



UNIVERSIDAD NACIONAL DE COLOMBIA

# **Zonificación agroecológica para el cultivo de arroz de riego (*Oryza Sativa* L.) en Colombia**

**Camilo Barrios Pérez  
C.C. 1130945442**

Universidad Nacional de Colombia  
Facultad de Ciencias Agropecuarias,  
Escuela de Posgrados  
Palmira – Colombia  
2016

# **Zonificación agroecológica para el cultivo de arroz de riego (Oryza Sativa L.) en Colombia**

**Camilo Barrios Pérez**

Tesis de grado para optar al título de Magister en  
CIENCIAS AGRARIAS CON ÉNFASIS EN SUELOS

Director:

Ph.D., Alexandre Bryan Heinemann

Codirector:

Ph.D., Carlos Ivan Cardozo Conde

Línea de Investigación:

Modelación de Cultivos y Clima

Grupo de Investigación:

Grupo de Modelación de Cultivos y Clima

Centro Internacional de Agricultura Tropical (CIAT)

Universidad Nacional de Colombia

Facultad de Ciencias Agropecuarias,

Escuela de Posgrados

Palmira, Colombia

2016

***Dedicatoria:***

*A Dios por darme la oportunidad de crecer cada día como persona y como profesional.*

*A mis padres por todo su amor, esfuerzo y dedicación.*

*A mi esposa por su amor y compañía incondicional.*

*A mis hermanos, por ser los mejores cómplices de la vida.*

*Los Amo con todo mi corazón!*

*Camilo Barrios Pérez*

## **Agradecimientos**

Al Centro de Investigación de Agricultura Tropical (CIAT), al Programa del CGIAR sobre el Cambio Climático, la Agricultura y la Seguridad Alimentaria (CCAFS); y al programa de Toma de Decisiones y Análisis de Políticas (DAPA), por haberme dado la confianza y oportunidad de desarrollar este trabajo de investigación.

A la Universidad Nacional de Colombia por la institucionalidad, orden, y compromiso de su cuerpo docente, así como a su Escuela de Posgrados en dirección del profesor Juan Carlos Menjivar, quien estuvo prestante a suplir cualquier necesidad presentada en este proceso.

A la Federación Nacional de Arroceros (FEDEARROZ) por su apoyo y colaboración durante el desarrollo de esta investigación.

A los Doctores Julian Ramirez y Camila Rebolledo (Investigadores del CIAT), y Alezandre Bryan Heinemann (Investigador de EMBRAPA), por su amabilidad y total disponibilidad al momento de brindarme su asesoría y conocimiento, con el fin de avanzar en el desarrollo de este trabajo de grado.

Al profesor Carlos Ivan Cardoso, por todo su apoyo y colaboración durante la maestría.

A mis compañeros de trabajo, David Arango, Diego Agudelo, Jefferson Rodríguez y Lizeth Llanos, por los momentos de discusión y retroalimentación, los cuales me permitieron refinar a diario las diferentes metodologías de análisis implementadas.

Detrás de este trabajo se encuentran muchas más personas a las cuales hago extensivo mis agradecimientos.

**A todos ustedes muchas gracias y que Dios los bendiga!**

La facultad y los jurados de tesis no se harán responsables de las ideas emitidas por el autor.

Artículo 24, resolución 04 de 19



## Resumen

El presente estudio se llevó a cabo con el objetivo principal de realizar una zonificación y caracterización ambiental del sistema de arroz de riego en los departamentos de Tolima, Huila y Córdoba en Colombia, aplicando un enfoque de simulación espacial a través del modelo de cultivo ORYZA2000. Además, se buscó identificar el estrés abiótico más relevante que limita el rendimiento del cultivo en la región de estudio.

Se evaluaron cuatro variedades de arroz (Fedearroz 2000, Fedearroz 60, Fedearroz 733 y CT21375) en 38 localidades con información climática histórica entre 1984 a 2012, se consideraron 18 fechas de siembra y seis diferentes tipos de suelos, en las regiones Centro y Bajo Cauca de Colombia. Sobre los rendimientos simulados, se implementó un método de agrupamiento jerárquico aglomerativo, con el fin de identificar los grupos ambientales. Para determinar el número óptimo de grupos se utilizaron indicadores de eficiencia y estabilidad de agrupamiento. Posteriormente, se produjeron mapas de distribución y frecuencia de ocurrencia de los ambientes.

Para identificar el principal factor de estrés que afectó el rendimiento del cultivo, se desarrolló un modelo de Red Neural Artificial supervisada en cada ambiente, el cual relacionó 16 variables ambientales explicativas con una variable de respuesta (rendimiento de cultivo). Finalmente, para identificar la importancia relativa (positiva o negativa) de cada variable predictora sobre la variable de respuesta, se aplicó el algoritmo de Garson a cada modelo de red neuronal.

Los resultados mostraron que el modelo ORYZA2000 pudo predecir la duración de cada uno de los estados fenológicos de los diferentes cultivares, en ambientes contrastantes, con un error no superior a 4 días. De igual manera fue lo suficientemente preciso para simular el rendimiento, el índice de área foliar (LAI) y biomasa de los órganos de los cultivares considerados. En promedio, los valores de RMSEn variaron entre 22% - 32% para la biomasa total, 27% - 37% para la biomasa de hoja verde, 27% - 40% para la biomasa de tallo, 30% - 40% para la biomasa de panícula y 30% - 35% para el LAI, y Los 11% - 15% para el rendimiento. Los resultados obtenidos revelaron que el modelo ORYZA2000 puede aplicarse como una herramienta valiosa para evaluar el desempeño del arroz bajo diferentes condiciones ambientales, y estrategias de manejo en las regiones del Centro y Bajo Cauca de Colombia.

Con base en los rendimientos simulados de cada grupo ambiental, estos se clasificaron de la siguiente manera: Altamente favorables (HFE), Favorables (FE) y Menos favorable (LFE). El HFE se caracterizó por tener los mayores rendimientos promedios ( $9.140 \text{ kg ha}^{-1}$ ), y representó el 18,5% del total del área de estudio. Por su parte, el FE tuvo un rendimiento promedio de  $7.578 \text{ kg ha}^{-1}$ , y ocupó el 50,2% de toda la zona de producción; Mientras que el LFE se caracterizó por ser el ambiente de menor rendimiento ( $6.000 \text{ kg ha}^{-1}$ ), y por ser el segundo ambiente con la mayor probabilidad de ocurrencia (31.3%).

Espacialmente, el HFE se distribuyó en dos pequeñas regiones ubicadas en el centro norte del departamento de Tolima y al sur del Huila. Este ambiente se caracterizó por tener una alta estabilidad ambiental (87% - 100%). Por otra parte, el FE ocupó una gran región que comprende los municipios del centro y norte de los departamentos de Tolima y Huila, dicha región se caracterizó por tener una mediana estabilidad ambiental (54% - 86%). En cuanto al ambiente LFE, este se distribuyó en el departamento de Córdoba, y presentó una mediana estabilidad ambiental (54% -86%).

Con base en el Análisis de Modelos de Redes Neuronales y la relevancia métrica, se identificó la temperatura mínima como el principal factor abiótico limitante que tiene un efecto negativo en el rendimiento del arroz. La fase comprendida entre la iniciación panicular y el final de la floración, fue la etapa más sensible al aumento de las temperaturas nocturnas. La temperatura mínima en el peor ambiente (LFE) osciló entre  $23^{\circ}\text{C}$  -  $25^{\circ}\text{C}$ , afectando el rendimiento, el número de espiguillas y la producción total de biomasa de los cuatro cultivares considerados en el estudio.

**Palabras clave:** *Caracterización ambiental, modelación de cultivos, Oryza Sativa, Factor limitante, Redes Neuronales.*

## Abstract

The aim of this study was carry out an environmental zonation and characterization of irrigated rice system across Tolima, Huila and Cordoba departments in Colombia, applying a spatial simulation approach through the ORYZA2000 crop model, as well as, identify the most relevant abiotic stress that limit the crop yield over the study region. Under this research, we seek to generate information that can support the new breeding programs strategies, in their effort to develop new germplasms with good yield stability, tolerant to certain abiotic stresses and resistant to interannual climate variability.

Four rice Colombian varieties (Fedearroz 2000, Fedearroz 60, Fedearroz 733 and CT21375) were evaluated across 38 locations with historical climate information from 1984 to 2012, considering 18 planting dates and six different soil types, in the Central and Bajo Cauca regions. An agglomerative hierarchical clustering method was employed on modeled attainable rice yields, in order to identify the environmental groups. Clustering efficiency and stability indicators were used to determine the optimal number of environments for the study region. Using the clustering results, maps of environments distribution and frequency were produced.

To identify the main stress factor that affected crop yield, a supervised Artificial Neural Network model was developed in each environment, which related 16 explicative environmental variables with one response variable (crop yield). Finally, to identify the relative importance (positive or negative) of each explicative variable on the response variable, the Garson algorithm was applicate to each Neural Network Model.

The results showed that ORYZA2000 model could predict well the duration of each phenological stage, however, consistently underestimated the length of growing period from 2 to 4 days. The range in normalized root mean square error (RMSEn) values for each phenological stage was between 3% and 6%. From the evaluation, was concluded that ORYZA2000 was sufficiently accurate in simulation of yield for each cultivar, leaf area index (LAI) and biomass of crop organs over time and across the regions. On average, RMSEn values were 22%–32% for total biomass, 27% – 37% for green leaf biomass, 27% – 40% for stem biomass, 30 % – 40 % for panicle biomass and 30% – 35% for LAI. The RMSEn values for final yield were 11.0 % – 15.0 %. The results obtained revealed that ORYZA2000 model can be applicate as a valuable tool to support the research oriented to

assess the rice performance under different environmental conditions, and management strategies developed across the Centro and Bajo Cauca regions in Colombia.

According to the yields obtained in each environmental group, the production environments were classified as follows: Highly Favorable (HFE), Favorable (FE) and Less Favorable (LFE). The HFE was the environment with the highest average yields (9,140 kg ha<sup>-1</sup>) and represented 18.5% of the total study area. For its part, the FE, had an average yield of 7,578 kg ha<sup>-1</sup>, occupying 50.2% of the entire production area; while the LFE was characterized by the lowest yielding environment (6,000 kg ha<sup>-1</sup>), and it was the second environment with the highest probability of occurrence (31.3%) within the study region.

Spatially, the HFE was distributed in two small regions located at north center Tolima department and to the southern Huila. This environment was distinguished to have a high environmental stability (87.0 % - 100.0 %). On the other hand, the FE occupied a large region comprising the municipalities across the central and northern Tolima and Huila, this region was distinguished by its medium environmental stability (54% - 86%). Regarding the LFE environment, it was distributed in Cordoba department, which presented a medium-environmental stability (54% - 86%).

Based on the Neural Network Model Analysis and the metric relevance, the minimum temperature was identified as the main limiting abiotic factor that has a negative effect on the rice yield. From the panicle initiation to the end of the flowering stage, was the most sensitive stage to the increase of the night temperatures. The minimum temperature in worst environment was between 23°C – 25°C, which reduced the yield, spikelet number and total biomass production of the four cultivars considered in the study.

**Keywords:** *Environmental characterization, crop modeling, Oryza Sativa, limiting factor, Neural networks.*

# Contenido

	Pág.
<b>Introducción .....</b>	<b>16</b>
<b>Capítulo 1.....</b>	<b>20</b>
<b>Marco Referencial .....</b>	<b>20</b>
1.1    El cultivo de arroz en las zonas Centro y Bajo Cauca de Colombia.....	21
1.1.1    Zona Centro.....	21
1.1.2    Zona Bajo Cauca.....	22
1.2    Importancia de caracterizar ambientes de producción de cultivos .....	23
1.3    Caracterización ambiental con base en la modelación de cultivos – Simulación de una amplia gama de ambientes.....	24
1.4    Modelos de simulación de cultivos .....	24
1.4.1    Clasificación de los modelos de cultivos .....	25
1.4.2    Modelos de simulación para el cultivo de arroz.....	30
1.4.3    Modelo de cultivo ORYZA2000.....	37
1.5    Análisis clúster .....	48
1.5.1    Clúster jerárquico aglomerativo .....	49
1.6    Redes Neuronales Artificiales (RNA).....	49
1.6.1    Métodos de aprendizajes de las RNA .....	51
1.6.2    Principales aplicaciones de las RNA.....	52
1.6.3    Análisis de Relevancia Métrica .....	53
<b>Capítulo 2.....</b>	<b>56</b>
<b>Materiales y métodos .....</b>	<b>56</b>
2.1    Caracterización del área de estudio .....	56
2.1.1    Departamento del Tolima.....	56
2.1.2    Departamento del Huila .....	59
2.1.3    Departamento de Córdoba .....	61
2.2    Datos observados .....	63
2.2.1    Datos de clima.....	63
2.2.2    Datos de suelos.....	65
2.2.3    Datos de los ensayos de arroz de riego.....	67
2.3    Parametrización y evaluación del modelo de cultivo.....	73
2.3.1    Modelo de cultivo Oryza2000.....	73
2.3.2    Parametrización del modelo de cultivo.....	75
2.3.3    Evaluación del modelo de cultivo.....	78
2.4    Identificación y caracterización ambiental .....	79

<b>Capítulo 3.</b> .....	<b>84</b>
<b>Resultados</b> .....	<b>84</b>
3.1 Evaluación del modelo Oryza2000.....	85
3.1.1 Tasas de desarrollo fenológico (DVR) .....	85
3.1.2 Parámetros de producción de biomasa y desarrollo foliar .....	91
3.1.3 Dinámica de Biomasa y Área Foliar (Datos de Calibración).....	96
3.1.4 Dinámica de Biomasa y Área foliar (Datos de Validación) .....	105
3.2 Identificación y Caracterización de ambientes .....	112
3.3 Determinación del mayor estrés abiótico y su influencia sobre el rendimiento	121
<b>Capítulo 4.</b> .....	<b>129</b>
<b>Discusión</b> .....	<b>129</b>
4.1 Consistencia entre los ensayos experimentales y las simulaciones de ORYZA2000. ....	129
4.2 Efecto de la alta temperatura mínima sobre el cultivo de arroz, en los diferentes grupos ambientales. ....	132
4.3 Efecto de las altas temperaturas y la alta humedad relativa sobre los rendimientos del ambiente menos favorable (LFE). ....	136
<b>Capítulo 5.</b> .....	<b>138</b>
<b>Conclusiones</b> .....	<b>138</b>
<b>Capítulo 6.</b> .....	<b>140</b>
<b>Recomendaciones</b> .....	<b>140</b>
<b>Bibliografía</b> .....	<b>146</b>

## Lista de figuras

	Pág.
Figura. 1. Regiones productoras de arroz en Colombia.....	21
Figura. 2. Factores jerárquicos de crecimiento (van Ittersum et al., 2003). .....	27
Figura. 3. Función de respuesta: Tasa de desarrollo vs temperatura media.....	38
Figura. 4. Relación entre el número de espigas vs crecimiento del cultivo desde la iniciación panicular hasta la floración . .....	45
Figura. 5. Grado días de enfriamiento vs porcentaje de esterilidad de la espiga .....	46
Figura. 6. Temperatura máxima media durante la floración vs la fertilidad de la espiga .	47
Figura. 7. Arquitectura de una Red Neuronal Artificial del tipo Perceptron Multicapa.....	50
Figura. 8. Aprendizaje supervisado .....	51
Figura. 9. Aprendizaje no supervisado. ....	52
Figura. 10. Parámetros de una RNA utilizados en las técnicas de relevancia métrica. ..	55
Figura. 11. Ubicación geográfica del área de estudio. ....	57
Figura. 12. Características climáticas del departamento del Tolima .....	59
Figura. 13. Características climáticas del departamento del Huila .....	61
Figura. 14. Características climáticas del departamento Córdoba .....	63
Figura. 15. Distribución espacial de las estaciones meteorológicas.....	66
Figura. 16. Ecuaciones para determinar las propiedades hidráulicas del suelo .....	68
Figura. 17. Ubicación geográfica de los sitios experimentales.....	69
Figura. 18. Diseño de parcelas experimentales.....	71
Figura. 19. Matriz de datos: Rendimientos para cada escenario de simulación.....	80
Figura. 20. Evaluación de la fenología del cultivo - Datos de calibración.....	90
Figura. 21. Evaluación de la fenología del cultivo - Datos de validación .....	90
Figura. 22. Calibración de la dinámica de distribución de asimilados en la planta .....	93
Figura. 23. Calibración de la dinámica del Área Foliar Específica (SLA). ....	95
Figura. 24. Dinámica de la producción de biomasa seca total (WAGT), hojas verdes (WLVG), panículas (WSO), y tallos (WST) - Datos de Calibración en Saldaña.....	98
Figura. 25. Dinámica del Índice de Área Foliar (LAI) - Datos de calibración Saldaña . ...	99
Figura. 26. Evaluación del modelo - Producción de biomasa, Índice de Área Foliar y Rendimiento del cultivo - Datos de calibración en Saldaña .....	100
Figura. 27. Dinámica de la producción de biomasa seca total (WAGT), hojas verdes (WLVG), panículas (WSO), y tallos (WST) - Datos de Calibración en Monteria.....	101

Figura. 28. Dinámica del Índice de Área Foliar (LAI) - Datos de calibración Monteria ...	102
Figura. 29. Evaluación del modelo - Producción de biomasa, Índice de Área Foliar y Rendimiento del cultivo - Datos de calibración en Monteria .....	103
Figura. 30. Evaluación del modelo - Producción de biomasa, Índice de Área Foliar y Rendimiento del cultivo - Datos de validación en Aípe. ....	108
Figura. 31. Evaluación del modelo - Producción de biomasa, Índice de Área Foliar y Rendimiento del cultivo - Datos de validación en Saldaña.. ....	109
Figura. 32. Evaluación del modelo - Producción de biomasa, Índice de Área Foliar y Rendimiento del cultivo - Datos de validación en Monteria.....	110
Figura. 33. Identificación de grupos ambientales - Clústering jerárquico .....	112
Figura. 34. Frecuencia de ocurrencia de los grupos ambientales.....	113
Figura. 35. Distribución espacial de los grupos ambientales .....	114
Figura. 36. Frecuencia de ocurrencia de las variedades en los grupos ambientales .....	116
Figura. 37. Frecuencia de ocurrencia de los diferentes tipos de suelos .....	118
Figura. 38. Rendimientos históricos simulados en los diferentes grupos ambientales...	120
Figura. 39. Arquitecturas de Red Neuronal Artificial.....	122
Figura. 40. Diagramas de Interpretación Neuronal (DIN) para el modelo RNA obtenido en el grupo ambiental LFE. ....	124
Figura. 41. Importancia relativa de las variables predictoras en los modelos neuronales de los diferentes ambientes .....	128
Figura. 42. Respuesta del rendimiento simulado y del número de espigas, al cambio de temperatura mínima media .....	133
Figura. 43. Perfil de distribución del rendimiento vs el cambio de la temperatura mínima media durante el inicio y el final de la floración .....	134

## Lista de tablas

	<b>Pág.</b>
Tabla 1. Modelos de simulación para el cultivo de arroz que son utilizados con mayor frecuencia en la actualidad.....	34
Tabla 2. Variables forzantes y procesos fisiológicos que caracterizan los diferentes modelos de simulación para el cultivo de arroz. ....	35
Tabla 3. Características de algunas técnicas de relevancia métrica. ....	54
Tabla 4. Ubicación de las estaciones meteorológicas .....	64
Tabla 5. Características de los suelos en las estaciones experimentales. ....	70
Tabla 6. Datos experimentales para la calibración y validación del modelo. ....	74
Tabla 7. Variables para la construcción de los modelos de Red Neuronal.....	81
Tabla 8. Tasas de desarrollo fenológico. ....	86
Tabla 9. Comparación entre la fenología observada y simulada – Datos de Calibración. ....	87
Tabla 10. Comparación entre la fenología observada y simulada – Datos de validación. ....	88
Tabla 11. Evaluación del desempeño del modelo ORYZA2000 - Datos de calibración....	89
Tabla 12. Evaluación del desempeño del modelo ORYZA2000 - Datos de validación....	89
Tabla 13. Fracción de materia seca en los cultivares seleccionados. ....	91
Tabla 14. Tasas de crecimiento relativo de hojas. ....	94
Tabla 15. Valores de Área Foliar Específica.....	95
Tabla 16. Parámetros cuantitativos de bondad de ajuste para evaluar el desempeño del modelo ORYZA2000 - Datos de calibración .....	104
Tabla 17. Parámetros cuantitativos de bondad de ajuste para evaluar el desempeño del modelo ORYZA2000 - Datos de calibración. ....	111

## Introducción

El cultivo de arroz (*Oryza Sativa L.*) es considerado la principal fuente alimenticia en muchas de las regiones tropicales de América Latina y el Caribe (Zorrilla et al., 2012), juega un papel importante dentro de la ingesta calórica diaria, aporta un 20.0% más de calorías que otros cultivos (trigo, maíz, yuca o papa) ampliamente cultivados en la región; además de ser una importante fuente de proteínas para más del 20.0% de la población pobre de Latino América (McLean, Hardy, & Hettel, 2013).

La producción de arroz tiene una gran relevancia social y económica. Contribuye de manera significativa en la generación de empleo y al Producto Interno Bruto (PIB) de los países tropicales de la región (Espinosa, 2002).

Colombia se ubica en el puesto 25 dentro del grupo de países productores de arroz en el mundo, y es el tercero en Latino América después de Brasil y Perú (~ 1.8 millones de toneladas en 2014 - FAOSTAT, 2016). El cultivo de arroz en Colombia es considerado, después del maíz, el segundo cultivo transitorio con mayor importancia en la agricultura del país. El arroz procesado juega un papel importante en la dieta y en la canasta familiar de los hogares colombianos, y en particular, en la canasta de consumo de los hogares más pobres. Según cifras del Departamento Administrativo Nacional de Estadística (DANE, 2015) y la Federación Nacional de Arroceros (FEDEARROZ), el cultivo de arroz paddy representa el 12.0% del área cosechada en Colombia y el 30.0% de los cultivos transitorios. Su producción representa el 6.0% del valor de la producción agropecuaria y el 10.0% de la actividad agrícola Colombiana.

A nivel nacional, el arroz se cultiva en cinco regiones geográficas (La región Centro, Bajo Cauca, Llanos Orientales, Costa Norte y Santanderes), en las cuales se destacan dos sistemas de producción diferenciados: Mecanizado y tradicional (manual), predominando el sistema mecanizado en un 94.0% del total de área arrocera, y en la cual más del 70.0% de la producción corresponde a arroz bajo riego (McLean et al., 2013). La región Centro, una de las regiones objetivo de este estudio, la conforman los departamentos de Tolima y Huila, concentra la mayor cantidad de agricultores, en ella se cultivan cerca de 150,000 hectáreas de arroz (alrededor de un tercio del total área cultivada en el país), con una

producción promedio de 900,000 toneladas/año y rendimientos promedios de 6,000 kg/ha. Por su parte, la región del Bajo Cauca (nuestra segunda región de interés) cuenta con una superficie cultivada de 60,000 hectáreas (representa el 15% de la superficie total cultivada), tiene producción promedio de 200,000 toneladas/año y un rendimiento promedio de 3,500 kg/ha. La región de los Llanos Orientales está representada principalmente por los departamentos de Meta y Casanare, en ella se cultiva aproximadamente el 40.0% del total de área dedicada al cultivo de arroz en Colombia; los agricultores de esta región se establecen bajo un sistema de siembra bajo seco, y producen rendimientos promedios alrededor de 4,000 kg/ha. Por último, las regiones de los Santanderes y Costa Norte, se distinguen por presentar las menores tasas de producción (FEDEARROZ, 2016).

