



UNIVERSIDAD
NACIONAL
DE COLOMBIA

**Evaluación de la respuesta del genotipo al
ambiente en rendimiento y calidad de papa
amarilla diploide (*Solanum tuberosum*, Grupo
Phureja)**

Harverth Silva Herrera

Universidad Nacional de Colombia

Facultad de Ciencias Agrarias

Bogotá, Colombia

2022

Evaluación de la respuesta del genotipo al ambiente en rendimiento y calidad de papa amarilla diploide (*Solanum tuberosum*, Grupo Phureja)

Harverth Silva Herrera

Trabajo de grado para optar por el título de:

Magister en Ciencias Agrarias

Línea de investigación en Genética y Fitomejoramiento

Director:

PhD. José Miguel Cotes Torres

Codirector:

PhD. Luis Ernesto Rodríguez Molano

Universidad Nacional de Colombia

Facultad de Ciencias Agrarias

Bogotá, Colombia

2022

A mis padres Jarber y Maria, mis hermanas Marcela y Cindy, mi hermano Bryan y mi sobrino Samuel, por su apoyo incondicional, acompañamiento y guía en mi proceso de formación y desarrollo profesional. Han sido testigos de mi disciplina académica y de mi pasión por ser un investigador con renombre internacional, contribuyendo a la ciencia y al avance de la humanidad.

*“Antes creíamos que nuestro futuro estaba en las estrellas. Ahora sabemos que está en nuestros genes”
James D. Watson*

Agradecimientos

Al Doctor José Miguel Cotes Torres, por su dedicación y acompañamiento en la ejecución de esta investigación y la escritura del presente documento.

Al Doctor Luis Ernesto Rodríguez Molano, por brindar la información y acompañamiento en la escritura y revisión del presente documento.

Al grupo de investigación de Mejoramiento de Especies Andinas y Tropicales por el aporte de información y enriquecimiento de la discusión del documento.

Al grupo de investigación en Mejoramiento Genético de Papa de la Universidad Nacional de Colombia sede Bogotá y la Universidad McGill de Canadá, quienes por medio del proyecto “Mejoramiento de la Seguridad Alimentaria a través de la Producción de Papa”

A los colegas Juliana Gómez González y Diego Felipe Conejo, por su ayuda en los procesos de redacción y análisis de la información. A mi amigo Felipe Ramírez por el apoyo y entusiasmo brindado.

A las facultades de Ciencias Agrarias de la sede Bogotá y Ciencias Agropecuarias de la sede Medellín de la Universidad Nacional de Colombia por brindarme el espacio y conocimiento requerido para la ejecución de esta investigación

Resumen

Evaluación de la respuesta del genotipo al ambiente en rendimiento y calidad de papa amarilla diploide (*Solanum tuberosum*, Grupo Phureja)

La IGA es un modelo teórico utilizado para explicar la respuesta diferencial de los genotipos al ambiente. Varios métodos estadísticos han sido desarrollados para descomponer la respuesta fenotípica basados principalmente en las medias generales, en el efecto del genotipo y su interacción con el ambiente, siendo tratada la varianza ambiental como factor de confusión. Sin embargo, la estimación del efecto ambiental a partir de los rasgos evaluados establecería una dependencia al genotipo, resultando en un menor ajuste y potencial predictivo del modelo. Se ha propuesto la integración de covariables ambientales a los modelos que asocien características edafoclimáticas con los rasgos de interés, con el propósito de aumentar el potencial predictivo y la varianza contenida por los modelos. El objetivo de este estudio fue evaluar la sensibilidad de los rasgos de interés en cultivos de papa amarilla diploide a covariables ambientales, seleccionando las covariables más relevantes como parámetros en modelos empíricos de regresión múltiple a partir de la varianza ambiental. Los resultados mostraron una alta variabilidad del rendimiento debida a covariables del componente hídrico, mientras que los rasgos de calidad se vieron principalmente afectados por rasgos de los componentes energéticos y fisicoquímicos del suelo. Los modelos ajustados explicaron la varianza debida intrínsecamente al ambiente, alcanzando ajustes superiores al 20%. Por lo tanto, se concluye que los rasgos presentan una alta sensibilidad fenotípica y la incorporación de covariables ambientales a los modelos de análisis de interacción genotipo por ambiente, podrían mejorar la comprensión de la estabilidad y adaptabilidad de los cultivares a partir de los datos obtenidos en pruebas multiambiente.

Palabras clave: Interacción genotipo ambiente, covariables ambientales, modelos lineales mixtos generalizados, respuesta fenotípica, epigenética.

Abstract

Evaluation of genotype response to environment on yield and quality of diploid yellow potato (*Solanum tuberosum*, Phureja Group).

The IGA is a theoretical model used to explain the differential response of genotypes to the environment. Several statistical methods have been developed to decompose the phenotypic response based mainly on the general means, on the effect of the genotype and its interaction with the environment, treating the environmental variance as a confounding factor. However, the estimation of the environmental effect from the traits evaluated would establish a dependency on the genotype, resulting in a lower fit and predictive potential of the model. The integration of environmental covariates to the models that associate edaphoclimatic characteristics with the traits of interest has been proposed, with the purpose of increasing the predictive potential and the variance contained by the models. The aim of this study was to evaluate the sensitivity of the traits of interest in diploid yellow potato crops to environmental covariates, selecting the most relevant covariates as parameters in empirical multiple regression models based on environmental variance. The results showed a high yield variability due to covariates of the water component, while the quality traits were mainly affected by traits of the energetic and physicochemical components of the soil. The adjusted models explained the variance due intrinsically to the environment, reaching adjustments greater than 20%. Therefore, it is concluded that the traits have a high phenotypic sensitivity and the incorporation of environmental covariates to the genotype-by-environment interaction analysis models could improve the understanding of the stability and adaptability of cultivars from the data obtained in multi-environment trials.

Key words: genotype by environment interaction, environmental covariates, generalized linear mixed models, phenotypic response, epigenetics.

Contenido

	Pág.
Resumen y Abstract.....	IX
Lista de tablas	XIII
Lista de Símbolos y abreviaturas.....	XV
Introducción	1
Referencias	3
Objetivos.....	7
Objetivo General.....	7
Objetivos Específicos.....	7
Capítulo 1. Marco teórico.....	10
1.1 Respuesta Fenotípica	10
1.2 Análisis de estabilidad y adaptabilidad fenotípica.....	11
1.3 Métodos lineales.....	11
1.4 Métodos bilineales.....	13
1.5 Métodos mixtos	15
1.6 Caracterización del ambiente en covariables	16
1.7 Interacción genotipo – ambiente en papa amarilla diploide	17
1.8 Referencias.....	18
Capítulo 2. Conocimiento del microambiente y su efecto en la respuesta fenotípica de papa amarilla diploide (Solanum tuberosum, Grupo Phureja).....	25
Resumen	25
Abstract.....	26
2.1 Introducción	27
2.2 Materiales y métodos	28
2.2.1 Base de Datos.....	28
2.2.2 Material vegetal	31
2.2.3 Componente hídrico y energético	31
2.2.4 Métodos Analíticos	38
2.3 Resultados y discusión	41
2.3.1 Sensibilidad Fenotípica del Rasgo al Ambiente	41
2.3.2 Modelos Ajustados	55
2.4 Conclusión	59
2.5 Referencias.....	59

XII Evaluación de la respuesta del genotipo al ambiente en rendimiento y calidad de
cultivares de papa amarilla diploide (*Solanum tuberosum*, Grupo Phureja)

- A. Anexo: Código para la importación y preparación de bases de datos de respuesta fenotípica, información edafoclimática de cada localidad y fechas de siembra y cosecha.....67**
- B. Anexo: Código para la preparación de la información ecofisiológica de cada una de las localidades evaluadas.....69**
- C. Anexo: Código para la verificación de interacción genotipo - ambiente por cambio de rango (COI) o heterogeneidad de varianzas (HV).....77**
- D. Anexo: Código para la ejecución de correlaciones, análisis de componentes principales, regresiones simples y múltiples79**

Lista de tablas

	Pág.
Tabla 1. Ubicación de las pruebas multiambiente de papa amarilla diploide en 11 localidades de los departamentos de Nariño y Cundinamarca, para los años 2012-2013.	21
Tabla 2. Genealogía de los genotipos de papa amarilla diploide evaluados en la prueba multiambiente.	22
Tabla 3. Parámetros químicos del suelo de las localidades de la prueba multiambiente de papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013.	24
Tabla 4. Parámetros físicos del suelo de las localidades de la prueba multiambiente de papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013.	25
Tabla 5. Componente hídrico y energético de las localidades de la prueba multiambiente de papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013.	27
Tabla 6. Cuantificación del coeficiente de determinación (R^2) entre el rasgo fenotípico y las covariables ambientales en papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013. Coloraciones verdes indican mayor ajuste del modelo, rojizas menor ajuste del modelo al interior de cada rasgo.	32
Tabla 7. Estimación de la diferencia predictiva de la raíz cuadrada media (RMSPD) entre el rasgo fenotípico y las covariables ambientales en papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013. Coloraciones verdes indican menor precisión del modelo, rojizas mayor precisión del modelo al interior de cada rasgo.	34
Tabla 8. Correlación lineal entre las covariables del componente hídrico de la varianza ambiental.	38
Tabla 9. Correlación lineal entre las covariables del componente energético de la varianza ambiental.	40
Tabla 10. Matriz de valores propios del ACP entre las covariables del componente fisicoquímico del suelo de la varianza ambiental. Coloraciones verdes indican mayor peso de la	

XIV Evaluación de la respuesta del genotipo al ambiente en rendimiento y calidad de
cultivares de papa amarilla diploide (*Solanum tuberosum*, Grupo Phureja)

covariable en la CP con signo positivo, rojizas mayor peso de la covariable en la CP.
Coloraciones amarillas indican un bajo peso de la covariable en la CP..... 42

Tabla 11. Coeficientes de determinación y diferencia predictiva de la raíz cuadrada media para
3 modelos de estimación de la respuesta fenotípica en cultivos de papa amarilla diploide
ajustados a partir de diferentes parámetros ambientales..... 44

Lista de Símbolos y abreviaturas

Símbolos con letras latinas

Símbolo	Término
\mathbf{b}	Vector de parámetros fijos
\mathbf{I}	Matriz de identidad (1's)
\mathbf{u}	Vector de parámetros aleatorios
\bar{x}	Valor estimado para la media del genotipo
\mathbf{X}	Matriz de diseño para genotipos
\mathbf{Z}	Matriz de diseño para localidades
Z_r	Profundidad radical
Z_i	Eje del componente principal

Símbolos con letras griegas

Símbolo	Término
α	Vectores singulares de los genotipos (Eigenvectors)
α_j	Efecto principal del componente hídrico
β	Coefficiente de regresión
β_j	Efecto principal del componente energético
γ	Vectores singulares de las localidades (Eigenvectors)
δ	Desviación de la regresión
$\hat{\delta}$	Valor estimado de la desviación de la regresión
δ_j	Efecto principal del componente físico químico del suelo
ε	Error residual
θ	Efecto principal de las localidades
θ_{FC}	Contenido de humedad a capacidad de campo
θ_{WP}	Contenido de humedad en el punto de marchitez permanente
λ	Valores singulares (Eigenvalues)

Símbolo	Término
μ	Media general
ρ_{xy}	Intensidad de la correlación lineal entre las variables x y
$\sigma_{d_i}^2$	parámetro de estabilidad de Eberhart y Russell
σ_x	Desviación estándar de la covariable x
σ_y	Desviación estándar de la covariable y
τ	Efecto principal del genotipo
ϕ	Pesos ponderados de la variable x en la componente principal Z

Subíndices

Subíndice	Término
i	I-ésimo término de un vector de genotipos
j	J-ésimo término de vector de localidades
k	K-ésimo término del vector SVD
x	Variable independiente-explicativa
y	Variable dependiente-explicada

Abreviaturas

Abreviatura	Término
ACP	Análisis de componentes principales
AFA	Agua fácilmente aprovechable
AH ₂ O	Almacenamiento de agua en el suelo
AI	Acides intercambiable
AMMI	Modelo de efectos aditivos e interacción multiplicativa
ANOVA	Análisis de varianza
AR	Azúcares reductores
BLUE	Mejor estimador lineal insesgado
BLUP	Mejor predictor lineal insesgado
CAC	Consumo de agua del cultivo
CAP	Consumo de agua de las plantas
CF	Calidad de fritura
CICE	Capacidad de intercambio catiónico efectiva
CO	Carbono orgánico
COI	interacción cruzada

Abreviatura	Término
COV_{xy}	Covarianza entre las variables x y
CP	Componente principal
DSAD	Días sin agua disponible
e_{ij}	Error residual del modelo
Ga_{ij}	Interacción genotipo por ambiente
GE	Gravedad específica
G_i	Efecto principal del genotipo i
HV	Heterogeneidad de varianzas
IAF	Índice de área foliar
IGA	Interacción genotipo por ambiente
I_j	Índice ambiental de la j-ésima localidad
K_e	Coefficiente de extinción
LOO	Leave one out
MLG	Modelos lineales generalizados
MLGM	Modelos lineales generalizados mixtos
MS	Materia seca
NPS	Numero de periodos de sequia
PARAi	Radiación fotosintéticamente activa incidente en el dosel
PM	Prueba multiambiente
PYVV	Virus del amarillamiento de venas de la papa
R	Coefficiente de correlación de Pearson
R^2	Coefficiente de determinación
R^2_{Ad}	Coefficiente de determinación ajustado
RADA	Radiación acumulada en el ciclo de cultivo
RMSPD	Diferencia predictiva de la raíz cuadrada media
S^2_{di}	Estimador del parámetro de estabilidad de Eberhart y Russell
SC	Suma de cuadrados
SCR	Suma de cuadrados de la regresión
SCT	Suma de cuadrados totales
SREG	Regresión por sitios
SVD	Descomposición de valores singulares
\hat{y}	Valor estimado de la respuesta fenotípica
y_{ij}	Respuesta fenotípica del i-ésimo genotipo en el j-ésimo ambiente
y_{Pre}	Respuesta fenotípica predicha por los modelos ajustados
y_{Obs}	Respuesta fenotípica observada por los modelos ajustados

Introducción

En los programas de fitomejoramiento la selección de cultivares superiores como parentales, se realiza evaluando diferentes genotipos en las zonas de interés para su producción (ambiente de la población objetivo), estas evaluaciones son conocidas comúnmente como evaluaciones de rendimiento o pruebas multiambientes (PM) (Theobald *et al.* 2002; Lozano *et al.* 2009; Silva *et al.* 2015; Pour *et al.* 2022). En las PM se recopila información referente a la respuesta media fenotípica de los cultivares en diferentes regiones, de esta forma, se puede establecer cuantitativamente la variación de la respuesta fenotípica debida al ambiente, esta variación debida al ambiente es conocida en los programas de fitomejoramiento como interacción genotipo por ambiente (IGA) (Silva *et al.* 2015; Omer *et al.* 2015; Omer & Singh, 2017; Dias *et al.* 2022). La IGA es un modelo estadístico que permite descomponer la respuesta fenotípica de los cultivares en valores cuantitativos que representan los efectos principales debidos intrínsecamente al genotipo o a su interacción con el ambiente, siendo de gran importancia en los programas de mejoramiento, ya que esta puede ocasionar retrasos en los procesos de avance genético (Crossa & Cornelius 1997; Crossa *et al.* 2011; Omer, & Singh, 2017). Sin embargo, este modelo estadístico carece de soporte real desde la teoría genética convencional, es decir, el ¿por qué se producen los desvíos de la media del genotipo al ambiente (Falconer, 1970) sigue siendo un desafío por esclarecer en la genética moderna.

Varios modelos estadísticos han sido desarrollados para la descomposición de la varianza y modelamiento de la respuesta fenotípica, los primeros modelos estadísticos desarrollados mediante el uso de métodos de regresión, fueron los métodos lineales en donde el rendimiento es la variable de interés y acorde al valor obtenido en cada localidad se asigna una valoración numérica al ambiente (índice ambiental), la cual será más grande de acuerdo a la magnitud de afectación y estará acompañada de un símbolo positivo si es favorable para un cultivar, en caso contrario será negativo, seleccionando de esta forma cultivares con alto rendimiento y capacidad adaptativa en subregiones que

presentan condiciones climáticas similares (Finlay & Wilkinson, 1963). Posteriormente, estos modelos se complementaron utilizando la varianza entre localidades para determinar la estabilidad de cada cultivar (Eberhart & Russell, 1966). Para cuantificar la adaptabilidad, se desarrollaron modelos que mediante la descomposición de la interacción genotipo-ambiente (IGA), permitieron determinar la estabilidad en los diferentes cultivares evaluados (Shukla, 1972). Otros modelos estadísticos desarrollados para la evaluación de datos de PM son los modelos fijos bilineales, tales como AMMI y SREG, que permiten trabajar con tablas de 2 entradas (genotipos y localidades) y además brindan la posibilidad de observar estos datos gráficamente (Theobald *et al.* 2002; Josse *et al.* 2014; Pour *et al.* 2022). Recientemente estos métodos han sido desarrollados con un enfoque bayesiano, brindando mayor flexibilidad en la modelación de datos desbalanceados, con mayor potencia predictiva (Viele & Srinivasan 2000; Crossa *et al.* 2011; Josse *et al.* 2014; Omer *et al.* 2015; Omer & Singh, 2017).

Los modelos estadísticos disponibles para el análisis de PM, pueden ser confusos para los fitomejoradores y no explican la varianza asociada a las variables ambientales, siendo utilizados para selección de cultivares superiores en localidades específicas o macroambientes (Yan *et al.* 2007; Camargo *et al.* 2011), estos a su vez, clasifican la respuesta al ambiente como factor de confusión, disminuyendo la capacidad predictiva de los modelos ajustados (Gauch & Zobel 1988). La variación debida al ambiente podría ser descompuesta en los componentes hídrico (precipitación, humedad relativa, evapotranspiración), energético (altitud, radiación neta, radiación incidente, radiación fotosintéticamente activa) y propiedades físico-químicas del suelo (pH, textura, estructura, contenido de micro y macronutrientes) describiendo en detalle la varianza no explicada de la respuesta fenotípica (Aldabe & Dogliotti, 2009; Tolhurst *et al.* 2021). La plasticidad descrita por estos componentes, determinaría el grado de susceptibilidad de cada rasgo de interés, permitiendo conocer las características específicas que afectarían en mayor o menor proporción la respuesta fenotípica y aumentando la precisión de los modelos ajustados (Lee *et al.* 2016).

En la zona intertropical, la altitud es un factor que influye en diferentes variables ambientales tales como, precipitación, concentración de CO₂, radiación y temperatura; las cuales determinan el crecimiento y desarrollo de las plantas. En Colombia se han desarrollado pruebas multiambiente con especies pertenecientes al Grupo Phureja, las

cuales se destacan por su alta adaptabilidad a diferentes condiciones ambientales, siendo cultivadas en rango altitudinal entre 2500-3600 msnm (UPRA, 2016). Los genotipos comerciales del Grupo Phureja se destacan por el rendimiento de tubérculo y resistencia al tizón tardío de la papa (Aldabe & Dogliotti, 2009; Ibañez *et al.* 2017). Para su uso en procesamiento industrial, las características de calidad deseadas son tubérculos con alto contenido de materia seca (MS), gravedad específica (GE) mayor a 1.080 y bajos niveles de azúcares reductores (AR) que determinan un buen color de fritura (sin pardeamiento) (Estrada, 1996; Tinjacá & Rodríguez, 2015).

Este estudio evaluó sobre diez cultivares del Grupo Phureja en tres localidades de la región del Altiplano Cundiboyacense y ocho del Nudo de los Pastos en los años 2012 – 2013, diferentes modelos de regresión lineal, los cuales utilizaron como parámetros las covariables con mayor capacidad predictiva de cada componente ambiental (hídrico, energético, propiedades físico-químicas de suelos), para seleccionar el modelo que mejor explicase la descomposición de la varianza ambiental en los rasgos de rendimiento y calidad de tubérculos.

Referencias

Aldabe, L. & Dogliotti, S. (2009). Bases fisiológicas del crecimiento y desarrollo del cultivo de papa (*Solanum tuberosum* L). Curso de Fisiología de Cultivos. Ciclo de Formación Central Agronómica. Universidad de la República. Uruguay. 16 p.

