BRAZILIAN JOURNAL OF BIOMETRICS

ISSN:2764-5290

ARTICLE

Special issue on (bio)statistics and biometrics in the age of the digital revolution

Rodrigo R. Pescim★,[1] and Luiz R. Nakamura[2]

[1]Departamento de Estatística, Universidade Estadual de Londrina, Londrina, Brazil
[2]Departamento de Estatística, Universidade Federal de Lavras, Lavras, Brazil
★Corresponding author. Email: rrpescim@uel.br,luiz.nakamura@ufla.br

Abstract

This special issue explores the thematic areas presented at the 67th Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBras) and 20th Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

Keywords: Biometrics; Data science; Statistics.

It is with great pleasure that we present to you this special issue of the Brazilian Journal of Biometrics (BJB), which we had the privilege of coordinating as Guest Editors. The seven papers in this issue are dedicated to the thematic areas presented at the 67th Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBras) and 20th Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO), which took place in Londrina, Paraná, Brazil, from 24th to 28th July, 2023.

Barros *et al.* (2024) compared the predictive performance of Logistic Regression, Decision Tree and Random Forest (three important machine learning methods) in identifying sugarcane areas using Landsat images within a study area of approximately 306,000 ha. Their conclusion was based on those methods performance and the Random Forest was the most recommended method among those investigated for sugarcane mapping, considering the spectral indices collected during a 17-month period, from January 2015 to May 2016.

Bueno *et al.* (2024) considered two polynomial item response theory models, the graded response model (GRM) and the generalised partial credit model (GPCM), to analyse the environmental sustainability perception of 2,519 residents in Paraná Basin III, which includes Cascavel, Foz do Iguaçu, and Toledo cities. The GRM outperformed the GCPM, providing a better estimation

for the 32 items in the dataset according to the Akaike and Bayesian Information criteria. Key findings include differences in sustainability perception among the three cities in study, with male having higher levels of perception, as well as a positive association between higher remuneration and increased sustainability perception.

Lima *et al.* (2024) evaluated the presence of Hardy-Weinberg equilibrium in a population of 49 piracanjuba fish (*Brycon orbignyanus*) by estimating the allele proportions and expected genotype proportions using frequentist, Bayesian, and sequential Bayesian approaches. The authors showed that sequential strategy was especially noteworthy, achieving remarkable efficiency with just 55.1% of the data (i.e., 27 fish), reducing sample size and optimising the procedure.

Oliveira *et al.* (2024) showed the effectiveness of using both univariate and bivariate linear mixed models to analyse the weight and total solid feed intake behaviour of 35 Holstein calves over eight weeks across three different diets, considering a notably unbalanced data scenario. Their findings highlight the advantages of the bivariate approach, which produces estimates with smaller standard errors. Furthermore, the authors claim that least confounded residuals are preferable to both studentised marginal and conditional residuals, since they presented lesser biased behaviour.

Portela *et al.* (2024) proposed a new frailty model to analyse failure times of repairable systems subject to unobserved heterogeneity. They considered the non-central gamma distribution for the frailty random effect using data from sugarcane harvesters and literature systems. They also ran a simulation study across multiple scenarios to verify the quality and behaviour of the obtained estimators. Their conclusion highlight the capability of the proposed approach in providing useful information for decision-making in complex systems, such as forecasting till the first failure.

Rodrigues *et al.* (2024) used the Bayesian approach to compare the Logistic and Gompertz nonlinear models for describing the diameter and length growth curves of the Asian pear tree (Shinseiki cultivar). The predictive accuracy of these models was assessed using the expected log pointwise predictive density (elpd) measure. Their investigation that both models exhibited good chain convergence, however, according to the elpd criterion, the Logistic model was chosen as the best. The adequacy of the chosen model was also confirmed by residual analysis.

Teixeira *et al.* (2024) used three different two-parameter distributions (gamma, Gaussian, and log-normal) to model monthly coffee bag prices from January 1981 to December 2022, with the last 10 years serving as test data. Using the Kolmogorov-Smirnov and chi-square tests, as well as the mean absolute percentage error, they determined that the log-normal distribution consistently outperformed its alternatives across all months. Notably, the authors identify January, February, and March as the months with the highest probability of experiencing raised average prices, placing them as ideal times to engage in coffee trading activities.

We would like to express our heartfelt appreciation to all of the authors for their outstanding work and dedication, as well as the reviewers for their careful inspection and useful comments. Without the dedication of these individuals, this volume would not have met its high-quality standards.

We hope that this special issue inspires and enlighten people interested in biometrics and applied statistics. May the insights offered here inspire more research and major advances in our fields of study.

References

1. Barros, A. C. A. V. B., Silva, M. A., Marques, R. D., Villegas, C & Luciano, A. C. S. Classification of sugarcane areas in Landsat images using machine learning algorithms. *Brazilian Journal of Biometrics* **42,** 147–157 (2024).

2. Bueno, R. B., Silva, M. A., Possan, E & Vincenzi, S. Analysis of environmental sustainability perception in Paraná Basin III residents via Item Response Theory. *Brazilian Journal of Biometrics* **42,** 158–170 (2024).

3. Lima, I. S., Brighenti, C. R. G. & Yazbeck, G. M. Sequential Bayesian approach for genetic diversity analysis of the piracanjuba fish (*Brycon orbignyanus*). *Brazilian Journal of Biometrics* **42,** 171–181 (2024).

4. Oliveira, A. Q., Poczynek, M, Bittar, C. M. M. & Lima, C. G. Linear mixed-effects models and least confounded residuals in the modelling of Holstein calves' performance. *Brazilian Journal of Biometrics* **42,** 103–118 (2024).

5. Portela, A. C. T., Brito, E. S., Tomazella, V. L. D., Diniz, C. A. R. & Ferreira, P. H. Reliability of repairable systems with Non-Central Gamma frailty. *Brazilian Journal of Biometrics* **42,** 182–201 (2024).

6. Rodrigues, G. M., Villegas, C & Savian, T. V. Bayesian analysis of the growth of pears in the cultivate Shinseiki using non-linear models. *Brazilian Journal of Biometrics* **42,** 131–146 (2024).

7. Teixeira, O. M., Santos, J. A. & Liska, G. R. Probabilistic Modeling of the Coffee Market in Brazil. *Brazilian Journal of Biometrics* **42,** 119–130 (2024).