En los últimos 15 años el total de área dedicada al cultivo de arroz en el país ha disminuido en un 17.0%, pasando de 447 mil hectáreas en el 2000 a 373 mil a finales del 2014; las regiones Centro y Bajo Cauca han sido las más afectadas al disminuir el área cultivada en 15.0% y 29.0%, respectivamente. De igual forma, la producción nacional pasó de cerca de 2 millones doscientos mil toneladas a 1 millón 730 mil toneladas, que equivale a una disminución del 22.0%. Por su parte, el rendimiento promedio nacional ha bajado en un 12.0%, pasando de 4.9 toneladas por hectárea a 4.3 toneladas en el mismo periodo de tiempo (DANE, 2014; FEDEARROZ, 2016). Los altos costos de producción y de insumos agrícolas (Mazuera Fernandez & Neira Rodriguez, 2009), las variaciones en los precios del mercado de una época de siembra a otra (Areiza Segura, 2013), y los acuerdos comerciales entre países de la región que perjudican la industria arrocera nacional, son aspectos socio-económicos por los cuales muchos agricultores de oportunidad están saliendo del cultivo, lo que demuestra así la disminución de área cultivada y producción de arroz en Colombia.

No obstante, la variabilidad climática que actualmente se presenta en el país, producto de la presencia de fenómenos como EL Niño y La Niña, se ha convertido en la principal causa que golpea a la mayoría de los agricultores en el campo (Lau, Jarvis, & Ramírez, 2010). El cultivo de arroz no ha sido ajeno a este fenómeno, el aumento de las temperaturas máximas y mínimas, y las prolongadas sequias han afectado los rendimientos del cultivo en departamentos como Tolima y Huila. De igual forma, en el departamento de Córdoba, el aumento de la temperatura mínima y humedad relativa, la mala distribución de las lluvias, la disminución de la radiación solar y la intensa sequía, ha aumentado los niveles de

vaneamiento en el cultivo (FEDEARROZ, 2012). Una de las soluciones ante este tipo de condiciones desfavorables en el ambiente, es el desarrollo de nuevos cultivares que puedan contrarrestar los efectos negativos en el rendimiento a causa de un determinado tipo de estrés abiótico o biótico (A.B. Heinemann et al., 2016). Sin embargo, las altas interacciones Genotipo x Ambiente que ocurren producto de la variabilidad del clima, dificultan la comprensión de tales interacciones e impiden el progreso en el desarrollo de nuevos genotipos adaptados a este tipo de regiones (Y. S. Chauhan & Rachaputi, 2014; Y. Chauhan, Wright, Rachaputi, & Mccosker, 2008).

La identificación y caracterización de ambientes de producción para el cultivo de arroz permite tener un entendimiento de su naturaleza, del nivel de ocurrencia de eventos agroclimáticos extremos, de las implicaciones que estos tienen para el crecimiento y desarrollo de la cultivo, así como del manejo apropiado que se le debe dar (Karine Chenu, 2015). De igual forma, ayuda a definir la frecuencia de ocurrencia de áreas potenciales para la selección y evaluación de nuevos cultivares (Sadras et al., 2009).

Hasta el momento se han propuesto varias metodologías que permiten clasificar y caracterizar ambientes de producción de cultivos (Karine Chenu, 2015). Muchas de ellas se han enfocado en identificar grupos ambientales a partir del establecimiento de ensayos experimentales en múltiples sitios, a través de varios años, caracterizándolos de forma cualitativa (ej. Presencia o ausencia de plagas y enfermedades) y cuantitativa (midiendo indicadores agroclimáticos, rendimiento promedio de diversos genotipos o la intensidad y frecuencia de estreses (biótico o abiótico) a los que está expuesta la planta; ej. Y. Chauhan et al., 2008; Chenu, Deihimfard, & Chapman, 2013; van Wart et al., 2013; Yang, Blade, Crossa, Stanton, & Bandara, 2005). No obstante, a pesar de que estas metodologías ayudan a interpretar la interacción entre la planta y el ambiente en que se desarrolla, sólo se llevan a cabo con información experimental de un número limitado de años y sitios en específico, sin considerar las variaciones a nivel espacial y temporal de factores agroclimáticos (suelo y clima), lo cual dificulta la caracterización (momento de ocurrencia e intensidad) de los patrones de estrés, y plantea un mayor desafío al momento de caracterizar, a través del tiempo, grandes regiones productoras.

Ante estos desafíos, las herramientas de modelación abren la puerta para posicionarse como una nueva alternativa que integra factores como el clima, el suelo, el cultivo y el

---

manejo agronómico; para delimitar y caracterizar ambientes de producción, en múltiples sitios y largos periodos de tiempo. Permitiendo, en primer lugar, tener mayor información espaciotemporal que los ensayos experimentales convencionales (ej. Casadebaig et al., 2016; Heinemann et al., 2016); en segundo lugar, cuantificar la intensidad de los diferentes estreses a los cuales la planta se encuentra expuesta (ej. K. Chenu et al., 2011; Alexandre Bryan Heinemann et al., 2015); y por último, ser de gran utilidad para el planteamiento de nuevas estrategias de mejoramiento, que buscan desarrollar nuevos germoplasmas con alto potencial de rendimiento y con mayor resiliencia a los estreses abióticos (Holzworth et al., 2014).

En este estudio, se analizó la identificación y caracterización agroambiental del sistema productivo de arroz de riego en los departamentos de Tolima, Huila y Córdoba, para el periodo 1984 – 2012 (28 años); usando un modelo de simulación de cultivo, con el fin de generar información que pueda apoyar las nuevas estrategias de los programas de mejoramiento, en su esfuerzo por desarrollar nuevos germoplasmas con buena estabilidad en el rendimiento, tolerantes a determinados estreses abióticos y resistentes a las variaciones climáticas interanuales (A.B. Heinemann et al., 2016). Los objetivos específicos fueron: 1) Calibrar y evaluar el desempeño del modelo de cultivo Oryza2000 para cuatro variedades Fedearroz 2000, Fedearroz 60, Fedearroz 733 y CT21375; en las regiones productoras Centro (Tolima y Huila) y Bajo Cauca (Córdoba) de Colombia; 2) Clasificar grupos ambientales en dichas regiones productoras, a partir de los rendimientos alcanzables de las variedades seleccionadas y su variabilidad a través de los años; y 3) Identificar, para cada grupo ambiental, el mayor patrón de estrés abiótico, y su influencia sobre el rendimiento del cultivo.

## Capítulo 1.

### Marco Referencial

“Atiende al consejo y acepta la corrección, y llegarás a ser sabio”

**Proverbios 19:20 (NVI)**

Este capítulo se divide en seis secciones las cuales están divididas de la siguiente manera: Inicialmente se presenta el contexto geográfico del cultivo de arroz en las regiones productoras Centro (Departamentos de Tolima y Huila) y Bajo Cauca (Córdoba). Posteriormente, se plantea un concepto general sobre la importancia de identificar y caracterizar ambientes de producción, y el rol de los modelos de simulación de cultivos dentro de la clasificación de grupos ambientales. Después, se presenta una definición sobre los modelos de simulación de cultivos y se detalla los componentes más relevantes del modelo ORYZA2000 para el cultivo de arroz. Finalmente, se explica el método jerárquico aglomerativo de análisis clúster, el cual es aplicado sobre los rendimientos simulados para la clasificación de los ambientes; y se define el concepto de las redes neuronales artificiales, y su utilidad para asociar la influencia de las variables de entrada de un modelo sobre una variable de respuesta, a través de técnicas de relevancia métrica.

## 1.1 El cultivo de arroz en las zonas Centro y Bajo Cauca de Colombia

### 1.1.1 Zona Centro

El cultivo de arroz en la zona Centro de Colombia se distribuye entre los departamentos de Tolima, Huila, Caquetá, Cundinamarca, Valle del Cauca y Cauca (ver Figura. 1 ). En promedio, se cultivan cerca de 150,000 hectáreas al año, predomina el sistema de siembra bajo riego mecanizado, con una producción promedio anual de 865,000 toneladas de arroz paddy seco (42.0% de la producción total nacional anual), que se traduce en un rendimiento promedio de 6.0 toneladas por hectárea ( $\text{ton ha}^{-1}$ ) (FEDEARROZ, 2016).

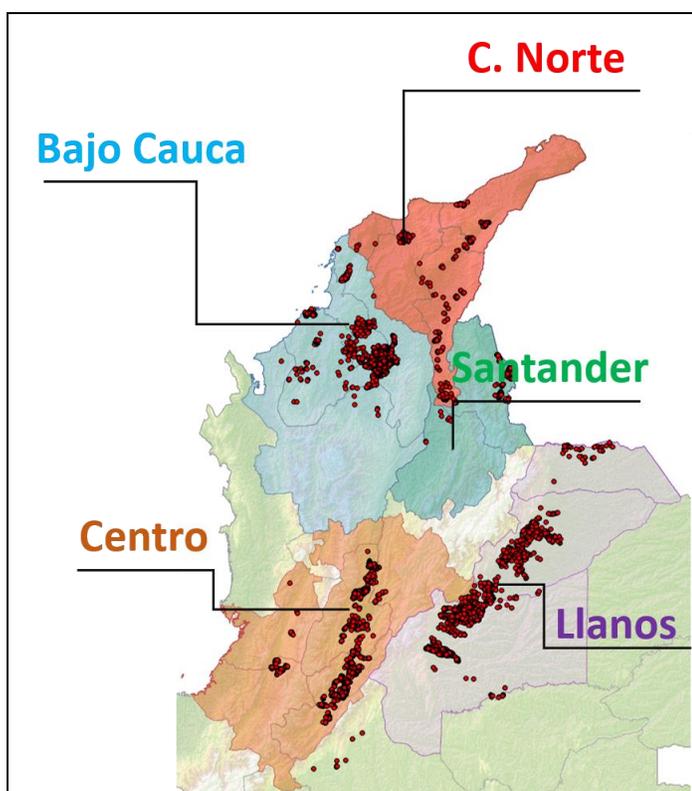


Figura. 1. Regiones productoras de arroz en Colombia

Según el III Censo Nacional Arrocero (FEDEARROZ, 2008b), la zona Centro presenta un total de 5,194 productores, distribuidos en 8,783 Unidades Productoras de Arroz (UPA),

las cuales cultivan cerca de 70,000 hectáreas durante la primera época de siembra del año (Abril - Julio) y 80,000 hectáreas en la segunda época (Octubre - Enero). Respecto al tamaño y a la tenencia de las UPA, el 77.0% son menores de 10 ha, los propietarios tienen el 42% de participación, y los arrendatarios el 57%. En términos del área sembrada, en promedio, la participación de los propietarios es de 45% y el de los arrendatarios es de 53%.

La zona Centro concentra el mayor porcentaje de área sembrada en los departamentos de Tolima y Huila, los cuales representan el 72% y 23% del total de área sembrada en el año, respectivamente.

### **1.1.2 Zona Bajo Cauca**

La región productora del Bajo Cauca cubre los departamentos de Antioquia, Bolívar, Choco, Córdoba y Sucre (ver Figura. 1); esta región del país es la de mayor diversidad, tiene los tres sistemas de producción de arroz (riego, secano mecanizado y secano manual). Presenta todo tipo de nivel tecnológico, desde la recolección con hoz hasta con combinada, y el manejo agronómico es igual de diverso (FEDEARROZ, 2012).

El secano manual es de gran importancia en la zona, este sistema es de pequeños productores, el área promedio de arroz es de 1.6 hectáreas, tienen un nivel tecnológico muy bajo y en su mayoría (~74.0%) produce arroz para auto consumo. Del total de agricultores que destinan sus cosechas para el autoconsumo (~9,000), cerca de 3,668 se ubican en Sucre, seguido por Choco (2,828), Córdoba, Antioquia y Bolívar. El otro 24.0% (3,400) utiliza algún tipo de maquinaria en la producción de arroz (FEDEARROZ, 2008a). En arroz mecanizado, Sucre cuenta con aproximadamente 2,100 productores, Bolívar con 650, Córdoba con 506 y Antioquia con 130. La zona del Bajo Cauca participa con 27% de los productores y 21% de las UPA en relación al total del país.

Respecto al área sembrada, en el Bajo Cauca se siembran aproximadamente 58,000 hectáreas al año, de las cuales el 26.0% se cultivan bajo un sistema de secano manual, y el 74% restante se cultiva bajo un sistema mecanizado; el 13% del total de área bajo arroz mecanizado corresponde a arroz de riego, y el 87.0% se dedica a secano mecanizado. La producción promedio anual de la zona equivale a 194,000 toneladas, el 83% lo aporta el arroz mecanizado y el resto de la producción se deriva del sistema secano manual. El rendimiento promedio anual del Bajo Cauca es de 3.4 ton ha<sup>-1</sup>, lo que la cataloga como la

zona de menor rendimiento en relación a las demás zonas productoras del país (FEDEARROZ, 2016).

El Bajo Cauca se ha caracterizado por ser una región de pequeñas explotaciones arroceras, cerca del 93% de las unidades productoras son menores a 10 hectáreas. Respecto al arroz mecanizado, el 79% de las UPA son menores a 10 hectáreas y aportan el 30% del área sembrada; mientras que los predios entre 10 y 50 hectáreas participan con el 40% del total de área sembrada (FEDEARROZ, 2008a).

## **1.2 Importancia de caracterizar ambientes de producción de cultivos**

Muchos programas de mejoramiento de cultivo suelen incluir dentro de sus estrategias de investigación, la evaluación de nuevos genotipos en múltiples sitios, bajo condiciones agroclimáticas específicas, los cuales son considerados como una muestra representativa de todo un ambiente de producción. Como resultado de esta estrategia, muchos cultivares suelen adaptarse muy bien a condiciones bajo las cuales fueron evaluados, generalmente durante un largo periodo de tiempo. No obstante, cuando se enfrentan a condiciones agroambientales totalmente opuestas, suelen presentar mayor incertidumbre en cuanto a su capacidad de adaptación (Karine Chenu, 2015).

Por ejemplo, si se desea evaluar un grupo de cultivares que tengan buenos rendimientos en suelos con excesiva humedad, durante varios años consecutivos, en un ambiente que generalmente se caracteriza por presentar condiciones de sequía, esto puede sesgar la selección de los mejores cultivares para todo el ambiente de producción, ya que solo van a responder positivamente ante condiciones que no son las características de dicho ambiente. Bajo años malos (sequia) seguramente los cultivares evaluados no van a tener la capacidad de adaptarse o responder ante tal condición, y es básicamente porque fueron seleccionados para condiciones diferentes a las características del ambiente de interés.

Para evitar este sesgo, los mejoradores deben considerar si realmente el ensayo de evaluación que están llevando a cabo durante su proceso de evaluación, es representativo para todo el ambiente de producción, y para esto, se requiere entender y conocer todos los aspectos agroclimáticos relacionados con el mismo.

### **1.3 Caracterización ambiental con base en la modelación de cultivos – Simulación de una amplia gama de ambientes**

Las herramientas de modelación ofrecen una nueva vía para la caracterización de ambientes de producción, ya que permiten integrar la interacción existente entre condiciones de suelos, clima histórico, manejo agronómico, y características del genotipo para tener una mejor comprensión de la respuesta de los cultivos a una mayor escala espacial; la cual no es posible lograr con los ensayos experimentales que se realizan en sitios puntuales (por ejemplo, Y. S. Chauhan & Rachaputi, 2014; Hammer & Jordan, 2007; A.B. Heinemann et al., 2016). Además de la producción, estos autores también han considerados características fenológicas de los cultivos (ejemplo: tiempo floración y madurez fisiológica) para explicar el impacto de la fenología en el rendimiento.

Por otro lado, también es importante reconocer que los modelos de cultivos presentan ciertas limitaciones al momento de ser aplicados en estudios de caracterización ambiental, estas limitantes están mayormente relacionadas con la incapacidad de ciertos modelos para cuantificar el efecto de estreses bióticos sobre los cultivos, ya que la gran mayoría no están parametrizados para este fin. Adicionalmente, otra de las limitaciones se debe a que no se cuenta con una parametrización para todos los genotipos de interés para los usuarios. No obstante, este vacío puede ser complementado con información experimental de múltiples genotipos, tal como lo reporta K. Chenu et al. (2011).

La identificación de ambientes a partir de rendimientos simulados puede caracterizarse por clases ambientales que son espacialmente contiguas (Y. S. Chauhan & Rachaputi, 2014; Alexandre Bryan Heinemann et al., 2015), destacando la probabilidad de ocurrencia de los diferentes grupos ambientales, así como el principal factor ambiental que puede generar algún tipo de limitación en el crecimiento y desarrollo del cultivo.

### **1.4 Modelos de simulación de cultivos**

Existen muchas maneras de interpretar el término de Modelos de Simulación, que van desde modelos simples y empíricos, hasta modelos muy complejos. Un modelo es la representación simplificada de un sistema, y un sistema es una parte limitada de la realidad

que contiene elementos interrelacionados (B. A. M. Bouman, Van Keulen, Van Laar, & Rabbinge, 1996). Por lo tanto, un modelo de simulación de cultivos, no es nada más que la simplificación de un sistema vegetal, el cual simula o imita el comportamiento real de un cultivo, prediciendo el crecimiento y desarrollo de sus componentes, tales como raíces, hojas, tallos y granos (Radha & Murthy, 2004). Los modelos, en general no sólo simulan el estado final de la biomasa y rendimiento de los cultivos, sino que también contienen información cuantitativa sobre de los principales procesos involucrados en el crecimiento y desarrollo de la planta (Jame & Cutforth, 1996).

### 1.4.1 Clasificación de los modelos de cultivos

Según la escuela de modelación agro-ecológica de la universidad de Wageningen (<http://models.pps.wur.nl/>), reconocida a nivel mundial por su larga trayectoria en el desarrollo y aplicación de modelos de simulación en investigaciones relacionadas con el campo de la agro-ecología, existen tres criterios fundamentales para caracterizar y clasificar los distintos tipos de modelos de cultivos. Estos son:

- Jerarquía en los factores de crecimiento.
- Objetivos del modelo y nivel de detalle.
- Dominios de aplicación, es decir, investigación, educación y apoyo a la toma de decisiones.

#### 1.4.1.1 Jerarquía en los factores de crecimiento

En 1982, los científicos de Wit y Penning de Vries propusieron una clasificación de los sistemas de simulación de cultivos, en tres situaciones de producción (Figura. 2):

- *Situación 1 - Producción potencial:* Está definida como la máxima producción que puede alcanzar un cultivo que crece y se desarrolla sin ningún tipo de limitación. Bajo situación de producción potencial, la planta depende, en gran medida, de las características físicas del ambiente (por ejemplo, la intensidad de la radiación solar, la concentración de dióxido de carbono, la temperatura del aire), de sus atributos fisiológicos, de la ejecución adecuada de las prácticas de manejo agronómico (tales

como, la selección de la fecha de siembra, la densidad de siembra, el cultivar adecuado, el suministro de agua y nutrientes necesarios), y de la adecuada protección contra malezas, plagas y enfermedades.

- *Situación 2 - Producción alcanzable:* Esta situación de producción se presenta cuando el cultivo se desarrolla de manera limitada, debido a la escasez de agua o nutrientes. Bajo esta condición, el manejo agronómico resulta ser la mejor estrategia para proveer al cultivo la cantidad necesaria de agua y nutrientes, que puedan aumentar su producción hasta niveles potenciales.
- *Situación 3 - Producción actual:* El sistema de producción actual considera factores bióticos (por ejemplo, malas hierbas, plagas y enfermedades) y abióticos (por ejemplo, contaminantes y toxicidad por Aluminio) que reducen o dificultan el crecimiento de los cultivos. El manejo agronómico bajo este sistema de producción, está direccionado principalmente al manejo efectivo de estos agentes contaminantes. En la situación producción actual, la productividad obtenida suele ser el resultado de una combinación de factores limitantes por disponibilidad de agua y nutrientes, sumado a la presencia de factores bióticos y abióticos.

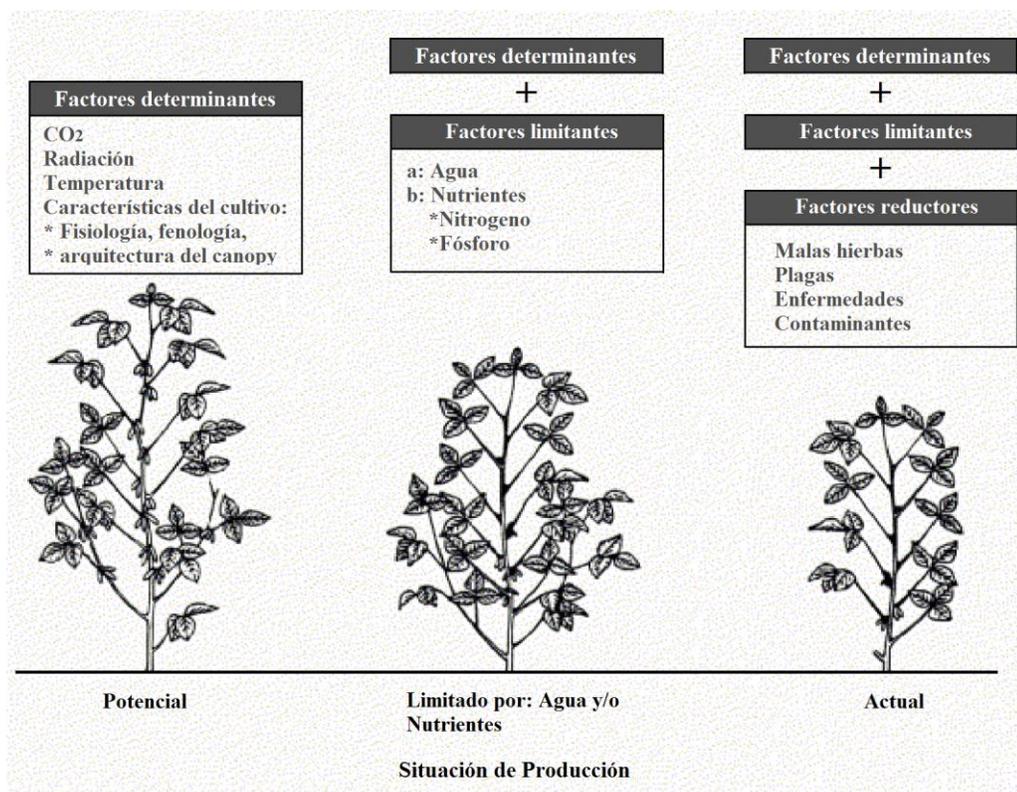


Figura. 2. Factores jerárquicos de crecimiento, situaciones de producción y niveles de producción asociados (van Ittersum et al., 2003).

### 1.4.1.2 Objetivo del modelo y nivel de detalle

Este segundo criterio se ha usado para clasificar los modelos agro-ecológicos dependiendo del propósito para el cual han sido diseñados, y el nivel de aplicabilidad que pueden tener. Aspectos como el tipo de análisis, nivel de predicción, sistema de cultivo, y la escala espacial (parcela agrícola o una región), entre otros, determinan el tipo de modelo a utilizar (van Ittersum et al., 2003). Diferentes categorías de modelos se han establecido para este fin, entre ellos se destacan los descriptivos o empíricos y los explicativos o mecanísticos.

- *Modelos descriptivos:* Estos modelos simulan un sistema de manera simplificada, muestran la existencia de relaciones entre los elementos del sistema sin dar explicación acerca de los mecanismos que modulan tal relación. Algunos ejemplos de este tipo de modelos son aquellos que utilizan datos experimentales para

relacionar la respuesta del rendimiento de un cultivo a la aplicación de fertilizantes, la relación entre el área foliar y el tamaño de hoja de una especie de planta dada, y la producción de biomasa en función de la intercepción de luz de un cultivo (Oteng-Darko, Yeboah, Addy, Amponsah, & Danquah, 2013).

- *Modelos explicativos*: Describen cuantitativamente los mecanismos y los procesos que guían el funcionamiento del sistema. Estos modelos tienen la capacidad de imitar procesos físicos, químicos o biológicos relevantes y describir cómo y por qué se produce una respuesta en particular. Durante el desarrollo de los modelos explicativos, el modelador suele comenzar con un cierto empirismo y, a medida que se adquieren conocimientos, se introducen parámetros adicionales y se introducen variables para explicar el rendimiento del cultivo.

Muchos de los modelos que actualmente se utilizan para fines de investigación son altamente explicativos o mecanicistas, mientras que los modelos descriptivos son ampliamente aplicados con propósitos predictivos (van Ittersum et al., 2003).