Camargo, I., Intire, Q. & Gordón, R. (2011). Identifying megaenvironments to enhance the use of superior rice genotypes in Panama. *Pesquisa Agropecuária Brasileira*, 46(9), 1061-1069. <http://www.scielo.br/pab>

Crossa, J. & Cornelius, P. (1997). Sites regression and shifted multiplicative model clustering of cultivar trials site under heterogeneity of error variance. *Crop sci.* 37: 405 – 415. <https://doi.org/10.2135/cropsci1997.0011183X003700020017x>

Crossa, J., Perez, S., Jarquin, D., Cotes, J., Viele, K., Liu, G. & Cornelius, P. (2011). Bayesian estimation of the additive main effects and multiplicative interaction model. *Crop Science*, 51(4), 1458-1469. <http://dx.doi.org/10.2135/cropsci2010.06.0343>

De Silva, C., Oliveira, L., Nuvunga, J., Pamplona, A., & Balestre, M. (2015). A Bayesian shrinkage approach for AMMI models. *Plos one*, 10(7), e0131414. <https://doi.org/10.1371/journal.pone.0131414>

Dias, K., Dos Santos, J., Krause, M., Piepho, H., Guimarães, L., Pastina, M. & Garcia, A. (2022). Leveraging probability concepts for cultivar recommendation in multi-environment trials. *Theoretical and Applied Genetics*, 1-15. <https://doi.org/10.1007/s00122-022-04041-y>

Eberhart, S. & Russell, W. (1966). Stability parameters for comparing varieties. *Crop Science*, 6: 36-40. <https://doi.org/10.2135/cropsci1966.0011183X000600010011x>

Estrada, N. (1996). Los recursos genéticos en el mejoramiento de la papa en los países andinos. In: UNIPAPA-ICA-CORPOICA, *Papas colombianas con el mejor entorno ambiental*. Comunicaciones y Asociados Ltda (1st ed., pp. 1-14). <http://hdl.handle.net/20.500.12324/32445>

Finlay, K. & Wilkinson, F. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal Agricultural Resource*, 14: 742-754. https://pdf.usaid.gov/pdf_docs/PNAAS139.pdf

Gauch, H. & Zobel, W. (1988) Predictive and postdictive success of statistical analyses of yield trials. *Theor Appl Genet* 76: 1-10. <https://doi.org/10.1007/BF00288824>

Ibáñez, V., Berli, F., Masuelli, R., Bottini, R. & Marfil, C. (2017). Influence of altitude and enhanced ultraviolet-B radiation on tuber production, seed viability, leaf pigments and morphology in the wild potato species *Solanum kurtzianum* Bitter & Wittm collected from an elevational gradient. *Plant Science*, 261, 60-68. Doi: 10.1016/j.plantsci.2017.04.014 <https://doi.org/10.1016/j.plantsci.2017.04.014>

Josse, J., Van Eeuwijk, F., Piepho, H. & Denis, J. (2014). Another look at Bayesian analysis of AMMI models for genotype-environment data. *Journal of Agricultural, Biological, and Environmental Statistics*, 19(2), 240-257. <http://www.jstor.org/stable/26452870>

- Lee, E., Deen, W., Hooyer, M., Chambers, A., Parkin, G., Gordon, R. & Singh, A. (2016). Involvement of year-to-year variation in thermal time, solar radiation and soil available moisture in genotype-by-environment effects in maize. *Crop Science*, 56(5), 2180-2192. <https://doi.org/10.2135/cropsci2015.04.0231>
- Lozano, A., Zamora, V., Ibarra, L., Rodríguez, S., De la Cruz, E. & de la Rosa, M. (2009). Análisis de la interacción genotipo-ambiente mediante el modelo Ammi y potencial de producción de triticales forrajeros (X Triticosecale Wittm.). *Universidad y ciencia*, 25(1), 81-92. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0186-29792009000100006&lng=es&nrm=iso
- Omer, S., Abdalla, A. W. H., Mohammed, M. H., & Singh, M. (2015). Bayesian estimation of genotype-by-environment interaction in sorghum variety trials. *Communications in Biometry and Crop Science*, 10, 82-95. https://www.researchgate.net/publication/282328349_Bayesian_estimation_of_genotype-by-environment_interaction_in_sorghum_variety_trials
- Omer, S. & Singh, M. (2017). Comparing Bayesian and Frequentist Approaches for GGE Bi-plot Analysis in Multi-Environment Trials in Sorghum. *Eur Exp Biol*, 7(6), 40. <http://dx.doi.org/10.21767/2248-9215.100040>
- Pour, A., Khalili, M., Poczai, P. & Olivoto, T. (2022). Stability Indices to Deciphering the Genotype-by-Environment Interaction (GEI) Effect: An Applicable Review for Use in Plant Breeding Programs. *Plants*, 11(3), 414. <https://doi.org/10.3390/plants11030414>
- Shukla, G. (1972). Some statistical aspects of partitioning genotype- environment components of variability. *Heredity* 29:237-245. <https://doi.org/10.1038/hdy.1972.87>
- Theobald, C. M., Talbot, M. & Nabugoomu, F. (2002). A Bayesian approach to regional and local-area prediction from crop variety trials. *Journal of Agricultural, Biological, and Environmental Statistics*, 7(3), 403-419. <https://doi.org/10.1198/108571102230>

Tinjacá, S. & Rodríguez, L. (2015). Catálogo de papas nativas de Nariño Colombia (1era ed.). Universidad Nacional de Colombia.

http://www.seguridadalimentarianarino.unal.edu.co/sites/default/files/pdf-componentes/C.Mej-Gen_Catalogo%20papas%20Nativas.pdf

Tolhurst, D., Gaynor, R., Gardunia, B., Hickey, J. & Gorjanc, G. (2021). Genomic Selection Using Environmental Covariates Within an Integrated Factor Analytic Linear Mixed Model. <http://dx.doi.org/10.21203/rs.3.rs-995887/v1>

Unidad de Planeación Rural Agropecuaria - UPRA. (2016). Cultivo Comercial de Papa - identificación de Zonas Aptas en Colombia, a escala 1:100.000. Imprenta Nacional. <http://hdl.handle.net/11438/8630>

Viele, K. & Srinivasan, C. (2000) Parsimonious estimation of multiplicative interaction in analysis of variance using Kullback-Leibler information. *J. Stat. Plan. Infer.* 84:201– 219. [https://doi.org/10.1016/S0378-3758\(99\)00151-2](https://doi.org/10.1016/S0378-3758(99)00151-2)

Yan, W., Kang, M., Ma, B., Woods, S. & Cornelius, P. (2007) GGE Biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci* 47:643-655. <https://doi.org/10.2135/cropsci2006.06.0374>.

Objetivos

Objetivo General

Evaluar la respuesta de genotipos al ambiente en el rendimiento y calidad de cultivares de *Solanum tuberosum* Grupo Phureja en localidades de las subregiones Nudo de los Pastos y Altiplano Cundiboyacense.

Objetivos Específicos

- Caracterizar covariables ambientales informativas mediante la evaluación de cultivares en pruebas multiambiente.
- Determinar covariables que descompongan la varianza ambiental con la mayor capacidad predictiva en los componentes hídrico, energético y físico-químico de suelos.
- Seleccionar localidades superiores para los rasgos fenotípicos de rendimiento y calidad.

Capítulo 1. Marco teórico

1.1 Respuesta Fenotípica

Los procesos de mejoramiento de las plantas domesticadas han sido relevantes para la mejora y aseguramiento de la alimentación, estos son realizados de forma empírica, por campesinos y comunidades raizales, o por programas de mejoramiento técnico realizados por empresas agrícolas y centros de investigación regionales (Frutos, 2011; Dias *et al.* 2022). Uno de estos procesos de mejoramiento de cultivos, son las pruebas multiambientes (PM), las cuales son la fase final de los procesos de selección que se realizan para la recomendación y registro de cultivares comerciales (Bandera & Perez, 2018). Las PM se realizan para evaluar la respuesta fenotípica de los cultivares en diferentes localidades donde sea potencial su producción, de esta forma, se puede determinar el grado de variación de la respuesta fenotípica debida al ambiente, denominada generalmente por los fitomejoradores como Interacción Genotipo-Ambiente (IGA) (Finlay & Wilkinson, 1963; Delacy *et al.* 1996; Rueda & Cotes, 2009; Frutos, 2011). Los rasgos usualmente están controlados por múltiples genes con varios tipos de efectos genéticos que son afectados por el ambiente. El uso del operador multiplicativo para modelar la interacción genotipo x ambiente ha sido propuesto por varios investigadores (Piepho & Möhring, 2005; Crossa *et al.* 2006).

La IGA es un método teórico que permite descomponer el valor de la respuesta fenotípica en las diferentes fuentes de variación a partir de modelos estadísticos, dichos modelos buscan explicar la mayor varianza con el objetivo de desarrollar algoritmos predictivos de alta precisión (potencia) y parsimoniosos, que permitan a su vez, la selección y recomendación de cultivares superiores para ambientes específicos (Crossa *et al.* 2006; Bandera & Perez, 2018; Dias *et al.* 2022). Se han desarrollado múltiples métodos para el análisis de bases de datos colectadas en pruebas multiambientes, las cuales se enfocan principalmente en el rendimiento del órgano de interés (tubérculos, semillas, flores, etc.) como variable respuesta. Se detecta IGA mediante el

establecimiento de PM con al menos dos genotipos evaluados en dos localidades diferentes, descomponiendo la respuesta fenotípica de la siguiente ecuación:

$$y_{ij} = \mu + G_i + (GA)_{ij} + e_{ij} \quad [Ec.1]$$

Donde y_{ij} hace referencia a la respuesta fenotípica del i -ésimo genotipo evaluado en el j -ésimo ambiente, μ hace referencia a la gran media, G_i hace referencia a los efectos principales del i -ésimo genotipo, $(GA)_{ij}$ hace referencia a la interacción del genotipo con el ambiente, la cual es constante cuando el rendimiento del cultivar no depende del ambiente, o con interacción cuando el rendimiento presenta cambio de magnitud (interacción cruzada [COI]) o heterocedasticidad (HV); e_{ij} hace referencia al error residual (Crossa *et al.* 2004; Burgueño *et al.* 2008). Cuando la interacción no es significativa, se suele relacionar con estabilidad en sentido estático, indicando que el rendimiento no variará a pesar de las condiciones ambientales; en contraste, cuando se encuentra interacción con diferencias significativas, se habla de estabilidad en sentido dinámico (agronómica, Piepho, 1996) y por la cual se debe realizar la selección de cultivares a localidades específicas con características ambientales similares (macroambientes) (Oliveira *et al.* 2020; Tolhurst *et al.* 2021).

1.2 Análisis de estabilidad y adaptabilidad fenotípica

Actualmente existen muchas metodologías de análisis de estabilidad fenotípica. El método tradicional consiste en el análisis conjunto de los experimentos, considerando todos los ambientes y el posterior desdoblamiento de la suma de cuadrados de efectos de los ambientes y de la interacción genotipo–ambiente. Algunos métodos planteados para el análisis de conjuntos de datos recolectados en PM son los siguientes.

1.3 Métodos lineales

Métodos estadísticos de regresión lineal para el análisis de PM han sido propuestos por varios investigadores en diferentes programas de mejoramiento, un ejemplo de estos modelos, es el método planteado por Finlay & Wilkinson (1963) quienes propusieron realizar un análisis de regresión del rendimiento en 277 cultivares sobre el índice ambiental de cada localidad, con el fin de estimar la estabilidad de los genotipos. El

índice ambiental es una estimación del potencial de rendimiento de cada localidad y es calculado como la diferencia entre el rendimiento promedio de los cultivares en cada localidad y el promedio de los genotipos en todas las localidades evaluadas. Eberhart & Russell (1966) realizaron un trabajo paralelo, llegando casi al mismo modelo, pero incorporando las desviaciones respecto a la línea de regresión como un criterio de estabilidad, planteando el siguiente modelo:

$$y_{ij} = \mu_i + \beta_i I_j + \delta_{ij} \quad [Ec. 2]$$

Con $i = 1, \dots, v$ (genotipos); $j = 1, \dots, n$ (ambientes); μ_i es la media de cada cultivar y β_i son los coeficientes de regresión que miden la respuesta de cada genotipo al cambio en índice ambiental; Los coeficientes de regresión β_i son conocidos como el parámetro de adaptabilidad, mientras que la suma de las desviaciones δ_{ij} es la desviación de la regresión de la variedad i en el ambiente j . I_j es el índice ambiental obtenido como el promedio de todas las variedades en el j -ésimo ambiente menos la gran media (Eberhart & Russell, 1966)

$$I_j = \left(\sum_i \frac{y_{ij}}{v} \right) - \left(\sum_i \sum_j \frac{y_{ij}}{vn} \right) \quad [Ec. 3] \quad \sum_j I_j = 0$$

El primer parámetro de estabilidad es un coeficiente de regresión estimado de la siguiente forma:

$$\beta_i = \frac{\sum_j y_{ij} I_j}{\sum_j I_j^2} \quad [Ec. 4]$$

El rendimiento de cada variedad puede ser predicho utilizando los estimadores de los parámetros donde $\hat{y}_{ij} = \bar{x}_i + \beta_i I_j$, donde \bar{x}_i es un estimador de la μ_i . Mientras que las desviaciones ($\hat{\delta}_{ij} = y_{ij} - \hat{y}_{ij}$) pueden ser potenciadas al cuadrado y sumadas para proveer un estimado de otro parámetro de estabilidad (σ_{di}^2).

$$S_{d_i}^2 = \frac{\sum_j \hat{\delta}_{ij}^2}{(n-2)} - \frac{S_e^2}{r} \quad [Ec. 5]$$

Donde S_e^2/r , es el error residual del análisis de varianza (ANOVA) combinado (incluyendo todas las pruebas regionales) y:

$$\sum_j \hat{\delta}_{ij}^2 = \left[\sum_j y_{ij}^2 - y_i^2 \right] - \frac{(\sum_j y_{ij} l_j)^2}{\sum_j l_j^2} \quad [Ec. 6]$$

De esta forma la partición de la IGA de cada genotipo se da en dos partes, la variación de la respuesta fenotípica debido a variaciones de índice ambiental (SC regresión) y las desviaciones inexplicables de la regresión sobre el índice ambiental. Eberhart y Russell consideran estables los genotipos que presenten coeficientes de regresión cercanos a uno ($\beta = 1$), con las menores desviaciones ($S_{d_i}^2 = 0$). Además, permite la aplicación de la prueba F de significancia en la comparación de medias entre genotipos ($H_0: \mu_1 = \mu_2 = \mu_3 == \mu_v$) y la determinación de diferencias genéticas debidas al índice ambiental para cada genotipo, utilizando el coeficiente de regresión ($H_0: \beta_1 = \beta_2 = \beta_3 == \beta_v$).

1.4 Métodos bilineales

Otros métodos lineales desarrollados para el análisis de PM, son los modelos de efecto principal aditivo e interacción multiplicativa (AMMI) (Gauch, 1982; Gauch & Zobel 1988), utilizados para estudiar patrones de respuesta genotípica en distintos entornos y puede ser definido de la siguiente ecuación:

$$Y_{ij} = \mu + \tau_i + \theta_j + \sum_{k=1}^t \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma_A^2) \quad [Ec. 7]$$

Donde Y_{ij} representa la respuesta fenotípica (como el rendimiento) para los genotipos $i = 1, \dots, I$ y ambientes $j = 1, \dots, J$. La parte lineal del modelo es representado por la gran media μ , los parámetros para los efectos principales son denotados por τ_i y θ_j ; la

interacción multiplicativa (parte bilineal) es representada por el término $\sum_{k=1}^t \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij}$ donde k representa el k -ésimo vector propio y ε el error residual. Dichos modelos están sobre parametrizados y requieren agregar algunas restricciones para obtener estimaciones únicas para los parámetros. En la práctica se utilizan las siguientes restricciones (Gauch & Zobel 1988; Yan *et al.* 2007; Carmargo *et al.* 2011; Josse *et al.* 2014):

$$\begin{aligned} \mathbf{1}'_I \boldsymbol{\tau} = \mathbf{1}'_J \boldsymbol{\delta} = 0, \quad \mathbf{1}'_I \boldsymbol{\alpha}_k = \mathbf{1}'_J \boldsymbol{\gamma}_k = 0, \\ \boldsymbol{\alpha}' \boldsymbol{\alpha} = \boldsymbol{\gamma}' \boldsymbol{\gamma} = I_K, \quad (5) \\ \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K \geq 0, \\ \alpha_{1k} \geq 0 \quad \text{para } k = 1, \dots, t, \end{aligned}$$

Donde $\mathbf{1}'_I$ y $\mathbf{1}'_J$ son vectores de tamaño I y J que contienen unos; $\boldsymbol{\alpha}_k$ y $\boldsymbol{\gamma}_k$ son la k -ésima columna de la matriz $(\boldsymbol{\alpha}_{I \times t}, \boldsymbol{\gamma}_{J \times t})$ y I_K es la matriz identidad de tamaño K . Las restricciones $\lambda_K \geq 0$ y $\alpha_{1k} \geq 0$ permite evadir las indeterminaciones de signos en los términos de interacción (Camargo *et al.* 2011; Josse *et al.* 2014). Con las restricciones elegidas en la ecuación 5 y para un conjunto de datos completo, las estimaciones de mínimos cuadrados para la parte aditiva son las mismas que en el marco de análisis de varianza habitual, μ se estima por la gran media de los datos, $\boldsymbol{\tau}$ es el efecto principal debido al genotipo, siendo estimado como el vector de las medias de las columnas menos la gran media, $\boldsymbol{\delta}$ como el efecto principal de los ambientes y se estima como el vector de las medias de las filas menos la gran media (Gauch & Zobel, 1988). Las estimaciones de mínimos cuadrados para los términos de interacción vienen dadas por la descomposición de valores singulares (SVD) de la matriz de residuos de interacción (la matriz de datos centrada en filas y columnas), da los K valores singulares (λ_k) ($k = 1, \dots, t$) y los vectores singulares izquierdo y derecho en las matrices $\boldsymbol{\alpha}$ y $\boldsymbol{\gamma}$. Se considera conocido el número de dimensiones K subyacentes (Camargo *et al.* 2011; Josse *et al.* 2014). Este modelo es parsimonioso, ya que los parámetros de interacción se estiman a partir de la descomposición de valor singular de la matriz.

El método AMMI puede ser utilizado para predecir el valor de un genotipo presente en un ambiente específico, es decir, se aplica en el contexto de modelos con parámetros de

efectos fijos y, por lo tanto, requiere datos balanceados y varianza homogénea que no siempre se puede garantizar para datos de PM (Crossa *et al.* 2011). Estas limitaciones motivaron la realización de análisis PM en el contexto de modelos mixtos, que teóricamente tienen mayor flexibilidad para el tratamiento de datos faltantes, comúnmente observados en PM (Paderewski *et al.* 2016). En este sentido, podemos destacar las versiones de modelos bilineales mixtos AMMI propuestos inicialmente por Piepho (1997, 1998), Smith *et al.* (2001), Piepho y Mohring (2006). La aplicación de modelos lineales y bilineales de efectos fijos a datos no balanceados puede ser complicada, requiriendo software especializado para realizar los procedimientos de estimación de mínimos cuadrados alternos recomendados (Eeuwijk, 1995; Burgueño *et al.* 2011). Sin embargo, un método aproximativo simple puede consistir en ajustar primero un modelo mixto a una tabla incompleta de genotipo por ambiente, luego calcular una tabla completa de las mejores predicciones lineales insesgadas (BLUP) y finalmente aplicar análisis de modelos bilineales a esa tabla completa (Burgueño *et al.* 2011).

1.5 Métodos mixtos

Las PM comprenden factores fijos (dependientes del fitomejorador) y factores aleatorios que varían dependiendo de la naturaleza en los ambientes de selección durante la evaluación agronómica (Edward *et al.* 2012; Bandera & Perez, 2018). Estas limitantes disminuyen la bondad de ajuste de los modelos fijos, afectando la capacidad predictiva y disminuyendo la explicación de la varianza (Wang *et al.* 2015). Esto ha contribuido al desarrollo de nuevos modelos como los modelos mixtos, donde se contemple el factor estocástico relacionado a cada localidad y el periodo de ejecución de la PM. Entre los modelos desarrollados, se encuentran los modelos lineales generalizados mixtos (MLGM), los cuales aportan variaciones al modelo lineal, permitiéndole al mismo ser aplicado a un rango más amplio de datos (Bandera & Perez, 2018). Los datos de las PM usualmente no satisfacen las premisas de los modelos lineales generales (MLG), como el supuesto de normalidad de las variables, por lo que los modelos lineales generalizados mixtos proporcionan una vía de análisis, posibilitando que las mismas se puedan ajustar a una distribución de la familia exponencial (Bertan *et al.* 2007).

Los MLGM son utilizados cuando hay presencia de variables con efectos estocásticos (independientes del fitomejorador) donde los valores de una variable independiente

categoría se ven como una muestra aleatoria de todos los valores (Edward *et al.* 2012; Bandera & Perez, 2018); también son utilizados cuando hay presencia de variables jerárquicas, donde se miden más de un nivel o en conjuntos de datos de medidas repetidas, donde hay presencia de correlación independiente entre las observaciones. Los MLGM pueden ser descritos de la siguiente ecuación:

$$y = Xb + Zu + e \quad [Ec. 8]$$

Donde y es el vector con la información de la respuesta fenotípica, X y Z son matrices de diseño conocidas con m columnas (variables) y n filas (observaciones), b es un vector de parámetros fijos (genotipos), u es un vector de efectos aleatorios (covariables ambientales) y e es la matriz que contiene los vectores aleatorios no observables que describen el error residual. Las esperanzas matemáticas de u y e , son asumidas como cero (Laird & Ware, 1982; Bandera & Perez, 2018). Los MLGM permiten calcular el mejor estimador lineal insesgado (BLUE) de los genotipos y el mejor predictor lineal insesgado (BLUP) de los ambientes. El BLUP representa la esperanza condicional de los efectos estocásticos debido al vector de respuestas fenotípicas, y es considerado un estimador Bayesiano. El BLUP de una combinación lineal de efectos fijos y aleatorios, es la combinación lineal de los BLUE de efectos fijos y los BLUP de efectos aleatorios (Piepho *et al.* 2008).

1.6 Caracterización del ambiente en covariables

Si las covariables ambientales (precipitación, radiación, y tipo de suelo) están disponibles para las localidades o regiones específicas de las PM, pueden ser incorporadas a un modelo de regresión lineal para predecir el rendimiento en ubicaciones específicas (Piepho *et al.* 1998), inicialmente fue propuesto este enfoque dentro del marco de los modelos de efectos fijos, donde se consideraba el efecto de los genotipos y las localidades como fijos (Smith *et al.* 2005). Posteriormente, Piepho *et al.* (1998) planteó el uso de datos ambientales como covariables en modelos de efectos mixtos que permitan predecir el rendimiento en localidades en donde no se han realizado las PM, esto debido a la imposibilidad de controlar las variables ambientales y la necesidad de predecir con precisión la respuesta fenotípica de los rasgos de interés en localidades no evaluadas, a

partir de la información recolectada en el resto de localidades. Heslot *et al.* (2014) extendieron este enfoque a procesos de selección genómica, considerando regresiones aleatorias sobre covariables del ambiente, derivadas de observaciones meteorológicas diarias. Por su parte, Jarquín *et al.* (2014) adaptó un modelo de norma de reacción para un conjunto muy grande de covariables correlacionadas. Descubrieron que la interacción de la respuesta fenotípica de los genotipos con las covariables ambientales explicaban el 23% de la variación genética entre entornos, aumentando la capacidad predictiva del modelo (Tolhurst *et al.* 2021).

1.7 Interacción genotipo – ambiente en papa amarilla diploide

La papa es reproducida asexualmente, muchos individuos con el mismo genotipo (clones) pueden ser producidos y plantados en diferentes condiciones ambientales, observándose efectos marcados del ambiente sobre la expresión fenotípica de estos (Bradshaw, 2021); debido a su plasticidad fenotípica se adapta con facilidad a diversos climas y sistemas de cultivo, motivo por el cual la producción mundial de papa va en continuo aumento, especialmente en los países más poblados (Núñez, 2006). En Colombia su adaptabilidad va desde los altiplanos de 2.000 a 4.000 msnm hasta las zonas bajas o valles de la zona tórrida de 1.000 a 2.000 msnm (Estrada, 2000; Rodríguez & Núñez, 2011; Núñez, 2018). Es una especie que posee gran diversidad de genotipos, y por ende presenta variantes en sus componentes fisiológicos y fenológicos, los cuales determinan las variables de eficiencia biológica y de productividad, así como las prácticas de manejo agronómico apropiadas para mejorar el desempeño del cultivo (Aldabe & Dogliotti, 2009).

