



CLAUDIO CARLOS FERNANDES FILHO

**BIOMETRICAL METHODS FOR ANALYSIS OF MULTI-
HARVEST FORAGE (*Urochloa spp.*, *Panicum maximum* and
Medicago sativa) BREEDING TRIALS**

**LAVRAS - MG
2023**

CLAUDIO CARLOS FERNANDES FILHO

**BIOMETRICAL METHODS FOR ANALYSIS OF MULTI-HARVEST FORAGE
(*Urochloa spp.*, *Panicum maximum* and *Medicago sativa*) BREEDING TRIALS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração em Genética e Melhoramento de Plantas, para obtenção do título de Doutor.

Prof. Dr. José Airton Rodrigues Nunes
Orientador

Prof. Dr. Sanzio Carvalho Lima Barrios
Coorientador

**LAVRAS – MG
2023**

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Filho, Claudio Carlos Fernandes.

Biometrical methods for analysis of multi-harvest forage(*Urochloa spp.*, *Panicum maximum* and *Medicago sativa*)breeding trials / Claudio Carlos Fernandes Filho. - 2022.
122 p. : il.

Orientador(a): José Airton Rodrigues Nunes.

Coorientador(a): Sanzio Carvalho Lima Barrios.

Tese (doutorado) - Universidade Federal de Lavras, 2022.

Bibliografia.

1. Melhoramento de pastagens. 2. Genótipo x Ambiente. 3. Modelos Mistos. I. Nunes, José Airton Rodrigues. II. Barrios, Sanzio Carvalho Lima. III. Título.

CLAUDIO CARLOS FERNANDES FILHO

**BIOMETRICAL METHODS FOR ANALYSIS OF MULTI-HARVEST FORAGE
(*Urochloa spp.*, *Panicum maximum* and *Medicago sativa*) BREEDING TRIALS**


**MÉTODOS BIOMÉTRICOS PARA ANÁLISE DE EXPERIMENTOS DE
FORRAGEIRAS (*Urochloa spp.*, *Panicum maximum* and *Medicago sativa*) SOB
MÚLTIPLAS COLHEITAS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, área de concentração em Genética e Melhoramento de Plantas, para a obtenção do título de Doutor.

APROVADA em 18 de agosto de 2022.

Dr. José Airton Rodrigues Nunes
Dr. Sanzio Carvalho Lima Barrios
Dr. Esteban Fernando Rios
Dr. Júlio Sílvio de Sousa Bueno Filho
Dr. Tiago de Souza Marçal

UFLA
Embrapa Gado de Corte
University of Florida
UFLA
UFLA


Dr. José Airton Rodrigues Nunes
Orientador

**LAVRAS – MG
2023**

Dedico à minha mãe, Fabrícia, pela vida,
cuidados, ensinamentos e amor incondicional.

AGRADECIMENTOS

À minha mãe Fabrícia por todo o apoio, pela educação que me foi dada, pelo exemplo de prosseguir na área acadêmica, pela paciência e confiança. Mãe, seu carinho e atenção foram fundamentais para alimentar a esperança e prosseguir.

Ao meu irmão Gabriel pelo incentivo, principalmente por ser o meu melhor amigo.

A todos os meus familiares, agradeço-lhes pelo apoio, carinho, torcida e orações.

À Universidade Federal de Lavras e ao Programa de Pós-Graduação em Genética e Melhoramento de Plantas, pela oportunidade de cursar o Doutorado.

Ao Programa de Pós-graduação em Genética e Melhoramento de Plantas, sediado no Departamento de Biologia da Universidade Federal de Lavras, por todo o suporte para a realização do doutorado.

Ao professor e orientador, José Airton, pela orientação, confiança, ensinamentos transmitidos, disponibilidade e, sobretudo, pela amizade construída.

Aos professores de Pós-Graduação, pelos ensinamentos e conselhos.

Às agências de fomento CAPES e FAPEMIG pela concessão da bolsa de estudos, essencial para possibilitar o curso do Doutorado e pela possibilidade de realização do doutorado sanduíche na University of Florida pelo Programa CAPES-PrInt-UFLA.

À Embrapa Gado de Corte pela oportunidade de utilizar os experimentos desta tese e por toda a infraestrutura e apoio.

À Universidade da Florida, em especial, ao “Forage Breeding and Genomics Lab”, pela oportunidade de crescimento profissional e, principalmente, pela amizade construída entre os pesquisadores e estudantes.

Ao professor Esteban F. Rios, pela confiança, ensinamentos transmitidos e apoio.

Aos colegas do Núcleo de Estudos em Genética e Melhoramento de Plantas “GEN”, em geral, aos amigos que compartilharam ansiedade e, ao mesmo tempo, colaboração durante as disciplinas e eventos realizados; agradeço-lhes pelo prazeroso trabalho em equipe e pela amigável convivência.

Aos meus amigos Mario, Rebecca e Pablo, do laboratório Forage Breeding and Genomics Lab, pelo apoio, conselhos e pela amigável convivência dentro e fora do ambiente de trabalho.

Aos meus amigos conhecidos em Gainesville-FL, Jéssica, Cecília, Renato e Bruna, pelo acolhimento, amizade e convivência.

Aos amigos das repúblicas Carandiru e Último Gole, por proporcionarem momentos inesquecíveis, também pelos conselhos e risadas.

Às minhas companheiras de grupo sob orientação do professor José Airton, Brena e Maiara pela amizade e convivência.

Aos membros da banca examinadora, pela disponibilidade e sugestões apresentadas.

Enfim, a todas as pessoas que me auxiliaram, direta ou indiretamente, para que os meus objetivos fossem concretizados. Meu muito obrigado!

ABSTRACT

This study focuses on the optimization of statistical methods in forage breeding trials, with a goal to improve the efficiency of the breeding process and increase the rate of genetic gain. The study is conducted on three forage species: *Medicago sativa*, *Panicum maximum*, and *Urochloa* spp. The first chapter of the study evaluates the use of spatial analysis in breeding trials, taking into consideration the spatial variation and correlations within a trial and between repeated measurements. The results of this chapter provide insight into the effectiveness of spatial analysis in forage breeding trials. The second chapter of the study focuses on the application of random regression and factor analytic mixed models to deal with longitudinal data generated in forage breeding trials. These models account for temporal correlations between repeated measurements and allow for the appropriate modeling of genetic effects over time. The results of this chapter highlight the usefulness of these methods in analyzing data from forage breeding trials. In the final chapter of the study, genomic selection is performed in alfalfa, incorporating enviromic-based data. This chapter highlights the potential of genomic selection in reducing breeding cycles and increasing the rate of genetic gain in perennial forage species. The results of this study provide valuable information for forage breeders and plant breeders in general, regarding the use of various statistical methods in breeding trials, and their potential impact on the efficiency of the breeding process and the rate of genetic gain. The findings of this study have the potential to contribute to the improvement of forage production, which is crucial for the supply of nutrient-dense food such as meat and milk, particularly in developing countries where forages are a primary source of nutrition for most ruminant livestock.

Keywords: Forage breeding; Genotype by Environment; Mixed Models; Spatial Analysis; Enviromics; Longitudinal Data.

RESUMO

Este estudo se concentra na otimização de métodos estatísticos em ensaios de melhoramento de forrageiras, com o objetivo de melhorar a eficiência do processo de melhoramento e aumentar a taxa de ganho genético. O estudo é realizado em três espécies de forrageiras: *Medicago sativa*, *Panicum maximum* e *Urochloa* spp. O primeiro capítulo do estudo avalia o uso da análise espacial em ensaios de melhoramento, levando em consideração a variação espacial e as correlações dentro de um ensaio e entre medições repetidas. Os resultados deste capítulo fornecem informações sobre a efetividade da análise espacial em ensaios de melhoramento de forrageiras. O segundo capítulo do estudo se concentra na aplicação de modelos mistos de regressão aleatória e análise fatorial para lidar com conjuntos de dados longitudinais gerados em ensaios de melhoramento de forrageiras. Esses modelos levam em consideração as correlações temporais entre medições repetidas e permitem o adequado modelamento dos efeitos genéticos ao longo do tempo. Os resultados deste capítulo destacam a utilidade desses métodos na análise de dados de ensaios de melhoramento de forrageiras. No último capítulo do estudo, é realizada a seleção genômica em alfafa, incorporando dados baseados em ambiente. Este capítulo destaca o potencial da seleção genômica na redução dos ciclos de melhoramento e no aumento da taxa de ganho genético em espécies de forrageiras perenes. Os resultados deste estudo fornecem informações valiosas para melhoradores de forrageiras e melhoradores de plantas em geral, sobre o uso de vários métodos estatísticos em ensaios de melhoramento, e seu potencial impacto na eficiência do processo de melhoramento e na taxa de ganho genético. As descobertas deste estudo têm o potencial de contribuir para a melhoria da produção de forrageiras, o que é crucial para o fornecimento de alimentos ricos em nutrientes, como carne e leite, especialmente em países em desenvolvimento onde as forrageiras são a principal fonte de nutrição para a maioria dos animais de pecuária ruminantes.

Palavras-chave: Melhoramento de pastagens; Genótipo x Ambiente; Modelos Mistos; Análise Espacial; Ambientômica; Dados Longitudinais.

SUMÁRIO

GENERAL INTRODUCTION	12
REFERENCES	13
Accounting for spatial trends over harvest in forage breeding trials	15
Abstract	15
1 INTRODUCTION.....	15
2 MATERIAL AND METHODS	16
2.1 Data set.....	16
2.2 Statistical analysis	18
2.2.1 Base model.....	18
2.2.2 AR1 X AR1 and AR1 X AR1 + nugget models	18
2.2.3 SpATS model.....	20
2.2.4 Model comparison.....	21
3 RESULTS	23
3.1 Selecting the best AR1 model for each harvest	23
3.2 Comparison between Base, best AR1 and SpATS in the second stage analysis.	24
3.2.1 BIC, relative efficiency and heritability.....	25
3.2.2 Impact on the selection of the best genotypes.....	27
4 DISCUSSION	30
5 CONCLUSION	33
REFERENCES.....	33
Modeling genotype by harvest interaction by random regression (RRM) and factor analytic (FAMM) models.....	35
Abstract	35
1 INTRODUCTION.....	35
2 MATERIAL AND METHODS	38
2.1 Data set.....	38
2.2 Statistical analysis	38
2.2.1 First stage: estimating the genotypes' BLUEs and weights accounting for spatial field trends	39
2.2.2 Second stage: modeling the genotype by harvest interaction	40
3 RESULTS	46
3.1 Overall description of RRM and FAMM across trials.....	46

3.2	Genotype selection and GxH interpretation for RRM and FAMM	49
3.2.1	Alfalfa (<i>Medicago sativa</i> L.) breeding trial – T1	49
3.2.2	<i>Urochloa brizantha</i> advanced breeding trial – T5	57
3.2.3	<i>Urochloa decumbens</i> advanced breeding trial – T7.....	64
4	DISCUSSION	70
4.1	Random regression models (RRM).....	71
4.1.1	Goodness of fit evaluation	71
4.1.2	Genotype by harvest interaction.....	72
4.1.3	Genotypes’ reaction norm, adaptability and stability	73
4.2	Factor analytic mixed models (FAMM)	74
4.2.1	Goodness of fit evaluation	74
4.2.2	Genotype by harvest interaction.....	74
4.2.3	Genotypes’ reaction norm adaptability and stability	75
4.3	Comparison between RRM and FAMM for longitudinal data	76
5	CONCLUSION	77
	REFERENCES.....	77
	Genomic prediction for complex traits across multiples harvests in alfalfa (<i>Medicago sativa</i> L.) is enhanced by enviromics.....	35
	Abstract	82
1	INTRODUCTION.....	82
2	MATERIAL AND METHODS	86
2.1	Reference population	86
2.2	Genotypic data	86
2.3	Experimental design and phenotyping	87
2.4	Weather data collection.....	88
2.5	Weather data processing	89
2.6	Computing the environmental covariable matrices (E and Eg).....	90
2.7	Computing environmental and genomic similarity kernels (K_E , K_{Eg} , K_g).....	92
2.8	Statistical models	92
2.8.1	G-BLUP Model (M0).....	93
2.8.2	GE-BLUP Model (M1)	93
2.8.3	GE-BLUP- K_E (M2) and GE-BLUP- K_{Eg} (M3)	94
2.8.4	RN-BLUP- K_E (M4) and RN-BLUP- K_{Eg} (M5).....	94
2.9	Persistence, adaptability and stability evaluation.....	94

2.10	Cross-validation and predictive ability	95
3	RESULTS	97
3.1	Genetic and non-genetic parameters	97
3.2	Predictive ability of the models.....	98
3.2.1	CV2: predicting untested families in observed harvests	99
3.2.2	CV0: predicting tested families in unobserved harvests	101
3.2.3	CV1: predicting untested families in observed harvests	104
3.2.4	CV00: predicting untested families in observed harvests	105
4	DISCUSSION	106
4.1	Variance components	106
4.2	Predictive ability through cross-validation scenarios	107
4.3	Enviromic-based models can improve accuracy under unbalanced multi-harvest trials 107	
4.4	Reducing the number of harvests evaluation in alfalfa breeding trials.....	108
4.5	Multi-harvest genomic selection	109
4.6	Persistence, broad adaptability and stability	110
4.7	Differences between K_E and K_{Eg} on predictive ability.....	111
5	CONCLUSION	111
6	CONFLICT OF INTEREST	111
7	FUNDING INFORMATION.....	111
	REFERENCES.....	112
	SUPPLEMENTARY MATERIAL.....	119

GENERAL INTRODUCTION

Forages (grasses and legumes) are the principal source of nutrition for most ruminant livestock in developing countries, thus contributing to the supply of nutrient-dense foods like meat and milk (FUGILIE et al., 2021). The Brazilian cattle herd has increased, approximately, from 150 million heads in 1990 to 214 million in 2018, at the same time there was a reduction of pasture area, from 194 million hectares to less than 162 million (ABIEC, 2019). The meat production was increased by 139% in the same period (ABIEC, 2019). This increase in yield can be attributed by the reduction/substitution of areas with native forage species, about 18.7% between 2006 and 2017, by improved forage species (IBGE, 2017). The improved forage cultivars present higher yield and higher nutritive value, consequently resulting in greater carrying capacity and provide a differentiation in the meat in terms of competitiveness (lower production costs) and product quality, thus contributing to increase the GDP (gross domestic product) of Brazilian livestock (JANK et al., 2011). During the last 4-5 decades, plant breeders have made important contributions to livestock productivity by developing high yielding forage varieties with tolerances to biotic and abiotic stresses (MILES et al., 2006; MILES; HARE, 2007; AGUIRRE et al., 2013; CARDOSO et al., 2013; RAO et al., 2016; HERNANDEZ et al., 2017; ABD EL-NABYHAFEZ; HASHEM, 2018).

The most used forage grasses species (*Urochloa spp.* and *Panicum maximum*) and legume species (*Medicago sativa*) are perennial species, thus the breeding cycles in such species are longer and phenotypic evaluation are more expensive when compared to annual crops since evaluations must be made in several harvests during several years. Therefore, the application of proper statistical methods to identify the best genotypes to be selected and increasing the rate of genetic gain is critical (SMITH; SPANGENBERG, 2014). Statistical methods for analyzing data from perennial forage variety selection trials need to account for the spatial variation and correlations within a trial and temporal correlation between repeated measurements. The methods also need to appropriately model the genetic effects over time (SMITH et al., 2007; DE FAVERI et al., 2015). Furthermore, new breeding methods such as genomic selection (MEUWISSEN et al., 2001) can reduce the breeding cycles in perennial forage species (ANNICHIARICO et al., 2015; LI et al., 2015; SIMEÃO et al., 2021; ANDRADE et al., 2022; AONO et al., 2022) leading to higher genetic gains over time. This study was divided into three chapters. The first chapter, we evaluated the use of spatial analysis in three forage species breeding trials (*Medicago sativa*, *Panicum maximum* and *Urochloa spp.*). In the second chapter

we investigated the use of random regression and factor analytic mixed models in dealing with longitudinal data set generated in forage breeding trials. Finally, in the last chapter we performed genomic selection in alfalfa by incorporating environmic-based data.

REFERENCES

- ABD EL-NABY, Z. N.; HAFEZ, Wafaa Abd El-Karim; HASHEM, Hanan Ahmed. (2019). Remediation of salt-affected soil by natural and chemical amendments to improve berseem clover yield and nutritive quality. **African Journal of Range & Forage Science**, Scottsville, v. 36, n. 1, p. 49-60, 2018.
- ASSOCIAÇÃO BRASILEIRA DAS INDÚSTRIAS EXPORTADORAS DE CARNES (ABIEC). **Beef REPORT**: perfil da pecuária no Brasil. 2019. Disponível em: <https://www.beefpoint.com.br/beef-report-per%EF%AC%81-da-pecuaria-no-brasil/>. Acesso em: 10 maio 2019.
- AGUIRRE, L. M. *et al.* Characterization of resistance to adult spittlebugs (Hemiptera: Cercopidae) in *Brachiaria* spp. **Journal of Economic Entomology**, Oxford, v. 106, n. 4, p. 1871-1877, 2013.
- AONO, A. H. *et al.* A joint learning approach for genomic prediction in polyploid grasses. **Scientific Reports**, London, v. 12, n. 1, p. 1-17, 2022.
- CARDOSO, J. A. *et al.* Advances in improving tolerance to waterlogging in *Brachiaria* grasses. **Tropical Grasslands**, Brisbane, v. 1, n. 2, p. 197-201, 2013.
- DE FAVERI, J. *et al.* Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. **Crop and Pasture Science**, Victoria, Australia, v. 66, n. 9, p. 947-962, 2015.
- HERNANDEZ, L. M. *et al.* Phenotyping *Brachiaria* genotypes to assess *Rhizoctonia* resistance by comparing three inoculum types. **Plant Disease**, St. Paul, v. 101, n. 6, p. 916-923, 2017.
- JANK, L. *et al.* Breeding tropical forages. **Crop Breeding and Applied Biotechnology**, [S.l.], v. 11, p. 27-34, 2011.
- MILES, J. W.; HARE, M. D. Plant breeding and seed production of apomictic tropical forage grasses. **Bioforsk Fokus**, [S.l.], v 2, p. 74-81, 2007.
- MILES, J. W. *et al.* Recurrent selection in a synthetic *brachiaria* grass population improves resistance to three spittlebug species. **Crop Science**, Madison, v. 46, n. 3, p. 1088-1093, 2006.
- RAO, I. M. *et al.* Root adaptations to soils with low fertility and aluminium toxicity. **Annals of Botany**, Oxford, v. 118, n. 4, p. 593-605, 2016.
- SIMEÃO, R. M. *et al.* Genomic selection in tropical forage grasses: current status and future applications. **Frontiers in Plant Science**, [S.l.], v. 12, p. 665195, 2021.

SMITH, A. B. *et al.* Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. **Euphytica**, Wageningen, v. 157, p. 253-266, 2007.

SMITH, K. F.; SPANGENBERG, G. Forage breeding for changing environments and production systems: an overview. **Crop & Pasture Science**, Victoria, Austrália, v. 65, p. i-ii, 2014.

ACCOUNTING FOR SPATIAL TRENDS OVER HARVEST IN FORAGE BREEDING TRIALS

ABSTRACT

Selection of the best genotypes is the main goal in plant breeding, and the use of proper statistical method and its interpretation are very important for the genotype selection process. Accounting for spatial across harvests in forage breeding trial can improve the genetics gain and experimental precision. Therefore, in this study we investigated the use of spatial models in multi-harvest forage breeding trials for dry matter yield by applying a two-stage analysis. We evaluated the use of spatial models for three perennial forage species (*Urochloa spp.*, *Panicum maximum*, and *Medicago sativa*). In this study we compared three different models: (i) base model, without any spatial correction; (ii) first-order autoregressive (AR1) models; (iii) models based in tensor product (SpATS). The model comparisons were based in the Bayesian Information Criteria (BIC), by evaluating the fit of the BLUEs obtained for each harvest in a second stage analysis. Our results showed that the pattern of spatial variation changed across harvests, and the use of spatial correlation between plots must be evaluate in order to select the proper model for each harvest. Spatial models were effective in controlling local and global errors and achieved greater accuracy and efficiency over the base model. The spatial models also led to differences in the genotypes ranking and consequently in the selection of the best genotypes. Therefore, the spatial analysis can lead to greater selection gains because of increased heritability as well as decreased experimental errors. Ignoring the spatial variations can lead to mistakes in selecting the best genotypes. The SpATS model can be used as a standard method to spatially correct the genotypes BLUEs and generates weights to be used in a multi-harvest forage breeding trials.

1 INTRODUCTION

Phenotypic selection in perennial forage species is usually based on multi-harvest trials in which cover a potentially large and heterogeneous area. Multi-harvest in perennial trials can take several years and can be costlier when compared to annual crops. Therefore, the application of proper statistical methods to identify the best genotypes to be selected and increasing the rate of genetic gain is critical (SMITH and SPANGENBERG, 2014). Early generation breeding trials are composed by a large number of genotypes and incomplete block designs, as augmented block design, are used. However, these experimental designs do not control efficiently the microenvironmental variation. Furthermore, due to the large number of genotypes large areas are needed to conduct the trials increasing the probability of spatial trends on the field (VELAZCO et al., 2017). These spatial trends can affect the prediction of genotypic values leading to a mistake on the selection process (ANDRADE et al., 2020).

Several papers have reported the use of spatial models in perennial forage species. Nearest neighbor methods (PAPADAKIS, 1937) were applied by Casler (1999), Smith and Kearney (2002), and Smith and Casler (2004). Recently, De Faveri et al (2015) applied a first-order autocorrelation structure on rows and columns in the residual covariance matrix (AR1 x AR1), as described by Gilmour, Cullis and Verbyla (1997) to model spatial trends in alfalfa (*Medicago sativa*) breeding trials, and also to model the temporal correlation between harvests. However, the authors applied a unique AR1 structure across all harvests, therefore the model assumed common spatial parameters over harvests which may not always be the case (DE FAVERI et al., 2015). Nonetheless, fitting different spatial structures for each harvest may not be possible in a single stage analysis leading to convergence problems, mainly when the number of harvests and genotypes are large, and convergence can be more difficult when also modeling the genotype by harvest interaction, as the number of parameters to be estimated by the model increases substantially (WELHAM et al., 2010). To overcome this problem, the two-step analysis have been used extensively in the Australian National Variety Trials for annual crops (GOGEL; SMITH; CULLIS, 2018), since Smith (2001) proposed the two-stage analysis by applying weights from individual trial analysis into the multi-environment analysis.

In a two-stage analysis, genotypes' best linear unbiased estimates (BLUES) of individual trials in stage one are combined in a weighted across trials using a mixed model analysis in the stage two, where the weights provide a measure of relative uncertainty of the estimated genotypes' BLUES for each trial (SMITH 2001; MÖHRING and PIEPHO, 2009; GOGEL; SMITH; CULLIS, 2018). The weights from the individual analysis account for both within-trial heterogeneity and differing number of replications (WELHAM et al., 2010), by using the weights in the multi-trial analysis on the stage two, the residual heterogeneity across trials are also taking in account in the model, since the uncertainty measures are unique for each trial.

In this study, we investigated the changes in the spatial trends across harvests that may occur in multi-harvest trials for four forage species, and the application of spatial analysis in the selection of genotypes by applying weights from each spatial model in a two-stage multi-harvest analysis.

2 MATERIAL AND METHODS

2.1 Data set

In this study, we used data from three forage species conducted from 2015 to 2020 in three different locations (Florida – USA, Mato Grosso do Sul – Brazil and Brasilia – Brazil) by

using different experimental designs (augmented row-column design – ARCD, alpha-lattice design – ALD, and randomized complete block design – RCBD) for total dry matter yield (DMY, kg.ha⁻¹) (Table 1). The number of genotypes evaluated varied from 9 (T6) to 182 (T1). The data sets are composed by early breeding trials (T1, T2, T3, T8 and T10) in which a large number of genotypes were tested and advanced breeding trials (T6, T7 and T9) (Table 1). The number of harvests in each trial varied from 6 (T7, T8 and T10) to 13 (T6) (Figure 1).

Table 1 - Description of experimental layout for the three forage specie evaluated from 2015 to 2020.

Trial	Specie	Origin	Year	Location	Design	Genotypes	Harvests	Columns	Rows	Plots
T1	Alfalfa (<i>Medicago sativa</i>)	University of Florida	2018 - 2019	Citra - FL - USA	ARCD	182	11	32	14	405
T2	<i>Panicum maximum</i>	Embrapa	2016 - 2019	Campo Grande - MS - BR	ALD	110	9	22	20	330
T3	<i>Panicum maximum</i>	Embrapa	2016 - 2019	Campo Grande - MS - BR	ALD	110	9	22	20	330
T6	<i>Urochloa decumbens</i>	Embrapa	2018 - 2019	Campo Grande - MS - BR	RCBD	9	13	9	4	36
T7	<i>Urochloa decumbens</i>	Embrapa	2018 - 2019	Brasilia - MS - BR	RCBD	12	6	12	4	48
T8	(Inter) <i>Urochloa</i>	Embrapa	2015 - 2016	Campo Grande - MS - BR	RCBD	99	6	8	50	396
T9	(Inter) <i>Urochloa</i>	Embrapa	2019 - 2020	Campo Grande - MS - BR	RCBD	15	10	6	11	60
T10	<i>Urochloa decumbens</i>	Embrapa	2015 - 2016	Campo Grande - MS - BR	RCBD	36	6	3	50	144

Source: from the author (2022).

2.2 Statistical analysis

2.2.1 Base model

All trials were analyzed considering a linear mixed model assuming independent errors as shown in the equations 1 (ARCD), 2 (RCBD) and 3 (ALD), that is, considering only the global errors through the design used in the experiment for each harvest.

$$y = \mu + X_g \beta_g + Z_r u_r + Z_c u_c + e \quad (1);$$

$$y = \mu + X_g \beta_g + Z_b u_b + e \quad (2);$$

$$y = \mu + X_g \beta_g + X_r \beta_r + Z_b u_b + e \quad (3);$$

where, y is the vector of phenotypic data (e.g. DMY) per plot; μ is a vector of 1's representing the intercept; β_g is the vector of the fixed effects of genotypes in all trials; β_r is the vector of fixed effects of replications in the ALD (T2 and T3); u_r and u_c are the vectors of random effects of rows and columns, respectively in the ARCD (T1, Table 1), where $u_r \sim N(0, I_r \sigma_r^2)$ and $u_c \sim N(0, I_c \sigma_c^2)$; u_b is the vector of random effects of blocks in a randomized complete block design (T6, T7, T8, T9 and T10, Table 1), or the vector of the random effects of the blocks inside replication in the alpha-lattice design (T2 and T3, Table 1), in which $u_b \sim N(0, I_b \sigma_b^2)$; e is the vector of random errors, $e \sim N(0, I_e \sigma_e^2)$; σ_r^2 , σ_c^2 , σ_b^2 and σ_e^2 are the variance components associated to the random effects of rows, columns, blocks and errors. X , Z_r , Z_c and Z_b are design matrices for the fixed effects, the random effects of rows, columns and blocks, respectively. I_r , I_c , I_b , and I_e are identity matrices.

2.2.2 AR1 X AR1 and AR1 X AR1 + nugget models

In AR1 X AR1 models the errors are assumed to be autocorrelated through row and columns directions. These models were fitted by a first-order autoregressive process (GILMOUR; CULLIS; VERBYLA, 1997). For the AR1 X AR1, the model will be the same as described for the base model (1 and 2). The difference between the base model and AR1 X AR1

models consist in the (co)variance structure for the error matrices, in which in the AR1 X AR1 models, the errors are assumed $e \sim N(0, R\sigma_e^2)$, where three co(variance) structure presented in equations 3, 4 and 5 were tested:

$$R = \sigma_e^2 \sum_r (\rho_r) \otimes I_c \quad (4);$$

$$R = \sigma_e^2 I_r \otimes \sum_c (\rho_c) \quad (5);$$

$$R = \sigma_e^2 \sum_r (\rho_r) \otimes \sum_c (\rho_c) \quad (6);$$

where, R is the co(variance) structure of the errors; $\sum_r (\rho_r)$ and $\sum_c (\rho_c)$ represents the first-order autoregressive correlation matrices; ρ_r and ρ_c are the autocorrelation parameters along rows and columns, respectively; \otimes denotes the Kronecker product.

The AR1 X AR1 models can be extended by inserting the random term (η) representing the independent part of the error. The η term is commonly known as nugget or measurement error (GILMOUR; CULLIS; VERBYLA, 1997). The AR1 X AR1 + nugget model can be described as:

$$y = X\beta + Z_r u_r + Z_c u_c + e + \eta \quad (7);$$

$$y = X\beta + Z_b u_b + e + \eta \quad (8);$$

where, all terms in the models are the same as described for the base model; η is the vector of random independent errors, in which $\eta \sim N(0, I_\eta \sigma_\eta^2)$; σ_η^2 is the variance component associated to the independent errors; e can be described as the same in the AR1 X AR1 model, where $e \sim N(0, R\sigma_e^2)$ the co(variance) structure tested are described in equations 3, 4 and 5.

After fitting each model (AR1 X AR1 and AR1 X AR1 + nugget), they were compared by the Bayesian information criterion (BIC). Those models presenting smaller BIC were selected and compared to the Base Model and SpATS model in the second stage analysis, in this work we are designating the selected AR1 model as Best AR1.

2.2.3 SpATS model

In this model, field trends by a bivariate function of the spatial coordinates $f(r, c)$, represented by 2D P-splines. The cubic B-splines is used for second-order penalties considering the number of knots as the number of rows and columns in each trial. The SpATS model can be described as follow:

$$y = X\beta + X_s\beta_s + Z_s u_s + Z_c u_c + Z_r u_r + e \quad (9);$$

$$y = X\beta + X_s\beta_s + Z_s u_s + Z_b u_b + e \quad (10);$$

where all terms in the model 8 and 9 are the same as in models 1 and 2, except for $X_s\beta_s$ and $Z_s u_s$; β_s is the vector of fixed effects of the smooth spatial surface (unpenalized); u_s is the vector of random effects of the penalized part of the smooth surface (penalized). The fixed term ($X_s\beta_s$ - unpenalized) and random term ($Z_s u_s$ - penalized) describe the mixed model expression of the smooth spatial surface ($f(r, c) = X_s\beta_s + Z_s u_s$), where the random spatial vector u_s has (co)variance matrix S .

The SpATS model uses the P-spline ANOVA (PS-ANOVA) (LEE; DURBÁN; EILERS, 2013) to describe the 2D-splines in the mixed model framework. The X_s , Z_s incidence matrices, and the (co)variance matrix S are described by Lee, Durbán and Eilers (2013) and Rodríguez-Álvarez et al. (2018). The PS-ANOVA parametrization can be decomposed as a linear sum of the univariate and bivariate smooth functions (VELAZCO et al., 2017):

$$f(r, c) = \beta_{s1}r + \beta_{s2}c + \beta_{s3}rc + f_1(r) + f_2(c) + h_3(r)c + rh_4(c) + f_5(r, c) \quad (10)$$

where the spatial surface is represented by: linear trends across row and columns ($\beta_{s1}r + \beta_{s2}c$) and a linear interaction trend ($\beta_{s3}rc$); two main smooth trends through rows and columns [$f_1(r) + f_2(c)$]; two linear-by-smooth interaction terms, where the slope of a linear trend along one covariate (c or r) is allowed to vary smoothly as function of other covariate [$h_3(r)c + rh_4(c)$]; $f_5(r, c)$ is the pure smooth-by-smooth interaction between column and row trends.

The random vector of the penalized spatial trends (u_s) has five independent sub-vectors u_{sk} , with k representing the additive components from the interactions. Therefore, the spatial

covariance matrix S is the direct sum of the S_k matrices, where each parameter of sk depends on a specific smoothing parameter (λ_{sk}). These smoothing parameters can be estimated via residual maximum likelihood (REML) as the ratio between the residual variance of the corresponding spatial effect, $\lambda = \sigma_e^2 / \sigma_{sk}^2$. The spatial surface is adjusted by five parameters in an anisotropic process dealing with two variations, global and local.

2.2.4 Model comparison

The model comparisons were done in a second stage analysis, where BLUEs (Best Linear Unbiased Estimation) from the base model, best AR1 X AR1 model, and SpATS model. For each model, we adjusted the BLUEs by using the weights from each spatial model in a second stage analysis to account for the spatial trends and variance heterogeneity and modeling the genotype by harvest interaction with a compound symmetry (co)variance structure:

$$y = X\beta + Zg + e \quad (11)$$

where, y is the genotypes BLUEs in each harvest; β is the vector of fixed effects of the intercept and harvests; g is the vector of random effects of genotypes, in which $u_g \sim N(0, H \otimes I_g \sigma_g^2)$; e is the random effects of the errors associated with the estimation of y , where $e \sim N(0, \Sigma \sigma_e^2)$. H is genotype by harvest (co)variance structure matrix considered as a compound symmetry matrix, where the diagonal terms are composed by the $\sigma_g^2 + \sigma_{gh}^2$ and the off diagonal is composed by genetic variance component (σ_g^2); σ_{gh}^2 is variance component associated with the genotype by harvest interaction. Σ is the variance matrix of the errors, where $\Sigma = \text{diag}(\Sigma_1 \cdots \Sigma_j)$, in which j is the harvest. X and Z are the incidence matrix of the fixed and random effects.

In practice, Σ_j are unknown and replaced by an estimate $\hat{\Sigma}_j$ from each trial. It is sometimes not feasible to store and use the full matrix $\hat{\Sigma}_j$ from each harvest, and so a vector of approximate weights is required. We used the weights proposed by Smith, Cullis and Gilmour (2001), where the weights are based on the diagonal elements of $\hat{\Sigma}_j^{-1}$ designated as Π , in which:

$$\Pi = \text{diag}(\pi_1^T \cdots \pi_j^T) \quad (12)$$

π_j consists of the diagonal elements of $\hat{\Sigma}_j^{-1}$. This simple approximation reflects the uncertainty in each estimated BLUE, accounting for both within-trial heterogeneity and differing replication.

After fitting the BLUEs from the base model, Best AR1 model, and SpATS model by using the weights from each model and harvest in the model 11, we compared the models based on the BIC, generalized heritability (H^2), relative efficiency (RE) and coincidence index (CI).

2.2.4.1 Heritability

The heritability was computed based on the Cullis, Smith and Coombes (2006) estimator:

$$H^2 = 1 - \frac{\bar{V}_{BLUP}}{2\sigma_g^2} \quad (13)$$

where, \bar{V}_{BLUP} is the mean prediction error variance between two best linear unbiased prediction (BLUP) of two genotypes, and σ_g^2 is the genetic variance component.

2.2.4.2 Relative efficiency

The relative efficiency between the models were computed to evaluate the improvement on accuracy by using the \bar{V}_{BLUP} estimated from model 11:

$$RE_{Base/BestAR1} = \frac{\bar{V}_{BLUP_0}}{\bar{V}_{BLUP_1}} \times 100;$$

$$RE_{Base/SpATS} = \frac{\bar{V}_{BLUP_0}}{\bar{V}_{BLUP_2}} \times 100$$

where, \bar{V}_{BLUP_0} , \bar{V}_{BLUP_1} and \bar{V}_{BLUP_2} are the estimated error variance estimated by using BLUEs and weights in model 11 from the Base model, Best AR1 model and SpATS model, respectively. The RE cannot be used as a criterion to select the best model, therefore these estimates were used as supplementary information.

2.2.4.3 Impact in the selection

To infer about the spatial correction on the genotype selection we used five selection intensities, 10%, 20%, 30%, 40% and 50%, and the percentage of agreement of the best genotypes selection was computed by using the coincidence index (QIAO et al., 2000):

$$C = \frac{a}{a+b}$$

where, a is the number of genotypes selected by both models, and b is the number of genotypes selected only by the base model.

