresso foi de baixa magnitude, quando a estratégia de melhoramento consistia da introdução de linhagens de outros programas; a segunda fase caracterizou-se por ganhos mais expressivos, quando, majoritariamente as linhagens foram obtidas por seleção em populações segregantes; e na fase posterior o progresso genético observado tendeu a um declínio à semelhança do ocorrido em Chiorato et al. (2010), refletindo as mudanças nos objetivos do programa com maior atenção a atributos relacionados à qualidade de panificação

(CARGNIN; SOUZA; FRONZA, 2008). Para o cultivo de sequeiro, a estimativa para a produtividade foi realizada por Cargnin, Souza e Fronza (2009), que obtiveram incremento de $37kg\,ha^{-1}$, ou 47% representando o progresso genético do potencial produtivo dos cultivares.

Para a cultura da aveia no Brasil, Barbosa Neto et al. (2000) estimaram o progresso genético da produtividade de grãos, obtendo $44kg\,ha^{-1}$ ($^1/_2$% da média do ciclo inicial). O período vegetativo teve um incremento de 1,5 dias ao ano ($^1/_2$%) enquanto que o peso do hectolitro teve um progresso equivalente a 0,3% ao ano. Pela análise de correlação os autores inferiram que nenhum caráter, em especial, influenciou o incremento do rendimento médio de grãos. Ao considerar diferentes condições de manejo, variando adubação nitrogenada e aplicação de defensivos, não foi possível discriminar o comportamento dos genótipos de diferentes fases, ou seja, os genótipos responderam de modo semelhante a mudanças do ambiente (BARBOSA NETO et al., 2000).

Na Itália, 40 anos de melhoramento de aveia proporcionaram progressos, ao ano, de $36kg\,ha^{-1}$, que representa cerca de 1% da média dos cultivares ancestrais (REDAELLI et al., 2008). Assim, como em Barbosa Neto et al. (2000), Redaelli et al. (2008) observaram que as cultivares mais recentes caracterizam-se pela maior produtividade, por serem mais precoces e possuírem menor estatura, mas fundamentalmente por apresentarem maior adaptação ao ambiente, uma característica que assegura boa produtividade sob variadas condições de cultivo.

Os resultados encontrados na literatura acerca de estudos do progresso genético das culturas agrícolas no Brasil e no mundo são resumidos na tabela 1. São apresentados: o país de origem do estudo; o tempo de duração do programa de melhoramento avaliado; o progresso genético anual da produtividade de grãos, em valores absolutos e em porcentagem; e quando apresentados, a estimativa do progresso de outro caráter avaliado em cada estudo.

Tabela 1 – Resumo de estimativas do progresso genético publicadas para algumas culturas agrícolas no Brasil e no Mundo

País	Cultura	Tempo	Progresso		Outro			Referência
		(anos)	$kg\,ha^{-1}$	(%)	Caráter	Progresso	(%)	
Brasil	Arroz	10	82,50	2	-	-	-	Soares, Ramalho e Sousa
		20	23,82	1	-	-	-	Soares et al.
		50	27,60	2	altura	-21cm^*	-	Souza et al.
		25	44,99	1	altura	-0.52cm	-0,49	Breseghello et al.
	Trigo	29	48,03	2	-	-	-	Cargnin, Souza e Fronza
		29	37,00	7	-	-	-	Cargnin, Souza e Fronza
	Feijão	28	14,52	2	-	-	-	Abreu et al.
		18	21,04	1	-	-	-	Chiorato et al.
	Soja	20	71,50	4	-	-	-	Lange e Federizzi
	Aveia	40	44,00	$^1/_2$	dias p/ floresc.	1,50 dias	0,52	Barbosa Neto et al.
EUA	Arroz	48	23,98	$^1/_2$	dias p/ floresc.	-0,21 dias	1,24	Tabien, Samonte e McClung
	Trigo	40	30,40	1	altura	-0,14cm	0,13	Underdahl et al.
	Soja	15	14,40	1	óleo / proteína**	-	-	Ustun, Allen e English
Filipinas	Arroz	30	75,00	1	-	-	-	Peng et al.
África	Soja	16	23,61	2	altura	-0,76	1,10	Tefera et al.
Itália	Aveia	40	0,36	1	-	-	-	Redaelli et al.

* Relativo ao progresso total obtido no período.

** Ajustou-se o modelo quadrático de progresso relativo a ambos caracteres.

Alguns estudos, dos aqui apresentados, fazem interpretações da associação dos progressos envolvendo vários caracteres (SOUZA et al., 2007; UNDERDAHL et al., 2008; TEFERA et al., 2010; BRESEGHELLO et al., 2011). Não foram encontradas na literatura formas canônicas de computar o progresso genético simultâneo dos vários caracteres de interesse. Essa estimativa vai de encontro às expectativas dos melhoristas, trazendo um aporte interpretativo interessante, de acordo com a sugestão de Lange e Federizzi (2009). Decorre diretamente desse fato a possibilidade de se utilizar meios para reunir a informação de mais de um caráter tal qual um índice de seleção. Todavia, um índice que reúna tais propriedades não seria útil apenas e tão somente ao propósito do estudo do progresso genético simultâneo.

2.1.3 Metodologias de Índices de Seleção

A seleção que se baseia em um caráter apenas pode conduzir a um produto final superior em relação a esse caráter, mas que todavia, pode não apresentar desempenho tão favorável no tocante aos demais caracteres que foram desconsideradas na seleção. Os desenvolvimentos teóricos acerca de índices de seleção mostram-se como uma eficiente alternativa nesse mister. Fundamentalmente, os índices se propõem a reunir e combinar, múltiplas informações, tal que a seleção baseie-se em um complexo de variáveis, que agrupe vários atributos de importância para o melhoramento (CRUZ; REGAZZI; CARNEIRO, 2001, capítulo 6).

Os índices de seleção são uma alternativa à seleção em *tandem*, ou seja, a seleção de um caráter por vez. Esse processo entretanto, pressupõe que as médias dos caracteres que não estejam sob seleção permaneçam inalteradas ou que sejam melhoradas indiretamente. Outra alternativa é o emprego de níveis independentes de eliminação, em que são estabelecidos limiares (de mínimos ou máximos) para cada caráter, restringindo a seleção apenas aos genótipos cujo desempenho atenda aos critérios preestabelecidos (BERNARDO, 2010, capítulo 13).

Diferentemente que realizar seleção para caracteres individuais, em *tandem* ou por níveis independentes de eliminação, um índice de seleção constitui-se como um caráter adicional e artificial, que estabelece uma combinação "ótima" dos vários caracteres e permite incrementar a eficiência da seleção simultânea para múltiplos caracteres (HAZEL; LUSH, 1942). Na figura 2, são considerandos dois caracteres hipotéticos x_1 e x_2 e exemplificados os três métodos de seleção, ou seja: *tandem*, níveis independentes de eliminação e emprego de um índice como função linear dos caracteres x_1 e x_2 de modo que $I = \beta_1 x_1 + \beta_2 x_2$. Da representação gráfica, Wricke e Weber (1986, capítulo 12) mostram que o limiar de seleção dado pela seleção com o índice (linha pontilhada em C, figura 2) é uma reta com coeficiente angular dado por $-\frac{\beta_1}{\beta_2}$.

Elston (1963) desenvolveu um modo de ordenar indivíduos com relação a observação de diferentes caracteres sem fazer qualquer pressuposição, o que, em certa medida, caracteriza o método como livre de pesos e parâmetros. A função que Elston (1963) apresenta para ordenar

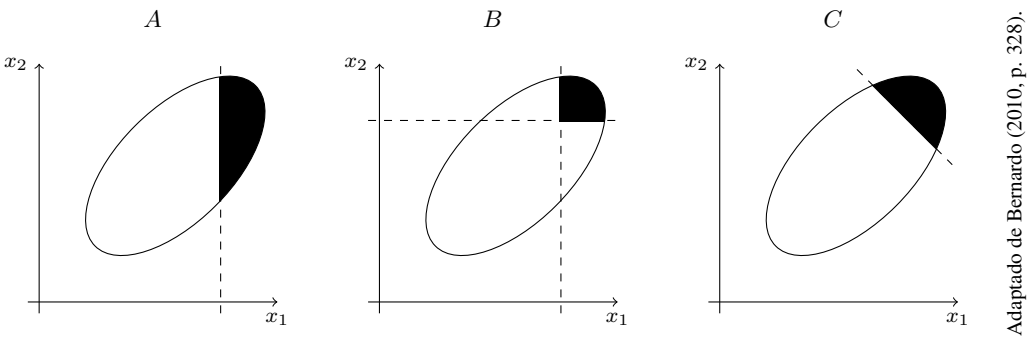

Figura 2 – Representação de dois caracteres hipotéticos x_1 e x_2 em três diferentes métodos de seleção: (A) seleção em *tandem*; (B) níveis independentes de eliminação; e (C) índice de seleção ($I = \beta_1 x_1 + \beta_2 x_2$)

os indivíduos é uma maneira *naïf*[1] de se derivar um índice de seleção e pode ser representado por:

$$I = \sum_{i=1}^{n} \omega_i = \omega_1 + \omega_2 + \omega_3 + \ldots + \omega_n \,,$$

em que $\omega_i = x_i - l_i$; x_i sendo o valor observado do $i-$ésimo caráter no indivíduo p e l_i é um limiar, pré-determinado de mínimo ou máximo para o referido caráter. Segundo Lin (1978) esse processo é uma sistematização do método de níveis independentes de eliminação, os resultados obtidos por ambos meios de seleção são comparáveis especialmente ao assumir $\omega_i = 0$, quando x_i estiver aquém (ou além) do limiar de mínimo (ou máximo) estabelecido.

Uma outra opção para a obtenção de um ordenamento dos indivíduos levando em consideração mais de um caráter foi proposta por Mulamba e Mock (1948), onde os p genótipos classificam-se em relação a cada um dos n caracteres sob avaliação. O ordenamento é o somatório dos postos dos referidos genótipos para todos caracteres n, ou seja:

$$I = \sum_{i=1}^{n} \phi_i = \phi_1 + \phi_2 + \ldots + \phi_n \,,$$

em que ϕ é o posto ou *ranking* do invidíduo para o caráter i. Essa proposta resulta em um novo "caráter" e pode ser tomada como um índice de seleção, apesar de que, da mesma forma que o método de Elston não há qualquer pressuposição, atribuição de pesos econômicos e/ou estimação de parâmetros (CRUZ; REGAZZI; CARNEIRO, 2001, capítulo 6).

De posse das observações das n variáveis avaliadas, Mendes, Ramalho e Abreu (2009) empregaram o somatório de tais variáveis padronizadas (estandartizada) tal qual um índice. Tais valores são obtidos por parcela, o que permite submete-los à análise estatística e também verificar qual a contribuição de cada variável padronizada no valor do somatório (índice) por meio de um gráfico polar (MENDES; RAMALHO; ABREU, 2009). Este procedimento foi sugerido como uma alternativa ao emprego de índices de seleção.

[1] O termo aparece no vocabulário, como sinônimo de ingênuidade e/ou originalidade.

Os métodos de Elston (1963), Mulamba e Mock (1948) e Mendes, Ramalho e Abreu (2009) relacionam-se apenas com as observações fenotípicas e desconsideram o relacionamento entre os caracteres. A crítica ao emprego de índices com essas características é que a seleção pode se tornar ineficiente (WRICKE; WEBER, 1986, capítulo 12). Quando se almeja realizar seleção sobre alguma medida não observável (genótipo) deve-se levar em conta a influência que o erro experimental exerce sobre as estimativas (variação não herdável), e foi isso, essencialmente, que Smith (1936) resolveu (ELSTON, 1963).

Smith (1936) abordou os casos em que se almeja selecionar genótipos por meio de vários fenótipos correlacionados sendo, todavia, desconhecido o verdadeiro valor objeto da seleção. Fisher (1936) demonstra como determinar uma funcão de diferentes variáveis, que melhor discrimina fatores com relação a uma variável subjacente. Tal procedimento maximiza a variância entre fatores e minimiza a variância dentro de cada fator. A partir do modelo de Fisher, Smith determina qual função discriminante dos n fenótipos que melhor indica o verdadeiro valor genotípico dos indivíduos sob avaliação.

Smith (1936) partiu do seguinte modelo na forma escalar $x_{ij} = g_{ij} + \varepsilon_{ij}$, em que x_{ij} é uma observação fenotípica de um caráter i no indivíduo j; g_{ij} é o efeito genotípico do caráter i que incide sobre o indivíduo j e ε_{ij} é termo do erro de medida, ou seja, inclui o efeito ambiental e a interação do efeito genotípico i com o ambiente, que se assume independente dos efeitos genéticos. Se g_{ij} é o valor genotípico de um caráter, o agregado genotípico do indivíduo j é mensurado como $H_j = \sum_{i=1}^{n} a_i g_{ij}$, em que a_i é uma constante conhecida que pondera a importância do caráter i no agregado genotípico, ou seja, o peso econômico do caráter i (SMITH, 1936).

Entretanto, de acordo com Smith (1936), tais quantidades H_j não podem ser diretamente avaliadas devido às variações não herdáveis, ou seja, pela incidência do termo ε, que por pressuposição é tomado como aleatório, independente e identicamente distribuído $\varepsilon_{ij} \cap N(0, \sigma^2_{\varepsilon_i})$. Seguindo o mesmo raciocínio, o "índice", ou valor fenotípico discriminante do índividuo j, é representado por $I_j = \sum_{i=1}^{n} \beta_i x_{ij}$, sendo β_i um coeficiente desconhecido. O problema consiste portanto, em estimar os valores de $\boldsymbol{\beta}$ nesta relação com $\boldsymbol{I}$ que melhor discriminem fenotipicamente os indivíduos com os melhores agregados genotípicos $\boldsymbol{H}$.

De acordo com Falconer e Mackay (1996, capítulo 11), o progresso genético com a seleção é expresso como $R = i\sigma_G\rho$, em que i é o diferencial de seleção estandartizado; σ_G é a raiz quadrada da variância genética e ρ a correlação entre os valores genéticos e fenotípicos, e que equivale à raiz quadrada da herdabilidade (h^2). Smith (1936) estendeu tal expressão do ganho, para a seleção simultânea de n caracteres. O progresso genético dos valores fenotípicos discriminantes (Δ_I), baseado na seleção pelo índice $\boldsymbol{I}$, dada uma intensidade de seleção de κ, é dada como:

$$\Delta_I = \kappa\sigma_H\rho_{IH} \,. \tag{2}$$

Os termos dessa expressão guardam uma relação com a expressão do ganho, a saber: κ é

o diferencial de seleção estandartizado, no conjunto dos n caracteres; σ_H é o desvio padrão do agregado genotípico; e ρ_{IH} é a correlação entre os valores fenotípicos discriminantes ($\boldsymbol{I}$) e os agregados genotípicos ($\boldsymbol{H}$) dos indivíduos sob avaliação. Lin (1978) argumenta que, apesar da correspondência, ρ_{IH} não se refere à raiz quadrada do coeficiente de determinação genotípico do índice, tal qual a herdabilidade.

Bulmer (1985, capítulo 11) teoriza acerca da pressuposição de que os valores fenotípicos e genéticos dos n caracteres, nos p indivíduos, possuem distribuição normal conjunta com uma função de densidade $f(x, g)$. Assim, para estimar a regressão dos valores do agregado genotípico nos valores fenotípicos estima-se a esperança de g dado x. Isto posto, o índice clássico de Smith constitui-se na regressão múltipla dos valores do agregado genotípico $\boldsymbol{H}$ sobre os valores fenotípicos $\boldsymbol{x}$ (BULMER, 1985, capítulo 11).

Hazel (1943) avaliou a expressão do ganho (eq. 2) e enfatiza que κ e σ_H são constantes na população e assim, para maximizar os ganhos, pode-se maximizar a correlação ρ_{IH}. Wricke e Weber (1986, capítulo 12) provam que o estimador obtido maximizando-se a correlação equivale ao obtido minimizando a variância da diferença entre os valores fenotípicos discriminantes e o agregado gentípico ($\sigma^2_{(\boldsymbol{I}-\boldsymbol{H})}$).

O relacionamento entre as quantidades fenotípicas e genéticas constituem o sistema:

$$\left\{\begin{array}{rcl}
\sigma^2_{P_1}\frac{\sigma_{P_1}}{\sigma_H}\rho_{P_1H} + \rho_{P_{12}}\frac{\sigma^2_{P_1}\sigma_{P_2}}{\sigma_H}\rho_{P_2H} + \cdots + \rho_{P_{1n}}\frac{\sigma^2_{P_1}\sigma_{P_n}}{\sigma_H}\rho_{P_nH} & = & \sigma^2_{G_1}a_1 + \rho_{G_{12}}\sigma_{G_1}\sigma_{G_2}a_2 + \cdots + \rho_{G_{1n}}\sigma_{G_1}\sigma_{G_n}a_n \\
\rho_{P_{12}}\frac{\sigma_{P_1}\sigma^2_{P_2}}{\sigma_H}\rho_{P_1H} + \sigma^2_{P_2}\frac{\sigma_{P_2}}{\sigma_H}\rho_{P_2H} + \cdots + \rho_{P_{2n}}\frac{\sigma^2_{P_2}\sigma_{P_n}}{\sigma_H}\rho_{P_nH} & = & \rho_{G_{12}}\sigma_{G_1}\sigma^2_{G_2}a_1 + \sigma^2_{G_2}a_2 + \cdots + \rho_{G_{2n}}\sigma_{G_2}\sigma_{G_n}a_n \\
 & \vdots & \\
\rho_{P_{1n}}\frac{\sigma_{P_1}\sigma^2_{P_n}}{\sigma_H}\rho_{P_nH} + \rho_{P_{2n}}\frac{\sigma_{P_2}\sigma^2_{P_n}}{\sigma_H}\rho_{P_nH} + \cdots + \sigma^2_{P_n}\frac{\sigma_{P_n}}{\sigma_H}\rho_{P_nH} & = & \rho_{G_{1n}}\sigma_{G_1}\sigma^2_{G_n}a_1 + \rho_{G_{2n}}\sigma_{G_2}\sigma^2_{G_n}a_2 + \cdots + \sigma^2_{G_n}a_n
\end{array}\right.$$

Se $\rho_{P_iH}\frac{\sigma_{P_i}}{\sigma_H}$ é o coeficiente angular da regressão do caráter i sobre o agregado genotípico $\boldsymbol{H}$, ou seja β_i e sendo $\rho_{*_{ij}}\sigma_{*_i}\sigma_{*_j}$ a covariância, genética ou fenotípica com índices G ou P, respectivamente, entre os caracteres i e j, substituindo os termos reescrevemos o sistema, de acordo com Falconer e Mackay (1996, capitulo 19), conforme segue:

$$\left\{\begin{array}{rcl}
\boldsymbol{P}_{11}\boldsymbol{\beta}_1 + \boldsymbol{P}_{12}\boldsymbol{\beta}_2 + \cdots + \boldsymbol{P}_{1n}\boldsymbol{\beta}_n & = & \boldsymbol{G}_{11}\boldsymbol{a}_1 + \boldsymbol{G}_{12}\boldsymbol{a}_2 + \cdots + \boldsymbol{G}_{1n}\boldsymbol{a}_n \\
\boldsymbol{P}_{21}\boldsymbol{\beta}_1 + \boldsymbol{P}_{22}\boldsymbol{\beta}_2 + \cdots + \boldsymbol{P}_{2n}\boldsymbol{\beta}_n & = & \boldsymbol{G}_{21}\boldsymbol{a}_1 + \boldsymbol{G}_{22}\boldsymbol{a}_2 + \cdots + \boldsymbol{G}_{2n}\boldsymbol{a}_n \\
 & \vdots & \\
\boldsymbol{P}_{n1}\boldsymbol{\beta}_1 + \boldsymbol{P}_{n2}\boldsymbol{\beta}_2 + \cdots + \boldsymbol{P}_{nn}\boldsymbol{\beta}_n & = & \boldsymbol{G}_{n1}\boldsymbol{a}_1 + \boldsymbol{G}_{n2}\boldsymbol{a}_2 + \cdots + \boldsymbol{G}_{nn}\boldsymbol{a}_n
\end{array}\right.$$

$\boldsymbol{P}$ e $\boldsymbol{G}$ são as matrizes ($n \times n$) de variâncias e covariâncias fenotípicas e genéticas respectivamente entre os n caracteres; $\boldsymbol{\beta}$ é o vetor de estimativas dos coeficientes do índice ($\hat{\boldsymbol{\beta}}' = \left[\hat{\beta}_1; \hat{\beta}_2; \ldots; \hat{\beta}_n\right]$); e $\boldsymbol{a}$ é o vetor de constantes, ou pesos econômicos ($\boldsymbol{a}' = [a_1; a_2; \ldots; a_n]$). Assim o sistema na forma matricial consiste em $\boldsymbol{P\beta} = \boldsymbol{Ga}$. Considerando $\boldsymbol{P}$ como uma matriz, simétrica positiva definida, não singular com inversa única, obtem-se a solução do índice de Smith e Hazel como:

$$\hat{\boldsymbol{\beta}} = \boldsymbol{P}^{-1}\boldsymbol{G}\boldsymbol{a}\,. \tag{3}$$

Hazel e Lush (1942) compararam a eficiência dos três métodos de seleção avaliando os respectivos progresso genéticos. A razão entre os progressos do índice e da seleção em *tandem* resulta na raiz quadrada do número de caracteres, ou seja, a eficiência do índice é $\sqrt{n}$ maior que o da seleção em *tandem*, mantidas as mesmas intensidades de seleção[2]. Níveis independetes de eliminação situam-se em eficiência, sempre entre os demais (HAZEL; LUSH, 1942).