El crecimiento depende de múltiples factores ambientales, dentro de los que se destacan la precipitación (agua disponible en el suelo), temperatura (grados día) y la intensidad de radiación solar (radiación solar absorbida por el cultivo); el primero modula la corriente de transpiración y el movimiento de nutrientes, el segundo la velocidad de las reacciones enzimáticas y la tercera está relacionada directamente con la actividad fotosintética de las plantas (Cabezas, 2002). Estos factores ambientales varían con la altitud y el cultivo de papa se encuentra diseminado en regiones que van desde los 2.200 hasta los 3.500 msnm, con un rango de temperatura desde 10°C hasta 15°C, es de esperarse que las

plantas de papa presenten diferencias en los procesos fisiológicos de crecimiento y desarrollo (Ñustez, 2006; Ñustez, 2018). En condiciones de déficit hídrico, el cultivo de papa tiende a reducir el crecimiento del follaje, impactado a su vez, en la producción de fotoasimilados y reduciendo el periodo vegetativo, lo cual se ve reflejado en una menor producción de tubérculos (Taiz & Zeiger, 2006; Marschner, 2011; Muthoni & Kabira, 2016); por otra parte, las altas temperaturas afectan la tuberización, inhibiéndola significativamente y la partición de fotoasimilados en los tubérculos se reduce considerablemente (Rykaczewska, 2015), afectando drásticamente en el rendimiento y calidad interna de tubérculo.

1.8 Referencias

Bandera, E. & Pérez, L. (2018). Generalized Linear Mixed Models. Its application in plant breeding. *Cultivos Tropicales*, 39(1), 127-133.
https://www.researchgate.net/publication/327751933_Review_Generalized_linear_mixed_models_Its_application_in_plant_breeding

Bertan, I., Carvalho, F. & Oliveira, A. (2007). Parental selection strategies in plant breeding programs. *Journal of crop science and biotechnology*, 10(4), 211-222.
https://www.researchgate.net/publication/254000574_Parental_Selection_Strategies_in_Plant_Breeding_Programs

Bradshaw, J. (2021). Potato breeding: theory and practice. Springer International Publishing. ISBN: 978-3-030-64414-7

Burgueño, J., Crossa, J., Cornelius, P. & Yang, R. (2008). Using factor analytic models for joining environments and genotypes without crossover genotype x environment interaction. *Crop Sci.* 48:1291-1305. <https://doi.org/10.2135/cropsci2007.11.0632>

Burgueño, J., Crossa, J., Cotes, J., Vicente, F. & Das, B. (2011). Prediction assessment of linear mixed models for multienvironment trials. *Crop Science*, 51(3), 944-954.
<https://doi.org/10.2135/cropsci2010.07.0403>

Cabezas, M., (2002). Estimación de la interceptación de la radiación solar en papa criolla (*Solanum phureja*, Juz, et Buk.) en tres localidades colombianas. Trabajo de Grado Maestría. Facultad de Agronomía. Universidad Nacional de Colombia. 122 p. https://www.researchgate.net/publication/262446787_Estimating_yellow_potato_Solanum_phureja_Juz_et_Buk_solar_radiation_interception_in_three_Colombian_localities

Camargo, I., Intire, Q. & Gordón, R. (2011). Identifying megaenvironments to enhance the use of superior rice genotypes in Panama. *Pesquisa Agropecuária Brasileira*, 46(9), 1061-1069. https://www.researchgate.net/publication/289230956_Identifying_mega-environments_to_enhance_the_use_of_superior_rice_genotypes_in_Panama

Crossa, J., Perez, S., Jarquin, D., Cotes, J., Viele, K., Liu, G. & Cornelius, P. (2011). Bayesian estimation of the additive main effects and multiplicative interaction model. *Crop Science*, 51(4), 1458-1469. <http://dx.doi.org/10.2135/cropsci2010.06.0343>

Crossa, J., Yang, R. & Cornelius, P. (2004). Studying crossover genotype x environment interaction using linear-bilinear models and mixed models. *J. Agric. Biol. Environ. Stat.* 9:362-380. <https://www.jstor.org/stable/1400487>

Delacy, I., Basford, K., Cooper, M. & Bull, J. (1996) Analysis of multienvironment trials an historical perspective. *Plant Adaptation and Crop Improvement*. Eds. M. Cooper and G. L. Hammer. CAB international, UK. <https://www.cabdirect.org/cabdirect/abstract/19961610824>

Días, K., Dos Santos, J., Krause, M., Piepho, H., Guimarães, L., Pastina, M. & Garcia, A. (2022). Leveraging probability concepts for cultivar recommendation in multi-environment trials. *Theoretical and Applied Genetics*, 1-15. <https://doi.org/10.1007/s00122-022-04041-y>

Eberhart, S. & Russell, W. (1966). Stability parameters for comparing varieties. *Crop Science*, 6: 36-40. <https://doi.org/10.2135/cropsci1966.0011183X000600010011x>

Edward, E., Gbur, E., Walter, W., Stroup, W., McCarter, K. & Durham, S. (2012). Analysis of generalized linear mixed models in the agricultural and natural resources sciences. Madison, Wis: American Society of Agronomy; 277 p. ISBN: 978-0-891-18182-8

Estrada, N. (2000). La biodiversidad en el mejoramiento genético de la papa. Fundación Promoción e Investigación de los Productos Andinos. PROINPA: Centro Internacional de la Papa. CIP: Centro de Información para el Desarrollo. CID. La Paz. Bolivia. 372. <http://hdl.handle.net/20.500.12324/32445>

Finlay, K. & Wilkinson, F. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal Agricultural Resource*, 14: 742-754. https://pdf.usaid.gov/pdf_docs/PNAAS139.pdf

Frutos, E. (2011). Interacción genotipo-ambiente: GGE Biplot y modelos AMMI. Departamento de Estadística. Universidad de Salamanca. (Disertación doctoral). <http://hdl.handle.net/10366/121900>

Gauch, H. (1982). Noise reduction by eigenvector ordinations. *Ecology*, 63(6), 1643-1649. <https://doi.org/10.2307/1940105>

Gauch, H. & Zobel, W. (1988) Predictive and postdictive success of statistical analyses of yield trials. *Theor Appl Genet* 76: 1-10. <https://doi.org/10.1007/BF00288824>

Heslot, N., Akdemir, D., Sorrells, M. & Jannink, J. (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theoretical and applied genetics*, 127(2), 463-480. <https://doi.org/10.1007/s00122-013-2231-5>

Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J. & de los Campos, G. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and applied genetics*, 127(3), 595-607. <https://doi.org/10.1007/s00122-013-2243-1>

Josse, J., van Eeuwijk, F., Piepho, H. & Denis, J. (2014). Another look at Bayesian analysis of AMMI models for genotype-environment data. *Journal of Agricultural, Biological, and Environmental Statistics*, 19(2), 240-257. <https://www.jstor.org/stable/26452870>

Laird, N. & Ware, J. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4):963-74. doi:10.2307/2529876. <https://doi.org/10.2307/2529876>

Marschner, H. 2011. *Mineral Nutrition of Higher Plants*, 3rd ed. New York ISBN: 9780080571874

Muthoni J. & Kabira J. (2016). Potato production under drought conditions: Identification of adaptive traits. *International Journal of Horticulture*, 6(12): 1-10. <http://dx.doi.org/10.5376/ijh.2016.06.0012>

Ñústez, C. (2011). Estudios fenotípicos y genéticos asociados a la calidad de fritura en *Solanum phureja* Juz et Buk. Tesis de doctorado en Ciencias Agropecuarias. Facultad de Agronomía. Universidad Nacional de Colombia. 204 p. <https://repositorio.unal.edu.co/handle/unal/7859>

Ñústez, C. (2018). Papas diploides: Un legado ancestral para la agricultura en Colombia. Universidad Nacional de Colombia. <http://www.papaunc.com/blog/papas-diploides-un-legado-ancestral-para-la-agricultura-en-colombia>

Oliveira, I., Guilhen, J., Oliveira, P., Gezan, S., Schaffert, R., Simeone, M. & Pastina, M. (2020). Genotype-by-environment interaction and yield stability analysis of biomass sorghum hybrids using factor analytic models and environmental covariates. *Field Crops Research*, 257, 107929. <https://doi.org/10.1016/j.fcr.2020.107929>

Paderewski, J., Gauch, H., Mađry, W. & Gacek, E. (2016). AMMI analysis of four-way genotypex locationx managementx year data from a wheat trial in Poland. *Crop Science*, 56(5), 2157-2164. <https://doi.org/10.2135/cropsci2015.03.0152>

Piepho, H. 1997. Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics*. 53:761—767. <https://doi.org/10.2307/2533976>

Piepho, H. 1998. Empirical best linear unbiased prediction in cultivar trials using factor analytics variance-covariance structures. *Theor. Appl. Genet.* 97:195-201. <https://doi.org/10.1007/s001220050885>

Piepho, H. & Möhring, J. (2006). Selection in cultivar trials- Is it ignorable? *Crop Sci.* 46: 192-201. <https://doi.org/10.2135/cropsci2005.04-0038>

Piepho, H., Möhring, J., Melchinger, A. & Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica*, 161(1), 209-228. <https://doi.org/10.1007/s10681-007-9449-8>

Rodríguez, L., Ñustez, C. & Estrada, N., (2009). Criolla Latina. Criolla Paisa y Criolla Colombia. nuevos cultivares de papa criolla para el departamento de Antioquia (Colombia). *Agron. Colomb.* 27(3). 289-303. <https://www.redalyc.org/articulo.oa?id=180316242002>

Rueda, J. & Cotes, J. (2009). Evaluación de dos métodos de estabilidad fenotípica a través de validación cruzada. *Revista Facultad Nacional de Agronomía-Medellín*, 62(2), 5111-5123. <https://repositorio.unal.edu.co/handle/unal/37088>

Rykaczewska, K. (2015). The effect of high temperature occurring in subsequent stages of plant development on potato yield and tuber physiological defects. *American Journal of Potato Research*, 92(3), 339-349. <https://doi.org/10.1007/s12230-015-9436-x>

Smith, A., Cullis, B. & Thompson, R. (2001). Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics*. 57: 1138-1147. <https://doi.org/10.1111/j.0006-341X.2001.01138.x>

Smith, A., Cullis, B. & Thompson, R. (2005). The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. *J. Agric. Sci.* 143:1-14. <https://doi.org/10.1017/S0021859605005587>

Taiz, L. & Zeiger, E. (2006). *Plant Physiology*, 4th ed. ISBN: 9780878935079

Tolhurst, D., Gaynor, R., Gardunia, B., Hickey, J. & Gorjanc, G. (2021). Genomic Selection Using Environmental Covariates Within an Integrated Factor Analytic Linear Mixed Model. <https://doi.org/10.21203/rs.3.rs-995887/v1>

Van Eeuwijk, F. (1995). Linear and bilinear models for the analysis of multi-environment trials: I. An inventory of models. *Euphytica*, 84(1), 1-7. <https://doi.org/10.1007/BF01677551>

Wang, T., He, P., Ahn, K., Wang, X., Ghosh, S. & Laud, P. (2015) A re-formulation of generalized linear mixed models to fit family data in genetic association studies. *Frontiers in Genetics* [Internet]. <https://doi.org/10.3389/fgene.2015.00120>

Yan, W., Kang, M., Ma, B., Woods, S. & Cornelius, P. (2007) GGE Biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci* 47:643-655. <https://doi.org/10.2135/cropsci2006.06.0374>

Capítulo 2. Conocimiento del microambiente y su efecto en la respuesta fenotípica de papa amarilla diploide (*Solanum tuberosum*, Grupo Phureja)

Resumen

La caracterización ambiental de las localidades se logra directamente mediante la medición de variables ambientales e indirectamente mediante la evaluación de pruebas multiambientes (PM), así, se puede determinar el grado de variación de la respuesta fenotípica debida al ambiente, denominada como Interacción Genotipo-Ambiente (IGA). La IGA permite descomponer el valor de la respuesta fenotípica en las diferentes fuentes de variación a partir de modelos estadísticos, actualmente la mayoría de los modelos se enfocan en determinar la varianza asociada a los efectos principales del genotipo y la IGA, mientras la estimación de la varianza ambiental se obtiene a partir del rendimiento o rasgo de interés, disminuyendo la capacidad predictiva de los modelos. El objetivo de este estudio fue evaluar diferentes covariables ambientales dentro de modelos de regresión, seleccionando el modelo más parsimonioso para describir la varianza fenotípica debida al ambiente. Se utilizó el conjunto de datos de las PM de 10 cultivares de papa amarilla diploide, en dos semestres consecutivos durante los años 2012-2013, los rasgos fenotípicos evaluados fueron rendimiento, azúcares reductores, calidad de fritura, materia seca y gravedad específica. Se estimó la Sensibilidad Fenotípica de los rasgos al ambiente mediante covariables ambientales del componente fisicoquímico del suelo, hídrico y energético. Se relacionó la respuesta fenotípica de los rasgos de interés, con las covariables explicativas de cada componente ambiental, utilizando la asociación lineal en los componentes hídrico y energético y ACP para el componente fisicoquímico del suelo, finalmente se realizó el ajuste de modelos acumulativos para predecir la varianza fenotípica asociada al ambiente. Las covariables más informativas fueron

CAC, PP, PARAi, RADA, altitud, pH y CICE. De las tres combinaciones distintas como parámetros de entrada en los modelos de regresión múltiple, se determinó que el mejor modelo para la predicción de la respuesta fenotípica de los rasgos evaluados fue el modelo que incluyó los siguientes parámetros PP, altitud, pH y CICE.

Abstract

The environmental characterization of locations is achieved directly by measuring environmental variables and indirectly by evaluating multi-environment trials (MET), thus, the degree of variation of the phenotypic response due to the environment can be determined and is called Genotype by Environment Interaction (GEI). The GEI allows to decompose the value of the phenotypic response in different sources of variation by statistical models, currently most models focus on determining the variance associated with the main effects of genotype and GEI, while the estimation of the environmental variance is obtained from the yield or trait of interest, decreasing the predictive capacity of the models. The aim of this study was to evaluate different environmental covariates within regression models, selecting the most parsimonious model to describe phenotypic variance due to environment. The dataset was composed of 10 diploid yellow potato cultivars, evaluated in two consecutive semesters during the years 2012-2013, the phenotypic traits evaluated were yield, sugar reducing, fry quality, dry matter, and specific gravity. Phenotypic sensitivity of the traits to the environment was estimated using environmental covariates of soil physicochemical, water and energy components. The phenotypic response of the traits of interest was related to the explanatory covariates of each environmental component, using linear association in the water and energy components and PCA for the soil physicochemical component. Finally, cumulative models were fitted to predict the phenotypic variance associated with the environment. The most informative covariates were CAC, PP, PARAi, RADA, altitude, pH and CICE. From three different combinations of the parameters in the multiple regression models, it was determined that the best model for the prediction of the phenotypic response of the evaluated traits was the model that included the following parameters PP, altitude, pH and CICE.

2.1 Introducción

La caracterización ambiental de las localidades objetivo se logra de forma directa mediante la medición de variables ambientales, que pueden ser de origen físico, biológico o químico (Rodríguez *et al.*, 2005; Tolessa & Gela, 2014; Bornhofen *et al.*, 2018); y de forma indirecta mediante la evaluación de pruebas multiambientes (PM), las cuales caracterizan los ambientes favorables mediante el uso del rendimiento y constituyen la fase final de los procesos de selección que se realizan para evaluar la respuesta fenotípica de los cultivares en diferentes localidades potenciales para su producción, de esta forma, se puede determinar el grado de variación de la respuesta fenotípica debida al ambiente, denominada generalmente por los fitomejoradores como Interacción Genotipo-Ambiente (IGA) (Finlay & Wilkinson, 1963; Delacy *et al.*, 1996; Rueda & Cotes, 2009; Frutos, 2011; Oliveira *et al.*, 2020).

La IGA es un método estadístico que permite descomponer el valor de la respuesta fenotípica en las diferentes fuentes de variación a partir de modelos estadísticos, dichos modelos buscan explicar la mayor varianza con el objetivo de desarrollar algoritmos predictivos de alta precisión (potencia) y parsimoniosos, que permitan a su vez, la selección y recomendación de cultivares superiores para ambientes específicos (Crossa *et al.*, 2006; Bandera & Perez, 2018). Varios métodos han sido propuestos para el análisis de la IGA, uno de estos métodos fue planteado por Finlay & Wilkinson (1963) quienes propusieron realizar un análisis de regresión sobre el índice ambiental de cada localidad a partir del rendimiento, con el fin de estimar la estabilidad de los genotipos. Por otra parte, Eberhart & Russell (1966) realizaron un trabajo paralelo, llegando casi al mismo modelo, pero incorporando las desviaciones respecto a la línea de regresión como un criterio de estabilidad. Los modelos multivariados como los efectos principales aditivos, la interacción multiplicativa (Gauch, 1988) y los métodos biplot de genotipo más IGA (Yan *et al.*, 2007), se han utilizado ampliamente para definir las zonas de producción y evaluar la estabilidad del genotipo en todos los entornos (Crossa, 2012; Malosetti *et al.*, 2013).

El conocimiento sobre IGA proporciona información para estimar la estabilidad del genotipo en diferentes ambientes, seleccionar genotipos en localidades específicas,

definir zonas de reproducción (o mega-ambientes) e identificar localidades para optimizar la ganancia genética en los programas de mejoramiento (Dias *et al.*, 2022). La comprensión adecuada de la IGA es uno de los mayores desafíos que enfrentan los fitomejoradores, actualmente la mayoría de modelos descritos, se enfocan en determinar la varianza asociada a los efectos principales del genotipo y la IGA, mientras la estimación de la varianza ambiental se obtiene a partir del rendimiento o rasgo de interés, valor que no depende intrínsecamente del ambiente, siendo afectado por el genotipo y, en consecuencia, disminuyendo la capacidad predictiva de los modelos ajustados (Sellam & Poovammal, 2016). La descomposición de la varianza ambiental se podría realizar en las siguientes fuentes de variación: componentes hídrico, energético y bioquímico (Brancourt *et al.*, 2000), estos componentes pueden ser descritos a partir de métodos de regresión lineal simple entre las covariables ambientales y los rasgos de interés como variables dependientes (Betancur *et al.*, 2012).

El presente estudio tuvo como objetivo evaluar diferentes covariables ambientales dentro de modelos de regresión lineal, seleccionando el modelo más parsimonioso para describir la varianza ambiental en los rasgos de rendimiento y calidad de tubérculos.

2.2 Materiales y métodos

2.2.1 Base de Datos

El conjunto de datos de las pruebas multiambiente (PM) de cultivares de papa amarilla diploide (*Solanum tuberosum* grupo Phureja), ejecutadas de conformidad con la resolución No. 04000 del Instituto Colombiano Agropecuario (ICA), fue obtenido gracias al programa de Fitomejoramiento en Papa de la Universidad Nacional de Colombia. Las PM se desarrollaron en ocho localidades de la región del Nudo de los Pastos en el departamento de Nariño y tres localidades del Altiplano Cundiboyacense en el departamento de Cundinamarca (

Tabla 1), en dos semestres consecutivos (replicas) durante los años 2012-2013. Los rasgos fenotípicos evaluados fueron, rendimiento de tubérculo ($\text{Kg}\cdot\text{ha}^{-1}$), porcentaje de azúcares reductores (AR), calidad de fritura (CF), porcentaje de materia seca (MS) y gravedad específica (GE).

Tabla 1. Ubicación de las pruebas multiambiente de papa amarilla diploide en 11 localidades de los departamentos de Nariño y Cundinamarca, para los años 2012-2013.

Localidad	Año	Replic a	Geo-referenciación		Altitud (msnm)	Temp. Media (°C)
			N	W		
Carlosama-Nariño	2012	1	0°52'4.7"	77°43'30.3	3027	12
Carlosama-Nariño	2013	2	0°51'7.0"	77°44'26.7	3073	12
Cumbal-Nariño	2012	1	0°54'26.5'	77°49'30.0	3273	10
Cumbal-Nariño	2013	2	0°54'36.4'	77°47'43.5	3046	10
Ipialpud Alto-Nariño	2012	1	0°56'19.0'	77°43'50.3	3040	10
Ipialpud Alto-Nariño	2013	2	0°56'31.9'	77°44'13.7	3082	10
Jamondino-Nariño	2012	1	1°10'49.7'	77°15'46.5	2605	10
Jamondino-Nariño	2013	2	1°10'39.5'	77°15'46.5	2605	10
Mosquera- Cundinamarca	2012	1	4°41'2.3"	74°12'53.8	2539	13
Mosquera- Cundinamarca	2013	2	4°40'50.1'	74°12'48.9	2541	13
Obonuco-Nariño	2012	1	1°11'33.4'	77°18'45.0	2794	12
Obonuco-Nariño	2013	2	1°11'31.2'	77°18'54.3	2889	12
San Ramón-Nariño	2012	1	0°58'28.4'	77°42'10.4	3039	10
San Ramón-Nariño	2013	2	0°58'22.4'	77°41'19.3	3074	10
Santa Bárbara-Nariño	2012	1	1°05'46.2'	77°16'43.1	2880	11
Santa Bárbara-Nariño	2013	2	1°03'47.8'	77°15'59.7	2978	11
Subachoque- Cundinamarca	2012	1	4°54'55.1'	74°09'5.9"	2663	11
Subachoque- Cundinamarca	2013	2	4°55'14.2'	74°11'4.3"	2680	11
Túquerres-Nariño	2012	1	1°11'8.4"	77°38'39.4	2662	13
Túquerres-Nariño	2013	2	1°10'36.7'	77°38'38.8	2913	13
Zipaquirá- Cundinamarca	2012	1	5°00'25.2'	74°00'4.3"	2650	12
Zipaquirá- Cundinamarca	2013	2	5°00'25.1'	74°01'7.0"	2643	12

2.2.2 Material vegetal

Se evaluaron ocho genotipos avanzados obtenidos mediante hibridación y selección en diferentes ambientes por el programa de Fitomejoramiento en Papa de la Facultad de Agronomía de la Universidad Nacional de Colombia-Sede Bogotá (Tabla 2), a través del proyecto IV-1344, financiado por el Ministerio de Agricultura y Desarrollo Rural, como testigos se utilizaron los cultivares Criolla Colombia y Criolla Galeras.

Criolla Colombia (Col), es la principal variedad de papa amarilla diploide cultivada en Colombia, es un cultivar de porte medió, adaptado entre los 2400 y 3200 msnm, con follaje verde ligeramente claro, floración abundante, es un cultivar precoz (120 días a 2600 msnm) y susceptible a PYVV; su potencial de rendimiento en condiciones óptimas de cultivo es de 15 a 25 t•ha⁻¹ y no posee periodo de reposo. (Ñustez, 2011; Rodríguez y Ñustez, 2012; Ñustez, 2018). El cultivar Criolla Galeras (Gal), es un cultivar de porte medio, adaptado entre 2500 y 3000 msnm, con follaje verde ligeramente claro, floración intermedia, es un cultivar precoz (127 días a 2600 msnm) y susceptible a PYVV; su potencial de rendimiento en condiciones óptimas de cultivo es de 30 t•ha⁻¹ y posee un periodo de reposo de 20 días (Ñustez, 2011; Rodríguez y Ñustez, 2012; Ñustez, 2018).

Tabla 2. Genealogía de los genotipos de papa amarilla diploide evaluados en la prueba multiambiente.

