

**University of São Paulo
“Luiz de Queiroz” College of Agriculture**

**Unraveling the stability of *Panicum maximum* in multiple
harvest-location trials using a probabilistic model**

Vitória Bizão Murakami

Dissertation presented to obtain the degree of Master
in Science. Area: Genetics and Plant Breeding

**Piracicaba
2024**

Vitória Bizão Murakami
Biotechnologist

**Unraveling the stability of *Panicum maximum* in multiple
harvest-location trials using a probabilistic model**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:

Prof. Dr. **ANTONIO AUGUSTO FRANCO GAR-
CIA**

Dissertation presented to obtain the degree of Master
in Science. Area: Genetics and Plant Breeding

Piracicaba
2024

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

Murakami, Vitória Bizão

Unraveling the stability of *Panicum maximum* in multiple harvest-location trials using a probabilistic model / Vitória Bizão Murakami. -- versão revisada de acordo com a resolução CoPGr 6018 de 2011. -- Piracicaba, 2024 .
44 p.

Dissertação (Mestrado) -- USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Melhoramento de forrageiras 2. Recomendação de cultivar 3. Variância-covariância 4. Modelos Bayesianos . I. Título.

RESUMO

Desvendando a estabilidade do *Panicum maximum* em ensaios de múltiplos ambientes e cortes usando um modelo probabilístico

A avaliação de cultivares em diferentes condições ambientais e anuais é uma fase crítica do melhoramento de culturas perenes. Esses ensaios de múltiplos locais e cortes oferecem informações a respeito das interações genótipo-ambiente (GEI) e genótipo-corte (GHI), que são as principais causas da resposta fenotípica diferencial entre locais e anos. Nesse contexto, os modelos lineares mistos e os modelos bayesianos são úteis para capturar a diversidade da expressão gênica entre locais e colheitas. Dessa forma, os objetivos deste estudo foram: (i) avaliar diferentes estruturas de variância-covariância em um ensaio de múltiplos locais e cortes, e (ii) explorar as interações genótipo-ambiente (GEI) e genótipo-corte (GHI) para avaliar a adaptabilidade e a estabilidade do *Panicum Maximum*. Os dados fenotípicos de matéria seca foliar (LDM) foram medidos em 23 genótipos em um delineamento de blocos completos casualizados com até dezessete cortes em cinco locais. As estruturas de covariância dos efeitos aleatórios foram modeladas e sua conformidade foi testada pelos critérios de informação de Akaike e Bayesiano. A partir do modelo selecionado, foram estimados os componentes de variância, os parâmetros genéticos e as médias ajustadas. Os modelos que levaram em conta a heterogeneidade nas estruturas de variância-covariância foram indicados como os melhores ajustados. Além disso, ajustamos quatro modelos bayesianos com desvios padrão residuais homogêneos (M1, M3) e heterogêneos (M2, M4). Com base no modelo selecionado (M2) pelo WAIC², os genótipos PM40, MASS e PM41 apresentaram as maiores probabilidades globais e pareadas de desempenho superior para massa seca foliar (LDM). Analisando o desempenho entre os locais, o genótipo PM32 apresentou adaptação ao local AC. Observamos no gráfico da norma de reação que a interação genótipo-corte (GHI) apesar de complexa e significativa não conseguiu alterar o genótipo em mais de duas posições na classificação, refletindo a homogeneidade da probabilidade de desempenho ao longo dos cortes. Com relação à estabilidade entre os locais, os genótipos TANZ, PM44 e PM42 foram os melhores. A representação gráfica das probabilidades ofereceu uma compreensão clara dos padrões de adaptação dos genótipos em vários locais e cortes, permitindo a comparação do desempenho dos genótipos. Portanto, nossos resultados auxiliam os processos de tomada de decisão na recomendação de genótipos, reduzindo os riscos de carregar genótipos com baixo desempenho para a próxima fase de melhoramento.