Para Pesek e Baker (1969) muitos melhoristas ficam seguros em arbitrar pesos econômicos aos caracteres, não obstante, outros preferem especificar ganhos desejados para tais caracteres, já que estes ganhos são menos empíricos e intuitivos e de fácil determinação do que pesos econômicos. Diante disso, Pesek e Baker (1969) propuseram utilizar o vetor de ganhos desejados como pesos econômicos para os caracteres do índice. A parte as manipulações algébricas obtem-se $\hat{\boldsymbol{\beta}} = \boldsymbol{G}^{-1}\boldsymbol{\Delta}$. Portanto, é possível estimar coeficientes assinalando ganhos desejados para cada caráter ao invés de considerar pesos econômicos.

Mesmo não necessitando da determinação dos pesos econômicos *a priori*, a partir do índice já estabelecido, tais ponderadores (intuitivamente empregados) podem ser obtidos por intermédio da relação inversa $\hat{\boldsymbol{a}} = \boldsymbol{G}^{-1}\boldsymbol{P}\hat{\boldsymbol{\beta}}$ (CRUZ; REGAZZI; CARNEIRO, 2001, capítulo 6). A importância que o melhorista atribui ao caráter reflete-se por meio do ganho desejado (intensidade de seleção praticada). Deste modo, reduz-se a influência que os pesos econômicos exercem sobre o índice. Todavia as variâncias e covariâncias influenciarão este índice de forma semelhante ao que acontece com o índice *natural*.

Williams (1962) propôs o estabelecimento de um índice como somatório das combinações lineares dos valores médios fenotípicos ponderados pelos seus respectivos pesos econômicos, de forma que $\boldsymbol{x}$ e $\boldsymbol{a}$ são vetores $(n \times 1)$, cujos elementos (x_i e a_i) são, respectivamente, as médias e os pesos econômicos dos caracteres estudados. Na situação em que as variâncias e covariâncias fenotípicas forem de causas majoritariamente genéticas, o índice de Williams resume-se ao índice clássico, pois $\boldsymbol{P}^{-1}\boldsymbol{G} = \boldsymbol{I}^*$, ($\boldsymbol{I}^*$ é uma matriz identidade), assim $\boldsymbol{\beta} = \boldsymbol{P}^{-1}\boldsymbol{G}\boldsymbol{a} = \boldsymbol{a}$. A proposta de Williams (1962) reduz a influência que o erros na estimativa da matriz de variâncias e covariâncias genéticas exercem sobre o índice clássico (WILLIAMS, 1962).

Os pesos econômicos podem ser atribuídos de modo a expressar a preferência do melhorista, basear-se em demandas e retornos econômicos, fundar-se em retrospecto ou ainda assumir valores constantes arbitrários. O ideal é que os pesos exprimam o benefício marginal do melhoramento dos caracteres (MAGNUSSEN, 1990). Assim, Magnussen (1990) apresenta um algoritmo, próprio para o índice clássico, que estima pesos econômicos assegurando o máximo progresso simultâneo em todos caracteres. A solução envolve operações matriciais entre as formas quadráticas das matrizes de variâncias e covariâncias genéticas e fenotípicas.

[2]A justificativa dessa superioridade advém de uma analogia geométrica: em um quadrado de lado um, a distância de um vértice a outro pelos lados é a $\sqrt{2}$; em um cubo, três lados distam $\sqrt{3}$ vezes a diagonal oposta a eles.

Conhecida a teoria acerca do índice de selecão clássico, Tallis (1962) argumenta que pouca atenção foi despendida ao que ocorre na prática da seleção condicional. Kempthorne e Nordskog (1959) consideram o caso em que ao melhorista interessa maximizar o progresso do agregado genotípico, porém deseja que em r ($r < n$) caracteres não ocorram mudanças, ou seja, com restrição de progresso zero para os r caracteres restritos. A partir da expressão do ganho esperado com a seleção pelo índice clássico (equação 2), Kempthorne e Nordskog (1959) determinam uma função linear de modo que a correlação entre os r caracteres e os valores fenotípicos discriminantes em $\boldsymbol{I}$ seja zero, enquanto que seja máxima para os demais caracteres, resultando no seguinte estimador:

$$\hat{\boldsymbol{\beta}} = \left[\boldsymbol{I}^* - \boldsymbol{P}^{-1}\boldsymbol{G}_r(\boldsymbol{G}_r'\boldsymbol{P}^{-1}\boldsymbol{G}_r)^{-1}\boldsymbol{G}_r'\right]\boldsymbol{P}^{-1}\boldsymbol{G}\boldsymbol{a}\,,$$

em que $\boldsymbol{I}^*$ refere-se a uma matriz identidade de ordem n; e $\boldsymbol{G}_r$ é uma submatriz de $\boldsymbol{G}$ com as colunas dos r caracteres restritos. Prova-se que quando não se impõem restrições o estimador se reduz ao índice clássico. Quanto maior o número de caracteres sujeitos a restrições, a correlação entre o índice e o agregado genótipo decai; pontua-se também que ao restringir o ganho em um dado caráter, o peso econômico associado a tal caráter torna-se irrelevante (KEMPTHORNE; NORDSKOG, 1959).

Generalizando Kempthorne e Nordskog, Tallis (1985) considera o caso em que de r caracteres restritos sejam requeridos progressos genéticos proporcionais a uma constante pré-determinada. A derivação da solução algébrica desta situação constitui-se em um índice que maximiza os ganhos no agregado genotípico sujeito a r restrições de que a covariância entre os r caracteres restritos e os valores fenotípicos discriminantes sejam proporcionais ao ganho desejado nesses caracteres, ou seja $\boldsymbol{G\beta} = \boldsymbol{\theta q}$, em que $\boldsymbol{\theta}$ é um vetor de constantes determinado *a posteriori* e $\boldsymbol{q}$ é o vetor com os ganhos desejados para os r caracteres restritos.

Ao resolver o sistema com as restrições, Tallis (1985) encontra o seguinte estimador:

$$\hat{\boldsymbol{\beta}} = \left[\boldsymbol{I}^* - \boldsymbol{P}^{-1}\boldsymbol{G}_r(\boldsymbol{G}_r'\boldsymbol{P}^{-1}\boldsymbol{G}_r)^{-1}\boldsymbol{G}_r'\right]\boldsymbol{P}^{-1}\boldsymbol{G}_{(n-r)}\boldsymbol{a}_{(n-r)} + \theta\boldsymbol{P}^{-1}\boldsymbol{G}_r(\boldsymbol{G}_r'\boldsymbol{P}^{-1}\boldsymbol{G}_r)^{-1}\boldsymbol{q}\,, \quad (4)$$

em que $\boldsymbol{G}_{(n-r)}$ refere-se a submatriz das colunas de $\boldsymbol{G}$ apenas dos caracteres sem restrição enquanto que $\boldsymbol{a}_{(n-r)}$ são apenas os pesos econômicos dos caracteres sem restrição. Os demais termos já foram definidos.

O processo de Tallis (1985) generaliza a imposição da restrição de Kempthorne e Nordskog; pela solução apresentada, quando os ganhos são restritos a zero, a expressão reduz-se ao índice de Kempthorne e Nordskog; quando se impõem restrições em todos os caracteres $r = n$, obtém-se $\boldsymbol{\beta} = \boldsymbol{G}^{-1}\boldsymbol{\Delta}$, que é o índice de ganhos desejados de Pesek e Baker; e no caso degenerado, quando nenhum restrição é imposta $r = 0$, resulta no índice ordinário.

Entretanto, Mallard (1972 apud ITOH; YAMADA, 1987) argumenta que impor restrições para que os ganhos sejam proporcionais é possível através de restrições de ganho zero para

alguma função linear de caracteres. Portanto o método de Kempthorne e Nordskog (1959) é mais generalista e pode ser aplicável nesse intento, bastando apenas reparametrizar as restrições.

Harville (1975) usa uma abordagem diferente com o mesmo intento, maximiza a correlação entre o agregado genotípico e os valores fenotípicos discriminantes, assumindo que a variância do índice seja unitária ($\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta} = 1$), o resultado é um índice de seleção que maximiza os progressos e independe dos pesos econômicos.

Itoh e Yamada (1987) mostram que o vetor de soluções do processo de Tallis (eq. 4) pode ser decomposto em dois termos: $\boldsymbol{\beta} = \boldsymbol{\beta}_{K\&N} + \boldsymbol{\beta}_H$; $\boldsymbol{\beta}_{K\&N}$ representando os coeficientes do índice restrito de Kempthorne e Nordskog, enquanto que os coeficientes de $\boldsymbol{\beta}_H$ maximizam os progressos, independentemente de pesos econômicos (método derivado por Harville). O que fica evidente é que ambos os índices de Harville e Tallis, guardam uma equivalência com o método de Kempthorne e Nordskog, a diferenca entre ambos relaciona-se apenas às diferenças de escala e parametrização (ITOH; YAMADA, 1987).

Os processos de estimação dos índices restritos apresentam meios de impor restrições de apenas uma natureza (LIN, 2005b). O índice de Kempthorne e Nordskog restringe nos r caracteres um progresso genético nulo e o processo de Tallis, dado um vetor $\boldsymbol{q}$ $(r \times 1)$ com os progressos almejados aos r caracteres restritos, restringe os ganhos proporcionalmente a $\boldsymbol{q}$ (LIN, 1978), como já comentado. A diferença entre ambos os métodos relaciona-se à parametrização.

Deste modo, Lin (2005b) deriva um modo de impor restrições simultâneas: seja o caso, que em r caracteres impõem-se a restrição de Tallis, enquanto que em s caracteres são impostos progressos desejados "fixos" ($\boldsymbol{c}$). Essa restrição comporta inclusive a restrição de Kempthorne e Nordskog, sendo $(r + s) \leq n$. Como apresentado por Mallard (1972 apud ITOH; YAMADA, 1987), além da proporção dos ganhos, ou ganhos "fixos", é possível impor restrições com forma $z_i\Delta G_i + z_j\Delta G_j = v$, em que z_i, z_j e v são constantes pré determinadas, ou seja, é possível restringir uma combinação linear para os progressos nos caracteres i e j (LIN, 2005b).

Lin (2005b) minimiza a variância da diferença entre $\boldsymbol{I}$ e $\boldsymbol{H}$, diferenciando simultaneamente em relação a $\boldsymbol{\beta}$ e aos multiplicadores de *Lagrange* das restrições. Igualando as derivadas parciais a zero, e resulta o seguinte sistema:

$$\begin{bmatrix} \boldsymbol{P} & \boldsymbol{G}_r & & \boldsymbol{G}_s & z_i\boldsymbol{G}_i + z_j\boldsymbol{G}_j \\ \boldsymbol{G}_r' & & -\boldsymbol{q} & & \\ & -\boldsymbol{q}' & & & \\ \boldsymbol{G}_s' & & & & \\ z_i\boldsymbol{G}_i' + z_j\boldsymbol{G}_j' & & & & \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\lambda} \\ \boldsymbol{\theta} \\ \boldsymbol{\delta} \\ \xi \end{bmatrix} = \begin{bmatrix} \boldsymbol{Ga} \\ \\ \\ \boldsymbol{c} \\ v \end{bmatrix} \tag{5}$$

Tal sistema, desconsideradas as restrições fica reduzido ao caso ordinário de Smith. Ao retirar a segunda, terceira e quinta restrições do sistema (eq. 5) resulta no método proposto por Tallis (1985), o mesmo ocorrendo ao se deixar apenas a segunda e terceira restrições. Mantendo-se apenas o segundo conjunto de restrições tem-se o processo de Kempthorne e Nordskog (LIN, 2005b).

Pelo processo de Tallis (1985) e no que é empregado em Lin (2005b) (eq. 4 e 5), ao impor para um caráter i a restrição $q = 1/2$, decorre que o progresso neste caráter restringe-se a 0,5 desvios padrões dos valores fenotípicos discriminantes (σ_I). Isso é equivalente a impor que $\boldsymbol{G}_i'\hat{\boldsymbol{\beta}} = 1/2$, assumindo uma intensidade de seleção igual a um (LIN, 2005a).

Lin (2005a) estende o trabalho anterior (Lin (2005b)) com relação à imposição de restrições constantes, pois tais restrições só tem significado prático quando associadas a uma intensidade de seleção (κ). Em Lin (2005b), o método impõe restrições "fixas", constantes em termos de desvio padrão do índice, enquanto que em Lin (2005a) tais restrições são constantes relativas ao progresso genético almejado e estão associadas a uma certa intensidade de seleção.

O progresso genético do caráter i após a seleção baseada no índice é estimado por $R_i = \kappa\frac{\boldsymbol{G}_i'\boldsymbol{\beta}}{\sigma_I}$, em que: R_i é o progresso genético do caráter i; κ é a intensidade de seleção praticada na seleção por meio do índice; $\boldsymbol{G}_i$ é a coluna da matriz $\boldsymbol{G}$ referente ao caráter i; $\boldsymbol{\beta}$ é o vetor de coeficientes do índice; e σ_I é o desvio padrão dos valores fenotípicos discriminantes. Ao obter o mínimo da função da variância da diferença sujeita a restrições do progresso dos r caracteres, Lin (2005a) apresenta:

$$\begin{bmatrix} \boldsymbol{P} & \dfrac{\kappa\boldsymbol{G}_r}{\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}} - \dfrac{\boldsymbol{P}\boldsymbol{\beta}\boldsymbol{c}'}{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}} \\ \dfrac{\kappa\boldsymbol{G}_r'}{\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}} & \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}\boldsymbol{a} \\ \boldsymbol{c} \end{bmatrix} \tag{6}$$

Algumas particularidades caracterizam tal sistema (eq. 6), pois a matriz dos coeficientes contém o próprio vetor de soluções $\boldsymbol{\beta}$; assim, a solução para os coeficientes do índice deve ser obtida por processo iterativo. Lin (2005a) observa que a matriz dos coeficientes possui posto completo $(n + r)$ e tem inversa única, portanto iterações sobre o sistema garantem a convergência da solução. O sistema depende da intensidade de seleção κ, sabe-se que o progresso genético é função da intensidade de seleção, ou seja, aumentando-se a intensidade aumenta-se o progresso genético. Todavia, fica evidente que sob tal restrição o relacionamento entre o progresso genético e a intensidade de seleção não é linear (LIN, 2005a).

Ao impor como restrições um vetor de zeros o sistema se reduz ao processo de Kempthorne e Nordskog (1959). Desconsiderando a quantidade $\frac{\kappa}{\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}}$, por não afetar a proporcionalidade obtem-se um sistema idêntico ao apresentado por Cunningham, Moen e Gjedrem (1970) e com solução que independe da intensidade de seleção. Pelo apresentado, é possível conjugar todos os tipos de restrições em um mesmo sistema, e ao manipular as restrições é possível verificar o relacionamento entre as diferentes abordagens da construção de índices restritos (LIN, 2005a).

Para efeito da derivação Ceron-Rojas et al. (2006), reparametriza o índice clássico, normalizando os coeficientes de $\boldsymbol{\beta}$, de modo que $\boldsymbol{\beta}'\boldsymbol{\beta} = 1$. Ao avaliar o mínimo da variância da diferença sujeita a essa restrição, estes autores obtiveram:

$$(\boldsymbol{P} - \lambda\boldsymbol{I}^*)\boldsymbol{\beta} = 0\,. \tag{7}$$

A relação assim encontrada (eq. 7) é análoga à decomposição espectral de $\boldsymbol{P}$, em que λ é um autovalor e $\boldsymbol{\beta}$ é o respectivo autovetor de λ (FERREIRA, 2008). Como já visto, $\boldsymbol{P}$ é singular, pelas propriedades da decomposição espectral. Segundo Gentle (2007), todos os autovalores (λ) de $\boldsymbol{P}$ são diferentes entre si. Como o vetor $\boldsymbol{\beta}$ está sendo interpretado como um autovetor de $\boldsymbol{P}$, a variância do índice é o autovalor correspondente ($\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta} = \lambda$). O estimador equivalente à eq. 2, expresso como função da intensidade de seleção (κ), da variância do índice e da matriz de variâncias e covariâncias fenotípicas é igual a $\Delta_I = \kappa\sqrt{\lambda}$.

O método de obtenção do índice exposto por Ceron-Rojas et al. (2006) é superior ao original de Smith, por ser mais direto e não necessitar da obtenção de variâncias e covariâncias genéticas e por independer dos pesos econômicos. Neste sentido, a solução dos coeficientes não considera qualquer importância "econômica" específica, empírica do selecionador diante dos caracteres (CERON-ROJAS et al., 2006).

Na prática dos programas de melhoramento, emprega-se usualmente uma mistura dos métodos de seleção: a seleção, para alguns caracteres, pode ser realizada em *tandem*; níveis independentes de eliminação, por sua vez, podem servir para descartar indivíduos em estágios iniciais; enquanto que para outros caracteres, a seleção pode ser feita utilizando diferentes pesos econômicos intuitivamente em um índice (WRICKE; WEBER, 1986, capítulo 12).

Por essência, fundamento e desenvolvimento, os diferentes índices proporcionam diferentes resultados, apresentando-se como alternativa uns aos outros. Mas todos eles mostram-se como um meio de classificar genótipos promissores e adequados aos propósitos do melhoramento genético (CRUZ; REGAZZI; CARNEIRO, 2001, capitulo 6). A melhor técnica a ser utilizada depende da natureza dos dados e dos objetivos do pesquisador. Fisher (1936) apresenta a teoria das funções discriminantes para discriminar diferentes espécies do gênero *Iris*; Smith (1936) estende essa teoria para predizer um dado caráter não observável como função de outros caracteres e Hazel (1943) aplica tais funções discriminantes com o propósito da seleção. Assim, pode-se pensar em adaptar a teoria dos ídices de seleção quando o propósito, ao invés de combinar fenótipos de múltiplos caracteres, é combinar progressos genéticos de múltiplos caracteres.

2.2 Material e Métodos

2.2.1 Progressos Genéticos Univariados

Com o objetivo de se mensurar o progresso genético obtido ao longo de 30 anos da pesquisa com tabaco, realizou-se no ano agrícola de 2011/2012 a comparação entre cultivares do grupo varietal *Virginia* quanto aos caracteres de interesse econômico da cultura. Tais tratamentos genéticos são provenientes de diferentes "eras" do programa de melhoramento de tabaco.