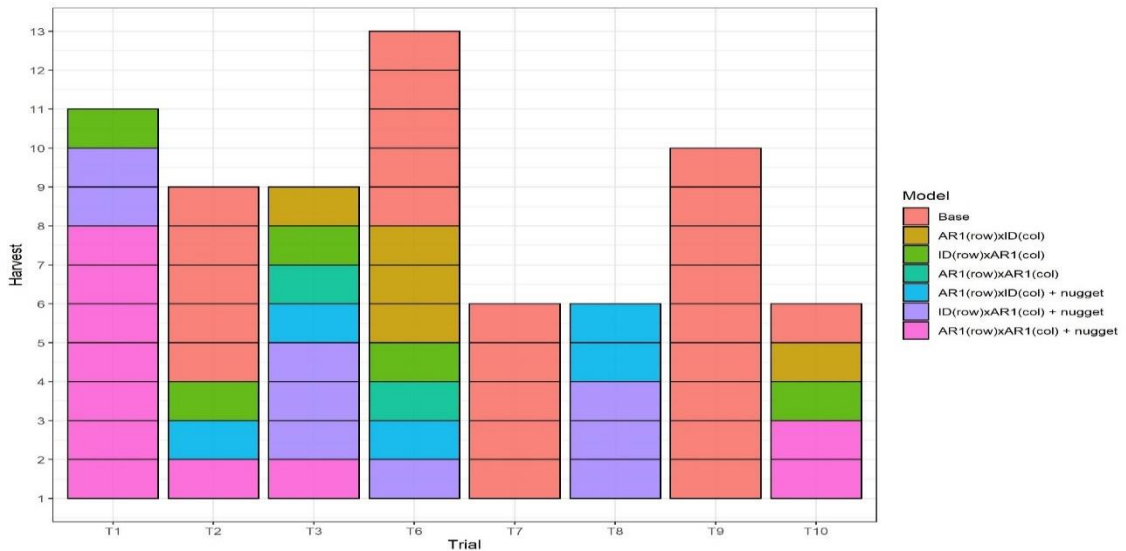
All the analyses were performed in R 4.1.1 (R CORE TEAM, 2021). The AR1 X AR1 models and second stage analysis were performed by the ASREML-R 4.1 package (BUTLER, 2021), and the SpATS model was performed by using the SpATS package (RODRIGEZ-ALVAREZ et al., 2018). All the figures were done by using the ggplot2 package (WICKHAM et al., 2016).

3 RESULTS

3.1 Selecting the best AR1 model for each harvest

The Figure 1 shows the best model selected for the 70 harvests and trials combinations. The spatial models were superior to the base model in approximately 64% of the harvests and trials combinations (smaller BIC values, data not shown). These results show the importance of evaluating the spatial correlation between plots in forage breeding trials. When observing the trials individually, there were changes in the best model across harvests, except for trials T7 and T9. This result evidenced that there was no unique model that would be appropriate for all harvests in a given trial (Figure 1). Only for the trials T7 and T9 the base model was superior for all harvests, indicating no presence of spatial trends in these trials (Figure 1). The use of the measurement error (nugget) was necessary in 71% of the cases where spatial modeling was needed, indicating the presence of variation inside the plots in the forage species evaluated (Figure 1).

Figure 1 - AR1 x AR1 and AR1 x AR1 + nugget models selected for each harvest and trial by using the smaller value of BIC for DMY.



Source: from the author (2022).

The trials T2 and T3 have the same genotypes and design, however these trials were evaluated in different conditions of fertility, in which the trial T2 was conducted under low fertility conditions and T3 under low fertility conditions. It was observed that unfertilized trials need to account for spatial trends in all harvests evaluated, whereas under high fertility levels only three harvests presented the need for spatial corrections (Figure 1). Therefore, spatial models can handle with heterogeneity present on field trials reducing the variance between plots. There was no clear pattern between the size of the trials and the need of using spatial corrections, where we observed that the smallest trial (T6 – Table 1) needed to account for spatial trends in five of the 13 harvests evaluated. Moreover, one of the largest trials (T2 – Table 1) did not need spatial corrections for most of the harvests evaluated (Figure 1).

3.2 Comparison between Base, best AR1 and SpATS in the second stage analysis.

Perennial forage breeders are always interested in the genotypes' performance across several harvests of evaluation. In this study we evaluated the use of proper estimated weights and BLUEs from each model in a multi-harvest analysis based on BIC, relative efficiency (RE), heritability (H^2), coincidence index (C) and Spearman correlation.

3.2.1 BIC, relative efficiency and heritability

By the BIC estimates, it was observed that the analysis in which used weights from models accounting to the spatial trends had smaller BIC values, except for trial T9 (Table 2). When comparing the best AR1 model to the SpATS model, it was observed better adjustment for the analysis in which used the weights form the SpATS model led to smaller BIC, except for trial T2 and T9 (Table 2).

Table 2 - BIC estimates for the in the second stage analysis by using weights from each harvest and model evaluated.

Trial	Base	Best AR1	SpATS
T1	26498	26372	26212*
T2	12621	12579*	12598
T3	12144	12142	12108*
T8	7987	8009	7944*
T6	1207	1204	1203*
T7	1013	1015	1009*
T9	1831*	1835	1843
T10	2807	2804	2768*

*Lowest BIC Between models for each trial.

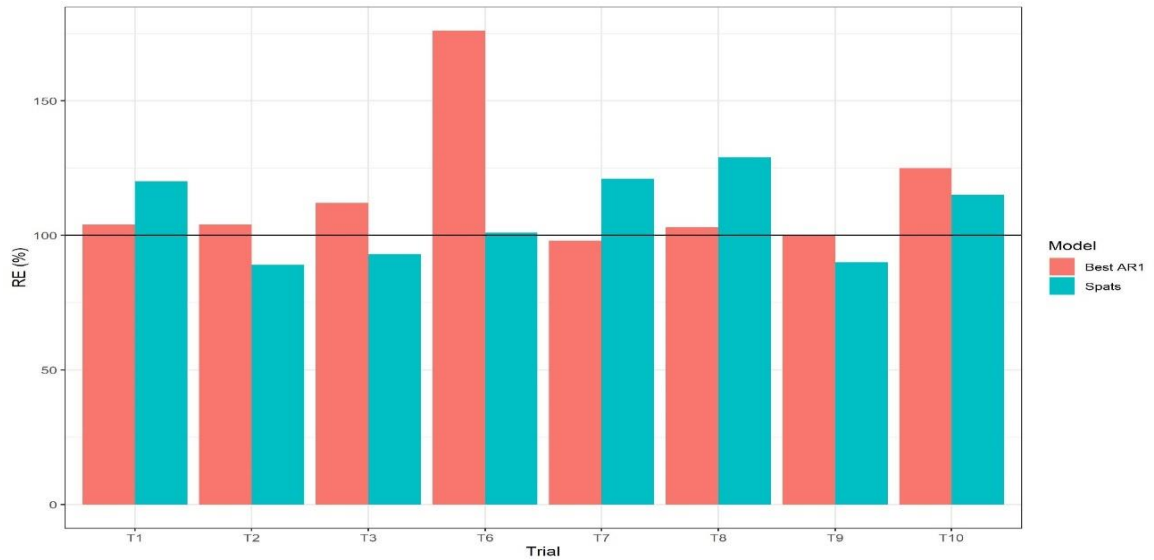
Source: from the author (2022).

RE was calculated based on the mean prediction error variance between two BLUPs (\bar{V}_{BLUP}), therefore the higher the RE lower the \bar{V}_{BLUP} and better the experimental precision. The RE varied from 89% (SpATS – T2) to 176% (Best AR1 – T6) (Figure 2). Although SpATS models tended to have smaller BIC, this did not reflect in greater RE where SpATS model had lower RE than the base model for trials T2, T3 and T9 (Figure 2 and Table 2). However, the best AR1 models always had RE at least equal to the base model when BIC was higher for best AR1 model (T7 – RE = 98% and T9 – RE = 100%) (Figure 2 and Table 2). The SpATS models had better RE than best AR1 model for trials T1, T7 and T8 (Figure 2).

The heritability estimates varied from 0.11 (T7 – SpATS) to 0.96 (T3 – SpATS) (Figure 3). Similar heritabilities were estimated when comparing the models, being slightly bigger for SpATS model (Figure 3). The main difference on heritability estimation ca be observed between SpATS model and best AR1 model for trial T6 where greater heritability was observed by using SpATS model. The AR1 model performs better in reducing the prediction error, since

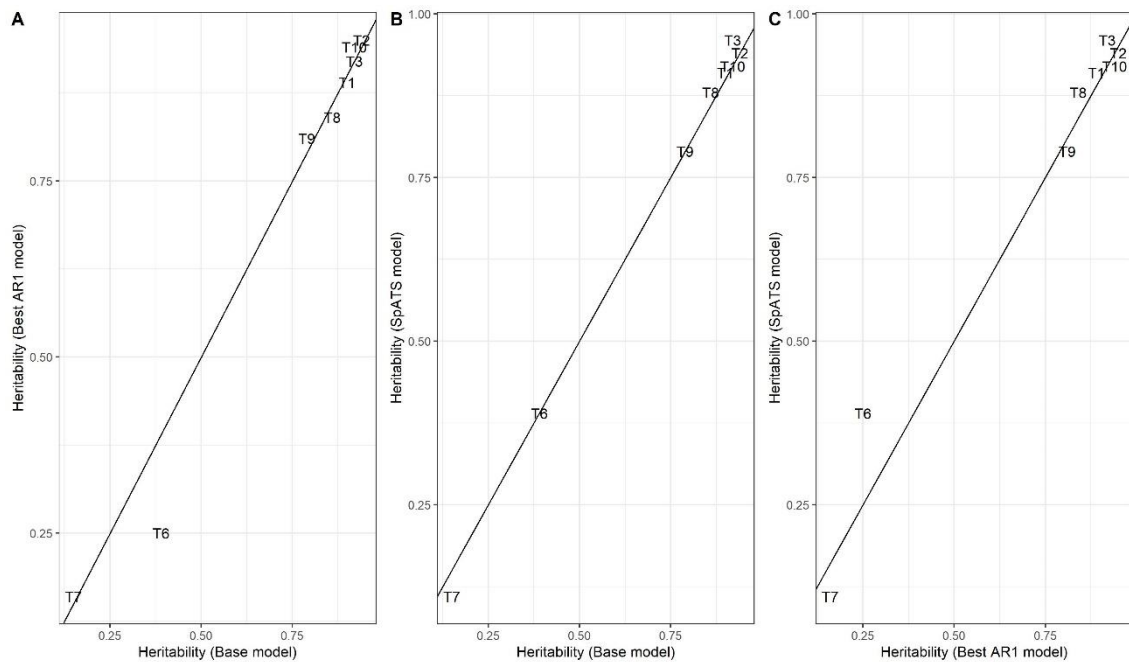
this models most always led to an increase in the RE (Figure 2). However, the SpATS model tend to increase the genetic variance, since this model had lower RE but the heritabilities were similar to the other models (Figure 2 and Figure 3).

Figure 2 - Relative efficiency based on the average of the mean prediction error variance between two BLUPs. The horizontal black line represents the threshold to show when spatial models were better (above the line) or worse (below the line) than the base model.



Source: from the author (2022).

Figure 3 - Heritability estimates contrasting different models – base model X best AR1 model (A); base model X SpATS model (B); best AR1 model X SpATS model (C).



Source: from the author (2022).

3.2.2 Impact on the selection of the best genotypes

The practical use of spatial models was measured based on the Spearman correlation and on the coincidence index (CI), to infer about the changing in rank that may occur when the genetic effects are corrected by spatial trends. Therefore, this is the most important subject for breeders since they are always concerned in selecting genotypes mostly accurate as possible. The overall changing in genotypes' ranking varied from 0.85 (SpATS model/base model – T6) to 1.00 (best AR1 model/base model – T7; Best AR1 model/SpATS model – T2 and T3) (Table 3). Greater differences on genotypes' ranking were observed between base model and spatial models based on the Spearman's correlations mean. Only for the trial T9 the genotypes' ranking had more changes between spatial models (Table 3) where the base model was the best based on BIC (Table 2).

Table 3 - Spearman correlations between predicted BLUPs by using the BLUEs and weights from the base model, best AR1 model and SpATS model.

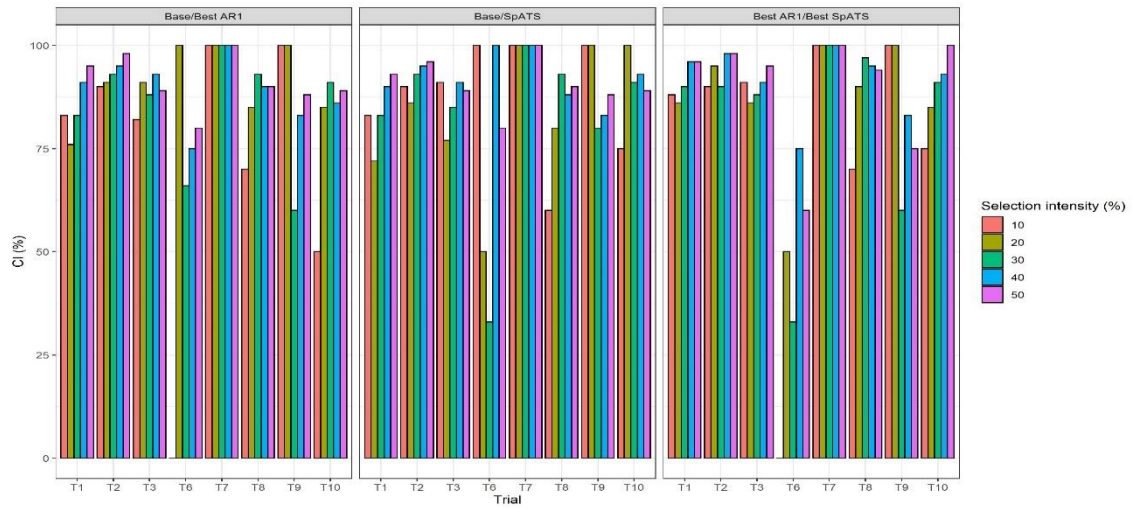
Trial	Base model/Best AR1 model	Base model/SpATS model	Best AR1 model/SpATS model
T1 ^a	0.97	0.96	0.99
T2 ^b	0.99	0.99	1.00
T3 ^a	0.95	0.94	1.00
T8 ^a	0.93	0.91	0.97
T6 ^a	0.87	0.85	0.90
T7 ^a	1.00	0.99	0.99
T9 ^c	0.95	0.97	0.91
T10 ^a	0.90	0.92	0.93
Mean	0.95	0.94	0.96

^aSpATS model had the smallest BIC; ^bBest AR1 model had the smallest BIC; and ^cBase model had the smallest BIC.

Source: from the author (2022).

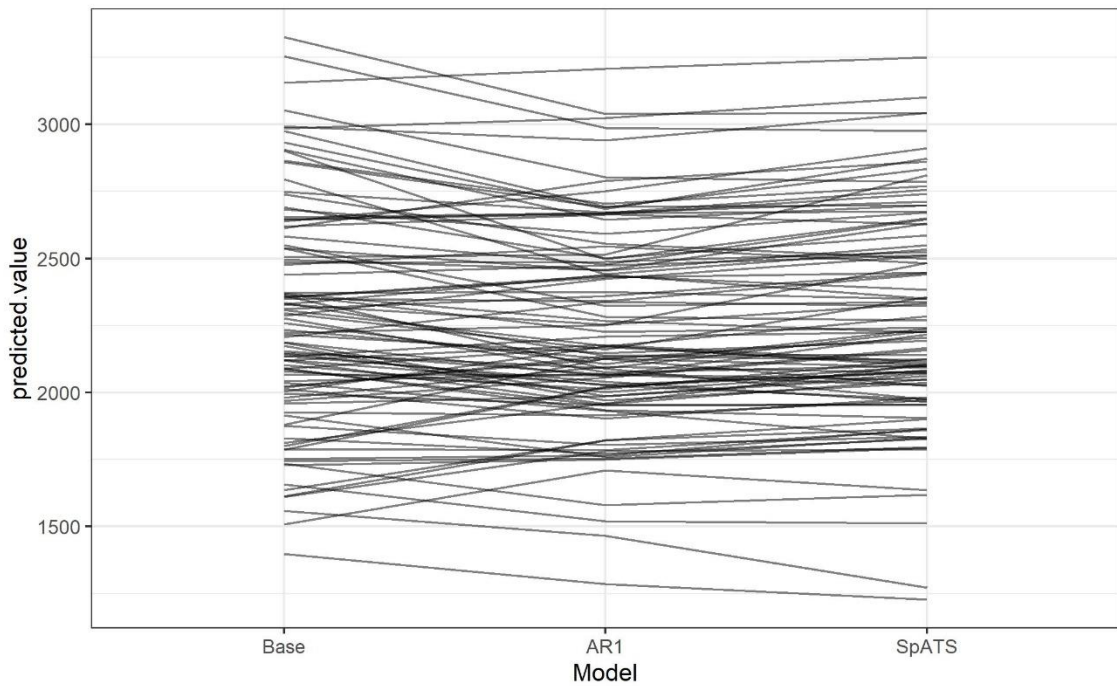
The greatest impact on the selection can be observed by the CI, where varied from 0% to 100% depending on the selection intensity and trial (Figure 4). As expected, the concordance between selected genotypes tended to increase as the selection intensity is less strict, on the average across all comparisons the coincidence between selected genotypes was from 78% using 10% of selection intensity to 90% using 50% of selection intensity, this trend can also be observed in most trials on Figure 4. The same genotypes were selected by using any model for the trial T7 (Figure 4). These results show the importance in using proper weights on the multi-harvest analysis. The changing in the ranking can also be verified on figure 5, where the predicted BLUPs are plotted for each model for trial T8.

Figure 4 - Coincidence between selected genotypes by using 10, 20, 30, 40 and 50% of selection intensity.



Source: from the author (2022).

Figure 5 - Predicted genotypes' BLUP in the second stage, by using BLUEs and weights from each model (Base, AR1 and SpATS) evaluated in the first stage of the analysis for trial T8.



Source: from the author (2022).

4 DISCUSSION

Selection of the best genotypes is the main goal in plant breeding, and the use of proper statistical method and its interpretation are very important for the genotype selection process. Accounting for spatial across harvests in forage breeding trial can improve the genetics gain and experimental precision. Therefore, in this study we investigated the use of spatial models in multi-harvest forage breeding trials by applying a two-stage analysis.

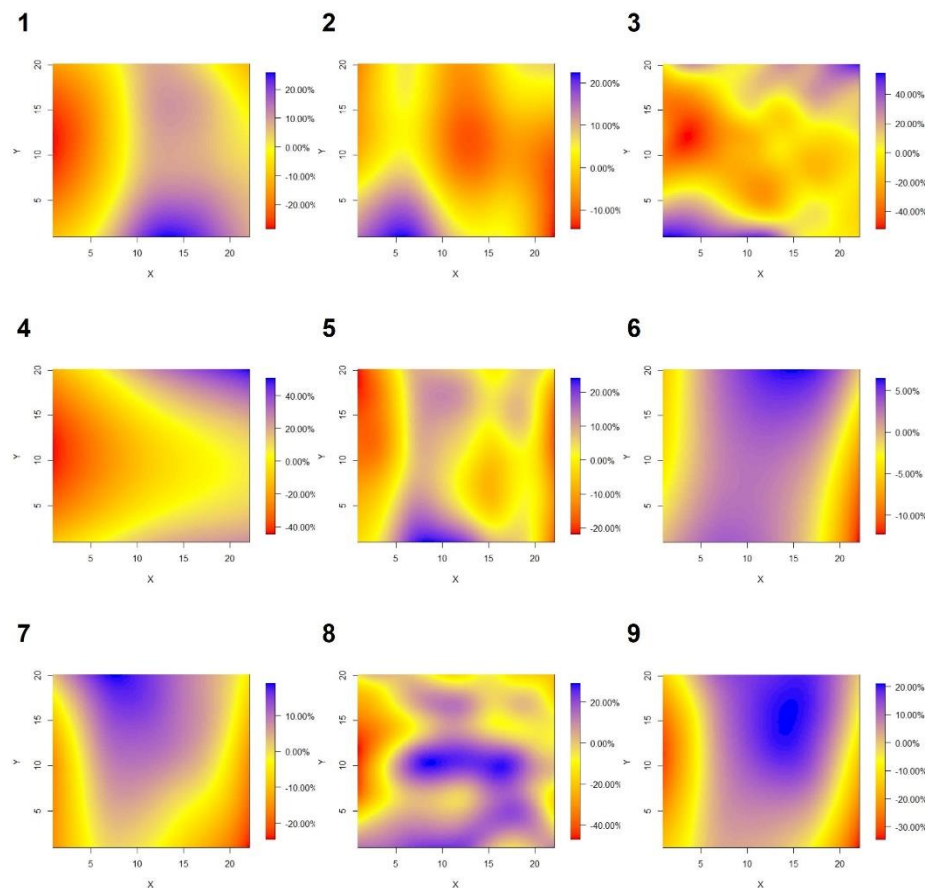
In plant breeding field trials, plots closer to each other commonly have spatial correlation (GILMOUR; CULLIS; VERBYLA, 1997; GEZAN; WHITE; HUBER, 2010). This correlation can lead to dependence between errors, consequently closer plots tend to have an increased rate of the type II error, whereas plots far apart the probability of type I error would be increased (VAN ES and VAN ES, 1993). Gilmour, Cullis and Verbyla (1997), De Faveri et al. (2015) and Andrade et al. (2020) demonstrated that no one spatial model will be always the better model for each trial and unique spatial trends and correlation might be present at each individual trial. In this study, different forms of spatial variation were found for each harvest and trial as demonstrated in Figure 1, where sometimes only spatial variation (AR1 x AR1) needed to be accounted in the model, and in some harvests the measure error (nugget) also needed to be incorporated in the model. The efficiency of AR1 models in controlling the spatial variations on field trials has been reported by several studies for several crops: wheat, sorghum, cotton, beet, barley, alfalfa, and potato (GILMOUR; CULLIS; VERBYLA, 1997; MÜLLER et al., 2010; LIU et al., 2015; DE FAVERI et al., 2015; VELAZCO et al., 2017; ANDRADE et al., 2020). In this study, the use of AR1 spatial models led to better efficiency in most of the harvests and trials, where 64% of the trials needed to account for local variation modeled by an AR1 process.

The inclusion of the nugget term in the model is used to represent an error inherent to the plot, this error is also called measurement error (DUTKOWSKI et al., 2006). De Faveri et al. (2015) found significant measurement error in the analysis of persistence for alfalfa. Sripathi et al. (2017), studying the efficiency of spatial variation in forage yield trials pointed out the importance of the measurement errors in forage, the authors reported that due to the larger plots frequently used in forage breeding trials, it causes the spatial variation most likely appears on a microscale at the plot level. Although larger plots control more efficiently the field heterogeneity, smaller plots can be more efficient than larger plots because of the ability of spatial analyses to capture variation on relative fine scale (CASLER, 2013). In this study, the

use of measurement error was needed in 71% of the trials in which AR1 modeling was necessary, so this type of error is very important on the forage breeding studied here.

Smith et al. (2007) and De Faveri et al. (2015), modeled spatial and temporal errors in a single stage analysis by using AR1 models. However, the three-way separable (harvest by column by row) structure assumed for spatial-temporal correlation may not always be appropriate (SMITH et al., 2007). De Faveri et al., (2015) reported that the model used by the authors assumes common spatial parameters over harvest times which may not always be the case. Our results indicate the spatial trends change across harvests and different magnitude of spatial correlation can be found and different types of AR1 model can be fitted (Figure 1 and Figure 6).

Figure 6 - Predicted spatial surface by the SpATS model for the trial T2 (Harvests 1 to 9).



Source: from the author (2022).

One way of considering the spatial trends at each harvest is using a weighting method in a two-stage analysis. Smith et al. (2001) proposed a weighting method based on the diagonal

elements of the inverse matrix of variance error difference between two BLUEs in wheat trials. The authors concluded that the weights in the combined analysis (second stage) not only accommodate variance heterogeneity between trials but also spatial variation and unequal replication within trials. In this study we evaluated the BLUEs and weights from each spatial model and non-spatial models. The SpATS model seemed to be very flexible in dealing with spatial variation that may occur over harvest, even when the trial did not need to account for AR1/nugget terms (trial T7 – Table 2) the model in which used the weights form the SpATS model had better BIC estimates.

When comparing the SpATS model to the best AR1 model, it was observed a tendency of AR1 models presenting smaller estimates of error variance components (Figure 2). This results, corroborate with the results showed by Velazco et al. (2017). Rodriguez-Álvarez et al. (2016) through simulation studies showed that autoregressive model tends to underestimate the residual variance term, whereas the SpATS model provides relatively accurate estimates of the random error variance. Small differences can be found on heritability estimates between the models evaluated. The differences in heritability estimation did not change much between the best AR1 model and SpATS model, however SpATS models heritabilities tended to be slightly bigger (Figure 3). These differences may occur because of the difficulties of the models in separating the genetic variance component from the spatial variation (VELAZCO et al., 2017). The changes in estimates of genetic variance indicate that by not modeling spatial trends, these values can be over or underestimated. Andrade et al. (2020) reported that SpATS model was equally efficient in controlling variation when compared with the best model based on first-order autoregressive process. The differences reported between the two approaches may be due to differences in the parametrization of the spatial variation (VELAZCO et al., 2017).

The SpATS model has some advantages over AR1 models. The process of adjusting the model is done directly, not requiring the numerous steps on which the AR1 models are based, such as graphical analysis and significance tests. Our results also showed when there is no spatial variation detected by the AR1 models, the SpATS model can yield similar results than the base model. Therefore, SpATS model can be used as a standard method to generate weights for a two-stage analysis in forage breeding trials.

5 CONCLUSION

This study shows that spatial trends can change over harvests time in forage breeding trials, and the variation presented in each harvest can be taken into account by using weighting methods in a two-stage analysis approach. The spatial analysis can lead to greater selection gains because of increased heritability as well as decreased experimental errors. Ignoring the spatial variations can lead to mistakes in selecting the best genotypes. The SpATS model can be used as a standard method to spatially correct the genotypes BLUEs and generates weights to be used in a multi-harvest forage breeding trials.

REFERENCES

- ANDRADE, Mario Henrique Murad Leite *et al.* Accounting for spatial trends to increase the selection efficiency in potato breeding. **Crop Science**, Madison, v. 60, n. 5, p. 2354-2372, 2020.
- CASLER, Michael D. Finding hidden treasure: a 28-year case study for optimizing experimental designs. **Communications in Biometry and Crop Science**, [S.l.], v. 8, n. 1, p. 23-38, 2013.
- CASLER, Michael D. Spatial variation affects precision of perennial cool-season forage grass trials. **Agronomy Journal**, Madison, v. 91, n. 1, p. 75-81, 1999.
- DE FAVERI, Joanne *et al.* Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. **Crop and Pasture Science**, Victoria, Australia, v. 66, n. 9, p. 947-962, 2015.
- DUTKOWSKI, Gregory W. *et al.* Spatial analysis enhances modelling of a wide variety of traits in forest genetic trials. **Canadian Journal of Forest Research**, Ottawa, v. 36, n. 7, p. 1851-1870, 2006.
- GEZAN, Salvador A.; WHITE, Timothy L.; HUBER, Dudley A. Accounting for spatial variability in breeding trials: a simulation study. **Agronomy Journal**, Madison, v. 102, n. 6, p. 1562-1571, 2010.
- GILMOUR, Arthur R.; CULLIS, Brian R.; VERBYLA, Arūnas P. Accounting for natural and extraneous variation in the analysis of field experiments. **Journal of Agricultural, Biological, and Environmental Statistics**, [S.l.], v. 2, n. 3, p. 269-293, 1997.
- GOGEL, Beverley; SMITH, Alison; CULLIS, Brian. Comparison of a one-and two-stage mixed model analysis of Australia's National Variety Trial Southern Region wheat data. **Euphytica**, Wageningen, v. 214, n. 2, p. 1-21, 2018.

- LIU, S. M. *et al.* Benefit of spatial analysis for furrow irrigated cotton breeding trials. **Euphytica**, Wageningen, v. 201, n. 2, p. 253-264, 2015.
- MÖHRING, J.; PIEPHO, H.-P. Comparison of weighting in two-stage analysis of plant breeding trials. **Crop Science**, Madison, v. 49, n. 6, p. 1977-1988, 2009.
- MÜLLER, Bettina U. *et al.* Comparison of spatial models for sugar beet and barley trials. **Crop Science**, Madison, v. 50, n. 3, p. 794-802, 2010.
- PAPADAKIS, J. S. Methode statistique pour des experiences sur champ. **Bulletin de l'Institute d'Amelioration des Plantes Salonique**, [S.l.], v. 23, p. 13-29, 1937.
- SMITH, Alison Barbara *et al.* Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. **Euphytica**, Wageningen, v. 157, n. 1, p. 253-266, 2007.
- SMITH, Alison; CULLIS, Brian; GILMOUR, Arthur. Applications: the analysis of crop variety evaluation data in Australia. **Australian & New Zealand Journal of Statistics**, Oxford, v. 43, n. 2, p. 129-145, 2001.
- SMITH, K. F.; KEARNEY, G. A. Improving the power of pasture cultivar trials to discriminate cultivars on the basis of differences in herbage yield. **Australian Journal of Agricultural Research**, Victoria, Australia, v. 53, n. 2, p. 191-199, 2002.
- SMITH, K. F.; SPANGENBERG, G. Forage breeding for changing environments and production systems: an overview. **Crop and Pasture Science**, Victoria, Australia, v. 65, n. 11, p. i-ii, 2014.
- SMITH, Kevin Francis; CASLER, M. D. Spatial analysis of forage grass trials across locations, years, and harvests. **Crop Science**, Madison, v. 44, n. 1, p. 56-62, 2004.
- SRIPATHI, Raghuv eer *et al.* Spatial variability effects on precision and power of forage yield estimation. **Crop Science**, Madison, v. 57, n. 3, p. 1383-1393, 2017.
- VAN ES, H. M.; VAN ES, C. L. Spatial nature of randomization and its effect on the outcome of field experiments. **Agronomy Journal**, Madison, v. 85, n. 2, p. 420-428, 1993.
- VELAZCO, Julio G. *et al.* Modelling spatial trends in sorghum breeding field trials using a two-dimensional P-spline mixed model. **Theoretical and Applied Genetics**, Berlin, v. 130, n. 7, p. 1375-1392, 2017.
- WELHAM, Sue J. *et al.* A comparison of analysis methods for late-stage variety evaluation trials. **Australian & New Zealand Journal of Statistics**, Oxford, v. 52, n. 2, p. 125-149, 2010.

MODELING GENOTYPE BY HARVEST INTERACTION BY RANDOM REGRESSION (RRM) AND FACTOR ANALYTIC (FAMM) MODELS

ABSTRACT

Genotype selection in perennial forage species is based on several repeated measures over time, seasons, and years. These repeated measurements in forage breeding trials generates longitudinal data set, in which must be properly analyzed giving useful interpretation in genotype selection process. In this study, we have presented methods of analysis for longitudinal DMY data generated from perennial forage breeding trials for ten trials and three different perennial species [*Medicago sativa* L. (T1), *Panicum maximum* (T2 and T3) and *Urochloa spp.* (T4 to T10)]. Two approaches were addressed in this paper the random regression models (RRM) and factor analytic mixed models (FAMM). We also proposed the estimation of adaptability based on the area under the curve and stability based on the curve coefficient of variation. Our results showed that FAMM is more a flexible model, since it approximates an unstructured (co)variance structure for the GxE effects, whereas RRM always approximated the (co)variance structure into an autoregressive pattern. However, RRM can offer more useful information about longitudinal data in forage breeding trials, where the breeder can select genotypes based on their seasonality by interpreting reaction norms. Therefore, we recommend the use of RRM for longitudinal traits in forage breeding trials.

1 INTRODUCTION

Genotype selection in perennial forage species is based on several repeated measures over time, seasons, and years. Therefore, the evaluation of multi-harvest forage breeding trials is time-consuming and expensive. In this context, the use of proper statistical methods that accurately predict the true genotypes' potential is crucial (SMITH and SPANGENBERG, 2014). As the genotypes experience different growth environmental conditions over time it is expected differential gene expression occurring over time. Therefore, in a multi-harvest trial a response variable can be treated as different traits in a multi-variate framework analysis (FALCONER and MACKAY, 1996). In this way, the genetic correlation between the traits is a measure of genotype by harvest (GxH) interaction (APIOLAZA; GARRICK, 2001; CROSSA; YANG; CORNELIUS, 2004; VAN EEUWIJK; BUSTOS-KORTS; MALOSETTI, 2016). Multi-harvests trials can be described as a special case of multi-environment harvest, in which the environments are the different times when the data were collected for in the same trial. The repeated measurement of the same trait over time generates a longitudinal data set, the sequential nature of measurements creates patterns of variation (HAND and CROWDER, 2017).

There are several models to deal with longitudinal data set, the most common and simple (co)variance structure is the autoregressive (AR1), where a single correlation parameter (ρ) is estimated. The model postulates a mechanism where the correlation between measurements j and k is $\rho^{|j-k|}$, where the genotypic value of the genotype is a function of genes acting in a given time plus genes acting on the new measurement (APIOLAZA and GARRICK, 2001). AR1 models is an appealing method for modeling (co)variance structure for phenotypes measured over time (APIOLAZA and GARRICK, 2001; YANG *et al.*, 2006; VANHATALO; LI; SILLA, 2019). However, AR1 model is recommended when the time between measurements is equally spaced, when the data set are unequally spaced nonlinear restriction should be imposed for parameters estimation (MCKENZIE, 2001). Irregular time series is frequently observed in perennial forage yield measurements over time, due to the yield seasonality. The yield seasonality is characterized by variation in forage availability and quality in response to climate conditions, in which do not allow the plants to have uniform growth rates during the whole year (REIS and ROSA, 2001). Therefore, under favorable climate conditions the plants grow faster, and the harvests are more frequent, whereas under unfavorable conditions harvests are less frequent. Thus, the time series for forage yield measures are naturally irregular.

Random regression models (RRM) were introduced by Henderson (1982) and Laird and Ware (1982). Schaeffer and Dekkers (1994) suggested their use in dairy cattle breeding for the analysis of test day production records. Since then, several papers were published using RRM to predict growth in sheep (LEWIS and BROTHERSTONE, 2002), body weight in beef cattle (ARANGO; CUNDIFF; VAN VLECK, 2004), body weight in swine (HUISMAN; VEERKAMP; VAN ARENDONK, 2002), egg production in layer (WOLC *et al.*, 2011). Recently random regression models have been applied to longitudinal data from perennial forage breeding trials for dry mater yield (DMY) in elephantgrass (*Pennisetum purpureum* Schmach.) (ROCHA *et al.*, 2018). The use of RRM has also been increasing for annual crops with the advent of high-throughput phenotyping, which generates longitudinal data set (SUN *et al.*, 2017, CAMPBELL *et al.*, 2018; MOREIRA *et al.*, 2021). RRM can deal with longitudinal data very well (SCHAEFFER, 2004) because it captures the change of a trait continuously over the trajectory with few parameters by covariance functions (eg. orthogonal polynomials and splines) (KIRKPATRICK; LOFSVOLD; BULMER, 1990; MEYER, 1998). Kirkpatrick, Lofsvold and Bulmer (1990) reported that RRM can deal with unequally time spaced measurement, relating that RRM should be the adequate model under this condition. Furthermore, there is a possibility to include environmental-dependent covariate in RRM e.g.,

temperature and humidity to study the genotypes' response to abiotic stress (BRÜGEMANN et al., 2011; BOHLOULI et al., 2013; MBUTHIA; MAYER; REINSCH, 2021).

Another way to model longitudinal data set is applying a multi-trait model, where (co)variance structure for the GxH interaction is modeled by an unstructured (co)variance structure (SMITH et al., 2007). This model is the most general form for the genetic (co)variance matrix and involves $t(t+1)/2$ unknown parameters, namely all the genotypes' variances and covariances (SMITH et al., 2007). Although this model is very informative, where all variances and covariances will be estimated, there may be computational difficulties and convergence problems associated with the estimation of such a structure, particularly if the number of harvests is large and/or the number of genotypes is small (SMITH; CULLIS; THOMPSON, 2001; THOMPSON et al., 2003; KELLY et al., 2007). In order to overcome these difficulties, Smith, Cullis and Thompson (2001) proposed the factor analytic mixed models (FAMM). This model with sufficient multiplicative terms (loadings and scores) has been found to provide a good parsimonious approximation to the unstructured form and is generally more computationally robust (THOMPSON et al., 2003). The genotypes adaptability and stability can be easily interpreted by latent regression, where the genotypes genotypic value for each harvest are regressed on environmental covariates (the loadings) but the covariates are estimated from the data rather than externally (SMITH; CULLIS; THOMPSON, 2005; CULLIS et al., 2014; SMITH and CULLIS, 2018). This has the advantage that regressions usually account for large proportion of GxH interaction (SMITH; CULLIS; THOMPSON, 2005). However, the environmental covariates are data-dependent so cannot be used for harvest prediction.

In this study, we investigated the use of RRM and FAMM for longitudinal data of dry matter yield in ten forage breeding trials for three different species [*Medicago sativa* L. (T1), *Panicum maximum* (T2 and T3) and *Urochloa spp.* (T4 to T10)] for genotype selection and genetic interpretation.

2 MATERIAL AND METHODS

2.1 Data set

In this study, we used data from three forage species conducted from 2015 to 2020 in three different locations by using different experimental designs (augmented row column design – ARCD, alpha-lattice design – ALD, and randomized complete block design – RCBD) for total dry matter yield (DMY, kg.ha⁻¹) (Table 1). The number of genotypes evaluated varied from 8 (T4 and T5) to 182 (T1), the data set is composed by early breeding trials (T1, T2, T3, T8 and T10) in which many genotypes are tested and advanced breeding trials (T4, T5, T6, T7 and T9) (Table 1). The number of harvests in each trial varied from six (T7, T8 and T10) to 16 (T4 and T5), the total number.