No	Genotipo	Código U.N.	Genealogía
1	UN 4	08-004-16	{Criolla Guaneña x Criolla Galeras}
2	UN 9	08-004-59	{Criolla Guaneña x Criolla Galeras}
3	UN 50	08-051-64	{Criolla Latina x Criolla Colombia}
4	UN 51	08-051-68	{Criolla Latina x Criolla Colombia}
5	UN 52	08-051-70	{Criolla Latina x Criolla Colombia}
6	UN 59	08-051-90	{Criolla Latina x Criolla Colombia}
7	UN 63	08-053-53	{Criolla Galeras x Criolla Guaneña}
8	UN 64	08-053-57	{Criolla Galeras x Criolla Guaneña}
9	Criolla Colombia	Testigo Comercial	Selección clonal genotipos redondos amarillos tipo yema de huevo
10	Criolla Galeras	Testigo Comercial	{Criolla Colombia x Variedad Amarilla Tumbay [<i>Solanum goniocalyx</i>]}

2.2.2.1. Covariables Ambientales

Componente fisicoquímico del suelo

Se tomó una muestra compuesta de suelo entre los 0 y 20 cm de profundidad, con base en la metodología descrita por Castro y Gómez (2013), en cada localidad previo a la siembra. Las muestras fueron analizadas en el Laboratorio de Suelos y Aguas en la Facultad de Ciencias Agrarias de la Universidad Nacional de Colombia sede Bogotá.

Los parámetros edáficos (

Tabla 3) se determinaron mediante las siguientes metodologías de análisis; pH, relación suelo:agua (1:1); porcentaje de carbono orgánico (CO), Método de Walkley-Black; N estimado a partir del CO (factor empleado: 0,0862); Ca, K, Mg y Na intercambiables, extracción con acetato de amonio [1M], pH 7 y valoración por absorción atómica; acidez intercambiable (AI), extracción con KCl [1M]; capacidad de intercambio catiónico efectiva (CICE), estimada por adición de bases y acidez intercambiable; P, método de Bray II; Cu, Fe, Mn, Zn, extracción con DTPA y valoración por absorción atómica; B, extracción con fosfato monocálcico y valoración colorimétrica (Tabla 3); porcentaje de arcillas (Ar), limos (L) y arenas (A), método de Bouyoucos, dispersión con Na-Hexametafosfato; AH₂O, almacenamiento de agua en el suelo calculado por el método de Allen *et al.* (2006) mediante el uso de la siguiente formula:

$$AH2O = 1000(\theta_{FC} - \theta_{WP})Z_r$$

Donde, θ_{FC} es el contenido de humedad a capacidad de campo (m³/m³), θ_{WP} es el contenido de humedad en el punto de marchitez permanente (m³/m³) y Z_r es la profundidad radical (m).

Tabla 3. Parámetros químicos del suelo de las localidades de la prueba multiambiente de papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013

Localidad	Replica	pH	CO	N	Ca	K	Mg	Na	Al	CICE	P	Cu	Fe	Mn	Zn	B
			%	meq/100g					ppm							
Carlosama	1	5.8	1.54	0.13	8.72	1.52	2.14	0.11	0	12.5	90.2	4.57	148	14.1	3.35	0.26
Carlosama	2	5.6	4.5	0.68	6.62	0.74	1.15	0.05	0	8.56	59.9	0.82	164	5.03	1.67	0.16
Cumbal	1	4.7	7.66	0.66	2.46	0.29	0.41	0.11	3.61	6.88	31.5	1.78	151	4.92	2.76	0.19
Cumbal	2	5.1	7.88	0.66	2.51	0.67	0.73	0.07	1.98	5.96	15.8	1.21	114	1.57	3.55	0.44
Ipialpud	1	5.3	4.84	0.42	7.35	2.08	1.33	0.2	0.54	11.5	116	1.65	194	8.1	3.17	0.228
Ipialpud	2	5.7	6.22	0.3	6.58	0.99	1.05	0.06	0	8.68	92.9	1.08	183	1.54	1.05	0.17
Jamondino	1	5.5	2.36	0.2	7.36	0.6	2.68	0.17	0	10.8	20.8	6.69	248	31.4	2.47	0.14
Jamondino	2	5.2	3.5	0.39	5.1	0.86	2.52	0.09	0.43	8.99	37.5	4.23	358	19.8	1.64	0.25
Mosquera	1	6.3	4.65	0.4	20.5	2.71	5.68	1.78	0	30.7	111	2.18	294	19.8	50.6	0.59
Mosquera	2	5.8	4.67	0.4	20.4	2.16	6.71	6.07	0	35.4	116	1.94	247	16.4	76.5	0.56
Obonuco	1	5.8	3.88	0.33	11.7	3.52	3.5	0.13	0	18.9	116	4.01	368	34.3	11.3	0.59
Obonuco	2	5.7	2.68	0.23	9	1.71	1.41	0.06	0	12.2	116	2.58	240	22.2	6.15	0.55
San Ramón	1	5	5.59	0.48	5.5	1.44	1.69	0.08	1.62	10.3	46.7	0.45	358	1.77	0.9	0.26
San Ramón	2	5.3	7.68	0.3	5.28	2.8	3.78	0.07	0.63	12.6	100	0.76	310	2.32	4.99	0.67
Santa Bárbara	1	4.8	15.9	1.37	2.72	0.19	0.6	0.15	3.86	7.52	24.6	0.04	281	8.47	0.25	0.12
Santa Bárbara	2	4.7	11	0.64	1.43	0.57	0.49	0.08	3.04	5.61	32.7	1.24	206	4.44	2.89	0.21
Subachoque	1	5.3	11.8	1.02	6.88	2.42	2.06	0.17	0.74	12.3	15.6	0.38	69.6	2.45	2.62	0.27
Subachoque	2	5.4	12.6	0.1	9.22	1.46	2.3	0.07	0.47	13.5	8.54	0.46	129	1.86	1.3	0.17
Túquerres	1	5.3	4.82	0.42	2.61	0.96	0.67	0.06	0.92	5.23	83.6	0.94	134	8.88	1.27	0.18
Túquerres	2	5.6	7.45	0.8	3.14	0.91	1.22	0.06	0	5.33	6.8	0.61	86.6	5.4	1.52	0.17
Zipaquirá	1	4.7	10.3	0.89	4.25	1.54	1.56	0.08	3.72	11.1	116	0.46	216	2	1.77	0.43
Zipaquirá	2	5.2	9.28	0.31	6.38	0.8	1.4	0.06	1.11	9.74	92.9	0.13	152	0.33	1.32	0.31

CO: Carbono Orgánico; **Al:** Acidez Intercambiable; **CICE:** Capacidad de intercambio catiónico efectivo.

Además, se cuantificaron los parámetros hidráulicos (**Fehler! Ungültiger Eigenverweis auf Textmarke.**), U, límite superior del estado 1 de evaporación acumulativa del suelo; k,

conductividad hidráulica a -0.1 Bar de potencial matricial del suelo; α , coeficiente calculado del suelo, obtenidos a partir de la información reportada por Ritchie (1972).

Tabla 4. Parámetros físicos del suelo de las localidades de la prueba multiambiente de papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013.

Localidad	Semestre	Ar	L	A	Textura	AH2O	U	k	α
		%				m3/m3	mm	cm/día	mm/día
Carlosama	1	25	30	45	F	0.13	12	0.15	5.08
Carlosama	2	8	29	63	F	0.13	12	0.15	5.08
Cumbal	1	8	35	57	FA	0.11	9	0.1	4.04
Cumbal	2	6	31	63	FA	0.11	9	0.1	4.04
Ipialpud	1	9	31	59	FA	0.11	9	0.1	4.04
Ipialpud	2	6	35	59	FA	0.11	9	0.1	4.04
Jamondino	1	21	36	43	F	0.13	12	0.15	5.08
Jamondino	2	18	41	41	F	0.13	12	0.15	5.08
Mosquera	1	23	47	29	F	0.13	12	0.15	5.08
Mosquera	2	20	45	34	F	0.13	12	0.15	5.08
Obonuco	1	13	31	55	F	0.13	12	0.15	5.08
Obonuco	2	24	33	43	F	0.13	12	0.15	5.08
San Ramón	1	7	30	63	FA	0.11	9	0.1	4.04
San Ramón	2	6	33	61	FA	0.11	9	0.1	4.04
Santa Bárbara	1	5	19	75	FA	0.11	9	0.1	4.04
Santa Bárbara	2	4	43	53	FA	0.11	9	0.1	4.04
Subachoque	1	7	38	55	F	0.13	12	0.15	5.08
Subachoque	2	12	41	47	F	0.13	12	0.15	5.08
Túquerres	1	6	29	65	FA	0.11	9	0.1	4.04
Túquerres	2	4	25	71	FA	0.11	9	0.1	4.04
Zipaquirá	1	15	29	55	FA	0.11	9	0.1	4.04
Zipaquirá	2	6	33	61	FA	0.11	9	0.1	4.04

F: Franco; **FA:** Franco Arcilloso; **AH2O:** Almacenamiento de agua en el suelo; **U:** Límite superior del estado 1 de evaporación acumulativa del suelo; **k:** Conductividad hidráulica a -0.1 Bar de Potencial Matricial del suelo; **α :** Coeficiente calculado del suelo (Ritchie, 1972).

2.2.3 Componente Hídrico y Energético

Los datos climáticos diarios de precipitación (PP), humedad relativa (HR), temperatura promedio (T_{PROM}), temperatura mínima (T_{MIN}), temperatura máxima (T_{MAX}) y radiación solar (RAD) para cada localidad (

Tabla 5) se obtuvieron a través del proyecto de Predicción de Recursos Energéticos Mundiales (POWER) del Centro de Investigación Langley (LaRC) de la NASA. A partir de estos datos y utilizando las metodologías descritas por Allen *et al.* (2006), se realizaron las estimaciones diarias de las variables del componente hídrico y energético, presentadas a continuación:

Días sin agua disponible (DSAD), los cuales corresponden a la sumatoria de los días donde el agua fácilmente aprovechable (AFA) es igual a 0, siendo AFA determinado de la siguiente forma:

$$AFA = p(AH_2O)$$

Donde p es la fracción promedio del total de agua disponible en el suelo (AH_2O) que puede ser agotada de la zona radicular antes de presentarse estrés hídrico, cuyo valor para el cultivo de papa es 0,4 de acuerdo con Allen *et al.* (2006). Número de periodos de sequía (NPS), es la sumatoria de los periodos de sequía por ciclo, cada periodo de sequía corresponde a más de un día consecutivo con $AFA = 0$.

Consumo de agua por las plantas (CAP) y consumo de agua del cultivo (CAC), determinados mediante la estimación de la transpiración (E_p) y la evapotranspiración diaria del cultivo (E) respectivamente, con base en la metodología planteada por Ritchie (1972).

Radiación acumulada en el ciclo (RADA), estimada mediante la siguiente ecuación:

$$RADA = \sum_i^f RAD_i$$

Donde RAD_i es el valor de radiación para el i -ésimo día del ciclo de cultivo; f es la duración del ciclo de cultivo.

Radiación fotosintéticamente activa incidente acumulada (PARAi), la cual se determinó utilizando el método planteado por Soto *et al.* (2018), a partir de la ecuación:

$$PARAi = \sum_{i=1}^f 1 - e^{(-k_e * IAF_i)}$$

Donde k_e es el coeficiente de extinción del cultivo de papa amarilla diploide (0.65); IAF_i es el índice de área foliar para el i-ésimo día del ciclo de cultivo.

Tabla 5. Componente hídrico y energético de las localidades de la prueba multiambiente de papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013.

Localidad	Replica	DSAD	NPS	CAP	CAC	RADA	PARAi
				mm/ciclo		MJ/m2ciclo	
Carlosama	1	0	0	112	336	1529	256
Carlosama	2	12	2	119	338	1519	262
Cumbal	1	36	5	136	363	1702	292
Cumbal	2	43	4	116	355	1557	275
Ipialpud	1	0	0	110	311	1414	247
Ipialpud	2	25	6	119	338	1542	262
Jamondino	1	8	2	130	382	1791	315
Jamondino	2	19	7	134	366	1677	300
Mosquera	1	26	3	156	411	1957	356
Mosquera	2	0	0	163	438	1975	367
Obonuco	1	8	2	130	382	1791	315
Obonuco	2	23	8	134	371	1682	301
San Ramón	1	9	2	118	349	1707	293
San Ramón	2	23	6	123	335	1568	274
Santa Bárbara	1	10	4	123	354	1738	311
Santa Bárbara	2	26	7	131	357	1714	308
Subachoque	1	19	3	159	427	2037	379
Subachoque	2	0	0	165	441	1984	370
Túquerres	1	12	4	122	353	1737	312
Túquerres	2	24	6	129	348	1705	307
Zipaquirá	1	5	2	164	470	2303	409
Zipaquirá	2	3	1	186	501	2292	432

DSAD: Días sin agua disponible; **NPS:** Número de periodos de sequía; **CAP:** Consumo de agua por las plantas (transpiración del cultivo); **CAC:** Consumo de agua del cultivo (evapotranspiración); **RADA:** Radiación acumulada en el ciclo de cultivo; **PARAi:** Radiación fotosintéticamente activa incidente en el dosel del cultivo (Ritchie, 1972).

2.2.4 Métodos Analíticos

Los métodos analíticos fueron aplicados utilizando el entorno de análisis de datos R (R Core Team, 2020), para la estimación de los análisis de correlación, análisis de componentes principales y regresiones lineales simples y múltiples.

2.2.5 Estimación de Sensibilidad Fenotípica del Rasgo al Ambiente

Se aplicó el método de regresión lineal, relacionando la respuesta fenotípica de los rasgos de interés (variables dependientes) con las covariables explicativas de cada componente ambiental, utilizando el siguiente modelo para cada regresión:

$$y = \beta_0 + \beta_1 X_j + \varepsilon$$

Donde Y es la respuesta fenotípica (variable dependiente); β_0 es el intercepto del modelo; β_1 es la pendiente o coeficiente angular de regresión; X_j es la j -ésima covariable ambiental (variable independiente); ε son los residuales del modelo.

En cada regresión lineal simple se estimó como criterio de bondad de ajuste el coeficiente de determinación R^2 con la siguiente ecuación:

$$R^2 = \frac{SCR}{SCT}$$

donde SCR es la suma de cuadrados explicada por la regresión; SCT es la suma de cuadrados totales. Además del ajuste del modelo, también se evaluó la potencia predictiva del mismo, mediante un esquema de validación cruzada, utilizando el método de *leave one out* (LOO por sus siglas en inglés), estimando la diferencia predictiva de la raíz cuadrada media (RMSPD por sus siglas en inglés) utilizando la siguiente expresión:

$$RMSPD = \sqrt{\frac{\sum_{i=1}^n (y_{Pred} - y_{Obs})^2}{n}}$$

Donde y_{Obs} es la respuesta fenotípica obtenida en la PM; y_{Pred} es la respuesta fenotípica predicha por los modelos ajustados; n es el número de observaciones. Se seleccionaron las covariables de mayor facilidad para su medición o estimación con mayor R^2 y menor RMSPD para cada uno de los rasgos evaluados como parámetros del modelo múltiple de regresión lineal para explicar la varianza del fenotipo debida al ambiente.

2.2.6 Estimación de la Asociación Lineal en Componentes

Ambientales

En los componentes ambiental hídrico y energético, las covariables de cada uno de estos, se evaluaron en pares para estimar la intensidad de la relación lineal utilizando la metodología de análisis de correlación lineal mediante la siguiente ecuación:

$$\rho_{xy} = \frac{Cov_{xy}}{\sigma_x \sigma_y}$$

Donde Cov_{xy} es la covarianza entre las variables x e y ; σ_x es la desviación estándar para la covariable x ; σ_y es la desviación estándar para la covariable y . Las covariables con el mayor valor absoluto de correlación lineal en cada componente, fueron seleccionadas como candidatas para cada uno de los modelos de regresión simple utilizados en la descomposición de la respuesta fenotípica.

2.2.7 Análisis de Componentes Principales en Covariables

Fisicoquímicas del Suelo

Las covariables del componente ambiental suelo, fueron sometidas al análisis de componentes principales para seleccionar las covariables con mayor explicación de la varianza ambiental asociada al rasgo de interés mediante el método de descomposición en componentes principales. Cada componente principal (Z_i) se obtiene por combinación

lineal de las covariables ambientales. La primera componente principal de un grupo de variables (X_1, X_2, \dots, X_p) es la combinación lineal normalizada de dichas variables que tiene mayor varianza:

$$Z_i = \phi_{11}X_1 + \phi_{21}X_2 + \dots + \phi_{p1}X_p$$

Donde puede variar tanto como sea posible para los individuos, sujeto a la siguiente restricción para realizar la normalización lineal:

$$\sum_{j=1}^p \phi_{j1}^2 = 1$$

Siendo los términos $\phi_{1i}, \dots, \phi_{pi}$ son los pesos ponderados de la variable X_i de la primera componente principal. Normalizadas ($0, \sigma$) se definió la varianza como:

$$\sum_{j=1}^p \text{Var}(X_j) = \sum_{j=1}^p \frac{1}{n} \sum_{i=1}^n x_{ij}^2$$

Donde $\text{Var}(X_j)$ es la varianza de la variable normalizada; p es la cantidad máxima de variables; n es el número de observaciones; x_{ij}^2 es el valor observado en la i-ésima observación en la j-ésima variable. Posteriormente se estimó la proporción de la varianza explicada con la siguiente formula:

$$IPCA = \frac{\left(\sum_{j=1}^p \phi_{jm} x_{ij} \right)^2}{\sum_{j=1}^p \sum_{i=1}^n x_{ij}^2}$$

Siendo ϕ_{jm} el peso ponderado para la m-ésima variable dentro la j-ésima componente principal; Se determinaron las covariables con mayor explicación de la varianza en las dos primeras componentes principales (IPCA2) y se seleccionaron como candidatas para la evaluación en los modelos a ajustar para la descomposición de la respuesta fenotípica.

2.2.8 Procedimiento de Ajuste de Modelos Acumulativos

Durante el procedimiento de ajuste de los modelos, se utilizó la inferencia formal para evaluar la importancia de los parámetros incluidos en los modelos de análisis de IGA sobre la respuesta fenotípica de los rasgos de interés. Se determinó un modelo de descomposición de la varianza fenotípica definido a continuación:

$$y_j = \mu + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_j + \varepsilon$$

Donde y_i es la respuesta fenotípica para el rasgo de interés; μ es la gran media; $\beta_1 X_1$ es el efecto principal del componente hídrico; $\beta_2 X_2$ es el efecto principal del componente energético; $\beta_3 X_3$ es el efecto principal del componente físico químico del suelo; β_j es el residuo de la regresión; ε es el error residual con $\varepsilon \sim N(0, \sigma^2)$. Se emplearon diagnósticos residuales en todo momento para verificar los supuestos de distribución, en contraste con el proceso de selección de covariables. Se seleccionaron los parámetros más informativos, utilizando como criterios de desviación la estimación del RMSPD y de ajuste, el coeficiente de determinación ajustado. (R^2_{ad}) propuesto por Payne *et al.*, 1993):

$$R^2_{ad} = 1 - \left[\left(\frac{n-1}{n-k-1} \right) \right] * (1 - R^2)$$

Donde R^2 es el coeficiente de determinación de regresión lineal simple; n es el número de observaciones y k es el número de parámetros utilizados en el modelo.

2.3 Resultados y discusión

2.3.1 Sensibilidad Fenotípica del Rasgo al Ambiente

2.3.1.1 Rendimiento

Los cultivares pueden mostrar adaptaciones al entorno local de su área de cultivo tradicional (Keller *et al.*, 2011; Wolkovich *et al.*, 2014; Sanchez *et al.*, 2021), siendo la variabilidad genética determinante de la adaptación de los genotipos a diferentes ambientes de producción (Lande y Shannon, 1996; Sanchez *et al.*, 2021). Los coeficientes de determinación para las regresiones lineales simples entre las covariables

ambientales y los rasgos de interés (Tabla 6), muestran el alto grado de sensibilidad de los rasgos a ciertas covariables ambientales. El rendimiento se vio afectado por las covariables de los tres componentes que describen la varianza ambiental, principalmente por las covariables del componente hídrico y energético, presentando la mayor relación lineal con la precipitación ($R^2 = 0.2640$) y la altitud ($R^2 = 0.2074$). Esto indicaría un alto grado de dependencia del rendimiento del cultivo de papa amarilla diploide a las condiciones de humedad y radiación, siendo hasta el 26% de la varianza del rasgo explicada por condiciones hídricas y el 20% explicada por la altitud, sugiriendo que la variabilidad genética no es el único factor determinante, ya que un porcentaje importante de la variabilidad de los cultivares, es descrita por los componentes ambientales.

Tabla 6. Cuantificación del coeficiente de determinación (R^2) entre el rasgo fenotípico y las covariables ambientales en papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013. Coloraciones verdes indican mayor ajuste del modelo, rojizas menor ajuste del modelo al interior de cada rasgo.

Componente Ambiental	Covariable	R^2				
		Rendimiento	AR	CF	MS	GE
Hídrico	HR	0.2606	0.1538	0.0864	0.0726	0.0082
	PP	0.2640	0.0848	0.0830	0.0000	0.0054
	DSAD	0.0302	0.0092	0.0019	0.0153	0.0013
	NPS	0.0336	0.0128	0.0007	0.1224	0.1659
	CAP	0.2200	0.0465	0.0347	0.0054	0.0044
	CAC	0.1763	0.0613	0.0327	0.0111	0.0021
Energético	T _{PROM}	0.1453	0.1023	0.0156	0.1895	0.0635
	T _{MIN}	0.1223	0.1010	0.0122	0.1876	0.0646
	T _{MAX}	0.1253	0.0890	0.0108	0.2243	0.0778
	Altitud	0.2074	0.2102	0.1910	0.0558	0.0411
	RAD	0.1063	0.0362	0.0194	0.0002	0.0067
	RADA	0.1063	0.0362	0.0194	0.0002	0.0067
	PARAi	0.1531	0.0373	0.0278	0.0012	0.0062
Química de Suelo	pH	0.1748	0.2520	0.2088	0.1665	0.2108
	N	0.0008	0.0268	0.0171	0.0761	0.0820
	Ca	0.1373	0.0507	0.0474	0.2478	0.3245
	K	0.0001	0.0798	0.0348	0.2619	0.3617
	Mg	0.0611	0.0362	0.0609	0.2156	0.2701
	Na	0.1250	0.0018	0.0418	0.0993	0.1271
	Al	0.1477	0.2989	0.2829	0.0596	0.0722
	CICE	0.0874	0.0145	0.0263	0.2556	0.3347

P	0.0065	0.0268	0.0060	0.0488	0.1147
Cu	0.0154	0.0233	0.0384	0.0539	0.0177
Fe	0.1314	0.1013	0.0856	0.0003	0.0157
Mn	0.0493	0.0199	0.0741	0.1236	0.0665
Zn	0.0995	0.0000	0.0325	0.2042	0.2647
B	0.0229	0.0320	0.0253	0.2679	0.2637
Ar	0.0194	0.0991	0.0392	0.0488	0.0589
L	0.0837	0.0008	0.0011	0.0802	0.0638
A	0.0661	0.0315	0.0201	0.0969	0.0954
ArL	0.0645	0.0320	0.0202	0.0905	0.0880

HR: Humedad Relativa; **PP:** Precipitación **DSAD:** Dias sin agua disponible; **NPS:** Número de periodos de sequía; **CAP:** Consumo de agua por las plantas (transpiración del cultivo); **CAC:** Consumo de agua del cultivo (evapotranspiración); **RADA:** Radiación acumulada en el ciclo de cultivo; **PARA_i:** Radiación fotosintéticamente activa incidente en el dosel del cultivo (Ritchie, 1972); **CICE:** Capacidad de intercambio catiónico efectivo; **AI:** Acidez intercambiable; **Ar:** Arcilla; **A:** Arena; **L:** Limo.