O ensaio constou de quatro experimentos instalados nas localidades a saber: duas no município de Mafra, estado de Santa Catarina, tendo como localização geográfica 26° 04' 59,8"S; 49° 58' 51,2"W com 880m de altitude (*M1*) e 25° 57' 28,8"S; 49° 50' 23,4"W e 785m de altitude (*M2*); uma no município de Rio Negro, 25° 57' 28,8"S; 49° 50' 23,4"W com

785m de altitude, aqui denominado *RN* e outro no município de Piên localizado a 26° 06' 27,5"S; 49° 30' 50,2"W com 830m de altitude (*PN*) ambos no estado do Paraná. Tais ambientes foram escolhidos por representarem a região de cultivo da cultura no país.

Os nove tratamentos foram cultivares selecionados que se distribuem em quatro fases do programa de melhoramento de tabaco do grupo varietal *Virginia*. Os cultivares (*K326* e *MN373*) são introduções de germoplasma de tabaco provenientes dos Estados Unidos da América e representam o início do período estudado, caracterizando a "era" 0. Dois outros cultivares foram obtidos como resultado dos primeiros esforços praticados no melhoramento da cultura aqui no país (*CSC06* e *CSC07*) e representam a "era" 1. Os cultivares (*CSC439*, *CSC444* e *CSC447*) fazem parte da "era" 2, sendo linhagens que foram obtidas pela seleção de genótipos de tabaco já previamente melhorados nas condições edafoclimáticas locais enquanto que os demais cultivares ("era" 3), *CSC4703* e *CSC4704*, representam os últimos esforços do programa em termos de melhoramento genético. Todos os genótipos empregados no experimento são cultivares registrados pela companhia detentora, e quando não, são porque se destacam em cultivos comerciais e/ou experimentais.

Os experimentos foram instalados na época ideal, e conduzidos de acordo com as recomendações preconizadas para a cultura para que se possa avaliar o máximo potencial produtivo. Todos os tratos culturais foram os recomendados para o cultivo comercial do tabaco do referido grupo varietal. Mudas de cada tratamento dos experimentos foram produzidas usando sementes renovadas no sistema *floating* em casa de vegetação e transplantadas em linhas que distavam em 1,20m. O espacejamento entre plantas dentro de linhas foi de 0,50m. As parcelas foram constituídas de duas linhas com dez plantas, perfazendo 20 plantas por parcela, ou 720 plantas por experimento. Não foram consideradas bordaduras por parcela, mas alocou-se uma linha de bordadura ao redor do experimento. O delineamento experimental foi o casualizado em blocos completos com quatro repetições.

Na fase imediatamente anterior à capação[3] os seguintes caracteres foram avaliados em cinco plantas competitivas de cada parcela: altura (*alt*), medida em metros; circunferência do caule no colo da planta (ccc), expresso em centímetros (cm); o número de folhas expandidas (*nfe*); e a largura e comprimento da terceira folha, respectivamente *l3f* e *c3f*, ambos também medidos em centímetros (cm). Além do número de plantas sobreviventes por parcela (estande).

Conforme o amadurecimento das folhas, as mesmas foram colhidas manualmente e curadas em estufa de ventilação forçada. Os dados brutos do rendimento de folhas curadas foram corrigidos para o estande ideal pelo método da covariância (RAMALHO; FERREIRA; OLIVEIRA, 2005, capítulo 13). Assim, obteve-se o peso de folhas curadas (*pfc*) no total de todas as colheitas praticadas, expresso em quilogramas por hectare ($kg\,ha^{-1}$).

As folhas de tabaco curadas de cada parcela passaram por um processo de avaliação conforme a Instrução Normativa de número 10 do Ministério da Agricultura Pecuária e Abastecimento (BRASIL, 2007) que regula a classificação comercial do tabaco curado

[3]Procedimento padrão de manejo da cultura que elimina a inflorescência.

resultando no índice da qualidade geral, sigla *iqg*, e consequentemente no valor comercial, ou a remuneração que o produtor receberia em *US$* em um processo de venda regulamentado. Os valores foram posteriormente extrapolados para a produção de um hectare (*$ha*). Esses caracteres podem ser interpretados como "caracteres índices".

Os dados foram gentilmente cedidos pelo Dr. Carlos Eduardo Pulccinelli (egresso do programa), que é Pesquisador Científico Sênior da Companhia Souza Cruz, líder nacional na produção e exportação de tabaco.

Com a finalidade de obter informações úteis para o melhoramento, os dados foram submetidos a análise de variância em duas etapas, a primeira contemplando o delineamento experimental de acordo com o seguinte modelo:

$$Y_{ijk} \sim \mu + a_i + r_{j/i} + t_k + (at)_{ik} + \varepsilon_{ijk}\,. \tag{8}$$

em que:
Y_{ijk} é a observação fenotípica da parcela no local i, na repetição j, e tratamento k;
μ é a a constante inerente a todas as observações, que pela restrição imposta é a média geral dos dados;
a_i é o efeito fixo da local i, com $i = 1, 2, \ldots, 4$;
$r_{j/i}$ é o efeito aleatório aninhado das repetições j dentro de locais i, com $j = 1, 2, \ldots, 4$;
t_k é o efeito fixo do genótipo k, com $k = 1, 2, \ldots, 9$;
$(at)_{ik}$ é o efeito da interação entre o local i e o genótipos k; e
ε_{ijk} é o resíduo relativo à parcelas ijk, com $\varepsilon \cap N(0, \sigma^2_\varepsilon)$.

Para as variáveis em que foram tomadas mais de uma amostra por parcela: (*alt*, *ccc*, *nfe*, *l3f*, *c3f*), incluiu-se o efeito do resíduo dentro de parcelas δ_{ijkl}, em que $l = 1, 2, \ldots, 5$ é o índice que identifica as amostras (plantas) dentro de cada parcela, assumido também como normal e identicamente distribuído, ou seja, $\delta \cap N(0, \sigma^2_\delta)$.

Afim de averiguar se a análise de variância realizada contemplou suas pressuposições inerentes, estimou-se a estatística K de Bartlett para testar a homegeneidade das variâncias dos experimentos instalados nos quatro locais. Essa estatística segue distribuição de χ^2_{k-1} e sua função de probabilidade foi usada para testar a hipótese nula de que as variâncias residuais são homogêneas (STEEL; TORRIE; DICKEY, 1996).

Dada a homogeneidade de variâncias, o pressuposto da normalidade dos resíduos foi verificado por meio da estimativa da estatística W de *Shapiro–Wilk* para testar se os efeitos residuais da análise conjunta eram normais e identicamente distribuídos (ROYSTON, 1995). A estimativa de W será apresentada acompanhada do valor da significância do teste. A pressuposição da aditividade do modelo, por sua vez, será testada por meio do teste F de *Tukey*, sobre o quadrado médio da não aditividade entre os efeitos de locais e genótipos, conforme Ramalho, Ferreira e Oliveira (2005, capítulo 3). A título de verificar a precisão/qualidade do ajuste do modelo, estimou-se também o coeficiente de variação ($CV(\%)$) e a acurácia seletiva ($Ac(\%)$).

No segundo momento da análise de variância almejou-se desdobrar no modelo (eq. 8), a fonte de variação genótipos (t_k) e as respectivas interações com locais ($(at)_{ik}$) nas diferenças que podem ser devido às diferentes "eras" e aquela que representa a variação média entre genótipos dentro de cada "era". A fonte de variação devida a "eras" foi por sua vez decomposta na componente linear que afere o progresso genético e os desvios do respectivo ajuste linear (PIMENTEL GOMES, 2009).

Dessa forma o modelo 8 foi estendido incorporando-se os efeitos de interesse ao estudo do progresso genético, resultando o seguinte resumo esquemático da análise de variância empregada, apresentada na Tabela 2:

Tabela 2 – Esquema da análise de variância, fontes de variação, graus de liberdade e respectivas esperanças dos quadrados médios com desdobramentos complementares ao estudo do progresso genético

F.V.		*g.l.*	*E[Q.M.]*
Ambientes	3	$(a-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + rg\phi_a + g\sigma^2_{r/a}$
Repetições / Ambientes	12	$a(r-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + g\sigma^2_{r/a}$
Genótipos	8	$(g-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + ra\phi_g$
Eras	3	$(p-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + ra\phi_p$
Progresso Linear	1	1	$n\sigma^2_\delta + \sigma^2_\epsilon + ra\phi_l$
Desvios	2	$(p-1-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + ra\phi_d$
Genótipos / Eras	5	$(g-p-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + ra\phi_{g/p} + r\phi_{ag/p}$
Ambientes × Genótipos	24	$(a-1)(g-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + r\phi_{ag}$
Ambientes × Eras	9	$(a-1)(p-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + r\phi_{ap}$
Ambientes × Progresso	3	$(a-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + r\phi_{al}$
Ambientes × Desvios	6	$(a-1)(p-1-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + r\phi_{ad}$
Ambientes × Genótipos / Eras	15	$(a-1)(g-p-1)$	$n\sigma^2_\delta + \sigma^2_\epsilon + r\phi_{ag/p}$
Resíduo	96	$a(1+rg-r-g)$	$n\sigma^2_\delta + \sigma^2_\epsilon$
Erro Dentro	576	$arg(n-1)$	σ^2_δ

As inferências serão realizadas pela interpretação das estimativas dos componentes de variâncias e dos parâmetros do modelo: médias dos tratamentos, médias das eras e estimativa do coeficiente angular da regressão que estima o progresso genético e pelo valor de significância dos respectivos testes. Todas as análises foram feitas e/ou implementadas usando o *software* `R` (R Development Core Team, 2014).

2.2.2 Índice de Seleção com Diferenciais de Seleção Restritos

Seja, como definido por Smith (1936), um índice $\boldsymbol{I}$, que sumariza e classifica p genótipos em relação a n valores fenotípicos:

$$\boldsymbol{I} = \sum_{i=1}^{n} \boldsymbol{x_i}\boldsymbol{\beta}_i = \boldsymbol{X}\boldsymbol{\beta}\,,$$

em que:
n e p são definidos como, respectivamente o montante de informações fenotípicas e o número de genótipos avaliados;
$\boldsymbol{\beta}$ o vetor de dimensão $n \times 1$ dos coeficientes de ponderação do índice; e
$\boldsymbol{X}$ a matriz $p \times n$ das médias dos n caracteres considerados de p genótipos (vetor linha representado por $\boldsymbol{x_i}$).

Seja ainda $\boldsymbol{H}$ um vetor de agregados genotípicos, que, segundo Kempthorne e Nordskog (1959) referem-se somente a efeitos aditivos, de modo que $\boldsymbol{H}$ denota o vetor de valores reprodutivos (*breeding values*):

$$\boldsymbol{H} = \sum_{i=1}^{m} \boldsymbol{z}_i \boldsymbol{a}_i = \boldsymbol{Z}\boldsymbol{a}\,,$$

em que:
$\boldsymbol{a}$ um vetor de dimensão $m \times 1$ dos "valores econômicos" (pré determinados); e
$\boldsymbol{Z}$ é a matriz $p \times m$ de m informações genéticas desconhecidas de p genótipos avaliados (vetor linha representado por $\boldsymbol{z_i}$).

As quantidades $\boldsymbol{P}$ e $\boldsymbol{G}$ são respectivamente as matrizes $n \times n$ e $m \times m$ de variâncias e covariâncias dos $\boldsymbol{x_i}$ e $\boldsymbol{z_i}$, ou seja, as matrizes de variâncias e covariâncias das informações fenotípicas e genéticas. Lin (1978) oportunamente mostra que, por definição, a covariância entre $\boldsymbol{x_i}$ e $\boldsymbol{z_i}$ é representada por $\boldsymbol{C}$, uma matriz $n \times m$ com as covariâncias entre os n valores fenotípicos e as m informações genéticas desconhecidas. Quando, $m = n$, $\boldsymbol{C} = \boldsymbol{G}$, já que a covariância entre fenótipo e genótipo é função apenas da variância genética.

Estas quantidades, $\boldsymbol{I}$ e $\boldsymbol{H}$, são funções lineares em $\boldsymbol{\beta}$ e $\boldsymbol{a}$, respectivamente. Pelas propriedades apresentadas em Searle (1997), decorre que a forma quadrática (variância) do índice ($\boldsymbol{I}$) é $\sigma^2_{\boldsymbol{I}} = \boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}$; e a variância dos agregados genotípicos ($\boldsymbol{H}$) é $\sigma^2_{\boldsymbol{H}} = \boldsymbol{a}'\boldsymbol{G}\boldsymbol{a}$. A covariância entre $\boldsymbol{I}$ e $\boldsymbol{H}$ é definida como $\sigma_{\boldsymbol{IH}} = \boldsymbol{\beta}'\boldsymbol{G}\boldsymbol{a}$, considerando $m = n$. Por outro lado, a correlação entre os valores do índice e os agregados genotípicos nos p genótipos é:

$$\rho_{\boldsymbol{IH}} = \frac{\boldsymbol{a}'\boldsymbol{G}\boldsymbol{\beta}}{\sqrt{\boldsymbol{a}'\boldsymbol{G}\boldsymbol{a}}\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}}\,.$$

A expressão do ganho genético para a seleção em um dado caráter (FALCONER; MACKAY, 1996, página 186, capítulo 11) é:

$$R = i\sigma_G\rho\,, \tag{9}$$

em que:
i é intensidade de seleção estandartizada, ou razão entre o diferencial de seleção e o desvio padrão genotípico $\left(\frac{\Delta}{\sigma_P}\right)$;
σ_G é o desvio padrão genotípico (raiz quadrada da variância genética); e
ρ é a acurácia seletiva, medida como a correlação entre os valores genéticos e fenotípicos (raiz quadrada da herdabilidade).

Outra apresentação da expressão do ganho (eq 9) é conhecida como *breeder's formula* ($R = h^2\Delta$), que mensura o ganho como função da herdabilidade (h^2) e do diferencial de seleção (Δ). Analogamente, a resposta à seleção $R_{\boldsymbol{I}}$, praticada seleção pelo índice $\boldsymbol{I}$, dada uma intensidade de seleção do índice (κ), é tida como:

$$R_{\boldsymbol{I}} = \kappa\sigma_{\boldsymbol{H}}\rho_{\boldsymbol{IH}}\,.$$

Segundo Hazel (1943), $\sigma^2_{\boldsymbol{H}}$ depende unicamente de κ, e $\boldsymbol{a}$ é um vetor de constantes. Portanto, maximizar os ganhos é o mesmo que maximizar a correlação entre o agregado genotípico ($\boldsymbol{H}$) e o índice ($\boldsymbol{I}$). Wricke e Weber (1986) provam que o estimador obtido maximizando-se a correlação equivale ao obtido minimizando a variância da diferença entre os valores fenotípicos discriminantes e o agregado genotípico $\left(\sigma^2_{(\boldsymbol{I}-\boldsymbol{H})}\right)$. Procedendo-se desse modo, obtêm-se:

$$\sigma^2_{(\boldsymbol{I}-\boldsymbol{H})} = \sigma^2_{\boldsymbol{I}} + \sigma^2_{\boldsymbol{H}} - 2\sigma_{\boldsymbol{IH}} = \boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta} + \boldsymbol{a}'\boldsymbol{G}\boldsymbol{a} - 2\boldsymbol{\beta}'\boldsymbol{G}\boldsymbol{a}\,. \tag{10}$$

Derivando-se tal expressão (eq. 10) em relação a $\boldsymbol{\beta}$ e igualando a zero resulta que $\boldsymbol{P}\boldsymbol{\beta} = \boldsymbol{G}\boldsymbol{a}$.

A título de demonstração apresenta-se a derivação do ponto de máximo do logaritmo da correlação ($\rho_{\boldsymbol{IH}}$) como prova da equivalência entre as soluções.

$$\begin{aligned}
\frac{\partial}{\partial\boldsymbol{\beta}}\log_e(\rho_{\boldsymbol{IH}}) &= \frac{\partial}{\partial\boldsymbol{\beta}}\log_e(\boldsymbol{a}'\boldsymbol{G}\boldsymbol{\beta}) - \frac{1}{2}\frac{\partial}{\partial\boldsymbol{\beta}}\log_e(\boldsymbol{a}'\boldsymbol{G}\boldsymbol{a}) - \frac{1}{2}\frac{\partial}{\partial\boldsymbol{\beta}}\log_e(\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}) \\
&= \frac{\boldsymbol{G}\boldsymbol{a}}{\boldsymbol{a}'\boldsymbol{G}\boldsymbol{\beta}} - \frac{\boldsymbol{P}\boldsymbol{\beta}}{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}} \\
\boldsymbol{P}\boldsymbol{\beta} &= \frac{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}{\boldsymbol{a}'\boldsymbol{G}\boldsymbol{\beta}}\boldsymbol{G}\boldsymbol{a}\,.
\end{aligned} \tag{11}$$

A fração $\dfrac{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}{\boldsymbol{a}'\boldsymbol{G}\boldsymbol{\beta}}$ na eq. 11 é considerada constante e a correlação invariante a mudanças de escala. Sendo uma constante, o termo pode assumir qualquer valor no intervalo $\mathbb{R}$, inclusive 1 e deste modo, pode ser retirado sem afetar a proporcionalidade em $\boldsymbol{\beta}$.

Como $\boldsymbol{P}$ é simétrica, positiva definida e não singular (com inversa única), obtêm-se portanto, a solução do índice de Smith (1936): $\hat{\boldsymbol{\beta}} = \boldsymbol{P}^{-1}\boldsymbol{G}\boldsymbol{a}$ (eq. 3).

Depreende-se por ambas as demonstrações (eq. 10 e 11) que existe uma relação de igualdade, $\boldsymbol{P}\boldsymbol{\beta} = \boldsymbol{G}\boldsymbol{a}$. Outra observação pertinente é notar que, ao considerar $m = n$, a covariância entre os valores do índice e os agregados genéticos ($\sigma_{\boldsymbol{IH}}$) é igual à variância do índice ($\sigma^2_{\boldsymbol{I}}$). Por conseguinte, $\rho_{\boldsymbol{IH}} = \dfrac{\sigma_{\boldsymbol{I}}}{\sigma_{\boldsymbol{H}}}$ e deste modo, resultados equivalentes aos obtidos podem ser alcançados maximizando a variância do índice $\sigma^2_{\boldsymbol{I}}$.

$\hat{\boldsymbol{\beta}}$ é um vetor $n \times 1$ com os n coeficientes que ponderam os valores fenotípicos constituindo o índice; é a função linear dos valores fenotípicos que maximiza o ganho na média dos n valores fenotípicos empregados. O ganho em unidades do índice dada uma intensidade de seleção κ é

expresso como $R_{\boldsymbol{I}} = \kappa \dfrac{\boldsymbol{a}'\boldsymbol{G}\boldsymbol{b}}{\sigma_{\boldsymbol{I}}}$, enquanto que o diferencial de seleção, em unidades do índice ao praticar seleção por meio do mesmo é $\Delta_{\boldsymbol{I}} = \kappa \dfrac{\boldsymbol{P}\boldsymbol{\beta}}{\sigma_{\boldsymbol{I}}}$.

Estendendo tais estimadores, para o j–ésimo caráter que compõe o índice, o diferencial de seleção "realizado", feita a seleção pelo índice com intensidade κ, é obtido por:

$$\Delta_j = \kappa \frac{\boldsymbol{P}_j'\boldsymbol{\beta}}{\sigma_{\boldsymbol{I}}} \,,$$

em que $\boldsymbol{P}_j$ é uma submatriz de $\boldsymbol{P}$, ou seja, a j–ésima coluna da matriz de variâncias–covariâncias fenotípicas.

Tomando-se r, com ($r \leq n$), como o número de caracteres sujeitos a restrições nos diferenciais de seleção, compõe-se $\boldsymbol{c}$ como um vetor $r \times 1$ com diferenciais de seleção para tais r caracteres restritos.