Table 1 - Description of experimental layout for the three forage species evaluated from 2015 to 2020.

Trial	Specie	Year	Location	Design	Genotypes	Harvests	Columns	Rows	Plots
T1	<i>Medicago sativa</i>	2018 - 2019	Citra - FL - USA	ARCD	182	11	32	14	405
T2	<i>Panicum maximum</i>	2016 - 2019	Campo Grande - MS - BR	ALD	110	9	22	20	330
T3	<i>Panicum maximum</i>	2016 - 2019	Campo Grande - MS - BR	ALD	110	9	22	20	330
T4	<i>Urochloa brizanta</i>	2009 - 2011	Campo Grande - MS - BR	RCBD	8	16	-	-	32
T5	<i>Urochloa brizanta</i>	2009 - 2011	Terenos- MS - BR	RCBD	8	16	-	-	32
T6	<i>Urochloa decumbens</i>	2018 - 2019	Campo Grande - MS - BR	RCBD	9	13	9	4	36
T7	<i>Urochloa decumbens</i>	2018 - 2019	Brasilia - DF - BR	RCBD	12	6	12	4	48
T8	<i>Urochloa (Inter)</i>	2015 - 2016	Campo Grande - MS - BR	RCBD	99	6	8	50	396
T9	<i>Urochloa (Inter)</i>	2019 - 2020	Campo Grande - MS - BR	RCBD	15	10	6	11	60
T10	<i>Urochloa decumbens</i>	2015 - 2016	Campo Grande - MS - BR	RCBD	36	6	3	50	144

Source: from the author (2022).

2.1 Statistical analysis

In this study, the analyzes were performed based on a two-stage analysis by using the weighting method proposed by Smith, Cullis and Thompson (2001). In a two-stage analysis, genotypes' best linear unbiased estimates (BLUES) of individual trials in stage one were combined in a weighted across trials mixed model analysis in stage two, where the weights provide a measure of relative uncertainty of the estimated genotypes' BLUES for each trial (SMITH 2001; MÖHRING and PIEPHO, 2009; GOGEL et al., 2018). All the analyses were done by using ASRM-R (BUTLER, 2021) and SpATS (RODRIGUEZ-ALVAREZ et al., 2018) R packages, and the data summarization through graphs were done by ggplot2 (WICKHAM et al., 2016) R package. The scripts for the analysis can be found at github (<https://github.com/claudiocff/RRM-and-FAMM-asreml-two-step>).

2.1.1 First stage: estimating the genotypes' BLUES and weights accounting for spatial field trends

We obtained the BLUES and weights of the genotypes at each harvest and trials, using the SpATS R package (VELAZCO et al., 2017) in a mixed model framework:

$$y = X\beta + X_s\beta_s + Z_s u_s + Z_c u_c + Z_r u_r + e \quad (1);$$

$$y = X\beta + X_s\beta_s + Z_s u_s + Z_b u_b + e \quad (2);$$

where, y is the vector from measured DMV from each plot; β is the vector of the fixed effects of genotypes in all trials, and replication in the alpha-lattice design (T2 and T3-Table1); u_r and u_c are the vectors of random effects of rows and columns, respectively in the augmented row column design (T1 - Table 1), where $u_r \sim N(0, I_r \sigma_r^2)$ and $u_c \sim N(0, I_c \sigma_c^2)$; u_b is the vector of random effects of block effects in a randomized complete block design (T4, T5, T6, T7, T8, T9 and T10 - Table 1), or the vector of the random effects of the blocks inside replication in the alpha-lattice design (T2 and T3 - Table 1), in which $u_b \sim N(0, I_b \sigma_b^2)$; β_s is the vector of fixed effects of the smooth spatial surface (unpenalized); u_s is the vector of random effects of the penalized part of the smooth surface (penalized). The fixed term ($X_s\beta_s$ - unpenalized) and

random term ($Z_s u_s$ - penalized) describe the mixed model expression of the smooth spatial surface ($f(r, c) = X_s \beta_s + Z_s u_s$), where the random spatial vector u_s has (co)variance matrix S . The SpATS model uses the P-spline ANOVA (PS-ANOVA) (LEE; DURBÁN; EILERS, 2013) to describe the 2D-splines in the mixed model framework. The X_s , Z_s incidence matrices, and the (co)variance matrix S are described by Lee, Durbán and Eilers (2013) and Rodríguez-Álvarez et al. (2018). The PS-ANOVA parametrization can be decomposed as a linear sum of the univariate and bivariate smooth functions (VELAZCO et al., 2017); e is the vector of random errors, $e \sim N(0, I_e \sigma_e^2)$; σ_r^2 , σ_c^2 , σ_b^2 and σ_e^2 are the variance components associated to the random effects of rows, columns, blocks and errors. X , Z_r , Z_c and Z_b are incidence matrices for the fixed effects, the random effects of rows, columns and blocks, respectively. I_r , I_c , I_b , and I_e are identity matrices.

2.1.2 Second stage: modeling the genotype by harvest interaction

2.1.2.1 Random regression models (RRM)

For the statistical model described below, the BLUEs obtained for each genotype in each trial were regressed on a time gradient (days), where the first harvest of each trial were considered as day zero and the other harvest times are days after the first harvest. Therefore, random coefficients are computed for each genotype to describe the ‘genopes’ DMY trajectory over harvest time. Polynomial functions were used to model the longitudinal dimensions by using orthogonal Legendre polynomials (KIRKPATRICK; LOFSVOLD; BULMER, 1990). The orthogonal polynomials were obtained by rescaling the time points from -1 to 1 using the expression:

$$t_i = -1 + 2 \left(\frac{t_i - t_{\min}}{t_{\max} - t_{\min}} \right) \quad (3)$$

the Legendre polynomials are denoted by $P_n(t)$. Defining $P_0(t) = 1$, the polynomial $n+1$ is described by the recursive equation:

$$P_{n+1}(t) = \frac{1}{n+1} [(2n+1)tP_n(t) - nP_{n-1}(t)] \quad (4)$$

on the normalized form the Legendre polynomial can be described:

$$\phi_n(t) = \left(\frac{2n+1}{2}\right)^{0.5} P_n(t) \quad (5)$$

therefore, for a polynomial of order two we can obtain the following equations:

$$\begin{aligned} \phi_0(t) &= \left(\frac{1}{2}\right)^{0.5} P_0(t) = 0.7071 \\ \phi_1(t) &= \left(\frac{3}{2}\right)^{0.5} P_1(t) = 1.22467t \\ \phi_2(t) &= \left(\frac{5}{2}\right)^{0.5} \left(\frac{3}{2}t^2 - \frac{1}{2}\right) = -0.7906 + 2.3717t^2 \end{aligned} \quad (6)$$

considering $m = 2$ as the order of the covariance function to be used, the Legendre coefficient matrix Λ will have dimensions of $(m+1) \times (m+1)$ can be defined as:

$$\Lambda = \begin{bmatrix} 0.7071 & 0 & -0.7906 \\ 0 & 1.2247 & 0 \\ 0 & 0 & 2.3717 \end{bmatrix} \quad (7)$$

considering four different time points and an order two polynomial, the incidence matrix of time points (M) can be defined as:

$$M = \begin{bmatrix} 1 & t_0 & t_0^2 \\ 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \end{bmatrix} \quad (8)$$

where t_i is the time point scaled by the equation (3).

Finally the Legendre polynomials can be computed as $\Phi = M\Lambda$, where Φ is a matrix containing the normalized polynomials for harvest time; M store polynomials of standardized harvest times; Λ is the matrix of Legendre polynomial coefficients of order $m+1$, where m is the degree of fit (SHAEFFER, 2016). The random regression model can be defined as:

$$y = \Phi_1\beta + \Phi_2u_g + e \quad (9)$$

Where y is the vector of BLUEs estimated by the models (1) or (2); β is the vector of the fixed regression coefficients; u_g is the vector of random regression coefficients of genotypes, in which $u_g \sim N(0, K_g \otimes I_g \sigma_g^2)$; e is the vector of the errors, where $e \sim N(0, R\sigma_e^2)$. K_g is an

unstructured (co)variance matrix associated to the random regression coefficients. The matrix K_g can be described as:

$$K_g = \begin{bmatrix} \sigma_{g_0}^2 & \sigma_{g_0g_1} & \cdots & \sigma_{g_0g_m} \\ & \sigma_{g_1}^2 & \cdots & \sigma_{g_1g_m} \\ & & \ddots & \vdots \\ Sym. & & & \sigma_{g_m}^2 \end{bmatrix} \quad (10)$$

where, $\sigma_{g_m}^2$ is the variance component associated to the coefficient of order m ; $\sigma_{g_n g_m}$ is the covariance between the coefficient of order m and n .

R is the variance matrix of the errors, where $R = \text{diag}(R_1 \cdots R_j)$, in which j is the harvest. In practice, R_j are unknown and replaced by an estimate \hat{R}_j from each harvest. It is sometimes not feasible to store and use the full matrix \hat{R}_j from each harvest, and so a vector of approximate weights is required. We used the weights proposed by Smith, Cullis and Gilmour (2001), where the weights are based on the diagonal elements of \hat{R}_j^{-1} designated as Π , in which:

$$\Pi = \text{diag}(\pi_1^T \cdots \pi_j^T) \quad (11)$$

π_j consists of the diagonal elements of \hat{R}_j^{-1} . This simple approximation reflects the uncertainty in each estimated BLUE, accounting for within-trial heterogeneity, differing replication and spatial trends.

Based on Kirkpatrick, Lofsvold and Bulmer (1990), the following estimator was used to obtain the genetic variance and covariance components across harvest times ($\hat{\Sigma}_g$) on original scale:

$$\hat{\Sigma}_g = \Phi_2 K_g \Phi_2^T \quad (12)$$

where, Φ_2 is the incidence matrix of the Legendre polynomials associated to the random effects of genotypes; K_g is the (co)variance matrix associated to the random genotypes' coefficients, defined in (10).

The genotypic values for each genotype across harvest time can be estimated by the equation:

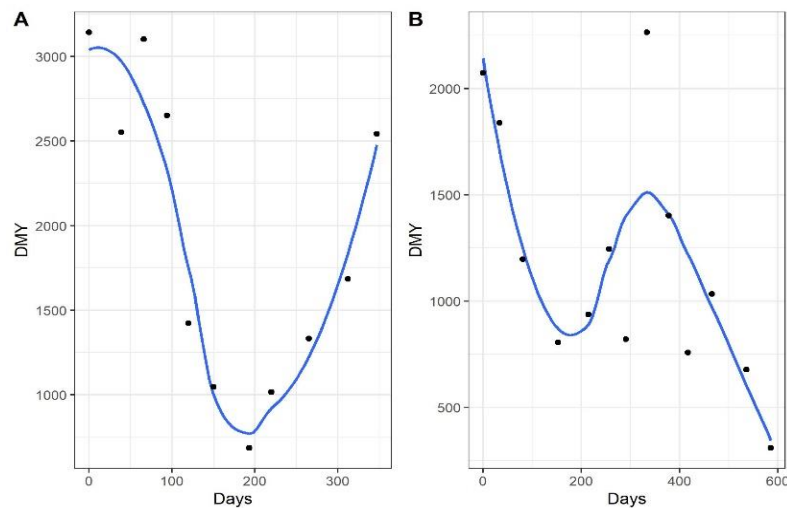
$$G_V = J\beta^T \Phi_1^T + U_g \Phi_2^T \quad (13)$$

where G_V is a $i \times j$ matrix of the genotypic values on the original scale, where i is the number of genotypes and j the number of harvest time points; J is a column vector of 1's size equal the

number of genotypes (i); β^T is the transposed vector of the fixed regression coefficients of size $1 \times (d+1)$, in which d is the degree of the polynomial fitted for the fixed regression; Φ_1^T is the transposed incidence matrix of the Legendre polynomials for each harvest time for the fixed regression with size $(d+1) \times j$; U_g is the genotypes's random coefficients matrix, size $i \times (m+1)$; Φ_2^T is the transposed incidence matrix of the Legendre polynomials for each harvest time for the random regression with size $(d+1) \times j$.

The polynomial function for the fixed regression was defined graphically by using a loess function, where the function order was determined by the number of curves (c) +1 in the mean DMY trajectory across harvest time. For example, for the trials T1 and T2 a polynomial of degree three were fitted, since two curves were observed on the mean DMY trajectory. The random polynomial regression degree was determined by the Bayesian information criteria (BIC).

Figure 1 - Mean DMY trajectory over time for trials T1 (Alfalfa - B) and T6 (*Urochloa decumbens*).



Source: from the author (2022).

2.1.2.1.1 Heritability, broad adaptability and stability for RRM

The heritability over harvest times was estimated by the expression:

$$H_j^2 = \frac{\text{diag}(\Sigma_g)}{\text{diag}(\Sigma_g) + \frac{\bar{\sigma}_e^2}{r}} \quad (14)$$

where H_j^2 is the broad sense heritability estimated at each harvest time; Σ_g is the genetic variance-covariance matrix estimated by the equation (12); $\bar{\sigma}_e^2$ is the mean error variance component across harvests; r is the number of replication in the trial.

The broad adaptability for each genotype was estimated based on the area under the DMY trajectory curve, in which reflects the total DMY accumulation over harvest time:

$$A_i = \int_{-1}^1 \underbrace{b_0 + b_1t + b_2t^2 + \dots + b_d t^d}_{\text{Fixed}} + \underbrace{g_{i0} + g_{i1}t + g_{i2}t^2 + \dots + g_{im}t^m}_{\text{Random}} d(t) \quad (15)$$

where A_i is the area under the trajectory curve of the genotype i ; b_0 is the fixed regression intercept; g_{i0} is the random regression intercept of the genotype i ; t is the harvest time point; d and m are the polynomial fitted degree for the fixed and random regression, respectively; b_d is the fixed regression coefficient of degree d ; g_{im} is the random regression coefficient of degree m for genotype i .

The genotypes' stability was calculated based on the trajectory curve's coefficient of variation (CV_c), in which reflect the genotypes' Type I stability, where the genotype is stable if present small variance between environments, also called biological stability (LIN et al., 1986):

$$CV_{c_i} = \frac{\sigma_{c_i}}{b_0 + g_{i0}} \quad (16)$$

where, σ_{c_i} is the standard trajectory curve deviation for genotype i ; $b_0 + g_{i0}$ is the overall performance for genotype i .

2.1.2.1.2 Genetic interpretation on random regression models

One of the advantages of random regression models is the use of eigenfunction (Ψ_k) of the genetic coefficient (co)variance matrix (10), in which can provide genetic insights about the studied trait, bases on Kirkpatrick, Lofsvold and Bulmer (1990):

$$\Psi_k = \sum_{m=0}^M (v_{\Psi_k})_m \Phi_m \quad (17)$$

where $(v_{\Psi_k})_m$ is the m^{th} element of the k^{th} eigenvector of K_g , and Φ_m is the normalized value of the m^{th} Legendre polynomial.

2.1.2.2 Factor analytic mixed model (FAMM)

The factor analytic mixed model can be described as a mixed model equation in the matrix notation:

$$y = X\beta + Zu_g + e \quad (18)$$

where: y is the vector of the genotypes' BLUES estimated by the model (1) or (2); β is the vector of fixed harvests effects added to the overall mean; g is the vector of the genotype (random) effects within sites within harvests, with $u_g \sim N(0, G \otimes I\sigma_g^2)$; e is the vector of random errors, with $e \sim N(0, R \otimes I\sigma_e^2)$; X and Z are incidence matrices for β and u_g , respectively; G is the variance and covariance matrix for the genotype effect within harvest-sites combinations; R is the variance and covariance matrix for the error, the R matrix is the same as used for the random regression model in the equation (11); I is an identity matrix.

To account for the covariances of the genetic effects between harvest in terms of small number, k , of (unknown) common factors, we used a factor analytic model (FA $_k$), proposed by Smith, Cullis and Thompson (2001).

$$G = (\Lambda^* \Lambda^{*T}) + \Psi \quad (19)$$

where Λ^* is the rotated loadings matrix ($p \times k$), in which p is the number of site-harvest combinations and k is the number of factors considered in the model. The rotation matrix (V_a) for Λ was obtained by the eigenvectors of $V_a = \Lambda^T \Lambda$. We obtained rotated estimated loadings as: $\Lambda^* = c\Lambda V_a$; where c is a constant is either 1 or -1. The sign is chosen to ensure the majority of first rotated loadings are positive than negative. Ψ is a diagonal matrix ($p \times p$) of the specific variances for environment p .

The percentage of the genetic variance (V_k) explained by the k factors from FA structure was calculated using:

$$v_k = 100 \times \text{tr}(\Lambda^* \Lambda^{*T}) / \text{tr}(\Lambda^* \Lambda^{*T} + \Psi) \quad (20)$$

The rotated scores of the genotypes (f_a^*) can be obtained as:

$$f_a^* = (cV_a^T \otimes I)f_a \quad (21)$$

where f_a is the vector of genotype scores in the k factors.

2.1.2.2.1 Heritability, adaptability and stability for FAMM

The broad sense heritability (H^2) for each harvest was estimated by the following equation:

$$H_j^2 = \frac{\text{diag}(G)}{\text{diag}(G) + \frac{\bar{\sigma}_e^2}{r}} \quad (22)$$

G is the genetic variance-covariance matrix; $\bar{\sigma}_e^2$ is the mean error variance across harvests; r is the number of replications for each trial.

The adaptability and stability were estimated by the latent regression as proposed by Cullis et al. (2014), where the genotypes' BLUPs for each harvest are regressed over the rotated environmental loadings estimated for the first factor. In this approach the broad adaptability is the regression intercept, and the stability can be measured by the determination coefficient R^2 . Furthermore, the genotype plasticity can be measured by the slope coefficient of the latent regression:

$$BLUP_{ij} = \mu_i + b_i \lambda_{rj}^* + \varepsilon_{ij}$$

where $BLUP_{ij}$ is the BLUP for genotype i in the j^{th} harvest; μ_i is the overall genotype performance across harvests (broad adaptability); b_i is the genotype plasticity; λ_{rj}^* is the rotated loading the j^{th} harvest; ε_{ij} is the regression errors.

3 RESULTS

3.1 Overall description of RRM and FAMM across trials

The degree of the polynomial fitted for the fixed part of the RRM varied from 2 (T10) to 5 (T2 and T3), there was no clear pattern between the number of harvests evaluated in each trial and higher order degrees (Table 1 and Table 2). The polynomial order for the fixed part of the model was determined by the number of contrasting seasons evaluated in each trial, these

contrasting seasons form the pits and peaks that needed to be model by the fixed polynomial. For the random part for the RRM there was a predominance of lower order polynomial, most of the trials the genotype by harvest interaction (GxH) a first order polynomial was needed, the second order polynomial was fitted for trials T4 and T8, and a third order polynomial was fitted for trial T1 (Table 2). The number of factors to be retained by the FAMM did not present a correlation between the number of harvests in each trial (Table 1 and Table 2).

The mean heritability across harvest varied from 0.16 (T6 – RRM) to 0.76 (T7 – FAMM) (Table 2). There was no big change in the mean heritability through the harvest between the two models, except for trials T6, T7 and T9. In general, FAMM tended to yield higher heritabilities. However, more variation for heritability can be observed for FAMM (Table 2). Otherwise, RRM tend to estimate higher genetic correlation between harvest then FAMM, and the variation for genetic correlation was also smaller for RRM (Table 2). These results can be explained by the different parametrization between the two models, where RRM uses a covariance function to estimate the variance and covariance components over harvest times, therefore these parameters are estimated in a smoother way in which RRM tend to approximate the genetic correlations in an autoregressive pattern (Figure 1). The FAMM approximates an unstructured genetic variance-covariance matrix, and more complex pattern of genetic correlation can be observed (Figure 1).

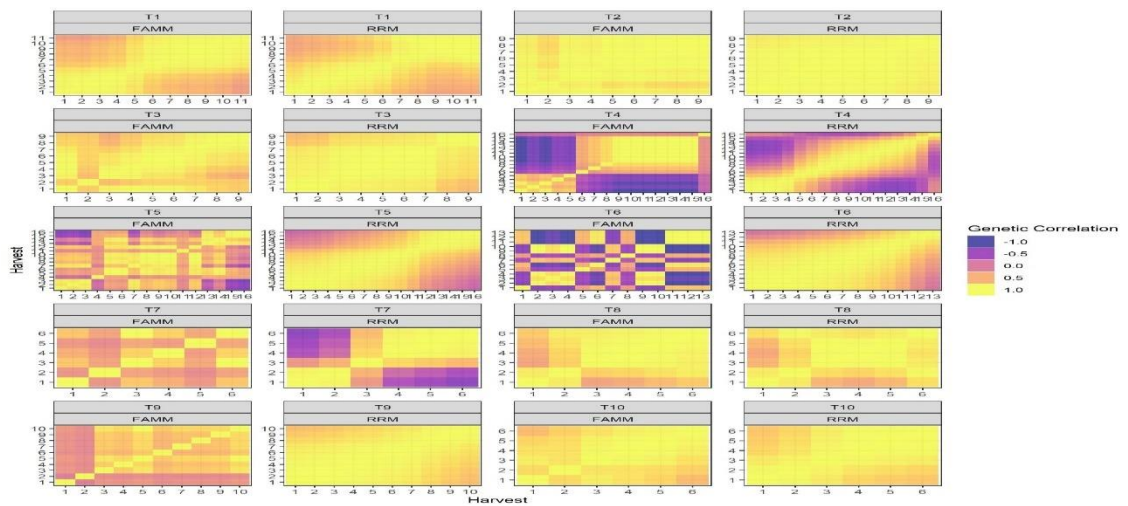
Table 2 - Fitted polynomial order for fixed and random regression for RRM, number of factors retained for each trial by FAMM, mean heritabilities and genetic correlation between harvest for each trial estimated by RRM and FAMM.

Trial	RRM				FAMM		
	Degree (F)	Degree (R)	H^2	ρ_g	Factors	H^2	ρ_g
T1	3	3	0.36 (0.11)	0.78 (0.19)	2	0.39 (0.14)	0.79 (0.21)
T2	5	1	0.63 (0.02)	0.98 (0.02)	2	0.65 (0.19)	0.95 (0.05)
T3	5	1	0.51 (0.01)	0.90 (0.10)	2	0.59 (0.19)	0.82 (0.14)
T4	3	2	0.34 (0.20)	0.26 (0.54)	1	0.44 (0.34)	0.06 (0.72)
T5	3	1	0.52 (0.10)	0.67 (0.33)	2	0.61 (0.25)	0.45 (0.43)
T6	3	1	0.16 (0.08)	0.75 (0.28)	1	0.42 (0.24)	-0.06 (0.76)
T7	4	1	0.42 (0.18)	0.29 (0.61)	1	0.76 (0.34)	0.48 (0.23)
T8	4	2	0.69 (0.02)	0.82 (0.18)	2	0.69 (0.06)	0.81 (0.20)
T9	3	1	0.31 (0.03)	0.89 (0.11)	1	0.53 (0.25)	0.51 (0.26)
T10	2	1	0.55 (0.07)	0.88 (0.11)	2	0.57 (0.10)	0.85 (0.11)

Source: from the author (2022).

H^2 is the mean heritability through harvest; ρ_g mean genetic correlation between harvest. The values between parenthesis represent the standard deviation of heritabilities and genetic correlations through harvests.

Figure 2- Genetic correlations between harvests for each trial, estimated by FAMM and RRM.



Source: from the author (2022).

In general, the FAMM yielded similar (co)variance structures as RMM, except for trials T5, T6 and T7 (Figure 1), this differences in the (co)variance structures reflected on the lack of correlation between adaptability parameters (area under the curve [A] – RMM and overall performance [OP] – FAMM), mainly for trial T6 ($\rho_{A-OP} = 0.22$, Table 3). However, all correlation between OP and A were higher ($\rho_{A-OP} > 0.74$, Table 3). It is worthy to note that when genetic correlations are higher (above 0.8, Table 2), the stability and adaptability parameters had higher correlations for both models, this fact can be observed for trials T2, T3 T8 and T9 (Table 2 and Table3). Therefore, the genotype selection can be done based only on adaptability. When complex interaction between harvest, i.e. there was a lack of genetic correlation between harvests, the correlation between adaptability and stability (ρ_{A-CV_c} , ρ_{A-R^2} , ρ_{OP-R^2} , and ρ_{OP-CV_c}) tended to be lower (Table 3), in these cases genotype selection should be done based on adaptability and stability.

Table 3 - Correlation between broad adaptability and stability parameter estimated for RRM (Area under the curve – A and curve coefficient of variation – CV_c) and for FAMM (overall performance – OP and latent regression R^2).

Trial	ρ_{A-CV_c}	ρ_{A-OP}	ρ_{A-R^2}	ρ_{CV_c-OP}	$\rho_{CV_c-R^2}$	ρ_{OP-R^2}
T1	-0.55	0.99	0.32	-0.5	0.47	0.38
T2	-0.97	0.99	0.94	-0.97	-0.96	0.94
T3	-0.98	0.99	0.97	-0.98	-0.97	0.97
T4	0.21	0.74	-0.37	0.19	-0.78	0.06
T5	-0.51	0.99	0.53	-0.47	0.42	0.58
T6	0.01	0.22	-0.19	-0.46	-0.49	0.85
T7	-0.26	0.82	0.24	-0.46	-0.78	0.64
T8	-0.92	0.91	0.92	-0.71	-0.89	0.73
T9	-0.85	0.95	0.79	-0.76	-0.75	0.85
T10	-0.61	0.99	0.12	-0.61	0.59	0.13
Mean	-0.54	0.86	0.43	-0.57	-0.41	0.61

Source: from the author (2022).

In the following sections of this study, we will present the models interpretation for genotype selection. The trials to be discussed later were chosen by presenting higher complex GxH, i.e., lower genetic correlation between harvests. Therefore, trials T1, T5 and T7 were chosen for interpretation (Table 2).

3.2 Genotype selection and GxH interpretation for RRM and FAMM

3.2.1 Alfalfa (*Medicago sativa* L.) breeding trial – T1

In this trial, 182 genotypes were evaluated for DMY through 11 harvests. The degree of Legendre polynomial fitted for this trial was three for both fixed and random part of the model, whereas for FAMM two factors were needed to model the GxH (Table 2). Twelve parameters were estimated for RRM and 35 were estimated by FAMM (data not shown).

3.2.1.1 Variance components, heritability, and genetic behavior – RRM

The polynomial genetic variances varied from 2,930 to 343,654 for g_3 and g_0 , respectively. The genotypes' intercept (g_0) retained most of the genetic variance, explaining 74% of the genetic variance, and the components related to the genotype's curve shape (g_1 , g_2 and g_3) accounted to 26% of the genetic variance (Table 4). The most important correlation

between the polynomial coefficients and the easiest to interpret is the correlation between g_0 and g_1 , it shows the genetic variance behavior over time. In this trial, the correlation between g_0 and g_1 was negative (-0.22, Table 4), indicating lower genetic variance can be observed over time.

Table 4 - Summary of genetic and non-genetic parameters estimated by RRM for trial T1.

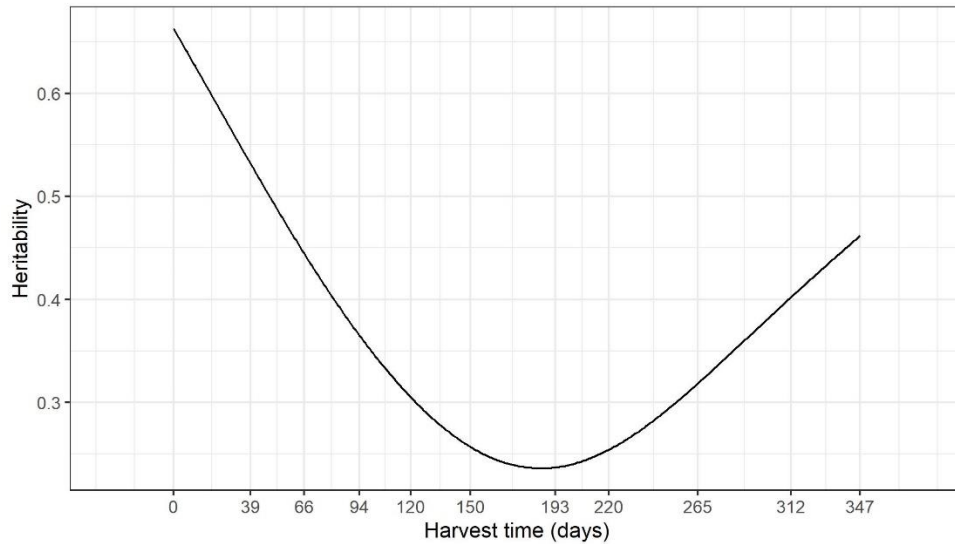
	g_0	g_1	g_2	g_3
g_0	343,654	-0.22	0.76	0.03
g_1		101,118	-0.61	-0.07
g_2			19,627	-0.63
g_3				2,930
Importance (%)	74	21	4	1
$\bar{\sigma}_e^2$				365,436

Source: from the author (2022).

g_0 , g_1 , g_2 and g_3 are the random regression intercept and first, second and third order coefficient, respectively. $\bar{\sigma}_e^2$ is the mean error variance across harvests. The diagonal elements of the table are the genetic variances associated with the intercept and polynomial coefficients ($\sigma_{g_0}^2, \sigma_{g_1}^2, \sigma_{g_2}^2, \sigma_{g_3}^2$); the off diagonal of the table are the correlations between intercept and polynomial coefficients ($\rho_{g_0g_1}, \rho_{g_0g_2}, \rho_{g_0g_3}, \rho_{g_1g_2}, \rho_{g_1g_3}, \rho_{g_2g_3}$).

One advantage of RRM is the heritability estimation in function of time, as well as pointed out above by the negative estimate of $\rho_{g_0g_1}$ the heritability tended to decrease over time as the genetic variance also decreased. The lower heritabilities ($H^2 < 0.30$) can be observed between in harvests done between 120 days and 220 days after the first harvest on they zero (Figure 3). It is worth noting, the harvests with lower heritabilities consist in the period from late summer to the late fall, where alfalfa plants are dormant (Figure 3). Therefore, selecting alfalfa non-dormant genotypes can be more challenging.

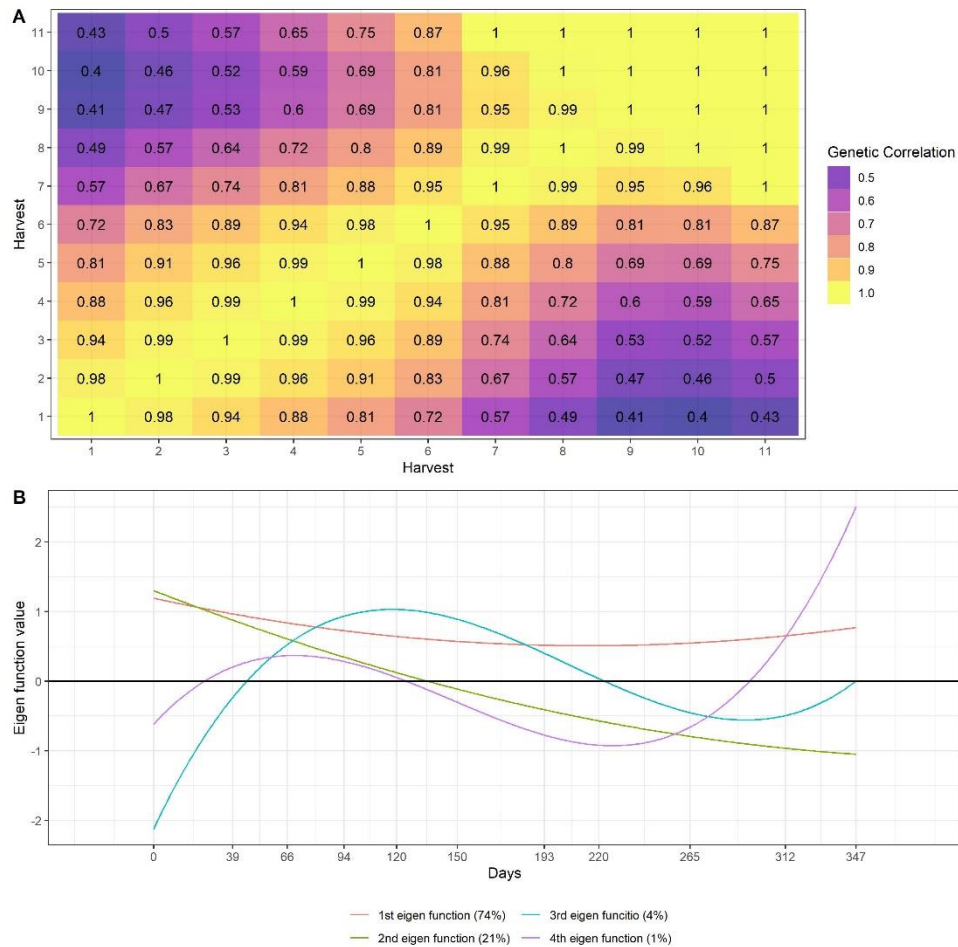
Figure 3- Heritability estimates over harvests time for trial T1.



Source: from the author (2022).

The genetic correlation varied from 0.43 to 1.00 (Figure 3A). The genetic correlation between harvests followed an autoregressive pattern, where harvest closer to each other tend to have higher genetic correlation and those harvests far apart from each other had lower genetic correlation (Figure 4A). The eigen functions can be used in RRM to infer about gene expression over time (Figure 4B). Where the first eigenfunction had a nearly constant behavior and explained 74% of the genetic variation, this variation represents a common gene pool that is being expressed over time, and explain the simple GxH interaction, since non differential expression was observed (Figure 4B). The second eigenfunction represents another gene pool, in which explained 21% of the genetic variation that shows differences in gene expression under different environment conditions, explaining most of the complex GxH interaction (Figure 4B). The third and fourth eigenfunctions explained only 4 and 1% of the genetic variation, also representing the complex GxH interaction, where differences on gene expression can be observed over time (Figure 4B).

Figure 4- Genetic correlation between harvests estimated by RRM for trial T1 (A); Estimation of the four eigenfunctions for trial T1 (B).



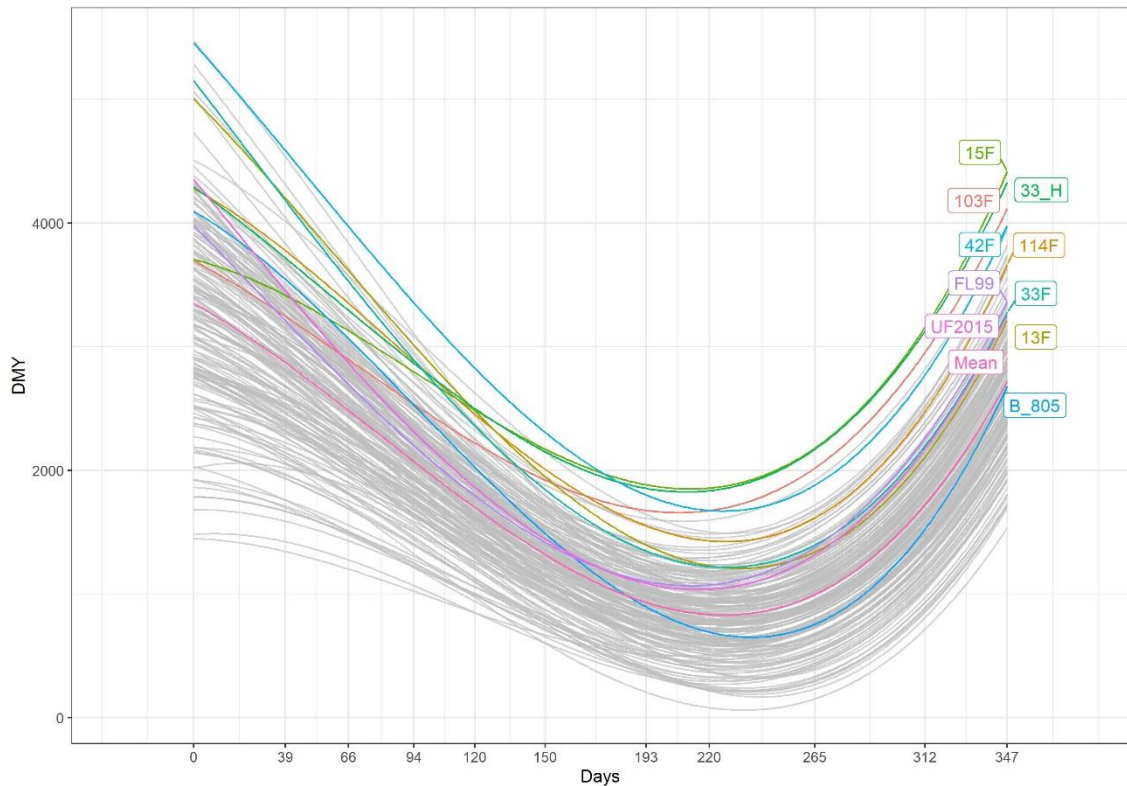
Source: from the author (2022).

