

Consolidating INIA's Rice Breeding Program Database, Phase I: Historical Indica Trials

S Scheffel^{1,5}, I Rebollo^{2,5}, F Pérez³ and JE Rosas^{4,5}

¹ Instituto Nacional de Investigación Agropecuaria (INIA), Uruguay, scheffelsheila@gmail.com

² Instituto Nacional de Investigación Agropecuaria (INIA), Uruguay, irebollo@inia.org.uy

³ Instituto Nacional de Investigación Agropecuaria (INIA), Uruguay, fperez@inia.org.uy

⁴ Instituto Nacional de Investigación Agropecuaria (INIA), Uruguay, jrosas@inia.org.uy

⁵ Facultad de Agronomía, Universidad de la República, Uruguay

Introduction

INIA's Rice Breeding Program (IRBP) is the major rice breeding program in Uruguay. To deliver improved cultivars, the performance of thousands of experimental lines is assessed in field and laboratory trials.

The joint analysis of all the experimental data, combining multiple years, locations and trials, allows better estimations of the genetic value of the breeding lines and thus optimizes the efficiency of the breeding program. Records of historic field and laboratory trials of the IRBP are fragmented in a multiplicity of media and formats that prevent their joint analysis.

This work describes the integration of a database corresponding to the Indica type materials, which constitutes approximately half of the IRBP germplasm. It comprises data from field and lab trials across 14 years (2006 to 2019) and three locations.

Objective

Present the methodology used to consolidate IRBP's database, as well as preliminary analysis' results.

Material and Methods

The strategy used was:

- 1) Implementation of uniform templates with standardized field names.
- 2) Manual review and quality control of data (variable identification and assignation to a standardized field, check units of measurement, check cell format, identification of missing data, and standardized annotation of qualitative observations).
- 3) Reading of the formatted spreadsheets of each trial with the R software, consolidating all trials in a single data frame with all the Indica IRBP data.
- 4) Standardization of field values (missing data and standardized levels for categorical variables).
- 5) Trial-wise quality control with statistical criteria: distribution of variables, trial heritability, best linear unbiased predictions (BLUPs) of genotypic breeding values. Trial heritability was estimated following Falconer and Mackay (1996) with the equation:

$$h^2 = \frac{V_G}{V_G + V_R/rep}$$

V_G : genetic variance, V_R : residual variance, rep: number of repetitions.

BLUPs were obtained fitting the linear mixed model for each trial:

$$y_{ij} = G_i + block_j + e_{ij}$$

y_{ij} : observed phenotypic value of the i^{th} genotype in the j^{th} block, G_i : random genotypic effect of the i^{th} genotype, $block_j$: fixed effect of the j^{th} block, e_{ij} : residual random effect of the i^{th} genotype in the j^{th} block.

Results

A database with more than 27,416 records was obtained, comprising 270 trials in 14 years, with 5,126 lines evaluated. A total of 35 variables related to phenology, yield and grain quality, disease incidence and pedigree were considered.

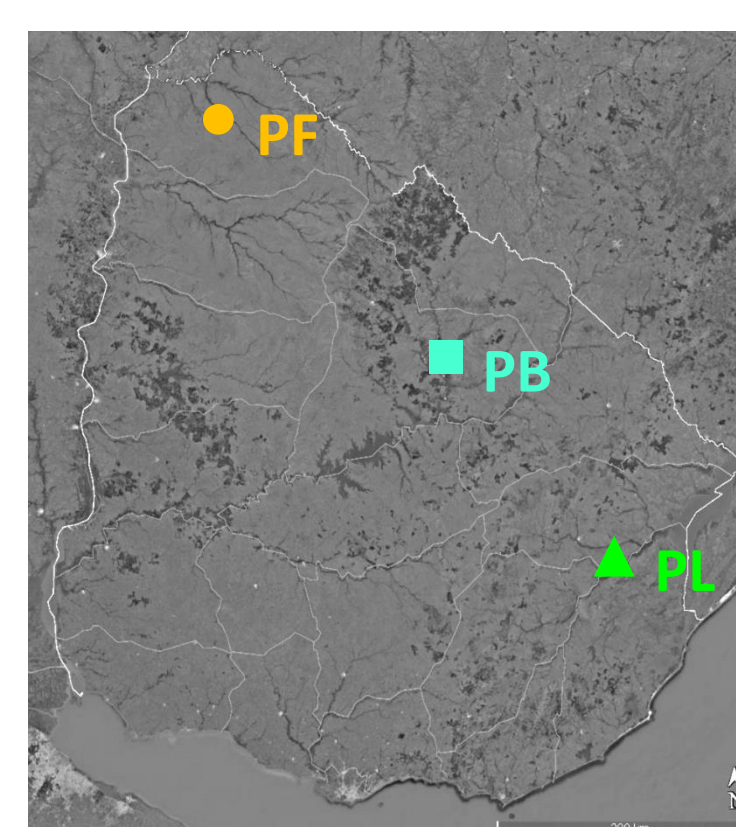


Figure 1. Locations of IRBP Experimental fields: *Paso Farías* (PF, advanced and final evaluations); *Pueblo del Barro* (PB, advanced and final evaluations); *Paso de la Laguna* (PL, all stages).