A solução de $\boldsymbol{\beta}$ que satisfaz à restrição é obtida utilizando-se multiplicadores de *Lagrange*. Obtendo o mínimo da função da variância da diferença do índice e do vetor de agregados genotípicos $\sigma^2_{(\boldsymbol{I}-\boldsymbol{H})}$, sujeita à restrição: $\kappa \dfrac{\boldsymbol{P}_{[r]}'\boldsymbol{\beta}}{\sigma_{\boldsymbol{I}}} = \boldsymbol{c}$, resulta que:

$$\begin{aligned} f &= \sigma^2_{(\boldsymbol{I}-\boldsymbol{H})} + \boldsymbol{\lambda}' \left(\kappa \frac{\boldsymbol{P}_{[r]}'\boldsymbol{\beta}}{\sigma_{\boldsymbol{I}}} - \boldsymbol{c} \right) \\ &= \boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta} + \boldsymbol{a}'\boldsymbol{G}\boldsymbol{a} - 2\boldsymbol{\beta}'\boldsymbol{G}\boldsymbol{a} + \boldsymbol{\lambda}' \left(\kappa \frac{\boldsymbol{P}_{[r]}'\boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}} - \boldsymbol{c} \right) . \end{aligned} \tag{12}$$

Derivando-se (eq. 12) em relação a $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ e igualando-se a zero obtêm-se:

$$\frac{\partial f}{\partial \boldsymbol{\beta}} = 2\boldsymbol{P}\boldsymbol{\beta} + 2\boldsymbol{G}\boldsymbol{a} + \left(\kappa \frac{\boldsymbol{P}_{[r]}}{\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}} - \frac{\boldsymbol{P}\boldsymbol{\beta}\boldsymbol{c}'}{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}} \right) \boldsymbol{\lambda} = 0 \text{ e} \tag{13}$$

$$\frac{\partial f}{\partial \boldsymbol{\lambda}} = \kappa \frac{\boldsymbol{P}_{[r]}'\boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}} - \boldsymbol{c} = 0 \,. \tag{14}$$

Rearranjando os termos em (eq. 13 e 14) sem afetar a proporcionalidade nas funções diferenciadas pode-se compor o seguinte sistema matricial:

$$\begin{bmatrix} \boldsymbol{P} & \kappa \dfrac{\boldsymbol{P}_{[r]}}{\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}} - \dfrac{\boldsymbol{P}\boldsymbol{\beta}\boldsymbol{c}'}{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}} \\ \kappa \dfrac{\boldsymbol{P}_{[r]}'}{\sqrt{\boldsymbol{\beta}'\boldsymbol{P}\boldsymbol{\beta}}} & \underset{\sim}{\boldsymbol{0}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}\boldsymbol{a} \\ \boldsymbol{c} \end{bmatrix} \tag{15}$$

Quando nenhum caráter recebe restrição, o sistema (eq. 15) se reduz à igualdade $\boldsymbol{P}\boldsymbol{\beta} = \boldsymbol{G}\boldsymbol{a}$. Constata-se que o sistema caracteriza-se como um caso especial do índice de Smith (1936). Se $r = 0$, nenhum caráter recebe restrição quanto ao diferencial de seleção. Nesse caso o emprego da solução obtida pelo sistema iguala-se ao que seria obtido empregando-se o índice ordinário de Smith (1936).

Para o emprego do índice apresentado é necessário que se estimem variâncias e covariâncias fenotípicas e se atribuam diferenciais de seleção para os diferentes caracteres. Vê-se que o sistema (equação 15) apresenta a matriz de variâncias e covariâncias genéticas, bem como o vetor de pesos econômicos ($\boldsymbol{Ga}$). Vale destacar que esse é o resultado obtido minimizando-se a variância da diferença. Quando por sua vez, se maximiza a variância do índice observa-se que $\boldsymbol{Ga}$ pode ser substituído por $\boldsymbol{P\beta}$ e assim, para estimar os coeficientes do índice, não é necessário estimar variâncias e covariâncias genéticas, nem atribuir pesos econômicos determinados *a priori*.

Diante do exposto, tal implementação do índice de Smith (1936), sujeito a restrições nos diferenciais de seleção, mostra-se como uma sistematização da seleção por níveis não independentes de eliminação, um avanço sobre *independent culling levels*, que, contudo, considera o relacionamento entre os caracteres (covariância fenotípica entre os caracteres). Outro ponto que merece destaque é o fato de o índice, em uma de suas formas, independer de pesos econômicos, pois a importância de cada caráter é atribuída pelo respectivo diferencial de seleção que se almeja impingir.

2.2.3 Exemplo de Aplicação

Diuturnamente o melhorista de plantas vislumbra a possibilidade de praticar a seleção para múltiplos caracteres. Porém, é sabido que, ao contrário do melhoramento animal, arbitrar pesos econômicos aos caracteres é extremamente subjetivo. Isso restringe em muito o emprego do índice ordinário de Smith (1936), que depende de pesos econômicos determinados *a priori*. Alternativamente, porém, o melhorista pode atribuir, para cada caráter, diferentes diferenciais de seleção. Tal processo tem a vantagem de que o diferencial de seleção, conforme a respectiva intensidade seletiva, é um modo de distinguir os diferentes caracteres quanto à importância que se almeja aplicar na seleção. Além disso, o diferencial é facilmente interpretável como unidades de ganho, visto que está na mesma unidade da variável sob análise.

Como resultado da análise de variância mostrada na seção 2.2.1, pode-se interpretar, quando o teste F da fonte de variação Progresso for significativo, que o coeficiente angular da regressão entre as médias das diferentes "eras" é diferente de zero e assim tomar a estimativa deste parâmetro como uma medida da alteração na média dos tratamentos que é devida à seleção praticada ao longo do período (progresso genético).

A fonte de variação de genótipos dentro de eras, por sua vez, mede a média da variação entre genótipos para cada era. Consequentemente, a correlação intraclasse desse componente quadrático é um análogo da herdabilidade média (h^2) nas "eras" avaliadas. Tais estimativas (R e h^2), frente à *breeder's formula* ($R = h^2\Delta$) podem ser manipuladas algebricamente resultando na estimativa do diferencial de seleção que fora realizado no período ($\Delta_{\text{realizado}} = R/\rho$[4]).

[4] ρ, refere-se a correlação intraclasse do componente quadrático associado ao efeito de genótipos dentro de eras.

Pode-se adaptar a teoria dos índices de seleção e utilizar o diferencial de seleção "realizado" com o intuito de reunir a informação dos vários caracteres avaliados em um índice que classifica os genótipos sob análise. Esse índice resultante, é portanto uma possibilidade promissora na interpretação do progresso genético quando mais de um caráter sofreu seleção. Assim, o desenvolvimento teórico (seção 2.2.2) mostra-se como uma elegante alternativa para se obter uma estimativa do progresso genético simultâneo.

Assim sendo, será aplicado o método proposto na seção 2.2.2 com base na estimativas do diferenciais de seleção "realizados" obtidos dos dados referentes ao estudo do progresso genético no melhoramento de tabaco, a fim de aferir a estimativa do progresso genético simultâneo obtido pelo programa durante o período sob análise.

2.3 Resultados

2.3.1 Progressos Genéticos Univariados

Algumas estatísticas referentes às análises realizadas nos diferentes caracteres são pormenorizadas na Tabela 3. Uma maneira de avaliar a qualidade de um dado ajuste é por meio do coeficiente de variação, sendo sua medida um critério de descarte de experimentos, vide Pimentel Gomes (2009). Nos caracteres avaliados, tal coeficiente nunca foi maior que 15%. O maior CV foi o do peso de folhas curadas (12%), enquanto que *nfe* possui o menor coeficiente (3%). Outra maneira de averiguar a qualidade do ajuste de um modelo com aplicação em genética e melhoramento é a acurácia seletiva (RESENDE; DUARTE, 2007). Nesse aspecto, o número de folhas expandidas foi o caráter com modelo preditivo mais acurado (99%).

Tabela 3 – Resultado do testes da homegeneidade de variâncias (K), normalidade dos resíduos (W) e aditividade dos efeitos (F); coeficiente de variação (CV) e acurácia seletiva (Ac) para os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*), número de folhas expandidas (*nfe*), circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*), Brasil 2011/2012

caráter		Bartlett[a]		Shapiro–Wilk		Tukey[b]			
	$\bar{x}$	K	P[c]	W	P	F	P	CV (%)	Ac (%)
pfc ($kg\,ha^{-1}$)	2909,28	6,71	0,08	0,99	0,48	0,34	0,56	12	97
alt (m)	1,12	4,33	0,23	0,99	0,24	7,10	0,01	6	99
nfe	29,29	9,75	0,02	0,99	0,34	13,61	< 0,01	3	99
ccc (cm)	8,48	7,79	0,05	0,97	0,01	4,12	0,05	7	99
l3f (cm)	28,07	17,76	< 0,01	0,97	< 0,01	4,65	0,04	8	98
c3f (cm)	60,28	23,19	< 0,01	0,90	< 0,01	1,20	0,28	8	95

[a] Graus de liberdade para o teste é 3.

[b] Graus de liberdade para o teste é 1 / 23.

[c] Significância do teste.

Observa-se que para os caracteres *pfc* e *alt*, existem fortes evidências para aceitar que os ambientes possuem efeitos residuais homogeneamente distribuídos. Se todavia, considerar-se um limiar de significância de 1%, para *l3f* e *c3f* seria rejeitada a hipótese nula. A significância do teste sobre a hipótese de normalidade dos efeitos residuais (*Shapiro–Wilk*) mostra que, tanto para *pfc*, *alt* e *nfe* não existem evidências suficientes para descartar tal premissa. O teste da aditividade dos efeitos de Locais e Tratamentos, apenas para o caráter *nfe*, apresentou significância abaixo do limiar de 1%.

O resumo das análises de variância dos caracteres *pfc*, *alt*, *nfe*, *ccc*, *l3f* e *c3f* estão dispostos nas Tabelas 4 e 5. Para todos caracteres, os testes sobre os efeitos do modelo inicial da análise (Locais, Tratamentos e a interação entre ambos) mostraram valores de significância menores que o limiar de 1%. A exceção ocorreu nos efeito da interação entre locais e tratamentos genéticos para o caráter *c3f*; o que é um indicativo para considerar que, em média, o desempenho dos tratamentos foi relativamente coincidente nos quatro locais avaliados (Tabela 5). Em todos os caracteres que foram tomadas mais de uma amostra por parcela, observou-se que a estimativa da variância entre parcelas foi menor que a variância devida ao erro dentro (dados não mostrados).

Tabela 4 – Análise de variância e desdobramentos complementares ao estudo do progresso genético para os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*) e número de folhas expandidas (*nfe*), Brasil 2011/2012

F.V.	G.L.	*pfc* ($kg\,ha^{-1}$)		*alt* (m)		*nfe*	
		Q.M.	P	Q.M.	P	Q.M.	P
Locais	3	61380694	< 0,01	7,845	< 0,01	676,40	< 0,01
Blocos / Locais	12	538745	< 0,01	0,097	< 0,01	14,40	< 0,01
Tratamentos	8	2006382	< 0,01	1,679	< 0,01	2222,14	< 0,01
Eras	3	3969096	< 0,01	4,459	< 0,01	5734,30	< 0,01
Progresso Linear	1	11845424	< 0,01	13,374	< 0,01	15829,20	< 0,01
Desvios	2	30932	0,79	0,001	0,93	686,85	< 0,01
Tratamentos / Eras	5	828753	0,04	0,011	0,91	114,80	< 0,01
Locais × Tratamentos	24	636283	< 0,01	0,193	< 0,01	81,25	< 0,01
Locais × Eras	9	1265239	< 0,01	0,447	< 0,01	190,80	< 0,01
Locais × Progresso	3	2719532	< 0,01	0,988	< 0,01	392,20	< 0,01
Locais × Desvios	6	538093	< 0,01	0,177	< 0,01	90,05	< 0,01
Locais × Tratamentos / Eras	15	258909	0,02	0,040	0,02	15,50	< 0,01
Residuo	96*	129291		0,019	< 0,01	5,20	< 0,01
Erro Dentro	576			0,005		2,90	

* para o caráter *pfc*, devido ao ajuste das médias pelo estande, perde-se um G.L..

No segundo momento da análise, desdobraram-se os efeitos de Tratamentos, bem como o efeito de sua interação com Locais em diferenças entre Eras e entre Tratamentos dentro de Eras. A fonte de variação Tratamentos dentro de Eras afere a variação média entre os tratamentos

de cada era. Levados em consideração os aspectos pontuados na seção 2.2.3, o componente quadrático dessa fonte de variação tem interpretação útil na estimativa do progresso simultâneo. Para os caracteres *pfc*, *nfe* e *l3f* o teste resultou em valores que fomentam a decisão de considerar significativa essa variação (Tabelas 4 e 5).

Em todos os caracteres avaliados, verificaram-se diferenças significativas entre as diferentes Eras do programa de melhoramento. Isso pode ser tomado como um indicativo de que existem argumentos para inferir que as médias dos tratamentos guardam relação com a Era da qual aqueles tratamentos são originários. Quando se parametriza as diferenças devidas a Eras com um efeito linearmente associado, todos os caracteres mostraram diferenças no desempenho que podem estar associadas à seleção praticada ao longo das eras do programa de melhoramento. Os desvios desse ajuste linear são uma medida da adequação do modelo. Para os caracteres *pfc*, *alt* e *c3f*, tais desvios não podem ser tomados como significativos (Tabelas 4 e 5).

Tabela 5 – Análise de variância e desdobramentos complementares ao estudo do progresso genético para os caracteres circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*), Brasil 2011/2012

F.V.	G.L.	*ccc* (*cm*)		*l3f* (*cm*)		*c3f* (*cm*)	
		Q.M.	P	Q.M.	P	Q.M.	P
Locais	3	688,8	< 0,01	3569,9	< 0,01	8718,8	< 0,01
Blocos / Locais	12	22,6	< 0,01	61,6	< 0,01	174,7	0,15
Tratamentos	8	257,7	< 0,01	617,3	< 0,01	1365,8	< 0,01
Eras	3	682,0	< 0,01	1433,9	< 0,01	3569,5	< 0,01
Progresso Linear	1	1887,3	< 0,01	3204,0	< 0,01	10481,9	< 0,01
Desvios	2	79,4	< 0,01	548,9	< 0,01	113,3	0,39
Tratamentos / Eras	5	3,1	0,67	127,4	0,03	43,6	0,96
Locais × Tratamentos	24	17,0	< 0,01	125,8	< 0,01	206,1	0,03
Locais × Eras	9	37,5	< 0,01	231,6	< 0,01	174,9	0,17
Locais × Progresso Linear	3	71,2	< 0,01	446,1	< 0,01	413,4	0,02
Locais × Desvios	6	20,6	< 0,01	124,3	< 0,01	55,7	0,83
Locais × Tratamentos / Eras	15	4,8	< 0,01	62,3	< 0,01	224,8	0,03
Residuo	96	1,8	< 0,01	22,6	< 0,01	118,9	< 0,01
Erro Dentro	576	1,0		6,4		24,0	

A significância dos efeitos de interação de Locais com Tratamentos, Eras, Progresso, Desvios e entre Tratamentos dentro de Eras podem ser utilizadas para verificar se os resultados marginais, na média dos quatro locais, são bons preditores dos resultados obtidos nos quatro locais isoladamente. Apenas para o caráter *c3f* todos esses efeitos podem ser tomados como não significativos ao nível de 1% de probabilidade. Desse modo, para os demais caracteres, os efeitos relacionados às alterações na performance média dos genótipos não foram os mesmos

para os quatro locais de experimentação, havendo necessidade de considerar isso na interpretação dos resultados (Tabelas 4 e 5).

Na Figura 3 estão representados as médias dos tratamentos, seus intervalos de predição de (90%) além da predição pontual e intervalar, com 90% de probabilidade, da regressão linear. Tendo como variável independente as Eras e como variável dependente as médias dos diferentes caracteres. Uma representação do progressos genéticos ou mudança na performance dos tratamentos genéticos que pode ser atribuída à seleção praticada. Afim de facilitar a interpretação e comparação dos resultados, optou-se por representar as estimativas do progresso e efeitos de tratamentos em relação ao intercepto do modelo (β_0).

O caráter que apresentou o maior progresso relativo foi *ccc*, indicando que a circunferência no caule da planta aumentou 1,52cm por era. Este incremento representa um acréscimo de 25%. O acréscimo no comprimento da terceira folha foi de 3,58cm por era, ou 7% em relação aos 54,71cm estimados no início do programa de melhoramento. Ao se considerar que o período entre eras é de cerca de 10 anos, pode-se estimar o progresso anual na produtividade do peso de folhas curadas como de aproximadamente 27,53$kg\,ha^{-1}$, representando acréscimos da ordem de 1% ao ano (Figura 3).

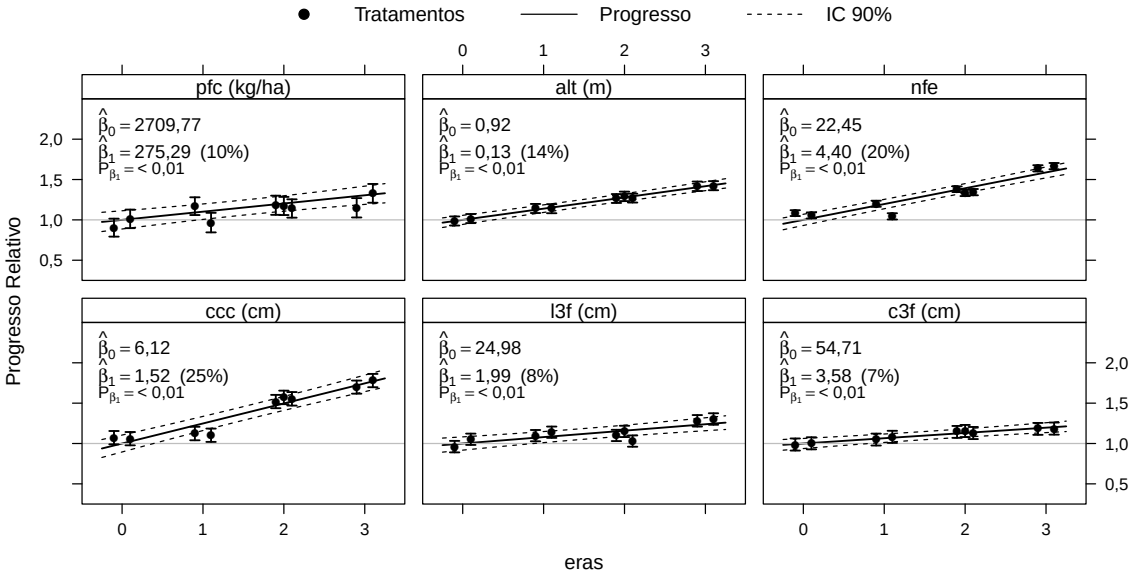

Figura 3 – Progressos genéticos na média dos quatro locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*), número de folhas expandidas (*nfe*), circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*), Brasil 2011/2012

A inferência sobre a significância do progresso genético se dá através do teste t sobre a estimativa do coeficiente angular ($\hat{\beta}_1$). O resultado desse teste é análogo ao teste F realizado sobre a fonte de variação Progresso apresentado nas Tabelas 4 e 5. É possível verificar pelo teste que, em todos os caracteres avaliados, houve progressos genéticos significativos com valores de probabilidade menores que 1% (Figura 3).

Assim como as estimativas das médias dos tratamentos, as predições do progresso genético foram feitas a um nível de probabilidade de 90%. Pelos intervalos de predição, pode-se observar que mesmo apresentando progressos equivalentes, 10% para *pfc* e 14% para *alt*, tais intervalos evidenciam diferentes precisões associadas a cada caráter. A mesma observação pode ser feita em relação aos caracteres *l3f* e *c3f* (Figura 3). A respeito das médias dos tratamentos, pode-se observar a sobreposição dos intervalos das médias dentro de cada era para todos os caracteres à exceção de *nfe*. Essa observação se coaduna com o teste sobre as diferenças entre Tratamentos dentro de Eras nas Tabelas 4 e 5.