Palavras-chave: Melhoramento de forrageiras, Recomendação de cultivar, Variância-covariância, Modelos Bayesianos

ABSTRACT

Unraveling the stability of *Panicum maximum* in multiple harvest-location trials using a probabilistic model

Evaluating cultivar response under different environmental and annual conditions is a critical phase of perennial crop breeding. These multi-harvest location trials allow access to genotype-by-environment (GEI) and genotype-by-harvest (GHI) interactions, which are the main causes of differential phenotypic response across location and years. In this context, linear mixed models and Bayesian models are useful to capture the expression of genotype diversity across location and harvests. Therefore, the objectives of this study were (i) to evaluate different variance-covariance structures for multiple harvest-location trials, and (ii) to explore the genotype-by-environment (GEI) and genotype-by-harvest (GHI) interactions to assess the adaptability and stability of *Panicum Maximum*. Dry leaf matter phenotypic data were measured in 23 genotypes in a complete randomized block design with up to seventeen harvests in five locations. The covariance structures of the random effects were modeled and their adequacy was tested by the Akaike and Bayesian information criteria. From the selected model, variance components, genetic parameters and adjusted means were estimated. Models that accounted for heterogeneity in the variance-covariance structures were best fitted. We fitted four Bayesian models with homogeneous (M1, M3) and heterogeneous (M2, M4) residual standard deviations. Based on the model selected by WAIC2 (M2), genotypes PM40, MASS, and PM41 had the highest global and pairwise probability of superior performance for LDM. When analyzing the performance within environments, the genotype PM32 showed an adaptation for the site AC. On the reaction norm plot, we observed that the genotype-by-harvest had a complex significant interaction but could not change more than two positions in the rank, reflecting the homogeneity of the probability of performance along harvests. In terms of stability across locations, genotypes TANZ, PM44 and PM42 were the best. The visual representation of probabilities provided straightforward insights into genotype adaptation patterns across environments and harvests, allowing comparison of genotype performance. Therefore, our results support decision making processes when recommending genotypes and reduce the risk of carrying poor performing genotypes into the next breeding phase.

Keywords: Forage breeding, Cultivar recommendation, Variance-Covariance, Bayesian models

1 INTRODUCTION

Over this decade, we have reached a population of 7.9 billion people (ROSER *ET AL.*, 2013), and as the human population grows, we demand more animal protein fonts. In this context, Brazil consolidated itself as the second-largest beef producer in the world, being only behind the United States (FAO, 2021). In 2021, Brazil raised around 200 million cattle heads in a pasture area of 163,1 Mha (ABIEC, 2022). One of the main reasons that brought Brazil to this position was its forage breeding programs with highly adapted and productive forages (JANK *ET AL.*, 2011).

However, even with all this success, only a few commercially available forage cultivars exist. Of the five principal one's, *Brachiaria brizantha* cv. Marandu and *Panicum maximum* (syn. *Megathyrsus maximus* (Jacq.) B. K. Simon & S. W. L. Jacobs) cv. Mombaça represented 70% of all 2018/2019 seed production (OMOTE *ET AL.*, 2021). Additionally, most cultivated forages have apomictic reproduction, a clonal propagation by seed, which roots out the generation of new variations. These factors raise concerns about the pasture's vulnerability to the uprising of new diseases and pests or resistance breaking down (JANK *ET AL.*, 2011). Therefore, developing and using new cultivars is one way to diversify Brazilian's pastures and alleviate those fears. Evaluations of *Panicum maximum* showed that the species has much potential variation to be explored and desirable agronomic traits, such as high productivity and quality, ease of management, great ground cover, and some tolerances (JANK *ET AL.*, 2008).

Brazilian Agricultural Research Corporation (Embrapa) Beef Cattle is responsible for the *P. Maximum* germplasm bank and its breeding program (JANK *ET AL.*, 2014). The program counts with a representative collection of *Panicum maximum*'s natural variation, composed of 426 apomictic accessions and 417 sexual plants. Embrapa Beef Cattle received this collection in 1982 from the French Institution ORSTOM (Institut Français de Recherche Scientifique pour le Développement en Coopération - IRD), which gathered it in 1967 and 1969 in East Africa (JANK *ET AL.*, 2008). Thus, the breeding program is relatively new and has the same goals as the consolidated programs of the major crops: high yield and quality, pest resistance, and adaptation to edaphic and climatic stress. Moreover, forage breeding also aims for high conversion efficiency to animal products (DO VALLE *ET AL.*, 2009), especially for dry leaf matter, because of its digestibility attribute. Specifically for *Panicum maximum*, the breeding target is leaf and seed yield, adaptation, regrowth, and *Bipolaris maydis* resistance (DO VALLE *ET AL.*, 2009; JANK *ET AL.*, 2011).

Currently, the development and release of a new cultivar passes through three stages. In the first stage, breeders test 100 to 200 accessions for desirable characteristics, and the top 20-25 are selected for the next stage. The second stage evaluates the adaptation and performance in target environments and harvests, and only 1 to 3 genotypes pass to the final stage. In the last stage, the genotypes are checked for animal efficiency (JANK *ET AL.*, 2014). Thus, the base of the second stage is multiple harvest-location trials (MHLT), which breeders commonly use when it is desirable to identify cultivars that have consistent efficiency from location to location and/or harvest to harvest (SMITH *ET AL.*, 2005). Genotype-by-environment (GEI) and genotype-by-harvest (GHI) interactions are the main causes of inconsistency in performance per location and harvest (KANG *ET AL.*, 2004).