### **1.4.1.3 Dominios de aplicación**

Actualmente los modelos de simulación están teniendo mayor aplicabilidad en diferentes disciplinas, tales como la investigación, la enseñanza, actividades agrícolas, manejo de recursos agrícolas, análisis de políticas y pronósticos de producción de los cultivos, entre otras. A continuación se describen algunos de los dominios de aplicación de los modelos de simulación de cultivos:

- *Investigación científica*: Generalmente, el objetivo del uso de los modelos de simulación en la investigación agrícola, es integrar el conocimiento de los diferentes procesos que intervienen en el crecimiento y desarrollo de un cultivo, y explicar el comportamiento del todo un sistema. El conocimiento sobre de las diferentes características físicas y fisiológicas de una planta, adquirido del trabajo experimental, se integra en modelos de simulación con el fin de ser usados para probar hipótesis bajo diferentes condiciones o escenarios.

- 
- *Evaluación del cultivo en sitios específicos:* Los modelos de cultivo se pueden utilizar para predecir el rendimiento del cultivo en regiones donde no han sido cultivados, o en sitios donde no han crecido bajo condiciones óptimas.
  - *Análisis del rendimiento:* Los modelos de simulación estructurados bajo una base de conceptos fisiológicos, permiten estimar las causas de reducción del rendimiento del cultivo, debido a factores climáticos y no climáticos (por ejemplo, siembra, espaciamiento de los cultivos, fertilidad del suelo, plagas y enfermedades). De igual forma, dan la posibilidad de separar las ganancias en el rendimiento, en componentes relacionadas con las condiciones climáticas, y los mejoramientos genéticos y tecnológicos.
  - *Proyección de los cultivos bajo cambio climático:* La aplicación de modelos de simulación para estudiar el impacto potencial de la variabilidad y el cambio climático ha sido ampliamente usado en muchos países. La producción de cultivos depende en gran medida de las condiciones climáticas y, por lo tanto, cualquier cambio que se pueda presentar en este aspecto, tendrá efectos importantes sobre los rendimientos de las plantas y su productividad. La temperatura y el dióxido de carbono afectan procesos biológicos como la respiración, la fotosíntesis, el crecimiento de las plantas, la reproducción, el uso del agua, etc. Una comprensión adecuada de las distintas variables que conforman el clima ayudará a guiar a los agricultores hacia la toma de decisiones acertadas en el manejo de cultivos, las fechas de siembra, la programación del riego, entre otros.
  - *Prácticas adecuadas en el manejo de los cultivos:* Las herramientas de simulación pueden ser ampliamente utilizadas para determinar las mejores prácticas de manejo de un determinado sistema de cultivo. En el pasado, la investigación agronómica se centró principalmente en la producción de cultivos. No obstante, en la actualidad, además del nivel de rentabilidad de la producción, la calidad del medio ambiente también se ha convertido en un tema importante que los productores agrícolas deben abordar. Los agricultores requieren nuevas estrategias para optimizar la rentabilidad de sus cosechas, al tiempo que deben garantizar la calidad del suelo y minimizar la degradación ambiental. Las soluciones a este nuevo desafío requieren considerar la forma de interacción de diferentes componentes que afectan el crecimiento de la planta.

Para lograr este objetivo, las futuras investigaciones agrícolas requerirán considerablemente más esfuerzo y mayores recursos que las actuales actividades de investigación. Ante esto, las herramientas de modelación pueden ser usadas para determinar, a nivel de sitio específico, las mejores prácticas de manejo que los agricultores deberán realizar a largo plazo. Decisiones de inversión como la compra de sistemas de riego, estrategias de fertilización y planes de cosecha, entre otras, se pueden tomar con base a los resultados obtenidos con herramientas de simulación.

- *Pronósticos del rendimiento:* Las estimaciones del rendimiento de los cultivos antes de la cosecha, ya sea a nivel de finca o en grandes áreas, son de inmenso valor tanto para el investigador como para el agricultor en términos de planificación. Bajo este enfoque, el uso de modelos de cultivos y modelos de pronósticos climáticos son determinantes para proyectar la respuesta de la planta a mediano o largo plazo.
- *Mejoramiento y desarrollo de nuevas variedades de cultivos:* El desarrollo y la liberación de una variedad es un proceso complejo que puede extenderse durante un período de 5 a 15 años. Dado que el enfoque de sistemas de modelación de cultivos integra diferentes componentes de los ecosistemas agrícolas, este podría ser utilizado para llevar a cabo experimentos de simulación en sitios específicos con el fin de comprender la interacción de un determinado genotipo con el ambiente. Tales estudios pueden ayudar a reducir el número de sitios requeridos para la evaluación de campo y así aumentar la eficiencia del proceso de desarrollo de nuevas variedades.

### **1.4.2 Modelos de simulación para el cultivo de arroz**

Los primeros trabajos sobre la modelación del crecimiento y desarrollo del cultivo de arroz iniciaron hace más de 30 años (B. A. M. Bouman & Van Laar, 2006). En 1983, el Centro Internacional de Investigación sobre el Arroz (IRRI) publicó el modelo de producción potencial RICEMOD para sistemas bajo secano (Mennamy & O'toole, 1983). A pesar de que este modelo incluía aspectos físicos y fisiológicos del cultivo, se consideró una

herramienta bastante simple y no recibió mucha atención. Horie, Nakagawa, & Centeno (1995) desarrollaron un modelo de procesos para el cultivo de arroz bajo condición de producción potencial, SIMRIW, el cual se usó para la predicción del rendimiento en algunas regiones de Japón y para predecir el efecto del cambio climático sobre el cultivo. Fukai et al. (2000, 1995) desarrolló un modelo simplificado para el cultivo de arroz, RLRice, específicamente para la variedad KDML105, cultivada en el norte de Tailandia. Otros dos modelos de crecimiento que integran múltiples procesos fisiológicos para el cultivo de arroz son WOFOST (Hijmans, Lens, & Diepen, 1994) y MACROS (Svirezhev, 1992). Ambos modelos simulan el crecimiento y desarrollo del cultivo en condiciones de producción potencial y bajo limitaciones de agua. Sin embargo, hasta el momento, no se ha publicado ninguna evaluación específica para el modelo MACROS, mientras que para WOFOST, Roetter et al. (1998) reportaron algunos resultados de calibración y validación del modelo, principalmente bajo condiciones de producción potencial utilizando datos experimentales de Filipinas.

Ceres-Rice es un modelo dinámico de simulación del cultivo de arroz, que hace parte de la plataforma DSSAT (Jones et al., 2003). Este modelo contiene una descripción detallada del crecimiento del cultivo bajo producción potencial, y bajo limitaciones de agua y Nitrógeno. Ceres-Rice ha sido ampliamente utilizado a nivel mundial, en investigaciones que involucran procesos de calibración y validación utilizando datos experimentales, así como estudios de evaluación de prácticas de manejo agronómico y el impacto de la variabilidad y el cambio climático sobre el cultivo de arroz (Hoang, Ngoc, & Maskey, 2016; JIANG & JIN, 2009; Kadiyala, Jones, Mylavarapu, Li, & Reddy, 2015; Satapathy, Swain, & Herath, 2014).

A mediados de la década de los 90, el IRRI y la Universidad de Wageningen desarrollaron la serie de modelos ORYZA con el fin de simular el crecimiento y desarrollo del arroz de riego en zonas tropicales (Boling, Bouman, Tuong, Murty, & Jatmiko, 2006). El primer modelo que se implementó fue ORYZA1 para producción potencial (Kropff, H.H. Laar, & Matthews, 1994), seguido por ORYZA\_W para sistemas de producción bajo limitaciones de agua (Wopereis et al., 1996), y después por ORYZA-N (Drenth, ten Berge, & Riethoven, 1994) y ORYZA1N (Aggarwal, Kropff, Cassman, & ten Berge, 1997) para sistemas de producción limitados por Nitrógeno. Kropff et al (1994) reportaron un caso de evaluación de ORYZA1, que se centró en la comparación entre los valores observados y simulados del rendimiento y la biomasa total del cultivo. Por su parte, Wopereis et al (1996) reportaron

un caso de validación del modelo ORYZA\_W utilizando un set de experimentos realizados en el IRRI, el cual les permitió concluir que dicho modelo se comportaba de manera satisfactoria. En cuanto a ORYZA-N, este modelo se parametrizó usando siete ensayos experimentales, para los cuales se consideraron diferentes variedades, años y sitios de evaluación. El primer caso de evaluación de ORYZA-N fue reportado por Drenth et al (1994). El modelo ORYZA1 se derivó parcialmente de ORYZA-N, el primer caso de evaluación fue reportado por Aggarwal et al (1997), quienes utilizaron tres ensayos experimentales realizados en el IRRI. En el año 2001, se lanzó una nueva versión de la serie de modelos ORYZA, la cual integró todas las versiones anteriores en un solo modelo llamado ORYZA2000 (B. a M. Bouman et al., 2001). Reconociendo la importancia de evaluar este nuevo modelo, se llevaron a cabo un gran número de ensayos experimentales considerando varios genotipos, distintos sistemas de producción (Potencial, limitados por agua y limitados por Nitrógeno), así como diversas condiciones ambientales (Filipinas, Indonesia y China).

Actualmente el desarrollo de modelos de sistemas productivos para el cultivo de arroz continúa evolucionando gracias al esfuerzo de muchas organizaciones en todo el mundo. Los investigadores están cada vez más interesados no solo en contribuir a la ciencia de los sistemas agrícolas (por ejemplo, el Proyecto de Intercomparación y Mejoramiento de Modelos Agrícolas (AgMIP - <http://www.agmip.org/>), algunos programas del Consorcio Global para la Investigación Agrícola (CGIAR - <http://www.cgiar.org/>), y el programa sobre Cambio Climático, Agricultura y Seguridad Alimentaria (CAAFS - <https://ccafs.cgiar.org/>)), sino al uso de los modelos de cultivos para proveer información confiable, útil y de fácil acceso, tanto a agricultores como al sector privado y a entidades gubernamentales, con el fin de apoyar la toma de decisiones a diferentes niveles de la sociedad.

Un estudio reportado por Li et al. (2015) muestra algunos de los resultados alcanzados por el equipo de modelación de arroz del programa AgMIP, el cual tuvo como objetivo general estimar el nivel de incertidumbre en la predicción del rendimiento y la biomasa, y validar la hipótesis acerca del uso de un ensamble de varios modelos de simulación para reducir el nivel de incertidumbre. Para dicha investigación se emplearon los 13 modelos de simulación para el cultivo de arroz que son utilizados con mayor frecuencia en la actualidad (Tabla 1). A pesar de que estos modelos han sido desarrollados teniendo en cuenta los principios fisiológicos que modulan el crecimiento y desarrollo del cultivo, muchos de ellos difieren en su estructura, variables forzantes y parámetros de entrada (Tabla 2). Dentro de

las razones que explican su alta tasa de usabilidad entre la comunidad científica y académica, se destaca el hecho de que todos estos modelos generan información relacionada con la duración de la etapa de crecimiento, la dinámica en producción de biomasa, y el rendimiento final del cultivo. Cada uno de estos 13 modelos ha sido clasificado de acuerdo a la forma como modelan los siguientes tres subprocesos, que son de gran relevancia para determinar la respuesta del cultivo a las condiciones agroambientales en las que se desarrolla:

- *Fenología o duración del periodo de crecimiento:* Todos los modelos usan la temperatura como principal factor forzante para simular el progreso del estado de desarrollo del cultivo, no obstante, la forma funcional que utilizan para determinar las tasas de desarrollo es diferente. Algunos de ellos consideran una función de respuesta a la temperatura de tipo lineal o no lineal, con un límite superior o asintótico para la tasa de desarrollo. Por su parte, existen otros que consideran una función con una temperatura óptima, para la cual la tasa de desarrollo disminuye proporcionalmente con el aumento de la temperatura.
- *Producción de biomasa:* Generalmente la producción de biomasa en estos modelos es modulada por la radiación solar y la tasa de CO<sub>2</sub>, sin embargo, algunos de ellos también consideran la temperatura (Tabla 2). Teniendo en cuenta el nivel de complejidad de los procesos fisiológicos que existen dentro de la estructura modular para simular la producción de biomasa, estos modelos se han clasificado en tres categorías, de la siguiente manera: 1) Modelos basados en la eficacia del uso de la radiación (RUE, por sus siglas en inglés) por parte del dosel del cultivo (Monteith & Moss, 1977). 2) Modelos basados en la curva de respuesta de la fotosíntesis a la intensidad de luz recibida por una hoja y su integración en el dosel (LRC) (Woodhead, 1979). 3) Aquellos basados en el modelo bioquímico de Farquhar-von Caemmerer-Berry sobre la fotosíntesis de la hoja y su integración en el dosel (FvCB) (Farquhar, Caemmerer, & Berry, 1980).
- *Fertilidad de la espiga en respuesta a la temperatura:* La sensibilidad del rendimiento puede variar substancialmente, dependiendo de la manera en que el modelo considera la esterilidad de la espiga, ya que esta es una de las principales fuentes de variación en el rendimiento bajo condiciones climáticas extremas. Por esta razón los modelos de simulación se clasifican teniendo en cuenta si consideran

o no la esterilidad en la espiga debido a un cambio inducido en la temperatura (Tabla 2). Es de esperarse que la esterilidad de la espiga disminuya el tamaño del grano y el índice de cosecha en el cultivo de arroz (Li et al., 2015).

Tabla 1. Nombre del modelo, institución de afiliación, y referencia a seguir para una explicación más detallada de cada uno de los modelos.

<b>Modelo</b>	<b>Afiliación</b>	<b>Contacto</b>	<b>*Ref.</b>
APSIM-ORYZA	CSIRO Ciencias de los ecosistemas, Australia.	Donald Gaydon	1
CERES-Rice	Instituto internacional de desarrollo de fertilizantes, USA.	Upendra Singh	2
DNDC-Rice	Instituto nacional para las ciencias agro-ambientales, Japón.	Tamon Fumoto	3
GECROS	Universidad de Wageningen, Holanda.	Xinyou Yin	4
GEMRICE	Organización nacional de investigaciones en agricultura y alimentos, Japón.	Hiroe Yoshida	5
McWLA	Academia China de las Ciencias, Instituto de ciencias geográficas y de investigación de los recursos naturales, China.	Fulu Tao	6
Hasegawa/Horie (H/H)	Instituto nacional para las ciencias agro-ambientales, Japón.	Toshihiro Hasegawa	7
ORYZA2000	Centro Internacional de Investigación sobre el Arroz, Filipinas.	Tao Li	8
RiceGrow	Universidad Agrícola de Nanjing, China.	Liang Tang	9
SAMARA	CIRAD, UMR AGAP, Francia.	Philippe Oriol	10
SIMRIW	Organización nacional de investigaciones en agricultura y alimentos, Japón	Hiroshi Nakagawa	11
STICS	Instituto nacional Francés para la investigación agrícola, Francia.	Francoise Ruget	12
WARM	Universidad de Milán, Italia.	Roberto Confalonieri	13

\* 1: D. Gaydon et al. (2012 a,b); 2: Singh et al. (1993); 3: Fumoto et al. (2008); 4: Yin & van Laar (2005); 5: Yoshida & Horie (2010); 6: Tao & Zhang (2013); 7: Hasegawa & Horie (1997); 8: B. A. M. Bouman & Van Laar (2006); 9: Tang et al. (2009); 10: Dingkuhn et al. (2016); 11: Horie et al. (1995); 12: Brisson et al. (1998); 13: Confalonieri et al. (2009).

Tabla 2. Variables forzantes y procesos fisiológicos simulados en cada modelo.

**N.A.** – Indica los modelos que no aplican para una determinada variable forzante; **RUE, LRC, FvCB** – Los tres métodos principales usados para simular la producción neta de biomasa foliar: 1) Uso eficiente de la radiación, 2) la curva de respuesta de la fotosíntesis a la intensidad de luz recibida por una hoja y su integración en el dosel, 3) modelo bioquímico de Farquhar-von Caemmerer-Berry sobre la fotosíntesis de la hoja y su integración en el dosel; **ND y PD** – Dos tipos de funciones que relacionan el desarrollo fenológico con la temperatura – 1) Función de respuesta a la temperatura de tipo lineal o no lineal, con un límite superior o asintótico para la tasa de desarrollo, 2) Función que considera una temperatura óptima, para la cual la tasa de desarrollo disminuye proporcionalmente con el aumento de la temperatura. **S y NS** – Presentan formación del rendimiento – 1) Modelos que cuantifican la esterilidad de la espiga debido a bajas y altas temperaturas, así como a condiciones de sequía, 2) No cuantifican la esterilidad; **PET y AET** – Denota la evapotranspiración potencial y actual, respectivamente; **N** – Es el Nitrógeno.

Modelo	CO2	Radiación Solar	Temperatura	Humedad	Velocidad del viento	Precipitación y/o Riego	Entrada de Nitrógeno
AFSIM-ORYZA	Fotosíntesis (LRC).	Fotosíntesis, PET, máximo valor del área foliar.	Fenología (PD), área foliar, respiración, Fotosíntesis, rendimiento (S).	PET, balance de agua en el suelo.	PET, balance de agua en el suelo.	Balance de agua en el suelo, absorción de agua, transpiración actual, crecimiento de la hoja y senescencia, fotosíntesis, fertilidad de la espiga.	Absorción de Nitrógeno (N), fotosíntesis, senescencia de la hoja.
CERES-RICE	Fotosíntesis (RUE), transpiración (resistencia estomática de la hoja).	Fotosíntesis, PET.	Fenología (PD), área foliar, producción de biomasa, transpiración, número de tallos, senescencia de la hoja, rendimiento (S).	PET	PET	Balance de agua en el suelo, transpiración actual indirectamente afecta la producción de biomasa.	Absorción de N, indirectamente afecta la producción de biomasa, área foliar.
DNDC-Rice	Fotosíntesis (FvCB), conductancia estomática.	Fotosíntesis, balance de calor en el suelo y en el agua.	Fenología (PD), área foliar, respiración, fotosíntesis, tasa de crecimiento del grano (NS).	Se asume una humedad relativa constante.	Transpiración, balance de calor en el suelo y en el agua.	Balance de agua en el suelo, absorción de agua, transpiración actual.	Absorción de N, fotosíntesis, área foliar.
GECCROS	Fotosíntesis (FvCB), transpiración.	Fotosíntesis, PET.	Fenología (PD), fotosíntesis, respiración, transpiración, NS.	Conductancia estomática, transpiración, fotosíntesis.	Conductancia de la capa límite, transpiración, temperatura de la hoja.	Balance de agua en el suelo.	Absorción de N, área foliar, senescencia, distribución raíz: vástago, fotosíntesis, respiración, transpiración.
GEMRICE	Fotosíntesis (LRC), fenología, susceptibilidad a altas temperaturas que generan esterilidad en la espiga.	Fotosíntesis	Fenología (ND), área foliar, respiración, fotosíntesis, absorción de N, tasa de crecimiento de los órganos sumideros (S).	N.A.	N.A.	N.A.	Absorción de N, área foliar, fotosíntesis, respiración, número de espigas, acumulación de carbohidratos no estructurados.

MCWLA	Fotosíntesis (FvCB), conductancia estomática, transpiración.	Fotosíntesis, PET.	Fenología (ND), área foliar, respiración, fotosíntesis, formación del rendimiento (S).	PET, balance de agua en el suelo, conductancia estomática.	PET, balance de agua en el suelo.	Balance de agua en el suelo, absorción de agua, transpiración actual.	Fotosíntesis, área foliar.
H/H	Fotosíntesis (FvCB), conductancia estomática.	Fotosíntesis, balance de calor en el suelo y en el agua.	Fenología (ND), área foliar, respiración, fotosíntesis, tasa de crecimiento del grano (NS).	Conductancia estomática, balance de calor en el suelo y en el agua.	Balance de calor en el suelo y en el agua.	N.A.	Absorción de N, área foliar, fotosíntesis.
ORYA2000	Fotosíntesis (LRC).	Fotosíntesis, PET, máximo valor del área foliar.	Fenología (PD), área foliar, respiración, fotosíntesis, formación del rendimiento (S).	PET, balance de agua en el suelo.	PET, balance de agua en el suelo.	Balance de agua en el suelo, absorción de agua, transpiración actual, crecimiento de la hoja y senescencia, fotosíntesis, fertilidad de la espiga	Absorción de N, fotosíntesis, senescencia de la hoja.
RiceGrow	Fotosíntesis (LRC).	Fotosíntesis.	Fenología (PD), área foliar, respiración, fotosíntesis, formación del rendimiento (S).	N.A.	N.A.	Balance de agua en el suelo, absorción de agua, transpiración actual, distribución de asimilados, fotosíntesis.	Absorción de N, fotosíntesis, distribución de asimilados.
SAMARA	Fotosíntesis (RUE), transpiración.	Fotosíntesis, PET, PET también afecta el enrollamiento de la hoja.	Fenología (PD), área foliar, respiración, fotosíntesis, transpiración y evaporación (PET), formación del rendimiento (S).	AET a través de PET.	AET a través de PET.	Balance de agua del suelo, transpiración actual, fotosíntesis, enrollamiento de la hoja, fenología, esterilidad de la espiga.	N.A.
SIMRIW	Fotosíntesis (RUE), fenología, susceptibilidad a altas temperaturas que inducen a la esterilidad de la espiga.	Fotosíntesis.	Fenología (ND), área foliar, índice de cosecha (S).	N.A.	N.A.	N.A.	N.A.
STICS	Fotosíntesis (RUE), transpiración (Conductancia estomática).	Fotosíntesis.	Fenología (PD), establecimiento del índice de área foliar, producción de biomasa, número de granos, velocidad del llenado de grano (NS).	PET, balance de agua en el suelo.	ET potencial, balance de agua en el suelo.	Balance de agua en el suelo, absorción de agua y estrés, absorción de N.	Absorción de N, producción de biomasa, establecimiento del índice de área foliar.
WARM	Fotosíntesis (RUE).	Fotosíntesis, meteorología dentro de la cobertura vegetal.	Fenología (PD), fotosíntesis, área foliar, formación del rendimiento (S).	Meteorología dentro de la cobertura vegetal.	Meteorología dentro de la cobertura vegetal.	Balance de agua en el suelo, transpiración actual.	N.A.

### 1.4.3 Modelo de cultivo ORYZA2000

Los modelos de cultivo ORYZA1 (para producción potencial), ORYZA-W (para producción bajo limitaciones de agua), y ORYZA-N (para producción bajo limitación de Nitrógeno) están integrados en un solo modelo llamado ORYZA2000 (B. a M. Bouman et al., 2001). Los componentes de este modelo, para condiciones de producción potencial, están basados en el concepto SUCROS, que consiste en un módulo para producción potencial, con o sin módulos que cuantifican las limitaciones de agua, nitrógeno, y/o los efectos de plagas, enfermedades y malezas (Spitters & Keulen, 1989). Dentro de las características especiales incorporadas en ORYZA2000, se destacan la capacidad de cuantificar los efectos del trasplante y el estrés hídrico en el desarrollo fenológico, producción de biomasa, área foliar y rendimiento. Respecto al cálculo del balance hídrico, ORYZA incluye los siguientes tres modelos: **SAHEL** (Específico para suelos bajo sistemas de secano, sin encharcamiento y con drenaje libre hasta un determinado nivel freático; Penning de Vries, 1989), **LOWBAL** (considera suelos encharcados, para sistemas de arroz bajo riego), y **PADDY** (desarrollado para cualquier tipo de suelo dedicado al cultivo de arroz, como también considera el proceso de ascenso capilar de agua en el suelo). Adicionalmente, ORYZA tiene en cuenta el balance total de N en el cultivo de arroz, para poder simular su respuesta ante cualquier tipo de limitación en el cultivo, a causa de este elemento.