Dada a significância do teste da interação Locais $\times$ Progresso, obteve-se a estimativa dos progressos para cada um dos locais estudados (Figuras 4 e 5). Observam-se para todos os caracteres, diferenças relativas ao progresso em cada ambiente, bem como, diferenças relativas à estimativa do desempenho das linhagens iniciais (intercepto, ou $\hat{\beta}_0$). Entrementes, em nenhum dos caracteres avaliados ocorreram discordâncias quanto às tendências dos progressos genéticos comparando-se os ambientes, obtendo-se progressos ora em uma direção, ora em outra, conforme o local.

Tomando o caráter *pfc* como exemplo, observa-se que a estimativa do progresso variou entre locais. Para os locais *PN* e *M1* não houve progressos significativos. Em *RN* o progresso representou incremento em produtividade de folhas curadas da ordem de 16% por era, ao passo que em *M2* o incremento em produtividade de folhas curadas foi 10% (Figura 4). Para o caráter *l3f* observa-se comportamento semelhante ao comentado (Figura 5). O progresso no ambiente *RN* mostrou-se não significativamente diferente de zero, todavia nos demais ambientes o progresso variou de 8% no ambiente *PN* a 15% em *M2*.

Assim como no caráter *pfc*, *l3f* também apresentou efeito significativo da interação Locais $\times$ Progresso e essa interação no entanto, pode ser interpretada como do tipo simples, variando apenas a magnitude do ganho e não a tendência do mesmo. Outro exemplo, mais ilustrativo desse mesmo fenômeno é o que ocorreu com os caracteres *nfe* e *ccc* em que, apesar de significativo o efeito da interação entre Locais e Progressos, os progressos nos quatro locais foram significativos e positivos. As estimativas do progresso variaram entre 11% a 26% para *nfe* e entre 14% e 35% para *ccc* (Figuras 4 e 5).

Quando se comparam os progressos nos caracteres *alt* e *ccc* entre os ambientes *RN* e *M1*, vê-se, em todas as situações, progressos significativos ($P \leq 0{,}01$). Todavia, ficam evidentes as diferenças quanto ao desempenho médio e principalmente ao intercepto ($\hat{\beta}_0$). O caráter *alt* apresentou progressos da ordem de 13% no ambiente *RN*, associado ao intercepto de 1,14m, enquanto que no ambiente *M1*, o progresso representou incrementos de 8% na altura das plantas, onde o desempenho das plantas iniciais foi de apenas 0,77m. No caráter *ccc* essa diferença é mais acentuada pois, no ambiente *M1* o intercepto estimado foi de 4,00cm, já no ambiente *RN* esse valor foi 1,74 vezes maior (6,97cm), mesmo ambos os ambientes apresentando progressos de magnitude equivalente: 27% e 26%, para *RN* e *M1* respectivamente (Figuras 4 e 5).

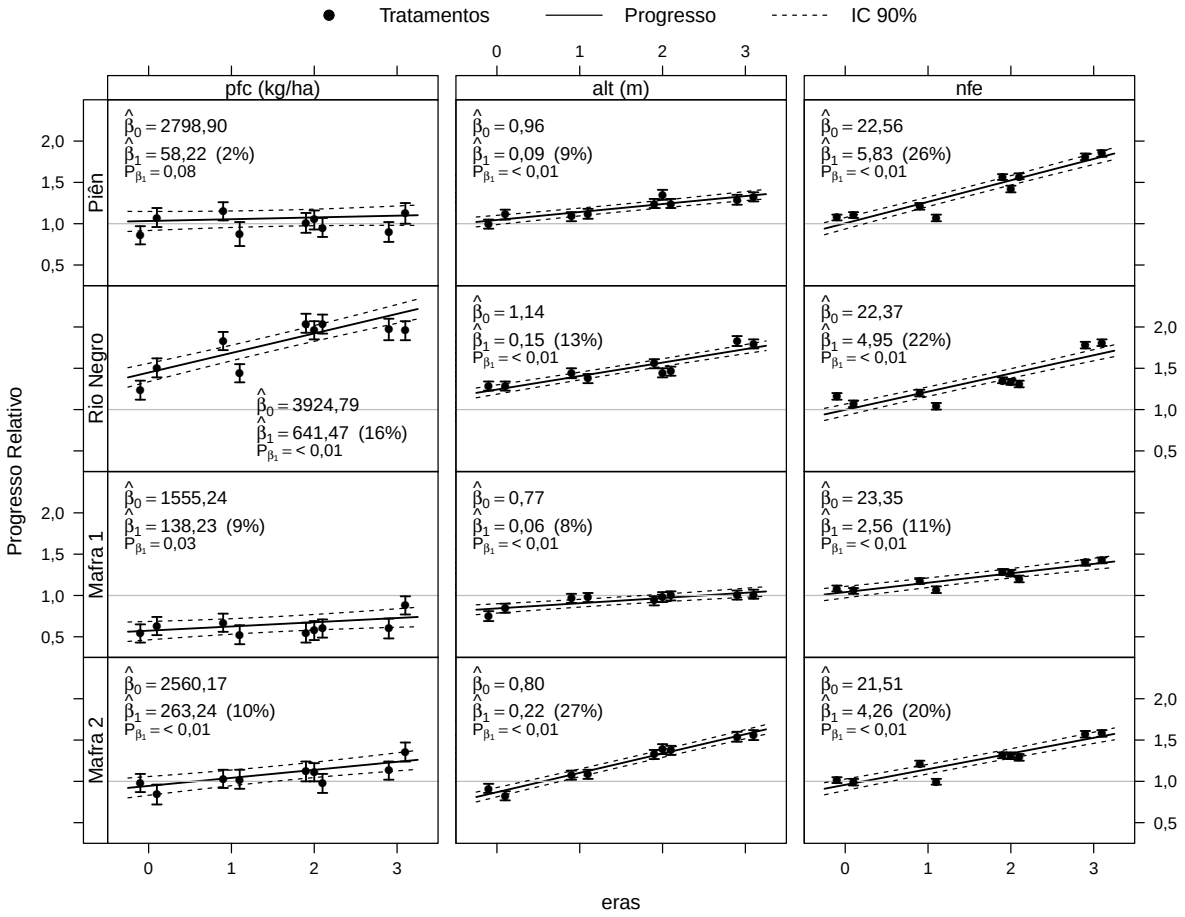

Figura 4 – Progressos genéticos por locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*) e número de folhas expandidas (*nfe*), Brasil 2011/2012

2.3.2 Progresso Genético Simultâneo

Como condição necessária à construção do índice, conforme proposto na seção 2.2.2, página 41, faz-se necessária a estimativa da matriz de variâncias e covariâncias fenotípicas. Conforme Steel, Torrie e Dickey (1996, capítulo 11), ao representar covariâncias como correlações, não há perda de informação. Deste modo, optou-se por apresentar esses relacionamentos entre caracteres por meio de coeficientes de correlação fenotípica ao nível de médias (Tabela 6).

Correlações, assim como as covariâncias, aferem o grau pelo qual dois caracteres se associam, sendo utilizadas com propósitos descritivos, por serem adimensionais e apresentarem um meio formal para testar sua significância. Nesse sentido, na diagonal inferior direita da Tabela 6 são apresentadas as significâncias do teste sobre das correlações estimadas. Se for considerado o limiar de probabilidade de 1%, apenas o relacionamento entre *pfc* e *l3f* seria considerado não significativo. Indicando que não há relação entre a produtividade de folhas curadas e a largura e comprimento da terceira folha.

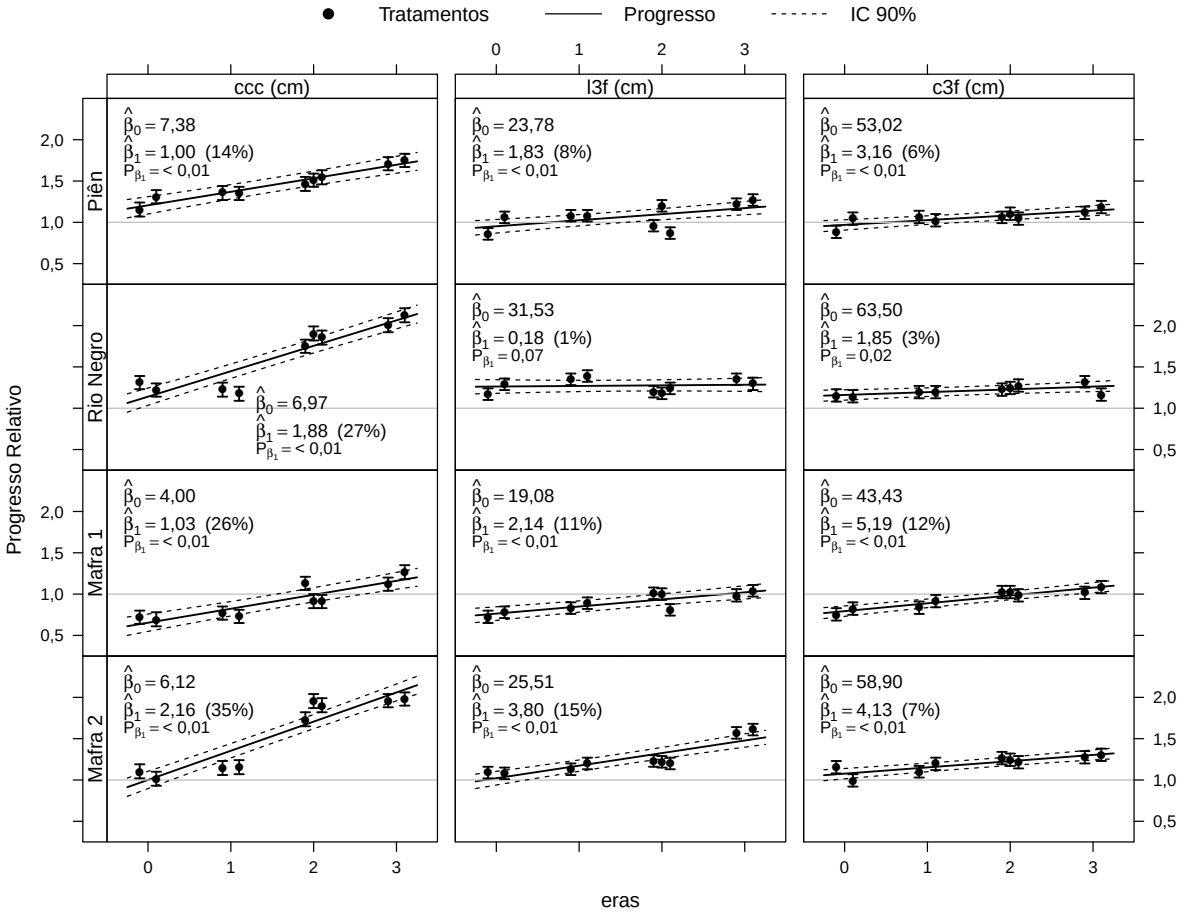

Figura 5 – Progressos genéticos por locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os caracteres circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*), Brasil 2011/2012

As demais estimativas de correlação variaram amplamente entre os pares de caracteres. Vale destacar que todas as correlações estimadas foram positivas, ou seja, o relacionamento entre os caracteres é diretamente proporcional; o aumento/diminuição em um caráter provoca aumento/diminuição em outro. A maior correlação observada foi entre *alt* e *ccc* (0,87), enquanto que a menor foi entre *c3f* e *pfc* (0,50). As correlações entre alguns caracteres possuem uma interpretação prática direta. A correlação entre *alt* e *nfe* foi de 0,83. Assim, como era de se esperar, plantas mais altas, possuem maior número de folhas. Outro exemplo é o que foi observado entre *alt* e *ccc*, evidenciando que plantas mais altas apresentam caules mais espessos e resistentes (Tabela 6).

Para a estimativa do progresso simultâneo, além das estimativas das variâncias e covariâncias fenotípicas, têm-se a necessidade de obter as estimativas do diferencial de seleção "realizado". Tais diferenciais serão então utilizados como restrição no índice. Essas estimativas, além de outras informações são apresentadas na Tabela 7.

Tabela 6 – Correlações fenotípicas na média dos quatro locais, entre os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*), número de folhas expandidas (*nfe*), circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*)

caráter	*pfc* ($kg\,ha^{-1}$)	*alt* (m)	*nfe* (cm)	*ccc* (cm)	*l3f* (cm)	*c3f* (cm)
pfc		0,586[a]	0,523	0,635	0,321	0,495
alt	< 0,01[b]		0,826	0,868	0,638	0,695
nfe	< 0,01	< 0,01		0,828	0,533	0,625
ccc	< 0,01	< 0,01	< 0,01		0,506	0,712
l3f	0,08	< 0,01	< 0,01	< 0,01		0,704
c3f	0,01	< 0,01	< 0,01	< 0,01	< 0,01	

[a] Diagonal superior direita refere-se aos coeficientes.
[b] Diagonal inferior esquerda refere-se a significância do teste.

Observam-se, entretanto, nas Tabelas 4 e 5 que em apenas três caracteres (*pfc*, *nfe* e *l3f*) a fonte de variação Tratamentos dentro de Eras se mostrou significativa. Assim, apenas para tais caracteres, é possivel estimar a componente quadrática associada a tal fonte de variação e consequentemente, configurar a restrição e obter o diferencial de seleção "realizado"[5].

As estimativas do componente quadrático $\phi^2_{g/p}$ para os caracteres *pfc*, *nfe* e *l3f* são apresentados na Tabela 7, onde se destacam também as elevadas correlações intraclasse (ρ) associados aos mesmos, variando entre 0,82 e 0,95. Os diferenciais de seleção "realizados", indicam, que os progressos obtidos para *pfc*, por exemplo, são fruto de seleção onde, em média, o limiar de seleção distou 326,17$kg\,ha^{-1}$ da média da população de seleção.

É preciso notar contudo, que pela particularidade dos dados e pela adaptação aos propósitos do emprego do índice, as estimativas referem-se à variância remanescente, após a seleção. Todavia, como será melhor discutido na seção 2.4, página 58, a variância remanescente guarda uma relação com a variância anterior à seleção e deste modo não configura um impedimento a demonstração que aqui se propõe realizar.

Utilizados os diferenciais de seleção para resolver o sistema do índice (eq. 15), após sete iterações o sistema converge para os coeficientes ($\hat{\beta}$) apresentados. Assim, o índice se constitui como uma função linear que pondera os caracteres *pfc*, *alt*, *nfe*, *ccc*, *l3f* e *c3f* com os respectivos coeficientes:

$$\hat{\boldsymbol{\beta}}' = [-0{,}01; 3{,}07; -1{,}04; 3{,}07; -1{,}96; 3{,}07]\ .$$

Os valores iniciais foram iguais à unidade em todos caracteres. Testes (não mostrados) demonstram uma proporcionalidade, pois dobrando-se os valores iniciais, dobram-se os coeficientes após a convergência. Observa-se também que, para os caracteres irrestritos (*alt*, *ccc* e *c3f*) os coeficientes são idênticos, que é outro argumento em favor da proporcionalidade do sistema. Para os caracteres restritos (*pfc*, *nfe* e *l3f*), os coeficientes foram todos negativos e menores em módulo que os demais.

[5]Pelo apresentado na seção 2.2.2, $r \leq n$, deste modo, é possível a construção do índice impondo restrições a um subconjunto das variáveis sem perda de generalidades.

Vale destacar o coeficiente associado ao caráter *pfc*, que foi de apenas -0,01. Isso é um indicativo de que, apesar da importância do caráter, no índice, incluídos os demais caracteres, o próprio caráter perde importância. Em contrapartida, quando se olha para o diferencial de seleção (Δ_{estimado}) que seria obtido e o ganho ($R|\boldsymbol{I}$) para o caráter, praticando-se a seleção pelo índice, reforça-se a idéia de que o os ganhos obtidos no caráter estão fortemente associados aos ganhos dos demais caracteres positivamente correlacionados a ele (Tabelas 6 e 7).

Algebricamente é possível estimar: os diferenciais de seleção para todos os caracteres e as intensidades de seleção que tais diferenciais representam apenas nos caracteres restritos, bem como o progresso genético (Tabela 7). Pode-se comprovar que o sistema converge para uma solução que satisfaz as restrições. Tais restrições, por sua vez, representam uma intensidade de seleção próxima dos 10% superiores para o caráter *pfc* e de cerca de 5% em relação a *nfe*.

Tabela 7 – Resumo das estimativas relativas ao índice reunindo os caracteres peso de folhas curadas (*pfc*), altura de plantas (*alt*), número de folhas expandidas (*nfe*), circunferência do caule no colo da plantas (*ccc*), e respectivamente, largura e comprimento da terceira folha (*l3f* e *c3f*)

	pfc ($kg\,ha^{-1}$)	*alt* (m)	*nfe*	*ccc* (cm)	*l3f* (cm)	*c3f* (cm)
$\bar{x}$	2909,28	1,12	29,29	8,48	28,07	60,28
$\phi^2_{g/p}$	43716,38	-*	6,85	-	6,55	-
ρ	0,84	-	0,95	-	0,82	-
$\Delta_{\text{realizado}}$	326,17	-	4,61	-	2,42	-
$\hat{\boldsymbol{\beta}}$	-0,01	3,07	-1,04	3,07	-1,96	3,07
Δ_{estimado}	326,17	0,19	4,61	3,03	2,42	10,90
$R\|\boldsymbol{I}$	275,29	-	4,40	-	1,99	-
i	1,56	-	1,76	-	0,94	-
P (%)	~ 10	-	~ 5	-		-

* F.V. Tratamentos / Eras não significativo (Tabelas 4 e 5).

Os caracteres *iqg* e *$ha* podem ser interpretados como índices empíricos "caracteres índice", enquanto que o índice construído, com base nos dados da Tabela 7, constitui-se em um super caráter que combina as informações dos demais caracteres. O resumo das análises de variâncias para tais variáveis são apresentado na Tabela 8.

A parte os comentários relativos à análise preliminar de variância, é necessário comentar os aspectos relativos às estimativas aplicadas aos propósitos do presente trabalho. Tanto para o índice construído como para *iqg* e *$ha* os efeitos de Eras foram significativos. Quando esse efeito é desdobrado em sua componente linear e nos desvios dessa regressão, verifica-se para *iqg* e para o índice que os progressos genéticos foram significativos ao limiar de 1% de probabilidade, enquanto que os desvios do ajuste linear podem ser considerados inexpressivos apenas para *iqg*. Quando se observam os efeitos de progresso que interagem com locais, verifica-se em todos os caracteres índice, que tais efeitos não podem ser desconsiderados; por

serem significativos, indicam que os progressos obtidos nos quatro locais diferem entre si (Tabela 8).

Pela tabela 8, observa-se que os efeitos associados à fonte de variação Tratamentos dentro de Eras foram não significativos para todos os caracteres índice. Ao passo que apenas a interação entre Locais e os efeitos de Tratamentos dentro Eras no caráter *$ha* foram significativos. Assim, na avaliação destes caracteres, as médias dos genótipos contemporâneos não diferiram significativamente, sugerindo uma homogeneidade no desempenho dos tratamentos provenientes de uma mesma fase do programa de melhoramento.