3.2.1.2 Genotypes' adaptability, stability and yield trajectory over time – RRM

One of forage breeders' interest is evaluate the genotypes behavior over time, for this RRM can be a very useful tool where genotypes' reaction norms can be plotted (Figure 5). There was great variability on the genotypes' reaction norm, and the main changes on ranking occurred between harvests realized from 39 to 220 days. On Figure 5, we highlighted seven genotypes in which presented higher area under the curve (A), i.e., higher broad adaptability, the checks (UF2015, FL99 and B_805), and the mean yiled trajectory curve. Among these genotypes, we can highlight the genotype 15F, 103F, 33_H and 42F, these genotypes hap better performance under the fall dormancy period (from day 150 to the day 265) and could be selected

as non-dormancy alfalfa genotypes (Figure 5). All seven genotypes with higher A performed better than the checks for most of the harvest evaluated (Figure 5).

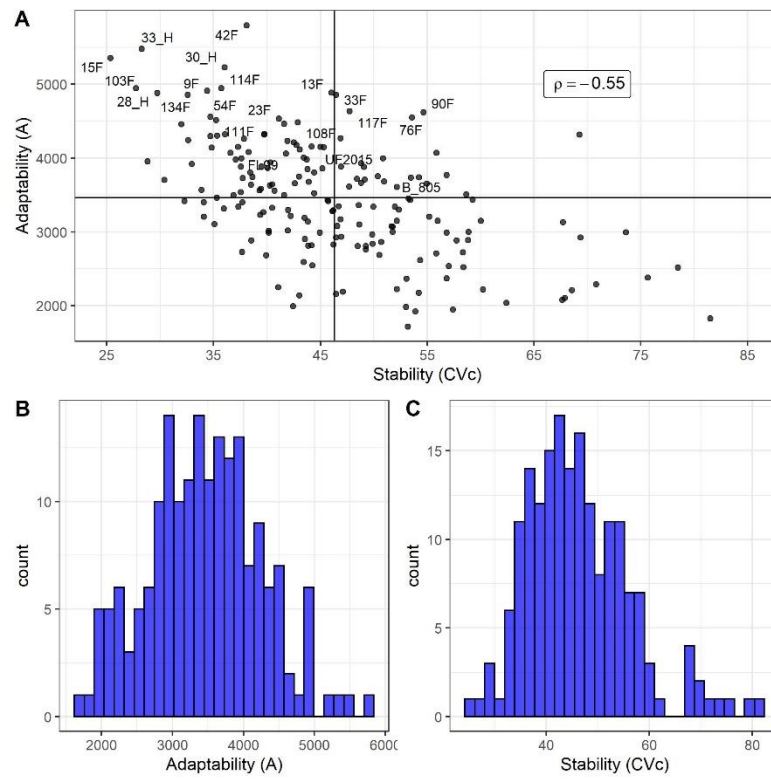
Figure 5 - Genotypes' DMY trajectory over harvest time for trial T1. The Highlighted genotypes represents the better genotypes based on the area under the curve, the checks (B_805, UF2015 and FL99) and the mean DMY trajectory.



Source: from the author (2022).

Although very informative, the genotypes selection based only on the reaction norms is not feasible mainly when the higher number of genotypes are evaluated and the complex GxH interaction plays an important role. To overcome this difficulty, we proposed the genotype selection based on the area under the curve (A) and in the curve coefficient of variation (CV_c), these two parameters represent genotypes' adaptability and stability for DMY, respectively. It was observed high variability for A and CV_c between genotypes (Figures 6A and 6B), and the correlation between the two parameters was of median magnitude (-0.55, Figure 6A), thus it is possible to select genotypes with higher A and lower CV_c , i.e., genotypes presenting high adaptability and stability. For section purposes we did a scatter plot showing the genotype's A and CV_c values, where the genotypes to be selected are on the superior left quadrant of the graph (Figure 6A). The 10% genotypes with higher A are highlighted on Figure 6A.

Figure 6- Stability versus adaptability scatter plot, the solid black lines are the mean values for A and CV_c for trail T1(A); histogram for adaptability (B); histogram for stability (C).



Source: from the author (2022).

3.2.1.3 Variance components and heritability – FAMM

The FA2 model explained most of the genetic variance (99%, Table 5). The first factor explained 78% of the total genotypic variance, whereas the second factor explained 21% (Table 5). The specific variance explained only 1% of the genetic variance and was different from zero only for harvest 1 (Table 5). The higher values of common variance (or communalities) showed that the two factors retained by the model explained a high amount of the genetic variance, showing a good fitting of the model to the data set (Table 5). Since communalities explained most of the genetic variance, there is a clear pattern between the percentage of the common variance in each harvest and heritability, where the higher percentage higher heritabilities were estimated (Table 5). It is worth noting that for the first factor only positive loadings were estimated, so this factor is explained only the simple GxH interaction (Table 5). Otherwise, for the second factor there are positive and negative values, indicating that this factor is explaining the complex GxH interaction (Table 5). The same pattern was observed in this trial for RRM and FAMM, where 74% of the genetic variance was explained for simple GxH as showed by

the eigenfunction and the other 26% was explained by complex GxH interaction (Figure 4B). The genetic correlation between harvest followed the same pattern as estimated by RRM (Figure 4A), where an autoregressive pattern was observed (Figure 7).

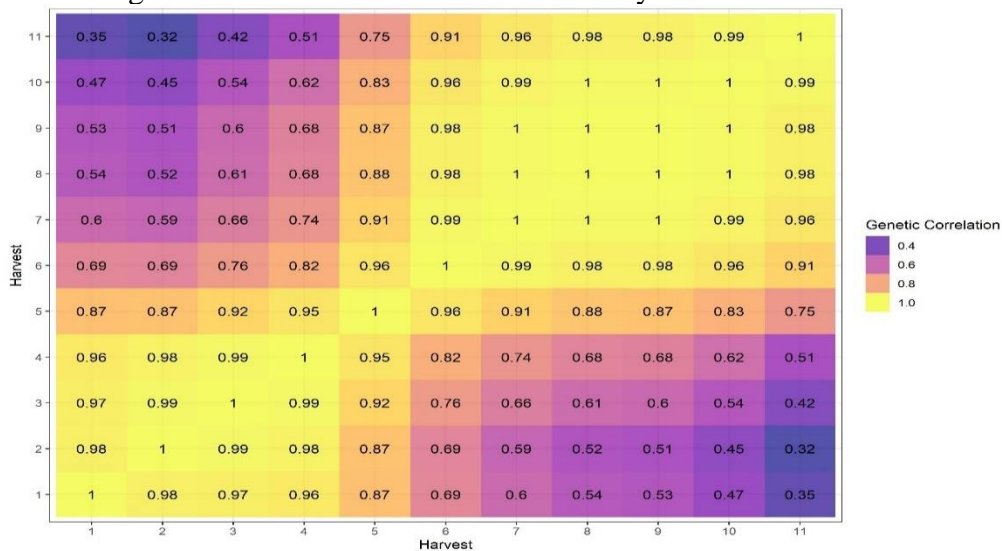
Table 5 - Estimated rotated loadings, common variances (communalities) and specific variances for trial T1 in each harvest.

Harvest	λ_1^*	λ_2^*	H^2	Common Variance	Specific Variance
1	758.7	0	0.62	575,743 (19)	28,846 (1)
2	668.5	23.5	0.55	447,556 (15)	0
3	601.0	-39.4	0.50	362,829 (12)	0
4	532.5	-89.7	0.44	291,671 (10)	0
5	337.0	-174.8	0.28	144,163 (5)	0
6	253.4	-250.1	0.26	126,833 (4)	0
7	175.6	-225.6	0.18	81,766 (3)	0
8	185.2	-278.4	0.23	111,846 (4)	0
9	267.3	-413.2	0.40	242,268 (8)	0
10	258.5	-472.6	0.44	290,213 (10)	0
11	183.9	-481.8	0.42	266,017 (9)	0
Importance (%)	78.0	21.0	-	99.0	1.0

Source: from the author (2022).

$0 \lambda_1^*$ are the estimated rotated loadings for each harvest for the first factor; λ_2^* are the estimated rotated loadings for each harvest for the second factor; the numbers between parenthesis represents the amount of genetic variance explained by common and specific variances.

Figure 7 - Genetic correlation estimated by FAMM for trial T1.

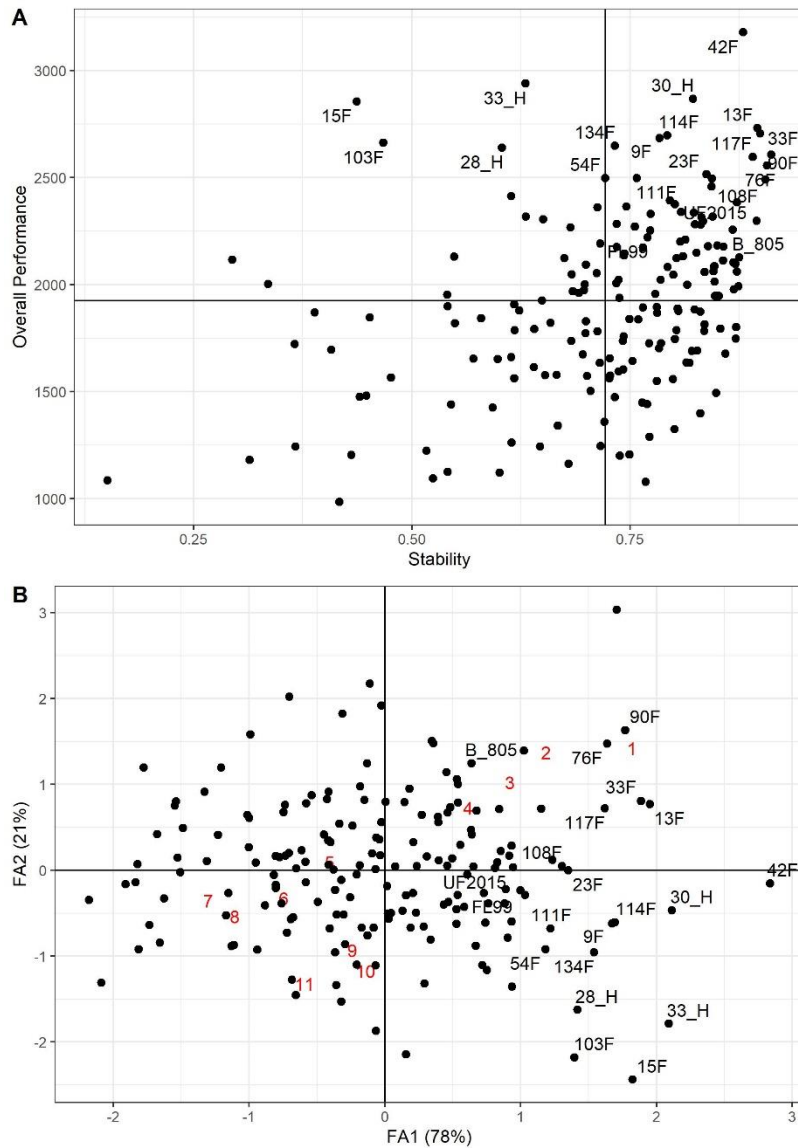


Source: from the author (2022).

3.2.1.4 Genotypes' stability and adaptability – FAMM

Genotype selection in FAMM can be easily by using latent regression, where the genotype's predicted values are regressed over the environmental loadings for the first factor. On latent regressions the genotype slope is related to the genotypes' response to the improvement in environment quality, the regression intercept is related to the broad sense adaptability (overall performance – *OP*) over harvests and the R^2 is the parameter for stability, in which reflects the genotype's predictability. On Figure 1A the R^2 and *OP* were plotted in a scatter plot to select genotypes with higher *OP* and R^2 . The genotypes that presented high *OP* and R^2 can be observed on the superior right quadrant on Figure 8A, where the 10% of the genotypes with higher *OP* were highlighted. In this trial, where the all the loadings for the first factor had positive values (Table 5), indicating that this factor explained the simple GxH interaction, the correlation between the first factor genotypes' scores with *OP* was high (0.99) and R^2 was closed related with the second factor genotypes' scores with correlation 0.76. Therefore, the genotypes' adaptability and stability can be also verified by the FAMM biplot (Figure 8B). The biplot can also be used to observe how harvest are grouped, the harvests from 1 to 4 had higher genetic correlation between each other, whereas harvests 5 to 11 were clustered in another group, these results agree with the estimated genetic correlations (Figure 7 and Figure 8B).

Figure 8 - Overall performance versus stability scatter plot (A); FARM biplot (B), red numbers are the harvests. The 10% best genotypes for overall performance and checks are highlighted.



Source: from the author (2022).

3.2.2 *Urochloa brizantha* advanced breeding trial – T5

In this trial, 9 genotypes were evaluated for DMV through 16 harvests. The degree of Legendre polynomial fitted for this trial was three for fixed and one for random part of the model, whereas for FARM two factors were needed to model the GxH (Table 2). Seven parameters were estimated for RRM and 49 were estimated by FARM (data not shown).

3.2.2.1 Variance components, heritability, and genetic behavior – RRM

For this trial only the intercept and first order polynomial coefficient were needed to model the GxH interaction. The intercept (g_0) and slope (g_1) explained 69 and 31% of the genetic variance (Table 6). The negative correlation ($\rho_{g_0g_1} = -0.41$) between g_1 and g_0 indicates that genetic variance decreased over time. The highest heritabilities estimates occurred between harvests realized between 0 to 71 days ($H^2 > 0.70$), and the lowest heritability occurred between harvests realized between 450 and 619 days ($H^2 < 0.43$) (Figure 8). The first drought season (71 to 238 days – May to October/2009) occurred presented heritability varying from 0.69 to 0.57, whereas the second drought season (450 to 619 – May to October/2010) presented the lower heritabilities varying from 0.40 to 0.43 (Figure 8). The higher heritability occurred on the first drought season can be explained by number of preceding harvests, where only three harvests occurred before for the first drought season, whereas for the second drought season ten harvests had occurred before (Figure 8).

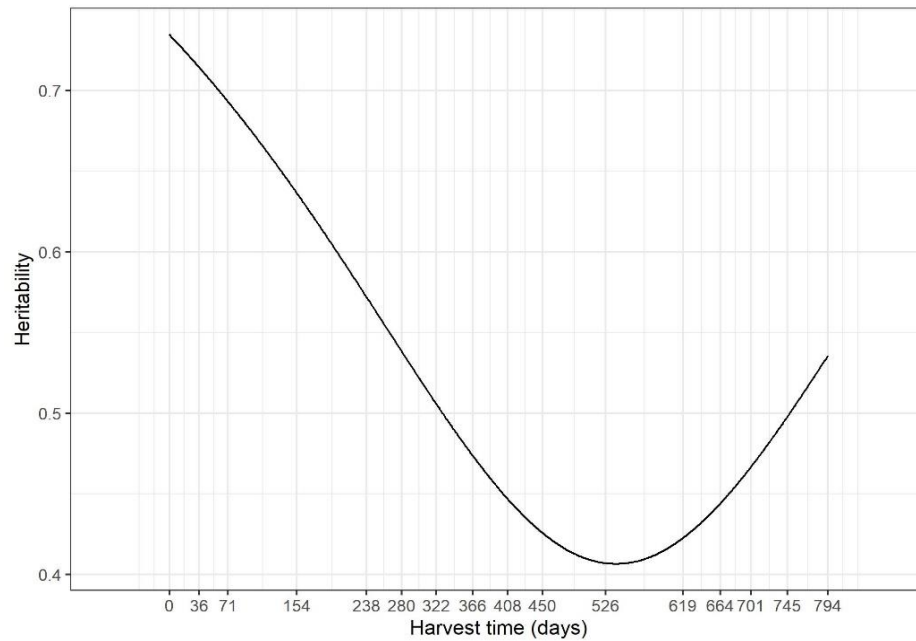
Table 6 - Summary of genetic and non-genetic parameters estimated by RRM for trial T5.

g_0	g_0	g_1
	117,153	-0.41
g_1		53,359
Importance (%)	69	31
$\bar{\sigma}_e^2$	282,442	

Source: from the author (2022).

g_0 and g_1 are the random polynomial intercept and first order coefficient. The diagonal elements of the table are the variance components associated with the polynomial intercept and coefficient; and the off diagonal is the correlation between intercept and regression coefficient ($\rho_{g_0g_1}$). $\bar{\sigma}_e^2$ is the mean error variance through harvests.

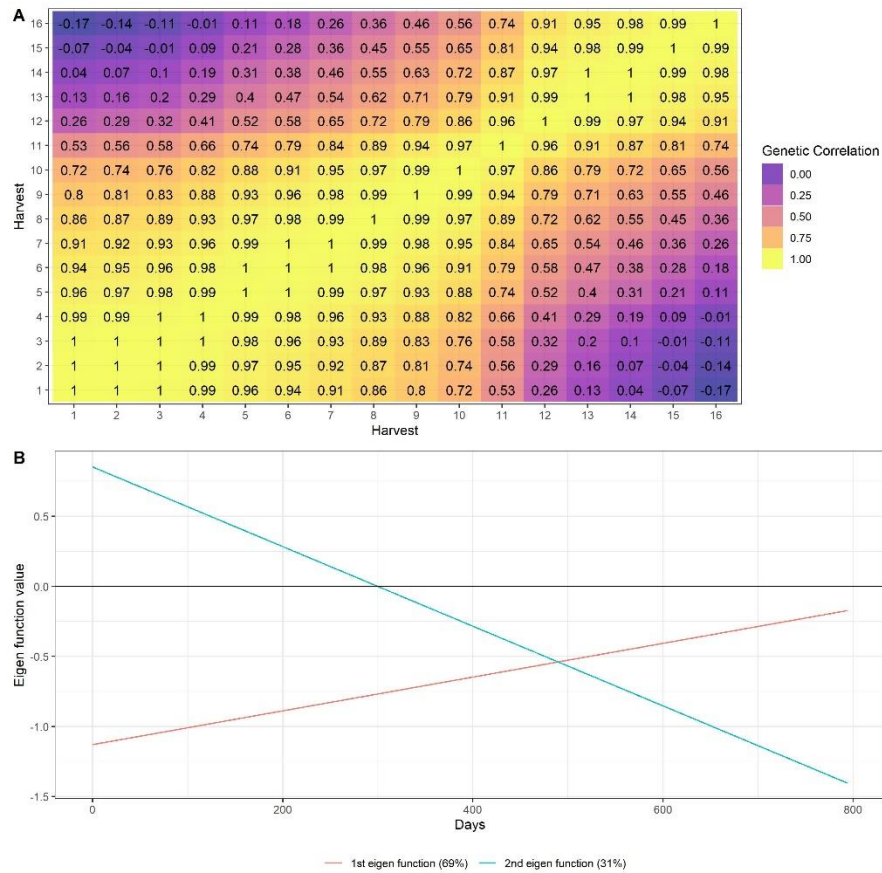
Figure 9 - Heritability estimates over harvests time for trial T5.



Source: from the author (2022).

The genetic correlation across harvests varied from -0.17 to 1.00 (Figure 9A). As occurred for trial T1 (Figure 3A) the genetic correlations followed an autoregressive structure (Figure 9A). Differently than occurred for trial T1, there was not a common factor explaining the GxH interaction for all harvest, since genetic correlations below zero occurred in this trial (Figure 9A). This fact can also be explained by the eigenfunctions, where the two eigenfunctions varied over time (Figure 9B). Both first and second eigenfunctions explained the complex GxH interaction, where the gene expression varied over time for the two eigenfunction (Figure 9B).

Figure 10 - Genetic correlation between harvests estimated by RRM for trial T5 (A); Estimation of the two eigenfunctions for trial T5 (B).



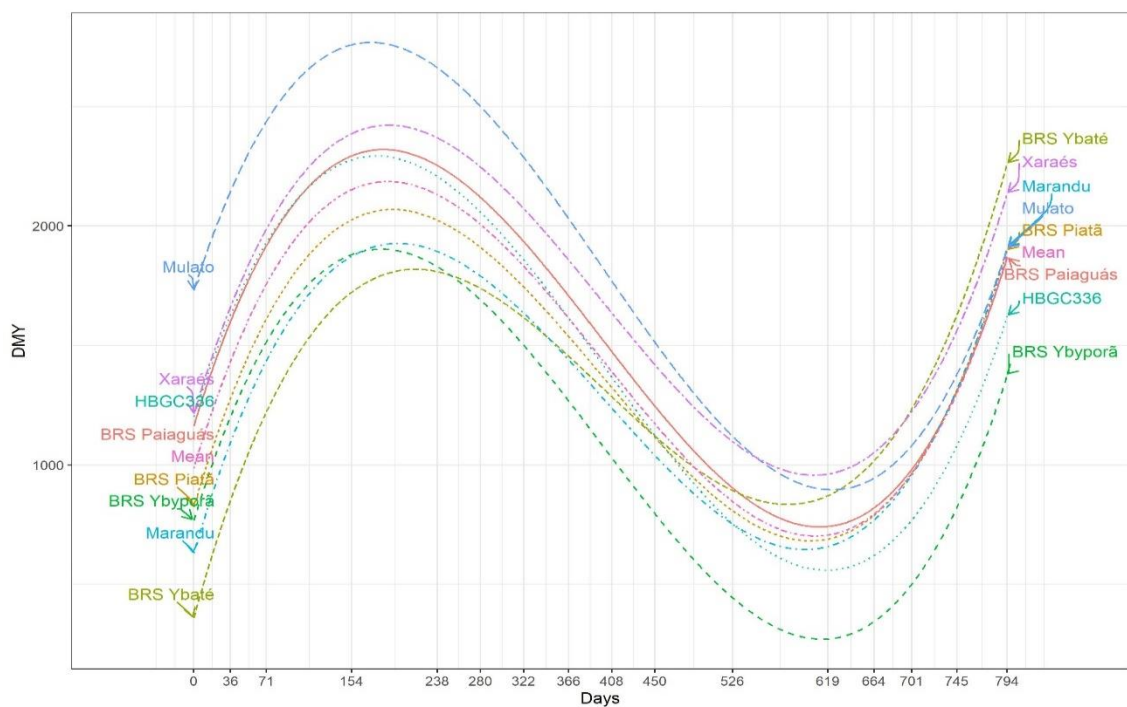
Source: from the author (2022).

3.2.2.2 Genotypes' adaptability, stability and yield trajectory over time – RRM

The genotypes' reaction norms (Figure 11) showed that most changes in genotypes raking occurred between days 238 and 664. The maximum DMY was reached between the first drought season (154 days) and beginning of the second rainy season (238 days) (Figure 11). This is an atypical behavior and can be explained by the time interval between harvests, reflecting on a longer period of dry matter accumulation (Figure 11). Another factor is the number of harvests realized before this period where only three harvests were done until the first drought season (Figure 11). Furthermore, atypical climate condition could happen in this season. As expected, the lowest DMY occurred at the end of the second drought season (619 days) (Figure 11). By the atypical behavior of the genotypes on first drought season, the selection of tolerant genotypes to this condition should be done by looking at genotypic values in the period from 450 to 619 days (Figure 11). Under drought conditions the genotypes BRS

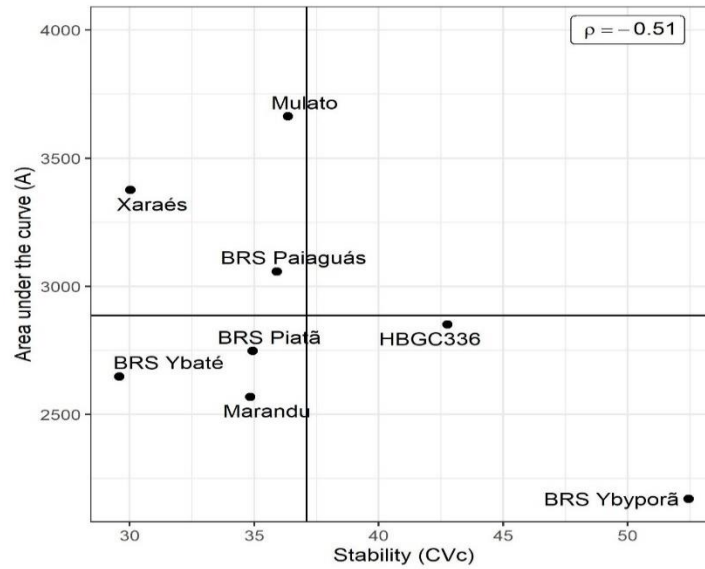
Ybapé, Xaraés and Mulato had the best performance (Figure 11). However, the genotype BRS YBAPÉ had the lowest DMY at the first harvest, indicating a poor establishment (Figure 11). By the reaction norms the genotype Mulato had the best performance across harvest with good establishment as well as good performance under the drought season (Figure 11). The correlation between adaptability and stability was -0.54, indicating that selection can be done for both parameters (Figure 11). Three genotypes can be select by presenting higher stability and adaptability (Mulato, Xaraés and BRS paiaguás, Figure 12).

Figure 11 - Genotypes' DMY trajectory over harvest time for trial T5.



Source: from the author (2022).

Figure 12 - Genotypes' stability versus adaptability scatter plot.



Source: from the author (2022).

3.2.2.3 Variance components and heritability – FAMM

The FA2 model explained a large amount of the genetic variance by communalities (87%), indicating good fit of this model to the data set (Table 7). In this trial, the specific variances explained a considerable amount of the genetic variance (13%), indicating particular gene expression under certain environmental conditions (harvests 1, 2, 4 and 11) (Table 7).

Table 7 - Estimated rotated loadings, common variances (communalities) and specific variances for trial T5 in each harvest.

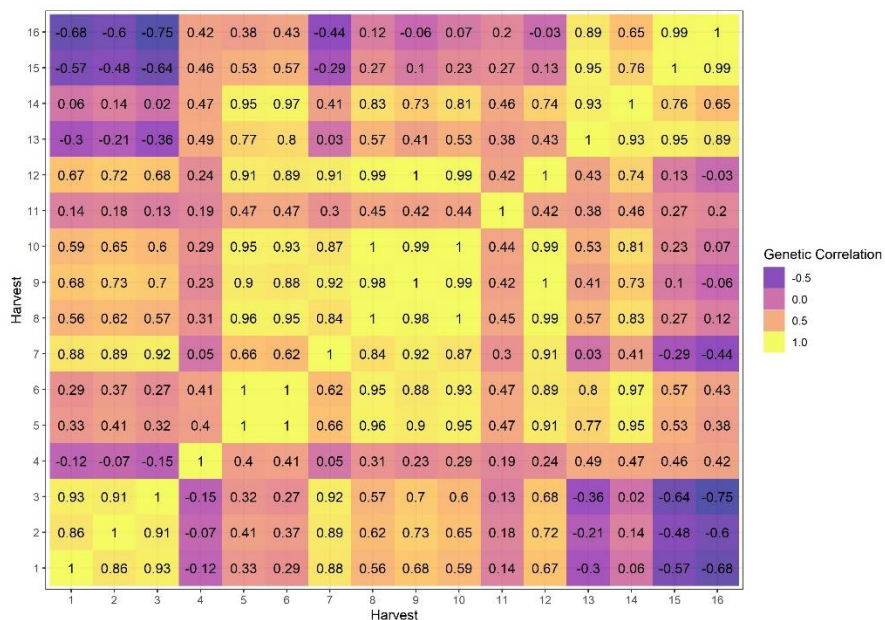
Harvest	λ_1^*	λ_2^*	H^2	Common variance	Specific variance
1	1,147.1	300.5	0.95	1,406,362 (24.7)	202,470 (3.5)
2	1,205.4	193.02	0.96	1,490,354 (26.2)	265,992 (4.7)
3	321.6	98.5	0.61	113,140 (2.0)	0
4	-0.7	-277.7	0.81	77,144 (1.3)	240,748 (4.2)
5	422.6	-589.8	0.88	526,498 (9.2)	0
6	280.9	-437.8	0.79	270,697 (4.8)	0
7	641.6	-66.4	0.85	416,111 (7.3)	0
8	386.0	-307.1	0.77	243,342 (4.3)	0
9	167.3	-90.7	0.33	36,236 (0.6)	0
10	193.7	-140.9	0.44	57,399 (1.0)	0
11	42.30	-64.4	0.27	5,938 (0.1)	20,475 (0.4)
12	228.8	-131.3	0.49	69,617 (1.2)	0

13	-18.5	-260.1	0.49	68,048 (1.2)	0
14	101.1	-310.2	0.6	106,483 (1.9)	0
15	-91.7	-220.0	0.44	56,831 (1.0)	0
16	-59.6	-96.3	0.15	12,832 (0.2)	0
Importance (%)	67	20		87	13

Source: from the author (2022).

When observing the drought season harvests (harvests 4 and 11), the largest amount of genetic variance accounted to the specific variance, in these harvests the amount of specific variance was three times greater than common variance (Table 7). Therefore, specific gene pool was expressed only under drought conditions. Specific variance had also great importance for harvest 1 and 2, indicating specific gene pool being expressed on the plants' establishment phase (Table 7). There was non-common factor acting over all the 16 harvests since both factors had positive and negative values (Table 7). This result indicates that the two factors are explaining complex GxH interaction, thus genetic correlations below zero are expected (Figure 13). Differently than occurred for trial T1 where the genetic correlations followed an autoregressive pattern for both models (RRM and FMM, Figures 3A and 6), in this trial the genetic correlations pattern estimated by FMM followed a more complex structure (Figure 13). The genetic correlation varied from -0.75 to 1.00, indicating a strong complex GxH interaction (Figure 13).

Figure 13 - Genetic correlation estimated by FMM for trial T5.

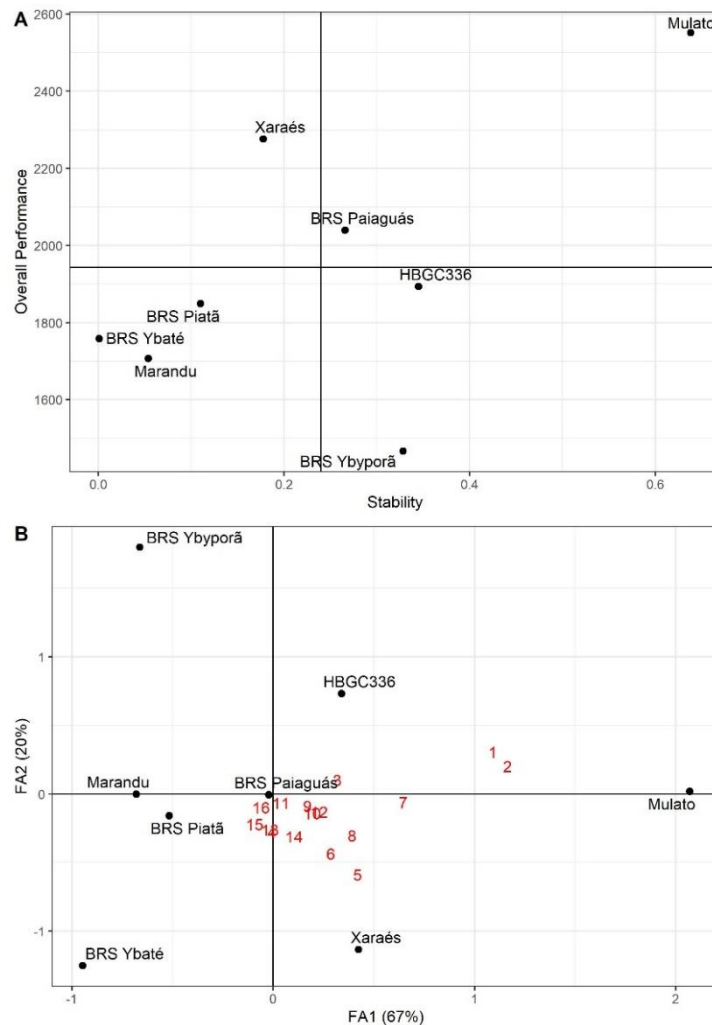


Source: from the author (2022).

3.2.2.4 Genotypes' stability and adaptability – FAMM

The correlation between overall performance (OP) and the first factor genotypes scores was high (0.89), indicating that these scores explained the genotypes adaptability. However, the correlation between the second factor scores and the stability parameter (R^2), was low (0.44). The genotype Mulato and BRS PAIAGUÁS had the better performance for adaptability and stability (Figure 14A), and this result agrees with the best genotypes selected by RRM adaptability and stability parameters (Figure 12). By the FAMM biplot the harvests were separated into four main groups given the distance between harvests in the biplot. The genotypes Mulato, HBGC336 and Xaraés responded positively to latent covariables explained by the first factor (Figure 14B). Only the genotype HBGC331 and HBGC336 responded positively for latent covariables explained by the second factor (Figure 14B).

Figure 14 - Overall performance versus stability scatter plot (A); FAMM biplot (B), red numbers are the harvests.



Source: from the author (2022).

3.2.3 *Urochloa decumbens* advanced breeding trial – T7

In this trial, 12 genotypes were evaluated for DMY through six harvests. The degree of Legendre polynomial fitted for this trial was four for fixed and one for random part of the model, whereas for FAMM one factor was needed to model the GxH (Table 2). Eight parameters were estimated for RRM and 19 were estimated by FAMM (data not shown).

3.2.3.1 Variance components, heritability, and genetic behavior – RRM

Differently than occurred for trial T1 and T5, the intercept explained lower amount (35%) of genetic variance than the first order polynomial coefficient in which explained 65% of the genetic variance (Table 8). Furthermore, the correlation between slope and intercept was positive, indicating that genetic variance increased over time (Table 8). The heritabilities estimates varied from 0.20 at the third harvest (drought season – August/2018) to 0.82 at the last harvest (beginning of drought season – June/2019) (Figure 15).

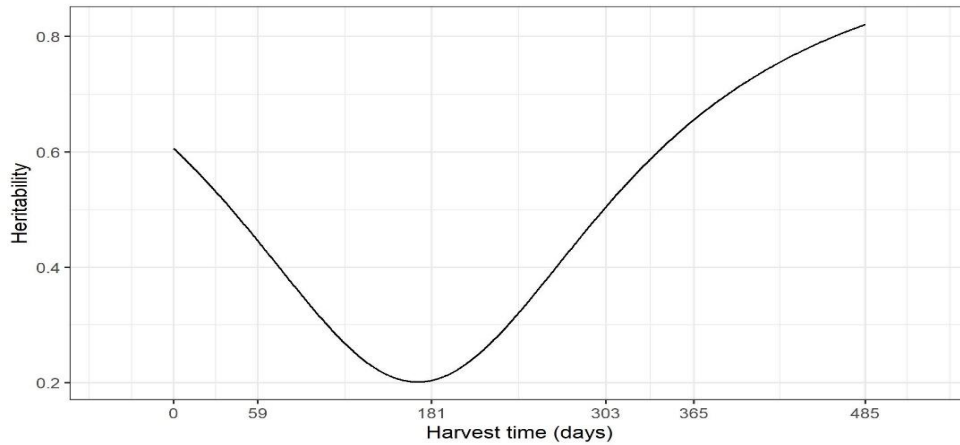
Table 8 - Summary of genetic and non-genetic parameters estimated by RRM for trial T7.

	g_0	g_1
g_0	165,648	0.68
g_1		299,828
Importance (%)	35	65
$\bar{\sigma}_e^2$	694,316	

Source: from the author (2022).

g_0 and g_1 are the random polynomial intercept and first order coefficient. The diagonal elements of the table are the variance components associated with the polynomial intercept and coefficient; and the off diagonal is the correlation between intercept and regression coefficient ($\rho_{g_0g_1}$). $\bar{\sigma}_e^2$ is the mean error variance through harvests.