#### 1.4.3.1 Desarrollo fenológico en ORYZA2000

La etapa de desarrollo (DVS, por sus siglas en inglés) de una planta define su edad fisiológica y está caracterizada por la formación de los diversos órganos y su apariencia. El cambio fenológico más importante es el que se presenta entre la fase vegetativa a la reproductiva, ya que determina el cambio en la distribución de la materia seca sobre los diferentes órganos. Debido a que muchos procesos fisiológicos y morfológicos cambian de acuerdo a la fase fenológica de la planta, la cuantificación exacta del desarrollo fenológico es esencial en cualquier modelo que simula el crecimiento de un cultivo. Las etapas clave del desarrollo fenológico del arroz son la emergencia (DVS= 0), la iniciación de la panícula (DVS= 0.65), la floración (DVS= 1) y la madurez fisiológica (DVS= 2) (B. a M. Bouman et al., 2001).

**Temperatura efectiva para el desarrollo fenológico:**

La temperatura es considerada el mayor factor que impulsa el desarrollo fenológico (Van Oort, De Vries, Yoshida, & Saito, 2015). Se ha observado en muchos cultivos que la velocidad de desarrollo está linealmente relacionada con la temperatura media diaria, superior a una temperatura base hasta alcanzar una temperatura óptima. Posteriormente, la velocidad de desarrollo disminuye linealmente hasta alcanzar una temperatura máxima (Boote, Jones, White, Asseng, & Lizaso, 2013). Para temperaturas por debajo de la temperatura base o superiores a la temperatura máxima, la tasa de desarrollo es cero. Por lo tanto, se pueden identificar tres temperaturas "cardinales": Temperatura base ( $T_{base}$ , TBD, °C), temperatura óptima ( $T_{opt}$ , TOD, °C) y temperatura máxima ( $T_{max}$ , TMD; °C). Para el arroz, estos valores son típicamente 8, 30 y 42°C, respectivamente (Sánchez, Rasmussen, & Porter, 2014). Esta respuesta "bilineal" sólo se observa cuando las temperaturas diarias son constantes (por ejemplo, en un entorno controlado); si la temperatura oscila entre un valor mínimo y un valor máximo, como ocurre en el campo, la respuesta se vuelve más curvilínea, particularmente cerca de cada temperatura cardinal (Figura. 3).

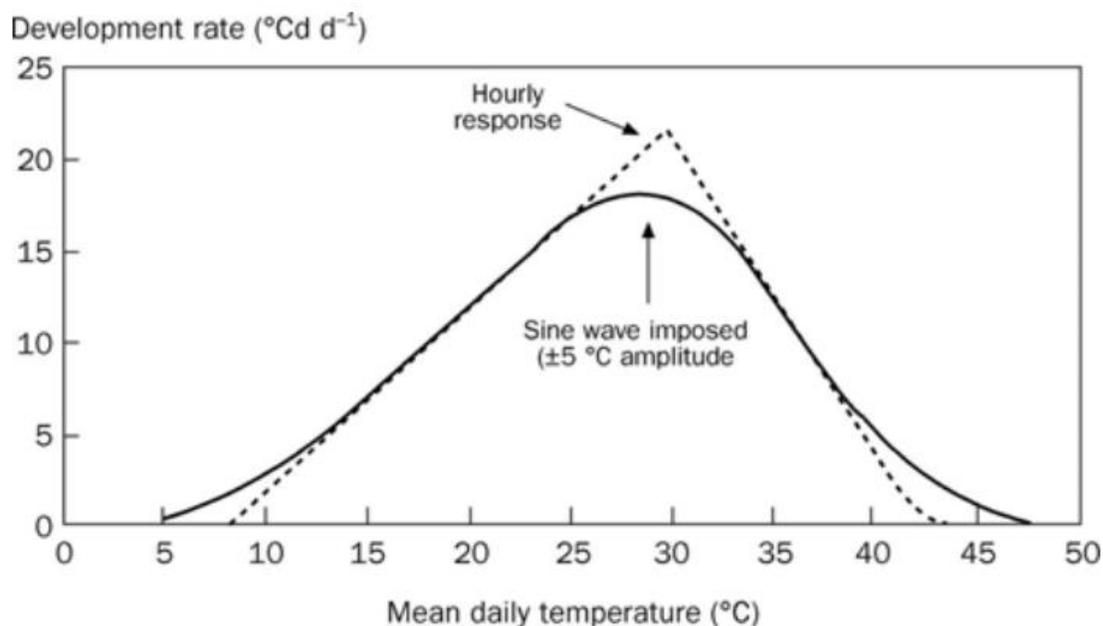


Figura. 3. Función de respuesta de la tasa de desarrollo fenológico a la temperatura media diaria. Simulación con  $T_{base} = 8^{\circ}\text{C}$ ,  $T_{opt} = 30^{\circ}\text{C}$ , y  $T_{max} = 42.5^{\circ}\text{C}$ .

Aunque esta respuesta curvilínea a la temperatura media diaria puede ser descrita por ecuaciones exponenciales complejas (por ejemplo, Xinyou 1996), en ORYZA1 se implementó un método sencillo usado por Matthews & Hunt (1994) en su modelo de yuca.

### ***Tasa de desarrollo fenológico:***

La tasa de desarrollo del cultivo se calcula con base en las constantes de velocidad de desarrollo para las diferentes etapas fenológicas, el incremento diario en unidades térmicas (HU; ° C día<sup>-1</sup>) y el fotoperiodo.

El ciclo de vida del cultivo de arroz se divide en cuatro fases fenológicas principales:

1. *La fase vegetativa básica (BVP):* Desde la emergencia (DVS = 0) hasta el inicio de la fase sensible al fotoperiodo (DVS = 0.4). La constante de velocidad de desarrollo en esta fase se denomina como DVRJ.
2. *Fase sensible al fotoperiodo (PSP):* Desde el final de la fase vegetativa básica hasta la iniciación de la panícula (DVS = 0,65). La constante de velocidad de desarrollo en esta fase se define como DVRI.
3. *Fase de formación de la panícula (PFP):* Desde la iniciación de la panícula hasta la floración (50%) (DVS = 1). La constante de velocidad de desarrollo en esta fase se denomina DVRP.
4. *Fase de llenado de grano (GFP):* Desde (50%) floración hasta la madurez fisiológica (DVS = 2). La constante de velocidad de desarrollo es DVRR.

El fotoperiodo se calcula a partir de la duración del día para tener en cuenta el efecto de niveles bajos de radiación después del atardecer y antes del amanecer. Cada una de estas cuatro fases tiene una constante de velocidad de desarrollo específica para una determinada variedad, la cual se calcula como la inversa de la suma de temperatura requerida para completar una fase específica, bajo condiciones óptimas de fotoperiodo. Las diferencias entre las variedades respecto a la duración total de la etapa de crecimiento, suelen ser causadas por diferencias en la duración de la fase vegetativa básica (B. a M. Bouman et al., 2001).

En el arroz trasplantado, la situación se vuelve más complicada debido al estrés ocasionado por el trasplante, esta situación provoca un retraso en el desarrollo fenológico. A través de ensayos experimentales diseñados para estudiar dicho efecto, se encontró que el retraso en el desarrollo fenológico depende de la edad de las plántulas que se trasplantan, la cual se expresa en grados días (TSTR; °Cd). En ORYZA1, el retraso se expresa en grados días (TSHCKD; °Cd), indicando el período después del trasplante durante el cual no se produce ningún desarrollo.

### **Estrés por sequía y su efecto en las tasas de desarrollo**

La sequía en la etapa vegetativa del cultivo retrasa la floración (Turner, O'Toole, Cruz, Namuco, & Ahmad, 1986; S. Yoshida, 1981). Wopereis et al (1996) encontraron que el retraso en la floración disminuyó cuando la sequía se presentó durante las últimas etapas del cultivo. Adicionalmente, identificaron que el aplazamiento de la floración estuvo acorde con el número de días entre la fecha de expansión de la hoja cero y la recuperación de la sequía, indicando que, si el suelo está demasiado seco para producir nuevas hojas, la tasa de desarrollo del cultivo también se ve interrumpida. Por lo tanto, en el modelo ORYZA1, el factor de reducción de la expansión foliar LESTRS se utiliza para simular el efecto de la sequía sobre el retraso de la floración. El valor de LESTRS es 1 en condiciones de producción potencial. Cuando ORYZA2000 se ejecuta considerando el balance hídrico para simular condiciones de producción limitadas por agua, el valor de LESTRS se calcula mediante la subrutina WSTRESS de ORYZA1 (B. a M. Bouman et al., 2001).

El retraso en la floración se simula a través de un factor de reducción (DVEW) que actúa sobre la tasa de desarrollo (DVR). Después de la floración (DVS = 1), DVEW es 1 y no se produce más reducción en la expansión de las hojas. Antes de la floración, DVEW es igual al factor de reducción de la expansión foliar LESTRS, el cual toma valores más bajos a medida que el estrés por sequía ocurre cerca a la floración.

### 1.4.3.2 Distribución de materia seca

En ORYZA2000, cualquier materia seca producida por el cultivo es distribuida entre la raíz y la parte aérea de acuerdo a un coeficiente de distribución ( $P_c$ , Kg de materia seca de un órgano  $\text{Kg}^{-1}$  materia seca del cultivo) definido como función del estado fenológico del cultivo ( $D$ ):

$$P_{cK} = f(D) \text{ Ecu. 1}$$

FSH es la fracción de materia seca distribuida a la parte aérea de la planta y FRT es la fracción de materia seca distribuida en las raíces. El coeficiente de distribución FSH se calcula a partir de la interpolación realizada con los datos experimentales de materia seca de los diferentes órganos. Esta información es suministrada en una tabla de registros (FSHTB) que lee ORYZA1 y la procesa usando la función LINT2.

La distribución de carbohidratos entre la parte aérea de la planta y la raíz, bajo estrés hídrico, se altera generalmente a favor de la biomasa en las raíces. Cuando las hojas dejan de expandirse, la fotosíntesis continúa y el nivel de carbohidratos de reserva aumenta, lo que los hace más disponibles para el crecimiento del sistema radicular (B. a M. Bouman et al., 2001). Por lo tanto, en ORYZA1, el efecto de la sequía sobre la distribución de asimilados entre la parte aérea y la raíz se calcula a partir del factor de expansión foliar, LESTRS. El valor de LESTRS es 1 en condiciones de producción potencial. Cuando ORYZA2000 se ejecuta considerando balance hídrico para situaciones de producción limitadas por agua, el valor de LESTRS se calcula mediante la subrutina WSTRESS.

La materia seca asignada a la parte aérea del cultivo se divide en diferentes órganos de la siguiente manera: FLV es la fracción de materia seca dividida en hojas, FST en los tallos y FSO en las panículas (todos como una fracción de la materia seca total "aérea" de la planta).

### 1.4.3.3 Pérdida de hojas verdes y reservas en los tallos

La tasa de pérdida de hojas se calcula teniendo en cuenta el peso de hojas verdes (WLVG, kg de hojas ha<sup>-1</sup>) y una tasa de mortalidad relativa de las hojas. Esta tasa de pérdida relativa es afectada por el nivel de nitrógeno del cultivo (Diepen, Keulen, & Vries, 1987). En ORYZA1, este proceso se calcula multiplicando la tasa de pérdida de hojas por el denominado factor de estrés por Nitrógeno (NSLLV), el cual acelera la muerte foliar. El valor de NSLLV es 1 bajo condiciones de producción potencial. Cuando ORYZA2000 se ejecuta considerando un balance de nitrógeno con el fin de simular una producción bajo limitaciones de este elemento, el valor de NSLLV se calcula a partir del contenido de nitrógeno del cultivo en la subrutina NCROP.

La tasa de pérdida de reservas en los tallos (LSTR, kg de biomasa de tallos ha<sup>-1</sup> d<sup>-1</sup>) inicia en la floración del cultivo, y se calcula dividiendo el peso de tallos (WST, kg de tallos ha<sup>-1</sup>) por un coeficiente de tiempo para la pérdida relativa de reservas en dichos órganos (TCLSTR).

### 1.4.3.4 Tasa de crecimiento del cultivo

La tasa de crecimiento neto diario del cultivo de arroz (Gp, GCR, Kg de materia seca ha<sup>-1</sup> d<sup>-1</sup>) se calcula de la siguiente manera:

$$G_p = \frac{A_d * \left(\frac{30}{44}\right) - R_m + R_t}{Q} \quad \text{Ecu. 2}$$

Donde Ad es a tasa neta de asimilación de CO<sub>2</sub> (DTGA; kg CO<sub>2</sub> ha<sup>-1</sup> d<sup>-1</sup>), Rm es la tasa de la respiración de mantenimiento (RMCR, kg CH<sub>2</sub>O ha<sup>-1</sup> d<sup>-1</sup>), Rt es la cantidad de reservas disponibles en el tallo para el crecimiento (LSTR; kg CH<sub>2</sub>O ha<sup>-1</sup> D<sup>-1</sup>), y Q son los asimilados requeridos para la producción de materia seca (CRGCR, kg de CH<sub>2</sub>O kg<sup>-1</sup> de materia seca).

La cantidad de reservas en el tallo que es redistribuida a otros órganos se multiplica por LRSTR (= 0.947) con el fin de considerar las pérdidas (5.3%) que se generan al momento de ser reasignadas (Penning de Vries, 1989). Estas reservas se expresan en CH<sub>2</sub>O, como resultado de la multiplicación entre la fracción de carbono presente en las reservas de los

tallos (FCSTR) y la relación de los pesos moleculares de CH<sub>2</sub>O y C (30/12), para convertir el carbono en asimilados que están disponibles para el crecimiento (Producción de materia seca). Finalmente, la tasa de crecimiento diario neto (NGCR, kg de materia seca ha<sup>-1</sup> d<sup>-1</sup>) es el valor total de asimilados menos la cantidad de reservas en el tallo que han sido redistribuidas.

### 1.4.3.5 Tasa de crecimiento de los órganos del cultivo

La tasa de crecimiento de un grupo de órganos de cultivo  $k$  ( $G_{p,k}$ ) se obtiene multiplicando la tasa de crecimiento total del cultivo ( $G_p$ , Ecu. 2, GCR) por la fracción asignada a ese grupo de órganos  $k$  ( $pc_k$ ):

$$G_{p,k} = pc_k * G_p \quad \text{Ecu. 3}$$

La materia seca disponible para el crecimiento se distribuye inicialmente hacia las raíces (GRT; kg de materia seca ha<sup>-1</sup> d<sup>-1</sup>), teniendo en cuenta la reducción de peso debido al trasplante (GRT1). La fracción de materia seca disponible para la parte aérea de la planta se divide entre hojas (GLV), tallos (GST) y órganos de almacenamiento (GSO) utilizando los factores de partición (FLV, FSH y FSO, respectivamente).

Para las hojas verdes y tallos, se tiene en cuenta la reducción en el peso debido al trasplante (RWLVG1 y GST1, respectivamente). La tasa de crecimiento neto de las hojas (RWLVG, kg de materia seca ha<sup>-1</sup> d<sup>-1</sup>) se determina a partir de la tasa de crecimiento del total de las hojas (GLV) menos su tasa de mortalidad (LLV). La tasa de crecimiento de los tallos se multiplica por (1 - FSTR) y proporciona información de la fracción de materia seca que es distribuida hacia los tallos (FST). La tasa de crecimiento del pool de reserva de los tallos (GSTR) se calcula de manera similar a la del FSTR. Justo antes de la floración (después de la etapa de desarrollo, DVS, 0.95), se asume que toda la materia seca distribuida en los órganos de almacenamiento (GSO), es redistribuida hacia los granos (GGR).

### 1.4.3.6 Espigas y formación del grano

La subrutina SUBGRN calcula la tasa de formación de las espigas (GNSP, número de espiguillas  $\text{ha}^{-1}$  suelo  $\text{d}^{-1}$ ) y granos (GNGR, número de granos  $\text{ha}^{-1}$  de suelo  $\text{d}^{-1}$ ), así como la fertilidad de las espigas en función de la temperatura alrededor de la floración del cultivo. La versión actual de esta subrutina también cuantifica el efecto de la sequía sobre la fertilidad de la espiga.

En las gramíneas, la producción de carbohidratos durante el período de llenado del grano puede ser mayor que la capacidad de almacenamiento de los mismos, la cual está determinada por el número de granos por  $\text{m}^2$  y la tasa de crecimiento máximo de los granos. Este proceso puede generar una sobre acumulación de asimilados en las hojas, causando tasas reducidas de asimilación de  $\text{CO}_2$  (Barnett & Pearce, 1983). Este proceso es de gran importancia para el cultivo de arroz cuando se cultiva en ambientes extremos con bajas o altas temperaturas antes de la floración, las cuales pueden inducir a la esterilidad de las espiguillas, lo que se traduce en una baja capacidad de almacenamiento de carbohidratos en las espigas y una menor tasa de producción de granos (S. Yoshida, 1981).

En el cultivo de trigo, se ha encontrado que el tamaño de la espiga durante la floración es proporcional al número de granos que se forman, y está estrechamente relacionado con la tasa de crecimiento del cultivo durante el período de formación de las espigas (Fischer, 1985). La cantidad de crecimiento de la planta durante este período depende tanto de la duración del período, que es influenciada por la temperatura, como de la tasa de crecimiento del cultivo, que está influenciada por la temperatura y la radiación. Para el cultivo de arroz se han encontrado relaciones similares (S. Yoshida & Parao, 1976) que han sido utilizadas para asociar el rendimiento del arroz con el coeficiente foto térmico “Q”, el cual se define como la relación entre la radiación solar incidente durante los últimos 30 días antes de la floración, y la temperatura media durante el mismo periodo, menos la temperatura base del cultivo (Islam & Morison, 1992).

A través de experimentos realizados en el IRRI, se encontró una buena relación entre el crecimiento total del cultivo, desde la iniciación de la panícula hasta el inicio de la floración, y el número de espigas en la floración (Figura. 4). Esta relación se mantiene durante las estaciones húmedas y secas, como también para niveles de aplicación de nitrógeno entre 0 a  $285 \text{ kg ha}^{-1}$ , para densidades de siembra entre 25 y  $125 \text{ plantas m}^{-2}$ , y para condiciones

severas de estrés por sequía. Una relación similar se encuentra a nivel de los tallos, ya que el número de espigas por cada tallo depende de la tasa de crecimiento del tallo durante el período de formación de la panícula. Los efectos negativos de la radiación solar, la temperatura, el nitrógeno, y el agua sobre la formación de las espigas, tienen mayor repercusión si un cambio severo se presenta durante el periodo de formación de la panícula. En ORYZA1, la pendiente que se obtiene de la Figura. 4 se conoce como el factor de formación de la espiga ( $\gamma$ ). En este modelo se hace un seguimiento de la tasa de crecimiento de la espiga, desde la iniciación de la panícula (es decir, a  $DVS = 0.65$ ) hasta el 50% de floración del cultivo ( $DVS = 1$ ), el número de espigas formadas ( $S$ , GNSP; número de espigas  $ha^{-1}$  de suelo  $día^{-1}$ ) se calcula como el producto entre esta tasa de crecimiento ( $G$ , GCR;  $kg$  de materia seca  $ha^{-1}$  suelo  $día^{-1}$ ) y el factor de formación de la espiga ( $\gamma$ , SPGF; número  $kg^{-1}$ ) (todo esto para un determinado día  $i$ ):

$$S_i = \sum_{i=P}^F (G_i * \gamma) \quad \text{Ecu. 4}$$

Donde  $P$  y  $F$  son las fechas de iniciación de panícula hasta la floración al 50%, respectivamente.

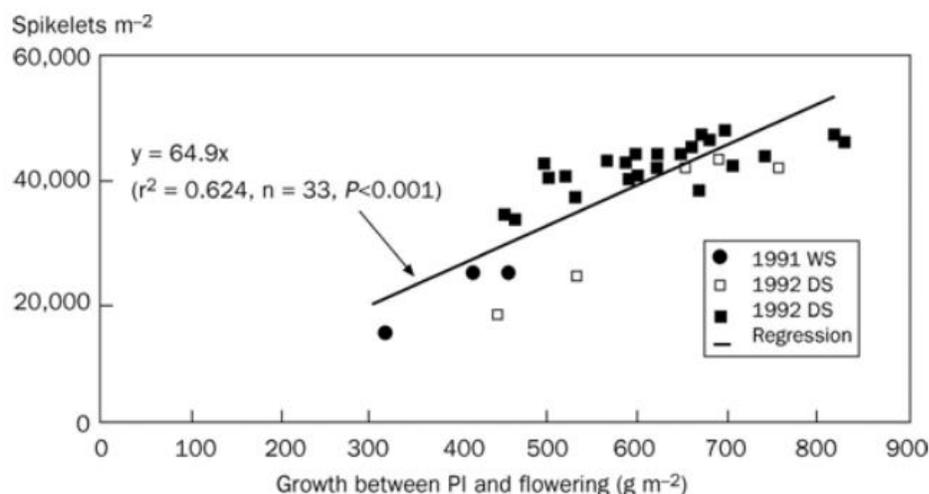


Figura. 4. Relación entre el número de espigas  $m^{-2}$  y el crecimiento del cultivo desde la iniciación panicular hasta la floración (datos de Kropff, Cassman, Torres, and Liboon (Símbolos ●, □) and S. Peng (■), IRRI, cultivar IR72) Tomado de B. a M. Bouman et al. (2001).

A medida que crece el cultivo, las espigas se convierten en granos. Sin embargo, algunas espiguillas se vuelven estériles debido a temperaturas demasiado altas o demasiado bajas y no se llenan. La esterilidad de la espiga se calcula de acuerdo con Horie et al (1992). La esterilidad causada por bajas temperaturas se basa en el concepto de "grados días de enfriamiento" (UCHIJIMA, 1976). Los grados días de enfriamiento ( $SQ_t$ , COLDTT; °C día) se calculan de la siguiente manera:

$$SQ_t = \sum (22 - T_d)$$

**Ecu. 5**

Donde  $T_d$  es el promedio de temperatura (corregido por el incremento de temperatura causado por la sequía). La sumatoria de  $SQ_t$  se determina para el periodo de alta sensibilidad de la panícula a bajas temperaturas ( $0.75 \leq DVS \leq 1.2$ ). La relación entre el porcentaje de esterilidad causado por el frío ( $S_c$ ; SF1) y la suma de los grados días de enfriamiento ( $SQ_t$ ) puede ser aproximado a través de la siguiente ecuación (Figura. 5):

$$S_c = 1 - \frac{4.6 + 0.054 * SQ_t^{1.56}}{100}$$

**Ecu. 6**

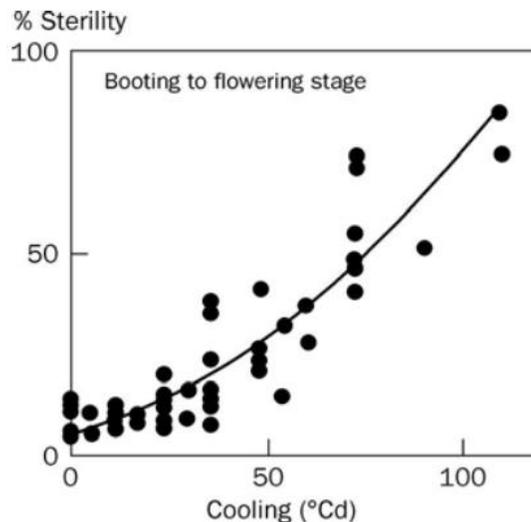


Figura. 5. Relación entre los grado días de enfriamiento y el porcentaje de esterilidad de la espiga en la variedad de arroz "Eiko", entre iniciación de panícula y floración (Horie et al., 1995).