En adición al parámetro de ajuste, un segundo criterio estimado para la validación de los modelos fue el RMSPD (

Tabla 7), el cual es un estimador de la potencia predictiva del modelo. En el componente hídrico las variables que presentaron los valores más bajos de RMSPD fueron PP con 2049.25 kg•ha¹, HR con 2111.07 kg•ha⁻¹ y CAP con 2146.87 kg•ha⁻¹, estas covariables también presentaron los mayores R², siendo seleccionadas para explicar la varianza ambiental asociada al componente hídrico en los 3 modelos evaluados. Algunos estudios han comparado el uso de agua de cultivares localmente restringidos frente a ampliamente difundidos (Medrano *et al.*, 2003; Satriani *et al.*, 2015), pero los mecanismos funcionales que subyacen a esta variabilidad entre cultivares son poco conocidos y entendidos. Sanchez *et al.*, (2021), evaluaron el efecto de la variación en la disponibilidad de agua en el cultivo de ajo, encontraron que cultivares con mayor potencial de rendimiento eran menos afectados que los cultivares con bajo potencial de rendimiento, por condiciones de exceso de agua; en contraste, cuando los cultivares fueron evaluados bajo condiciones de sequía, no presentaron variaciones significativas en los rendimientos alcanzados en todos los cultivares, indicando un alto grado de varianza debida al ambiente para este rasgo. Resultados similares fueron reportados en cultivos de girasol (Sandras *et al.*, 2009) y arveja (Sandras *et al.*, 2013).

Tabla 7. Estimación de la diferencia predictiva de la raíz cuadrada media (RMSPD) entre el rasgo fenotípico y las covariables ambientales en papa amarilla diploide en los departamentos de Nariño y Cundinamarca, para los años del 2012-2013. Coloraciones verdes indican menor precisión del modelo, rojizas mayor precisión del modelo al interior de cada rasgo.

Componente Ambiental	Covariable	RMSPD				
		Rendimiento	AR	CF	MS	GE
Hídrico	HR	2111.0734	0.0059	0.1146	0.5088	0.0022
	PP	2049.2535	0.0062	0.1136	0.5360	0.0022
	DSAD	2363.6490	0.0064	0.1192	0.5467	0.0022
	NPS	2391.0803	0.0065	0.1201	0.4857	0.0019
	CAP	2146.8725	0.0062	0.1165	0.5266	0.0022
	CAC	2242.4915	0.0062	0.1167	0.5259	0.0022
	Energético	T _{PROM}	2306.4765	0.0062	0.1208	0.4888
T _{MIN}		2343.6786	0.0063	0.1216	0.4906	0.0021
T _{MAX}		2327.5673	0.0063	0.1210	0.4786	0.0021
Altitud		2132.7216	0.0058	0.1067	0.5269	0.0022
RAD		2361.5444	0.0062	0.1156	0.5294	0.0021
RADA		2361.5444	0.0062	0.1156	0.5294	0.0021
PARAi		2264.4039	0.0062	0.1156	0.5258	0.0021
Química de Suelo	pH	2193.6406	0.0058	0.1110	0.4884	0.0020
	N	2400.2445	0.0066	0.1233	0.5001	0.0020
	Ca	2255.3413	0.0064	0.1184	0.4598	0.0018
	K	2441.3468	0.0062	0.1164	0.4635	0.0017
	Mg	2380.3993	0.0064	0.1153	0.4704	0.0019
	Na	2197.2700	0.0065	0.1253	0.7517	0.0038
	Al	2205.6556	0.0059	0.1048	0.5174	0.0021
	CICE	2317.8661	0.0064	0.1191	0.4655	0.0019
	P	2419.1355	0.0064	0.1215	0.5267	0.0021
	Cu	2371.9587	0.0063	0.1201	0.5218	0.0022
	Fe	2361.0389	0.0061	0.1168	0.5495	0.0022
	Mn	2348.6556	0.0063	0.1141	0.5134	0.0021
	Zn	2312.1814	0.0063	0.1184	0.5148	0.0022
	B	2456.9609	0.0063	0.1177	0.4638	0.0019
	Ar	2395.5612	0.0061	0.1196	0.5189	0.0021
	L	2330.7324	0.0067	0.1246	0.5125	0.0022
A	2348.6590	0.0063	0.1227	0.5041	0.0021	
ArL	2347.6730	0.0063	0.1227	0.5062	0.0021	

HR: Humedad Relativa; **PP:** Precipitación **DSAD:** Días sin agua disponible; **NPS:** Número de periodos de sequía; **CAP:** Consumo de agua por las plantas (transpiración)

del cultivo); **CAC**: Consumo de agua del cultivo (evapotranspiración); **RADA**: Radiación acumulada en el ciclo de cultivo; **PARA_i**: Radiación fotosintéticamente activa incidente en el dosel del cultivo (Ritchie, 1972); **CICE**: Capacidad de intercambio catiónico efectivo; **AI**: Acidez intercambiable; **Ar**: Arcilla; **A**: Arena; **L**: Limo.

En el componente energético, las covariables más informativas seleccionadas para los modelos de estimación de la varianza fenotípica, por presentar el valor más bajo de RMSPD (Tabla 7), fueron la altitud 2.132, seguida de la PARA_i con 2.264 y la T_{PROM} con 2.306 kg, estas covariables también presentaron los mayores R². El rendimiento de los cultivos está estrechamente relacionado con la duración del ciclo de cultivo; a una mayor duración del ciclo del cultivo las hojas pueden capturar más radiación, lo cual implica más energía que resulta en un alto rendimiento del dosel (Sinclair & Horie, 1989; Rahman *et al.*, 2019). Estudios como el realizado por Rahman *et al.*, (2019), donde evaluaron la geometría de siembra en el cultivo de arroz, la cual altera la cantidad de radiación incidente en el área foliar y la temperatura interna del cultivo, reportaron que la selección activa utilizando la geometría de plantación puede detectar genotipos de alta plasticidad fenotípica respecto al peso de las panículas, observando un alto grado de variabilidad en el rasgo debido a la radiación incidente en las láminas foliares en las diferentes condiciones de siembra evaluadas. Resultados similares fueron reportados en girasol y vid (Sandras *et al.*, 2009).

2.3.1.2 Calidad de Tubérculo

Existe una correlación directa entre los AR y la CF, ya que a medida que aumenta la concentración de los AR, aumenta el pardeamiento de la hojuela, característica cualitativa analizada en la medición de la CF, disminuyendo la calidad del producto en el procesamiento (Cambouris *et al.*, 2016). Se observó un bajo grado de dependencia lineal de los AR con las covariables del componente hídrico, siendo descrito solo el 15% de su varianza por la humedad relativa. Las covariables del componente fisicoquímico del suelo como el pH y AI describieron el 25% y 29% de la varianza del rasgo y presentaron los valores de RMSPD más bajos del componente, 0.0058 y 0.0059 respectivamente; esta misma tendencia se presentó en el rasgo de CF. Observando un buen ajuste del modelo con alta precisión al utilizar las covariables de pH y AI, indicando que estos rasgos podrían ser afectados principalmente por las propiedades de fertilidad de los suelos. Respecto al componente energético, los rasgos AR y CF, presentaron una relación lineal con la

covariable de altitud, explicando el 21% y 19% de la varianza ambiental y un RMSPD igual a 0.0058 y 0.1067 respectivamente, con la mayor precisión de predicción para estos rasgos. Por lo anterior se seleccionó como primer parámetro informativo del componente fisicoquímico del suelo al pH para ser utilizado en los modelos de regresión múltiple.

La MS presentó una mayor dependencia a las covariables del componente fisicoquímico del suelo, encontrando un alto grado de dependencia lineal del rasgo a covariables relacionadas con algunos macronutrientes como Ca, K y Mg, con valores de R^2 que oscilan entre 21-26%; y micronutrientes como B y Zn con valores de R^2 del 20 y 26% respectivamente, los RMSPD del modelo también fueron los menores para estas mismas covariables. Apreciándose una alta sensibilidad de este rasgo a las condiciones de fertilidad del suelo y una alta precisión en las predicciones obtenidas a partir de estas covariables como parámetros de estimación. La mayor parte de la información de la varianza ambiental de GE se asoció a las covariables del componente fisicoquímico del suelo, con un 36% para el contenido de K y valores superiores al 30% para la CICE y el Ca, indicando que el rasgo es afectado por las condiciones de fertilidad del suelo, lo cual podría implicar una dependencia lineal de la GE a los contenidos de cationes en el suelo.

En nuestro estudio se observó una alta explicación de la varianza de los rasgos de calidad por parte de las covariables relacionadas con el componente fisicoquímico del suelo, indicando que las condiciones de fertilidad del suelo son determinantes en la calidad de los tubérculos. Esta información puede ser trascendente en diferentes campos, como en la industria de procesamiento de papas, donde se ha analizado que la concentración de MS del tubérculo está altamente correlacionada con la GE, una alta concentración de MS evita la absorción de grasa durante el procesamiento en productos como papas fritas (Fontes et al, 2010) y a su vez el contenido de aceite de las papas fritas disminuye con el aumento de la gravedad específica (Schippers, 1976; Lulai y Orr 1979; Cambouris *et al.*, 2016).

Fontes *et al.*, (2010), evaluó el efecto de diferentes tasas de fertilización con nitrógeno sobre características de calidad de tubérculos en cuatro cultivares de papa. Observaron que las diferentes tasas de nitrógeno afectaban las enzimas involucradas en la calidad final del tubérculo, como la fosforilasa y la invertasa, responsables de la descomposición

del almidón en glucosa (Davies, 1990); y la hidrólisis de la sacarosa en monosacáridos, glucosa y fructosa, está regulada por el pH del tubérculo (Sowokinos, 1990).

Usualmente los valores de pH bajos o una alta concentración de Al generan una baja GE (Sun *et al.*, 2017), además, aumenta el contenido de azúcares reductores, indicando un patrón de sensibilidad de los rasgos de calidad a las características químicas del suelo. Patrones de respuesta fenotípica similares fueron reportados en batata dulce (Liu *et al.*, 2017).

2.3.1.3 Asociación Lineal en Componentes Ambientales

Componente Hídrico

En la Tabla 8 se aprecian los valores del coeficiente de correlación de Pearson con sus intervalos de confianza obtenidos para cada par de covariables ambientales explicativas del componente hídrico; los coeficientes de correlación más altos los presentó la covariable CAC con CAP y PP, con coeficientes R de 0,9677 y 0,8050 respectivamente, indicando una relación fuerte con pendiente positiva con ambas covariables; en contraste, se observó una correlación inversa fuerte entre CAC y la HR con un R de -0,8244; con base en estos resultados, el CAC se seleccionó por su asociación lineal con la mayor cantidad de covariables ambientales informativas del componente hídrico para el modelo 1. La siguiente covariable con mayor asociación lineal con el resto de covariables del componente hídrico fue la PP, con coeficientes de correlación de 0,8244 con CAP, 0,8050 con CAC y -0,7578 con HR, indicando una relación directamente proporcional con las transpiración y evapotranspiración, mientras presenta una relación inversamente proporcional con la HR; con base en esta información, la PP fue seleccionada para los modelos 2 y 3 por su asociación lineal con las covariables del componente hídrico y la fácil consecución de esta información en las localidades, en contraste con el CAC, que requiere un procesamiento de datos dependiente de varios parámetros de entrada.

El CAC determinado por Ritchie (1972) y la evapotranspiración (ET) estimada por el método de Penman – Monteith descrito por Allen *et al.*, (2006), son la combinación de los procesos por los cuales el agua se libera de la superficie del suelo por evaporación y del cultivo por transpiración. La evaporación de la atmósfera es expresada como evapotranspiración del cultivo de referencia (ET_o), representada por una superficie vegetada estandarizada, es un parámetro importante para comprender el balance hídrico

regional, los requisitos de riego, la estabilidad de la capa límite atmosférica y el pronóstico del tiempo. Es difícil medir la ETo directamente, por lo cual en la mayoría de las aplicaciones se estima mediante modelos (Beven, 1979).

Varios autores han demostrado que los parámetros meteorológicos pueden tener una correlación positiva o negativa con la evapotranspiración y pueden influir significativamente en ella según las condiciones climáticas del lugar. Bably (2003) demostró que la temperatura del aire, las horas de sol y la radiación solar, tenían una correlación positiva fuerte con la evapotranspiración, pero la precipitación y la humedad relativa presentaban una correlación negativa, estimando un coeficiente $r = -0.78$, valor cercano al -0.8244 obtenido en este estudio entre CAC y HR. Además, Isikwue *et al.*, (2015) reportaron una correlación negativa estrecha entre la nubosidad (alta asociación lineal con la precipitación) y la ETo evaluadas en tres ciudades de Nigeria. Resultados similares fueron reportados por Singh *et al.*, (1995) y Shrivastava *et al.*, (2001), quienes reportaron que la velocidad del viento, las horas de sol, la temperatura media del aire y la radiación solar se correlacionaron positivamente con la evaporación, mientras que la humedad relativa y la precipitación se correlacionaron negativamente con la evaporación.

Tabla 8. Correlación lineal entre las covariables del componente hídrico de la varianza ambiental.

Correlación	r	Límite inferior	Límite superior
DSAD-CAP	-0.2430	-0.3005	-0.1838
DSAD-CAC	-0.3125	-0.3673	-0.2554
DSAD-NPS	0.7433	0.7142	0.7698
DSAD-HR	0.0611	-0.0009	0.1226
DSAD-PP	-0.6072	-0.6449	-0.5665
CAP-CAC	0.9677	0.9635	0.9714
CAP-NPS	-0.2668	-0.3235	-0.2083
CAP-HR	-0.8526	-0.8687	-0.8348
CAP-PP	0.8244	0.8035	0.8433
CAC-NPS	-0.3590	-0.4118	-0.3038
CAC-HR	-0.8270	-0.8457	-0.8064
CAC-PP	0.8050	0.7820	0.8258
NPS-HR	0.2821	0.2240	0.3381
NPS-PP	-0.4747	-0.5214	-0.4253
HR-PP	-0.7578	-0.7830	-0.7301

HR: Humedad Relativa; **PP:** Precipitación **DSAD:** Dias sin agua disponible; **NPS:** Número de periodos de sequía; **CAP:** Consumo de agua por las plantas (transpiración del cultivo); **CAC:** Consumo de agua del cultivo (evapotranspiración);

Componente Energético

En la se observan los coeficientes de correlación lineal de Pearson con sus intervalos de confianza obtenidos para cada par de covariables ambientales informativas del componente energético. La covariable $PARA_i$ presentó el mayor grado de asociación lineal con las demás covariables ambientales con valores R de 0.9893 para RADA y 0.7239 para T_{PROM} , indicando una correlación positiva muy fuerte con ambas covariables, a diferencia de la altitud, con la cual presento un coeficiente R de -0.7670, indicando una correlación inversa fuerte entre las covariables; con base en lo anterior, se seleccionó la $PARA_i$ como parámetro informativo del componente energético para el modelo 1. La RADA fue la siguiente covariable con mayor asociación lineal con las demás covariables del componente energético, siendo seleccionada como parámetro para el modelo 2, con una correlación positiva fuerte con la T_{PROM} con un R de 0.7335 y una correlación inversa fuerte con la altitud con un R de -0.7118. Finalmente, para el modelo 3, la covariable seleccionada fue la altitud, la cual presentó coeficientes de correlación R de -0.7403 y -0.7118 con las covariables $PARA_i$ y RADA respectivamente, cabe resaltar que la altitud es una característica dependiente e invariante de la localidad evaluada.

La radiación solar transmitida tiene tres flujos principales entrantes según su rango de longitud de onda. El componente ultravioleta (rango espectral: 290 a 400 nm) es la porción de alta energía del espectro solar, es absorbido por la capa de ozono junto a otros constituyentes atmosféricos y tiene efectos nocivos en muchos sistemas biológicos (Parisi & Wong, 2000). El segmento espectral PAR, que cubre tanto los términos de fotones como de energía, se encuentra entre 400 y 700 nm en el espectro solar (McCree, 1972), siendo utilizado por las plantas en el proceso bioquímico de la fotosíntesis (Frisina, 2002).

Estudios han utilizado los modelos de regresión lineal simple para cuantificar la cantidad de varianza ambiental explicada por parámetros meteorológicos, como la

radiación global acumulada y la PAR_{Ai} . Uno de estos estudios fue el realizado por Escobedo *et al.*, (2011), quienes reportaron coeficientes de determinación y correlación superiores al 90% entre las variables de radiación solar global y radiación PAR, indicando que se podría seleccionar una sola variable para explicar la varianza contenida en ambas, sin pérdida de información.

Tabla 9 se observan los coeficientes de correlación lineal de Pearson con sus intervalos de confianza obtenidos para cada par de covariables ambientales informativas del componente energético. La covariable PAR_{Ai} presentó el mayor grado de asociación lineal con las demás covariables ambientales con valores R de 0.9893 para RADA y 0.7239 para T_{PROM} , indicando una correlación positiva muy fuerte con ambas covariables, a diferencia de la altitud, con la cual presentó un coeficiente R de -0.7670, indicando una correlación inversa fuerte entre las covariables; con base en lo anterior, se seleccionó la PAR_{Ai} como parámetro informativo del componente energético para el modelo 1. La RADA fue la siguiente covariable con mayor asociación lineal con las demás covariables del componente energético, siendo seleccionada como parámetro para el modelo 2, con una correlación positiva fuerte con la T_{PROM} con un R de 0.7335 y una correlación inversa fuerte con la altitud con un R de -0.7118. Finalmente, para el modelo 3, la covariable seleccionada fue la altitud, la cual presentó coeficientes de correlación R de -0.7403 y -0.7118 con las covariables PAR_{Ai} y RADA respectivamente, cabe resaltar que la altitud es una característica dependiente e invariante de la localidad evaluada.

La radiación solar transmitida tiene tres flujos principales entrantes según su rango de longitud de onda. El componente ultravioleta (rango espectral: 290 a 400 nm) es la porción de alta energía del espectro solar, es absorbido por la capa de ozono junto a otros constituyentes atmosféricos y tiene efectos nocivos en muchos sistemas biológicos (Parisi & Wong, 2000). El segmento espectral PAR, que cubre tanto los términos de fotones como de energía, se encuentra entre 400 y 700 nm en el espectro solar (McCree, 1972), siendo utilizado por las plantas en el proceso bioquímico de la fotosíntesis (Frisina, 2002).

Estudios han utilizado los modelos de regresión lineal simple para cuantificar la cantidad de varianza ambiental explicada por parámetros meteorológicos, como la radiación global acumulada y la PAR_{i} . Uno de estos estudios fue el realizado por Escobedo *et al.*, (2011), quienes reportaron coeficientes de determinación y correlación superiores al 90% entre las variables de radiación solar global y radiación PAR, indicando que se podría seleccionar una sola variable para explicar la varianza contenida en ambas, sin pérdida de información.

Tabla 9. Correlación lineal entre las covariables del componente energético de la varianza ambiental.

Correlación	r	Límite inferior	Límite superior
PAR_{i} -RADA	0.9893	0.9879	0.9906
PAR_{i} - T_{PROM}	0.7239	0.6930	0.7522
PAR_{i} -Altitud	-0.7403	-0.7670	-0.7109
RADA- T_{PROM}	0.7335	0.7035	0.7609
RADA-Altitud	-0.7118	-0.7411	-0.6799
T_{PROM} -Altitud	-0.5770	-0.6169	-0.5341

RADA: Radiación acumulada en el ciclo de cultivo; **PAR_{i} :** Radiación fotosintéticamente activa incidente en el dosel del cultivo.

Componentes Principales en Covariables Físicoquímicas del Suelo

El análisis de componentes principales (ACP), es una técnica para reducir el número de variables y eliminar las relaciones entre las variables de entrada, mediante el desarrollo de un conjunto de índices o componentes principales (CP) que son funciones lineales de las variables originales (Manly & Alberto, 2016). En la figura 1 se muestra la distribución de las coordenadas para el componente físicoquímico del suelo en el biplot de ACP, se pueden distinguir 3 grupos de covariables que presentan un alto grado de correlación lineal entre sí; un primer grupo ubicado a la derecha, compuesto por la Al y el N, indicando una alta correlación lineal entre sí, afectados principalmente por la CP1; el segundo grupo de variables con alta correlación lineal entre sí, está compuesto por la textura del suelo y el pH; el tercer grupo está compuesto por los demás nutrientes del suelo. En la Tabla 10 se muestran los vectores propios (eigenvectors) de los dos primeros componentes principales (CP), los valores propios (eigenvalues), 9.3299 para el CP1 y 2.3809 para el CP2, el

primer componente principal, por tanto, tiene una varianza de 9.3299 variables originales y el segundo componente tiene una varianza de 2.3809 variables originales, siendo la varianza explicada por cada componente del 77.4% para la CP1 y del 9.44% para la CP2, con una varianza acumulada del 87% conteniendo la información de 11.7108 de las 18 covariables originales evaluadas, dicho de otra forma el 65.05% de las covariables contiene el 87% de la varianza del componente fisicoquímico del suelo.