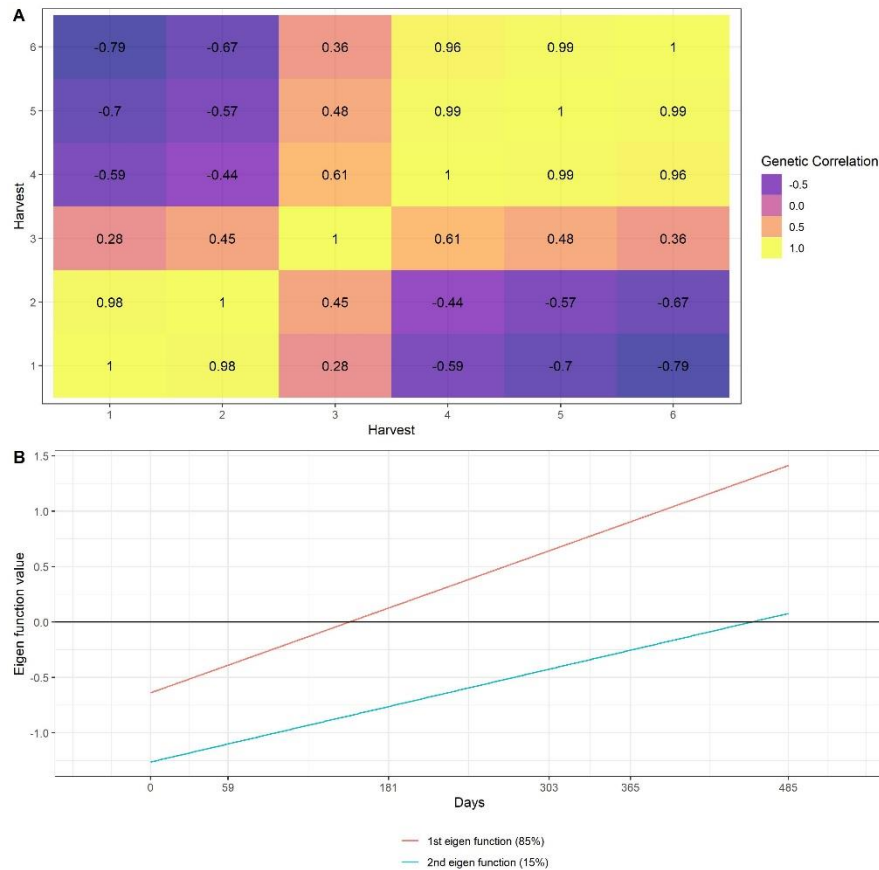
Figure 15 - Heritability estimates over harvests time for trial T7.



Source: from the author (2022).

The genetic correlations across harvests varied from 0.99 to -0.79, indicating a high and complex GxH interaction effect (Figure 16A). As in other trials previously studied the RRM approximated the (co)variance structure into an autoregressive structure, where more close harvests are more correlated than those far apart (Figure 16A). The two estimated eigenfunctions varied across time, indicating there was not a common gene pool expressing in the same way for all harvests evaluated (Figure 16B). Both gene pools represented by the eigenfunctions are expressing differentially across harvest, explaining the strong complex GxH interaction occurred for this trial (Figure 16B).

Figure 16 - Genetic correlation between harvests estimated by RRM for trial T7 (A); Estimation of the two eigenfunctions for trial T7 (B).

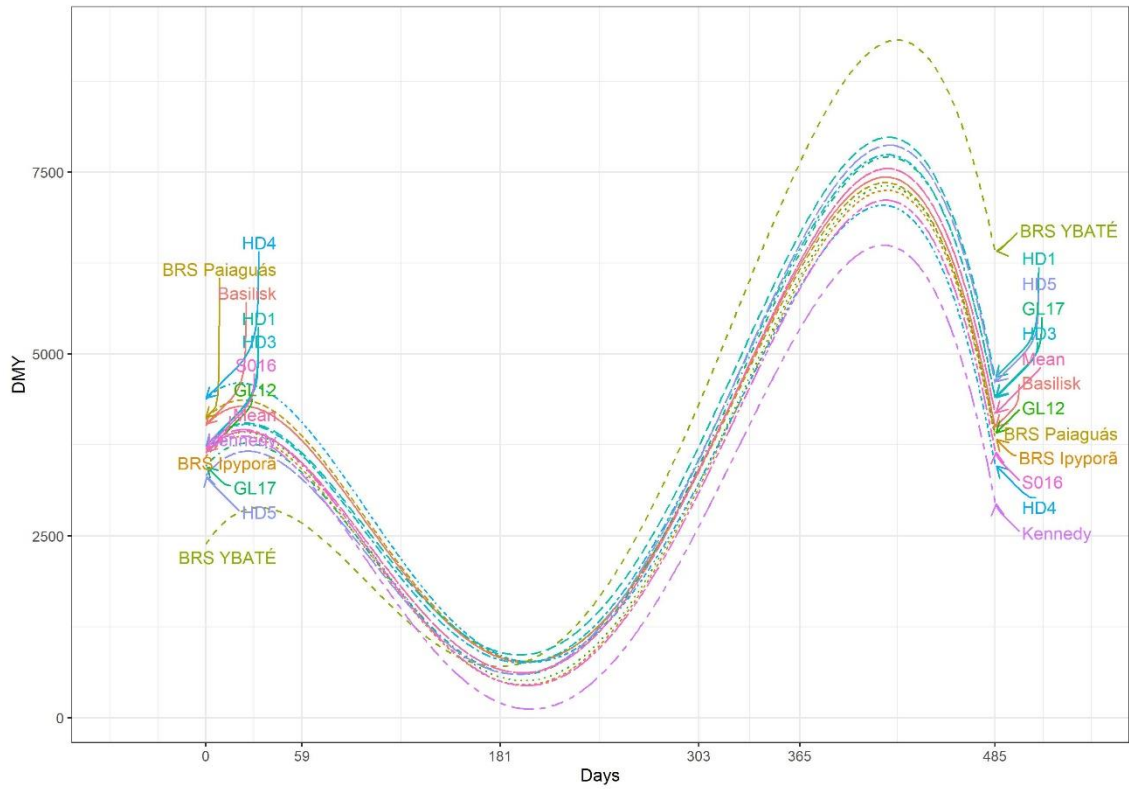


Source: from the author (2022).

3.2.3.2 Genotypes' adaptability, stability and yield trajectory over time – RRM

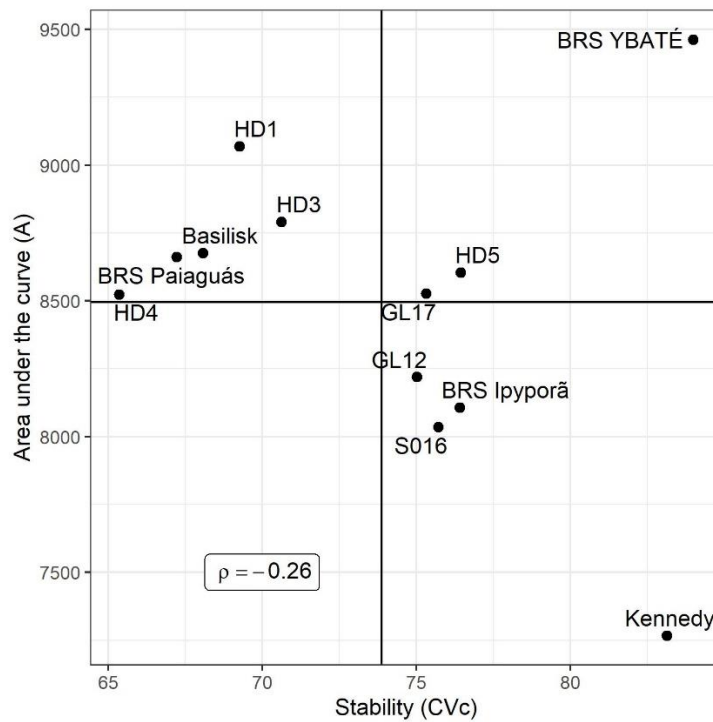
The most changes in genotypes raking occurred after the first drought season at harvest realized on day 181 (Figure 17). The highest DMY occurred at harvest realized on day 365 (rainy season), and the lowest DMY occurred at the first drought season (day 181) (Figure 17). The genotype BRS YBAPÉ, also evaluated in trial T5, had the same behavior, where it had a poor establishment (day 0 to 59) and a good recovery after the first drought season, being the best genotype in all harvests after the drought season (Figure 17). The correlation between CV_c and A was -0.25, indicating the possibility of selecting adaptable and stable genotypes (Figure 17). Although the genotype BRS YBAPÉ had the best adaptability, it was one of the most instable genotype with greater variation in DMY across harvests (Figure 17). Five genotypes were identified having good stability and adaptability (R086, X067, Basilisk, Paiaguás and 254-1) (Figure 17).

Figure 17 - Genotypes' DMY trajectory over harvest time for trial T7.



Source: from the author (2022).

Figure 18 - Genotypes' stability versus adaptability scatter plot.



Source: from the author (2022).

3.2.3.3 Variance components and heritability – FAMM

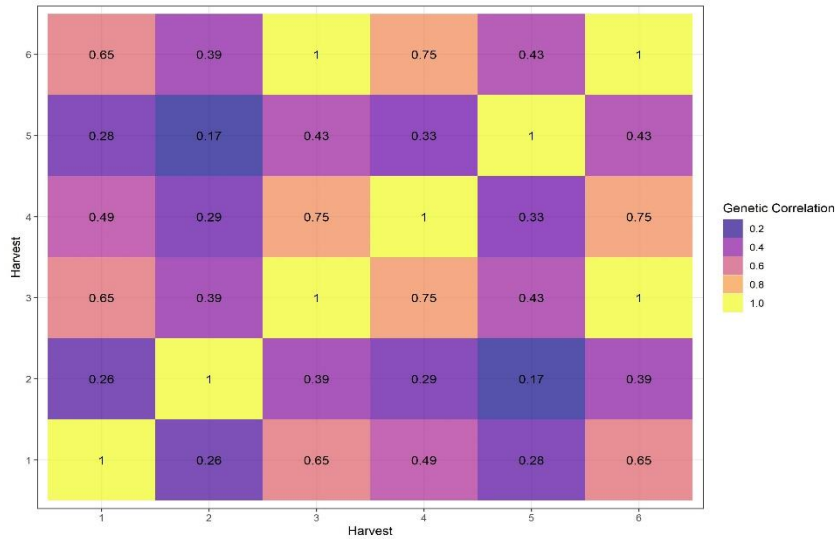
For this trial the best FAMM fitted, only one factor was needed to model the GxH interaction (Table 9). Only positive loadings were observed, indicating this factor is explaining the simple GxH interaction effects (Table 9). The complex GxH interaction can be explained by the specific variances, in which explained 50.7% of the genetic variance, indicating expression of specific gene pools in certain harvest (Table 9). The commonalities explained only 49.3% of the genetic variance, indicating a lack of fitness of the model to the data set (Table 9). The heritabilities estimates varied from 0.96 in harvest 4 to 0.05 in harvest 3, the lowest heritability occurred on the first drought season indicating more difficulty in selecting genotypes under this condition (Table 9). The genetic correlations varied from 0.17 to 0.75 (Figure 19). There was a common factor acting across all harvests since no genetic correlation below zero was observed (Figure 19).

Table 9 - Estimated rotated loadings, common variances (communalities) and specific variances for trial T7 in each harvest.

Harvest	λ_1^*	H^2	Common variance	Specific variance
1	1,194	0.95	1,426,537 (12.1)	1,927,396 (16)
2	410	0.86	168,393 (1.4)	921,458 (7.8)
3	104	0.05	10,828 (0.9)	0
4	1,596	0.96	2,547,977 (21.7)	1,977,999 (16.8)
5	522	0.89	273,390 (2.3)	1,179,174 (10.1)
6	1,133	0.88	1,283,690 (10.9)	0
Importance (%)	49.3		49.3	50.7

Source: from the author (2022).

Figure 19 - Genetic correlation estimated by FMM for trial T5.

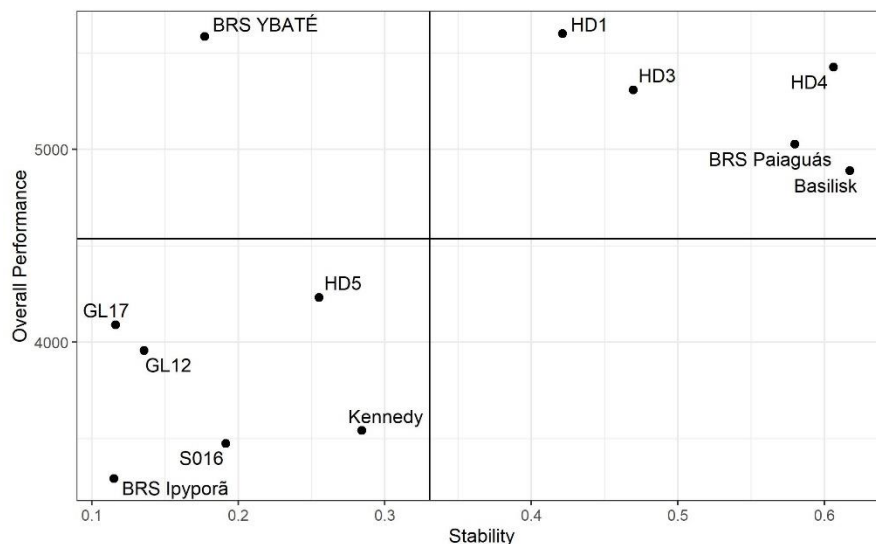


Source: from the author (2022).

3.2.3.4 Genotypes’ stability and adaptability – FMM

Although different (co)variance structures were estimated for RRM and FMM (Figure 16A and Figure 19), the same genotypes were selected based on broad adaptability and stability (Figure 20). The genotype BRS YBAPÉ had the best overall performance, however it is an instable genotype (Figure 20). Five genotypes were observed to have higher adaptability and stability (R086, X067, 254-1, Paiaguás and Basilisk) (Figure 20).

Figure 20- Overall performance versus stability scatter plot.



Source: from the author (2022).

4 DISCUSSION

In this study we have presented methods of analysis for longitudinal DMY data generated from perennial forage breeding trials for ten trials and four different perennial species. Two approaches were addressed in this paper the random regression models (RRM) and factor analytic mixed models (FAMM). These two approaches were interpreted in sense of variance components estimation and genotype selection for DMY adaptability and stability.

4.1 Random regression models (RRM)

4.1.1 Goodness of fit evaluation

Statistical methods for analyzing yield data of forage perennial species need to model the genetic effects over time properly (DE FAVERI et al., 2015). RRM is commonly used to deal with longitudinal records in animal breeding (SCHAEFFER, 2004; KRANIS et al., 2007). In all trials analyzed in this study, the DMY records were obtained unequally spaced in time, due to the seasonality present in forage grown under tropical and subtropical climate regions (REIS and ROSA, 2001). Kirkpatrick, Lofsvold and Bulmer (1990) reported that one advantage of infinite-dimensional method such as random regression, it can handle with unequally spaced records. This is because yield trajectories are continuous functions of time, so that a trait in an individual requires an infinite rather than finite number of measurements to fully describe (Kirkpatrick and Heckman, 1989).

The random polynomial order choosing in RRM can be easily done by using goodness of fit and parsimony criteria (CORRALES; MUNILLA; CANTE, 2015). In this study, the random polynomial order was selected by the BIC as suggested by Rocha et al. (2018). Some difficulties can arise when in choosing the phenotypic curve shape (fixed function). When the overall trajectory is linear fitting the fixed part of the model by a function can be simple. However, fluctuation on forage DMY always occur due to the climate conditions changes over time. Therefore, DMY trajectory will always follow a non-linear pattern. In this case, the fixed part of the model can be treated as factor variables (SCHAEFFER, 2004). However, factor variables take up more degree of freedom. In this study we suggested to establish a fixed function based on a smooth loess function, where the polynomial order can be chosen graphically (Figure

1). By using a mathematical function instead of factors, gave us smooth trajectories over time regardless of the number of observations (Figures 5, 11 and 17).

4.1.2 Genotype by harvest interaction

In RRM, the genotype by harvest interaction can be studied by ‘genotypes reaction norms’, genetic correlations estimated by covariance functions and/or eigenfunctions across time. When interpretation is done by reaction norms (Figures 5, 11 and 17), the GxH interaction occurs when reaction norms are not parallel, i.e., they intersect, diverge or converge (VAN EEWIJK et al., 2016). Divergence and convergence occur when simple GxH is acting over time, meaning that the genetic variance is increasing (diverging) or decreasing (converging) (CROSSA; YANG; CORNELIUS, 2004). The complex GxH interaction occur when the reaction norms intersect, meaning that is a lack of genetic correlation between measurements (ELIAS et al., 2016; VAN EEUWIJK; BUSTOS-KORTS; MALOSETTI, 2016). In all trials analyzed, the estimated genetic correlations tended to follow an autoregressive pattern, where harvests closer to each other had higher correlations, whereas harvests far apart had lower genetic correlation (Figures 4A, 11A and 16A). This autoregressive pattern is very common in longitudinal data in perennial species (APIOLAZA and GARRICK, 2001; YANG et al., 2006; GIRI et al., 2019; PEIXOTO et al., 2020; BORNHOFEN et al., 2022; ANDRADE et al., 2022) and have a satisfactory biological explanation, indicating that genes are expressing differently according to the environmental conditions and genotypes’ age (GAUCH and ZOBEL, 1996; FALCONER and MACKAY, 1996). Another way to interpreting the GxH interaction in RRM is through eigenfunctions. Eigen functions are analogous to eigenvectors (principal components). Each eigenfunction is a continuous function that represents a possible evolutionary deformation of the mean yield trajectory (KIRKPATRICK; LOFSVOLD; BULMER, 1990). When the eigenfunction is nearly constant, it means that the eigenfunction captured a gene pool that was equally expressed over time (KIRKPATRICK; LOFSVOLD; BULMER, 1990; ROCHA et al., 2018; PEIXOTO et al., 2020). On trial T1, the first eigenfunction had a constant behavior and it is explaining the general adaptability gene pool equally expressed over time and the positive genetic correlation (Figure 4B). The other eigenfunctions from trial T1 are explaining the lack of genetic correlation, where the gene pools had differential expression over time (Figure 4B). For trials T5 and T7, there was negative genetic correlation over time indicating that there was not a common gene pool expressing equally through the harvests, therefore none eigenfunctions had a constant behavior (Figures

11B and 16B). As demonstrated in this study, eigenfunctions can explain the GxH interaction as introduced Falconer and Mackay (1996), where genotype by environment interaction can be considered a pleiotropic effect of a trait evaluated across environments.

4.1.3 Genotypes' reaction norm, adaptability and stability

A reaction norm defines a genotype-specific function that translates environmental inputs into a phenotype (VAN EEUWIJ; BUSTOS-KORTS; MALOSETTI, 2016). The differential genotypes' response to the environment (genotype plasticity) generates the genotype by environment interaction. In this study, it was observed a higher variation for genotypes' reaction norms for all trials interpreted (Figures 5, 11 and 17), indicating strong complex GxH interaction. Reaction norms are very informative for analyzing perennial forage over time, it allowed us to observe the behavior of each genotype, identify periods where seasonality occurred, and identify those genotypes which respond better to environmental stress (Figures 5, 11 and 17). However, when the number of genotypes is too high interpreting each reaction norm can be difficult. Therefore, breeders usually use specific-genotype parameters, such as intercepts, slopes, curvatures, and variances. These specific-genotype parameters are called sensitivity, adaptability and stability parameters in plant breeding literature and they facilitate the modeling of complex genotype by environment interaction (FINLAY and WILKINSON, 1963; EBERHART and RUSSEL, 1966; LIN and BINS, 1988; PIEPHO, 1998; SLAFER et al., 2014). Another way to select genotype based on adaptability and stability is computing an index regarding the predicted genotypic values across all environments (KELLY et al., 2007; FAVERI et al., 2015; ROCHA et al., 2018). In this study, we proposed as adaptability measure the genotype's area under the curve, in which have a closer meaning to the total DMY across all harvest (A) and overall genotype's performance. For stability parameter we proposed the use of the curve coefficient of variation (CV_c). Stability can also be referred as risk in variety adoption, where the most stable genotype should have lower variance across environments, meaning that the genotype is more predictable (EBERHART and RUSSEL, 1966; WRICKE, 1966, LIN and BINNS, 1988). The coefficient of variation is a broad used and easy to interpret parameter in different disciplines. In time series, mainly in economics it is frequently used to infer about the risk and uncertainty in shares on the stock market exchange (CURTO and PINTO, 2009). Therefore, we used this concept to infer about

DMY stability in genotype selection, in which genotypes having lower CV_c will have higher stability and will also have lower variation across harvests.

4.2 Factor analytic mixed models (FAMM)

4.2.1 Goodness of fit evaluation

The factor analysis is a multivariate method (LAWLEY and MAXWELL, 1971; COMREY and LEE, 2013) that allows the multivariate correlated data simplification as well as the principal component analysis (PCA). The FAMM can be considered as an extension of the PCA which approximates the (co)variance structure into an unstructured structure (Figure 3), generating parsimonious models (SMITH; CULLIS; THOMPSON, 2005). In the FAMM described by Smith, Cullis and Thompson (2001) and Resende and Thompson (2003; 2004), the fit of the model, i.e., the number of multiplicative terms needed can be tested by the likelihood ratio test (LRT). In this study, we tested the number of factors to be retained by the FAMM by using the LRT test, the maximum number of factors needed through the trial was two, and six trials needed two factors whereas four trials needed only one factor (Table 2). The fit of the model can also be evaluated by the explanation of the genetic variance by common variances (communalities), i.e., the amount of genetic variance explained by the factors (RESENDE et al., 2014). When observing the interpreted trials in this study the factors explained higher amount of the genetic variance for trials T1 (99%) and T5 (87%) was explained by the communalities, whereas for T7 only 49% of the genetic variance was explained by communalities (Tables 5, 7 and 9).

4.2.2 Genotype by harvest interaction

The FAMM allows the GxH interpretation through the genetic correlations between harvests, and by the estimated common and specific variances. Differently than occurred for RRM where all the (co)variance structures followed an autoregressive pattern, in FAMM as reported by Smith, Cullis and Thompson (2001) the factors will approximate the (co)variance structure into an unstructured matrix (Figure 3). In this study, the variance structure estimated by FAMM followed an autoregressive pattern for trial T1 (Figure 7). However, for trial T5 and

T7 there was not a clear pattern in the genetic correlations (Figures 13 and 18). These results demonstrate that FAMM are very flexible and deal with a complex pattern of genetic correlation between harvests. The FAMM were capable to capture either complex and simple interaction, where harvests with higher genetic correlation had simpler GxH interaction and those in which the genetic correlation was smaller had more complex GxH interaction (Figures 7, 13 and 18). The amount of complex and simple interaction captured by the factors can be investigated based on the rotated estimated environmental loadings. Cullis et al. (2014) pointed out that if the environmental loadings from a factor are predominantly positive, it means that most of the genotype slopes in the latent regression will be positive and is expected no changes in genotypes ranking. However, if the environmental loadings for a factor have positive and negative values, it is expected positive and negative slopes for genotypes latent regressions, therefore crossover interaction is expected. For trial T1 and T7, the environmental loadings for the first factor were all positives, indicating that this factor is explaining most of the simple GxH interaction. Therefore, the factors in FAMM can have an analogous interpretation as the eigenfunctions in RRM. The common variances represent a common gene pool that is expressing in all harvest, when the loadings are positive for a factor it means that this gene pool are expressing equally over the harvests, whereas when positive and negative values are present it means a differential gene expression over harvests. By the nature of the G×E in this study, exploring the specific variance components is not feasible, since time is not a repeatable environment (Yan, 2016). Therefore, forage breeders should focus in exploring the variance captured by the factors in FAMM.

4.2.3 Genotypes' reaction norm adaptability and stability

In FAMM, adaptability and stability can be investigated by latent regression, in which the genotypic value for each harvest is regressed over an environmental gradient (environmental loadings) (CULLIS et al., 2014; SMITH and CULLIS et al., 2018; OLIVEIRA et al., 2020). In this context, the environmental rotated loadings will represent the environmental quality, where lower values represent lower environmental quality whereas higher values represent high quality environments. In this way, the latent regression represents the genotype's reaction norm in response to the environment. The advantage of this method is to represent the genotype's response with a simple linear regression and the interpretation can be done as proposed by Finlay and Wilkson (1963), where the slope reflect the genotypes sensitivity, the intercept is

the broad adaptability, and the stability can be interpreted based on the R^2 of the regression to infer about genotype's predictability. As suggested by Smith and Cullis (2018), the genotype selection can be done by scatter plots where adaptability and stability can easily be observed for each genotype (Figures 8A, 14A and 19). Another advantage of FAMM is the possibility of clustering harvests and select genotypes adaptable to specific environmental conditions through biplots (Figures 8A and 14A).

4.3 Comparison between RRM and FAMM for longitudinal data

Both models RRM and FAMM are very useful in dealing with GxH interaction in forage breeding trials and offer a detailed information about GxH interaction. The FAMM in all trials evaluates yielded lower BIC estimations than RRM (data not shown), which is expected since FAMM approximate an unstructured (co)variance matrix and this structure is the most general variance model, and therefore the model will provide the best fit (in a likelihood sense) to the data (SMITH; CULLIS; THOMPSON, 2005). However, the interpretability of RRM for forage yield data regarding on the genotype's reaction norm is more realist, where breeder can select the genotypes based in their trajectory over a continuous time. The environmental loadings (covariates) in FAMM are estimated from the data rather than externally (SMITH; CULLIS; THOMPSON, 2005; CULLIS et al., 2014; SMITH and CULLIS, 2018) and the correlation between environmental loadings and external covariates (temperature, humidity, latitude, etc.) can be done *ad hoc* (OLIVEIRA et al., 2020). In this way, RRM can be improved by incorporating not only time, but also environmental covariates in which can generate more complex variance structure when combining different (co)variance structures estimated for each covariate. Bohlouli et al, 2019, reported higher prediction accuracies of longitudinal traits in dairy cattle when using simultaneously a temperature-humidity index as covariate in RRM. Therefore, a way to use FAMM and RRM jointly would be select the most important covariate using the correlation between the estimated loadings by FAMM and environmental covariates to choose the most important environmental covariates. These most important covariates could be then studied in RRM, to evaluate how these covariates is shaping the genotypes' reaction norms. The greater advantage of RRM over the FAMM, is the possibility of generating yield trajectory over time, which is very appealing for forage breeding, to identify genotypes that genotypes having less seasonality.

RRM can be very useful on genomic selection, since the predictions are based on regression parameters (slope, intercept, and curvatures), the genomic prediction will be based in forecasting the genotypes over time which can be very informative for forage breeding. Another advantage is the possibility of QTL mapping over time (YANGET *al.*, 2006), where different QTLs can be investigated in each stage of the genotype growth season.

5 CONCLUSION

In this study we showed the application of FAMM and RRM applied in longitudinal data generated from forage breeding trials. Both models are powerful in dealing with longitudinal data set giving valuable information about GxH interaction. We also proposed the estimation of adaptability and stability based in the area under the reaction norm curve and the curve coefficient of variation. In longitudinal data context the RRM allowed better interpretation, where is possible to investigate genotypes seasonality by the predicted reaction norms. Therefore, we recommend the use of RRM for longitudinal traits in forage breeding trials.

REFERENCES

- ANDRADE, Mario Henrique *et al.* Genomic prediction for canopy height and dry matter yield in alfalfa using family bulks. **The Plant Genome**, [S.l.], p. e20235, 2022.
- APIOLAZA, Luis A.; GARRICK, Dorian J. Analysis of longitudinal data from progeny tests: some multivariate approaches. **Forest Science**, Bethesda, v. 47, n. 2, p. 129-140, 2001.
- ARANGO, J. A.; CUNDIFF, L. V.; VAN VLECK, L. Dale. Covariance functions and random regression models for cow weight in beef cattle. **Journal of Animal Science**, Champaign, v. 82, n. 1, p. 54-67, 2004.
- BOHLOULI, Mehdi *et al.* The relationship between temperature-humidity index and test-day milk yield of Iranian Holstein dairy cattle using random regression model. **Livestock Science**, Amsterdam, v. 157, n. 2-3, p. 414-420, 2013.
- BORNHOFEN, Elesandro *et al.* Leveraging spatio-temporal genomic breeding value estimates of dry matter yield and herbage quality in ryegrass via random regression models. **Plant Genome**, Madison, v. 15, n. 4, p. e20255, 2022.
- BRÜGEMANN, K. *et al.* Genetic analyses of protein yield in dairy cows applying random regression models with time-dependent and temperature x humidity-dependent covariates. **Journal of Dairy Science**, Champaign, v. 94, n. 8, p. 4129-4139, 2011.

BUTLER, D. G. *et al.* **ASReml-R reference manual**. Brisbane: Queensland Government, 2021.

CAMPBELL, Malachy; WALIA, Harkamal; MOROTA, Gota. Utilizing random regression models for genomic prediction of a longitudinal trait derived from high-throughput phenotyping. **Plant Direct**, Oxford, v. 2, n. 9, p. e00080, 2018.

COMREY, Andrew L.; LEE, Howard B. **A first course in factor analysis**. London: Psychology, 2013.

CORRALES, J. D.; MUNILLA, S.; CANTET, R. J. C. Polynomial order selection in random regression models via penalizing adaptively the likelihood. **Journal of Animal Breeding and Genetics**, Hamburg, v. 132, n. 4, p. 281-288, 2015.

CROSSA, José; YANG, Rong-Cai; CORNELIUS, Paul L. Studying crossover genotype× environment interaction using linear-bilinear models and mixed models. **Journal of Agricultural, Biological, and Environmental Statistics**, [S.l.], v. 9, n. 3, p. 362-380, 2004.

CULLIS, Brian R. *et al.* Factor analytic and reduced animal models for the investigation of additive genotype-by-environment interaction in outcrossing plant species with application to a *Pinus radiata* breeding programme. **Theoretical and Applied Genetics**, Berlin, v. 127, n. 10, p. 2193-2210, 2014.

CURTO, J. D.; PINTO, J. C. The coefficient of variation asymptotic distribution in the case of non-iid random variables. **Journal of Applied Statistics**, Abingdon, v. 36, n. 1, p. 21-32, 2009.

DE FAVERI, Joanne *et al.* Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. **Crop and Pasture Science**, Victoria, Australia, v. 66, n. 9, p. 947-962, 2015.

EBERHART, S. A.; RUSSEL, W. A. Stability parameters for comparing varieties. **Crop Science**, Madison, v. 6, p. 36-40, 1966.

ELIAS, Ani A. *et al.* Half a century of studying genotype× environment interactions in plant breeding experiments. **Crop Science**, Madison, v. 56, n. 5, p. 2090-2105, 2016.

FINLAY, K. W.; WILKINSON, G. N. The analysis of adaptation in a plant-breeding programme. **Australian Journal of Agricultural Research**, Victoria, Australia, v. 14, n. 6, p. 742-754, 1963.

GAUCH, H. G., Jr., ZOBEL, R. W. AMMI analysis of yield trials. In: GAUCH, H. G., Jr., ZOBEL, R. W. **Genotype-by-environment interaction**. Boca Raton: CRC, 1996. p. 85-122.

HAND, David; CROWDER, Martin. **Practical longitudinal data analysis**. London: Routledge, 2017.

HENDERSON, J. R.; CHARLES, R. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. **Biometrics**, Washington, p. 623-640, 1982.

HUISMAN, A. E.; VEERKAMP, R. F.; VAN ARENDONK, J. A. M. Genetic parameters for various random regression models to describe the weight data of pigs. **Journal of Animal Science**, Champaign, v. 80, n. 3, p. 575-582, 2002.

KELLY, Alison M. *et al.* The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. **Crop Science**, Madison, v. 47, n. 3, p. 1063-1070, 2007.

KIRKPATRICK, Mark; HECKMAN, Nancy. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. **Journal of Mathematical Biology**, New York, v. 27, n. 4, p. 429-450, 1989.

KIRKPATRICK, Mark; LOFSVOLD, David; BULMER, Michael. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, Austin, v. 124, n. 4, p. 979-993, 1990.

KRANIS, Andreas *et al.* The application of random regression models in the genetic analysis of monthly egg production in turkeys and a comparison with alternative longitudinal models. **Poultry Science**, Champaign, v. 86, n. 3, p. 470-475, 2007.

LAIRD, Nan M.; WARE, James H. Random-effects models for longitudinal data. **Biometrics**, Washington, v. 38, p. 963-974, 1982.

LAWLEY, Derrick Norman; MAXWELL, Albert Ernest. **Factor analysis as a statistical method**. London: Butterworths, 1971.

LEE, Dae-Jin; DURBÁN, María; EILERS, Paul. Efficient two-dimensional smoothing with P-spline ANOVA mixed models and nested bases. **Computational Statistics & Data Analysis**, Amsterdam, v. 61, p. 22-37, 2013.

LEWIS, R. M.; BROTHERSTONE, S. A genetic evaluation of growth in sheep using random regression techniques. **Animal Science**, Penicuik, v. 74, n. 1, p. 63-70, 2002.

LIN, C. S.; BINNS, M. R. A method of analyzing cultivar x location x year experiments: a new stability parameter. **Theoretical and Applied Genetics**, Berlin, v. 76, n. 3, p. 425-430, 1988.

MACKAY, Trudy F. C. DS Falconer and introduction to quantitative genetics. **Genetics**, Austin, v. 167, n. 4, p. 1529-1536, 1996.

MBUTHIA, J. M.; MAYER, M.; REINSCH, N. Modeling heat stress effects on dairy cattle milk production in a tropical environment using test-day records and random regression models. **Animal**, Cambridge, v. 15, n. 8, p. 100222, 2021.

MCKENZIE, David J. Estimation of AR (1) models with unequally spaced pseudo-panels. **The Econometrics Journal**, Oxford, v. 4, n. 1, p. 89-108, 2001.

MEYER, Karin. Estimating covariance functions for longitudinal data using a random regression model. **Genetics Selection Evolution**, Paris, v. 30, n. 3, p. 221-240, 1998.

- MOREIRA, Fabiana *et al.* High-throughput phenotyping and random regression models reveal temporal genetic control of soybean biomass production. **Frontiers in Plant Science**, Lausanne, v. 12, p. 715983, 2021.
- OLIVEIRA, Isadora Cristina Martins *et al.* Genotype-by-environment interaction and yield stability analysis of biomass sorghum hybrids using factor analytic models and environmental covariates. **Field Crops Research**, Amsterdam, v. 257, p. 107929, 2020.
- PEIXOTO, Marco Antônio *et al.* Random regression for modeling yield genetic trajectories in *Jatropha curcas* breeding. **PLoS One**, San Francisco, v. 15, n. 12, p. e0244021, 2020.
- REIS, R. A.; ROSA, B. Suplementação volumosa: conservação do excedente das pastagens. In: SIMPÓSIO SOBRE MANEJO DA PASTAGEM, 18., 2001 Piracicaba. **Anais...** Piracicaba: FEALQ, 2001. p. 193-232.
- RESENDE, M. D. V. *et al.* Estatística matemática, biométrica e computacional. Viçosa: UFV, 2014.
- RESENDE, Marcos Deon Vilela; THOMPSON, Robin. Factor analytic multiplicative mixed models in the analysis of multiple experiments. **Revista de Matemática e Estatística**, São Paulo, v. 22, n. 2, p. 31-52, 2004.
- RESENDE, Marcos Deon Vilela; THOMPSON, Robin. **Multivariate spatial statistical analysis of multiple experiments and longitudinal data**. Colombo: Embrapa Florestas, 2003.
- ROCHA, João Romero do Amaral Santos de Carvalho *et al.* Genetic insights into elephantgrass persistence for bioenergy purpose. **PLoS one**, San Francisco, v. 13, n. 9, p. e0203818, 2018.
- RODRIGUEZ-ALVAREZ, Maria Xose *et al.* Correcting for spatial heterogeneity in plant breeding experiments with P-splines. **Spatial Statistics**, Amsterdam, v. 23, p. 52-71, 2018.
- SCHAEFFER, L. R., DEKKERS, J. C. M. Random regressions in animal models for test-day production in dairy cattle. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 5., 1994, Canadá. **Proceedings...** Guelph, 1994. v. 18, p. 443-446.
- SCHAEFFER, Lawrence R. Application of random regression models in animal breeding. **Livestock Production Science**, Amsterdam, v. 86, n. 1-3, p. 35-45, 2004.
- SMITH, A. B.; CULLIS, Brian R.; THOMPSON, R. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. **The Journal of Agricultural Science**, v. 143, n. 6, p. 449-462, 2005.
- SMITH, Alison B.; CULLIS, Brian R. Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. **Euphytica**, Wageningen, v. 214, n. 8, p. 1-19, 2018.

SMITH, Alison Barbara *et al.* Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. **Euphytica**, Wageningen, v. 157, n. 1, p. 253-266, 2007.

SMITH, Alison; CULLIS, Brian; THOMPSON, Robin. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. **Biometrics**, Washington, v. 57, n. 4, p. 1138-1147, 2001.

SMITH, K. F.; SPANGENBERG, G. Forage breeding for changing environments and production systems: an overview. **Crop and Pasture Science**, Victoria, Australia, v. 65, n. 11, p. i-ii, 2014.

SUN, Jin *et al.* Multitrait, random regression, or simple repeatability model in high-throughput phenotyping data improve genomic prediction for wheat grain yield. **The plant genome**, Madison, v. 10, n. 2, p. plantgenome2016.11.0111, 2017.

THOMPSON, Robin *et al.* A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. **Australian & New Zealand Journal of Statistics**, Oxford, v. 45, n. 4, p. 445-459, 2003.