Tabela 8 – Análise de variância e desdobramentos complementares ao estudo do progresso genético para o índice de qualidade geral (*iqg*), para a remuneração paga pelo tabaco em *US$* por hectare (*$ha*) e para o índice construído (índice), Brasil 2011/2012

F.V.	G.L.	*iqg*		*$ha*		índice	
		Q.M.	P	Q.M.	P	Q.M.	P
Locais	3	3472,70	< 0,01	782362196	< 0,01	4931,4	< 0,01
Blocos / Locais	12	33,00	0,12	5013190	< 0,01	170,7	0,48
Tratamentos	8	223,90	< 0,01	6952149	< 0,01	1381,2	< 0,01
Eras	3	536,30	< 0,01	6621519	< 0,01	3220,2	< 0,01
Progresso Linear	1	1519,10	< 0,01	6351521	0,03	5666,9	< 0,01
Desvios	2	44,90	0,13	6756518	0,01	1996,9	< 0,01
Tratamentos / Eras	5	36,40	0,08	7150528	0,09	277,8	0,36
Locais × Tratamentos	24	58,50	< 0,01	4788744	< 0,01	304,7	0,03
Locais × Eras	9	131,20	< 0,01	7656243	< 0,01	421,9	0,02
Locais × Progresso	3	288,40	< 0,01	10059777	< 0,01	757,1	< 0,01
Locais × Desvios	6	52,55	0,03	6454477	< 0,01	254,3	0,20
Locais × Tratamentos / Eras	15	14,90	0,78	3068245	0,01	234,4	0,19
Resíduo	96	21,30		1269935		174,6	

Ao observar a representação gráfica dos progressos genéticos para os caracteres índice (*iqg*, *$ha* e índice com restrições), algumas observações fazem-se pertinentes. Na média dos quatro locais a variável que reúne aspectos da qualidade das folhas de tabaco apresentou um decréscimo no período estudado. As linhagens iniciais apresentaram um desempenho médio de 42,69*iqg*[6], enquanto que os decréscimos por era representam uma redução de 10%, ou 4,21*iqg*'s (Figura 6).

De acordo com a Figura 6, os ganhos estimados para *$ha* não se mostraram favoráveis. Tal caráter mensura a remuneração média do produtor com a venda regulamentada das folhas curadas, visto que se constataram progressos relativos à produtividade de folhas curadas (ver Tabela 4 e Figura 3). Concomitante com uma redução do *iqg*, pode-se argumentar que, sendo

[6]Por tratar-se de um "índice" empírico, o caráter não apresenta uma unidade, sendo portanto representado por *iqg*'s.

$ha uma função da produtividade e da qualidade, o decréscimo observado em *iqg* compensa os ganhos em *pfc*.

O objetivo maior do emprego do índice é reunir as estimativas dos vários caracteres mensurando o progresso simultâneo. Isto posto, ao observar o progresso avaliado pelo índice (Figura 6), pode-se argumentar que os ganhos reunidos dos vários caracteres (*pfc*, *alt*, *nfe*, *ccc*, *l3f* e *c3f*) representam acréscimos da ordem de 6% por era. À semelhança do que foi enfatizado em relação ao *iqg*, essa variável configura-se como um super caráter, e desse modo, a unidade do índice não posssui uma interpretação prática. O progresso relativo (6%), no entanto, reflete o ganho em todos caracteres que participam do índice, considerando o relacionamento entre os caracteres e portanto as respostas correlacionadas à seleção praticada.

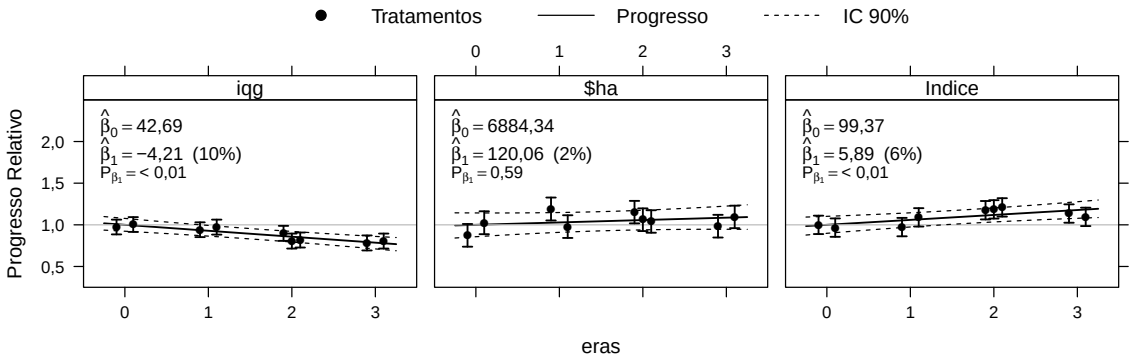

Figura 6 – Progressos genéticos na média dos quatro locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os "caracteres índice" índice de qualidade geral (*iqg*), remuneração ao produtor em *US$* por hectare (*$ha*) e para o índice construído (índice), Brasil 2011/2012

Deslindar os progressos genéticos nos quatro locais diante da significância da interação Locais $\times$ Progressos mostra-se oportuno e esclarecedor. Em um local (*PN*), os ganhos em qualidade (*iqg*) mostraram-se não significativos, como decorrência de um desempenho médio bem aquém do observado para o caráter na média nos diferentes locais. Nas localidades de *RN* e *M2*, os descrésimos em qualidade foram equivalentes e de 10% em ambos. Frisa-se contudo, que o desempenho médio em *RN* mostrou-se substancialmente maior ao observado em *M2* (Figura 7).

A respeito de *$ha*, na média dos quatro locais, os progressos não se mostraram significativos a 1% de probabilidade (Figura 6), todavia, para *RN* observaram-se progressos que representam um incremento no retorno financeiro de 905,59 *US$* por era. Nos demais locais as estimativas de $\hat{\beta}_1$ não foram significativamente diferentes de zero pelo teste t.

Contrapor os resultados no local *RN* dos caracteres *pfc*, *iqg* e *$ha* reitera o relacionamento entre ambos. Observaram-se ganhos de 641,47$kg\,ha^{-1}$ para *pfc*, o maior progresso entre os quatro locais, devido ao melhor desempenho médio das linhagens. Foi também em *RN*, que se observou o maior decréscimo na qualidade (*iqg*), que foi todavia, devido ao maior desempenho médio. Ao se comparar tal condição com o progresso na remuneração do produtor (*$ha*), *RN* foi

o único ambiente com progressos significativos, evidenciando que, mesmo frente à redução da qualidade, o progresso de *$ha* foi significativamente positivo como resultado do alto progresso em produtividade e ao alto desempenho expresso pelos genótipos nesse ambiente (Figuras 4 e 7).

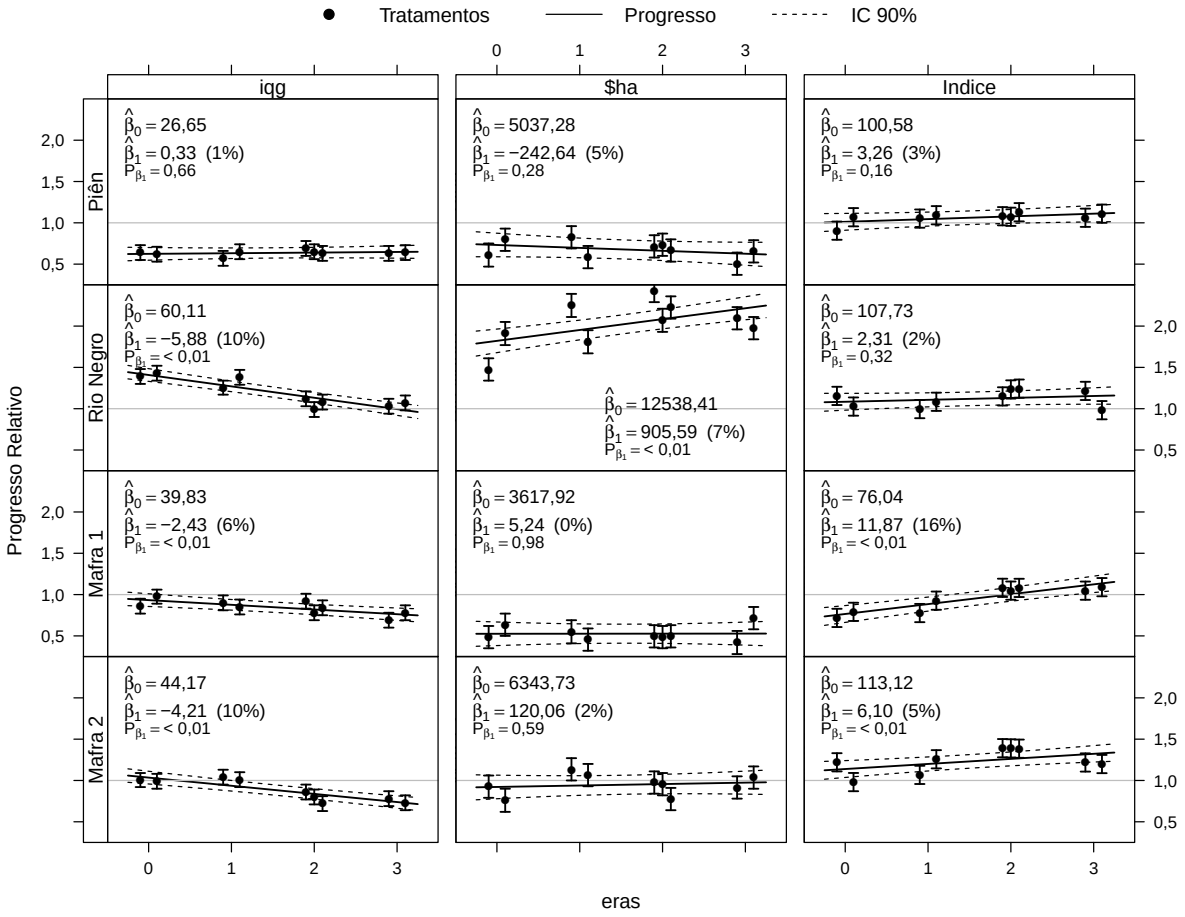

Figura 7 – Progressos genéticos por locais, médias e intervalos de predição (90%) dos tratamentos em relação ao intercepto do modelo para os "caracteres índice" índice de qualidade geral (*iqg*), remuneração ao produtor em *US$* por hectare (*$ha*) e para o índice construído (índice), Brasil 2011/2012

Para o índice, em dois locais (*PN* e *RN*) os ganhos não foram significativos. Em *M1* e *M2* os progressos representam, respectivamente, incrementos de 16% a 5% no conjunto de todos os caracteres utilizados no índice. O desempenho médio dos genótipos em *M1*, apesar do maior ganho, mostrou-se menor. Mesmo considerando significativos os efeitos da interação Locais $\times$ Tratamentos (Tabela 8), observa-se, pelos progressos nos quatro locais que, independentemente da significância, todas as estimativas de $\hat{\beta}_1$ são positivas, o que demonstra incrementos na medida do índice como função da seleção praticada e sobretudo que houve coincidência no desempenho das linhagens nos quatro ambientes (Figura 7).

2.4 Discussão

2.4.1 Progressos Genéticos Univariados

Ajustados os modelos de análise de variância, indubtavelmetne almeja-se estimar diferenças entre efeitos de interesse, aplicar testes de significância e obter estimativas da precisão ou acurácia dos estimadores (erros padrão e/ou intervalos fiduciais). Todavia, alguns pressupostos devem ser atendidos, como: aditividade dos efeitos, resíduos independentes, normal e identicamente distribuído (homogeneidade de variâncias). Quando entretanto, algum pressuposto não é atendido, algumas consequências devem ser esperadas (PIMENTEL GOMES, 2009).

Detectaram-se argumentos em favor da não aditividade dos efeitos de Locais e Tratamentos em *alt* e *nfe*. Tais desvios dos pressupostos, segundo Cochran (1947), acarretam principalmente em perda de informação, mas podem ser negligíveis já que boa aproximação dos efeitos são alcançadas pelo método de estimativa (quadrados mínimos). A respeito da normalidade e independência dos efeitos residuais, é consenso que não há sérias consequências sobre os testes de significância quando o interesse é na estimativa de efeitos (COCHRAN, 1947). Desse modo, podem ser relevados os argumentos observados nos caracteres *l3f* e *c3f* em detrimento a transformação ou descarte de tais variáveis.

Tendo como consequência deficiências na estimativa das médias, a heterogeneidade de variâncias, acarreta perda de sensibilidade nos testes de significância (COCHRAN, 1947). Enfatiza-se contudo, que os argumentos em favor da heterogeneidade de variâncias nos caracteres *l3f* e *c3f* não comprometem de sobremaneira as estimativas do progresso, visto que o teste de homogeneidade foi realizado como requisito da análise conjunta dos quatro locais e as médias e/ou comparações dos efeitos desses locais não são de interesse principal.

Pelo que foi visto, ao romper-se qualquer um dos pressupostos tem-se consequências que podem ser relativas às estimativas dos efeitos de interesse, sobre os testes de significância, ou na precisão dos estimadores. Conforme o pressuposto, algumas propriedades podem ser afetadas; todavia, pelo que foi discutido, mesmo diante da não adequação dos dados a alguns pressupostos, não há comprometimentos sérios aos objetivos maiores da análise.

Da mesma forma que em relação as inferências como dos teste sobre aos pressupostos da análise, ao longo de todo o texto, buscou-se fazer a inferência com menção ao valor de significância (P valor) sob o conceito original de Fisher, ou seja, a significância de um teste é a probabilidade, sob nulidade, de se obter a estatística daquele teste igual ou mais extrema que aquela observada. Assim, a significância é usada como um "guia" aproximado da evidência contra a hipótese nula (H_0). Ao considerar determinado teste significativo, interpreta-se que esse resultado é digno de atenção na forma de necessitar mais resultados experimentais, mas não como prova em si a favor de alguma hipótese alternativa (LEHMANN, 1993).

A inferência com base no conceito estrito de significância evita falsas interpretações, como associar os valores de significância com a probabilidade de a hipótese de nulidade ser verdadeira; que a significância informa sobre a magnitude da estatística referente ao teste; ou

ainda, considerar que o limiar de singificância faz qualquer referência à probabilidade de erro tipo I (rejeitar H_0 quando esta é verdadeira) (LEHMANN, 1993).

No melhoramento, a precisão experimental é comumente avaliada por meio do coeficiente de variação (CV) (PIMENTEL GOMES, 2009). Resende e Duarte (2007) enfatizam que essa estatística sofre influência da média e assim propõem o emprego da acurácia seletiva, que é um análogo da repetibilidade da performance dos tratamentos nas diferentes repetições. A acurácia corresponde à raiz quadrada do coeficiente de determinação genotípico e refere-se a correlação entre os valores preditos e observados.

Segundo Resende e Duarte (2007), valores de acurácia maiores que 70% indicam boa precisão. Enquanto que em conformidade com a legislação vigente, que regula a realização de ensaios de valor de cultivo e uso (*VCU*) do feijoeiro, por exemplo, usa o critério de descarte do CV, em que o limiar é de 25% (BRASIL, 2006). Para todos os caracteres, tanto avaliando o CV, quanto a acurácia, pode-se inferir que houve boa precisão nas estimativas das médias dos tratamentos e consequentemente boa precisão experimental.

Outro aspecto que se relaciona a precisão experimetal é a magnitude da variância dos desvios entre plantas dentro de parcelas e da variância entre parcelas que receberam o mesmo tratamento em diferentes repetições. Essas estimativas são importantes, já que a estimativa do ganho está relacionada à variância fenotípica, que por sua vez está relacionada às variâncias residuais, entre e dentro de parcelas. É possível verificar (dados não mostrados) que as estimativas da variância dentro de parcelas sempre foram maiores que a variância entre parcelas, ou seja, a maior variação ocorre dentro de parcelas. A variação entre plantas dentro de parcelas é toda de natureza ambiental, pois os tratamentos genéticos são linhagens endogâmicas. Já a variação entre parcelas, além de ambiental pode ser devida a erros de amostragem, quando as plantas dentro das parcelas não representam o genótipo.

A produtividade de folhas curadas (*pfc*) na média dos quatro locais aumentou de 2709,77$kg\,ha^{-1}$ para 3535,64$kg\,ha^{-1}$, isso representa um progresso genético de 275,29$kg\,ha^{-1}$ por era, ou 10%. Como o período entre eras compreende 10 anos, o acréscimo por ano é da ordem de 27,53$kg\,ha^{-1}$, ou 1%. Recente trabalho que almejou estimar o progresso genético do melhoramento de tabaco na Croácia entre 1973 e 2007 reporta aumento da produtividade de folhas curadas de 17,39$kg\,ha^{-1}$ por ano (SARCEVIC et al., 2013). Bowman et al. (1984) observou, ao longo de três décadas de melhoramento de tabaco, significativo incremento de 16$kg\,ha^{-1}$ por ano na produtividade de folhas curadas. Quando foram comparados o desempenho em produtividade entre os cultivares de tabaco estadunidense da década de 60 e 90 do século passado foi estimado um incremento de 5% (BOWMAN; SISSON, 2000).

Conforme Sarcevic et al. (2013), constatou-se de maneira geral, maior altura de plantas entre as cultivares modernas e significativa redução do número de folhas expandidas. Os resultados concordam com Sarcevic et al. (2013) em relação a altura de plantas, em que o incremento por era foi 13%, todavia, ao contrário do que foi comentado por Sarcevic et al.

(2013), observou-se maior número de folhas expandidas, com significativo aumento de 4,4 folhas por era.

As estimativas do progresso genético foram obtidas como o coeficiente angular da regressão linear. Para os caracteres *pfc*, *alt* e *c3f* os desvios do ajuste linear não foram significativamente diferentes de zero, indicando boa adequação do modelo proposto. Os resultados não mostram quaisquer sobressaltos entre as médias das diferentes eras ou evidência de estagnação dos progressos genéticos.

Deve-se frisar entretanto, que os diferentes métodos de estimativa possuem distintas propriedades em relação à capacidade de detectar inadequação do modelo linear. Soares, Ramalho e Sousa (1994) ao comparar o método seminal de Vencovsky et al. (1986) com a regressão de erros absolutos notaram equivalências entre as estimativas do progresso, mas frisam que a regressão de erros absolutos comporta-se melhor com a ocorrência de comportamentos aberrantes, que podem ser sobressaltos no progresso genético representados por desvios do ajuste linear.

Na ocorrência de indicativos da não linearidade do progresso genético, o interesse deve se voltar em elucidar as razões para tal intermitência. Breseghello et al. (2011) quando estimam o progresso genético do arroz, fazem a divisão do período sob a luz da história do programa de melhoramento, considerando pontos da descontinuidade que dividem o período em três fases distintas, que apresentaram diferentes comportamentos (progressos) e refletem diferentes estratégias empregadas. Abordagem semelhante foi adotada por Chiorato et al. (2010), em que se fez necessária uma divisão dos dados para contemplar as diferentes estratégias utilizadas em diferentes momentos do programa.

Breseghello et al. (2011) e Chiorato et al. (2010) atribuem às descontinuidades as diferentes estratégias de seleção. Peng et al. (2000), por sua vez, associam essas descontinuidades a mudanças nos objetivos do programa em que, num momento almejou-se a melhoria no índice de colheita e posteriormente na produção de biomassa. Ambas as estratégias tiveram reflexo positivo na produtividade, mas com diferentes magnitudes (PENG et al., 2000).

Se confrontados os resultados obtidos com outros estudos do progresso genético (BRESEGHELLO et al., 2011; CHIORATO et al., 2010; PENG et al., 2000), pode-se realçar que o adequado ajuste linear observado é um indicativo de que mesmo frente às alterações das estratégias ou dos objetivos da seleção, a eficiência do programa manteve-se constante ao longo do período. Essa argumentação é válida para todos os caracteres avaliados. Como também não foram observados indicativos de estagnação dos progressos em nenhum caráter, é possível vislumbrar a continuidade dos progressos em produtividade de tabaco, nos componentes da produção, bem como nos caracteres morfológicos, se forem seguidos os mesmos objetivos ou que as mesmas estratégias venham a ser utilizadas.

Um aspecto que merece atenção relaciona-se a interação genótipos por ambientes. Os resultados mostram que a estimativa do progresso na média dos quatro locais não é um bom preditor dos progressos por locais. Lange e Federizzi (2009) constataram comportamento

semelhante ao correlacionar as médias das cultivares em diferentes regiões de cultivo, o que permitiu inferir acerca do resultado da seleção em uma região alvo e o seu reflexo nas demais.