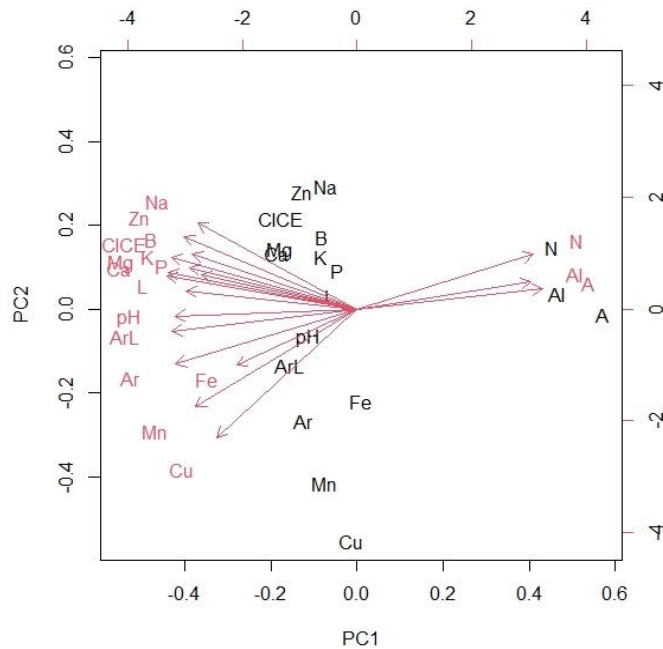


Figura 1. Plano factorial de las covariables del componente químico del suelo en la varianza ambiental.

Tabla 10. Matriz de valores propios del ACP entre las covariables del componente fisicoquímico del suelo de la varianza ambiental. Coloraciones verdes indican mayor peso de la covariable en la CP con signo positivo, rojizas mayor peso de la covariable en la CP. Coloraciones amarillas indican un bajo peso de la covariable en la CP.

Variable	CP1	CP2
pH	0.2439	-0.0986
N	-0.1794	0.3123
Ca	0.3056	0.1323
K	0.2127	0.1631
Mg	0.2979	0.1691
Na	0.2229	0.3194
Al	-0.1942	0.2125
CICE	0.2979	0.2373
P	0.1794	0.1329
Cu	0.1628	-0.4922
Fe	0.1343	-0.0319
Mn	0.2080	-0.3246
Zn	0.2638	0.3018
B	0.2098	0.2276
Ar	0.2550	-0.2546
L	0.2137	0.0101
A	-0.2850	0.1393
ArL	0.2817	-0.1519
Eigenvalue	9.3299	2.3809
Desv. Estándar	3.7334	1.3037
Prop. Varianza	0.7744	0.0944
Var. Acumulada	0.7744	0.8688

CICE: Capacidad de intercambio catiónico efectivo; **Al:** Acidez intercambiable; **Ar:** Arcilla; **A:** Arena; **L:** Limo. El coeficiente resaltado indica la covariable ambiental seleccionada para el CP.

Con base en el ACP, se observó que en el CP1 las covariables con mayores eigenvalues fueron Ca (0.3056), Mg (0.2979), CICE (0.2979) y ArL (0.2817) con una relación positiva, indicando que estas tienen un alto peso en la componente principal. El Ca y el Mg son cationes intercambiables, a partir de los cuales se estima la CICE, y ya que las arcillas y la materia orgánica del suelo se comportan como aniones, son capaces de retener o adsorber los cationes intercambiables (Garrido, 1993). Ya que este conjunto de covariables presentó una alta correlación entre sí, se seleccionó la CICE como

segundo parámetro informativo del componente fisicoquímico del suelo en los modelos de regresión múltiple. Resultados similares fueron observados en el estudio desarrollado por Gil *et al.* (2005), quienes realizaron la selección de propiedades del suelo mediante un análisis de componentes principales (ACP), en cultivos de cacao y banano; reportaron mayor peso de las covariables K y limo en la CP1, siendo contenida el 33.08% de la varianza; la CP2 y CP3, presentaron el mayor peso en las covariables Mg y Ca respectivamente, acumulando las 3 componentes un 94.89% de la varianza total de los datos originales. En el mismo sentido, Panishkan *et al.* (2012) estudiaron la clasificación de un conjunto de muestras de suelo pertenecientes a 67 localidades de 3 provincias en Tailandia, según sus propiedades mediante el análisis de componentes principales; las propiedades evaluadas fueron materia orgánica (OM), nitrógeno total (TN), capacidad de intercambio de catiónico (CEC), densidad aparente (BD), porcentaje de limos (Silt) y porcentaje de arcillas (Clay); estimando que el CP1 explicó el 41,57% de la variación total y el CP2 el 30,67%, acumulando un 72,24% de la varianza total. La CP1 mostró altos pesos para las propiedades de OM y TN con efecto positivo y BD con efecto negativo, la CP2 se asoció con CEC y Clay, ambos con efectos positivos.

2.3.2 Modelos Ajustados

Con base en la información obtenida en las secciones de sensibilidad fenotípica y asociación lineal, se determinaron las covariables más informativas para cada uno de los componentes ambientales. Estas fueron filtradas mediante el uso de criterios de correlación lineal (r), bondad de ajuste (R^2) y precisión (RMSPD); seleccionando las covariables con el mayor r , R^2 y menor RMSPD respecto al resto de covariables de cada componente, para cada uno de los rasgos evaluados. Las covariables seleccionadas se utilizaron en tres combinaciones distintas como parámetros de entrada en los modelos de regresión múltiple, para la predicción de la respuesta fenotípica a partir del efecto ambiental. A continuación, se presentan las ecuaciones propuestas para cada uno de los modelos:

$$M1 \rightarrow \hat{y} = \mu + CAC_j + PAR_{Ai_j} + (pH_j + CICE_j) + \varepsilon$$

$$M2 \rightarrow \hat{y} = \mu + PP_j + RADA_j + (pH_j + CICE_j) + \varepsilon$$

$$M3 \rightarrow \hat{y} = \mu + PP_j + Altitud_j + (pH_j + CICE_j) + \varepsilon$$

Donde \hat{y} es la respuesta fenotípica estimada para la observación del rasgo de interés; μ es la gran media; CAC_j y PP_j son el consumo de agua del cultivo y la precipitación diaria, covariables explicativas del efecto principal del componente hídrico de la j -ésima (1,2,...,j) localidad; $PARA_{ij}$, $RADA_j$ y $Altitud_j$, son la radiación fotosintéticamente activa incidente en el dosel, la radiación acumulada en el ciclo del cultivo y la altitud de la localidad, covariables explicativas del efecto principal del componente energético; $(pH_j + CICE_j)$, corresponden al pH y a la capacidad de intercambio catiónico efectiva, covariables explicativas del efecto principal del componente fisicoquímico del suelo; ε es el error residual con $\varepsilon \sim N(0, \sigma^2)$.

En la Tabla 11 se muestran los RMSPD y los R^2 ad, de cada uno de los modelos evaluados para la predicción de la respuesta fenotípica a partir del efecto ambiental. El modelo 3 se encuentra resaltado, ya que presenta el mayor potencial predictivo para todos los rasgos y el mejor ajuste para los rasgos AR y CF, explicando el 53.52% y el 40.31% respectivamente de la varianza fenotípica asociada al ambiente. El modelo 2 por su parte, presenta el mejor ajuste para los rasgos de rendimiento, MS y GE, sin embargo, la diferencia de los R^2 del modelo 3 respecto al modelo 2 no superan el 1% de la varianza explicada. Con base en esta información se podría determinar entonces que el mejor modelo para la predicción de la respuesta fenotípica de los rasgos evaluados es el modelo 3.

Tabla 11. Coeficientes de determinación y diferencia predictiva de la raíz cuadrada media para 3 modelos de estimación de la respuesta fenotípica en cultivos de papa amarilla diploide ajustados a partir de diferentes parámetros ambientales.

Modelo	Parámetro	Rendimiento	AR	CF	MS	GE
1	RMSPD	2443.2046	0.0058	0.1407	0.5198	0.0020
	R2	0.3987	0.5105	0.3365	0.3087	0.4190
2	RMSPD	2338.8765	0.0057	0.1416	0.4963	0.0017
	R2	0.4183	0.5025	0.3513	0.3837	0.5503
3	RMSPD	2212.6868	0.0052	0.1338	0.4849	0.0017
	R2	0.4084	0.5352	0.4031	0.3739	0.5483

Landau *et al.*, (2000), desarrollaron un modelo de regresión múltiple de los efectos del clima en el rendimiento del trigo, utilizaron un set de datos del rendimiento de cultivos de trigo de invierno en el Reino Unido, dividiendo el ciclo de cultivo en cinco fases. Dentro

de cada iteración del procedimiento, se evaluaron las expresiones de los efectos climáticos esperados por su potencial para explicar la variación en los rendimientos observados. Se compararon expresiones alternativas de un solo efecto climático esperado y las variables explicativas involucradas en la mejor expresión, si alguna de ellas estaba involucrada se agregaba al conjunto de posibles variables explicativas. El porcentaje de variación del rendimiento explicado por un modelo de regresión se midió mediante el coeficiente de determinación ajustado R^2 ad. A partir del modelo propuesto se determinó que la duración del ciclo del cultivo es afectada por la radiación interceptada y dependiendo de la fenología sus efectos pueden ser positivos o negativos sobre el rendimiento del cultivo, conteniendo un 12% de la variación explicada por el modelo; altas precipitaciones tienen un efecto negativo sobre el rendimiento, explicando el 6% de la variación.

Estudios recientes, como el desarrollado por Bhattacharyya *et al.*, (2021), utilizaron modelos de regresión lineal para estudiar los efectos de las variables del clima en el rendimiento de distintos cultivos, arroz, maíz, semillas oleaginosas, legumbres, verduras, piña, cítricos, chiles, jengibre y cúrcuma; los resultados mostraron que los rendimientos de los cultivos dependen significativamente de las variables climáticas, las precipitaciones influyen significativamente en el rendimiento del arroz, el maíz, la piña, las verduras y el jengibre y no mostraron ningún efecto en otros cultivos, ya que en su mayoría son cultivos de invierno, el rendimiento de estos últimos se explicó principalmente por la temperatura y la humedad relativa.

Coulibali *et al.*, (2020), desarrollaron un modelo óptimo para predecir los requerimientos de NPK para un alto rendimiento y calidad de los tubérculos (tamaño y gravedad específica, GE), según el impacto de las variables climáticas y de suelos mediante un algoritmo de machine learning; utilizaron un set de datos de 273 ensayos de campo realizados entre 1979 y 2017 en Canadá, encontrando que los modelos de predicción de GE fueron los más informativos, como lo demuestran los valores más altos de R^2 , oscilaron entre 0,72 (KNN) y 0,94 (RF). El algoritmo de selección de características mostró el impacto de los índices climáticos en la GE del tubérculo; la precipitación total y la temperatura arrojaron las puntuaciones más altas en orden decreciente.

Al Soboh *et al.* (2002) revisaron los factores que afectan la GE en cultivos de papa destinada a fritura y enfatizaron que el riego durante las primeras etapas de crecimiento aumenta el contenido de materia seca (MS) del tubérculo. La gravedad específica podría reducirse sustancialmente si se produjeran fuertes lluvias antes de la cosecha; y afirmaron que las papas cultivadas durante un período de aumento de la duración del día, la temperatura y la intensidad de la luz producen tubérculos con alta GE. Por otra parte, Coulibali *et al.*, 2020 reportó que dosis de aplicación excesivas de N y K junto con altos niveles en el suelo de cualquiera de los nutrientes pueden reducir la GE, mientras la aplicación de fósforo puede aumentar la materia seca cuando los niveles de P en las pruebas de suelo son bajos (Maier *et al.*, 1994).

En un estudio Nelson & Shaw (1976) informaron que el gradiente de temperatura del suelo en una colina gobierna el nivel de azúcares en las papas, especialmente en el período de madurez (Kumar *et al.*, 2003). Carlson (1970), encontró que los cambios en los azúcares eran menores a una temperatura de 10°C que a 4-5°C. Duarte-Delgado *et al.*, (2015) y Duarte-Delgado *et al.*, (2016) evaluaron un grupo de papas del grupo Phureja sembradas a alturas 3400 y 2905 msnm respectivamente, al evaluar los contenidos de AR en las variedades comerciales Colombia, Galeras y Guaneña, se contrasto entre los estudios, que a mayor altura de siembra mayor fue la concentración de AR en la cosecha. Sustentando que temperaturas bajas aumentan el contenido de azúcares reductores durante el desarrollo de los tubérculos.

Los azúcares reductores tienen un papel crítico en la determinación de la calidad de las papas destinadas a fritura (CF), ya que altas concentraciones de AR en los tubérculos dan como resultado productos fritos de color marrón y sabor amargo, características no deseadas en el procesamiento industrial. La coloración marrón también conocida como pardeamiento no enzimático, resulta de la reacción de Maillard, que se da por la interacción de los azúcares reductores con los aminoácidos, la cual tiene lugar en condiciones de alta temperatura y poca humedad (Kumar *et al.*, 2004). La información reportada coincide con resultados de este estudio, sustentado la influencia ambiental que existe sobre los rasgos de AR y CF.

2.4 Conclusión

El presente estudio reveló que los rasgos de papa amarilla diploide, rendimiento, AR, CF, MS y GE, se ven altamente influenciados por la variación en las condiciones ambientales entre los entornos de crecimiento. Con base en los resultados de los análisis de correlación y ACP, la precipitación, el consumo de agua del cultivo (componente hídrico); la altitud (componente energético); el pH y la CICE (componente fisicoquímico), fueron las covariables ambientales seleccionadas, por su alta varianza explicada y potencial predictivo, para ser utilizadas como parámetros en los modelos de regresión lineal evaluados. Cabe resaltar que estas covariables ofrecen una ventaja y es su fácil y práctica medición.

El modelo que mejor explicó la varianza fenotípica asociada al ambiente de los rasgos evaluados fue el modelo 3. En el cual se utilizaron como parámetros la precipitación, la altura, el pH y la CICE, los cuales no requieren un procesamiento previo de los datos y por lo cual acumulan una mayor varianza explicada del fenotipo debida al ambiente.

2.5 Referencias

Al Soboh, G., Sully, R., & Andreatta, S. (2002). Factors affecting specific gravity loss in crisping potato crops in Koo Wee Rup, Victoria. Department of Natural Resources and Environment. ISBN: 0 7341 0531 2.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (2006). Evapotranspiración del cultivo: guías para la determinación de los requerimientos de agua de los cultivos. Roma: FAO, 298(0). ISBN: 92-5-304219-2

Bandera, E. & Pérez, L. (2018). Generalized Linear Mixed Models. Its application in plant breeding. *Cultivos Tropicales*, 39(1), 127-133.
https://www.researchgate.net/publication/327751933_Review_Generalized_linear_mixed_models_Its_application_in_plant_breeding

Bably, A. (2003). Estimation of evapotranspiration using statistical model. Camarda, D. (ed.), Grassin, I. L. (Ed.), Local resources and global trades: Environments and agriculture in the Mediterranean region. CIHEAM. 441-449.

Beven, K. (1979). A sensitivity analysis of the Penman-Monteith actual evapotranspiration estimates. *Journal of Hydrology*, 44(3-4), 169-190. [https://doi.org/10.1016/0022-1694\(79\)90130-6](https://doi.org/10.1016/0022-1694(79)90130-6)

Bhattacharyya, B., Biswas, R., Sujatha, K. & Chiphang, D. (2021). Linear Regression Model to Study the Effects of Weather Variables on Crop Yield in Manipur State. *Int. J. Agricult. Stat. Sci*, 17(1), 317-320. ISSN: 0973-1903, e-ISSN: 0976-3392

Bornhofen, E., Todeschini, M., Stoco, M., Madureira, A., Marchioro, V., Storck, L. & Benin, G. (2018). Wheat yield improvements in Brazil: Roles of genetics and environment. *Crop Science*, 58(3), 1082-1093. <https://doi.org/10.2135/cropsci2017.06.0358>

Brancourt, M., Denis, J. & Lecomte, C. (2000). Determining environmental covariates which explain genotype environment interaction in winter wheat through probe genotypes and biadditive factorial regression. *Theoretical and Applied Genetics*, 100(2), 285-298. <https://doi.org/10.1007/s001220050038>

Betancur, L., Posada, S. & Solano, R. (2012). Application of the principal-component analysis in the evaluation of three grass varieties. *Revista Colombiana de Ciencias Pecuarias*, 25(2), 258-266. https://www.researchgate.net/publication/262505522_Application_of_the_principal-component_analysis_in_the_evaluation_of_three_grass_varieties

Cambouris, A., St. Luce, M., Zebarth, B., Ziadi, N., Grant & C., Perron, I. (2016). Potato response to nitrogen sources and rates in an irrigated sandy soil. *Agronomy Journal*, 108(1), 391-401. <http://dx.doi.org/10.2134/agronj2015.0351>

Carlson, H. (1970). Production of potatoes for chipping. Uppsala, Sweden: The Agricultural College of Sweden.
<https://www.cabdirect.org/cabdirect/abstract/19711702322>

Cecchi, H. (1999). Fundamentos teóricos e práticos em análise de alimentos (Theoretical and practical in food analysis) Campinas/SP. UNICAMP, Brazil.
https://issuu.com/editoraunicamp/docs/20pp_fundamentos_teoricos_e_pratico

Coulibali, Z., Cambouris, A. & Parent, S. (2020). Site-specific machine learning predictive fertilization models for potato crops in Eastern Canada. PloS one, 15(8),
<https://doi.org/10.1371/journal.pone.0230888>

Crossa, J., Burgueño, J., Cornelius, P., McLaren, G., Trethowan, R. & Krishnamachari, A. (2006). Modeling genotype x environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. *Crop science*, 46(4), 1722-1733. <https://doi.org/10.2135/cropsci2005.11-0427>

Crossa, J. (2012). From genotype x environment interaction to gene x environment interaction. *Curr. Genomics* 13:225–244. <https://doi.org/10.2174/138920212800543066>

Davies, H. (1990). Carbohydrate metabolism during sprouting. *Am J Potato* 67:815–827
Feibert EBG, Shock CC, Saunders LD (1998) Nitrogen fertilizer requirements of potatoes using carefully scheduled sprinkler irrigation. *HortScience* 33:262–265.
<https://doi.org/10.1007/BF03044293>

Delacy, I., Basford, K., Cooper, M. & Bull, J. (1996) Analysis of multi-environment trials an historical perspective. *Plant Adaptation and Crop Improvement*. Eds. M. Cooper and G. L. Hammer. CAB international, UK.
<https://www.cabdirect.org/cabdirect/abstract/19961610824>

Dias, K., Dos Santos, J., Krause, M., Piepho, H., Guimarães, L., Pastina, M., & Garcia, A. (2022). Leveraging probability concepts for cultivar recommendation in multi-environment trials. *Theoretical and Applied Genetics*, 1-15. <https://doi.org/10.1007/s00122-022-04041-y>

Duarte, D., Narváez, C., Restrepo, L., Kushalappa, A., & Mosquera, T. (2015). Development and validation of a liquid chromatographic method to quantify sucrose, glucose, and fructose in tubers of *Solanum tuberosum* Group Phureja. *Journal of Chromatography B*, 975, 18-23. <https://doi.org/10.1016/j.jchromb.2014.10.039>

Duarte, D., Núñez, C., Narváez, C., Restrepo, L., Melo, S., Sarmiento, F. & Mosquera, T. (2016). Natural variation of sucrose, glucose and fructose contents in Colombian genotypes of *Solanum tuberosum* Group Phureja at harvest. *Journal of the Science of Food and Agriculture*, 96(12), 4288-4294. <https://doi.org/10.1002/jsfa.7783>

Eberhart, S. & Russell, W. (1966). Stability parameters for comparing varieties. *Crop Science*, 6: 36-40. <https://doi.org/10.2135/cropsci1966.0011183X000600010011x>

Escobedo, J., Gomes, E., Oliveira, A. & Soares, J. (2011). Ratios of UV, PAR and NIR components to global solar radiation measured at Botucatu site in Brazil. *Renewable Energy*, 36(1), 169-178. <https://doi.org/10.1016/j.renene.2010.06.018>

Finlay, K. & Wilkinson, F. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal Agricultural Resource*, 14: 742-754. https://pdf.usaid.gov/pdf_docs/PNAAS139.pdf

Fontes, P., Braun, H., Busato, C. & Cecon, P. (2010). Economic optimum nitrogen fertilization rates and nitrogen fertilization rate effects on tuber characteristics of potato cultivars. *Potato Research*, 53(3), 167-179. <https://doi.org/10.1007/s11540-010-9160-3>

Frisina, V. (2002). Modelagem das radiações global, difusa e fotossinteticamente ativa em ambiente protegido e suas relações com o crescimento e produtividade da cultura de pimentão (*Capsicum annuum* L.). https://repositorio.unesp.br/bitstream/handle/11449/101753/frisina_va_dr_botfca.pdf?sequence=1&isAllowed=y

Frutos, E. 2011. Interacción genotipo-ambiente: GGE Biplot y modelos AMMI. Departamento de Estadística. Universidad de Salamanca. (Disertación doctoral). <http://hdl.handle.net/10366/121900>

Garrido, S. (1993). Interpretación de análisis de suelos. Hoja Divulgativa Nro, 5. https://www.mapa.gob.es/ministerio/pags/biblioteca/hojas/hd_1993_05.pdf

Gil, F., Trasar, C., Leirós, M. & Seoane, S. (2005). Different approaches to evaluating soil quality using biochemical properties. *Soil biology and biochemistry*, 37(5), 877-887. <https://doi.org/10.1016/j.soilbio.2004.10.003>

Isikwue, B., Audu, M. & Eweh, E. (2015). Correlation of evapotranspiration with climatic parameters in some selected cities in Nigeria. *Journal of earth sciences and Geotechnical Engineering*, 5(4), 103-115. https://www.researchgate.net/publication/316857440_Correlation_of_Evapotranspiration_with_Climatic_Parameters_in_some_selected_Cities_in_Nigeria

Jhajharia, D., Kithan, S., & Fancon, A. (2006). Correlation between pan evaporation and meteorological parameters under the climatic conditions of Jorhat (Assam). *Journal of Indian Water Resources Society*, 26(1-2), 39-42. https://www.researchgate.net/publication/235004367_Correlation_between_pan_evaporation_and_meteorological_parameters_under_the_climatic_conditions_of_Jorhat_Assam

Keller, S., Soolanayakanahally, R., Guy, R., Silim, S., Olson, M. & Tiffin, P., (2011). Climate-driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus balsamifera* L. (Salicaceae). *Am. J. Bot.* 98, 99–108. <https://doi.org/10.3732/ajb.1000317>

Kumar, D., Ezekiel, R., & Khurana, S. P. (2003). Effect of location, season and cultivar on the processing quality of potatoes. *J. Indian Potato Assoc*, 30(3-4), 247-251. https://www.researchgate.net/publication/269108056_Effect_of_location_season_and_cultivar_ob_the_processing_quality_of_potatoes

Kumar, D., Singh, B. P., & Kumar, P. (2004). An overview of the factors affecting sugar content of potatoes. *Annals of Applied Biology*, 145(3), 247-256. <https://doi.org/10.1111/j.1744-7348.2004.tb00380.x>

Landau, S., Mitchell, R., Barnett, V., Colls, J., Craigon, J. & Payne, R. (2000). A parsimonious, multiple-regression model of wheat yield response to environment. *Agricultural and forest meteorology*, 101(2-3), 151-166. [https://doi.org/10.1016/S0168-1923\(99\)00166-5](https://doi.org/10.1016/S0168-1923(99)00166-5)

Liu, T., Dodds, E., Leong, S., Eyres, G., Burritt & D., Oey, I. (2017). Effect of pulsed electric fields on the structure and frying quality of “kumara” sweet potato tubers. *Innovative food science & emerging technologies*, 39, 197-208. <https://doi.org/10.1016/j.ifset.2016.12.010>

Lulai, E. & Orr, P. (1979). Influence of potato specific gravity on yield and oil content of chips. *Am Potato J* 56:379–390. <https://doi.org/10.1007/BF02855348>

Maier, N., Dahlenburg, A. & Williams, C. (1994). Effects of nitrogen, phosphorus, and potassium on yield, specific gravity, crisp colour, and tuber chemical composition of potato (*Solanum tuberosum* L.) cv. Kennebec. *Australian Journal of Experimental Agriculture*, 34(6), 813-824. ISSN: 0816-1089

Malosetti, M., Ribault, J. & van Eeuwijk, F. (2013). The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. *Front. Physiol.* 4:1–17. <https://doi.org/10.3389/fphys.2013.00044>

Manly, B. & Alberto, J. (2016). *Multivariate statistical methods: a primer*. Chapman and Hall/CRC. ISBN: 9781498728966.