Genotype-by-environment interaction is a complex phenomenon where genotypes have different responses across environments (ROMAGOSA and FOX, 1993). This interaction significantly impacts the selection of true superior genotypes, resource allocation, breeding strategy, and development of varieties adapted to specific environments (KANG *ET AL.*, 2004; MALOSETTI *ET AL.*, 2013; DE LEON *ET AL.*, 2016; ELIAS *ET AL.*, 2016). Additionally, since *Panicum maximum* is a perennial species, breeders must regard the genotype-by-harvest interaction. Thus, breeding programs need proper analytical tools to understand and measure the GEI and capture the impacts of environmental seasonality, i.e., GHI, to select superior and stable genotypes (MALOSETTI *ET AL.*, 2013; CHAVES *ET AL.*, 2022; ELIAS *ET AL.*, 2016).

Researchers developed several statistical models to analyze genotype-by-environment interaction, such as the additive regression on the mean, the additive and multiplicative interactions (AMMI), and factorial regression (MALOSETTI *ET AL.*, 2013). Those models are generally used for modeling fixed effects; however, linear mixed models have become widely accepted to have some advantages in MHLT analysis (SMITH *ET AL.*, 2005; CROSSA, 2012). Mixed models are superior when dealing with incomplete data and provide great flexibility by fitting different variance-covariance structures for the random effects, allowing us to model genetic and residual correlations between environment and/or harvest (SMITH *ET AL.*, 2005; ELIAS *ET AL.*, 2016; DIAS *ET AL.*, 2018).

Another framework to explore GEI is using Bayesian models. The Bayesian strategy is advantageous because it can handle complex data scenarios with missing data and heterogeneity of variances (e.g., mixed models), can incorporate data from prior studies (JARQUÍN *ET AL.*, 2016; EDWARDS and JANNINK, 2006; DIAS *ET AL.*, 2022), and avoids overparameterization or restrictive assumptions that some models may have (CROSSA *ET AL.*, 2011; JOSSE *ET AL.*, 2014; ANTONIO DE OLIVEIRA *ET AL.*, 2015).

In this sense, DIAS *ET AL.* (2022) proposed a model based on probability concepts to better recommend adapted cultivars for a set of target locations. Therefore, the objectives of this dissertation were (i) to evaluate the modeling of different variance-covariance structures in multiple harvest-location trials, and (ii) to explore *Panicum Maximum* genotype-by-environment (GEI) and genotype-by-harvest (GHI) interactions, assessing its adaptability and stability with Dias et al.'s (2022) probabilistic approach.

2 CONCLUSIONS

In this research, mixed models and probability methods were used to understand the genotype by environment and harvest interactions in the context of Panicum breeding. Models that included heterogeneous variance-covariance structures showed the best goodness-of-fit compared to those that did not. Furthermore, the visual interpretations of probabilities presented are straightforward, providing information on genotypes adaptations patterns across different environments and harvests and comparative performance between genotypes, which can help the decision-making process for cultivar recommendations. The findings of this research emphasize the importance of considering two key aspects of breeding programs. First, select a suitable model that counts with interaction effects. Second, understanding the specific breeding scenario being targeted is essential to tailor the risks of carrying out genotypes with poor performances to the next breeding phases. Overall, this study highlights the significance of these methodologies in understanding and enhancing breeding efforts for Panicum.