Las espigas también son sensibles a la alta temperatura, particularmente durante el periodo de floración. Según Huang et al (2017) y Satake & Yoshida (1978) los daños en el polen ocurren cuando la temperatura durante la floración del cultivo supera los 35°C. La Figura. 6 representa la relación entre la fracción de espigas fértiles ( $S_h$ , SF2; -) y la temperatura máxima media diaria ( $T_{m,a}$  TFERT; °C) durante la floración del cultivo ( $0.96 \leq DVS \leq 1.22$ ), para la variedad Akihikari, la cual fue sembrada en una capsula experimental con variación gradual de la temperatura y la concentración de CO<sub>2</sub>. La Figura. 6 indica que la concentración de CO<sub>2</sub> no tiene algún efecto sobre la relación entre la temperatura y la fertilidad. Dicha relación se explica a través de la siguiente ecuación:

$$S_h = \frac{1}{1 + \exp(0.853(T_{m,a} - 36.6))} \quad \text{Ecu. 7}$$

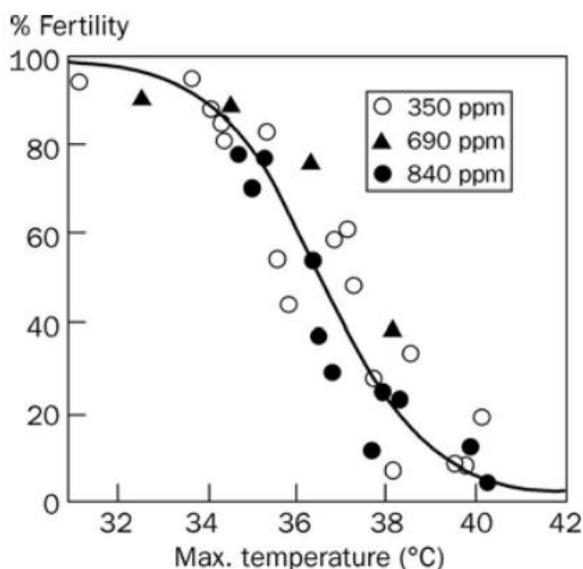


Figura. 6. Relación entre la temperatura máxima media diaria durante el periodo de floración y la fertilidad de la espiga en la variedad de arroz "Akihikari" a diferentes concentraciones de CO<sub>2</sub>.

El estrés por sequía durante la fase reproductiva del cultivo también aumenta el riesgo de que se presente esterilidad en la espiga, especialmente cuando se presenta cerca a la etapa de floración (EKANAYAKE, DATTA, & STEPONKUS, 1989).

Turner et al (1986) encontró una relación entre el incremento en la temperatura de la planta debido a la sequía y un aumento en la esterilidad de la espiga, además relacionó dicho aumento de temperatura con el porcentaje de hojas enrolladas. Ellos encontraron que un aumento de temperatura de 1.6°C podría generar el enrollamiento de todas las unidades foliares. En ORYZA2000 se usa esta relación existente entre hojas enrolladas, expresado a través de un factor de enrollamiento de las hojas ( $S_{r1}$ , LRSTRS; -), y el incremento de temperatura ( $T_i$ , TINCR; °C) para simular el incremento en la esterilidad de la espiga:

$$T_i = 5(1 - S_{r1})1.6 \quad \text{Ecu. 8}$$

El valor de LRSTRS es 1 bajo condiciones de producción potencial. Cuando ORYZA2000 se ejecuta considerando un balance hídrico para simular condiciones limitadas por agua, el valor de LRSTRS es calculado por la subrutina WSTRESS de ORYZA1 (B. a M. Bouman et al., 2001). La esterilidad actual de la espiga (SPFERT; -) se calcula como el mínimo valor de esterilidad causado ya sea por bajas o altas temperaturas. Finalmente, la tasa de crecimiento del número de granos (GNGR) se calcula multiplicando el factor de esterilidad de la espiga por la tasa de crecimiento del número de espigas (GNSP).

## 1.5 Análisis clúster

El Análisis Clúster (AC) es un método estándar del análisis multivariado que puede reducir una compleja cantidad de información a pequeños grupos o “clusters”, donde los elementos del mismo poseen características similares (Lin & Chen, 2006). El análisis clúster tiene por objeto formar grupos o clusters homogéneos en función de las similitudes o similitudes entre los objetos que los componen (Cuadras, 2014), los clasifica de tal forma que cada objeto es muy parecido a los que hay dentro del clúster con respecto a algún criterio de selección predeterminado (Heredia, Escobar, & Díaz, 2012).

El análisis clúster se considera una técnica eminentemente exploratoria que no utiliza ningún tipo de modelo estadístico para llevar a cabo el proceso de clasificación (Peterson, 2002), por ello se la podría calificar como una técnica de aprendizaje no supervisado, es

decir, una técnica muy adecuada para extraer información de un conjunto de datos sin imponer restricciones previas en forma de modelos estadísticos.

### **1.5.1 Clúster jerárquico aglomerativo**

Los llamados métodos jerárquicos tienen por objetivo agrupar clúster para formar uno nuevo, o separar alguno ya existente para dar origen a otros dos, de tal forma que, si sucesivamente se va efectuando este proceso de aglomeración o división, se minimice alguna distancia o bien se maximice alguna medida de similitud.

Los métodos aglomerativos, también conocidos como ascendentes, comienzan el análisis con tantos grupos como individuos haya. A partir de estas unidades iniciales se van formando grupos, de forma ascendente, hasta que al final del proceso todos los casos tratados están englobados en un mismo conglomerado (Luis & Villardón, 2007).

La forma en que generalmente operan estos métodos es bastante simple. Por ejemplo, en los métodos aglomerativos se parte de tantos grupos como individuos haya. A continuación se selecciona una medida de similitud, agrupándose los dos grupos o clúster con mayor similitud. Así se continúa hasta que:

1. Se forma un solo grupo.
2. Se alcanza el número de grupos prefijados.
3. Se detecta que hay razones estadísticas para no continuar agrupando clúster, ya que los más similares no son lo suficientemente homogéneos como para determinar una misma agrupación.

## **1.6 Redes Neuronales Artificiales (RNA)**

Una red neuronal artificial (RNA) es un sistema de procesamiento de información que se asemeja a la estructura del cerebro humano, inspirado en las neuronas biológicas. Su funcionamiento desde el punto de vista matemático, se basa en una gran cantidad de unidades de procesamiento simples (neuronas), trabajando en paralelo (ver Figura. 7).

Esta herramienta de análisis de información multidimensional o multivariada, tiene la capacidad de aprender a partir de datos históricos, inferir relaciones entre variables, reconocer patrones lineales y no lineales, tolerar datos erróneos, faltantes, o con presencia de ruido, y adaptándose a tendencias globales o locales, predecir el comportamiento del sistema en casos futuros o no conocidos. Es importante destacar que para su implementación no es necesario de un amplio conocimiento del fenómeno, ni de la relación entre sus variables (Barrios-Perez, 2012).

A diferencia de modelos determinísticos o incluso probabilísticos, donde es necesario un conocimiento detallado a priori del modelo matemático subyacente que controla el fenómeno, las redes neuronales son modelos tipo “caja negra”. Para poder aplicar preliminarmente una red neuronal, basta con disponer de una cantidad adecuada de datos (entradas del problema con salida conocida).

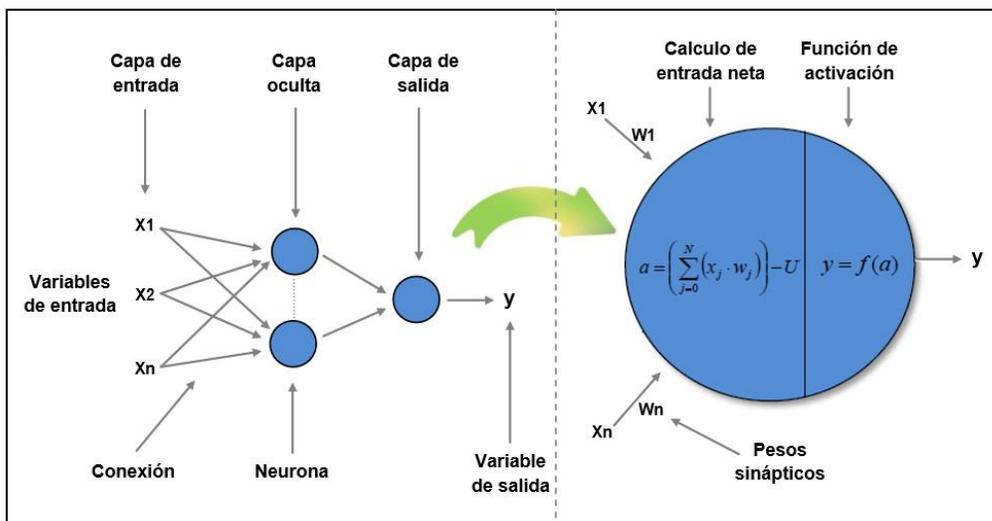


Figura. 7. Arquitectura de una Red Neuronal Artificial del tipo Perceptron Multicapa.

## 1.6.1 Métodos de aprendizajes de las RNA

### Aprendizaje supervisado:

El aprendizaje supervisado, cuyo esquema se presenta en la Figura. 8, se caracteriza porque el proceso de entrenamiento es controlado por un agente externo llamado supervisor o maestro. En la Figura. 8, el mundo real corresponde al problema a modelar que se presenta mediante un conjunto de datos de entrada y salida. El supervisor opera como un “maestro” que guía el aprendizaje en la red, y conoce las salidas deseadas correspondientes a las respectivas entradas.

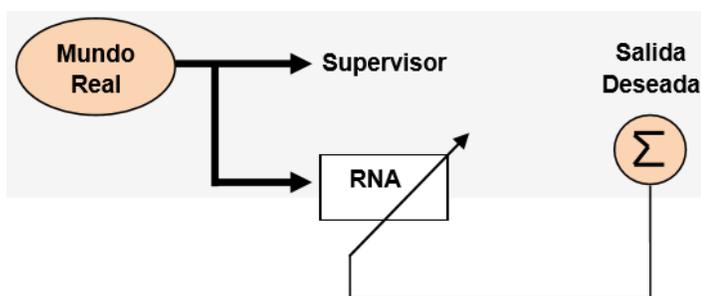


Figura. 8. Aprendizaje supervisado

### Aprendizaje no supervisado:

En el aprendizaje no supervisado, el vector de datos que describe el problema se le presenta directamente a la red, pero ahora ya no hay un supervisor o maestro que guía el aprendizaje. En este caso, los pesos de la red se calculan en función de la caracterización que se haga de la entrada que la red neuronal este recibiendo, de acuerdo a un objetivo específico que nos permite obtener el conocimiento de lo que queremos representar con la red. El esquema de este tipo de aprendizaje se presenta en la Figura. 9.

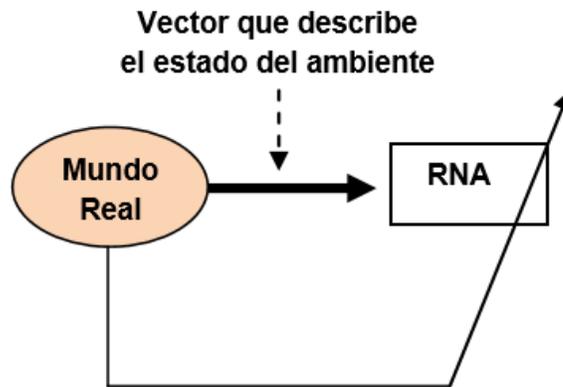


Figura. 9. Aprendizaje no supervisado.

### 1.6.2 Principales aplicaciones de las RNA

Existen 3 tipos principales de aplicaciones:

1. **Predicción:** Se puede entrenar una red neuronal utilizando datos conocidos, de forma que aprende y puede predecir el comportamiento del sistema ante otras entradas desconocidas. Esto es importante en numerosos problemas donde no se tiene un modelo confiable que describa el comportamiento del sistema pero sí se cuenta con datos, o sea, entradas para el problema con salida conocida.
2. **Clasificación y clústering:** Las RNA pueden ser entrenadas a partir de ejemplos, para dividir un conjunto en dos o más clases, de forma que puedan clasificar objetos nuevos como pertenecientes a una de las clases aprendidas. También se puede utilizar una red neuronal para dividir un conjunto en clases, sin información previa sobre la posición o composición de las clases (clústering).
3. **Interpolación y extrapolación de funciones:** Las redes son aproximadores universales. Pueden aproximar (casi) cualquier función sobre un dominio compacto, a la precisión deseada, soportando y filtrando leves imprecisiones y ruido en los datos.

### 1.6.3 Análisis de Relevancia Métrica

En la modelación con redes neuronales supervisadas se utilizan variables de entrada y salida para intentar describir el comportamiento de cualquier fenómeno. La concepción “caja negra” de las RNA no es un impedimento para conocer la relación entre variables, ni la importancia de una variable de entrada con respecto a una de salida. Las técnicas de relevancia métrica permiten calcular estas relaciones.

El uso de estas técnicas es similar al resultado de un análisis de componentes principales (ACP) en el sentido de disminuir el número de variables de entrada, pero difieren en que las técnicas de relevancia métrica determinan la importancia de variable de entrada con respecto a la variable de salida. Por su parte, el ACP calcula una combinación lineal de varias variables de entrada en los denominados componentes, sin tener en cuenta la influencia de éstas con la variable de salida (Barrios-Perez, 2012).

Varios investigadores han propuesto utilizar los parámetros de la arquitectura de las redes neuronales para inferir relaciones entre variables de entrada y salida. Satizábal & Pérez-Urbe (2007) hacen una extensa revisión de técnicas de relevancia métrica para reducir el número de variables de entrada. En la Tabla 3 se presenta un resumen de algunas técnicas de relevancia métrica. Por su parte, la Figura. 10 presenta los parámetros de la red neuronal utilizados en las técnicas de relevancia métrica. Es importante mencionar que la red neuronal debe estar entrenada, y su desempeño (en términos de ajuste y precisión) debe ser aceptable para que la aplicación de estas técnicas sea útil.

Según Satizábal & Pérez-Urbe (2007), existen dos grandes grupos de técnicas para determinar relevancia métrica de las variables: Uno que utiliza los parámetros internos de la red (pesos) para determinar la importancia de las conexiones entrada-salida, y otro que utiliza el set de datos de entrada para obtener una medida analítica o numérica de la sensibilidad de cada patrón de entrada con respecto a su salida. Las técnicas Garson's Algorithm, Overall Connections, y ANNIGMA utilizan los pesos de la arquitectura seleccionada para determinar la importancia de cada variable, mientras que la técnica Sensitivity Matrix utiliza los patrones de entrada y la arquitectura de la red para determinar la tasa de cambio de la variable de salida con respecto a cada variable de entrada.

Tabla 3. Características de algunas técnicas de relevancia

Nombre	Descripción	Formula	Resultado
Garson's Algorithm.	Suma fracciones de las diferentes conexiones que hay entre cada variable de entrada y la de salida con respecto a las demás conexiones entrada-salida de las variables de entrada. Este método involucra los pesos entre la capa de entrada y la oculta.	$R_{ik} = \sum_{j=1}^{nh} \left( \frac{ w_{ij}   w_{jk} }{\sum_{i'=1}^{ni} ( w_{i'j}   w_{jk} )} \right) = \sum_{j=1}^{nh} \left( \frac{ w_{ij} }{\sum_{i'=1}^{ni} ( w_{i'j} )} \right)$ <p>Donde <math>ni</math> y <math>nh</math> son el número de variables entradas y neuronas en la capa oculta respectivamente, <math>w_{ij}</math> es el peso entre la variable de entrada <math>i</math> y la neurona <math>j</math> de capa oculta, <math>w_{jk}</math> es el peso entre la neurona <math>j</math> de la capa oculta de la neurona de salida <math>k</math> (en este caso siempre <math>k = 1</math>).</p>	Una matriz de cantidades numéricas mayores a cero, de dimensión $k \times ni$ , donde $k$ es el número de variables de salida (en este caso siempre $k = 1$ ) y $ni$ es el número de variables de entrada.
Overall Connections	Suma las contribuciones positivas y negativas de las conexiones entrada-salida de cada variable de entrada. El signo indica la dirección de la contribución de cada entrada	$R_{ik} = \sum_{j=1}^{nh} w_{ij} \cdot w_{jk}$ <p>Donde <math>nh</math> es el número de neuronas en la capa oculta, <math>w_{ij}</math> es el peso entre la variable de entrada <math>i</math> y la neurona <math>j</math> de capa oculta, <math>w_{jk}</math> es el peso entre la neurona <math>j</math> de la capa oculta de la neurona de salida <math>k</math> (en este caso siempre <math>k = 1</math>).</p>	Una matriz de cantidades numéricas positivas y negativas, de dimensión $k \times ni$ , donde $k$ es el número de variables de salida (en este caso siempre $k = 1$ ) y $ni$ es el número de variables de entrada.
ANNIGMA	Es el acrónimo de <i>Artificial Neural Net Input Gain Measurement Approximation</i> . Suma las contribuciones absolutas de las conexiones entrada-salida de cada variable de entrada.	$R_{ik} = \sum_{j=1}^{nh}  w_{ij} \cdot w_{jk} $ <p>Donde <math>nh</math> es el número de neuronas en la capa oculta, <math>w_{ij}</math> es el peso entre la variable de entrada <math>i</math> y la neurona <math>j</math> de capa oculta, <math>w_{jk}</math> es el peso entre la neurona <math>j</math> de la capa oculta de la neurona de salida <math>k</math> (en este caso siempre <math>k = 1</math>).</p>	Una matriz de cantidades numéricas positivas, de dimensión $k \times ni$ , donde $k$ es el número de variables de salida (en este caso siempre $k = 1$ ) y $ni$ es el número de variables de entrada.
Sensitivity Matrix.	Expresa la tasa de cambio de la variable de salida con respecto a una variable de entrada en específico. El concepto es similar al de primera derivada de una función.	$R_{ik} = \sum_{p \in \text{Set datos}} \left  \frac{\partial O_k^p}{\partial I_i} \right  \quad \text{y} \quad \frac{\partial O_k^p}{\partial I_i} = \frac{\partial f_k}{\partial \alpha_k^p} \sum_{j=0}^{nh} w_{jk} \frac{\partial f_j}{\partial \alpha_j^p} w_{ij}$ <p>Donde <math>nh</math> es el número de neuronas en la capa oculta, <math>w_{ij}</math> es el peso entre la variable de entrada <math>i</math> y la neurona <math>j</math> de capa oculta, <math>w_{jk}</math> es el peso entre la neurona <math>j</math> de la capa oculta de la neurona de salida <math>k</math> (en este caso siempre <math>k = 1</math>), <math>\frac{\partial f_j}{\partial \alpha_j^p}</math> es la primera derivada de la función de activación de la neurona <math>j</math> de la capa oculta, <math>\frac{\partial f_k}{\partial \alpha_k^p}</math> es la primera derivada de la función de activación de la neurona <math>k</math> de la capa de salida (en este caso siempre <math>k = 1</math> y si la función de activación es lineal, esta derivada es 1).</p>	Una matriz de cantidades numéricas positivas que expresan la tasa de cambio de la variable de salida con respecto a una variable de entrada, de dimensión $k \times ni$ , donde $k$ es el número de variables de salida (en este caso siempre $k = 1$ ) y $ni$ es el número de variables de entrada.

Tomado de (Erazo-Meza, 2011).

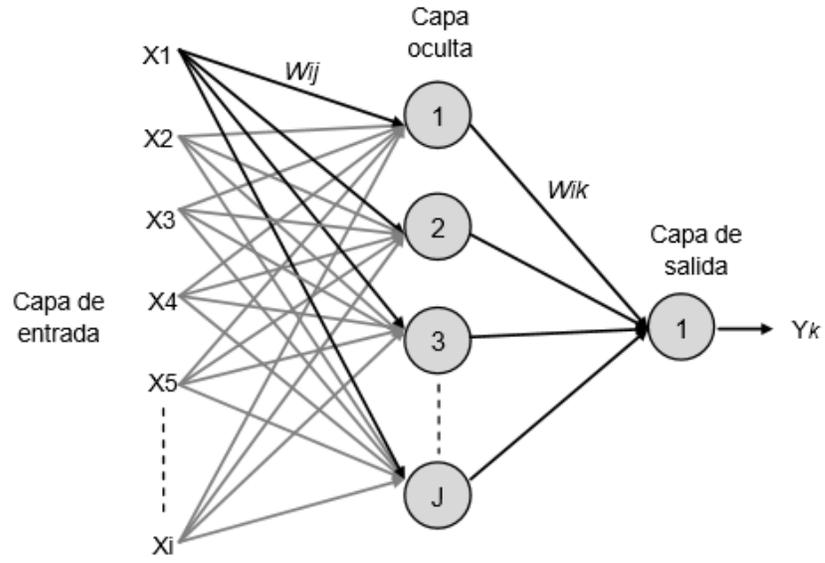


Figura. 10. Parámetros de una RNA utilizados en las técnicas de relevancia métrica.

## Capítulo 2.

### Materiales y métodos

“El que adquiere cordura a sí mismo se ama, y el que retiene el discernimiento prospera.”

**Proverbios 19:8 (NVI)**

#### 2.1 Caracterización del área de estudio

El área de estudio está distribuida entre los departamentos de Tolima y Huila en la zona Centro, y el departamento de Córdoba en la zona del Bajo Cauca. La región de estudio, ubicada entre los departamentos de Tolima y Huila, cubre un rango de altitud entre 250 y 1,300 metros sobre el nivel del mar (msnm), su ubicación astronómica varía entre 1°50' - 5°15' de latitud norte y 74°32' - 75°50' de longitud occidental. Por su parte, en el departamento de Córdoba la altitud de la región productora varía entre 5 y 500 msnm, y se distribuye entre 7°21' - 9°25' de latitud norte y 74°48' - 76°12' de longitud occidental (ver Figura. 11).

##### 2.1.1 Departamento del Tolima

En el departamento del Tolima las siembras del cultivo de arroz se centran entre los periodos Abril – Julio y Octubre – Enero. Durante el primer semestre del año, generalmente existe una participación cercana a 4,800 UPA de arroz, los cuales producen en promedio 375,000 toneladas de arroz, con un rendimiento de 7.0 ton/ha. En el segundo semestre,

el número de UPA aumenta a 5,200, de igual forma, la producción y el rendimiento aumentan a 381,000 ton y 7.2 ton/ha, respectivamente.

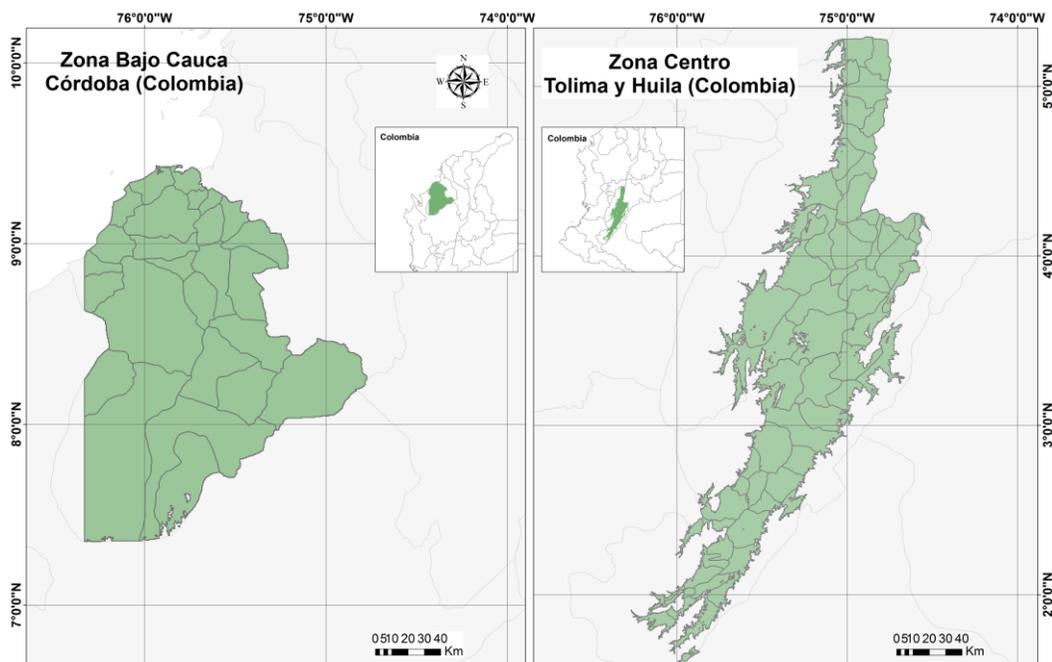


Figura. 11. Ubicación geográfica del área de estudio. Al lado izquierdo se encuentra la región productora del departamento de Córdoba perteneciente a la región del bajo Cauca. Sobre la derecha, se proyecta el área productora de los departamentos de Tolima y Huila, en la zona Centro.