Medrano, H., Escalona, J., Cifre, J., Bota, J. & Flexas, J. (2003). A ten-year study on the physiology of two Spanish grapevine cultivars under field conditions: effects of water availability from leaf photosynthesis to grape yield and quality. *Funct. Plant Biol.* 30, 607–619. <http://dx.doi.org/10.1071/FP02110>

Nelson, D. C., & Shaw, R. (1976). Effect of planting and harvest dates, location in the hill and tuber size on sugar content of Kennebec potatoes. *American Potato Journal*, 53(1), 15-21. <https://doi.org/10.1007/BF02854087>

Oliveira, I., Guilhen, J., Oliveira, P., Gezan, S., Schaffert, R., Simeone, M. & Pastina, M. (2020). Genotype-by-environment interaction and yield stability analysis of biomass sorghum hybrids using factor analytic models and environmental covariates. *Field Crops Research*, 257. <https://doi.org/10.1016/j.fcr.2020.107929>

Panishkan, K., Swangjang, K., Sanmanee, N. & Sungthong, D. (2012). Principal component analysis for the characterization in the application of some soil properties. *International Journal of Environmental and Ecological Engineering*, 6(5), 279-281. <https://doi.org/10.5281/zenodo.1058231>

Parisi, A. & Wong, J. (2000). An estimation of biological hazards due to solar radiation. *Journal of Photochemistry and Photobiology, Biology*. 54:126 – 30. [https://doi.org/10.1016/S1011-1344\(00\)00006-3](https://doi.org/10.1016/S1011-1344(00)00006-3)

Payne, R. (1993). *Genstat 5 release 3 reference manuals*. Oxford University Press. ISBN-13: 978-0198523123

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>

Rahman, M., Tareq, T., Sarker, P., Rashid, E., Yasmeen, R., Ali, M. & Shimono, H. (2019). Genetic variation of phenotypic plasticity in Bangladesh rice germplasm. *Field Crops Research*, 243. <https://doi.org/10.1016/j.fcr.2019.107618>

Rodríguez, J., Sahagún, J., Villaseñor, H., Molina, J. & Martínez, Á. (2005). Interaction genotype x environment in the characterization of rainfed areas for wheat. *Agrociencia*. https://www.researchgate.net/publication/238706096_Interaction_genotypeXenvironment_in_the_characterization_of_rainfed_areas_for_wheat

Rueda, J. & Cotes, J. (2009). Evaluación de dos métodos de estabilidad fenotípica a través de validación cruzada. *Revista Facultad Nacional de Agronomía-Medellín*, 62(2), 5111-5123. <https://repositorio.unal.edu.co/handle/unal/37088>

Sánchez, Á., Sadras, V. & Sánchez, D. (2021). Phenotypic plasticity in relation to inter-cultivar variation of garlic (*Allium sativum* L.) functional performance and yield-stability in response to water availability. *Scientia Horticulturae*, 285, 110128. <https://doi.org/10.1016/j.scienta.2021.110128>

Sadras, V., Reynolds, M., Vega, A., Petrie, P. & Robinson, R. (2009). Phenotypic plasticity of yield and phenology in wheat, sunflower and grapevine. *F. Crop. Res.* 110, 242–250. <https://doi.org/10.1016/j.fcr.2008.09.004>

Sadras, V., Lake, L., Leonforte, A., McMurray, L. & Paull, J. (2013). Screening field pea for adaptation to water and heat stress: associations between yield, crop growth rate and seed abortion. *F. Crop. Res.* 150, 63–73. <https://doi.org/10.1016/j.fcr.2013.05.023>

Satriani, A., Loperte, A. & Soldovieri, F. (2015). Integrated geophysical techniques for sustainable management of water resource. A case study of local dry bean versus commercial common bean cultivars. *Agric. Water Manag.* 162, 57–66. <http://dx.doi.org/10.1016/j.agwat.2015.08.010>

Schippers, P. (1976). The relationship between specific gravity and percentage dry matter in potato tubers. *Am Potato J* 53:111–122. <https://doi.org/10.1007/BF02854115>

Sellam, V. & Poovammal, E. (2016). Prediction of crop yield using regression analysis. *Indian Journal of Science and Technology*, 9(38), 1-5. <https://dx.doi.org/10.17485/ijst/2016/v9i38/91714>

Sinclair, T. & Horie, T. (1989). Leaf nitrogen, photosynthesis, and crop radiation use efficiency: a review. *Crop. Sci.* 29, 90-98. <https://doi.org/10.2135/cropsci1989.0011183X002900010023x>

A. Anexo: Código para la importación y preparación de bases de datos de respuesta fenotípica, información edafoclimática de cada localidad y fechas de siembra y cosecha

```
rm(list=ls(all=TRUE))
library(readxl)
#Fijando directorio de trabajo####
path <- "G:/Mi unidad/THARVERTH"
#path<-"/Datos/Insync/THARVERTH"

#setwd("C:/Users/toshiba/Desktop/MAESTRIA/MATERIAS/TESIS/SCRIPT")
setwd(paste0(path, "/BD"))

###Importando datos
PEA <- as.data.frame(read_excel("PEA_PAPA.xlsx"))
Date_PEA<-as.data.frame(read_excel("PEA_PAPA.xlsx",sheet=2))
Suelos <- read_excel("Suelos_pre.xlsx")
LaiBD<-read_excel("IAF.xlsx")

#Transformación de data frame a Listados####
PEA_list<-split(PEA,paste(PEA[["Nlocalidad"]],PEA[["Semestre"]],sep="-"))
Date_PEA_list<-
split(Date_PEA,paste(Date_PEA[["Nlocalidad"]],Date_PEA[["Semestre"]],sep="-"))
Suelos_list <- split(Suelos, paste(Suelos[["Nlocalidad"]],Suelos[["Semestre"]],sep="-"))

#Covariables climáticas para cada Localidad en Las fechas de siembra y
cosecha####
Nlocalidad<-excel_sheets("CONSOLIDADOCLIMA.xlsx")
clima<-lapply(Nlocalidad, function (X){
  tmp<-read_excel("CONSOLIDADOCLIMA.xlsx", sheet = X,na = "-99")
  tmp<-cbind(Nlocalidad=X,tmp)
  tmp$Fecha<-as.Date(tmp$Fecha)
  for (i in c("VV", "HR", "TPROM", "RAD", "TMAX", "TMIN")){
    tmp1<-tmp[[i]]
```

```

tmp2<-approx(x=1:length(tmp1), tmp1, n=length(tmp1))
tmp[is.na(tmp1),i]<-tmp2[["y"]][is.na(tmp1)]
}
return(tmp)
})
names(clima)<-Nlocalidad

NLoc_Clima<-unlist(lapply(PEA_list, function(X) X[["Nlocalidad"]][1]))
li<-do.call(c,lapply(Date_PEA_list, function(X) as.Date(X[["Siembra"]]))))
ls<-do.call(c,lapply(Date_PEA_list, function(X) as.Date(X[["Cosecha"]]))))

clima<-lapply(1:length(Date_PEA_list), function(X){
  tmp0<-clima[[NLoc_Clima[X]]]
  tmp0$Fecha<-as.Date(tmp0$Fecha)
  tmp0$Semestre<-as.numeric(unlist(strsplit(names(NLoc_Clima[X]), "-"))[2])
  out<-subset(tmp0,Fecha>=li[X] & Fecha<=ls[X])
  return(out)
})
names(clima) <- names(Date_PEA_list)

#Adhesión de información de suelos a La información climática####
clima_suelos<-list()
for(i in 1:length(clima)) clima_suelos[[i]]<-
merge(clima[[i]],Suelos_list[[i]],by=c("Nlocalidad","Semestre"))
names(clima_suelos)<-names(clima)
#test<-lapply(names(Suelos_list), function(i)
merge(clima[[i]],Suelos_list[[i]],by=c("Nlocalidad","Semestre")))
#test<-mapply(merge,clima,Suelos_list,by=c("Nlocalidad","Semestre"))

#Estimación del índice de área foliar teórico####
LaiM<-aggregate(mean.ilAF~iDDE,data=LaiBD,mean)
colnames(LaiM)<-c("DDE","Lai")
DDSmax<-nrow(clima[[1]])
DDEmax<-nrow(LaiM)
Lai<-c(rep(0,DDSmax-DDEmax),LaiM$Lai)
DDE<-c(rep(0,DDSmax-DDEmax),LaiM$DDE)
clima<-lapply(clima, function(X) cbind(X,DDS=1:DDSmax,DDE=DDE,Lai=Lai))

#Salida de datos climáticos para cada una de Las Localidades####
save(clima,
li,ls,Date_PEA,Date_PEA_list,PEA,PEA_list,Suelos,Suelos_list,clima_suelos,file="BD.RData")

```

B.Anexo: Código para la preparación de la información ecofisiológica de cada una de las localidades evaluadas

```
rm(list=ls(all=TRUE))
path <- "G:/Mi unidad/THARVERTH"
#path <- "/Datos/Insync/THARVERTH"
#setwd("C:/Users/toshiba/Desktop/MAESTRIA/MATERIAS/TESIS/SCRIPT")
setwd(paste0(path, "/BD"))
load("BD.RData")
NExp <- names(clima)

#Referencias para elaboración del algoritmo de Balance Hídrico (BH)####

#Ref.0 <- Model [or Predicting Evaporation Row Crop with Incomplete
Cover,
#Ritchie, 1972

#Ref.1 <- Evapotranspiracion del cultivo Guías para La determinacion de
Los
#requerimientos de agua de Los cultivos, FAO, 2006
#URL: http://www.fao.org/3/x0490s/x0490s.pdf

#Ref.2 <- https://biblioteca.cenicafe.org/handle/10778/859

#Transformación de formato fechas a días joulianos####
date_to_nday <- function(x){
  #x es una fecha en formato Date

  tmp <- as.numeric(format(x,"%Y"))
  tmp <- paste("01","01",tmp,sep = "/")
  tmp <- as.Date(tmp,"%d/%m/%Y")
  tmp <- x-tmp+1
  return(as.numeric(tmp))
}
#test <- date_to_nday(x)
#x <- as.Date("2020-10-06", format = "%Y-%d-%m")
#str(x)

#Estimacion De La Radiación Solar Día Despejado####
```

```

SRDso_f <- function(Latitud,Altitud,Fecha){

#SRDso <- Radiación solar día despejado (Mj/m2.día)
#Altitud <- Altitud sobre el nivel del mar de La Localidad (m)
#Latitud <- Latitud de La Localidad (°Dec)
#Fecha <- Fecha en formato as.Date()

#Ra <- Radiación astronomica (Mj/m2.día)
#Gsc <- Constante Radiación solar (Mj/m2.min)
#phi <- Latitud (rad)
#dr <- Distancia relativa inversa tierra-sol
#delta <- Declinacion solar (Rad)
#omega <- Angulo de Radiación a La puesta del sol (Rad)
#J <- Numero del día en el ano siendo 01 de enero el día 1

#Verificar función con Pag. 217 Ref. 1
Gsc <- 0.082 #Constante Pag. 21 Ref. 1
J <- date_to_nday(Fecha)
phi <- (pi/180)*Latitud #Conversion de grados decimales a radianes Eq. 22
Ref. 1
dr <- 1+0.033*cos((2*pi*J)/365) #Eq. 23 Ref. 1 (adimensional)
delta <- 0.409*sin(2*pi*J/365-1.39) #Eq. 24 Ref. 1 (Rad)
omega <- acos(-tan(phi * tan(delta))) #Eq. 25 Ref. 1 (Rad)
Ra <- ((24*60)/pi)*Gsc*dr*(omega*sin(phi)*sin(delta)+cos(phi)*cos(delta)*sin(omega)) #Eq.
21 Ref. 1 (MJ/m2.día)
SRDso <- (0.75+2E-5*Altitud)*Ra #Eq. 37 Ref. 1 (MJ/m2.día)
return(SRDso)
}

test <- SRDso_f(Latitud = Suelos_list["Mosquera-1"]$Latitud,Altitud =
Suelos_list["Mosquera-1"]$Altitud, Fecha = clima["Mosquera-1"]$Fecha)
#test <- SRDso(Latitud = Ubicacion$Latitud[[1]],Altitud =
Ubicacion$altitud[[1]],Fecha = as.Date(CAM1_c$Fecha, format =
"%d/%m/%Y"))
#test <- SRDso(Latitud = -20,Altitud = 0,Fecha = as.Date("03/09/2013",
format = "%d/%m/%Y"))

#Estimación De La Constante Psicrometrica

gamma_f <- function(Altitud,Tm){
#gamma <- Constante psicrometrica (kPa/°C)
#Altitud <- Altitud sobre el nivel del mar (m)
#Tm <- Temperatura media (°C)

#Cp <- Calor especifico del aire (kJ/kg.°C)
#patm <- Presion atmosferica (kPa)
#epsilon <- Relación del peso molecular del vapor de agua/aire seco
#lambda <- Calor latente de vaporización (MJ/kg)

Cp <- 1.013 #Constante pag. 223 Ref. 1 (kJ/kg.°C)
Cp <- Cp/1000 #Conversion kJ a MJ

```

```

patm <- 101.3 * ((293-(0.0065*Altitud))/293)^5.26 #Eq. 7 Ref. 1 (kPa)
epsilon <- 0.622 #Constante pag. 223 Ref. 1 (adimensional)
lambda <- 2.501-(2.361E-3*Tm) #Eq. 3-1 pag. 221 Ref. 1 (MJ/kg)
gamma<-(Cp*patm)/(epsilon*lambda) #ecuacion psicrometrica Eq.3-10 Pag. 223
Ref. 1 (kPa/°C) (pasar a mBar???)
return(gamma)
}
test <- gamma_f(Altitud = Suelos_list[["Mosquera-1"]]$Altitud,Tm = clima[["Mosquera-
1"]]$TPROM)
#test <- gamma_f(2000,18)

#Estimación de La radiación solar neta####

SRDn_f<- function(Albedo, SRD, Th, Tl, Tp, SRDso, RHD){
  #SRDn <- Radiación solar neta (Mj/m2.día)
  #Albedo <- coeficiente de reflexión del cultivo, que es 0,23
  # para el cultivo hipot??tico de referencia
  #SRD <- Radiación solar diaria en Mj/m2.días
  #Th <- Temperatura máxima diaria (°C)
  #Tl <- Temperatura mínima diaria (°C)
  #SRDso <- Radiación solar del día despejado (MJ/m2.día)
  #RHD <- Hudemad relativa diaria (%)

  #SRDnl <- radiación neta de onda larga (Mj/m2.día)
  #SRDns <- radiación neta solar o de onda corta (Mj/m2.día)
  #Tm <- Temperatura media (°C)
  #e0 <- presión de saturación de vapor a La temperatura del aire (kPa)
  #eA <- presión de vapor real media (kPa)

  e0<- 0.6108 * exp(17.27*(Tp)/(Tp + 237.3)) #Eq. 11 Ref. 1 (kPa)
  delta<- (4098*e0)/(Tp+237.3)^2 #pendiente curva de presion de saturacion de
vapor Eq. 3-9 Pag. 223 Ref. 1 (kPa/°C)
  #e0<- e0*10 #Conversion de kPa a mBar
  eA<- (RHD/100) * e0 #Eq. 19 Ref. 1 (kPa)
  SRDnl <- 4.903E-9 * ((Th+273.16)^4+(Tl+273.16)^4)*0.5*(0.34-
0.14*sqrt(eA))*(1.35*sapply(SRD/SRDso,function(x) min(x,1))-0.35) #Eq. 39 Ref. 1
(MJ/m2.día)
  SRDnl<- sapply(SRDnl,function(x) max(x,0.1))
  SRDns <-(1-Albedo)*SRD #Eq. 38 Ref. 1. (MJ/m2.día)
  SRDn <- SRDns - SRDnl #Eq. 40 Ref. 1

  #Devolver e0, eA y delta
  return(list(SRDn = SRDn,e0 = e0,eA = eA,delta = delta))
}

#Testeo de función
test <- SRDso_f(Latitud = Suelos_list[["Mosquera-1"]]$Latitud,Altitud =
Suelos_list[["Mosquera-1"]]$Altitud, Fecha = clima[["Mosquera-1"]]$Fecha)
SRDn_test <- SRDn_f(Albedo = 0.23, SRD = clima[["Mosquera-1"]]$RAD, Th =
clima[["Mosquera-1"]]$TMAX, Tl = clima[["Mosquera-1"]]$TMIN, Tp = clima[["Mosquera-

```

```

1"]}$TPROM, SRDso = test, RHD = clima[["Mosquera-1"]]$HR)
#SDR_test<- SRDn(Albedo =0.23, SRD=CAM1_c$RAD, Th = CAM1_c$TMAX, TL =
CAM1_c$TMIN,SRDso = test, RHD = CAM1_c$HR)

#Algoritmo de Balance Hídrico de Ritchie####
ritchie<-function(dds,Fecha,Altitud,Latitud,Tp,Th,Tl,SRD,RHD,wind,P,U,alpha,lai,Es1,Es2,t){

#dds <- días despues de siembra (entero)
#Fecha <- Fecha en formato As.Date()
#Altitud <- Altitud sobre el nivel del mar (m)
#Latitud <- Latitud de la Localidad (°Dec)
#Tp <- Temperatura promedio (°C)
#Th <- Temperatura máxima (°C)
#Tl <- Temperatura mínima (°C)
#SRD <- Radiación solar diaria (Mj/m2.día)
#RHD <- Humedad relativa media diaria (%)
#Wind <- Velocidad del viento en (m/s)
#P <- Precipitación diaria (mm/día)
#U <- Límite superior acumulativo de evaporacion (mm/día)
#alpha <- Coeficiente calculado del suelo (mm/sqrt(día))
#lai <- Índice de area foliar (adimensional)
#Es1 <- Valor null para la evaporación en el estado 1 de secado (mm)
#Es2 <- Valor null para la evaporación en el estado 2 de secado (mm)
#t <- días de sequia (día)

#Rs <- Radiación solar en evaporacion equivalente (mm/día)
###calculada
#Albedo = Albedo
#SRDso

if(is.null(Tp)) Tp<- (Th+Tl)/2#(°C)
gamma<-gamma_f(Altitud = Altitud,Tm = Tp)
SRDso<-SRDso_f(Latitud = Latitud, Altitud = Altitud, Fecha = Fecha)
Albedo<-0.14+0.25*(0.23-0.14)*lai #Eq2 Ref. 0
#Tabla 5.3 page 56 Ref. 2. Albedo suelo negro
SRDn<-SRDn_f(Albedo = Albedo, SRD = SRD, Th = Th, Tl = Tl, Tp= Tp, SRDso = SRDso, RHD)

###Conversiones de unidades para la función de Ritchie
Rno <- SRDn[["SRDn"]]*23.9006/59 #Conversion Mj/m2.día a Cal/cm2.día Cuadro
3 Ref. 1
#Conversion Cal/cm2.día a mm/día #Pag. 9 Ref. 0
Wind<-(wind/1000)*3600*24 #Conversion de unidades m/s a km/día
gamma<- gamma*10 #Conversion de kPa/°C a mBar/°C
delta<-SRDn[["delta"]]*10 #Conversion de kPa/°C a mBar/°C
e0<-SRDn[["e0"]]*10 #Conversion de kPa a mBar
eA<-SRDn[["eA"]]*10 #Conversion de kPa a mBar
Eo<-((delta/gamma)*Rno+0.262*(1+0.0061*Wind)*(e0-eA))/((delta/gamma)+1)#Eq. [1]
(kPa) (mBar???)
Eso<-(delta/(delta+gamma))*Rno*exp(-0.398*lai)#Eq. 5 Ref. 0

```



```

nodo1<-expression({
  Es1<-Es1+Eso
  if (Es1>U){
    Es<-Eso-0.4*(Es1-U)
    Es2<-0.6*(Es1-U)
    t<-(Es2/alpha)^2 #Eq. 7 ref. 0
  } else {
    Es<-Eso
  }
})

nodo2<-expression({
  Es2<-Es2+Es-P
  t<-(Es2/alpha)^2 #Eq. 7 ref. 0
})

#if (dds==1) {
  #Valores iniciales
  # Es1<-Es2<-0
  # t<-1
#}
if (Es1<U){
  Es1<-max(0,Es1-P)
  eval(nodo1)
} else {#Es1<U
  if(P>=Es2){
    P<-P-Es2
    Es1<-max(0,U-P)
    eval(nodo1)
  } else{#P>=Es2
    t<-t+1
    Es<-alpha*sqrt(t)-alpha*sqrt(t-1) #Eq. 8 Ref. 0
    if (P>0){
      Esx<-0.8*P
      if (Esx<=Es) Esx<-Es+P else {
        if (Esx>Eso) Esx<-Eso else Es<-Esx
        eval(nodo2)
      }
    } else {#P>0
      Es<-max(Es,Eso)
      eval(nodo2)
    }
  }
}

#Lai condicionales
if (lai>=0.1){
  if (lai>=2.7){
    Ep<-Eo-Es
  } else {
    Ep<-Eo*(-0.21+0.7*sqrt(lai)) #Eq. 9 Ref 0
  }
}

```

```

    Ep<-min(Ep,Eo-Es)
  }
} else Ep<-0
E<-Es+Ep

out<-data.frame(DDS=dds,Fecha=Fecha,Albedo=Albedo,SRDso=SRDso
,Rno=Rno,gamma=gamma,Wind=Wind,delta=delta,e0=e0,eA=eA,Eo=Eo,Eso=Eso,t=t,Es1=Es1,Es
2=Es2,Es=Es,Ep=Ep,E=E)

return(out)
}