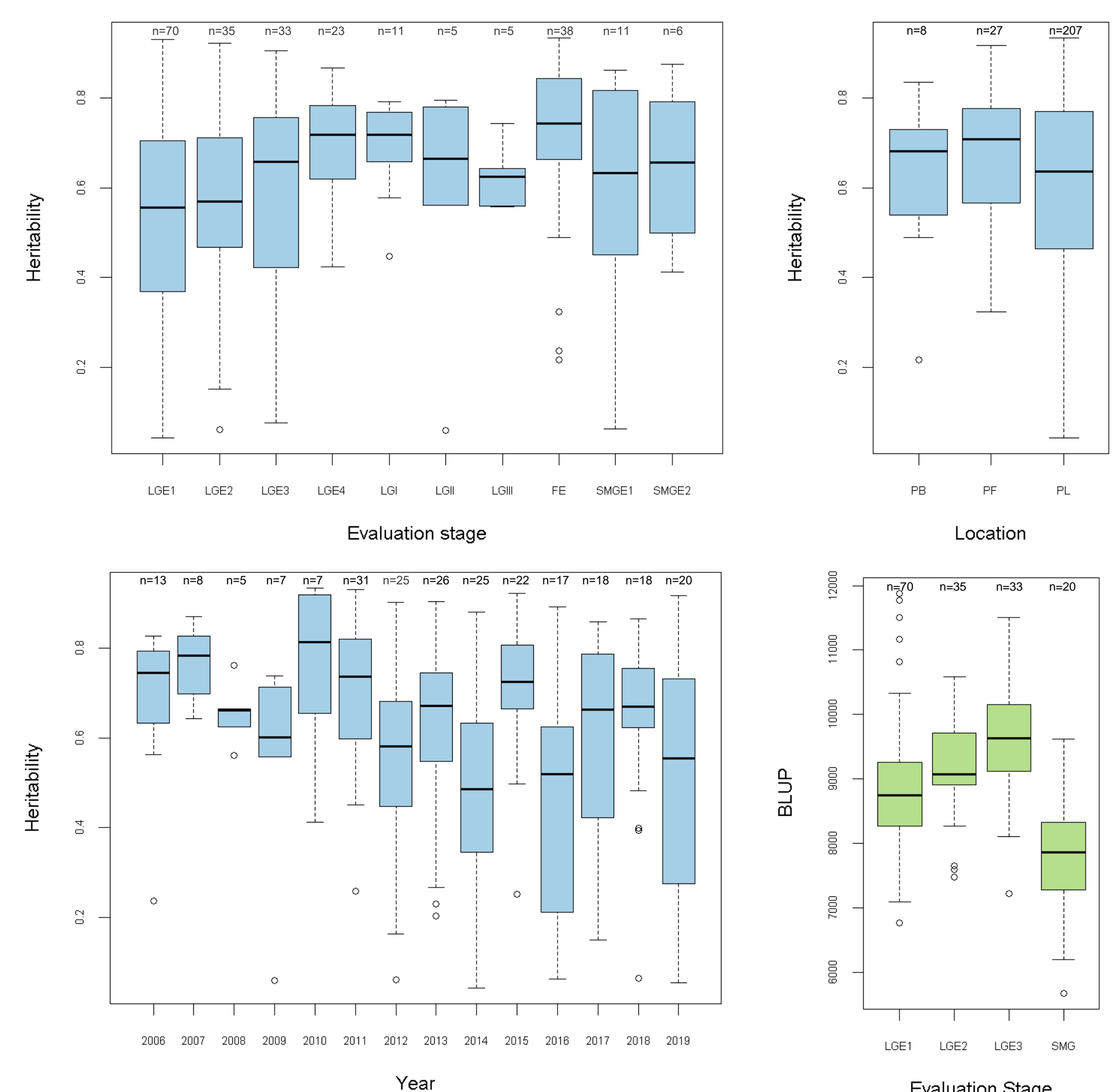


Figure 2. (a) Heritability by evaluation stage, (b) Heritability by location, (c) Heritability by year, (d) Trial-wise BLUPs by evaluation stage.

LGE1 to LGE4: Long grain (indica type) yield evaluations years 1st to 4th; FE: Final evaluation (indica type); SMGE1/E2: short medium grain (temperate japonica type) yield evaluation years 1st and 2nd; PL: Experimental Unit *Paso de la Laguna*; PB: Experimental field *Pueblo del Barro*; PF: Experimental field *Paso Farías*.

Conclusions and Perspectives

The trial-wise genetic parameters estimated in this work will be used for a quality control of trial data, to select good quality trials for the joint analysis. The consolidated data will be crucial for implementing joint modeling for multiple-environment analysis that will enhance selection accuracy and improve genetic gain in the IRBP.

References

Falconer, D.S. and Mackay, T.F.C. (1996) Introduction to Quantitative Genetics. 4th Edition, Prentice Hall, Essex.