A significância do teste relativo à fonte de variação Locais $\times$ Tratamentos evidencia um comportamento diferenciado dos genótipos nos diferentes locais de avaliação e isso têm consequências com a estimativa do progresso. Nos quatro locais, as estimativas do progresso dos caracteres *pfc* e *l3f* variaram de não significativo a respectivamente 16% e 15% por era. Nos demais caracteres, *alt*, *nfe*, *ccc* e *c3f* todas as estimativas do progresso foram significativas e positivas. Destaca-se também as diferenças quanto as estimativas do intercepto ($\hat{\beta}_0$), o que reforça a idéia do comportamento diferencial dos cultivares nos diferentes ambientes de cultivo.

Underdahl et al. (2008) também reportam ampla variação na estimativa do progresso em diferentes locais e anos, como resposta da interação dos cultivares com os ambientes. Ao avaliar o progresso sob diferentes manejos de Nitrogênio, Tabien, Samonte e McClung (2008) observaram progressos significativos sob todos os manejos. Todavia, o progresso de maior magnitude esteve associado à maior dose do fertilizante, o que evidencia progresso em responsividade dos genótipos. Barbosa Neto et al. (2000) entretanto, variando a adubação nitrogenada e o manejo de defensivos, não observaram progressos genéticos diferenciados em resposta à melhoria do ambiente na aveia.

Ao invés de avaliar a estimativa do progresso genético em diferentes locais ou sob manejos contrastantes, Ustun, Allen e English (2001) estimaram o progresso da "estabilidade" fenotípica e mostram que além do progresso genético da produtividade, as cultivares modernas são mais estáveis e responsivas às melhorias do ambiente. Ao longo de três décadas de melhoramento de tabaco, além de incremento na produtividade de folhas, Sarcevic et al. (2013) observam que as variedades mais recentes são mais adaptadas às alterações ambientais.

Além dos progressos individuais e da sua interação com os ambientes de cultivo, Lange e Federizzi (2009) argumentam que, de sobremaneira, grandes esforços são gastos na seleção de outros caracteres de interesse, como a resistência/tolerância à fatores bióticos e abióticos, características de qualidade e morfológicos. Devido a isso, faz-se necessária uma avaliação do relacionamento dos caracteres sob análise.

No tabaco, os caracteres morfológicos, segundo Sarcevic et al. (2013), podem ser utilizados para predizer o desempenho da produtividade de folhas curadas. Em relação à composição química das folhas, mostrou-se que os teores de açúcares redutores e de alcalóides totais são negativamente correlacionados com a produtividade, o que pode dificultar o melhoramento (LI; BASS; CORNELIUS, 2006; BOWMAN; WEEKS; WILKINSON, 1991). Estimativas recentes reforçam a correlação negativa entre produtividade e composição química das folhas, em que o progresso em produtividade esteve associado a decréscimos no teor de nicotina (alcalóide), sem alterações significativas no conteúdo de açúcares redutores (SARCEVIC et al., 2013). Os resultados mostram que a produtividade é positivamente correlacionada com todos os caracteres morfológicos, exceto a largura das folhas. Assim, o

progresso genético de *pfc* esteve associado a plantas mais altas, com maior número de folhas expandidas e folhas maiores em comprimento. Sarcevic et al. (2013) por sua vez, constatam que o progresso da produtividade de folhas curadas na Croácia esteve atrelado a redução da altura de plantas e largura e comprimento das folhas e incremento no número de folhas expandidas.

Deve-se fazer mensão a outros complicadores que podem reduzir os progressos em produtividade de folhas curadas. Critérios mínimos de qualidade química das folhas muitas vezes são exigidos, Além disso, essas análises de alto custo, se realizadas em estágios finais de avaliação, dificultam o sucesso da seleção tanto para produtividade quanto para tais características químicas (BOWMAN; WEEKS; WILKINSON, 1991).

2.4.2 Progresso Genético Simultâneo

No desenvolvimento do índice, utilizou-se o método dos multiplicadores de *Lagrange*, que é um procedimento matemático amplamente utilizado para otimizar funções sujeitas a restrições nas suas variáveis (HARVILLE, 1975). O índice impõe restrições constantes, da mesma maneira do que Lin (2005b) apresenta, ou seja, restringindo os diferenciais de seleção para valores predeterminados. O procedimento difere do índice restrito de Tallis (1985), pois se dá de modo simultâneo, derivando-se a função de variância da diferença ($\sigma^2_{(\boldsymbol{I}-\boldsymbol{H})}$) em relação a $\boldsymbol{\beta}$ e $\boldsymbol{\lambda}$ e igualando as derivadas parciais a zero. Deste modo, o procedimento torna-se mais flexível, acomodando inclusive restrições de outra natureza, assim como Lin (2005b) desenvolve.

É necessário enfatizar algumas particularidades que caracterizam o processo de estimação do índice. A intensidade de seleção do índice (κ) é uma variável do sistema, portanto, ao variar essa intensidade variam também os coeficientes em $\boldsymbol{\beta}$. Todavia, essa variação se dá de modo não linear. Outra característica do sistema é que a matriz dos coeficientes contém o próprio vetor de soluções e deste modo, um processo iterativo se faz necessário na obtenção da solução, do vetor $\hat{\boldsymbol{\beta}}$.

Avaliando o sistema (eq. 15), é oportuno frisar que a matriz dos coeficientes é composta pela matriz de variâncias e covariâncias fenotípicas ($\boldsymbol{P}$), que tem posto completo n e por uma submatriz de $\boldsymbol{P}$, também com posto completo r. Assim, a matriz dos coeficientes tem posto completo $n + r$, e deste modo, iterações garantem a convergência para uma solução única. A rápida convergência obtida, após sete iterações, é devida à dimensão do sistema, ao posto da matriz dos coeficientes e ao número de caracteres e restrições. Problemas de convergência, são comentados por Lin (2005b), e podem estar associados a uma inadequada definição da variável κ ou a imposição de diferenciais de seleção extremos ou fenotipicamente impossíveis. Assim sendo, a imposição de restrições e a escolha de κ devem ser biometricamente razoáveis.

No processo iterativo, iniciou-se o sistema com todos os coeficientes de $\boldsymbol{\beta}$ iguais à unidade. Essa escolha pode ser interpretada como considerar, *a priori*, que todos os caracteres tem igual contribuição no cômputo do índice. Após a convergência, tais coeficientes guardam

relação com o peso/importância que cada caráter exerceu ao longo do processo de seleção. As constantes utilizadas como restrição (diferenciais de seleção realizados) foram estimados por meio das estimativas do progresso genético e da correlação intraclasse associada a variação entre tratamentos dentro de eras. Manipulando-se algebricamente a *breeder's formula* foi possível estimar os diferenciais de seleção realizados para os caracteres *pfc*, *nfe* e *l3f*. Portanto, mais do que fenotipicamente possível, tais diferencias de seleção foram efetivamente realizados.

As restrições foram satisfeitas, já que os diferenciais de seleção estimados são iguais as restrições impostas ($\Delta_{\text{realizado}} = \Delta_{\text{obsevado}}$). Isso dá suporte teórico ao índice desenvolvido (eq. 15). Ao interpretarmos os resultados de modo retrospectivo, pode-se inferir que se tivesse sido realizada a seleção baseada no índice os mesmo progressos seriam obtidos. Vislumbra-se que, dando continuidade a seleção fazendo uso do índice, o mesmo padrão de ganho nos caracteres avaliados deve ser observado.

Interpretar os coeficientes do índice pode ser revelador. Para o caráter *pfc* o coeficiente após a convergência foi de apenas -0,01. Assim, considerando o relacionamento e os progressos dos caracteres, infere-se que o incremento observado da produtividade de folhas curadas foi obtido por meio do progresso concomitante dos demais caracteres correlacionados. Assim, tendo-se praticado seleção nos demais caracteres, foram obtidos progressos na produtividade. Essa observação se coaduna com o que Sarcevic et al. (2013) observa, ou seja, que os caracteres morfológicos podem ser utilizados para predizer o desempenho em produtividade.

Adaptou-se o desenvolvimento do índice com o propósito de reunir as estimativas dos progressos genéticos de mais de um caráter considerando o relacionamento entre eles como forma de expressar o progresso genético simultâneo. Se todavia, o procedimento for empregado com vistas à seleção, arbitrando-se restrições nos diferencias de seleção, os coeficientes do índice representam um modo de praticar seleção, impondo limiares de seleção que consideram o relacionamento entre os caracteres. Isso implica em praticar seleção em níveis seletivos não independentes, o que se configura como uma alternativa ao que Magnussen (1990) desenvolve.

É oportuno frisar que os cultivares avaliados são os genótipos "selecionados" de cada era, de natureza fixa. Portanto, a variação entre genótipos dentro de eras é um análogo da variância genética remanescente, depois da seleção. Para os progressos individuais, essa particularidade não traz consequências visto que a estimativa desses progressos é função das diferenças de desempenho entre eras e em certa medida independe da variação dentro de eras.

Com o objetivo de estimar o progresso simultâneo o ideal seria avaliar as populações das quais os genótipos foram selecionados em cada era e assim estimar as variância genéticas antes da seleção. Buscou-se contudo, utilizar os dados para a estimação dos progressos por caráter e os mesmo dados foram empregados como exemplo de aplicação na estimativa do progresso simultâneo, adaptando a teoria dos índices de seleção. Essa incongruência em relação à estimação e interpretação da variação entre genótipos dentro de eras, tem grande relevância no método que se pretende apresentar.

A parte as particularidade dos dados, deve-se mencionar que sob certas circunstâncias é razoável utilizar a variação entre os genótipos selecionados de cada era, posterior a seleção, como uma estimativa da variância anterior à seleção. Para Falconer e Mackay (1996, capítulo 11), apenas em um cenário é possível prever alterações na variância de uma população. Ao se praticar seleção truncada, a variação fenotípica entre os indivíduos além do limiar é menor que a observada no conjunto de toda população.

Tal fenômeno, foi primeiramente derivado por Bulmer (1985 apud FALCONER; MACKAY, 1996, capítulo 9). Ignoram-se erros de amostragem, que podem ser devido ao tamanho efeitvo (N_e), assim a redução da variação fenotípica é função da intensidade de seleção praticada, enquanto que a variância aditiva, reduz em razão do desequilíbrio gamético, sendo portanto, função da herdabilidade e intensidade de seleção. A variância dos efeitos de dominância é ignorada já que sofre ínfima influência do desequilíbrio gamético. Bulmer (1985) argumenta que metade do desequilíbrio gamético é perdido por recombinação, mas mais desequilíbrio é regenerado por seleções sucessivas. Portanto, uma relação assintótica de recorrência pode ser estabelecida, onde há um balanço entre o desequilíbrio gerado pela seleção e o desfeito por recombinação. Atingido esse ponto de equilíbrio, a variação remanescente é um bom estimador da variância da população antes da seleção (FALCONER; MACKAY, 1996, capítulo 11).

Falconer e Mackay (1996) consideram também que a variância residual não deve sofrer alteração com a seleção. Todavia, deve-se considerar que em avaliações futuras dos genótipos selecionados, a variância entre esses genótipos é a variância remanescente e em razão da maior homogeneidade desses genótipos, podemos estender o raciocínio e prever uma redução da variância devida a causas não genéticas como uma função da contraparte do coeficiente de determinação genético ($e^2 = (1 - h^2)$). Essas considerações corroboram com o entendimento do Dr. Ralph E. Comstock da Universidade de Minnesota (comunicação pessoal com o Professor Roland Vencovsky). Segundo ele, dado algum esquema recorrente de seleção, ao longo dos sucessivos ciclos, pode-se pressupor que a herdabilidade mantenha-se constante.

Os melhoristas assumem que a redução na variância genética na seleção recorrente pode ser desprezível, ou de forma mais tênue, ou seja, que não ocorrem alterações significativas na variância genética ao longo do processo de melhoramento. Em consonância com o argumento de Comstock, para a estimativa do progresso simultâneo, considerou-se que é desprezível a redução da variância genética dentro das eras além de assumir que, mesmo ocorrendo uma redução na variância genética remanescente, ocorre também uma diminuição da variância de causas não genéticas em razão de os genótipos selecionados serem mais homogêneos.

Avaliou-se a título de ilustração, a correspondência entre as estimativas da herdabilidade antes (h_0^2) e após a seleção (h_s^2). Considerou-se uma população hipotética, em que a redução da variância genética, está sujeita ao *efeito Bulmer* (FALCONER; MACKAY, 1996, capítulo 11). Contemplou-se tanto a descrição original de Falconer e Mackay, que ignora alterações na variância residual, quanto o cenário no qual a redução da variância aditiva é acompanhada por

uma redução da variância de causa não genética. Esse último caso, sob certas certas circunstâncias, tem comportamento análogo ao que Comstock preconiza (Figura 8).

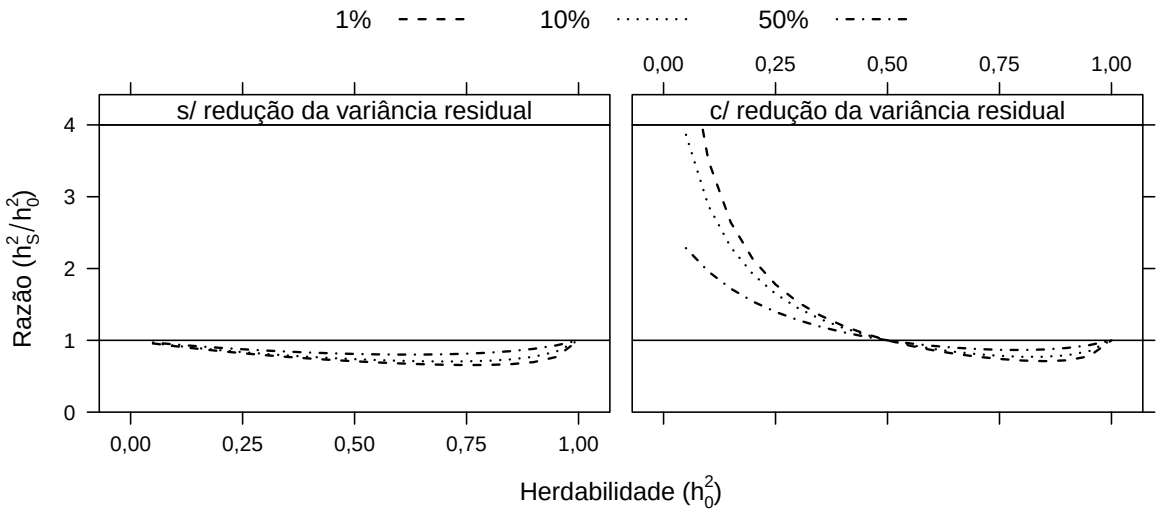

Figura 8 – Correspondência entre as estimativas da herdabilidade antes (h_0^2) e após (h_s^2) a seleção em função da herdabilidade e intensidade de seleção, com e sem redução da variância residual

Ao considerar negligível a redução da variância genética residual, observa-se que em herdabilidades extremas, não há diferenças quanto a estimar a variância da população de melhoramento antes ou após a seleção. Nos demais casos, sempre se incorre em subestimativas quando se utiliza apenas os indivíduos selecionados. Quanto mais intensa for a seleção, maior será o viés na subestimativa que utiliza apenas os indivíduos selecionados (Figura 8).

Em havendo redução da variância residual, é possível observar que fixada uma intensidade de seleção, quando a herdabilidade é igual a 0,5, ou quando toda variação é genética ($h^2 = 1$) não há diferença em obter essa estimativa de toda a população ou de sua parte selecionada. Coeficientes de herdabilidade maiores que 0,5 e menores que 1 apresentam uma subestimativa quando estimada por meio dos índivíduos selecionados. Quando a herdabilidade assume valores menores que 0,5, estimar a herdabilidade por meio desses indivíduos acarreta em sobrestimativas tanto maiores quanto menor for tal parâmetro. Ao fixar a herdabilidade, quanto mais branda a seleção praticada, menores são as magnitudes do desvio na sua estimativa obtida do conjunto de indivívuos além do limiar de seleção (Figura 8).

Contrapondo o que se observou graficamente com as estimativas da acurácia (análogo da herdabilidade) dos caracteres, observa-se que a variação entre indivíduos dentro de cada ciclo, após a seleção pode ser uma subestimativa da variância na população de melhoramento. Sob certas circunstâncias entretanto, o viés na estimativa pode ser previsível. Portanto, ao menos como exemplo de aplicação, a adaptação da teoria dos índices de seleção para a estimativa do progresso simultâneo é válida. Ela visa mensurar o progresso de uma supervariável que reúne

mais de um caráter, considera o relacionamento entre eles, mas impõe uma restrição a fim de representar os progressos individuais obtidos.

Na média dos quatro locais, o progresso simultâneo dos caracteres *pfc*, *alt*, *nfe*, *ccc*, *l3f* e *c3f* representou acréscimos significativos de 6% por era. Ou seja, no conjunto desses caracteres, em média obtiveram-se progressos de aproximadamente 0,6% por ano. Outros estudos no Brasil reportam progressos na produtividade de grãos em várias espécies, em torno de 1% ao ano (SOARES; RAMALHO; SOUSA, 1994; ABREU et al., 1994). A estimativa do progresso simultâneo em muito se aproxima do que se vem alcançando em outras culturas. Ao considerar que esforços foram despendidos não só em relação a *pfc*, o progresso simultâneo representa o sucesso do programa em melhorar simultaneamente todos os caracteres envolvidos.

Nos locais *M1* e *M2* os progressos simultâneos foram respectivamente de 16% e 5% por era. O maior progresso simultâneo observado em *M1* pode ser atribuído ao pior desempenho médio observado nesse local e aos substanciais ganhos estimados para os caracteres *ccc*, *l3f* e *c3f*. Nos locais *PN* e *RN* não houve progressos significativos. Vale frisar que, em média, nesses locais, melhores desempenhos médios foram observados, tanto no índice como em todos os caracteres utilizados na sua construção.

Sarcevic et al. (2013) estimaram decréscimos significativos no conteúdo de nicotina, mas não observaram alterações significativas nos teores de açúcares redutores. Os resultados reportam um decréscimo significativo na qualidade geral (*iqg*) do tabaco no Brasil. Como já comentado, esse fato pode ser devido a tais avaliações, que, devido ao custo, são apenas realizadas nas fases finais de avaliação. Assim, como reflexo das correlações negativas entre a produtividade e características químicas é obtido sucesso no incremento da produtividade em detrimento da qualidade do tabaco.

Não foram encontradas alterações significativas na remuneração da venda do tabaco *$ha*. Na Croácia, foram noticiados progressos da ordem de 55,42$US\$\, ha^{-1}$ por ano. Contrapondo esses resultados, pode-se inferir que o ganho em remuneração na Croácia advém tanto do incremento em produtividade quando da melhoria da composição química (teor de nicotina), enquanto que a constância da remuneração estimada pode ser reflexo da redução da qualidade geral das folhas compensada pelo maior incremento em produtividade (27,53$kg\, ha^{-1}$ por ano no Brasil, frente a 17,39$kg\, ha^{-1}$ ao ano na Croácia).

3 CONSIDERAÇÕES FINAIS

A cultura do tabaco, possui grande relevância no meio acadêmico, sendo utilizada como espécie modelo em diversas disciplinas, inclusive nos estudos de genética. A origem botânica do gênero foi estudada utilizando a citogenética (KENTOM et al., 1993). Marcadores genéticos foram empregados para avaliar a diversidade entre e dentro de subpopulações (BINDLER et al., 2007; LI; BASS; CORNELIUS, 2006). No melhoramento, técnicas ao nível do *DNA* mostraram-se ferramentas eficazes na seleção de caracteres relativos à resistência ao parasitismo, morfologia, qualidade e composição química das folhas (JULIO et al., 2006; CZUBACKA; DOROSZEWSKA, 2010; JULIO et al., 2008).