```

#Ritchie por Localidad####

#X<-NExp[1] #Nombre de la Localidad

PorEnsayo<-function(X){

#ADT <- Agua disponible total en la zona radical (mm)
#CAS <- Capacidad de almacenamiento de agua en el suelo (mm)
#AH20 <- Almacenamiento de agua en el suelo (m3/m3)
#DH20 <- Agua disponible (mm)
#ETP <- Evapotranspiracion potencial del cultivo (mm/dia)
#p <- Factor de agotamiento de agua en el suelo [0-1]
#AFA <- Agua facilmente aprovechable (mm)
#BH <- Balance hídrico
#0.35 <- Profundidad radical (m)

Evap<-resultados<-BH<-list()

for (dds in 1:nrow(clima[[X]])){

CAS<-1000*Suelos_list[[X]][["AH20"]]*0.35 #Eq.82 Ref1 (Cuadro 19 y Cuadro 22 Ref.1)

if (dds==1){

Evap[[dds]]<-ritchie(dds=dds,
Fecha=clima[[X]]\$Fecha[dds],
Altitud=Suelos_list[[X]]\$Altitud,
Latitud=Suelos_list[[X]]\$Latitud,
Tp=clima[[X]]\$TPROM[dds],
Th=clima[[X]]\$TMAX[dds],
Tl=clima[[X]]\$TMIN[dds],
SRD=clima[[X]]\$RAD[dds],
RHD=clima[[X]]\$HR[dds],
wind=clima[[X]][["VV"]][dds],
P=clima[[X]][["PP"]][dds],
U=Suelos_list[[X]][["U"]],
alpha=Suelos_list[[X]][["alfa"]],
lai=clima[[X]][["Lai"]][dds],
Es1=0,
Es2=0,
t=1)

ADT<-CAS #Se asume que para el primer día el cultivo está a CC.

```

} else {
  Evap[[dds]]<-ritchie(dds=dds,
    Fecha=clima[[X]]$Fecha[dds],
    Altitud=Suelos_list[[X]]$Altitud,
    Latitud=Suelos_list[[X]]$Latitud,
    Tp=clima[[X]]$TPROM[dds],
    Th=clima[[X]]$TMAX[dds],
    Tl=clima[[X]]$TMIN[dds],
    SRD=clima[[X]]$RAD[dds],
    RHD=clima[[X]]$HR[dds],
    wind=clima[[X]][["VV"]][dds],
    P=clima[[X]][["PP"]][dds],
    U=Suelos_list[[X]][["U"]],
    alpha=Suelos_list[[X]][["alfa"]],
    lai=clima[[X]][["Lai"]][dds],
    Es1=Evap[[dds-1]][["Es1"]],
    Es2=Evap[[dds-1]][["Es2"]],
    t=Evap[[dds-1]][["t"]])
  ADT<-BH[[dds-1]][["DH20"]]
}
ETP<-Evap[[dds]][["E"]]
P<-clima[[X]][["PP"]][dds]
DH2O<-max(min(P+ADT,CAS)-ETP,0)
p<-min(1,0.4+0.04*(5-ETP)) #Cuadro 22. Ref1.
BH[[dds]]<-data.frame(ADT=ADT,ETP=ETP,P=P,DH2O=DH2O,p=p,AFA=p*ADT)
}
BH<-do.call(rbind,BH)
Evap<-do.call(rbind,Evap)
resumen<-cbind(Evap,BH)
}

#Balance hídrico estimado####
BH<-lapply(NExp,PorEnsayo)
names(BH)<-NExp
DSAD<-lapply(BH, function (X) sum(X[["AFA"]]==0)) #Días sin agua disponible
CAP<-lapply(BH, function (X) sum(X[["Ep"]])) #Consumo de agua por plantas
CAC<-lapply(BH, function (X) sum(X[["E"]])) #Consumo de agua del cultivo
NPS<-lapply(BH, function (X) sum(diff(X[["AFA"]]==0)==1)) #Numero de periodos de sequia
PARAi<-lapply(clima, function (X) sum(X[["RAD"]]*0.44*(1-exp(-0.65*X[["Lai"]])))) #PAR incidente acumulada
RADA<-lapply(clima, function (X) sum(X[["RAD"]])) #Radiacion solar acumulada

#Matriz con covariables Ecofisiológicas y ambientales####
Ecofisiologia<-as.data.frame(cbind(DSAD=unlist(DSAD),CAP=unlist(CAP),CAC=unlist(CAC),
  ,NPS=unlist(NPS),PARAi=unlist(PARAi), RADA=unlist(RADA)))
climamean<-do.call(rbind,lapply(clima, function (X)
  colMeans(X[c("HR","TPROM","TMIN","TMAX","RAD","PP"]))))
Environmental<-cbind(Ecofisiologia,climamean,Suelos)
#Environmental<-within(Environmental,{

```

```
# NaP<-Na/CICE
# CaP<-Ca/CICE
# MgP<-Mg/CICE
# KP<-K/CICE
# AIP<-AI/CICE
# MO<-(1/0.58)*CO
# CaMgK<-(Ca+Mg)/K
#})
```

```
#Ambiente de R con información climática####
save(Ecofisiologia, Environmental, file = "clima.RData")
```

C.Anexo: Código para la verificación de interacción genotipo - ambiente por cambio de rango (COI) o heterogeneidad de varianzas (HV)

```
rm(list=ls(all=TRUE))
path <- "G:/Mi unidad/THARVERTH"
#path<-"/Datos/Insync/THARVERTH"
#setwd("C:/Users/toshiba/Desktop/MAESTRIA/MATERIAS/TESIS/SCRIPT")
setwd(paste0(path,"/BD"))
load("BD.RData")
load("clima.RData")
NExp<-names(clima)

#Vector de medias para el rasgo RT por genotipo y Localidad####

meanRT<-do.call(rbind,lapply(PEA_list, function (X)
tapply(X[["RendimientoTotal"]],X[["Genotipo"]],mean))) #Media RT por genotipo,
Localidad y semestre
#meanRC<-
tapply(PEA[, "Rcomercial"],list(paste(PEA[, "NLocalidad"],PEA[, "Semestre"]
,sep="-"),PEA[, "Genotipo"]),mean) #Media RT por genotipo, Localidad y
semestre
#meanRT[2,3]<-NA

#Algoritmo para verificación de COI o HV####

COI_list<-list()
k<-0
for (i in 1:21){
  for (j in (i+1):22){
    k<-k+1
    tmp1<-sign(unlist(sapply(1:9,function (X) diff(meanRT[i,],X))))
    tmp2<-sign(unlist(sapply(1:9,function (X) diff(meanRT[j,],X))))

    COI<-tcrossprod(tmp1,tmp2)
    COI_list[[k]]<-COI[upper.tri(COI,diag=TRUE)]
  }
}
```

}

#Transformar de lista a DF

COI<-**unlist**(COI_list)

#Porcentaje de COI####

pCOI<-**sum**(COI==**-1**,**na.rm**=**TRUE**)/**sum**(**!is.na**(COI))

D.Anexo: Código para la ejecución de correlaciones, análisis de componentes principales, regresiones simples y múltiples

```
rm(list=ls(all=TRUE))
library(readxl)
library(WriteXLS)
#Directorio de trabajo
path <- "G:/Mi unidad/THARVERTH"
setwd(paste0(path,"/BD"))

#Carga de informacion fenotipica y climatica
load("BD.RData")
load("clima.RData")
NExp<-names(clima)
Environmental$ArL <- Environmental[, "Ar"] + Environmental[, "L"]

###Estimacion de medias  $Y_i \sim N_{Localidad} + Semestre$ 
mediasPEA<-
aggregate(cbind(RendimientoTotal,AR,CF,MS,GE)~Nlocalidad+Semestre,mean,na.rm=TRUE,data=P
EA)

#Insercion de columna de empalme
mediasPEA$ENV<-paste(mediasPEA$Nlocalidad,mediasPEA$Semestre,sep="-")
Environmental$ENV<-paste(Environmental$Nlocalidad,Environmental$Semestre,sep="-")

###Union de matrices - Base de datos general
Datos<-merge(mediasPEA,Environmental,by=c("ENV"))

###Matriz de respuesta fenot?pica
respuesta<-as.list(Datos[,c("RendimientoTotal","AR","CF","MS","GE")])

###Matriz de covariables ambientales
covariables<-as.list(Datos[c(9:20,31,32,35:41,43,45:53)])

#Validación de Los modelos simples####
tmp<-lapply(covariables, function (X){
  out<-do.call(cbind,lapply(respuesta, function (Y) {
```

```

P<-sapply(1:length(Y), function (i){
  m<-lm(Y[-i]~X[-i])
  pred<-coefficients(m)%*%c(1,X[i])
})
RMSPD<-sqrt(sum((Y-P)^2))/length(Y)
Correlation<-cor(Y,P)
R2<-summary(lm(Y~X))["r.squared"]
return(c(RMSPD=RMSPD,CoorL00=Correlation,R2=R2))
}))
return(out)
#out<-cbind(Variable=names(respuesta_list)[i]
#summary(lm(Y~X))["r.squared"]))
})

#Estimacion de criterios de validacion####
RMSPD<-do.call(rbind,lapply(tmp,function (X) X[1,]))
CorrL00<-do.call(rbind,lapply(tmp,function (X) X[2,]))
R2<-do.call(rbind,lapply(tmp,function (X) X[3,]))

#Creacion de matriz de resultados - criterios de validacion
#write.csv(List(RMSPD=as.data.frame(RMSPD),
#CorrL00=as.data.frame(CorrL00),
#R2=as.data.frame(R2)),
#"AjustesSimples.csv",row.name=TRUE)

#Ajuste de R2, correlacion y RMSPD###
for (i in 1:length(tmp)){
  tmp[[i]]<-cbind(Variable=names(tmp)[i],
  Parameter=rownames(tmp[[i]]),
  tmp[[i]])
  rownames(tmp[[i]])<-NULL
  tmp[[i]]<-as.data.frame(tmp[[i]])
}
todo<-do.call(rbind,tmp)

#Modelo sin genotipos####
modSinGen<-function(variables){
  tmp<-do.call(cbind,lapply(respuesta, function (Y) {
  depend<-as.matrix(Datos[,variables])
  R2<-summary(lm(Y~depend))["r.squared"])
  P<-sapply(1:length(Y), function (i){
  m<-lm(Y[-i]~depend[-i,])
  pred<-coefficients(m)%*%c(1,depend[i,])
  })
  RMSPD<-sqrt(sum((Y-P)^2))/length(Y)
  Correlation<-cor(Y,P)
  return(c(RMSPD,Correlation,R2))
}))
rownames(tmp)<-cbind("RMSPD","Correlation","R2")

```



```

return(tmp)
}

#Modelos de regresion lineal fenotipo ~ ambiente####
modelo1<-modSinGen(c("CICE","CAC","PARAi","pH"))
modelo2<-modSinGen(c("CICE","PP","RADA","pH"))
modelo3<-modSinGen(c("CICE","PP","Altitud","pH"))

#Creacion de matriz con comparacion de modelos####
write.csv(list(M1=as.data.frame(modelo1),
              M2=as.data.frame(modelo2),
              M3=as.data.frame(modelo3)),
          "AjustesModeloSinGenotipo.csv",row.name=TRUE)

#Macro seccion del detalle - Medias de cultivares####
mediasPEAGEN<-
aggregate(cbind(RendimientoTotal,AR,CF,MS,GE)~Nlocalidad+Semestre+Genotipo,mean,na.rm=TRUE,
          data=PEA)
mediasPEAGEN$ENV<-paste(mediasPEA$Nlocalidad,mediasPEA$Semestre,sep="-")
mediasPEAGEN$Genotipo <- as.factor(mediasPEAGEN$Genotipo)
GEN <- split(mediasPEAGEN, mediasPEAGEN$Genotipo)

#Creacion de listado de covariables ambientales por genotipo####
Datos_list <- list()
for (i in 1:length(GEN))Datos_list[[i]]<- merge(GEN[[i]],Environmental, by = c("ENV"))
names(Datos_list)<- names(GEN)

#Creacion de listado de variables respuesta por genotipo####
respuesta_list <- list()
for (i in 1:length(Datos_list)) respuesta_list[[i]]<-
as.list(Datos_list[[i]][c("RendimientoTotal","AR","CF","MS","GE")])
names(respuesta_list)<- names(Datos_list)

#Creacion de listado de matriz de variables y covariables####
covariables_list <- list()
for (i in 1:length(respuesta_list)) covariables_list[[i]]<-
as.list(Datos_list[[i]][c(10:20,31,32,35:41,43,45:53)])
names(covariables_list)<- names(Datos_list)

#modelos de regresion lineal simple por genotipo####
modgen<-function(variables){
tmp3_list<- list()
for (i in 1:length(respuesta_list)) {
#depend<-
as.matrix(Datos_list[[i]][,"Fe"]+Datos_list[[i]][,"CAC"]+Datos_list[[i]]
[, "PARAi"]+Datos_list[[i]][,"pH"])
depend<-as.matrix(Datos_list[[i]][,variables])
tmp3_list[[i]]<-do.call(cbind,lapply(respuesta_list[[i]],

```

```

function (Y) {
  R2<-summary(lm(Y~depend))["r.squared"]
  P<-sapply(1:length(Y), function (i){
    m<-lm(Y[-i]~depend[-i,])
    pred<-coefficients(m)%*%c(1,depend[i,])
  })
  RMSPD<-sqrt(sum((Y-P)^2)/length(Y))
  #Correlation<-cor(Y,P)
  return(c(RMSPD=RMSPD,
    #CoorL00=Correlation,
    R2=R2))
  })
tmp3_list[[i]]<-cbind(Genotype=names(respuesta_list)[i],
  Parameter=rownames(tmp3_list[[i]]),
  tmp3_list[[i]])
rownames(tmp3_list)<-NULL
}
#return(tmp3_list)
#names(tmp3_list)<-names(Datos_List[[i]]_list)
tmp3<- as.data.frame(do.call(rbind,tmp3_list))
return(tmp3)
#tmp3<- apply(tmp3, 2, as.character)
#rownames(tmp3)<-names(tmp_list)
#rownames(tmp)<-c("RMSPD", "Correlation", "R2")
}
#Modelos de regression multiple fenotipo ~ genotipo + ambiente####
modelo1<-modgen(c("CICE", "CAC", "PARAi", "pH"))
modelo2<-modgen(c("CICE", "PP", "RADA", "pH"))
modelo3<-modgen(c("CICE", "PP", "Altitud", "pH"))

write.csv(list(M1=as.data.frame(modelo1),
  M2=as.data.frame(modelo2),
  M3=as.data.frame(modelo3)),
  "AjustesModeloPorGenotipo.csv")

#Análisis De Componentes Principales####
cov_modelo <- as.data.frame(covariables) #Transformacion a DF
mcor <-
cor(cov_modelo[,c("pH", "N", "Ca", "K", "Mg", "Na", "AI", "CICE", "P", "Cu", "Fe", "Mn", "Zn", "B", "Ar", "L", "A", "ArL")]) #matriz de correlaciones
#cor(as.data.frame(covariables))

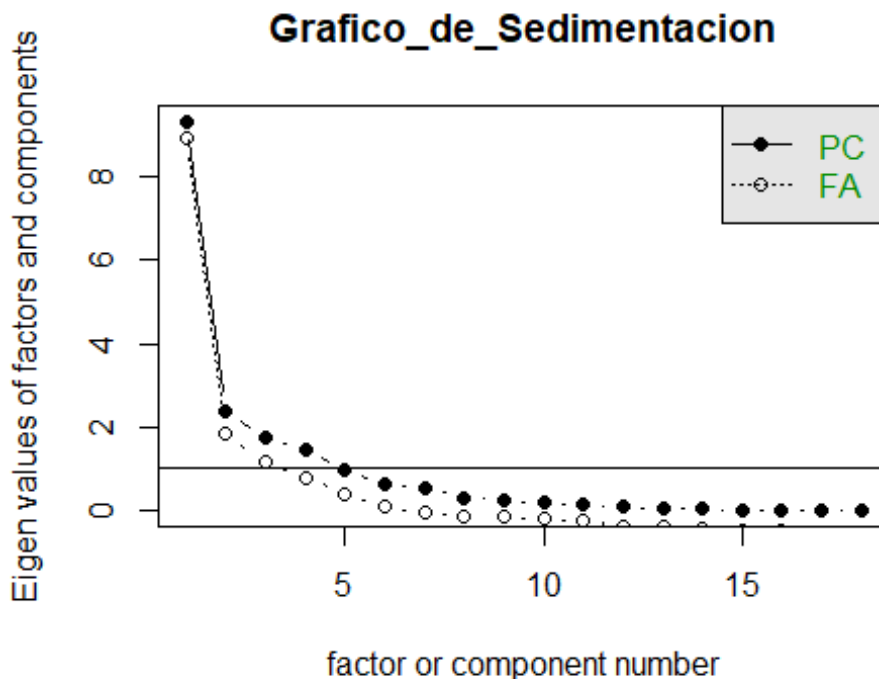
#Grafico de correlaciones
#library(corrplot)
#corrplot(mcor)
#library(PerformanceAnalytics)
#chart.Correlation(mcor, histogram = F) #Grafico de correlaciones y
significancia

```

#Análisis de factibilidad del ACP`library(psych)``KMO(mcor) #Kaiser-Meyer Olkin Factor Test``## Kaiser-Meyer-Olkin factor adequacy``## Call: KMO(r = mcor)``## Overall MSA = 0.5``## MSA for each item =`

	pH	N	Ca	K	Mg	Na	Al	CICE	P	Cu	Fe	Mn	Zn	B	ArL
##	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

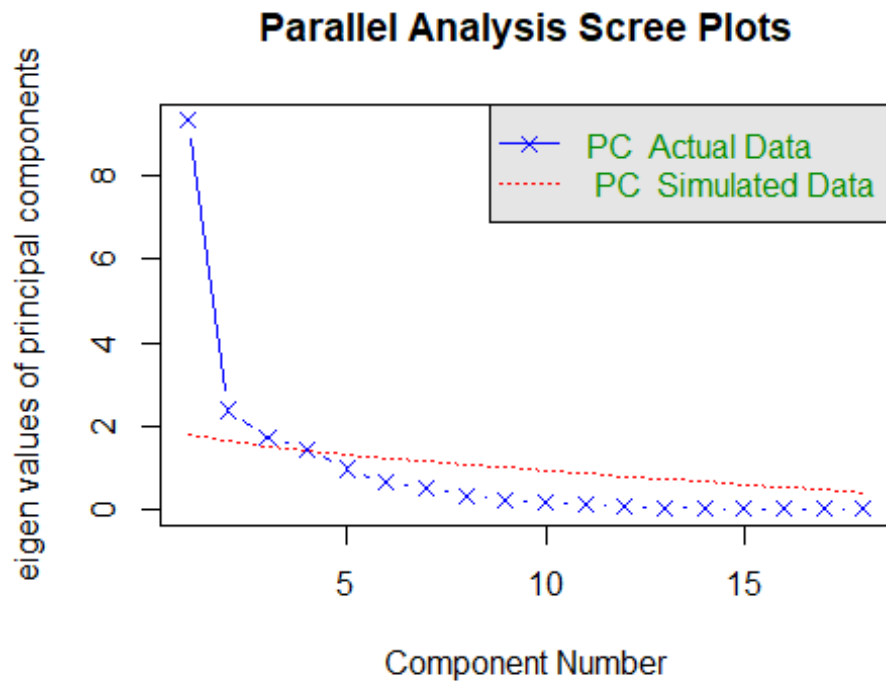
`## A ArL``## 0.5 0.5``scree(mcor, main = "Grafico de Sedimentacion") #Grafico de sedimentacion``## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done``## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done``## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was done``## Warning in cor.smooth(r): Matrix was not positive definite, smoothing was done``## Warning in fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs, :``## The estimated weights for the factor scores are probably incorrect. Try a ## different factor score estimation method.``## In factor.scores, the correlation matrix is singular, an approximation is used``## Warning in cor.smooth(r): Matrix was not positive definite, smoothing was done`



```
fa.parallel(mcor,fa="pc")# Analisis paralelo simulado
## Warning in fa.parallel(mcor, fa = "pc"): It seems as if you are using a
## correlation matrix, but have not specified the number of cases. The number
## of
## subjects is arbitrarily set to be 100
## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was
done

## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was
done

## Warning in cor.smooth(R): Matrix was not positive definite, smoothing was
done
## Warning in cor.smooth(r): Matrix was not positive definite, smoothing was
done
## Warning in fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs,
:
## The estimated weights for the factor scores are probably incorrect. Try a
## different factor score estimation method.
## In factor.scores, the correlation matrix is singular, an approximation is
used
## Warning in cor.smooth(r): Matrix was not positive definite, smoothing was
done
```



```
## Parallel analysis suggests that the number of factors = NA and the number
of components = 3
```

```
#Importancia de Los Componentes Principales####
```

```
Eigen_mcor <- eigen(mcor)
```

```
Eigenvalues <- Eigen_mcor$values
```

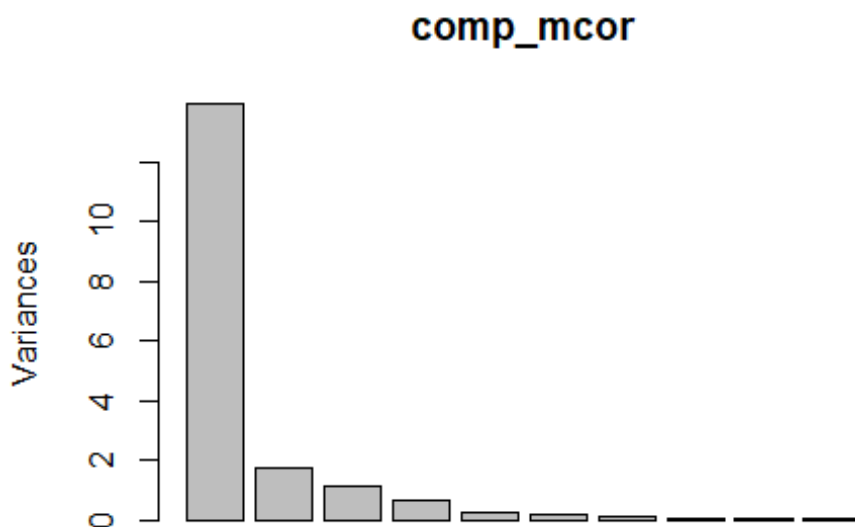
```
Eigenvector <- round(Eigen_mcor$vectors[,1:2],4)
```

```
#write.csv(x = Eigenvector, file = "EigenV.csv")
```

```
comp_mcor <- prcomp(mcor, scale. = T, center = T) #Estimacion de Las CP
```

```
a <- summary(comp_mcor)
```

```
plot(comp_mcor)
```



```
biplot(comp_mcor, scale = T)
```