## Características climáticas:

### Precipitación

El departamento del Tolima está comprendido en su totalidad sobre la cuenca alta del río Magdalena, entre las cordilleras Central y Oriental. Las mayores lluvias, entre 2,500 y 3,000 mm anuales, se observan al norte, en límites con Caldas, en los municipios de Fresno y Mariquita; y al extremo sur, en jurisdicción de los municipios de Chaparral y Río blanco, ambos sitios en estribaciones de la cordillera Central. Las menores cantidades anuales se localizan al oriente, en límites con Cundinamarca, al suroriente en Natagaima, y al occidente, en Cajamarca y Santa Isabel, con cantidades entre 1,000 y 1,500 mm al año.

El régimen de lluvias durante el año es de tipo bimodal. Presenta dos épocas secas en el año, la primera se registra durante el trimestre Junio – Julio - Agosto y la segunda en Diciembre – Enero - Febrero, siendo la principal la correspondiente a mitad de año. Esta última va siendo más marcada en dirección al centro y sur del departamento. Las temporadas de lluvia en la mayor parte del departamento, ocurren en los meses de Marzo – Abril - Mayo y Septiembre – Octubre - Noviembre.

El número de días con lluvia presenta un patrón marcado. La menor frecuencia se observa al oriente del departamento y oscila entre 50 y 150 días al año, y las mayores frecuencias ocurren al occidente, con valores alrededor de 200 días lluviosos al año, en promedio.

### **Temperatura**

Las mayores temperaturas medias se registran a lo largo de las riberas del río Magdalena, en municipios como Ambalema, sectores de Lérída, Espinal, Purificación, Coyaima y Natagaima, con valores medios superiores a los 28°C, constituyéndose en una de las regiones más calientes del país. Al sur del valle del río Magdalena, las temperaturas disminuyen ligeramente, y sobre las estribaciones, especialmente las correspondientes a la cordillera Central, en límites con Caldas, la temperatura disminuye rápidamente hasta alcanzar valores extremadamente fríos sobre los picos nevados.

### **Clasificación Climática**

El valle del Magdalena presenta clima cálido semiárido, y al extremo sur, cálido semi húmedo y húmedo. Los climas templados, aparecen en estrechas franjas a lo largo de las cordilleras. Los climas extremadamente fríos, se localizan al extremo occidental, en jurisdicción de Murillo y Santa Isabel.

La Figura. 12 muestra la variación espacial de las principales variables climatológicas del departamento del Tolima.

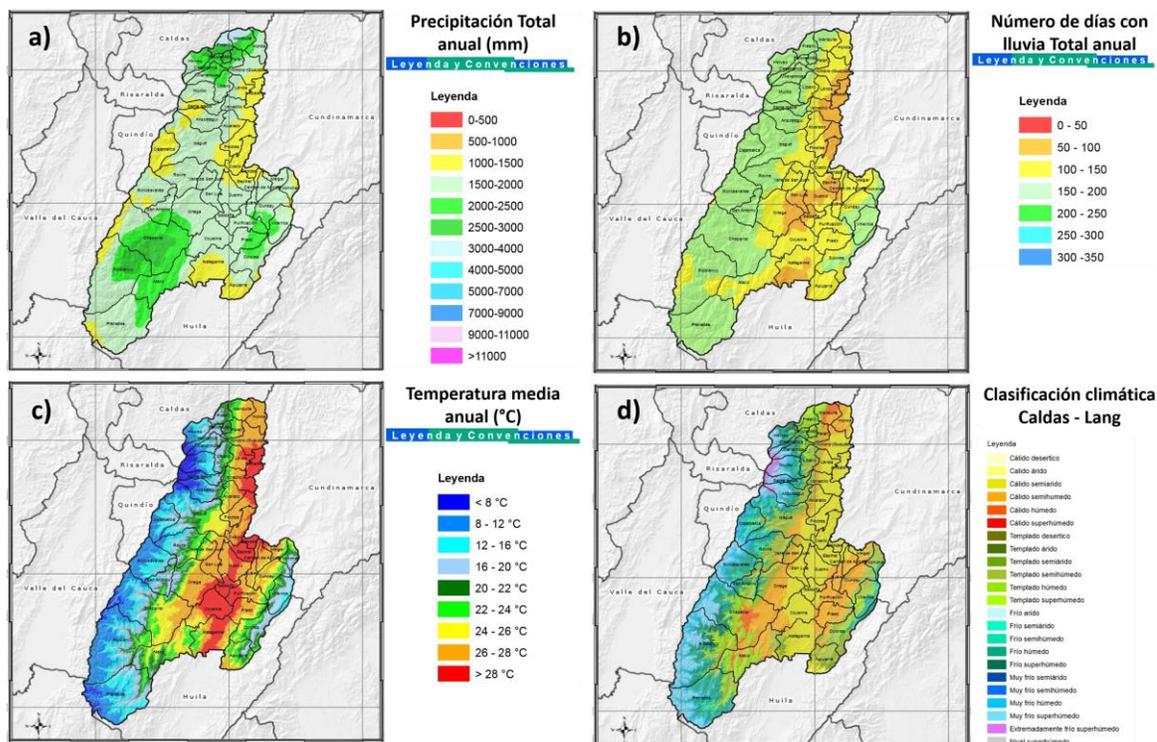


Figura. 12. Características climáticas del departamento del Tolima. a) Precipitación total anual; b) Número de días con lluvia total anual; c) Temperatura media anual; d) Clasificación climática según Caldas – Lang. Fuente: <http://atlas.ideam.gov.co/>

## 2.1.2 Departamento del Huila

Al igual que el departamento del Tolima, en el Huila las épocas de siembra del cultivo de arroz se centran durante los meses Abril – Julio y Octubre – Enero. Durante el primer semestre del año se tiene la participación de 1,431 UPA de arroz, que producen alrededor de 120,000 ton de arroz, y un rendimiento promedio de 7.0 ton/ha. Por su parte, en el segundo semestre las UPA aumentan a lo largo del departamento, pasando de 1,431 a 1,703 aproximadamente. No obstante, la producción y el rendimiento se mantienen igual a los valores reportados para el primer semestre del año (DANE, 2015).

## **Características climáticas:**

### **Precipitación**

La mayor parte del departamento comprende áreas del valle del alto Magdalena y las estribaciones de las cordilleras Central y Oriental. El valle del Magdalena registra las menores tasas de precipitación total anual, oscilando entre 1,000 y 1,500 mm. La vertiente occidental de la cordillera oriental presenta estos mismos volúmenes de precipitación. Por su parte, en la vertiente occidental de la cordillera Central las lluvias se incrementan hasta valores cercanos a los 2,000 mm. Al extremo sur, las lluvias se incrementan hasta los 2,500 mm en los municipios de Acevedo y Palestina.

El régimen de lluvias durante el año tiene una distribución bimodal, con una temporada seca principal muy marcada a mitad de año (Julio - Septiembre) y una secundaria, menos significativa entre Enero y Febrero. Las dos temporadas de lluvia se registran en el trimestre Marzo - Mayo y de Octubre - Diciembre. Al extremo sur, en el sector del municipio de San Agustín, se inicia la transición hacia el régimen de la Amazonia y las temporadas secas son poco marcadas.

El número de días con lluvia oscila entre 100 y 150 sobre el valle del Magdalena, no obstante, en áreas cercanas a los municipios de Aipe y Villa vieja pueden registrarse menos de 100. Sobre los piedemontes, los días lluviosos se incrementan ligeramente hasta alcanzar 200 días.

### **Temperatura**

Las mayores temperaturas se registran en áreas cercanas al municipio de Villa vieja, al extremo norte del valle del Magdalena, con valores mayores a los 28°C. En Neiva y Campoalegre, las temperaturas están entre 26 y 28°C, y hacia las estribaciones de las cordilleras los valores descienden en función de la elevación.

### **Clasificación Climática**

El clima cálido árido y semiárido caracteriza la mayor parte del valle del Magdalena. Hacia el piedemonte de la cordillera Oriental el clima es templado semiárido y semi húmedo. Sobre la cordillera Occidental, se localizan climas de tipo templado húmedo y frío húmedo. La Figura. 13 muestra la variación espacial de las principales variables climatológicas del departamento del Huila.

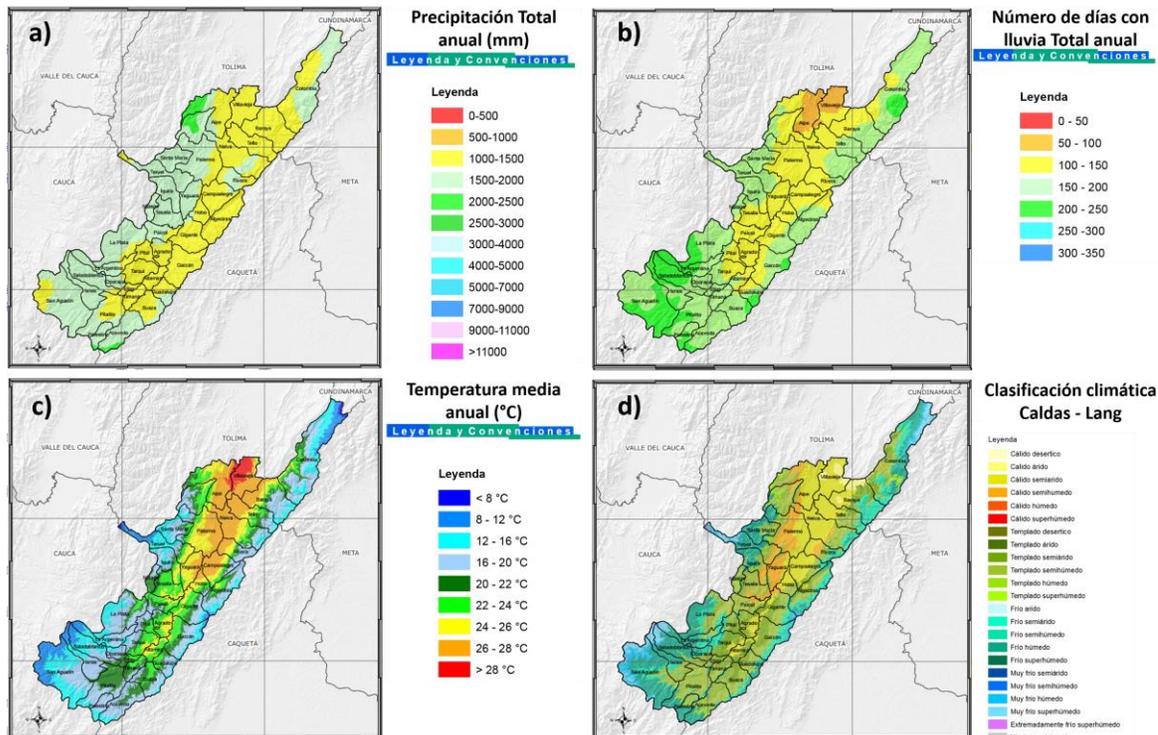


Figura. 13. Características climáticas del departamento del Huila. a) Precipitación total anual; b) Número de días con lluvia total anual; c) Temperatura media anual; d) Clasificación climática según Caldas – Lang. Fuente: <http://atlas.ideam.gov.co/>

### 2.1.3 Departamento de Córdoba

Córdoba se distingue por ser uno de los departamentos con mayor actividad agrícola de del Bajo Cauca, cerca del 90% del área sembrada está dedicada a los cultivos de Maíz y Arroz (Briceño, 2012). Durante el año se cultivan alrededor de 7,300 hectáreas de arroz mecanizado, en promedio se producen unas 34,000 toneladas, y se alcanza un rendimiento medio anual de 4.7 ton/ha, siendo superior al promedio general de la región. Respecto al número de productores y UPA de arroz en el departamento, en promedio hay una participación de 550 productores bajo sistema mecanizado y 600 UPA en el año (DANE, 2015).

## **Características climáticas:**

### **Precipitación**

El departamento de Córdoba registra un patrón muy variado en la distribución de las lluvias medias anuales. La franja costera presenta las menores lluvias con menos de 1,500 mm al año. El centro del departamento tiene entre 1,500 y 2,000 mm, y hacia el extremo sur, en cercanías a la cordillera, los volúmenes de lluvia pueden alcanzar los 4,000 mm. El régimen de lluvias durante el año, al centro y sur del departamento es mono modal con una única época seca de Diciembre a Marzo, y una temporada de lluvias que comienza a incrementarse a partir del mes de Abril. Durante el trimestre Julio – Agosto - Septiembre se alcanzan los máximos valores, y luego disminuyen progresivamente hasta el final de año. Este régimen es algo menos marcado en proximidades a la costa, donde se nota una leve disminución de lluvias a mitad de año. El número de días con lluvia está entre 50 y 100 al norte del departamento, y en dirección al sur se incrementa hasta alcanzar valores cercanos a los 200 días.

### **Temperatura**

La mayor parte del departamento registra temperaturas entre 26 y 28°C. En sectores aislados de Ayapel y Planeta Rica se pueden registrar valores levemente mayores a los 28°C. En dirección a la cordillera, en los municipios de Tierralta y Puerto Libertador, se localizan algunas franjas con temperaturas menores de 20°C.

### **Clasificación Climática**

Al norte de Montería, el clima predominante es de tipo cálido semiárido. La franja comprendida entre los municipios de Monte Líbano y Pueblo Nuevo posee un clima cálido semi húmedo, y más hacia el sur aparecen climas cálidos húmedos y algunas franjas de clima templado semi húmedo.

La Figura. 14 muestra la variación espacial de las principales variables climatológicas del departamento de Córdoba.

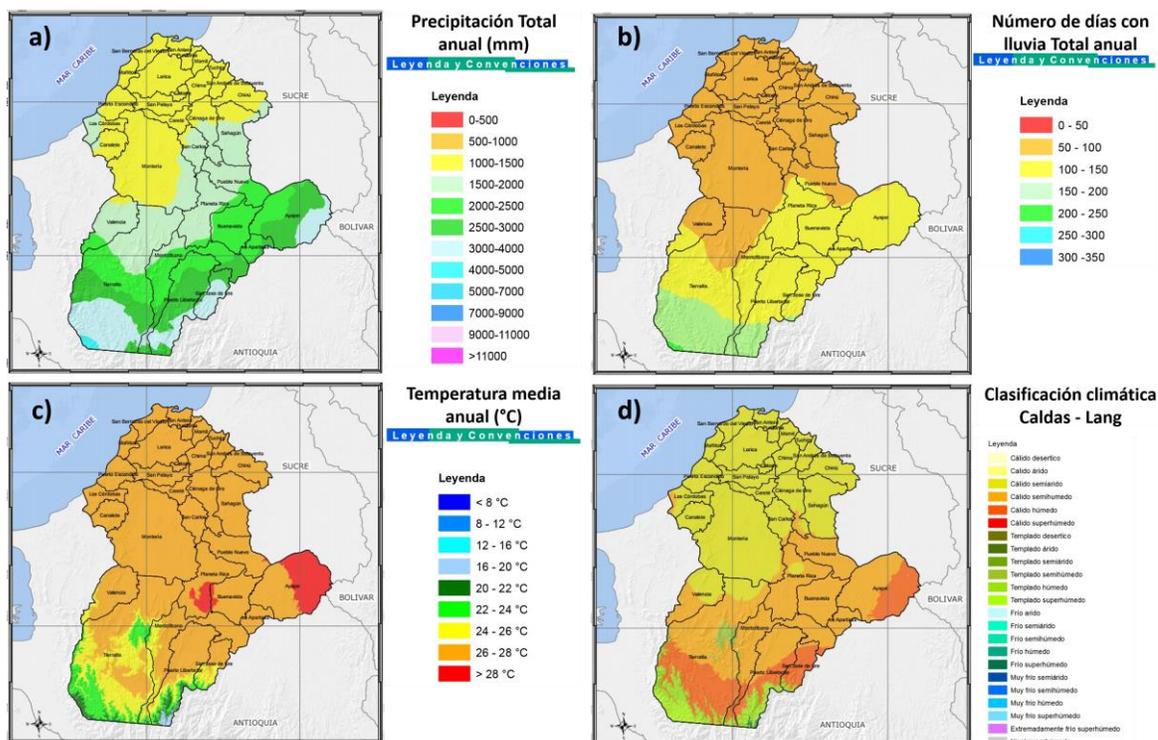


Figura. 14. Características climáticas del departamento Córdoba. a) Precipitación total anual; b) Número de días con lluvia total anual; c) Temperatura media anual; d) Clasificación climática según Caldas – Lang. Fuente: <http://atlas.ideam.gov.co/>

## 2.2 Datos observados

### 2.2.1 Datos de clima

Se seleccionaron un total de 38 estaciones meteorológicas distribuidas de manera estratégica en las áreas productoras de arroz de los departamentos de Tolima (17 estaciones), Huila (9 estaciones) y Córdoba (12 estaciones). Para cada una de las estaciones se utilizó información climática histórica diaria de 28 años de registro (1984 - 2012), de las variables precipitación, temperatura máxima y temperatura mínima. La información de las 38 estaciones fue suministrada por el Instituto de Hidrología, Meteorología y Estudios Ambientales de Colombia (IDEAM - <http://www.ideam.gov.co/>).

A cada una de las series climáticas observadas se le aplicó un análisis exploratorio y cuantitativo para verificar la calidad de los datos, siguiendo la metodología reportada por

Castro & Carvajal (2010). Con el fin de realizar un llenado de posibles datos faltantes dentro de las series (< 10.0 % del total de los datos registrados) se siguió la metodología aplicada por D'Afonseca et al (2013). Para cada una de las estaciones se estimaron los datos de radiación solar diaria a través del método de Richardson & Wright (1984). Con el objetivo de llevar a cabo escenarios de simulación espacial en de toda la región de estudio, se determinó el área de influencia para cada una de las estaciones meteorológicas a partir del trazado de polígonos de Thiessen (Figura. 15), tal como lo reporta Alexandre Bryan Heinemann et al. (2015).

Tabla 4. Identificación de la estación, latitud, longitud, departamento, municipio y tipo de suelo dentro de su área de influencia (ver Figura. 15)

<b>Código estación</b>	<b>Lat</b>	<b>Lon</b>	<b>Depto</b>	<b>Municipio</b>	<b>Texturas de suelos x área estación *</b>
1	2.70	-75.36	Huila	Campo Alegre	FAr; FArA
2	2.28	-75.71	Huila	Agrado	F; FAr; FArA
3	2.70	-75.30	Huila	Campo Alegre	F; FAr
4	3.25	-75.24	Huila	Villa Vieja	F; FAr; FArA
5	3.33	-75.18	Huila	Villa Vieja	F; FAr; FArA
6	2.83	-75.33	Huila	Palermo	F; FAr; FArA
7	2.71	-75.42	Huila	Campo Alegre	F; FAr; FArA; FA
8	3.13	-75.08	Huila	Baraya	F; FAr; FArA
9	2.95	-75.29	Huila	Neiva	F; FAr; FArA
10	3.22	-75.22	Huila	Villa Vieja	FAr
11	4.42	-75.09	Tolima	Ibagué	FAr; FArA
12	4.23	-75.00	Tolima	Coello	FArA
13	4.19	-74.96	Tolima	Espinal	FArA
14	4.82	-74.92	Tolima	Lérida	FAr; FArA
15	4.21	-74.61	Tolima	Melgar	F
16	3.86	-75.02	Tolima	Purificación	FAr; FArA
17	3.76	-74.93	Tolima	Prado	F; FAr
18	4.01	-74.98	Tolima	Guamo	F; FAr; FArA
19	4.76	-74.84	Tolima	Ambalema	FAr; FArA
20	3.77	-75.12	Tolima	Coyaima	F; FAr; FArA
21	4.19	-74.85	Tolima	Flandes	F; FAr; FArA
22	4.22	-75.10	Tolima	San Luis	FAr; FArA
23	4.48	-74.98	Tolima	Piedras	FAr; FArA
24	4.86	-74.89	Tolima	Lérida	FAr; FArA
25	4.96	-74.91	Tolima	Armero	FAr; FArA

---

26	5.18	-74.83	Tolima	Honda	FAR
27	8.30	-75.16	Córdoba	Ayapel	FAR; FArA; ArA
28	8.84	-75.80	Córdoba	Cerete	FAR; FArA
29	9.15	-75.62	Córdoba	Chima	FAR; FArA
30	8.91	-75.58	Córdoba	Ciénaga De Oro	FAR; FArA
31	8.00	-75.40	Córdoba	Monteriano	F; FAr; FArA
32	8.81	-75.85	Córdoba	Montería	FAR; FArA
33	8.66	-75.97	Córdoba	Montería	FAR; FArA
34	8.41	-75.89	Córdoba	Montería	FAR; FArA
35	8.79	-75.86	Córdoba	Montería	FAR; FArA
36	8.83	-75.83	Córdoba	Montería	FAR; FArA
37	8.18	-75.63	Córdoba	Planeta Rica	FAR; FArA; Ar; ArA
38	8.40	-75.58	Córdoba	Planeta Rica	FAR; FArA; Ar; ArA

---

\* F – Franco; FAR – Franco Arcilloso; FArA – Franco Arcillo Arenoso; FA – Franco Arenoso; Ar – Arcilloso; ArA – Arcillo Arenoso.

### 2.2.2 Datos de suelos

Uno de los obstáculos para llevar a cabo avanzados análisis que involucren el uso de modelos de cultivos a nivel espacial, es la disponibilidad de información observada de propiedades físicas, químicas e hidráulicas del suelo, a una detallada resolución espacial, debido al alto costo que esto genera y a la complejidad de las metodologías existentes para estimar dichas propiedades. No obstante, existen bases de datos secundarias, a escala global, que cuentan con información edáfica a una fina resolución, las cuales surgen como una alternativa ante la limitante mencionada anteriormente (Han, Ines, & Koo, 2015). SoilGrids1km (<http://soilgrids.org>) es una de estas importantes fuentes de datos, es un sistema de información de suelos a escala global en 3-D, con resolución espacial de 1 km, a diferentes niveles de profundidad (desde 0.0 hasta 200 cm), creada por el ISRIC y desarrollada a través de la plataforma GSIF (Global Soil Information Facilities) en colaboración abierta para el montaje, recopilación y producción de información de suelos (Hengl et al., 2014).

Para llevar a cabo de esta investigación, se utilizó la base de datos SoilGrids 1km con el fin de obtener información espacial de variables físicas y químicas del suelo, para las áreas productoras de arroz ubicadas en los departamentos de Tolima, Huila y Córdoba. En total se obtuvo información de 2,052 puntos

distribuidos de manera aleatoria en toda la región, para cada uno de ellos se extrajo información de variables como textura (% arenas, limo y arcillas), densidad aparente ( $\text{kg m}^{-3}$ ), carbono orgánico ( $\text{g kg}^{-1}$ ), capacidad de intercambio catiónico ( $\text{cmol kg}^{-1}$ ) y pH; a las siguientes profundidades: 20, 40 y 60 cm.

Con base en el sistema “nacional” de clasificación de suelos de los Estados Unidos, se determinó la clase textural de cada uno de los puntos muestreados, y se clasificaron los tipos de suelos más predominantes en toda la región de estudio (ver en la Tabla 4 y Figura. 15 la distribución espacial de los tipos de suelos).