VAN EEUWIJK, Fred A.; BUSTOS-KORTS, Daniela V.; MALOSETTI, Marcos. What should students in plant breeding know about the statistical aspects of genotype× environment interactions? **Crop Science**, Madison, v. 56, n. 5, p. 2119-2140, 2016.

VANHATALO, Jarno; LI, Zitong; SILLANPÄÄ, Mikko J. A Gaussian process model and Bayesian variable selection for mapping function-valued quantitative traits with incomplete phenotypic data. **Bioinformatics**, Oxford, v. 35, n. 19, p. 3684-3692, 2019.

VELAZCO, Julio G. *et al.* Modelling spatial trends in sorghum breeding field trials using a two-dimensional P-spline mixed model. **Theoretical and Applied Genetics**, Berlin, v. 130, n. 7, p. 1375-1392, 2017.

WICKHAM, Hadley *et al.* Package ‘ggplot2’. **Create Elegant Data Visualisations Using the Grammar of Graphics**, [S.l.], v. 2, n. 1, p. 1-189, 2016.

WOLC, Anna *et al.* Evaluation of egg production in layers using random regression models. **Poultry science**, v. 90, n. 1, p. 30-34, 2011.

YANG, Runqing; TIAN, Quan; XU, Shizhong. Mapping quantitative trait loci for longitudinal traits in line crosses. **Genetics**, Austin, v. 173, n. 4, p. 2339-2356, 2006.

Artigo aceito para publicação - DOI: 10.1002/tpg2.20306

GENOMIC PREDICTION FOR COMPLEX TRAITS ACROSS MULTIPLE HARVESTS IN ALFALFA (*Medicago sativa* L.) IS ENHANCED BY ENVIROMICS

ABSTRACT

Breeding for dry matter yield (DMY) and persistence in alfalfa (*Medicago sativa* L.) can take several years as traits must be evaluated under multiple harvests. Furthermore, DMY and persistence are complex quantitative traits that exhibit low to moderate heritability and high genotype by environment interaction (G×E). In this study, we measured persistence based on the phenotypic plasticity for DMY, i.e., the slope of the DMY over time regression. In this study, we investigated the implementation of environmental covariates into genomic prediction schemes for DMY, persistence, adaptability and stability in 177 alfalfa families evaluated across 11 harvests. Four cross-validation scenarios were tested: (i) predicting tested families in observed harvests (CV2); (ii) predicting untested families in observed harvests (CV1); (iii) predicting tested families in unobserved harvest (CV0); and (iv) predicting untested families in unobserved harvests (CV00). All scenarios were analyzed using six models (M0 to M6) in a Bayesian mixed model framework, and models were compared based on the Pearson correlation between predicted and observed values (predictive ability – PA). Our results demonstrate that models which accounted for enviromic data led to higher PA in a reduced number of harvests used in the training data set under CV0 and CV2. The Bayesian framework allowed to model G×E even using only one harvest in the training set. Models that accounted for enviromic data (M2 and M3) outperformed the base model (M0) for predicting adaptability across all scenarios. Models M2 and M3 also provided higher PA for persistence compared to M0, as predictions increased from 0 to 0.16, 0.20, 0.56 and 0.46 for CV00, CV1, CV0 and CV2. For DMY, there was a slight increase in PA between M0 and M2/M3 under CV1 from 0.22 to 0.24, and bigger increase for CV0 and CV2 (0.53 to 0.61, and 0.52 to 0.63 respectively). The results also demonstrate that GBLUP without enviromics has low power to predict persistence, thus the adoption of enviromics is a cheap and efficient alternative to increase accuracy and biological meaning.

1 INTRODUCTION

Alfalfa (*Medicago sativa* L.) is an herbaceous perennial forage crop grown throughout temperate and sub-tropical regions of the world (BRUMMER, 2003; LI et al., 2015; ANNICHIARICO et al., 2015). Alfalfa is allogamous, pollinated by bee species, and is characterized by inbreeding depression. Most cultivated alfalfa is a tetrasomic tetraploid ($2n=4x=32$) and cultivars are synthetic populations consisting of highly variable, heterozygous plants. Alfalfa growth, especially in late summer into autumn, is affected by fall dormancy, which refers to growth reduction and decumbent shoot orientation that typically occurs in late summer and early autumn as temperature declines and photoperiod shortens (CASTONGUAY

et al., 2006; MCKENZIE et al., 1988). Nondormant cultivars are widely adopted in the southern United States and other subtropical regions (BOUTON, 2012). In Florida, nondormant cultivars were developed for improved adaptation and greater dry matter yield (DMY) ['Florida 66' (HORNER, 1970), 'Florida 77' (HORNER; RUELKE, 1981), and 'Florida 99']. Breeding efforts are underway to develop new nondormant alfalfa cultivars (ACHARYA et al., 2020; ADHIKARI et al., 2019; ANDRADE et al., 2022; BISWAS et al., 2021). Alfalfa is used to supply forage production in the early spring and late fall in Florida, when warm-season perennial forages are not actively growing. Therefore, the breeding program at the University of Florida focuses in selecting nondormant germplasm that normally exhibits low persistent (ACHARYA et al., 2020; HOPPEN et al., 2019; HORNER, 1970; HORNER; RUELKE, 1981).

Alfalfa breeding is typically conducted as phenotypic recurrent selection using among and within half-sib family selection (CASLER; BRUMMER, 2008), although various breeding schemes have been proposed to improve DMY (ANNICCHIARICO; PECETTI 2021). Genetic gain for DMY has stagnated (BRUMMER; CASLER, 2014), although both DMY and quality improvement have been targeted recently (ACHARYA et al., 2020; ANDRADE et al., 2022; BISWAS et al., 2021; ADHIKARI et al., 2019; DOS SANTOS et al., 2018; SAKIROGLU; BRUMMER, 2017). Dry matter yield, quality, and persistence are complex and quantitatively inherited traits, and exhibit moderate to low heritability and high genotype by environment interaction ($G \times E$) (ACHARYA et al., 2020; BOWLEY; CHRISTIE, 1981; BRUMMER; CASLER, 2014; RIDAY; BRUMMER, 2004). Due to the perennial behavior in alfalfa, selection for DMY is usually based on measurements taken across multiple harvests (DE ASSIS et al., 2010; FAVERI et al., 2015), which leads to longer selection cycles, high evaluation costs, and consequently lower genetic gains (ANNICCHIARICO et al., 2010). Stand persistence is critical for alfalfa and it can be defined as the capacity of alfalfa plants to survive over time (BOUTON, 2012). Assessing persistence takes several harvests and must be evaluated over multiple years. Persistence is a complex trait as it depends on several plant characteristics and environmental factors such as drought, temperature, grazing/harvest pressure, aluminum-toxicity tolerance, fall dormancy, disease resistance, among other factors (IRWIN, 1977; LEACH; CLEMENTS, 1984; RIMI et al., 2014; HOPPEN et al., 2019). Due to the complexity to evaluate persistence, De Assis et al. (2010) proposed an indirect method to measure persistence by regressing yield over time. Therefore, persistence is estimated through the regression coefficient, and it is expected to be negative as the plant stand decreases over time.

Due to the low heritability estimates, high G×E effects, long and expensive selection cycles for DMY and persistence, it is critical to use selection procedures based on genomic information (marker SNPs) as a partial substitute of phenotypic selection (LI; BRUMMER, 2012). Recently, genomic selection (GS) has been applied in alfalfa for DMY (ANNICCHIARICO et al., 2015; LI et al., 2015; JIA et al., 2018), and nutritional value traits (BIAZZI et al. 2017). All these studies were based on measurements of phenotypic and/or genotypic data at the single plant level. More recently, Andrade et al. (2022) reported predictive ability (PA) for DMY and canopy height in alfalfa family bulks. However, none of these previous studies modeled environmental covariates of G×E interactions.

For any genotype-phenotype association study across multiple environments, there are strong nongenetic influences that can be taken into account in genomic prediction models to increase PA, and to reduce the number of harvests to be included in training populations. To understand the environmental impacts on crop performance, the concept of “envirotyping” is proposed as a third “typing” technology, complementing with genotyping and phenotyping (Xu, 2016). Envirotyping-based data consists in collecting, processing, and integrating environmental information in genetic and genomic studies (COSTA-NETO et al., 2020). The genotypes (G) determine the yield potential can be investigated by molecular markers (MORRELL et al., 2012; YUAN et al., 2017). For alfalfa, the phenotypes (P) can be obtained by collecting ground-based DMY (ACHARYA et al., 2020), and by the development of high-through-put phenotyping tools and methodologies (ARAUS; CAIRNS, 2014; BISWAS et al., 2021).

Recently, the environmental information, has been incorporated into genomic prediction models through envirotyping (E), (JARQUÍN et al., 2014; COSTA-NETO et al., 2020, 2021a and 2021b). The envirotyping aiming to increase the PA under multi-environmental trials by including the interaction between genomic features (G) and environmental factors (E) under different cross validation schemes (JARQUÍN et al., 2017) mimicking prediction scenarios of interest for breeders. All the previous implementations attempt to aid breeders to make more informed selections of outperforming genotypes when there are no available phenotypic records in a given environment, and/or to select genotypes that have never been phenotyped (tested) at any environment.

Some of the proposed models for including G×E effects were based on the genotype response over an environmental gradient (reaction norm) characterized by the environmental deviation from the mean value to determine the environmental quality (FINLAY;

WILKINSON, 1963). These models are based on a simple linear regression and can fail in predicting phenotypes under more complex G×E patterns, and the establishment of the environmental gradient in untested environment is difficult without any information about the environment. Costa-Neto et al (2021b) proposed the use of envirotyping-based data in the prediction of the environmental quality for the Finlay-Wilkinson model.

Authors found high correlations between observed and predicted environmental quality, and proposed that inferences about genomic plasticity can be made in untested environments. Gauch (1988) proposed the additive main effects and multiplicative interaction (AMMI) models, which consists of decomposing the multiplicative part of the G×E by principal component analysis, in order to capture the G×E patterns in a reduced number of latent variables. These latent variables (environmental loadings) can be correlated after with environmental variables in order to select environmental variables to be used in envirotyping-based models (RINCENT et al., 2019). Piepho (1998) and Smith et al. (2005) proposed factor analytic models to decompose the multiplicative interaction in a mixed model context, and this model allows the incorporation of genomic information, and the environmental loadings can be correlated with environmental covariates (ECs) to infer the importance of those ECs in the phenotypic expression. Furthermore, factor analytic models allow the interpretation about adaptability and stability by latent regressions (CULLIS et al., 2014; SMITH; CULLIS, 2018). All those models aim at explaining GE effects by using latent variables based on the measured phenotypic data records, and then inferences about the ECs can be made after predicting environmental loadings; therefore, predicting genomic estimated breeding values (GEBVs) in untested environments can be difficult. Jarquín et al (2014) used ECs to create envirotyping-based kinship matrices, enabling the establishment of putative environmental similarities that may drive a large amount of phenotypic variation, thus the information about the similarities between environments allow the prediction of genotypes in untested environments. To build the environmental kinship matrix, Costa-Neto et al. (2021a), developed an R package (*EnvRtype*). *EnvRtype* handles a very robust envirotyping pipeline for collecting, processing, computing environmental kinship matrices by using linear and nonlinear kernels, and analyzing enviromic-genomic models by using the Bayesian Genomic Genotype × Environmental Interaction R package (BGGE, GRANATO et al., 2018).

Previous studies in genomic prediction by using envirotyping information were performed in annual crops. In this study, we investigated the use of envirotyping-based data for alfalfa in order to estimate the PA of genomic-enviromic models for dry matter yield, by using

four cross-validation scenarios as proposed by Jarquín et al (2017): (i) predicting tested families in observed harvests (CV2); (ii) predicting untested families in observed harvests (CV1); (iii) predicting tested families in unobserved harvest (CV0); and (iv) predicting untested families in unobserved harvests (CV00). All these scenarios were evaluated by increasing the number of harvests used in the training population, to estimate the optimal number of harvests to accurately predict genomic estimated breeding values (GEBV) in each harvest for DMY, persistence, broad adaptability and stability based in the regression slope (b_1), intercept (b_0) and R^2 of the GEBVs over time, respectively. Two types of environmental kinship were also evaluated, one based in climate data during the field study, and another based on historical data from the last 34 years, in order to evaluate if historical climate data can be used in alfalfa genomic prediction.

2 MATERIAL AND METHODS

2.1 Reference population

The population consisted of 177 families (142 full-sib and 35 half-sib) derived from crosses between parental lines coming from 33 populations, from various germplasm sources selected based on high dry matter yield (DMY) and persistence (ACHARYA et al., 2020). The crosses were performed in a factorial mating design. Six male parents adapted to Florida conditions were used (four genotypes from the UF breeding program, and two commercial cultivars: “Bulldog 805” and “AmeriStand 915”), and 27 selected genotypes were used as female parents. All crosses were conducted as described by Acharya et al. (2020).

2.2 Genotypic data

The DNA extraction were done at family level as described by Andrade et al (2022) by bulking one leaf from 30 individuals belonging to each family. Genomic DNA was isolated using the DNeasy Plant Mini Kit (Qiagen) and quantified with a Quant-iT PicoGreen dsDNA assay kit (Life Technologies, P7589). The genotyping process was performed at Rapid Genomics, LCC (Gainesville, FL), employing targeted hybridization. For this, 30,000 120-mer biotinylated probes were designed based on the Cultivated Alfalfa at the Diploid Level genome (CADL), and first screened against the same genome, CADL. After this, the probes sequences

were further searched against a tetraploid alfalfa assembly (Dr. Maria Monteros, Noble Research Institute personal communication), and only probes that hit four or fewer times in the genome were kept. Finally, the probes were screened against the alfalfa chloroplast genome (KU321683.1), to avoid probes that might be associated with non-nuclear DNA. Finally, 17,707 target probes were used in the bulked family DNA samples, and sequencing was performed using the Illumina HiSeq 1x100 platform.

Subsequently the raw reads were filtered and trimmed using Trimmomatic v.0.39 (BOLGER et al., 2014). Filtered reads were mapped against the longest scaffolds of each of the eight homologous groups in the genome assembly of *Medicago sativa* cv. ‘XinJiangDaYe’ (CHEN et al., 2020) using the BWA v.0.7.17 software (LI; DURBIN, 2009). The alignment files were processed using SAMtools v.1.10 (LI et al., 2009) to convert format and sort files, and Picard v.2.21.2 (<http://picard.sourceforge.net/index.shtml>) to add groups and remove PCR duplicates. SNPs were called using FREEBAYES v.1.3.2 (GARRISON; MARTH, 2012).

SNPs were further filtered considering the following criteria: minimum base quality of 20; minimum mapping quality of 30; only biallelic locus; no monomorphic locus; maximum missing data of 10%; mean depth value over all samples greater than or equal 30; the ratio between the mapping quality of the alternative and the reference allele between 0.95 and 1.05; minor allele frequency bigger than 0.05 and smaller than 0.95. After filtering, 114,945 SNPs were considered for the downstream population genetics and genomic prediction analyses.

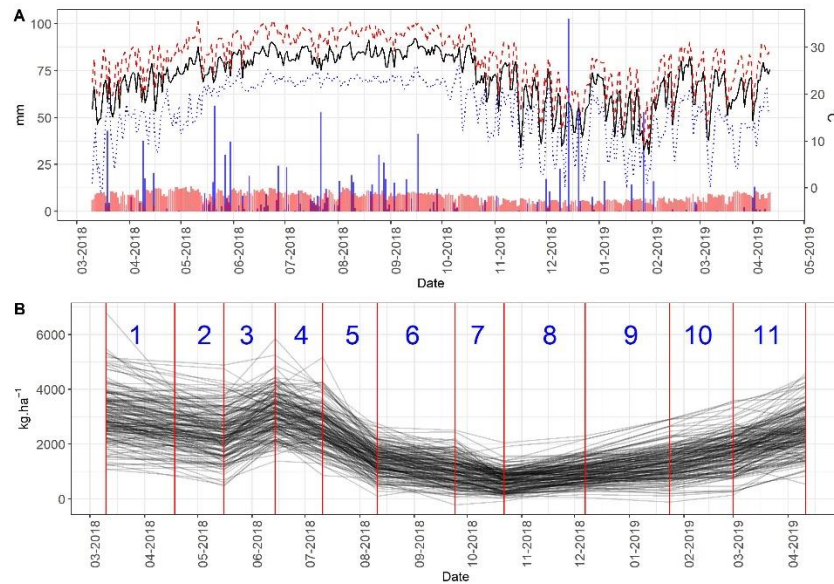
Alternative and reference read counts for each SNP and sample were extracted from the variant calling file using VCFtools v.0.1.16 (DANECEK et al., 2011). Genotype parametrization was performed considering the allele ratio $\#A/(\#A + \#a)$, where $\#A$ is the allele count of the alternative allele and $\#a$ is the allele count of the reference allele; in this parametrization, the dosage calling step was not performed and the data varied continuously from 0 to 1 as described in de Bem Oliveira et al. (2019, 2020).

2.3 Experimental design and phenotyping

The field trial was established at the UF/IFAS Plant Science Research and Education Unit (PSREU), Citra, FL (29°24'16'' N and 82°10'17'' W) and the phenotypic data were collected across 11 harvests from April 2018 to March 2019 (Figure 1). The trial was designed as an augmented row and column with three controls, the cultivars “Bulldog 805” and “Florida 99”, and the breeding line from UF breeding program “UF_AlfPers_2015”. Each plot was

composed by twenty seedlings per family, planted in two rows (1.82 x 0.12m). The cultural practices were done as described by Acharya et al (2020).

Figure 1 - (A) Weather information for the period from first (March 2018) to last (May 2019) harvest.



Blue bars represent the daily precipitation (mm); red bars are the daily evapotranspiration (mm); red dashed line is the maximum daily temperature (°C); black solid line is the mean daily temperature (°C); and the blue dashed line is the minimum daily temperature (°C). (B) Dry matter yield (DMY) of families collected across eleven harvests.

Plots were manually harvested when the control “UF_AlfPers_2015” reached 10% blooming, using a sickle, and total fresh weight was recorded. From the whole plot, a sample of approximately 500g was collected and it was dried at 65°C for seven days, to estimate DMY (kg.ha⁻¹).

2.4 Weather data collection

We obtained two sets of historical environmental data to compute environmental kinship. The first set was obtained from the Florida Automated Weather Network (FAWN) (<https://fawn.ifas.ufl.edu>), that is located at the PSREU (29°48'9''N; 82°24'39''W), in the period when the trial was conducted (03/10/2018 – 04/11/2019) (Figure 1). The second data set

was obtained by the satellite-based weather system named “NASA’s Prediction of Worldwide Energy Resources” (NASA POWER, <https://power.larc.nasa.gov/>), using the *EnvRtype* package (COSTA-NETO et al., 2021a) considering 34 years (1984 – 2017) by using the *get_weather()* function. The goal was to compare the two similarity kernels in the prediction models and find out if the generalization of the environmental kinship matrix based on a large historical data set (34 years), can be employed in genomic prediction in alfalfa. For the first data set from FAWN station we considered 12 weather variables: soil mean temperature (ST, °C.day⁻¹), daylight mean temperature (DTMEAN, °C.day⁻¹), daylight maximum temperature (DTMAX, °C.day⁻¹), daylight minimum temperature (DTMIN, °C.day⁻¹), night mean temperature (NTMEAN, °C.day⁻¹), night maximum temperature (NTMAX, °C.day⁻¹), night minimum temperature (NTMIN, °C.day⁻¹), dew point temperature (DP, °C.day⁻¹), relative humidity (RH, %), photoperiod (PP, h.day⁻¹), rain fall precipitation (PREC, mm.day⁻¹), solar radiation (SRAD, MJ.m⁻².day⁻¹). For the data obtained from NASA power, we considered the same variables except for ST, NTMEAN, NTMAX and NTMIN.

2.5 Weather data processing

The raw data from the two data set were used to calculate variables related to the eco-physiological interactions between soil, plant and atmosphere. The potential evapotranspiration (ETP, mm.day⁻¹) calculation was based on the Priestley-Taylor method (PRIESTLEY; TAYLOR, 1972). The slope of the curve of saturation vapor pressure (SVP) was calculated according to the Food and Agriculture Organization (FAO) manual (ALLEN et al., 1998). The difference between rainfall precipitation and crop evapotranspiration (PETP, mm⁻¹.day) was calculated by the difference between PREC and ETP. The effect of temperature on radiation uses efficiency (FRUE) (SOLTANI; SINCLAIR, 2012) and growing degree days (GDD) were calculated based on temperature cardinal points. The temperature cardinal points used in this study were the same as proposed by Malik et al. (2018) for alfalfa in the growing conditions of Northeast Spain: 3°C (Tb₁, base lower), 25°C (To₁, base optimum), 33°C (To₂, upper optimum) and 45°C (Tb₂, base upper). The weather variables above were calculated by using the function *processWTH()* using the R package *EnvRtype* (COSTA-NETO et al., 2021a).

By using the processed variables, we estimated the leaf area index (LAI) as a function of thermal-time accumulation by assuming an optimum leaf area expansion rate (LAER_{opt}) of 0.016 m².m⁻².degree-day (TEIXEIRA et al., 2007). Using the values of LAI for each day until

the harvest, intercepted photosynthetically active radiation (PAR_i) was also estimated, considering a single coefficient of diffuse PAR (kd) of 0.81 (TEIXEIRA et al., 2007) by the expression: $PAR_i = 1 - e^{-0.81 * LAI}$. The air-soil temperature ratio (ASR) was calculated once/after this variable is related with the partitioning between shoots and roots (TEIXEIRA et al., 2008; Engels, 1994). The variables LAI, PAR_i and ASR were used only to compute the similarity kernel based on the data from FAWN.

2.6 Computing the environmental covariable matrices (E and Eg)

For the first set of environmental data (FAWN), each variable was sampled within each environment. The environment was represented by the weather conditions between consecutive harvests (Figure 1A and 1B). In each environment, time intervals (growth stages) were determined as: 0 to 7, 7 to 14, 14 to 21, 21 to 28, 28 to 35 and 35 to 42 days after harvest. For each weather variable we established thresholds based on the Shelford's Law of Tolerance (SHELFORD, 1931) as suggested by Costa-Neto et al. (2021b). Furthermore, each weather variable was split in three thresholds: (i) stress by deficit, (ii) optimum growing conditions, and (iii) stress by excess, these thresholds were established according to the literature (Table 1). For variables without a clear threshold, we performed the discretization using histogram of percentiles (0-25, 26-50, 51-75 and 75-100%) as suggested by Costa-Neto et al (2021b). Finally, to compute the environmental covariable matrix (E) (Figure 2A) we considered the frequency of days within each environment for the three thresholds for each growth stage as an environmental covariable (EC).

For the historical data set, we assigned as an environment a two-month interval time, by overlapping the last month of an environment (harvest) to the first month of the next harvest. This was done using weather data across 34 years. The historical weather data was used to compute a generalized matrix of environmental covariates (Eg) (Figure 2B). This matrix intends to mimic the conditions that plants could encounter in future trials, since breeders do not have in advance the weather data for future trials and the exact dates for planting and harvesting. After processing the historical weather data, we computed the Eg matrix in the similar way as was done for the E matrix, but for computing Eg we set time intervals of 15 days (0 to 15, 15 to 30, 30 to 45 and 45 to 60). After filtering by the variance greater than zero for each EC, we retained 388 and 174 ECs for E and Eg respectively.

Figure 2 - (A) Environmental covariable matrix E , and (B) generalized environmental covariable matrix Eg . The columns represent each combination of growth stage and weather variable threshold, and the rows represent each harvest.

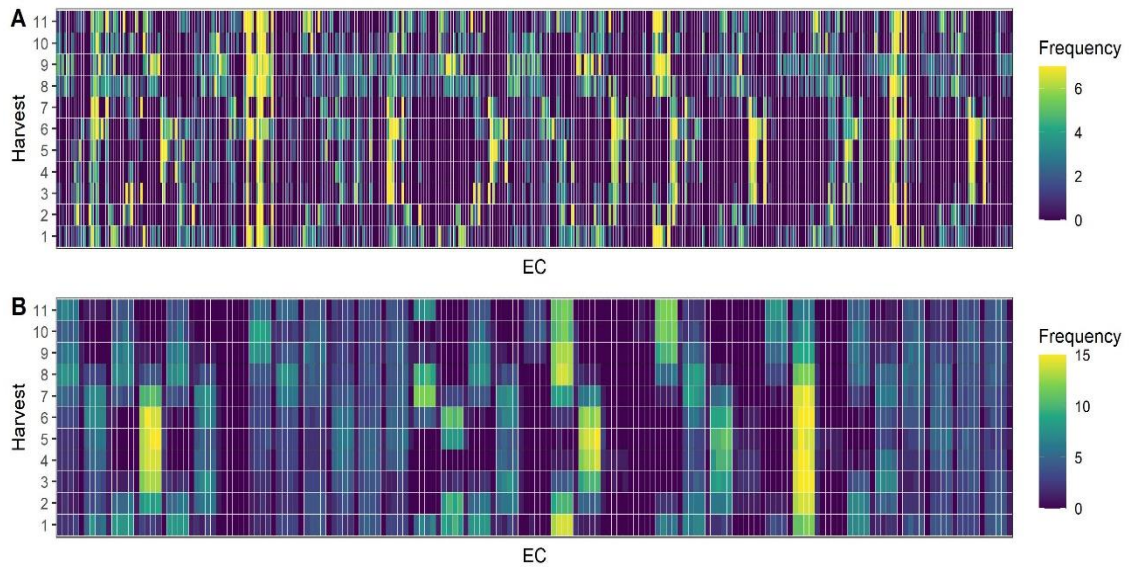


Table 1 - Weather variable thresholds based on the Shelford's Law of tolerance for daylight temperature, night temperature, photoperiod, PAR, LAI, and air-soil ratio temperature.

Variable	Stress by deficit	Lower sub-optimum	Optimum	Higher sub-optimum	Stress by excess	Reference
Daylight temperature	< 10°C	10°C to 24°C	25°C to 30°C	31°C to 34°C	> 34°C	Patterson (1993) Collino et al. (2005) Malik et al. (2018)
Night temperature	< 5°C	5°C to 14°C	15°C to 20°C	21°C to 25°C	>25°C	Robison et al. (1969) Patterson (1993)
Photoperiod	No leaf expansion	Linear leaf expansion		Maximum leaf expansion		Teixeira et al. (2007)
	< 10h	10h to 12.5h		>12.5h		
PAR	Linear interception		Maximum interception			Teixeira et al. (2007)
	<0.81		> 0.81			
LAI	< 3		> 3			
air-soil ratio temperature	No partitioning to the roots	Linear partitioning to the roots		Maximum partitioning to the roots		Teixeira et al. (2008)
	<0.8	0.8 to 1.3		>1.3		
PREC	Threshold established as quantile from data set					-
ETP	Threshold established as quantile from data set					-
PETP	Threshold established as quantile from data set					-
SRAD	Threshold established as quantile from data set					-
SPV	Threshold established as quantile from data set					-
GDD	Threshold established as quantile from data set					-
FRUE	Threshold established as quantile from data set					-

2.7 Computing environmental and genomic similarity kernels (K_E, K_{Eg}, K_G)

The environmental similarity kernels (K_E (1) and K_{Eg} (2)) were computed based on the linear variance-covariance matrix as proposed by Jarquín et al. (2014) and the genetic relationship kernel (K_G (3)) was computed as suggested by Bem Oliveira et al (2020):

$$K_E = \frac{EE'}{\text{tr}(EE')/r_E} \quad (1);$$

$$K_{Eg} = \frac{EgEg'}{\text{tr}(EgEg')/r_{Eg}} \quad (2); \quad K_G = \frac{MM'}{\text{tr}(MM')/r_M} \quad (3).$$

where, K_E and K_{Eg} are the environment and generalized environment similarity kernels between harvests, respectively; E and Eg are the environmental co-variance matrices (Figure 2) and their entries describe environmental similarities between pairs of harvests; r_E and r_{Eg} are the number of rows in the matrices E and Eg , respectively. K_G is the kinship matrix and its entries describe genomic similarities between pairs of genotypes; Z is the mean centered matrix of SNP frequency and r_M is the number of rows in the M matrix.

2.8 Statistical models

We obtained the best linear unbiased estimates (BLUEs) of the families at each harvest, using the following linear mixed model:

$$y = X\beta + Z_1r + Z_2c + e \quad (4),$$

where, y is the vector of data, β is the vector of the fixed effects of the intercept and families, r and c are the random effects of the rows and columns following independent and identically distributed (IID) normal densities such that $r \sim N(0, I\sigma_r^2)$ and $c \sim N(0, I\sigma_c^2)$; σ_r^2 and σ_c^2 are the variance components associated with the effect of rows and columns; and e is the vector of the residuals, with $e \sim N(0, I\sigma_e^2)$; σ_e^2 is the variance component of the residuals. X is the incidence matrix of the fixed effects; Z_1 and Z_2 are the incidence matrix of the effects of rows and columns, respectively; and I is an identity matrix.

After adjusting the families mean for each harvest by the design effects using the model (4), we tested six genomic prediction models and assess the effects of accounting for the family by harvest interaction term: (i) Model without account to the $G \times E$ effects (G-BLUP – M0), considered as baseline model; (ii) Model accounting to the $G \times E$ effects modeled as a block

diagonal matrix, without any environmental similarity kernel (GE-BLUP – M1); (iii) model with environmental kernel K_E , and families by harvests interaction (G×E) modeled as a block diagonal matrix (GE-BLUP- K_E – M2); (iv) model with environmental kernel K_{Eg} , and G×E interaction modeled as a block diagonal matrix (GE-BLUP- K_{Eg} – M3); (v) model with environmental kernel K_E , and G×E modeled as a Kronecker product between K_{Eg} and K_G (RN-BLUP- K_E – M4); and (vi) model with environmental kernel K_E , and G×E modeled as a Kronecker product between K_{Eg} and K_G (RN-BLUP- K_E – M5). All the kernel models were fitted using the BGGE R package (GRANATO et al., 2018) using 20,000 iterations, with 1,000 used as burn-in and using a thinning of 10. Models are described in the following sections.

2.8.1 G-BLUP Model (M0)

The model used as a baseline (M0) includes a fixed intercept for each harvest and random additive genetic effects. In this model, only the additive genetic effects were modeled by the use of genomic relationship matrix (K_g):

$$y = 1\mu + X_E\beta + Z_A u_A + \varepsilon \quad (6)$$

where, y is the vector of the families' BLUEs in each q harvest; μ is the intercept; β is the vector of fixed effect of harvests; u_A is the additive genetic effects, with $u_A \sim N(0, J_q \otimes K_G \sigma_A^2)$, where σ_A^2 is the variance component associated with the additive genetic effects; ε is the residual deviations assumed as $\varepsilon \sim N(0, I_n \sigma_\varepsilon^2)$. X_E and Z_A are incidence matrices of the fixed effects β and random effects u_A , respectively; J_q is a $q \times q$ matrix of 1s; I_n is an identity matrix of size n , where n is the number of observations; and \otimes denotes a Kronecker product.

2.8.2 GE-BLUP Model (M1)

The second model (M1) includes a fixed intercept for each harvest, random additive genetic effects, random harvest effects and, random additive genetic by harvest interaction. In this model, the interaction was modeled as the main effect of families plus genomic-by-harvest deviation (G + GE model) as follows:

$$y = 1\mu + X_E\beta + Z_A u_A + u_E + u_{AE} + \varepsilon \quad (7)$$

u_E is the main random effect of harvests with $u_E \sim N(Z_E\beta, I_q \otimes J_p\sigma_E^2)$, where J_p is a $p \times p$ matrix of 1s; u_{AE} is the additive genetic by harvest interaction effects modeled using a block diagonal matrix of the additive effects, built as $u_{AE} \sim N(0, I_q \otimes K_G\sigma_A^2)$; ε is the residual deviations assumed as $\varepsilon \sim N(0, I_n\sigma_\varepsilon^2)$.

2.8.3 GE-BLUP- K_E (M2) and GE-BLUP- K_{Eg} (M3)

The model accounting for the environmental similarities (K_E or K_{Eg}) can be described in the same way as equation (7) by adding a main environmental relatedness effect:

$$y = 1\mu + X_E\beta + Z_A u_A + u_E + u_{AE} + \varepsilon \quad (8)$$

with $u_E \sim N(Z_E\beta, K_E \otimes J_p\sigma_E^2)$, where K_E is the environmental similarity kernel; and σ_E^2 is the variance component from the environmental co-variance matrix E . For the model M3 we substituted the K_E by K_{Eg} kernel. In this model, the effects of G×E are modeled as a block diagonal (GE +E) as in the GE-BLUP model. Therefore, if no information about the environment is provided the expected value for harvests is given by $Z_E\beta$ as the G-BLUP model (COSTA-NETO et al., 2021a).

2.8.4 RN-BLUP- K_E (M4) and RN-BLUP- K_{Eg} (M5)

The model (8) can be expanded by modeling the G×E interactions as reaction norms (RN) (JARQUÍN et al., 2014) based on the Kronecker product between the environment and genomic kernels (MARTINI et al., 2020):

$$y = 1\mu + X_E\beta + Z_A u_A + u_E + u_{AE} + \varepsilon \quad (9)$$

with, $u_{AE} \sim N(0, K_E \otimes K_G\sigma_{AE}^2)$, where σ_{AE}^2 is the variance component associated with enviromic by additive genetic effects interaction. For the model M5 we replaced K_E by K_{Eg} .

2.9 Persistence, adaptability and stability evaluation

For each model, we adjusted a linear regression of the predicted GEBVs for each family over time, adjusted by the seasonality:

$$y_{ijk} = s_k + b_{0i} + t_j b_{1i} + e_{ijk} \quad (10)$$

where, y_{ijk} is the GEBVs for DMY of the family i in the time j and season k ; s_k is the effect of season k (spring, summer, fall and winter); b_{0i} is the intercept of the family i (adaptability); t_j is the time j , denoted as a regressor variable given by days after the first harvest (harvest 1 – 0 to harvest 11 – 396 days); and b_{1i} is the regression coefficient of the family i (persistence, DE ASSIS et al., 2010); e_{ijk} is the residual deviation, with $e_{ijk} \sim N(0, I\sigma_e^2)$. To measure the stability, we estimated the determination coefficient R^2 of the linear regression for each family in which reflects the type 3 stability, when the stable genotype has smaller deviation from the regression model, i.e. it has greater predictability (LIN et al., 1986).

2.10 Cross-validation and predictive ability

The cross-validations (CVs) were done by assigning, randomly, 70% of the families as training data set and the other 30% families were set as the validation set. Furthermore, we set a constraint on the CV given the presence of population structure in the breeding population, by sampling 70% of the families within each of the six genetic structured groups (ANDRADE et al., 2022). The CVs were repeated ten times, by assigning different training and testing populations and increasing the number of harvests used in the training set (Figure 3).

For the CV, we established four different scenarios (Figure 3) according to Jarquín et al. (2017):

- (i) CV2: predicting tested families in observed harvests;
- (ii) CV1: predicting untested families in observed harvests;
- (iii) CV0: predicting tested families in unobserved harvests, and;
- (iv) CV00: predicting untested families in unobserved harvests.

The PA was assessed in all the CVs scenarios by calculating the Pearson's correlation coefficient between predicted (GEBVs) and observed DMY values (BLUEs from model [4]). The CVs were also performed for the regression parameters b_0 , b_1 and R^2 , in which the observed values for b_0 , b_1 and R^2 were estimated by regressing on the BLUEs (4) over time by using the regression model (10). The standard error for the correlation (SE) was estimated following the expression presented by Carvalho et al (2020):

$$SE = SD \left(\frac{\sum \sqrt{\frac{1}{Nfp} + \frac{Nftrn}{Nfval}}}{n_t} \right) \quad (11)$$

where, SD is the standard deviation of the genomic-environmental prediction, Nfp is the size of the whole population; $Nftrn$ is the size of the training population; $Nfval$ is the size of the validation population; and n_t is the number of times each CV was sampled (10 times).

Figure 3 - Example for the CV schemes (CV0, CV00, CV1 and CV2) using only three harvests (H1, H2 and H3), three genotypes (G1, G2 and G3) and three folds (F1, F2 and F3), by increasing the number of harvests from one to two. The reduced number of harvests, genotypes and folds were used only for simplicity.