No Brasil, outros trabalhos, concomitantes a este, fomentam decisões relativas ao melhoramento da cultura. Toledo et al. (2013) fizeram uso de simulação computacional com intuito de estudar a melhor configuração de forma e tamanho das parcelas nas fases finais de avaliação de genótipos. Outro estudo, almejou dimensionar a rede suficiente de experimentos finais de avaliação para conferir precisão na seleção/recomendação de cultivares (PULCINELLI et al., 2014, no prelo).

Além do apelo científico, deve-se fazer referência à inegável importância que a cultura tem para a economia do país. O Brasil está atrás apenas da China em produção e é líder mundial nas exportações. Na última safra foram produzidas 730 mil toneladas e 88% da produção foi exportada. O grupo varietal *Virginia* responde por cerca de 85% da produção, ou seja, 606 mil toneladas produzidas numa área de 264 mil hectares em 2012/2013. Tais fatos reforçam e justificam os recursos despendidos nesta tese, visto que se trata de uma forma de mensurar o retorno do investimento na pesquisa com a cultura.

A pesquisa com a espécie, incluindo o melhoramento, vem conseguindo sucessos significativos, deve-se enfatizar que atualmente a produtividade média do grupo varietal *Virginia* é de $2297 kg\, ha^{-1}$ (POLL et al., 2013). Pode-se comparar a produtividade média nas lavouras comerciais com os resultados obtidos e avaliar quanto dos ganhos genéticos estimados experimentalmente são efetivamente transferidos ao produtor rural. Deve-se fazer esse comparativo para otimizar a eficiência do programa e avaliar as técnicas empregadas ou a serem implementadas, tanto as relacionadas a seleção e melhoramento, quanto de outras áreas da pesquisa com a espécie.

Se o objetivo é medir o sucesso da seleção praticada para múltiplos caracteres, ao adaptar a teoria dos índices de seleção, pode-se sumarizar estimativas dos progressos individuais considerando o relacionamento entre os caracteres. Para a produtividade de folhas curadas, altura de plantas, número de folhas expandidas, circunferência do caule no colo da planta, além de largura e comprimento de folhas, os resultados indicam progressos genéticos de 0,6% por ano. Esse resultado representa um sucesso da seleção em atender as expectativas dos melhoristas.

O índice apresentado, foi empregado com o objetivo de estimar o progresso genético simultâneo. O método apresentado, mostrou-se promissor em representar o progresso

genético, quando mais de um caráter sofreu seleção. Os coeficientes que ponderam os caracteres no índice também mostraram-se promissores em interpretar a importância relativa de cada caráter no processo de seleção realizado.

Além de ser utilizado para mensurar o progresso simultâneo, o índice pode também ser empregado com vistas a seleção para múltiplos caracteres. Quando o propósito é a seleção, pode-se arbitrar diferentes diferenciais de seleção. O método apresentado, ao contrário da seleção por níveis independentes de seleção, considera o relacionamento entre caracteres. Além da restrição ser expressa na mesma unidade do caráter, não há necessidade de arbitrar pesos econômicos no processo de estimação. A importância que o melhorista atribui a cada caráter é imposta pelos distintos diferenciais de seleção.

REFERÊNCIAS

ABREU, A. de F. B.; RAMALHO, M. A. P.; SANTOS, J. B.; MARTINS, L. A. Progresso do melhoramento genético do feijoeiro nas décadas de setenta e oitenta nas regiões sul e alto paranaíba em minas gerais. **Pesquisa Agropecuária Brasileira**, Brasília, v. 29, p. 105–112, 1994.

BARBOSA NETO, J. F.; MATIELLO, R. R.; CARVALHO, F. I. F.; OLIVEIRA, J. M. S.; PEGORARO, D. G.; SCHNEIDER, F.; SORDI, M. E. B.; VACARO, E. Progresso genético no melhoramento da aveia-branca no sul do Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, v. 35, p. 1605–1612, 2000.

BERNARDO, R. **Breeding for Quantitative Traits in Plants**. 2nd. ed. Hardbound: Stemma, 2010. 369 p.

BINDLER, G.; DER HOEVEN, R. van; GUNDUZ, I.; PLIESKE, J.; GANAL, M.; ROSSI, L.; GADANI, F.; DONINI, P. A microssatellite marker based linkage map of tobacco. **Theoretical and Applied Genetics**, Berlin, v. 114, p. 341–349, 2007.

BOWMAN, D. T.; SISSON, V. A historical overview of flue-cured tobacco breeding. **Tobbaco Science**, Raleigh, v. 44, p. 59–64, 2000.

BOWMAN, D. T.; WEEKS, W. W.; WILKINSON, C. A. Stability of alkaloid production in flue-cured tobacco. **Crop Science**, Madison, v. 31, p. 1121–1124, 1991.

BOWMAN, D. T.; WERNSMAN, E. A.; CORBIN, T. C.; TART, A. G. Contribution of genetics and production technology to long term gain in flue-cured tobacco. **Tobacco Science**, Raleigh, v. 28, p. 30–35, 1984.

BRASIL. Instrução Normativa n° 25 do MAPA, de 23 de maio de 2006. Estabelece critérios do valor de cultivo e uso (*VCU*) de feijão. **Diário Oficial [da] República Federativa do Brasil**, Brasília, 23 mai. 2006. Disponível em: <http://www.agricultura.gov.br/legislacao>. Acesso em: 10 mai. 2014.

____. Instrução Normativa n° 10 do MAPA, de 16 de abril de 2007. Institui a classificação de tabacco. **Diário Oficial [da] República Federativa do Brasil**, Brasília, 16 abr. 2007. Disponível em: <http://www.agricultura.gov.br/legislacao>. Acesso em: 10 mai. 2014.

BRESEGHELLO, F. **Ganhos para produtividade pelo melhoramento genético do arroz irrigado no nordeste do Brasil**. 1995. 100 p. Dissertação (Mestrado em Agronomia) — Universidade Federal de Goias, Goiânia, 1995.

BRESEGHELLO, F.; MORAIS, O. P. de; PINHEIRO, P. V.; SILVA, A. C. S. oes; CASTRO, E. M. de; aES, E. P. G.; CASTRO, A. P. de; PEREIRA, J. A.; LOPES, A. M.; UTUMI, M. M.; OLIVEIRA, J. P. de. Results of 25 years of ipland rice breeding in brazil. **Crop Science**, Madison, v. 51, p. 914–923, 2011.

BULMER, M. G. **Mathematical Theory of Quantitative Genetics**. Oxford: University Press, 1985. 252 p.

CARGNIN, A.; SOUZA, M. A. de; FRONZA, V. Progress in breeding of irrigated wheat for the cerrado region of brazil. **Crop Breeding and Applied Biotechnology**, Viçosa, v. 8, p. 39–46, 2008.

____. Genetic and environmental contributions to increased wheat yield in minas gerais, brazil. **Scientia Agricola**, Piracicaba, v. 66, p. 317–322, 2009.

CERON-ROJAS, J. J.; CROSSA, J.; SAHAGUN-CASTELLANOS, J.; SANTACRUZ-VARELASY, F. C.-G. an A. A selection index method based on eigenanalysis. **Cropscience**, Madison, v. 46, p. 1711–1721, 2006.

CHIORATO, A. F.; CARBONELL, S. A. M.; VENCOVSKY, R.; JR, N. S. F.; PINHEIRO, J. B. Genetic gain in the breeding program of commom beans at IAC from 1989 to 2007. **Crop Breeding and Applied Biotechnology**, Viçosa, v. 10, p. 329–336, 2010.

COCHRAN, W. G. Consequences when the assumptions for the analyis of variance are nor satisfied. **Biometrics**, Washington, v. 3, p. 22–38, 1947.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos Biométricos Aplicados ao Melhoramento Genético**. 2. ed. Viçosa: UFV, 2001. 390 p.

CUNNINGHAM, E. P.; MOEN, R. A.; GJEDREM, T. Restriction of selection index. **Biometrics**, Washington, v. 26, p. 67–74, 1970.

CZUBACKA, A.; DOROSZEWSKA, T. Estimating agronomic traits of transgenic tobacco lines. **Euphytica**, Wageningen, v. 172, n. 1, p. 35–47, 2010.

ELSTON, R. C. A weight-free index for the purpose of ranking or selection with respect to several triats at a time. **Biometrics**, Washington, v. 19, p. 85–97, 1963.

FALCONER, D. S.; MACKAY, T. F. C. **Introduction to Quantitative Genetics**. 4th. ed. Essex: Benjamin Cummings, 1996. 480 p.

FERNANDES, J. S. C.; FRAZON, J. F. Thyrty years og genetic progress in maize (*Zea mays* l.) in a tropical environment. **Maydica**, Bergamo, v. 42, p. 21–27, 1997.

FERREIRA, D. F. **Estatística Multivariada**. Lavras: UFLA, 2008. 662 p.

FISHER, R. A. The use of multiple measurements in taxonomic problems. **Annals of Eugenics**, London, v. 7, p. 179–188, 1936.

GENTLE, J. E. **Matrix Algebra**. New York: Springer-Verlag, 2007. 528 p.

HARVILLE, D. A. Index selection with proportionality constraints. **Biometrics**, Washington, v. 31, p. 223–225, 1975.

HAZEL, L. N. The genetics basis for constructing selection indexes. **Genetics**, Austin, v. 28, p. 476–490, 1943.

HAZEL, L. N.; LUSH, J. L. The efficiency of three methods of selection. **Journal of Heredity**, Washington, v. 33, p. 393–399, 1942.

ITOH, Y.; YAMADA, Y. Comparisons of selection indices achieving predetermined proportional gains. **Genetic Selection Evolution**, Paris, v. 19, p. 69–82, 1987.

JULIO, E.; DENOYES-ROTHAN, B.; VERRIER, J. L.; de Borne, F. D. Detection of qtls linked to leaf and smoke properties in *Nicotiana tabacum* based on a study of 114 recombinant inbred lines. **Molecular Breeding**, Dordrecht, v. 18, p. 69–91, 2006.

JULIO, E.; LAPORTE, F.; REIS, S.; ROTHAN, C.; BORNE, F. D. Reducing the content of nornicotine in tobacco via targeted mutation breeding. **Molecular Breeding**, Dordrecht, v. 21, p. 369–381, 2008.

KEMPTHORNE, O.; NORDSKOG, A. W. Restricted seclection indices. **Biometrics**, Washington, v. 15, p. 10–19, 1959.

KENTOM, A.; PAROKONNY, A. S.; GLEBA, Y. Y.; BENNET, M. D. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. **Molecular Genetics and Genomics**, Berlin, v. 240, p. 159–169, 1993.

LANGE, C. E.; FEDERIZZI, L. C. Estimation of soybean genetic progress in the south of brazil using multi-environmental yield trials. **Scientia Agricola**, Piracicaba, v. 66, p. 309–316, 2009.

LEHMANN, E. L. Fisher, neyman-pearson theories of testing hypothesis: One theory or two? **Journal of Applied Statistics**, Abingdon, v. 88, p. 1242–1249, 1993.

LI, B. C.; BASS, W. T.; CORNELIUS, P. L. Resistance to tobacco black shank in nicotiana species. **Crop Science**, Madison, v. 46, n. 2, p. 554–560, 2006.

LIN, C. Y. Index selection for genetic improvement of quantitative characters. **Theoretical and Applied Genetics**, Berlin, v. 52, p. 49–56, 1978.

____. An iterative procedure for deriving selection indexes with constant restrictions. **Journal of animal science**, Champaign, v. 83, p. 2313–2318, 2005.

____. A simultaneous procedure for deriving selection indexes with multiple restrictions. **Journal of animal science**, Champaign, v. 83, p. 531–536, 2005.

MACKAY, I.; HORWELL, A.; GARNER, J.; WHITE, J.; MCKEE, J.; PHILPOTT, H. Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time. **Theoretical and Applied Genetics**, Berlin, v. 122, p. 225–238, 2011.

MAGNUSSEN, S. Economic weights for maximum simultaneos genetic progress. **Theoretical and Applied Genetics**, Berlin, v. 79, p. 289–293, 1990.

MALLARD, J. La theorie et le calcul des index de selection avec restriction: Synthese critique. **Biometrics**, Washington, v. 28, p. 713–735, 1972.

MENDES, F. F.; RAMALHO, M. A. P.; ABREU, A. F. B. Índice de seleção para escolha de populações segregantes de feijoeiro-comum. **Pesquisa Agropecuária Brasileira**, Brasília, v. 44, n. 10, p. 1312–1318, 2009.

MILES, J. W.; PANDEY, S. Long term selection in plants in the developing world. In: JANICK, J. (Ed.). **Plant Breeding Reviews: Long Term Selection: Crops, Animals and Bacteria**. New York: John Wiley & Sons, 2004. p. 45–88.

MRODE, R. A.; THOMPSON, R. **Linear Models for the Prediction of Animal Breeding Values**. 2nd. ed. Oxfordshire: CABI, 2005. 368 p.

MULAMBA, N. N.; MOCK, J. J. Improvement of yield potential of the eto blanco maize (*Zea mays* l.) population by breeding for plant traits. **Egypt Journal of Genetics and Cytology**, Alexandria, v. 7, p. 40–51, 1948.

PENG, S.; LAZA, R. C.; VISPERAS, R. M.; SANICO, A. L.; CASSMAN, K. G.; KHUSH, G. S. Grain yield of rice cultivars and lines developed in the philippines since 1966. **Crop Science**, Madison, v. 40, p. 307–314, 2000.

PESEK, J.; BAKER, R. J. Desired improvement in relation to selection indices. **Canadian Journal Plant Science**, Ottawa, v. 49, p. 803–804, 1969.

PIMENTEL GOMES, F. **Um curso de estatistica experimental**. 15. ed. Piracicaba: FEALQ, 2009. 451 p.

POLL, H.; KIST, B. B.; SANTOS, C. E. dos; CARVALHO, C. de; REETZ, E. R. **Anuário Brasileiro do Tabaco**. Santa Cruz do Sul: Editora Gazeta, 2013. 160 p.

PULCINELLI, C.; BRUZI, A. T.; TOLEDO, F. H. R. B.; RAMALHO, M. A. P. Experimental strategies in carrying out VCU for the tobacco crop II: dimension of the experimental network. **Genetics and Molecular Research**, Ribeirão Preto, 2014. *No prelo*.

R Development Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2014.

RAMALHO, M. A. P.; ABREU, A. de F. B.; SANTOS, J. B. dos; NUNES, J. A. R. **Aplicações da Genética Quantitativa no Melhoramento de Plantas Autógamas**. Lavras: UFLA, 2012. 522 p.

RAMALHO, M. A. P.; FERREIRA, D. F.; OLIVEIRA, A. C. de. **Experimentação em Genética e Melhoramento de Plantas**. 2ª. ed. Lavras: UFLA, 2005. 322 p.

REDAELLI, R.; LAGANà, P.; RIZZA, F.; NICOSIA, O. L. D.; CATTIVELLI, L. Genetic progress of oats in italy. **Euphytica**, Wageningen, v. 164, p. 679–687, 2008.

RESENDE, M. D. V.; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, Goiânia, v. 37, p. 182–194, 2007.

RODRIGUES, J. A. S. **Progresso genético e potencial de risco da cultura do sorgo granífero (*Sorgum bicolor* L. Moench) no Brasil**. 1990. 171 p. Tese (Doutorado em Genética e Melhoramento de Plantas) — Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo, Piracicaba, 1990.

ROYSTON, P. A remark on algorithm as 181, the W–test for normality. **Journal of Applied Statistics**, Abingdon, v. 44, p. 547–552, 1995.

SARCEVIC, H.; GUNJACA, J.; BUDIMIR, A.; BOIC, M.; BOLARIC, S.; BUKAN, M.; LEWIS, R. S.; KOZUMPLIK, V. Long-term genetic improvement and genetic diversity of croatian flue-cured tobacco (nicotiana tabacum l.) cultivars. **Crop Science**, Madison, v. 3, p. 112–120, 2013.

SEARLE, S. R. **Linear Models**. New York: Wiley e Soon, 1997. 560 p.

SMITH, H. F. A discriminant function for plant selection. **Annals of Eugenics**, London, v. 7, p. 240–250, 1936.

SOARES, A. A.; RAMALHO, M. A. P.; SOUSA, A. F. de. Estimativa do progresso genético obtido pelo programa de melhoramento de arroz irrgado da EPAMIG, na época de oitenta. **Pesquisa Agropecuária Brasileira**, Brasília, v. 29, p. 97–104, 1994.

SOARES, A. A.; SANTOS, P. G.; MORAIS, O. P.; SOARES, P. C.; REIS, M. S.; SOUZA, M. A. de. Progresso genético obtido pelo melhoramento do arroz de sequeiro em 21 anos de pesqueisa em minas gerais. **Pesquisa Agropecuária Brasileira**, Brasília, v. 34, p. 415–424, 1999.

SOUZA, M. A.; MORAIS, O. P. de; HERáN, R. E. C.; CARGNIN, A.; PIMENTEL, A. J. B. Progresso genético do melhoramento de arroz de terras altas no período de 1950 a 2001. **Pesquisa Agropecuária Brasileira**, Brasília, v. 42, p. 371–376, 2007.

STEEL, R. G. D.; TORRIE, J. H.; DICKEY, D. A. **Principles and Procedures of Statistics: A Biometrical Approach**. 3th. ed. New York: McGraw-Hill, 1996. 672 p.

TABIEN, R. E.; SAMONTE, S. O. P. B.; MCCLUNG, A. M. Forty-eight years of rice improvement in texas since the release of cultivar bluebonnet in 1944. **Crop Science**, Madison, v. 48, p. 2097–2106, 2008.

TALLIS, G. M. A selection index for optimum genotype. **Biometrics**, Washington, v. 18, n. 1, p. 120–122, 1962.

____. Constrained selection. **Japanese Journal of Genetics**, Mishima, v. 60, p. 151–155, 1985.

TEFERA, H.; KAMARA, A. Y.; ASSAFO-ADJEI, B.; DASHIELL, K. Breeding progress for grain yield and associated traits in medium and late maturng promiscuous soybeans in nigeria. **Euphytica**, Wageningen, v. 175, p. 251–260, 2010.

TOLEDO, F. H. R. B.; RAMALHO, M. A. P.; PULCINELLI, C. E.; BRUZI, A. T. Experimental strategies in carrying out VCU for tobacco crop I: plots design and size. **Genetics and Molecular Research**, Ribeirão Preto, v. 12, p. 3766–4774, 2013.

TRACY, W. F.; TIEFENTHALER, I. L. G. A. E.; SCHABER, M. A. Trends in productivity of US crops and long-term selection. In: JANICK, J. (Ed.). **Plant Breeding Reviews: Long Term Selection: Crops, Animals and Bacteria**. New York: John Wiley & Sons, 2004. p. 89–108.

UNDERDAHL, J. L.; MERGOUM, M.; RANSON, J. K.; SCHATZ, B. G. Agronomic traits improvement and association in hard red spring wheat cultivars released in north dakota from

1968 to 2006. **Crop Science**, Madison, v. 48, p. 158–166, 2008.

USTUN, A.; ALLEN, F. L.; ENGLISH, B. C. Genetic progress in soybean of the US midsouth. **Crop Science**, Madison, v. 41, p. 993–998, 2001.

VENCOVSKY, R.; MORARES, A. R.; GARCIA, J. C.; TEIXEIRA, N. M. Progresso genético em vinteanos de melhoramento do milho no Brasil. In: EMBRAPA. **Anais do congresso nacional de milho e sorgo**. Sete Lagoas, 1986. p. 300–307.

VENCOVSKY, R.; RAMALHO, M. A. P. Contribuições do melhoramento de plantas no Brasil. In: PATERNIANI, E. (Ed.). **Ciência, Agricultura e Sociedade**. Brasília: Embrapa, 2006. p. 41–74.

WILLIAMS, J. S. The evaluation of a selection index. **Biometrics**, Washington, v. 18, p. 375–393, 1962.

WRICKE, G.; WEBER, W. E. **Quantitative Genetics and Selection in Plant Breeding**. New York: De Gruytier, 1986. 406 p.