References

- ABIEC, 2022 Beef report: perfil da pecuária no brasil.
- ANTONIO DE OLIVEIRA, L., C. PEREIRA DA SILVA, J. J. NUVUNGA, A. Q. DA SILVA, and M. BALESTRE, 2015 Credible intervals for scores in the ammi with random effects for genotype. *Crop Science* **55**: 465–476.
- CHAVES, S. F., J. S. EVANGELISTA, R. S. ALVES, F. M. FERREIRA, L. A. DIAS, R. M. ALVES, K. O. DIAS, and L. L. BHERING, 2022 Application of linear mixed models for multiple harvest/site trial analyses in perennial plant breeding. *Tree Genetics & Genomes* **18**: 1–12.
- CROSSA, J., 2012 From genotype x environment interaction to gene x environment interaction. *Current Genomics* **13**: 225–244.
- CROSSA, J., S. PEREZ-ELIZALDE, D. JARQUIN, J. M. COTES, K. VIELE, G. LIU, and P. L. CORNELIUS, 2011 Bayesian estimation of the additive main effects and multiplicative interaction model. *Crop Science* **51**: 1458–1469.
- DE LEON, N., J.-L. JANNINK, J. W. EDWARDS, and S. M. KAEPLER, 2016 Introduction to a special issue on genotype by environment interaction. *Crop Science* **56**: 2081–2089.
- DIAS, K. O., J. P. DOS SANTOS, M. D. KRAUSE, H.-P. PIEPHO, L. J. GUIMARÃES, M. M. PASTINA, and A. A. GARCIA, 2022 Leveraging probability concepts for cultivar recommendation in multi-environment trials. *Theoretical and Applied Genetics* **135**: 1385–1399.
- DIAS, K. O. D. G., S. A. GEZAN, C. T. GUIMARÃES, S. N. PARENTONI, P. E. D. O. GUIMARÃES, N. P. CARNEIRO, A. F. PORTUGAL, E. A. BASTOS, M. J. CARDOSO, C. D. O. ANONI, *ET AL.*, 2018 Estimating genotype \times environment interaction for and genetic correlations among drought tolerance traits in maize via factor analytic multiplicative mixed models. *Crop Science* **58**: 72–83.
- DO VALLE, C. B., L. JANK, and R. M. S. RESENDE, 2009 O melhoramento de forrageiras tropicais no brasil. *Revista Ceres* **56**: 460–472.
- EDWARDS, J. W. and J.-L. JANNINK, 2006 Bayesian modeling of heterogeneous error and genotype \times environment interaction variances. *Crop Science* **46**: 820–833.
- ELIAS, A. A., K. R. ROBBINS, R. DOERGE, and M. R. TUINSTRA, 2016 Half a century of studying genotype \times environment interactions in plant breeding experiments. *Crop Science* **56**: 2090–2105.

- FAO, 2021 Crops and livestock products. FAOSTAT .
- JANK, L., S. C. BARRIOS, C. B. DO VALLE, R. M. SIMEÃO, and G. F. ALVES, 2014 The value of improved pastures to brazilian beef production. *Crop and Pasture Science* **65**: 1132–1137.
- JANK, L., R. M. S. RESENDE, C. D. VALLE, M. D. RESENDE, L. CHIARI, L. J. CANÇADO, and C. SIMIONI, 2008 Melhoramento genético de *panicum maximum*. *Melhoramento de forrageiras tropicais* **1**: 55–87.
- JANK, L., C. D. VALLE, and R. RESENDE, 2011 Breeding tropical forages. *Crop Breeding and Applied Biotechnology* **11**: 27–34.
- JARQUÍN, D., S. PÉREZ-ELIZALDE, J. BURGUEÑO, and J. CROSSA, 2016 A hierarchical bayesian estimation model for multienvironment plant breeding trials in successive years. *Crop science* **56**: 2260–2276.
- JOSSE, J., F. VAN EEUWIJK, H.-P. PIEPHO, and J.-B. DENIS, 2014 Another look at bayesian analysis of ammi models for genotype-environment data. *Journal of Agricultural, Biological, and Environmental Statistics* **19**: 240–257.
- KANG, M. S., V. PRABHAKARAN, and R. MEHRA, 2004 Genotype-by-environment interaction in crop improvement. *Plant breeding: Mendelian to molecular approaches* pp. 535–572.
- MALOSETTI, M., J.-M. RIBAUT, and F. A. VAN EEUWIJK, 2013 The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Frontiers in physiology* **4**: 44.
- OMOTE, H., L. DE CASTRO, V. GRACIANO, R. D. SANTOS, J. DE SOUZA, A. VAZ, A. DE ARAUJO, A. SALMAN, L. BUENO, M. KOPP, *ET AL.*, 2021 Monitoramento tecnológico de cultivares de forrageiras no brasil. São Carlos, SP: Embrapa Pecuária Sudeste .
- ROMAGOSA, I. and P. FOX, 1993 Genotype× environment interaction and adaptation. *Plant breeding: principles and prospects* pp. 373–390.
- ROSER, M., H. RITCHIE, E. ORTIZ-OSPINA, and L. RODÉS-GUIRAO, 2013 World population growth. *Our World in Data* <https://ourworldindata.org/world-population-growth>.
- SMITH, A., B. R. CULLIS, and R. THOMPSON, 2005 The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science* **143**: 449–462.