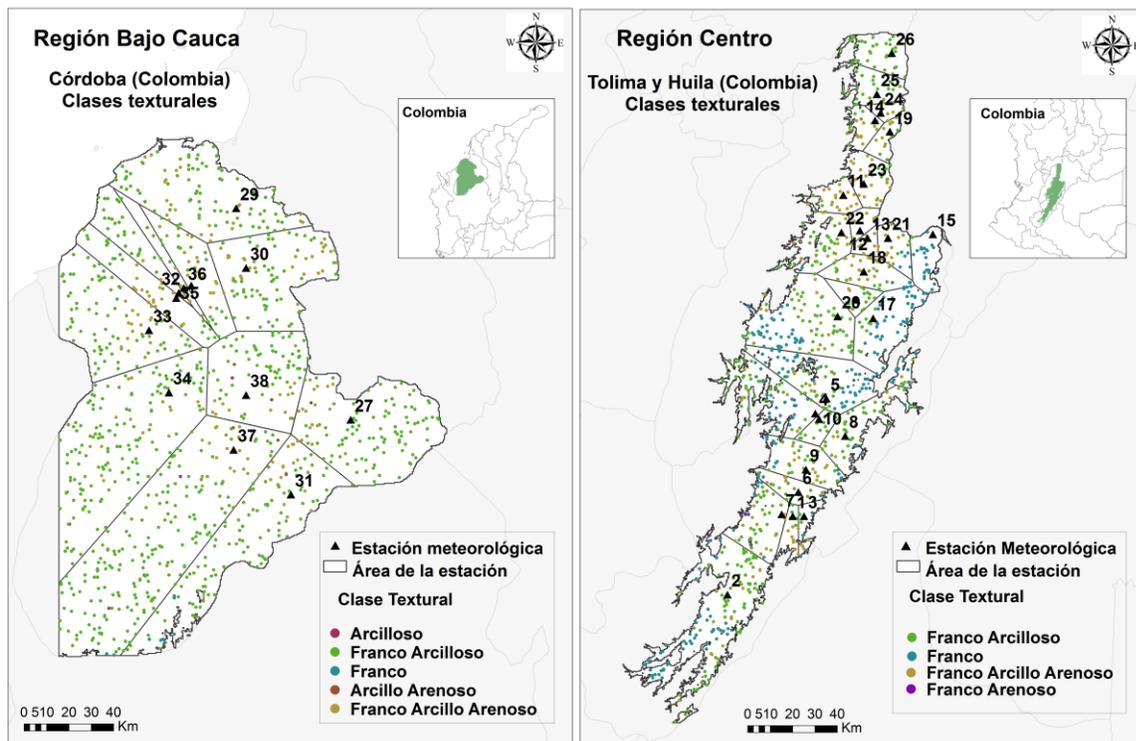


Figura. 15. Distribución espacial de las estaciones meteorológicas (triángulos) en las áreas productoras de Tolima, Huila y Córdoba; cada estación está delimitada por un área de influencia (polígonos), y los diferentes tipos de suelos (puntos). Los números representan el número de identificación de cada estación descrita en la Tabla 4.

Las propiedades hidráulicas de los suelos son otro de los factores importantes al momento de realizar estudios a través de modelos de cultivos. Sin embargo, comúnmente existe un vacío de información al respecto, ya que son propiedades que generalmente no están disponibles debido al grado de dificultad de su medición en campo y al alto costo económico asociado al analizarlas en el laboratorio. La estimación de propiedades hidráulicas del suelo a través del conocimiento de otras variables relevantes que han sido medidas, puede llevarse a cabo a través del uso de Funciones de Pedotransferencia (FPT) (Nebel et al., 2010).

Siguiendo los resultados obtenidos por Tomasella et al (2003, 2004) sobre la aplicación de FPT para suelos tropicales, en este estudio se determinaron propiedades hidráulicas, tales como el contenido de humedad volumétrico ( $m^3 m^{-3}$ ) a capacidad de campo ( $\theta_{UL}$ ), a punto de marchitez permanente ( $\theta_{LL}$ ) y el agua disponible en el suelo (ASW); en función de la textura (% arenas, limo y arcillas), la densidad aparente, el contenido de carbono orgánico y la capacidad de intercambio catiónico. En la Figura. 16 se muestran las ecuaciones utilizadas para determinar las propiedades hidráulicas de los suelos.

### **2.2.3 Datos de los ensayos de arroz de riego**

En el marco de dos proyectos de cooperación técnica y científica financiados por 1) El Ministerio de Agricultura y Desarrollo Rural (MADR), y 2) El Departamento Administrativo de Ciencia, Tecnología e Innovación (COLCIENCIAS); el CIAT y FEDEARROZ han estado diseñando, durante los últimos cuatro años (2012 – 2016), nuevas estrategias para la adaptación del sector arrocero de Colombia ante la variabilidad climática regional, a través de la evaluación de nuevos germoplasmas y líneas promisorias con capacidad de responder ante los nuevos desafíos climáticos, en las diferentes regiones productoras del país.

Dentro de dichos proyectos, entre los años 2013 y 2015 se llevaron a cabo ensayos experimentales en las estaciones de FEDEARROZ ubicadas en las localidades de Saldaña - Tolima (latitud: 3°54'49" N; longitud: 74°59'05" O; elevación: 415 m), Aipe - Huila (latitud: 3°15'11" N; longitud: 75°14'26" " O; elevación: 382 m) y Montería - Córdoba (latitud: 8°48'40" N; longitud: 75°51'14" O; elevación: 13 m) (Figura. 17). La Tabla 5 presenta algunas propiedades físico-químicas e hidráulicas de los suelos en los tres sitios experimentales.

$$x_{14} = -1.05501 + 0.0650857SI$$

$$x_{15} = -2.07588 + 0.0423954CL$$

$$x_{16} = -6.03402 + 4.80572Db$$

$$x_{17} = -2.18409 + 8.84963Me$$

$$z_9 = 0.175202 + 1.18513x_{17} - 0.0996042(x_{17})^2 + 0.327915x_{16} - 0.0758657(x_{16})^2$$

$$z_{10} = 0.929344z_9 + 0.132519x_{14}$$

$$\theta_{10kPa} = 0.339255 + 0.112526z_{10}$$

$$z_{11} = 0.191452 + 1.25652x_{17} - 0.079098(x_{17})^2 + 0.393814x_{16} + 0.152095x_{17}x_{16}$$

$$\theta_{33kPa} = 0.28951 + 0.103815z_{11}$$

$$z_{13} = 0.235084 + 0.33033x_{15} - 0.191838(x_{15})^2 + 0.0543679(x_{15})^3 + 0.977685x_{17} \\ + 0.304174x_{15}x_{17} - 0.218857(x_{17})^2 - 0.164373x_{15}(x_{17})^2 + 0.0415057(x_{17})^3 \\ + 0.373361x_{16} + 0.0811861x_{17}x_{16} - 0.0768087x_{15}x_{16}x_{17}$$

$$\theta_{1500kPa} = 0.214008 + 0.0862945z_{13}$$

Figura. 16. Ecuaciones aplicadas para determinar las propiedades hidráulicas del suelo. SI – Limo (%); CL – Arcilla (%); Db – Densidad aparente ( $Mg\ m^{-3}$ ); Me – Humedad equivalente ( $g\ g^{-1}$ );  $\theta_{10\ kPa}$ ,  $\theta_{33\ kPa}$ ,  $\theta_{1500\ kPa}$  – Contenido de humedad volumétrico a 10, 33 y 1500 kPa ( $m^3\ m^{-3}$ ).

Un total de 47 ensayos experimentales fueron considerados en este estudio, para llevar a cabo el proceso de calibración y validación de cuatro variedades (Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21374) dentro del modelo de cultivo Oryza2000. Los experimentos fueron clasificados en dos grupos de la siguiente manera:

1) **Datos para calibración del modelo:** 21 ensayos experimentales distribuidos entre las localidades de Saldaña (11 ensayos) y Montería (10 ensayos); de los cuales, 6 fueron para calibrar Fedearroz 2000, 5 para calibrar Fedearroz 733, 4 para calibrar Fedearroz 60 y 6 para calibrar CT21375.

2) **Datos para validación del modelo:** 26 ensayos distribuidos entre las localidades de Saldaña (8 ensayos), Aipe (16 ensayos) y Montería (2 ensayos); 8 ensayos fueron

seleccionados para validar Fedearroz 2000, 5 para validar Fedearroz 733, 5 para validar Fedearroz 60 y 8 para validar CT21375.

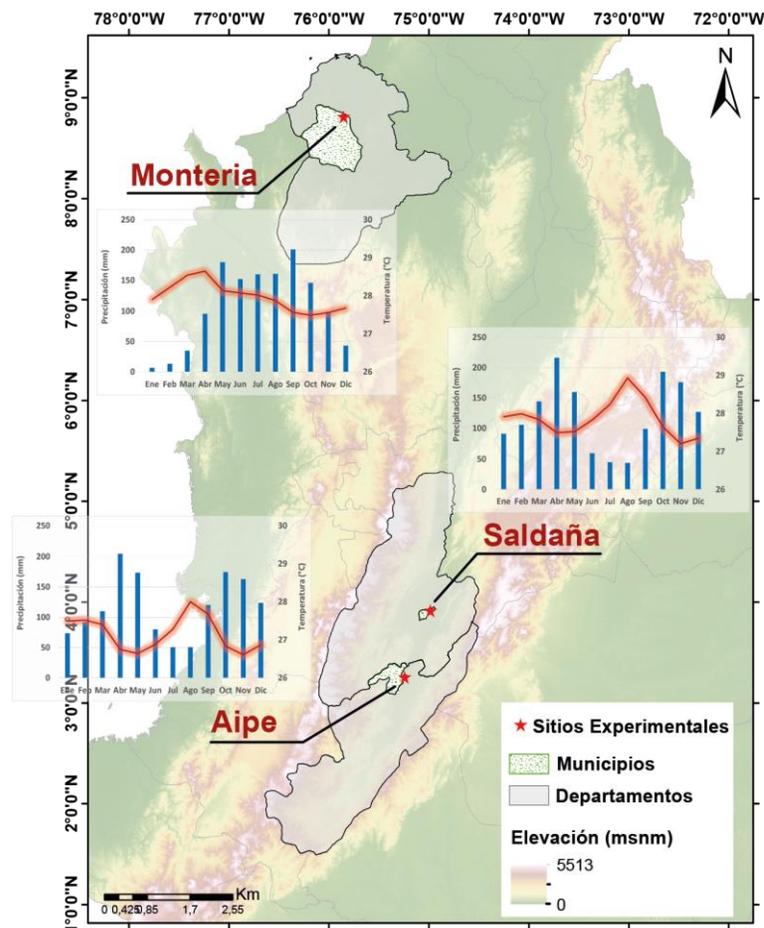


Figura. 17. Ubicación geográfica de los sitios experimentales y distribución promedio mensual de la precipitación (barras azules) y temperatura media (línea roja).

Respecto a las épocas seleccionadas para la ejecución de los ensayos de calibración, en Saldaña se realizaron entre las temporadas Diciembre 2013 – Marzo 2014 (fecha de siembra 5 de Diciembre), Febrero 2014 – Mayo 2014 (fecha de siembra 5 de Febrero) y Octubre 2014 – Enero 2015 (fecha de siembra 7 de Octubre). Por su parte, en Montería se los ensayos se llevaron a cabo en las épocas Mayo 2014 – Octubre 2014 (fecha de siembra 30 de Mayo), Julio 2014 – Noviembre 2014 (fecha de siembra 17 de Julio) y Noviembre 2014 – Marzo 2015 (fecha de siembra 7 de Noviembre).

En cuanto a los ensayos de validación, en la localidad de Saldaña se realizaron durante las temporadas Abril 2013 – Septiembre 2013 (fecha de siembra 29 de Abril), Julio 2014 – Noviembre 2014 (fecha de siembra 15 de Julio) y Marzo 2015 – Junio 2015 (fecha de siembra 3 de Marzo). En Aipe los ensayos se ejecutaron en las épocas Mayo 2013 – Octubre 2013 (fecha de siembra 28 de Mayo), Enero 2014 – Mayo 2014 (fecha de siembra 4 de Enero), Febrero 2014 – Junio 2014 (fecha de siembra 24 de Febrero) y Agosto 2014 – Diciembre 2014 (fecha de siembra 6 de Agosto). Por último, los dos únicos experimentos considerados para la validación del modelo en Montería, se llevaron a cabo entre Diciembre 2014 – Abril 2015 (fecha de siembra 23 de Diciembre) (Tabla 6).

Tabla 5. Propiedades físicas, químicas e hidráulicas (promedio) de los suelos en las estaciones experimentales de Saldaña, Aipe y Montería; para los ensayos realizados durante el periodo 2013 – 2015.

<b>Saldaña</b>										
Profundidad del suelo (cm)	DA (g cm <sup>3</sup> )	Textura (%)			CHS (cm h <sup>-1</sup> )	Humedad volumétrica (%)			pH	CIC (Cmol kg <sup>-1</sup> )
		Arena	Limo	Arcilla		Sat	CC	PMP		
0 - 20	1.59	40	47	13	16.92	51	31	19	5.6	7.9
20 - 40	1.70	58	27	15	23.16	44	24	15	6.4	4.7
40 - 60	1.56	81	5	14	36.71	49	20	10	6.5	3.1
<b>Aipe</b>										
0 - 20	1.36	12	27	61	1.22	54	44	35	6.7	27.1
20 - 40	1.47	16	17	66	1.46	58	43	32	7.0	28.8
40 - 60	1.44	14	21	70	1.43	55	42	34	7.2	27.7
<b>Montería</b>										
0 - 20	1.37	14	36	50	0.15	53	45	35	7.1	23.5
20 - 40	1.49	19	43	38	0.74	45	41	33	7.7	19.1
40 - 60	1.47	20	43	38	0.25	46	40	31	7.7	18.5

DA – Densidad aparente; CHS – Conductividad hidráulica saturada; pH – Potencial de hidrógeno; CIC – Capacidad de intercambio catiónico; Sat – Saturación; CC – Capacidad de campo; PMP – Punto de marchitez permanente.

Los ensayos mencionados anteriormente fueron seleccionados debido a que no tuvieron una afectación considerable, a causa de algún tipo de estrés biótico o abiótico; ya que se ejecutaron bajo condiciones cercanas a producción potencial, excluyendo cualquier tipo de limitación por agua, nutrientes, o alguna infestación por plaga, enfermedades o malezas.

Los diferentes experimentos se establecieron bajo un Diseño de Bloques Completos al Azar (DBCA) con tres repeticiones y cuatro tratamientos (correspondientes a las cuatro variedades consideradas en el estudio). Cada unidad experimental tuvo un área de 30 m<sup>2</sup> (5 m x 6 m) y estuvo compuesta de 24 surcos, con una distancia de 0.25 m entre surcos (Figura. 18).

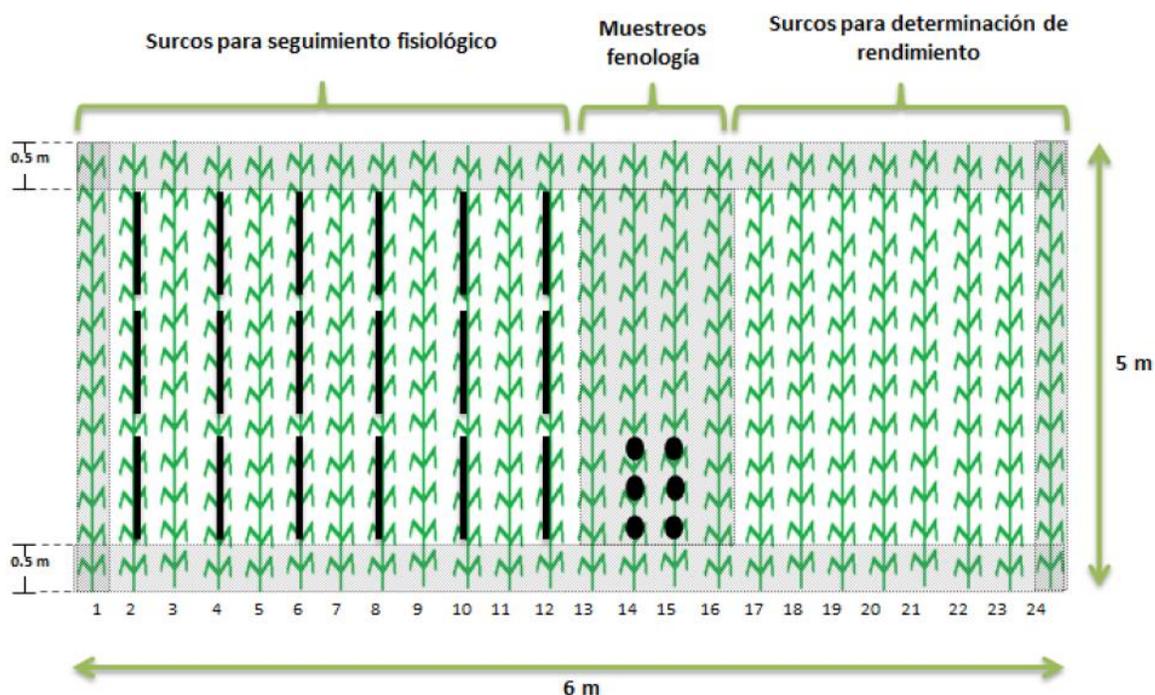


Figura. 18. Diseño experimental de las parcelas establecidas para el monitoreo de etapas fenológicas, mediciones secuenciales de biomasa y determinación del rendimiento final del cultivo. Cada ensayo estuvo conformado por 12 parcelas experimentales (4 tratamientos x 3 repeticiones).

En cada parcela experimental, los surcos 1 y 24, al igual que los bordes superior e inferior (cada uno de 0.5 m de ancho), no se consideraron en los muestreos secuenciales, ni en la cosecha, con el propósito de eliminar el efecto de borde.

En cada surco de 5 m se distribuyeron 8.25 gramos de semilla (7.5 gr + 10.0% por algún tipo de error en germinación), la densidad de siembra para el establecimiento de los ensayos fue de 60 kg ha<sup>-1</sup>.

En los surcos 14 y 15 de la parcela experimental, se seleccionó un metro lineal como unidad de muestreo para identificar las fechas de iniciación de panícula y la floración del cultivo al 50% y 100%. Por su parte, las fechas de siembra, emergencia y madurez fisiológica fueron las mismas que las identificadas entre los surcos 2 y 12; en los cuales se seleccionaron, cada dos semanas, 6 metros lineales de manera aleatoria, y se tomaron mediciones secuenciales de Índice de Área Foliar (LAI), pesos secos de la biomasa producida en hojas verdes, hojas muertas, tallos y panículas.

Los pesos secos se obtuvieron después de secar las muestras cosechadas en un horno, a una temperatura constante de 65°C hasta alcanzar un peso seco constante (aproximadamente 3 días). Al momento en que el cultivo se encontraba a madurez fisiológica y los granos alcanzaron una humedad entre el 22% y 24%, el área comprendida entre los surcos 17 y 24 fue cosechada con el fin de determinar el rendimiento potencial de cada variedad. Durante toda la etapa del cultivo, en cada uno de los experimentos se monitoreo las condiciones climáticas de precipitación, temperatura mínima, temperatura máxima y radiación solar; a través de estaciones meteorológicas Davis ubicadas en los diferentes centros experimentales de FEDEARROZ.

Respecto al manejo agronómico, en todos los experimentos, tanto de calibración como de validación, se llevó a cabo un manejo agronómico 100% preventivo, con el fin de evitar cualquier afectación por pestes o enfermedades. De igual forma, se evitó tener deficiencias nutricionales (fertilización adicional en caso de ser necesaria) y algún nivel de sequía (suministrando el riego necesario) que pudiera afectar al cultivo.

El manejo agronómico se estableció con base al programa de Adopción Masiva de Tecnología – AMTEC (FEDEARROZ, 2011), el cual permite la expresión del potencial de rendimiento de cada variedad, teniendo en cuenta los siguientes aspectos:

- 1) Buena adecuación y preparación del terreno, utilizando los implementos adecuados de acuerdo a las condiciones físicas del suelo.
- 2) Optima nutrición del cultivo, teniendo en cuenta las condiciones físicas, químicas y biológicas del suelo; así como los requerimientos nutricionales de la planta, las condiciones

---

ambientales, enmiendas orgánicas e inorgánicas, dosis y época de aplicación de los fertilizantes.

3) Manejo integrado de malezas, identificando las malezas predominantes y realizando un excelente control químico en las etapas de pre siembra, pre emergencia y pos emergencia, para garantizar que el cultivo quede libre de malezas durante los 30 primeros días.

4) Manejo integrado de insectos fitófagos, tratando las semillas con insecticidas y fungicidas con el fin de garantizar una buena protección durante la germinación y emergencia de la plántula, además de monitorear la presencia de insectos en el suelo, sogata, comedores de follaje y barrenadores.

5) Manejo integrado de enfermedades (hongos, bacterias y virus) durante las diferentes etapas del cultivo, implementando medidas fitosanitarias como la aplicación de fungicidas, preparación del suelo, manejo del riego, control de insectos, nutrición mineral etc.

6) Adecuado manejo del riego para garantizar una buena disponibilidad de nutrientes a la planta, un control de malezas, plagas y enfermedades. Durante las primeras semanas del cultivo, se mantuvo un riego a saturación del suelo y se aplicaron mojes periódicos hasta que la planta inició la etapa de macollamiento, posteriormente se estableció una lámina de riego permanente que no excedió los 5 cm, con el fin de evitar algún tipo de estrés por sequía.

La Tabla 6 presenta un resumen de los experimentos utilizados, las variables medidas, y su uso dentro de este estudio.

## **2.3 Parametrización y evaluación del modelo de cultivo**

### **2.3.1 Modelo de cultivo Oryza2000**

Oryza2000 es un modelo mecanístico y eco-fisiológico para el cultivo del arroz, desarrollado entre el IRRI y la Universidad de Wageningen. Este modelo simula, a escala diaria, el crecimiento, desarrollo, y balance de agua del cultivo de arroz en situación de producción potencial, limitación de agua y limitación por Nitrógeno (B. a M. Bouman et al., 2001; Feng et al., 2007). En todas las situaciones de producción, se asume que el cultivo está bien protegido contra plagas, pestes y enfermedades que puedan afectar el rendimiento del cultivo.

Tabla 6. Datos experimentales obtenidos en las localidades de Saldaña (SAL), Montería (MON) y Aipe (AIP); para las variedades Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21375.

Experimentos	Sitio	Cultivar	Descripción <sup>1</sup>	Siembra	VARIABLES MEDIDAS		
Calibración (21 Ensayos)	SAL	Fed2000	S2MADR	5-dic-13	Fenología: Emergencia, iniciación de panícula, floración, madurez fisiológica		
		Fed733	S3MADR	5-feb-14			
		CT21375	S3COL	7-oct-14			
	MON	Fed60	S2MADR S3MADR	5-dic-13 5-feb-14	Dinámica de biomasa: Hojas verdes, hojas muertas, tallos y panículas		
		Fed2000 CT21375	S3MADR S4MADR S1COL	30-may-14 7-nov-14 17-jul-14		Índice de área foliar (LAI) y rendimiento	
		Fed733 Fed60	S3MADR S4MADR	30-may-14 7-nov-14			
Validación (26 Ensayos)	SAL	Fed2000	S1MADR	29-abr-13	Fenología: Emergencia, iniciación de panícula, floración, madurez fisiológica		
		CT21375	S2COL S4COL	15-jul-14 20-abr-15			
	MON	Fed733 Fed60	S1MADR	29-abr-13	Dinámica de biomasa: Hojas verdes, hojas muertas, tallos y panículas		
		Fed2000 CT21375	S2COL	23-dic-14		Índice de área foliar (LAI) y rendimiento	
		AIP	Fed2000	S1MADR			28-may-13
			CT21375 Fed733 Fed60	S2MADR S3MADR S4MADR		4-ene-14 24-feb-14 6-ago-14	Dinámica de biomasa: Hojas verdes, hojas muertas, tallos y panículas
					Rendimiento		

<sup>1</sup> S1, S2, S3, S4 – Número de siembra realizada en cada localidad; MADR – Ensayos realizados en el marco del proyecto con el ministerio de agricultura de Colombia; COL – Ensayos realizados en el marco del proyecto financiado por COLCIENCIAS.

El modelo *Oryza2000* ha sido usado ampliamente para diversas aplicaciones en el campo de la investigación agrícola (Larijani, Sarvestani, Nematzadeh, Manschadi, & Amiri, 2011; Li et al., 2013; Sudhir-Yadav, Li, Humphreys, Gill, & Kukal, 2011; Wang et al., 2014; Xue et al., 2008; J. Zhang, Feng, Zou, & Liu, 2015). Este modelo tiene una gran habilidad para

predecir el crecimiento, desarrollo y rendimiento del cultivo de arroz, en función de las condiciones ambientales, prácticas agronómicas y rasgos propios del cultivo. Estas características ofrecen una gran oportunidad para caracterizar el desempeño del arroz en múltiples sitios bajo diferentes condiciones ambientales, además de poder cuantificar la incidencia de un determinado estrés abiótico sobre el cultivo, ya que tiene en cuenta las interacciones entre la planta y el ambiente en el cual se desarrolla (Karine Chenu et al., 2013).