CV0								CV1							
Genotype	1 Harvest			2 Harvests			Fold	Genotype	1 Harvest			2 Harvests			Fold
	H1	H2	H3	H1	H2	H3			H1	H2	H3	H1	H2	H3	
G1	1	0	0	1	1	0	F1	G1	1	0	0	1	1	0	F1
G2	1	0	0	1	1	0	F1	G2	0	0	0	0	0	0	F1
G3	1	0	0	1	1	0	F1	G3	1	0	0	1	1	0	F1
G1	0	1	0	0	1	1	F2	G1	0	0	0	0	0	0	F2
G2	0	1	0	0	1	1	F2	G2	0	1	0	0	1	1	F2
G3	0	1	0	0	1	1	F2	G3	0	1	0	0	1	1	F2
G1	0	0	1	1	0	1	F3	G1	0	0	1	1	0	1	F3
G2	0	0	1	1	0	1	F3	G2	0	0	1	1	0	1	F3
G3	0	0	1	1	0	1	F3	G3	0	0	0	0	0	0	F3

CV00								CV2							
Genotype	1 Harvest			2 Harvests			Fold	Genotype	1 Harvest			2 Harvests			Fold
	H1	H2	H3	H1	H2	H3			H1	H2	H3	H1	H2	H3	
G1	1	0	0	1	1	0	F1	G1	1	0	0	1	1	0	F1
G2	0	0	0	0	0	0	F1	G2	0	0	0	0	0	0	F1
G3	1	0	0	1	1	0	F1	G3	1	0	0	1	1	0	F1
G1	0	0	0	0	0	0	F2	G1	0	0	0	0	0	1	F2
G2	0	1	0	0	1	1	F2	G2	0	1	0	0	1	1	F2
G3	0	1	0	0	1	1	F2	G3	0	1	0	0	1	1	F2
G1	0	0	1	1	0	1	F3	G1	0	0	1	1	0	1	F3
G2	0	0	1	1	0	1	F3	G2	0	0	1	1	0	1	F3
G3	0	0	0	0	0	0	F3	G3	0	0	0	0	0	0	F3

	Training		Testing		0	Unobserved		1	Observed
--	----------	--	---------	--	---	------------	--	---	----------

Finally, we adjusted the Pearson's correlation coefficients over the number of harvests used in the training data set by using an asymptotic non-linear regression adjusted by the R function $nls()$, in order to estimate the optimal number of harvests after which PA will not increase significantly:

$$r_{ij} = a_i - (a_i - b_i) * e^{(-c_i X)} + \varepsilon_{ij} \quad (12)$$

where, r_{ij} is the Pearson correlation between predicted and observed values for model i with j number of harvests in the training set; a_i is the asymptotic point for model i ; b_i is the correlation where X equals to zero; c_i is the proportional in PA gains over number of harvests in the training set; X represents the number of harvests in the training set; and ε is the error. The optimal

number of harvests were calculated based on the inverse function of the equation (12), in which the lower limit of the asymptotic point (a_i) 95% confidence interval.

3 RESULTS

3.1 Genetic and non-genetic parameters

The inclusion of the environmental effects in Models 2 to 5 led to a better partition of the phenotypic variance, as the error variance (σ_ϵ^2) was decreased and the genetic variance (σ_A^2) increased, compared to Model 1 (Table 2). Models that accounted for the environmental relatedness kernel (M2, M3, M4 and M5) were able to capture more variance from the environment effects. These results showed that enviromic effects were an important component of the phenotypic variance (Table 2). When comparing the models M2 and M3 to the models M4 and M5, the inclusion of a more complex variance structure for including the G×E interaction term led to an increase of the variance component associated with it (i.e., σ_{AE}^2), indicating that the reaction norm kernel was more efficient in capturing the G×E pattern present in the phenotypic data (Table 2).

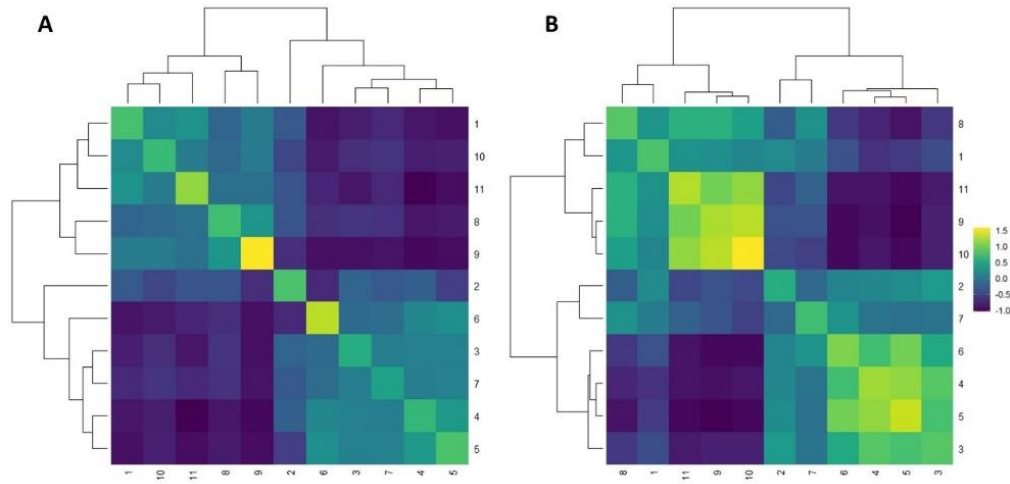
Table 2 - Summary of the variance components for the five models evaluated for alfalfa dry matter yield. The variance components were estimated by fitting the models with all recorded phenotypes across all harvests ($p = 177$ families and $q = 11$ harvests).

Variance component	Model					
	M0	M1	M2	M3	M4	M5
	441,448	741,648	736,784	747,041	765,795	759,882
	-	1,504,121	2,478,132	3,995,112	4,349,691	3,634,176
	-	122,295	123,572	121,919	246,610	230,930
	996,763	169,005	169,023	169,461	146,128	155,408

σ_A^2 , σ_E^2 , σ_{AE}^2 and σ_ϵ^2 are the variance components of the additive genomic effects, harvests effects, additive genomic by harvests interaction effects and error, respectively.

The magnitude of the σ_A^2 , σ_{AE}^2 and σ_ϵ^2 did not change significantly between models with different enviromic kernels (M2 and M3; M4 and M5) (Table 2). These results can be explained by the similar pattern of the environmental relatedness explained by K_E and K_{Eg} , in which the enviromics effects grouped the harvests in two main clusters, the first group constituted by harvests 1, 8, 9, 10, and 11 and the second by harvests 2, 3, 4, 5, 6 and 7 (Figure 4).

Figure 4 - Environmental similarity kernels K_E (A) and K_{Eg} (B) computed by the environmental covariables.



3.2 Predictive ability of the models

The PA varied from low (-0.08 – M4 in CV1 for R^2) to high (0.66 – M2, M3 and M4 in CV0 for b_0) (Table 3). Comparing all the tested scenarios, validation sets composed only by untested genotypes in all harvests (CV00 and CV1) had the lowest PA for all parameters estimated (Table 3). The M0 model showed similar PA compared to the other models evaluated for the CV00 and CV1 for GEBV and R^2 , and lower PAs were estimated for M0, M4 and M5 for b_0 and b_1 when compared to M1, M2 and M3 under CV00 and CV1 scenarios (Table 3). For the CV0 and CV2 scenarios, the models M0 and M1 showed similar PAs for GEBV, whereas models M2, M3, M4 and M5 had higher PAs for all parameters (Table 3). The model in which ignores the G×E effects (M0) had the lowest PA for b_0 , b_1 and R^2 when compared to other models under CV0 and CV2. When comparing the different enviromic relatedness kernels (M2 and M3, M4 and M5), similar PAs can be observed across all simulated scenarios.

Table 3 - Mean predictive ability across all number of harvests in the training population for GEBV, b_0 , b_1 and R^2 estimated by using M0, M1, M2, M3, M4 and M5 models for the CV00, CV1, CV0 and CV2 scenarios.

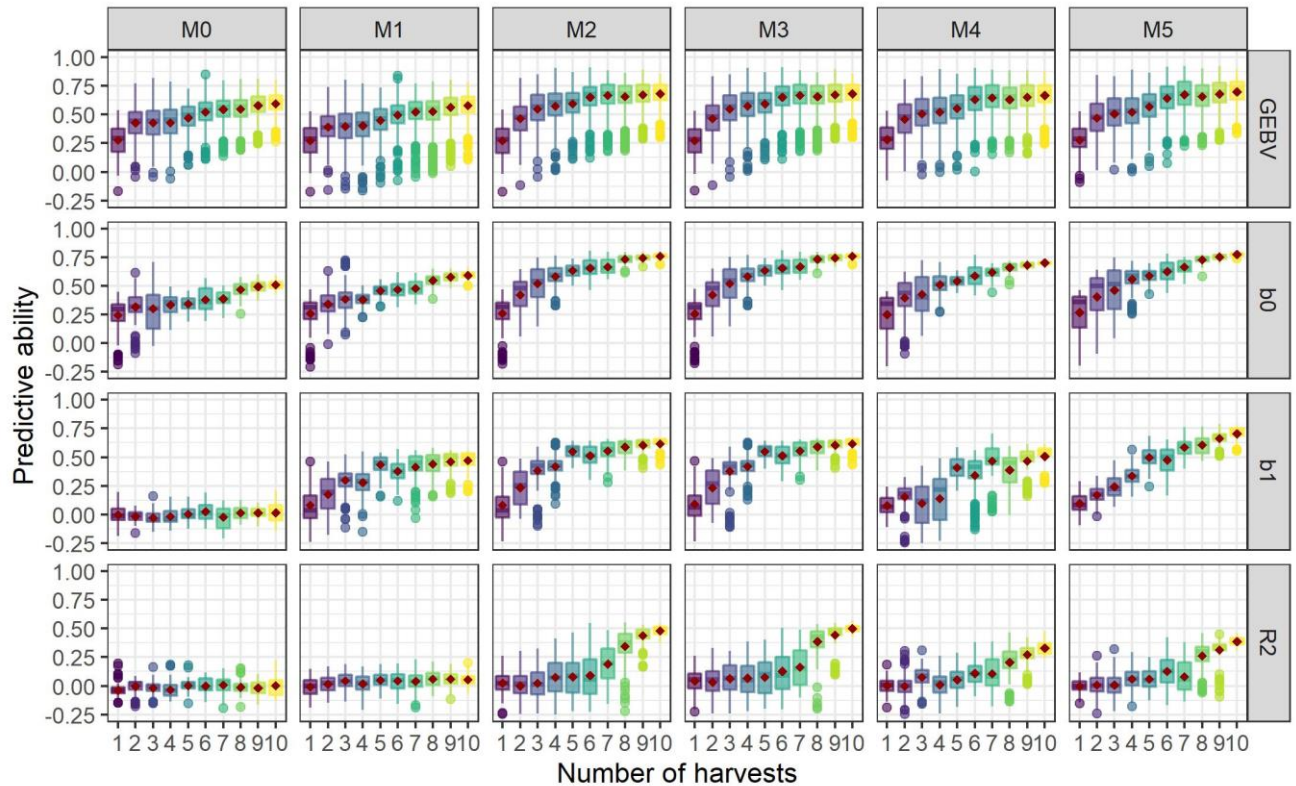
Parameter	Model	CV00	CV1	CV0	CV2
<i>GEBV</i>	M0	0.21 ± 0.03	0.22 ± 0.03	0.53 ± 0.07	0.52 ± 0.03
	M1	0.21 ± 0.03	0.24 ± 0.03	0.54 ± 0.02	0.51 ± 0.03
	M2	0.20 ± 0.03	0.24 ± 0.03	0.61 ± 0.02	0.63 ± 0.03
	M3	0.20 ± 0.03	0.24 ± 0.03	0.61 ± 0.02	0.63 ± 0.03
	M4	0.20 ± 0.03	0.20 ± 0.03	0.59 ± 0.02	0.61 ± 0.03
	M5	0.20 ± 0.03	0.20 ± 0.03	0.59 ± 0.02	0.63 ± 0.03
b_0	M0	0.08 ± 0.03	0.08 ± 0.01	0.42 ± 0.05	0.37 ± 0.03
	M1	0.12 ± 0.03	0.11 ± 0.02	0.53 ± 0.02	0.45 ± 0.03
	M2	0.11 ± 0.03	0.13 ± 0.02	0.66 ± 0.02	0.60 ± 0.03
	M3	0.11 ± 0.03	0.12 ± 0.02	0.66 ± 0.02	0.56 ± 0.03
	M4	0.06 ± 0.03	0.07 ± 0.03	0.64 ± 0.03	0.54 ± 0.03
	M5	0.07 ± 0.03	0.07 ± 0.03	0.66 ± 0.03	0.58 ± 0.03
b_1	M0	0.01 ± 0.03	0.01 ± 0.02	-0.02 ± 0.03	0.00 ± 0.03
	M1	0.13 ± 0.03	0.15 ± 0.02	0.43 ± 0.03	0.34 ± 0.03
	M2	0.16 ± 0.03	0.20 ± 0.02	0.56 ± 0.03	0.45 ± 0.04
	M3	0.16 ± 0.03	0.20 ± 0.02	0.54 ± 0.03	0.46 ± 0.04
	M4	-0.03 ± 0.02	-0.01 ± 0.02	0.58 ± 0.03	0.29 ± 0.04
	M5	-0.03 ± 0.02	-0.02 ± 0.02	0.56 ± 0.03	0.44 ± 0.04
R^2	M0	0.00 ± 0.03	0.01 ± 0.02	0.00 ± 0.03	0.01 ± 0.01
	M1	0.02 ± 0.03	0.04 ± 0.02	0.11 ± 0.01	0.04 ± 0.01
	M2	0.02 ± 0.03	0.08 ± 0.03	0.22 ± 0.04	0.16 ± 0.05
	M3	0.02 ± 0.03	0.08 ± 0.03	0.24 ± 0.04	0.18 ± 0.05
	M4	-0.06 ± 0.03	-0.08 ± 0.03	0.24 ± 0.03	0.12 ± 0.03
	M5	-0.04 ± 0.02	-0.07 ± 0.03	0.17 ± 0.03	0.13 ± 0.03

3.2.1 CV2: predicting untested families in observed harvests

The CV2 scenario imposes imbalanced data that may occur in repeated measures trials for perennial crops (missing data in some harvests for certain plots) to infer if the enviromic information can improve the PA. For this, 30% of the genotype-harvest combinations were withdraw and 70% of whole data were used as training population, and this was done by increasing the number of harvests in the training set. For this scenario, the PA for all the parameters increased as data from more harvests were included in the training population until reaching a plateau for all models, except for R^2 for all models, and for b_0 and b_1 for model M0 (Figure 5). When comparing the models M0 and M1, the inclusion of the G×E term led to higher PA for b_0 and b_1 . Models that accounted for the enviromic information (M2, M3, M4 and M5)

had higher PA when compared to the model M1, and this was more evident for R^2 (Figure 5 and Table S1).

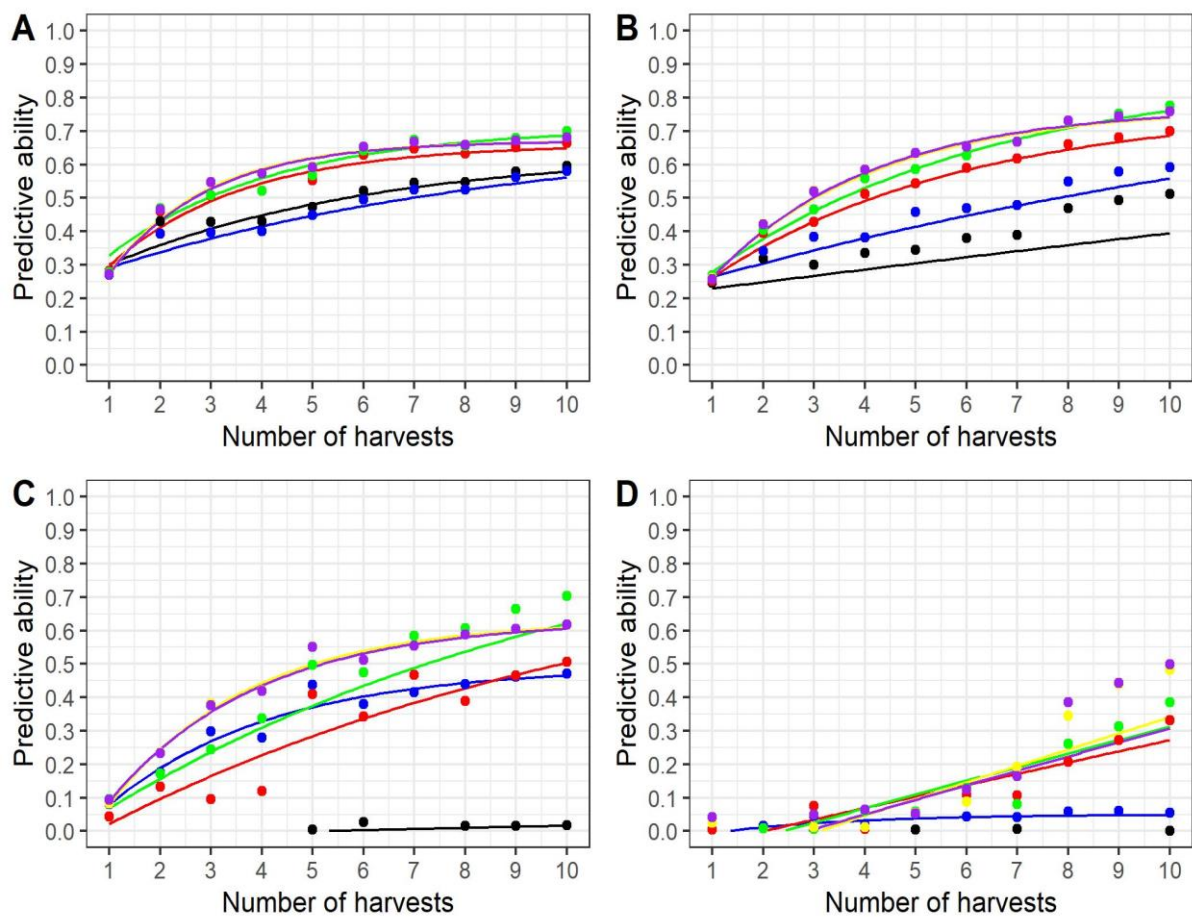
Figure 5 - Predictive ability for families GEBVs for DMY, and regression parameters of adaptability (b_0), persistence (b_1) and stability (R^2) by increasing the number of harvests in the training data set for the CV2.



The mean PA varied from 0.01 for M0 model (R^2) to 0.63 for M2, M3 and M5 (Table S1). The inclusion of the G×E effects in the model improved the PA for b_0 and b_1 , and the improvement was more evident for b_1 where the estimated PA increased from 0.00 (M0) to 0.34 (M1). Models that accounted for enviromic information (M2, M3, M4 and M5) improved the PAs by 19, 24, 17 and 73% for $GEBVs$, b_0 , b_1 and R^2 , respectively when compared to the model M1. Families' $GEBVs$, b_0 and b_1 had a good adjustment to the asymptotic regression, since most of the non-linear regression parameters (a , b and c) were significant for all models, except for b_0 and b_1 for model M0 (Table 3, Figure 6A, 6B and 6C). For model M0, only PAs for GEBVs showed an asymptotic behavior since this model ignores the G×E effects (Table S1 and Figure 6). The PAs for M1 reached the asymptotic point with less harvests and four, three and six harvests were needed for $GEBV$, b_0 and b_1 , respectively (Table S1). However, models that included enviromic information had higher PA with the same number of harvests than the model

M1, except for M4 (b_I), even though the enviromic based models did not reach the asymptotic point (Figure 6A, 6B and 6C). The models M2 and M3 performed better than the models M4 and M5 by reaching the asymptotic with less harvests in the training population with the similar PA for GEBV, b_0 and b_I (Table S1). By comparing different enviromic kernels (M2 and M3; M4 and M5), both kernels yielded similar results based on the asymptotic point (a) (Table S1).

Figure 6 - Non-linear asymptotic regressions fitted for predictive ability by number of harvests used in the training set for models M0 (black), M1 (blue), M2 (yellow), M3 (purple), M4 (red), and M5 (green) for GEBV (A), b_0 (B), b_I (C) and R^2 (D) in the scenario CV2.

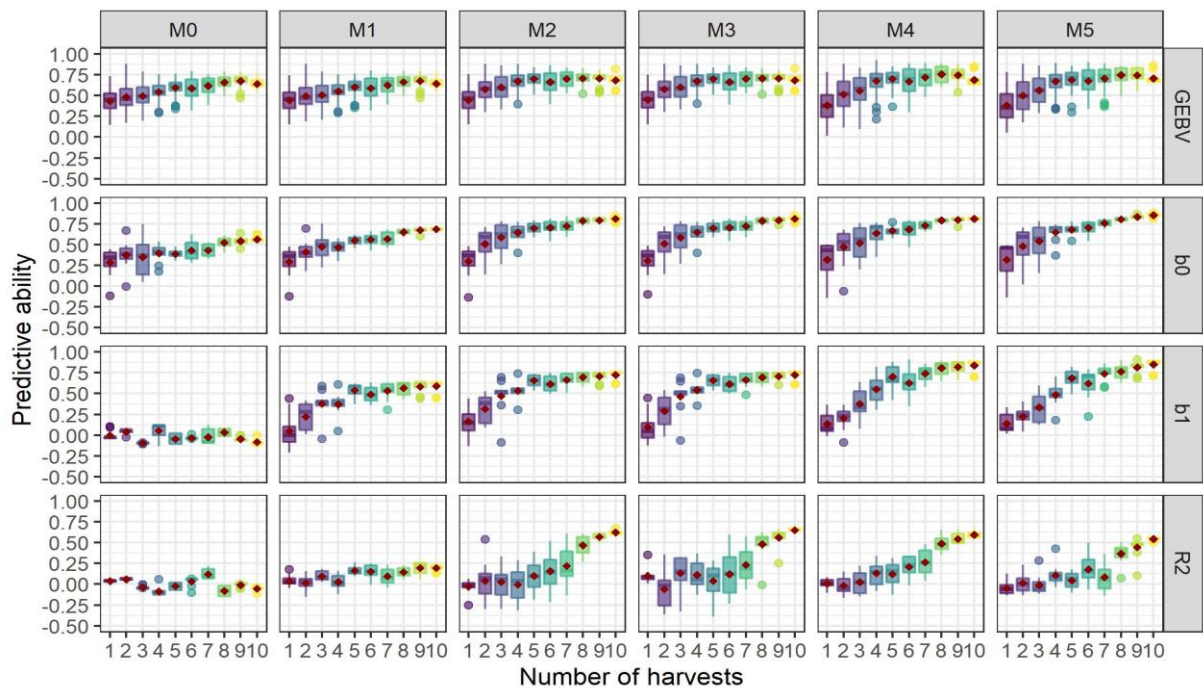


3.2.2 CV0: predicting tested families in unobserved harvests

The PA increased by increasing the number of harvests in the training data set for all models and parameters, except for b_I for model M0 and R^2 for model M0 and M1 (Figure 7). By including the G×E term in the model, the PA increased (Figure 7 – M0 and M1), and the greater difference was observed for b_I where the PA increased from zero (M0) to approximately 0.50 (M1) when using more than five harvests in the training population (Figure 7). Models

that accounted for the enviromic information outperformed model M1 for all parameters and number of harvests (Figure 7 and Table S2). The most expressive difference was observed for the stability parameter (R^2), as PA did not increase for M1 when more harvests were included in the training set, while it increased for models M2 to M5 (Figure 7). Furthermore, the number harvests increased the PA of models linearly until it reached a plateau for all parameters except for R^2 . For R^2 , models M2, M3, M4 and M5 resulted in greater PAs after five harvests were included in the training set (Figure 7).

Figure 7 - Predictive ability for families GEBVs, and regression parameters of adaptability (b_0), persistence (b_1) and stability (R^2) by increasing the number of harvests in the training data set for CV0.

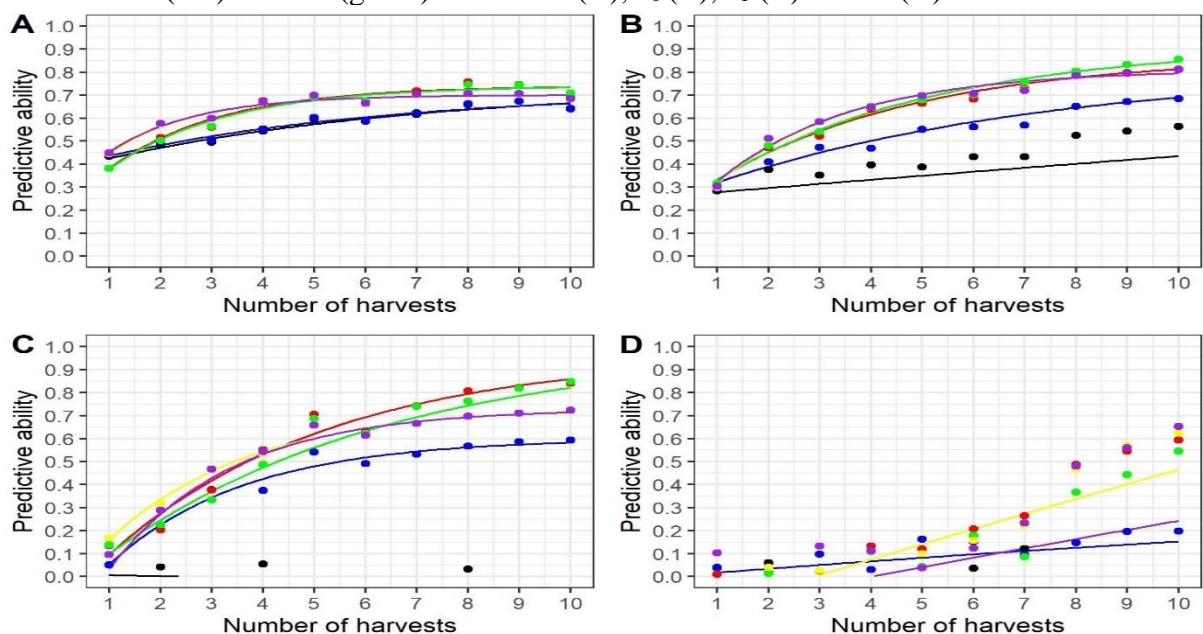


The mean PA through the different number of harvests varied among models and parameters, ranging from -0.02 for R^2 (M0) to 0.66 for b_0 (M2, M3 and M5) (Table S2). Higher PAs were observed for GEBV, b_0 and b_1 , whereas for R^2 the mean PA (\bar{r}) were lower evidencing greater difficulty in predicting the stability parameter. Furthermore, there was a lack of adjustment of the PAs to the asymptotic regression equation for R^2 (Figure 6D) since most of the parameters were not significant (Table S2). For GEBV, b_0 and b_1 the PAs had good adjustment to the asymptotic regression, except for M0 where only for GEBV the asymptotic regression had a good adjustment (Figures 8A, 8B and 8C).

By including the $G \times E$ term in the model the PA increased from 0.53 to 0.57, 0.42 to 0.53, -0.02 to 0.43 and 0.00 to 0.11, for GEBV, b_0 , b_1 and R^2 , respectively (Table S2). By using the enviromic information in the model, the models M2, M3, M4 and M5 increased PA by 10, 18, 22 and 47% for GEBVs, b_0 , b_1 and R^2 , respectively (Table S2). When comparing the models, minor changes in PAs occurred between the models that accounted for enviromic information, and the M0 model had the lowest PA for all parameters (Table S2). However, there were changes in the number of harvests to reach the PA asymptotic point (Table S2).

The reaction norm models (M4 and M5) had the highest asymptotic points, but these models needed for more harvests in the training population to reach the plateau, whereas for models M2 and M3 similar asymptotic patterns can be reached with less harvests in the training set (Table S2 and Figure 8). The K_E (M2 and M4) and K_{Eg} (M3 and M5) kernels did not change the number of harvests to reach the plateau, except for the EMDs models (M2 and M3) for b_0 in which the model M2 reached asymptotic point with seven harvests used in the training set whereas for M3 the number of harvests was eight (Table S2). Only for b_1 , the model M1 reached asymptotic point with less harvests (3) in the training set when compared with the models RN (M4 and M5) and EMDs (M2 and M3) (Table S2), however with three harvests the models EMDs and RN performed better than M1 (Figure 8). The models M2 and M3 had the highest values for the c parameter, i.e. these models had better linear gains in PA by increasing the number of harvests when compared to other models (Table S2 and Figure 8).

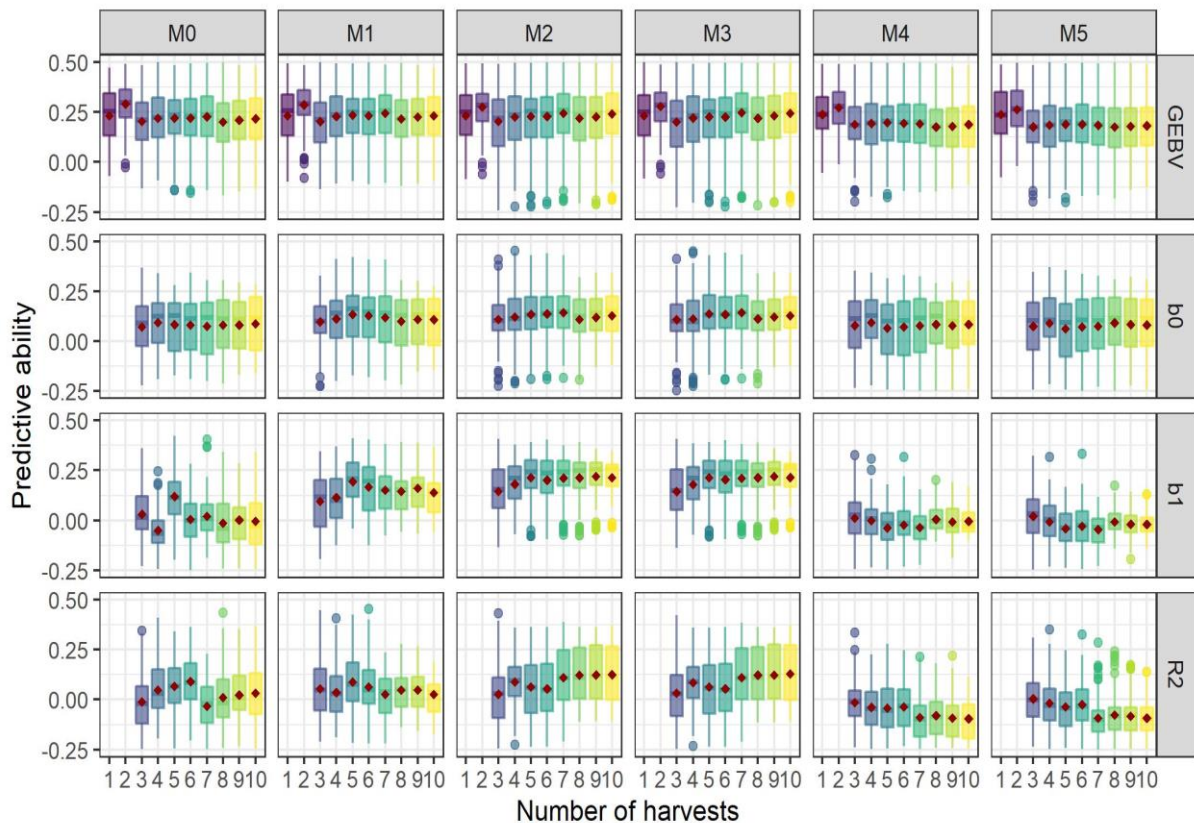
Figure 8 - Non-linear asymptotic regressions fitted for predictive ability by number of harvests used in the training set for models M0 (black), M1 (blue), M2 (yellow), M3 (purple), M4 (red) and M5 (green) for GEBV (A), b_0 (B), b_1 (C) and R^2 (D) in the scenario CV0.



3.2.3 CV1: predicting untested families in observed harvests

The mean predictive ability for the CV1 ranged from -0.08 (M4 for R^2) to 0.24 (M1, M2 and M3 for GEBVs), and PA was significantly lower than for CV0 and CV2 scenarios (Table 3). The increasing number of harvests in the training population led to a slightly increase in PA only for b_1 and R^2 for the models M2 and M3, where PA stabilized with five harvests in the training population (Figure 7). When including the G x E term in the model there is an increase in the PA from 0.08 (M0) to 0.11 (M1) and 0.01 (M0) to 0.15 (M1) for b_0 and b_1 , respectively (Table 3). There were non-significant differences for PA between M1 and models with enviromic information for b_0 and GEBVs (Figure 9). However, for b_1 there was a significant increase in PA when accounting for enviromic information in the models, as PA increased from 0.15 (M1) to 0.20 (M2 and M3) (Table 3). The models M4 and M5 had lower PA compared to the model M1 for b_0 and b_1 (Figure 9).

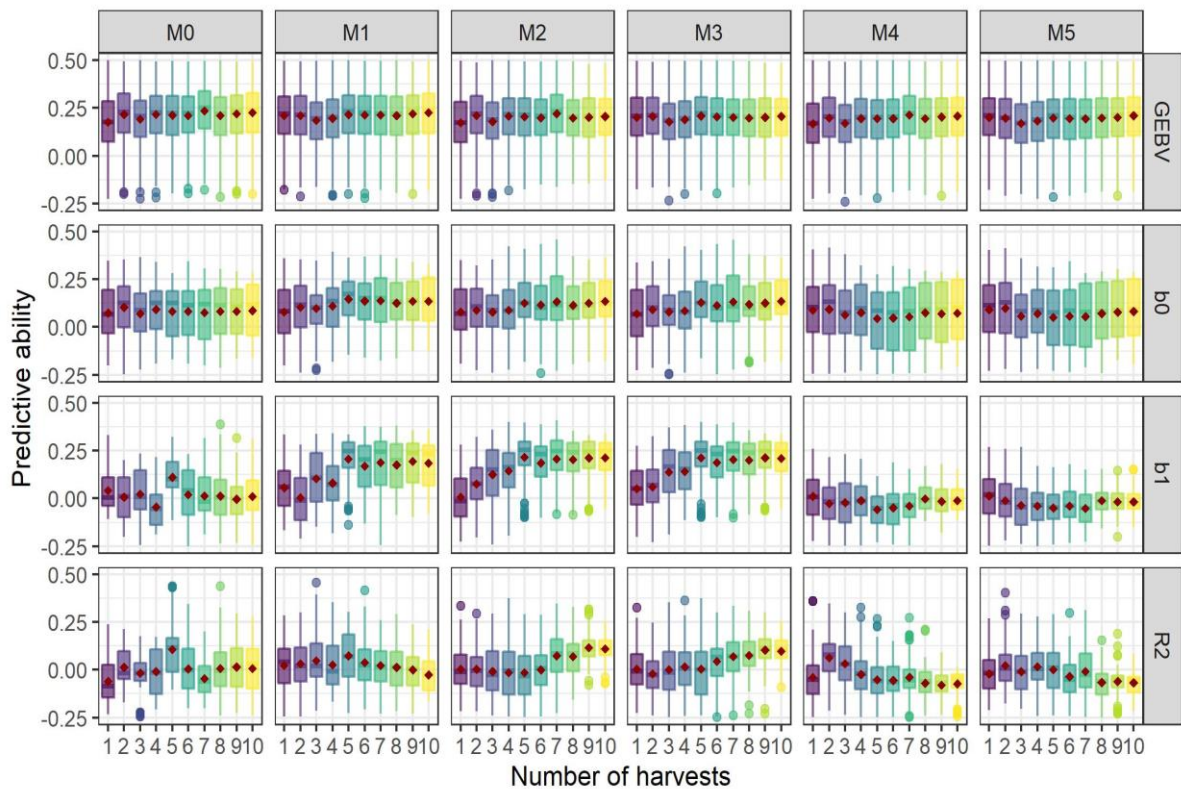
Figure 9 - Predictive ability for predicting untested families GEBVs in observed harvests (CV1), and predicting regression parameters of adaptability (b_0), persistence (b_1) and stability (R^2) by increasing the number of harvests in the training data set.



3.2.4 CV00: predicting untested families in observed harvests

The CV00 is the most challenging breeding scenario, where information for genotypes and harvests are missing. The lowest PA were observed for CV00 when comparing to other CV scenarios. The PA ranged from -0.06 (R^2 – M4) to 0.21 (GEBVs – M0 and M1), and the R^2 was the most difficult parameter to predict for all models (Table 3). The number of harvests used in the training population did not significantly increase PA for most parameters and models. Except for b_I with the M1, M2 and M3 models, where an increase in the PAs was observed when at least five harvests were included in the training set (Figure 8). When comparing M0 and M1, there was an increase in the PA for b_I by the inclusion of the G×E effects (Figure 8). In this scenario the differences between the models that used the enviromic information and the M1 model were not significant, except for b_I where the RN models (M4 and M5) had the lowest PA (Figure 10).

Figure 10 - Predictive ability for predicting untested families GEBVs in untested harvests (CV00), and predicting regression parameters of adaptability (b_0), persistence (b_I) and stability (R^2) by increasing the number of tested harvests.



4 DISCUSSION

Alfalfa breeding programs target the improvement of DMY, quality, persistence, and tolerance to biotic and abiotic stresses. As a perennial species, the phenotypic selection process for alfalfa can take several years, as traits must be evaluated under multi-harvests to find out which genotypes can maintain yield and quality over time. Therefore, genetic gain is lower compared to annual crops. Furthermore, due to the perennial behavior alfalfa breeders have to handle with G×E interaction from the beginning of the selection cycle. This can lead to lower heritability estimates when cross-over interactions are predominant across harvests. In addition, multi-harvest field trials are costly, forcing breeders to reduce the number of tested genotypes. Since whole-genome prediction was proposed (MEUWISSEN et al., 2001), genomic selection models became widely used in plant and animal breeding, aiming in accelerate the breeding process and increase genetic gains. Envirotyping has been incorporated in to whole-genome prediction aiming to better G×E modeling, enabling the prediction of novel genotypes in untested environments (Xu, 2016). In this study, we computed PAs based in breeding values for DMY (GEBVs), persistence (b_1), broad adaptability (b_0) and stability (R^2) for G-BLUP and for models that account for enviromic information through environmental relatedness kernel.

4.1 Variance components

Genotype by environment effects can be divided into complex and simple interactions effects. In this study, both simple and complex G×E were observed. The simple G×E occurs when there is heterogeneity of variances (CROSSA et al., 2004), and complex interaction is observed when there is a lack of genetic correlation between environments (VAN EEWIJK et al., 2016). However, the predominance of simple interaction was observed in this study where small amount of variance can be attributed to the G×E effects (σ_{AE}^2 , Table 2). Our results showed that the inclusion of enviromic information decreased the residual variance component, and the environment accounted for most of the phenotypic variance, and similar results were reported by Jarquín et al (2014), Jarquín et al. (2017) and Costa-Neto et al. (2021a). In addition, as pointed out by Costa-Neto et al. (2021a), the use of enviromic information impacts the model's ability in explaining phenotypic records, increasing the genomic variance, and the same pattern was found in our study (σ_A^2 , Table 2).

4.2 Predictive ability through cross-validation scenarios

The PAs ranged from low (-0.09) to high (0.63), depending on the used model, cross-validation scenario, and parameter. As expected, higher PAs were observed for CV0 and CV2 because the genotypes/families were already observed in any other harvest, whereas lower PAs were computed for CV00 and CV1. These results are in agreement with reports by Jarquín et al. (2017), by performing the same cross-validation scenarios in wheat (*Triticum aestivum*) for multi-environment trials. Therefore, these results evidenced that the phenotypic records in the training population for a given genotype has an important role improving the PAs. The RN models (M4 and M5) resulted in lower PAs under more unbalanced scenarios (CV00 and CV1) for all predicted parameters and had better results than M0 and M1 models when phenotypic records were available. This fact indicates an over fit from RN models to the data set, as they are more parametrized and can handle more complex G×E. For this dataset, we observed a predominance of non-crossover G×E, with a small proportion of variance captured by the interaction term (Table 2). The M1, M2 and M3 models resulted in similar PAs under CV00 and CV1 scenarios, but M2 and M3 models had better performance when phenotypic records were available (CV0 and CV2), therefore these model also can be implemented when non crossover G×E is predominant.

4.3 Enviromic-based models can improve accuracy under unbalanced multi-harvest trials

In early generation trials, a large number of families is evaluated in alfalfa field trials. Due to the large number of families, the use of incomplete block designs such as augmented block (FEDERER, 1956), p-rep (CULLIS et al., 2006) and row-column augmented block design (FEDERER; RAGHAVARAO, 1975) are used in order to evaluate the largest number of families as possible without increasing the number of plots. However, due to the lack of replication in those designs when a plot is lost in a given harvest, the information about that family will be missed, causing an imbalance at the subplot level (harvests), which can lead to a lower experimental precision. In this study, we simulated 30% of missing genotype-harvest combination (CV2) to assess if the enviromic information can help to improve accuracy under unbalanced multi-harvest trials. Our results showed that the incorporation of enviromic

information into genomic models increased the PA of predicted GEBV for DMY from 0.51 to 0.63 (Table 3).

Our results from CV2 opens the possibility in designing sparse multi-harvest trials, where not all genotypes are evaluated at each harvest, or even at each location (GONZÁLEZ-BARRIOS et al., 2018). This strategy makes feasible the evaluation of more families per breeding cycle. Jarquín et al. (2017) emphasized that the CV2 strategy can reduce the workload per harvest and increases the efficiency per plot. Although the accuracy can be lower in sparse multi-environment trials, the predictive ability can be recovered when more genotypes are tested across harvests (JARQUÍN et al., 2020; CRESPO-HERRERA et al., 2021). Montesinos-Lopez et al. (2022) proposed the use of incomplete block design to allocate wheat lines across the environments, and the authors reported an improvement on the PAs when compared to the random allocation of lines across the environment. Therefore, the approach proposed by Montesinos-Lopez et al. (2022) can be applied in alfalfa multi-harvest trials, in which the families to be harvested would be selected according to an incomplete block design.

4.4 Reducing the number of harvests evaluation in alfalfa breeding trials

A concern among forage breeders is the number of harvests needed to select superior genotypes by keeping a good balance between precision and resources allocated (FIGUEIREDO et al., 2019). Therefore, reducing the number of harvests is crucial to accelerate alfalfa breeding cycle, as well as increasing the number of families to be evaluated in early generation cycles. Resende et al. (2002) suggested the use of the repeatability coefficient as a parameter to infer about the optimal number of harvests, in which repeatability above 0.80 should be the ideal for phenotypic selection. By using repeatability coefficient, Souza Sobrinho et al. (2010) estimated the optimal number of harvests for DMY for *Urochloa ruziziensis* L., where the authors concluded that 7 to 8 harvests are sufficient for phenotypic selection. We suggested a genomic prediction approach based on cross-validation to determine the optimal number of harvests to predict GEBV for DMY for tested genotypes in untested harvests (CV0). Our study showed that the amount of information (phenotypic records) in the model followed an asymptotic trend. Costa-Neto et al. (2021b) pointed out that more phenotypic records do not always lead to better accuracies. Therefore, the optimal number of harvests would be the point in which the PA do not significantly increase on the asymptotic regression. The best models (M2 and M3) reached the asymptotic point by using two fewer harvests than M1 model, and

the PA increased from 0.62 to 0.68 for DMY. For b_0 and b_1 , the M1 model reached the asymptotic point by using fewer harvests (six harvests for b_0 and three for b_1) in the training population. However, despite that M1 reached the asymptotic point using fewer harvests, the models M2 and M3 had higher PA with the same number of harvests in the training population (Figure 8), and at the asymptotic point the M2 and M3 models reached PA of 0.77 (seven harvests) and 0.70 (seven harvests) for b_0 and b_1 , respectively.

4.5 Multi-harvest genomic selection

Genomic selection has been successfully applied in alfalfa breeding programs (ANNICHIARICO et al., 2015; LI et al., 2015; ANDRADE et al., 2022). Neglecting the genotype by harvest interaction in alfalfa can lead to bias prediction, as PA varies across harvests (ANDRADE et al., 2022). As described by Gauch and Zobel (1996) the G×E interaction can be explained as the expression of different traits across environment, where the correlation indicates if the same genes are being expressed or not. Therefore, different PAs across harvests indicates the presence of G×E.

By modeling the interaction using a heterogeneous autoregressive covariance structure for G×E, Andrade et al. (2022) reported moderate PA by using five cumulated harvests in a CV1 cross-validation scheme. As opposed to Andrade et al. (2022), we did not observe an increase in the PA by increasing the number of harvests in the training population. By using a Bayesian approach, we were able to include all the harvests in the model, in which some harvests were used as training and other as testing sets. Since we used all harvests, the models were able to capture the G×E even when only one harvest was used as the training population. For the CV00 and CV1 scenarios, the best enviromic-based models (M2 and M3) yielded similar results than M1, for predicting GEBVs and moderate PAs were observed (0.20 for CV00 and 0.24 for CV1). Moderate PA for DMY were also observed by Annichiarico et al. (2015), Li et al. (2015) and Andrade et al. (2022). The results found in our study can serve to optimize the alfalfa breeding program by reducing the time and resources necessary in the evaluation of training populations.

4.6 Persistence, broad adaptability and stability

There are three possible ways of dealing with the presence of G×E in plant breeding: to ignore it, to avoid it or to explore it (EISEMANN et al., 1990). By ignoring the G×E (M0), the selection for alfalfa adaptation can potentially be lost and might compromise yield potential in specific environments (Yan, 2016). By exploring the G×E, the target environments are grouped into mega-environments blocking the G×E effects, and selection is done within mega-environments allowing the maximum expression of yield potential for the selected genotypes (GAUCH, 1988; YAN et al., 2000; SMITH et al., 2001; CULLIS et al., 2014; DIAS et al., 2022). To avoid G×E, the selection of the genotypes has two components, high mean trait performance and high stability across environments (FINLAY; WILKINSON, 1963; EBERHART; RUSSELL, 1966). By the nature of the G×E in this study, the best strategy to deal with G×E is avoiding it as the harvests (time) are not repeatable (Yan, 2016).

We investigated the PA for the regression parameters over time, in which the parameters b_0 and R^2 can be interpreted as proposed by Eberhart and Russel (1966), b_0 as broad adaptability and R^2 as stability. In this study, we regressed the predicted GEBVs for DMV over time, therefore b_1 is not the genotypes' plasticity to the improved environmental quality as originally proposed by Eberhart and Russel (1966). As suggested by De Assis et al. (2010), the b_1 coefficient is the genotype's persistence, and genotypes with good persistence must have values close to zero. When ignoring G×E (M0), the PAs for b_0 , b_1 and R^2 were lower, and the results showed the importance of considering G×E in the genomic prediction models even when there is predominance of simple G×E. As expected, our results showed that the stability parameter (R^2) was the most difficult parameter to be predicted. PAs closer to zero can be observed for all models in the CV00 and CV1 scenarios, whereas b_0 and b_1 had PA greater than zero under CV00 and CV1. Under CV0 and CV2, R^2 needed more harvests to have reasonable PAs for the enviromic-based models, whereas PAs were low for the M0 model even with the increasing number of harvests in the training population. These results are aligned with reports by Mendes and Ramalho (2018), where the authors studied the repeatability coefficient of adaptability and stability parameters for 26 common beans assessed in 36 environments, and the authors reported greater repeatability estimates for b_0 and b_1 and very low repeatability for stability parameter.

The authors also reported that selection based on stability should be done in a greater number of environments. These results indicate the possibility of using genomic selection for persistence in alfalfa breeding programs with reduced number of harvests.

4.7 Differences between K_E and K_{Eg} on predictive ability

Our study showed that the use of historical data in computing the environmental relatedness kernel did not change PAs for the enviromic-based models in all CV scenarios and predicted parameters. Since breeders do not have weather data for the year in which the trials will be conducted, historical data can be used in genomic prediction models to predict new genotypes. The use of historical data has the advantage of being updated every year and can capture environmental trends driven by climate change.

5 CONCLUSION

Our findings suggest that G×E term must be taking into account in genomic prediction models and enviromic-based models can improve efficiency of alfalfa breeding programs, by reducing the number of harvests in the selection cycle (CV0) and increasing the experimental precision under unbalanced multi-harvest trials (CV2) for DMY and persistence. The weather historical data to compute environmental relatedness kernel can be used in enviromic-based model.

Furthermore, the use of a Bayesian regression kernels allowed moderate predictive ability in genomic predictions (CV1 and CV00) even when only one harvest was used in the training data set, and also when increasing the number of harvests in the training set the predictive abilities did not change.

6 CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 FUNDING INFORMATION

Claudio Carlos Fernandes Filho received funding from CAPES PrInt UFLA as PhD Visiting Scholar (Doctoral Exchange) at the University of Florida, Gainesville.

REFERENCES

- Acharya, J. P., Lopez, Y., Gouveia, B. T., De Bem Oliveira, I., Resende, M. F. R., Muñoz, P. R., et al. (2020). Breeding Alfalfa (*Medicago sativa* L.) Adapted to Subtropical Agroecosystems. *Agronomy* 10, 742. doi:10.3390/agronomy10050742.
- Adhikari, L., Makaju, S. O., and Missaoui, A. M. (2019). QTL mapping of flowering time and biomass yield in tetraploid alfalfa (*Medicago sativa* L.). *BMC Plant Biology* 19. doi:10.1186/s12870-019-1946-0.
- Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). FAO Irrigation and drainage paper No. 56. Rome: Food and Agriculture Organization of the United Nations, 56(97), e156.
- Andrade, M. H. M. L., Acharya, J. P., Benevenuto, J., de Bem Oliveira, I., Lopez, Y., Munoz, P., Resende Jr., M. F. R., Rios., E. F. Genomic prediction for canopy height and dry matter yield in alfalfa using family bulks.
- Annicchiarico P, Scotti C, Carelli M, Pecetti L. Questions and avenues for lucerne improvement. (2010). *Czech J Genet Plant Breed*, 46, 1–13.
- Annicchiarico, P., and Pecetti, L. (2021). Comparison among nine alfalfa breeding schemes based on actual biomass yield gains. *Crop Science* 61, 2355–2371. doi:10.1002/csc2.20464.
- Annicchiarico, P., Nazzicari, N., Li, X., Wei, Y., Pecetti, L., and Brummer, E. C. (2015). Accuracy of genomic selection for alfalfa biomass yield in different reference populations. *BMC Genomics* 16. doi:10.1186/s12864-015-2212-y.
- Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends in plant science*, 19(1), 52-61.
- Biazzi, E., Nazzicari, N., Pecetti, L., Brummer, E. C., Palmonari, A., Tava, A., et al. (2017). Genome-Wide Association Mapping and Genomic Selection for Alfalfa (*Medicago sativa*) Forage Quality Traits.. *PLOS ONE* 12, e0169234. doi:10.1371/journal.pone.0169234.
- Biswas, A., Andrade, M. H. M. L., Acharya, J. P., De Souza, C. L., Lopez, Y., De Assis, G., et al. (2021). Phenomics-Assisted Selection for Herbage Accumulation in Alfalfa (*Medicago sativa* L.). *Frontiers in Plant Science* 12. doi:10.3389/fpls.2021.756768.
- Bouton, J. H. (2012). Breeding lucerne for persistence. *Crop & Pasture Science* 63, 95. doi:10.1071/cp12009.
- Bowley, S. R., and Christie, B. R. (1981). Inheritance of dry matter yield in a heterozygous population of alfalfa. *Canadian Journal of Plant Science*, 61(2), 313-318. doi:org/10.4141/cjps81-044.
- Brummer C., E., and Casler, M. D. (2014). Cool-season forages. Yield gains in major US field crops, 33, 33-51. doi:org/10.2135/cssaspecpub33.c3.

- Brummer, E. C. Applying genomics to alfalfa breeding programs. *Crop science*, 44(6) 1904-2004.
- Casler, M. D., and Brummer, E. C. (2008). Theoretical Expected Genetic Gains for Among- and-Within-Family Selection Methods in Perennial Forage Crops. *Crop Science* 48, 890. doi:10.2135/cropsci2007.09.0499.
- Castonguay, Y., Laberge, S., Brummer, E. C., and Volenec, J. J. (2006). Alfalfa winter hardiness: a research retrospective and integrated perspective. *Advances in agronomy*, 90, 203-265. doi: 10.1016/S0065-2113(06)90006-6.
- Chen, H., Zeng, Y., Yang, Y., Huang, L., Tang, B., Zhang, H., et al. (2020). Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nature Communications* 11. doi:10.1038/s41467-020-16338-x.
- Collino, D. J., Dardanelli, J. L., De Luca, M. J., and Racca, R. W. (2005). Temperature and water availability effects on radiation and water use efficiencies in alfalfa (*Medicago sativa* L.). *Australian Journal of Experimental Agriculture* 45, 383. doi:10.1071/ea04050.
- Costa-Neto, G. M. F., Morais Júnior, O. P., Heinemann, A. B., De Castro, A. P., and Duarte, J. B. (2020). A novel GIS-based tool to reveal spatial trends in reaction norm: upland rice case study. *Euphytica* 216. doi:10.1007/s10681-020-2573-4.
- Costa-Neto, G., Crossa, J., and Fritsche-Neto, R. (2021b). Enviromic assembly increases accuracy and reduces costs of the genomic prediction for yield plasticity in maize. *Frontiers in plantscience*, 12.
- Costa-Neto, G., Galli, G., Carvalho, H. F., Crossa, J., and Fritsche-Neto, R. (2021a). EnvRtype: a software to interplay enviromics and quantitative genomics in agriculture. *G3*, 11(4).
- Crespo-Herrera, L., Howard, R., Piepho, H., Pérez-Rodríguez, P., Montesinos-Lopez, O., Burgueño, J., et al. (2021). Genome-enabled prediction for sparse testing in multi-environmental wheat trials. *The Plant Genome* 14. doi:10.1002/tpg2.20151.
- Crossa, J., Yang, R.-C., and Cornelius, P. L. (2004). Studying crossover genotype \times environment interaction using linear-bilinear models and mixed models. *Journal of Agricultural, Biological and Environmental Statistics* 9, 362–380. doi:10.1198/108571104x4423.
- Cullis, B. R., Jefferson, P., Thompson, R., and Smith, A. B. (2014). Factor analytic and reduced animal models for the investigation of additive genotype-by-environment interaction in outcrossing plant species with application to a *Pinus radiata* breeding programme. *Theoretical and Applied Genetics*, 127(10), 2193-2210.
- Cullis, B. R., Jefferson, P., Thompson, R., and Smith, A. B. (2014). Factor analytic and reduced animal models for the investigation of additive genotype-by-environment interaction in outcrossing plant species with application to a *Pinus radiata* breeding programme. *Theoretical and Applied Genetics* 127, 2193–2210. doi:10.1007/s00122-014-2373-0.

Cullis, B. R., Smith, A. B., and Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological and Environmental Statistics* 11, 381–393. doi:10.1198/108571106x154443.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., Depristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/bioinformatics/btr330.

De Assis, G. M. L., Ruggieri, A. C., Mercadante, M. E. Z., De Camargo, G. M. F., and Carneiro Júnior, J. M. (2010). Selection of alfalfa cultivars adapted for tropical environments with repeated measures using PROC MIXED of SAS® System. *Plant Genetic Resources: Characterization and Utilization* 8, 55–62. doi:10.1017/s1479262109990153.

De Bem Oliveira, I., Amadeu, R. R., Ferrão, L. F. V., and Muñoz, P. R. (2020). Optimizing whole-genomic prediction for autotetraploid blueberry breeding. *Heredity* 125, 437–448. doi:10.1038/s41437-020-00357-x.

De Bem Oliveira, I., Resende, M. F. R., Ferrão, L. F. V., Amadeu, R. R., Endelman, J. B., Kirst, M., et al. (2019). Genomic Prediction of Autotetraploids; Influence of Relationship Matrices, Allele Dosage, and Continuous Genotyping Calls in Phenotype Prediction. *G3 Genes|Genomes|Genetics* 9, 1189–1198. doi:10.1534/g3.119.400059.

Dias, K. O. G., Dos Santos, J. P. R., Krause, M. D., Piepho, H.-P., Guimarães, L. J. M., Pastina, M. M., et al. (2022). Leveraging probability concepts for cultivar recommendation in multi-environment trials. *Theoretical and Applied Genetics*. doi:10.1007/s00122-022-04041-y.

Eberhart, S. A., and Russell, W. A. (1966). Stability Parameters for Comparing Varieties 1.. *Crop Science* 6, 36–40. doi:10.2135/cropsci1966.0011183x000600010011x.

Eisemann, R.L., M. Cooper, and D.R. Woodruff. (1990). Beyond the analytical methodology—Better interpretation and exploitation of genotype-by-environment interaction in breeding. In: M.S. Kang, editor, *Genotype-by-environment interaction and plant breeding*. Louisiana State University, Baton Rouge. p. 108–117.

Engels, C. (1994) Effect of root and shoot meristem temperature on shoot to root dry matter partitioning and the internal concentrations of nitrogen and carbohydrates in maize and wheat. *Annals of Botany* 73, 211–219. doi: 10.1006/anbo.1994.1025

Faveri, J. D., Verbyla, A. P., Pitchford, W. S., Venkatanagappa, S., and Cullis, B. R. (2015). Statistical methods for analysis of multi-harvest data from perennial pasture variety selection trials. *Crop & Pasture Science* 66, 947. doi:10.1071/cp14312.

Federer, W. F. (1956). *Experimental design* (Vol. 81, No. 4, p. 334). LWW.

Federer, W. T., and Raghavarao, D. (1975). On Augmented Designs. *Biometrics* 31, 29. doi:10.2307/2529707.

- Figueiredo, U. J. D., Berchembrock, Y. V., Valle, C. B. D., Barrios, S. C. L., Quesenberry, K. H., Muñoz, P. R., & Nunes, J. A. R. (2019). Evaluating early selection in perennial tropical forages. *Crop Breeding and Applied Biotechnology*, 19, 291-299.
- Finlay, K., and Wilkinson, G. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research* 14, 742. doi:10.1071/ar9630742.
- Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv: 1207.3907.
- Gauch Jr, H. G. (1988). Model selection and validation for yield trials with interaction. *Biometrics*, 705-715.
- Gauch, H. G. Jr., & Zobel, R. W. (1996). AMMI analysis of yield trials. In *Genotype-by-environment interaction* (pp. 85–122). Boca Raton, FL: CRC Press.
- González-Barrios, P., Díaz-García, L., and Gutiérrez, L. (2019). Mega-Environmental Design: Using Genotype \times Environment Interaction to Optimize Resources for Cultivar Testing. *Crop Science* 59, 1899–1915. doi:10.2135/cropsci2018.11.0692.
- Granato, I., Cuevas, J., Luna-Vázquez, F., Crossa, J., Montesinos-López, O., Burgueño, J., and Fritsche-Neto, R. (2018). BGGGE: a new package for genomic-enabled prediction incorporating genotype \times environment interaction models. *G3: Genes, Genomes, Genetics*, 8(9), 3039-3047.
- Hoppen, S. M., Ta, H., Yang, X., Jáuregui, J., Neres, M. A., and Moot, D. (2019). Shoot and perennial organs yields of lucerne genotypes of three fall dormancy levels over five years. *Proceedings of...*, Wagga Wagga, Australia.
- Horner, E. S. Registration of Florida 66 Alfalfa (Reg. No. 48) (1970). *Crop Sci.*, 10, 456. doi: 10.2135/cropsci1970.0011183X001000040047x
- Horner, E.S. and Ruelke, O.C. Registration of Florida 77 Alfalfa (Reg. No. 99) (1981). *Crop Sci.*, 21, 797–797. <https://doi.org/10.2135/cropsci1981.0011183X002100050042x>
- Irwin, J. (1977). Factors contributing to poor Lucerne persistence in southern Queensland. *Australian Journal of Experimental Agriculture* 17, 998. doi:10.1071/ea9770998.
- Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., et al. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics* 127, 595–607. doi:10.1007/s00122-013-2243-1.
- Jarquín, D., Lemes Da Silva, C., Gaynor, R. C., Poland, J., Fritz, A., Howard, R., et al. (2017). Increasing Genomic-Enabled Prediction Accuracy by Modeling Genotype \times Environment Interactions in Kansas Wheat. *The Plant Genome* 10, plantgenome2016. doi:10.3835/plantgenome2016.12.0130.

- Jia, C., Zhao, F., Wang, X., Han, J., Zhao, H., Liu, G., & Wang, Z. (2018). Genomic prediction for 25 agronomic and quality traits in alfalfa (*Medicago sativa*). *Frontiers in plant science*, 1220. doi: 10.3389/fpls.2018.01220.
- Leach, G. J., and Clements, R. J. (1984). Ecology and grazing management of alfalfa pastures in the subtropics. *Advances in Agronomy*, 37, 127-154.
- Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi:10.1093/bioinformatics/btp324.
- Li, X., and Brummer, E. C. (2012). Applied Genetics and Genomics in Alfalfa Breeding. *Agronomy* 2, 40–61. doi:10.3390/agronomy2010040.
- Li, X., Wei, Y., Acharya, A., Hansen, J. L., Crawford, J. L., Viands, D. R., et al. (2015). Genomic Prediction of Biomass Yield in Two Selection Cycles of a Tetraploid Alfalfa Breeding Population. *The Plant Genome*, 8. doi:10.3835/plantgenome2014.12.0090.
- Lin, C. S., Binns, M. R., and Lefkovitch, L. P. (1986). Stability Analysis: Where Do We Stand? 1.. *Crop Science* 26, 894–900. doi:10.2135/cropsci1986.0011183x002600050012x.
- Malik, W., Boote, K. J., Hoogenboom, G., Cavero, J., and Dechmi, F. (2018). Adapting the CROPGRO Model to Simulate Alfalfa Growth and Yield. *Agronomy Journal* 110, 1777–1790. doi:10.2134/agronj2017.12.0680.
- McKenzie, J. S., Paquin, R., & Duke, S. H. (1988). Cold and heat tolerance. *Alfalfa and alfalfa improvement*, 29, 259-302. doi: 10.2134/agronmonogr29.c8.
- Mendes, M. H. S., and Ramalho, M. A. P. (2018). Repeatability of some phenotypic stability parameters - a resampling approach. *Crop Breeding and Applied Biotechnology* 18, 139–147. doi:10.1590/1984-70332018v18n2a20.
- Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157, 1819–1829. doi:10.1093/genetics/157.4.1819.
- Montesinos-Lopez, O. A., Montesinos-Lopez, A., Acosta, R., Varshney, R. K., Bentley, A., and Crossa, J. (2022). Using an incomplete block design to allocate lines to environments improves sparse genome-based prediction in plant breeding. *The Plant Genome*. doi:10.1002/tpg2.20194.
- Morrell, P. L., Buckler, E. S., and Ross-Ibarra, J. (2012). Crop genomics: advances and applications. *Nature Reviews Genetics* 13, 85–96. doi:10.1038/nrg3097.
- Patterson, D. T. (1993). Effects of Day and Night Temperature on Goatsrue (*Galega officinalis*) and Alfalfa (*Medicago sativa*) Growth. *Weed Science* 41, 38-45.
- Piepho, H. P. (1998). Methods for comparing the yield stability of cropping systems. *Journal of Agronomy and Crop Science*, 180(4), 193-213.

Priestley, C. H. B., and Taylor, R. J. (1972). On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Monthly Weather Review* 100, 81–92. doi:10.1175/1520-0493(1972)100<0081:otaosh>2.3.co;2.

Resende MDV (2002) Biometric genetics and statistics in perennial plant breeding (in Portuguese). Embrapa Informação Tecnológica, Colombo, Brazil. <https://www.bdpa.cnptia.embrapa.br/consulta/busca?b=pc&id=306061&biblioteca=vazio&busca=autoria:%22RESENDE,%20M.%20D.%20V.%20de.%22&qFacets=autoria:%22RESENDE,%20M.%20D.%20V.%20de.%22&sort=&paginacao=t&paginaAtual=1>

Riday, H., and Brummer, E. C. (2004). Performance of intersubspecific alfalfa hybrids in sward versus space planted plots. *Euphytica* 138, 107–112. doi:10.1023/b:euph.0000046755.07566.c3.

Rimi, F., Macolino, S., Leinauer, B., Lauriault, L. M., and Ziliotto, U. (2014). Fall Dormancy and Harvest Stage Impact on Alfalfa Persistence in a Subtropical Climate. *Agronomy Journal* 106, 1258–1266. doi:10.2134/agronj13.0495.

Rincent, R., Malosetti, M., Ababaei, B., Touzy, G., Mini, A., Bogard, M., ... and Van Eeuwijk, F. (2019). Using crop growth model stress covariates and AMMI decomposition to better predict genotype-by-environment interactions. *Theoretical and Applied Genetics*, 132(12), 3399-3411.

Robison, G. D., Robinson, G. D., and Massengale, M. A. (1969). Effect of Night Temperature on Growth and Development of Alfalfa (*Medicago sativa* L.). *Journal of the Arizona Academy of Science* 5, 227. doi:10.2307/40021957.

Sakiroglu, M., and Brummer, E. C. (2017). Identification of loci controlling forage yield and nutritive value in diploid alfalfa using GBS-GWAS. *Theoretical and Applied Genetics* 130, 261–268. doi:10.1007/s00122-016-2782-3.

Santos, I. G. D., Cruz, C. D., Nascimento, M., Rosado, R. D. S., and Ferreira, R. D. P. (2018). Direct, indirect and simultaneous selection as strategies for alfalfa breeding on forage yield and nutritive value. *Pesquisa Agropecuária Tropical* 48, 178–189. doi:10.1590/1983-40632018v48i51950.

Shelford, V. E. 1931. Some concepts of bioecology. *Ecology* 12:455–467.

Smith, A. B., and Cullis, B. R. (2018). Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. *Euphytica*, 214(8), 1-19.

Smith, A. B., Cullis, B. R., and Thompson, R. (2005). The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science*, 143(6), 449-462.

Soltani, A., Sinclair, TR. (2012). *Modeling Physiology of Crop Development, Growth and Yield* Wallingford, CA: CABInternational. pp. 322.

Souza Sobrinho, F., Borges, V., Léo, F. J. D. S., & Kopp, M. M. (2010). Repetibilidade de características agronômicas e número de cortes necessários para seleção de *Urochloa ruziziensis*. *Pesquisa Agropecuária Brasileira*, 45, 579-584.

Teixeira, E. I., Moot, D. J., and Brown, H. E. (2008). Defoliation frequency and season affected radiation use efficiency and dry matter partitioning to roots of lucerne (*Medicago sativa* L.) crops. *European Journal of Agronomy* 28, 103–111. doi:10.1016/j.eja.2007.05.004.

Teixeira, E. I., Moot, D. J., Brown, H. E., and Pollock, K. M. (2007). How does defoliation management impact on yield, canopy forming processes and light interception of lucerne (*Medicago sativa* L.) crops? *European Journal of Agronomy* 27, 154–164. doi:10.1016/j.eja.2007.03.001.

Van Eeuwijk, F. A., Bustos-Korts, D. V., and Malosetti, M. (2016). What Should Students in Plant Breeding Know About the Statistical Aspects of Genotype \times Environment Interactions? *Crop Science* 56, 2119–2140. doi:10.2135/cropsci2015.06.0375.

Xu, Y. (2016). Envirotyping for deciphering environmental impacts on crop plants. *Theoretical and Applied Genetics* 129, 653–673. doi:10.1007/s00122-016-2691-5.

Yan, W. (2016). Analysis and Handling of $G \times E$ in a Practical Breeding Program. *Crop Science* 56, 2106–2118. doi:10.2135/cropsci2015.06.0336.

Yan, W., Hunt, L. A., Sheng, Q., and Szlavnic, Z. (2000). Cultivar Evaluation and Mega-Environment Investigation Based on the GGE Biplot. *Crop Science* 40, 597–605. doi:10.2135/cropsci2000.403597x.

Yuan, Y., Bayer, P. E., Batley, J., and Edwards, D. (2017). Improvements in Genomic Technologies: Application to Crop Genomics. *Trends in Biotechnology* 35, 547–558. doi:10.1016/j.tibtech.2017.02.009.

SUPPLEMENTARY MATERIAL

Table S1 - Parameter estimations from fitting asymptotic regression to mean cross validation predictive ability by number of harvests in the training population for breeding values in each untested harvest (GEBV), broad adaptability (b_0), persistence (b_1) and stability (R^2) for CV2.

Predicted	Regression Parameter	M0	M1	M2	M3	M4	M5
GEBV	a	0.64*	0.70*	0.67*	0.67*	0.66*	0.71*
	b	0.23*	0.24*	0.01 ^{ns}	0.01 ^{ns}	0.13 ^{ns}	0.15 ^{ns}
	c	0.19 ^{ns}	0.12 ^{ns}	0.52*	0.51*	0.38*	0.31*
	\bar{r}	0.52	0.51	0.63	0.63	0.61	0.63
	95% asymptotic point	0.48 (5)	0.43 (4)	0.65 (7)	0.65 (7)	0.64 (7)	0.63 (6)
b_0	a	2.15 ^{ns}	0.97*	0.77*	0.77*	0.77*	0.88*
	b	0.21*	0.22*	0.08 ^{ns}	0.07 ^{ns}	0.15 ^{ns}	0.16 ^{ns}
	c	0.01 ^{ns}	0.06 ^{ns}	0.31*	0.32*	0.20*	0.18*
	\bar{r}	0.37	0.45	0.60	0.56	0.54	0.58
	95% asymptotic point	-	0.55 (3)	0.73 (9)	0.73 (9)	0.69 (10)	0.79 (11)
b_1	a	0.18 ^{ns}	0.49*	0.63*	0.63*	0.89 ^{ns}	1.00*
	b	-0.02 ^{ns}	-0.07 ^{ns}	-0.16 ^{ns}	-0.13 ^{ns}	-0.06 ^{ns}	-0.03*
	c	0.02 ^{ns}	0.31*	0.36*	0.34*	0.09 ^{ns}	0.10*
	\bar{r}	0.00	0.34	0.45	0.46	0.29	0.44
	95% asymptotic point	-	0.41 (6)	0.59 (8)	0.59 (8)	-	0.50 (7)
R^2	a	-0.01 ^{ns}	0.05*	5.1 ^{ns}	4.47 ^{ns}	5.81*	6.99 ^{ns}
	b	-0.05 ^{ns}	-0.03 ^{ns}	-0.16*	-0.13*	-0.07 ^{ns}	-0.10*
	c	0.62 ^{ns}	0.37 ^{ns}	0.01 ^{ns}	0.01 ^{ns}	0.006 ^{ns}	0.006 ^{ns}
	\bar{r}	0.01	0.04	0.16	0.18	0.12	0.13
	95% asymptotic point	-	-	-	-	-	-

a is the asymptotic predictive ability; b is the predictive ability with number of harvests in the training set equals zero; c is the linear predictive ability by the number of harvests in the training set; \bar{r} represents the mean predictive ability through the increasing number of harvests in the training set. *Significant by the t-test at 5% of probability. The numbers between parentheses represent the number of harvest estimates at 95% of asymptotic predictive ability.

Table S2 - Parameter estimations from fitting asymptotic regression equation $r_{ij} = a_i - (a_i - b_i) * e^{(-c_i X)}$ to mean cross validation predictive ability by number of harvests in the training population for breeding values in each untested harvest (GEBV), broad adaptability (b_0), persistence (b_1) and stability (R^2).

Predicted	Regression Parameter	M0	M1	M2	M3	M4	M5
GEBV	a	0.74*	0.72*	0.70*	0.70*	0.74*	0.74*
	b	0.37*	0.38*	0.23 ^{ns}	0.25 ^{ns}	0.17*	0.18*
	c	0.16 ^{ns}	0.18 ^{ns}	0.61*	0.60*	0.46*	0.44*
	\bar{r}	0.53	0.57	0.64	0.63	0.64	0.64
	95% asymptotic point		0.62 (7)	0.68 (5)	0.68 (5)	0.70 (6)	0.70 (6)
b_0	a	2.1 ^{ns}	0.82*	0.81*	0.81*	0.87*	0.92*
	b	0.26 ^{ns}	0.24*	0.09 ^{ns}	0.10*	0.18 ^{ns}	0.18 ^{ns}
	c	0.01 ^{ns}	0.15*	0.38*	0.38*	0.25*	0.23*
	\bar{r}	0.42	0.53	0.66	0.66	0.64	0.66
	95% asymptotic point	-	0.58 (6)	0.77 (8)	0.77(8)	0.81 (10)	0.84 (10)
b_1	a	0.21 ^{ns}	0.60*	0.74*	0.73*	0.97*	1.00*
	b	0.01 ^{ns}	-0.20*	-0.09*	-0.25*	-0.13 ^{ns}	-0.08*
	c	-0.02 ^{ns}	0.38*	0.36*	0.42*	0.23*	0.18*
	\bar{r}	-0.02	0.43	0.56	0.54	0.58	0.56

	95% asymptotic point	-					
	<i>a</i>	-0.10 ^{ns}	0.54 (3)	0.68 (7)	0.69 (8)	0.79 (8)	0.72 (8)
	<i>b</i>	0.02 ^{ns}	0.59 ^{ns}	6.80 ^{ns}	4.17 ^{ns}	10.68 ^{ns}	9.35 ^{ns}
	<i>c</i>	0.05 ^{ns}	0.00 ^{ns}	-0.20*	-0.17 ^{ns}	-0.16*	-0.14*
R^2	\bar{r}	0.00	0.03 ^{ns}	0.01 ^{ns}	0.01 ^{ns}	0.00 ^{ns}	0.00 ^{ns}
	95% asymptotic point	-	-	-	-	-	-

a is the asymptotic predictive ability; *b* is the predictive ability with number of harvests in the training set equals zero; *c* is the linear predictive ability by the number of harvests in the training set; \bar{r} represents the mean predictive ability through the increasing number of harvests in the training set. *Significant by the t-test at 5% of probability. The numbers between parentheses represent the number of harvest estimates at 95% of asymptotic predictive ability.