En este estudio se establecieron condiciones potenciales de Nitrógeno en el cultivo. La dinámica de agua se simuló a través del módulo de balance de agua en el suelo "PADDY". Este es un modelo unidimensional multicapa, que simula el balance de agua en el suelo para una variedad específica en condiciones inundadas o no inundadas. El modelo "PADDY" se encuentra descrito en mayor detalle por Boling, Bouman, Tuong, Murty, & Jatmiko (2006).

### 2.3.2 Parametrización del modelo de cultivo

Oryza2000 fue parametrizado para tres variedades comerciales de FEDEARROZ (Fedearroz 2000, Fedearroz 733 y Fedearroz 60) y una línea avanzada de mejoramiento del CIAT (CT21375).

**Fedearroz 2000** es una variedad de crecimiento inicial rápido, su vigor inicial mejora cuando se pre abona; presenta alto macollamiento, es tolerante al volcamiento, resistente al virus de hoja blanca y susceptible a *Piricularia* en hoja y cuello, *Burkholderia glumae* y *Rhizoctonia*; es de amplia adaptabilidad y estabilidad en ambos semestre del año.

**Fedearroz 733** se caracteriza por ser una variedad de alto vigor, presenta alto macollamiento, tiene una ligera tendencia al volcamiento, baja susceptibilidad a *Piricularia grisea*, *Rhizoctonia solani* y *Sarocladium*; y es tolerante al virus de Hoja Blanca.

**Fedearroz 60** tiene alto vigor y alto macollamiento, es resistente al volcamiento, aunque puede ser afectada por excesos de agua y de fertilización nitrogenada; es sensible a altas dosis de herbicidas hormonales, es tolerante a la *Piricularia* y susceptible a *Rhizoctonia* y Hoja Blanca en campo; tiene un alto potencial productivo en épocas de mayor oferta lumínica en el año (para mayor información consultar la página <http://www.fedearroz.com.co/>).

**CT21375** se distingue por ser una línea de alto rendimiento, de bajo macollamiento, panículas largas, tallos fuertes, altura intermedia, follaje verde oscuro con senescencia retrasada y alta fertilidad (González Calderón, 2015).

Los parámetros que fueron objeto de calibración dentro del modelo de cultivo fueron: 1) Tasas de desarrollo fenológico (DVR); 2) Factores de distribución de asimilados en hojas verdes, hojas muertas, tallos y panículas (FLV, DRL, FST y FSO, respectivamente); 3) Área foliar específica (SLA); 4) Tasa de crecimiento relativo del área de las hojas – Mínimo y Máximo (RGRLMN y RGRLMX, respectivamente). Todos los demás parámetros permanecieron iguales a los registrados en el archivo de cultivo estándar de Oryza2000, los cuales corresponden a la variedad tropical IR72.

### **Tasas de desarrollo fenológico (DVR)**

Los parámetros de desarrollo fenológico calibrados fueron: 1) DVRJ ( $^{\circ}\text{C dia}^{-1}$ ) – Fase vegetativa base, desde la emergencia hasta el inicio de la fase sensible al fotoperiodo; 2) DVRI ( $^{\circ}\text{C dia}^{-1}$ ) – Fase sensible al fotoperiodo, desde el final de la etapa vegetativa base hasta la iniciación de la panícula; 3) DVRP ( $^{\circ}\text{C dia}^{-1}$ ) – Fase de formación de la panícula, desde la iniciación de la panícula hasta la floración (50%) del cultivo; 4) DVRR ( $^{\circ}\text{C dia}^{-1}$ ) – Fase de llenado de grano, desde la floración (50%) hasta la madurez fisiológica. Para realizar la calibración de estos parámetros, se consideraron las fechas de emergencia, iniciación de la panícula, floración (50%), y madurez fisiológica reportada en los ensayos de calibración.

Para la determinación de la fecha de iniciación de la panícula, se consideró la etapa entre el inicio de la elongación de los entrenudos y 30 días antes de la fecha estimada de floración, se seleccionaron los cinco tallos de mayor longitud en ese momento, si se podía observar la presencia del primordio en cuatro de los cinco tallos, se consideraba esa fecha como el momento de iniciación de la panícula.

Para cada una de las variedades se obtuvo las DVR promedio, producto de un proceso iterativo que modificaba cada una de las tasas de desarrollo, hasta que los datos simulados de las etapas fenológicas se aproximaron con el menor error a las fechas fenológicas observadas en los experimentos en campo.

---

Para este proceso se consideraron las temperaturas cardinales reportadas por Bouman et al (2001) para la variedad IR72: Temperatura base ( $T_b$ ) de 8°C; Temperatura optima ( $T_o$ ) de 30°C; y Temperatura máxima ( $T_m$ ) de 42°C.

A pesar de que el modelo considera la etapa sensible al fotoperiodo dentro de las tasa de desarrollo fenológico, este parámetro no fue modificado ya que en Colombia las variedades en estudio no son afectadas por el fotoperiodo.

### **Factores de distribución de asimilados (FLV, DRL, FST y FSO) y desarrollo foliar (SLA, RGRLMN y RGRLMX)**

Para la obtención de los parámetros genéticos que modulan el crecimiento y desarrollo de cada una de las variedades consideradas, se aplicó el mismo método descrito por Li et al (2013). Después de calcular las tasas de desarrollo para los cuatro cultivares, se utilizó las mediciones de rendimiento, área foliar y producción de biomasa en los diferentes órganos, obtenidos en los ensayos de calibración (Tabla 6), como criterio para derivar los diferentes parámetros genéticos. La información registrada sobre el manejo agronómico establecido en campo, condiciones climáticas diarias y características específicas de los suelos, fue usada durante este proceso.

Para cada una de las variedades, la calibración se basó en un proceso iterativo, en el cual se modificaba los valores iniciales de los parámetros registrados para la variedad estándar IR72, hasta que las diferencias entre las mediciones obtenidas en campo y los resultados de la simulación fueron mínimas en un determinado número de iteraciones (para este caso fueron 10,000). El proceso finalizó cuando las diferencias entre los datos medidos (LAI, biomasa en hojas, tallos, panículas y biomasa total) y simulados estaban dentro del rango de desviaciones consideradas ( $\pm 10\%$  de los datos medidos). El proceso iterativo se basó en la minimización del error cuadrático medio y absoluto de las diferencias entre los datos observados y simulados.

Una vez se obtuvieron los parámetros genéticos para cada genotipo, el modelo fue validado con el fin de estimar el crecimiento y desarrollo del cultivo en una amplia gama de condiciones ambientales.

### 2.3.3 Evaluación del modelo de cultivo

El desempeño del modelo ORYZA2000 fue evaluado usando 26 ensayos experimentales, establecidos durante los años 2013 – 2015, en las localidades de Saldaña – Tolima, Aipe – Huila y Montería - Córdoba; para las variedades Fedearroz 2000 (8 experimentos), CT21375 (8 experimentos), Fedearroz 733 (5 experimentos) y Fedearroz 60 (5 experimentos).

Siguiendo la metodología de B. A. M. Bouman & Van Laar (2006), se utilizó una serie de gráficas y medidas estadísticas generalmente empleadas en la evaluación de modelos. Gráficamente se compararon los valores observados y simulados de la fenología, la producción de biomasa en los diferentes órganos, el LAI y el rendimiento del cultivo. Para las mismas variables, se calculó la pendiente ( $\alpha$ ), el intercepto ( $\beta$ ) y el coeficiente de determinación ( $R^2$ ) de la regresión lineal entre los valores simulados (Y) y observados (X). Además, se calculó la prueba de medias t-Student asumiendo varianza desigual ( $P(t^*)$ ), y el error medio cuadrado absoluto ( $RMSE_a$ ) y normalizado ( $RMSE_n$ ) entre los valores simulados y observados:

$$RMSE_a = \sqrt{\frac{\sum(Y_i - X_i)^2}{n}} \text{Ecu. 9}$$

$$RMSE_n = \frac{100 * RMSE_a}{\frac{\sum X_i}{n}} \text{Ecu. 10}$$

Donde  $n$  es el número de observaciones. Se consideró que el modelo tuvo una buena representación de los datos experimentales cuando  $\alpha$  fue cercano a 1,  $\beta$  cercano a 0,  $R^2$  cercano a 1,  $p(t)$  mayor a 0.05 y  $RMSE_a$  menor a la desviación estándar de los datos observados.

---

## 2.4 Identificación y caracterización ambiental

### Escenarios de simulación

Para llevar a cabo la simulación espacial del modelo de cultivo se consideró la información climática histórica (1984 - 2012) monitoreada por 38 estaciones meteorológicas, ubicadas de manera estratégica en toda la región de estudio (Figura. 15). Teniendo en cuenta los periodos tradicionales de siembra y cosecha del cultivo de arroz en la región, para cada año se consideraron 18 fechas de siembra, distribuidas entre los periodos 15 de Marzo – 3 Junio y 15 de Septiembre - 4 de Diciembre, para cada periodo se tuvo 9 fechas de siembra, separadas en intervalos de 10 días cada una. De igual forma, se consideraron 6 diferentes tipos de suelos (Tabla 4), y un manejo agronómico de acuerdo a lo recomendado por FEDEARROZ en cada una de los sitios simulados. Todas las simulaciones se configuraron para condiciones de arroz inundado, sin limitaciones por Nitrógeno, ni por factores bióticos. La tasa de evapotranspiración fue calculada con base en el método de Priestley-Taylor.

### Identificación de grupos ambientales

La identificación de los grupos ambientales se hizo con base en el rendimiento alcanzable obtenido para cada uno de los genotipos, a partir de los múltiples escenarios de simulación que combinaron diversas condiciones de clima\*suelo\*manejo agronómico. Todos los rendimientos simulados ( $\approx 184,000$  escenarios de simulación) desde 1984 hasta 2011, fueron organizados en una matriz de datos que combinó el año de simulación, fecha de siembra, sitio y tipo de suelo (Figura. 19). Posteriormente se aplicó un algoritmo de agrupamiento jerárquico aglomerativo (Luis & Villardón, 2007) sobre los rendimientos simulados, utilizando como medida de disimilaridad la distancia Euclídea al cuadrado, y el incremento de la suma de cuadrados como criterio de fusión (Ward, 1963).

Para identificar el número óptimo de grupos, se utilizó el paquete NbClust (Charrad, Ghazzali, Boiteau, & Niknafs, 2014) del programa estadístico R (R Development Core Team, 2008), el cual proporciona 30 índices para determinar el número óptimo de grupos y propone el mejor esquema de agrupamiento. Los análisis realizados se llevaron a cabo usando la librería FactoMineR (Husson, Josse, Le, & Maintainer, 2016) del programa estadístico R.

Soil	Dep	Cultivar	EMD	Lat	Lon	Yield.1984	Yield.1985	Yield.1986	Yield.1987	Yield.1988
Fra_Arc	HUI	FED2000	65	2.7	-75.36	7082.8	6920.5	6792	6831.9	6695.5
Fra_Arc	HUI	FED2000	75	2.7	-75.36	7093.5	7265.2	6926.8	6805.8	6947.5
Fra_Arc	HUI	FED2000	85	2.7	-75.36	7234	7458.2	7036.3	6991.1	7037.7
Fra_Arc	HUI	FED2000	95	2.7	-75.36	7227.6	7543.5	7396.3	7109.7	7236.8
Fra_Arc	HUI	FED2000	105	2.7	-75.36	7577.5	7516.8	7587.2	7100.3	7376.9
Fra_Arc	HUI	FED2000	115	2.7	-75.36	7934.8	7351.5	7768.4	6906.8	7447.5
Fra_Arc	HUI	FED2000	125	2.7	-75.36	7579.7	7269.8	7612.4	7027.4	7296.2
Fra_Arc	HUI	FED2000	135	2.7	-75.36	7469.1	7376.2	7446	7182.5	7123.2
Fra_Arc	HUI	FED2000	145	2.7	-75.36	7331.8	7474.2	7454.9	7187.8	7150.9
Fra_Arc_Are	HUI	FED2000	65	2.7	-75.36	7130.6	6955.9	6826.3	6867.2	6727.8
Fra_Arc_Are	HUI	FED2000	75	2.7	-75.36	7141.1	7304.7	6961.1	6840.4	6981
Fra_Arc_Are	HUI	FED2000	85	2.7	-75.36	7276.5	7502.6	7073.8	7028.6	7074.4

Figura. 19. Ejemplo de la Matriz de datos que agrupa los rendimientos obtenidos para cada escenario de simulación.

## Determinación del mayor estrés abiótico y su influencia sobre el rendimiento del cultivo

Para cada grupo ambiental identificado previamente, se determinó el principal factor de estrés abiótico a partir de un modelo de red neuronal artificial supervisado, el cual permitió identificar la importancia relativa de un grupo de variables explicativas (variables ambientales) sobre una variable de respuesta (rendimiento del cultivo).

## Preparación de los datos

En la Tabla 7 se muestran las variables explicativas y la variable de respuesta, que fueron consideradas para la construcción del modelo de red neuronal en cada uno de los grupos ambientales. Inicialmente, los datos obtenidos de los múltiples escenarios de simulación, se clasificaron de acuerdo al número de ambientes identificados. Luego, para cada una de las simulaciones se extrajo la información promedio de temperatura mínima, temperatura máxima y radiación solar, divididas en las etapas fenológicas mencionadas en la Tabla 7. En total, para cada una de las simulaciones realizadas ( $\approx 184,000$ ) se obtuvieron 16 variables explicativas vs el respectivo rendimiento como variable de respuesta.

Tabla 7. Variables explicativas y variable de respuesta para la construcción del modelo de red neuronal y análisis de relevancia.

Variables explicativas	Etapas fenológicas	Variable respuesta
Temperatura mínima media (°C)		
Temperatura máxima media (°C)	Eme – I. Pa	Rendimiento del cultivo
Radiación solar promedio (MJ m <sup>2</sup> día <sup>-1</sup> )	I. Pa - Ini. Flo	
	Ini. Flo - Fin. Flo	
	Fin. Flo - Mad. Fis	
Precipitación acumulada (mm)		

Etapas fenológicas: Eme – Ipa: Emergencia - Iniciación de panícula; Ipa - Ini.Flo: Iniciación de panícula - Inicio de la floración; Ini.Flo - Fin.Flo: Inicio de la floración - Fin de la floración; Fin.Flo - Mad. Fis: Fin de la floración – Madurez fisiológica.

Teniendo en cuenta que las variables seleccionadas tenían diferentes escalas debido a sus unidades de medida, se eliminó esta disparidad llevando a cabo una normalización o estandarización con la siguiente ecuación:

$$y = \frac{(y_{max} - y_{min}) * (x - x_{min})}{(x_{max} - x_{min})} + y_{min} \quad \text{Ecu. 11}$$

Donde  $y_{\min}$  y  $y_{\max}$  son el valor mínimo y máximo deseado (-1 y 1 en este caso) respectivamente,  $x_{\max}$  y  $x_{\min}$  son el valor mínimo y máximo del conjunto de datos,  $x$  es un valor del conjunto de datos,  $e$  y es su valor transformado o estandarizado.

## Algoritmo implementado – Red Neuronal Perceptrón Multicapa

El algoritmo backpropagation de red neuronal implementado, consta de los siguientes pasos:

1. Seleccionar el número de nodos en la capa de entrada, capa oculta y capa de salida.
2. Inicializar las tasas de aprendizaje y de interacción máxima (configurar los umbrales de variación de todos los pesos en cada una de las capas).
3. Asociar los vectores de entrada a los nodos de entrada y los vectores de salida al nodo de salida.
4. Calcular los valores de entrada a los nodos de la capa oculta a partir de la siguiente ecuación:

$$S_j = \sum_{i=1}^n x_i W_{ij} \quad \text{Ecu. 12}$$

Donde  $x_i$  es la variable de entrada al nodo  $i$  y  $W_{ij}$  es el peso del nodo de entrada  $i$  al nodo de la capa oculta  $j$ . Las salidas de los nodos de la capa oculta se derivan de acuerdo a la siguiente ecuación:

$$Y_j = f(S_j) = \frac{1}{1 + e^{-S_j}} \quad \text{Ecu. 13}$$

Donde  $Y_j$  es la variable de salida del nodo  $j$  de la capa oculta. Este mismo algoritmo fue empleado para calcular las entradas a los nodos de la capa de salida.

5. Calcular el error de simulación en la capa de salida.
6. Determinar el número óptimo de neuronas en la capa oculta de manera empírica, examinando diferentes arquitecturas hasta alcanzar el menor error cuadrático medio estandarizado (RMSEn), obtenido a partir de validación cruzada entre los valores observados y simulados.

Una vez obtenido el modelo de red neuronal con mejor desempeño en cada uno de los ambientes, se analizó cualitativamente la arquitectura obtenida y sus conexiones entre capas, a través de Diagramas de Interpretación Neuronal (NID por sus siglas en inglés) (Özesmi & Özesmi, 1999).

### **Importancia relativa de las variables predictoras**

Para determinar la importancia relativa de las variables predictoras, respecto a la variable de respuesta, se implementó el algoritmo de relevancia métrica propuesto por Garson & David (1991) y Goh (1995). Los detalles del algoritmo matemático utilizado se muestran en la Tabla 3.

## Capítulo 3.

### Resultados

“Pero tú, persevera hasta el fin y descansa, que al final de los tiempos te levantarás para recibir tu recompensa”

**Daniel 12:13 (NVI)**

En este capítulo se presenta inicialmente los resultados obtenidos durante el proceso de calibración y validación del modelo de cultivo Oryza2000, para las variedades Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21375.

Para el proceso de calibración, se detallan los valores finales de los parámetros genéticos que modulan, dentro del modelo, el desarrollo fenológico; los factores de distribución de asimilados en hojas verdes, hojas muertas, tallos y panículas; y el desarrollo foliar de cada una de las variedades. Posteriormente, se muestran los resultados obtenidos en el proceso de validación del modelo de cultivo, a partir de una serie de gráficas y medidas estadísticas generalmente son utilizadas para este propósito.

Como segundo punto, se presentan los ambientales identificados a partir de técnicas estadísticas multivariadas de agrupamiento sobre los rendimientos simulados; se presentan mapas que muestran la distribución espacial de los grupos ambientales y la estabilidad de cada uno de ellos; y se detallan sus principales características agroclimáticas.

---

Después, se presentan los modelos de red neuronal artificial supervisado que han sido utilizados para identificar la importancia relativa de diferentes variables ambientales sobre el rendimiento, además se presenta el principal estrés abiótico que limita el crecimiento, desarrollo y rendimiento del cultivo en cada grupo ambiental. Finalmente, se presentan algunos resultados gráficos que detallan el comportamiento entre el estrés abiótico identificado y algunas características fisiológicas del cultivo.

## **3.1 Evaluación del modelo Oryza2000**

### **3.1.1 Tasas de desarrollo fenológico (DVR)**

Las tasas de desarrollo se determinaron a partir de las fechas fenológicas registradas en cada uno de los experimentos considerados para el proceso de calibración del modelo (Tabla 6). En la Tabla 8 se muestran las DVR obtenidas para cada una de las variedades en estudio.

Con base en los resultados obtenidos, la variedad Fedearroz 2000 tuvo el menor valor de DVRJ y el mayor valor de DVRR, generando que la duración promedio de la etapa vegetativa fuera mayor en  $\pm 3$  días, y que la fase de llenado de grano se hiciera más corta ( $\pm 2$  días), en relación a los demás cultivares. Por su parte, la variedad Fedearroz 60 tuvo un comportamiento totalmente opuesto (mayor DVRJ, mayor DVRR y menor DVRR), ocasionando que tanto la etapa vegetativa como de desarrollo panicular, se completará en un menor número de días, y que la etapa reproductiva se retrasara en promedio 4 días.

Respecto a las variedades Fedearroz 733 y CT21375, las DVR obtenidas se ubican en el promedio del rango de valores extremos establecidos por las variedades Fedearroz 2000 y Fedearroz 60.

En general, las tasas de desarrollo obtenidas para cada variedad fueron bastante estables (CV <10%) a través de las diferentes condiciones climáticas para las cuales fueron calculadas. Por tanto, las DVR registradas en la Tabla 8 pueden ser consideradas para simular la fenología de las mismas variedades en una amplia gama de condiciones ambientales.

Tabla 8. Tasas de desarrollo fenológico promedio de las variedades Fedearroz 2000, Fedearroz 60, Fedearroz 733 y CT21375, para la calibración y validación del modelo.

<b>Variedad</b>	<b>DVRJ</b>	<b>DVRI</b>	<b>DVRP</b>	<b>DVRR</b>
Fed2000	0.0006413	0.0007576	0.0007175	0.0019237
CT21375	0.0006473	0.0007576	0.0007172	0.0017759
Fed60	0.0006865	0.0007576	0.0007261	0.0016241
Fed733	0.0006663	0.0007576	0.0007015	0.0017680

DVRJ, DVRI, DVRP y DVRR – Tasas de desarrollo fenológico ( $^{\circ}\text{C día}^{-1}$ ) en las fase vegetativa, sensible al fotoperiodo, desarrollo panicular y reproductiva, respectivamente.

Las diferencias entre los valores observados y simulados de la duración de las etapas fenológicas entre emergencia y: Iniciación de panícula, floración y madurez fisiológica, para el grupo de datos de calibración y validación, se observan en la Tabla 9 y la Tabla 10, respectivamente. Las métricas estadísticas que evalúan el desempeño del modelo para estimar la duración de cada una de las etapas fenológicas mencionadas, se muestran en la Tabla 11 y Tabla 12.

En general, el modelo predice relativamente bien la duración de cada una de los estados fenológicos, no obstante, consistentemente sub estima la duración total de todo el ciclo del cultivo con un error entre 2 – 4 días.

Tabla 9. Comparación entre la fenología observada y simulada, respecto a los días después de emergencia del cultivo – Datos de Calibración.

Variedad	Experimento	Días entre Eme - Ini. Pan		Días entre Eme - Flo		Días entre Eme - Mad. Fis	
		Obs	Sim	Obs	Sim	Obs	Sim
Fed2000	MON-S1COL	59	55	88	81	113	110
	MON-S3MADR	50	54	76	80	116	111
	MON-S4MADR	56	54	77	81	110	113
	SAL-S3COL	59	56	83	80	113	110
	SAL-S2MADR	53	53	89	84	108	108
	SAL-S3MADR	45	49	80	81	118	113
CT21375	MON-S1COL	58	53	87	83	113	112
	MON-S3MADR	52	54	78	81	118	115
	MON-S4MADR	54	53	76	81	109	112
	SAL-S3COL	59	56	83	79	113	112
	SAL-S2MADR	46	49	90	84	116	113
	SAL-S3MADR	50	52	78	80	118	114
Fed60	MON-S3MADR	49	52	76	79	118	114
	MON-S4MADR	56	53	78	79	109	112
	SAL-S2MADR	52	50	89	82	116	111
	SAL-S3MADR	47	50	74	78	118	114
Fed733	MON-S3MADR	50	53	77	81	117	113
	MON-S4MADR	55	52	78	81	110	114
	SAL-S3COL	60	55	83	79	113	112
	SAL-S2MADR	46	51	91	85	116	110
	SAL-S3MADR	50	52	82	80	118	115

SAL – Experimentos realizados en Saldaña (Tolima); MON – Experimentos realizados en Montería (Córdoba); MADR – Experimentos realizados en el proyecto con el Ministerio de Agricultura; COL – Experimentos realizados en el marco del proyecto con Colciencias.

Tabla 10. Comparación entre la fenología observada y simulada, respecto a los días después de emergencia del cultivo – Datos de validación.

Variedad	Experimento	Días entre Eme - Ini. Pan		Días entre Eme - Flo		Días entre Eme - Mad. Fis	
		Obs	Sim	Obs	Sim	Obs	Sim
Fed2000	AIP-S1MADR	48	53	95	90	137	125
	AIP-S2MADR	55	57	88	84	109	113
	AIP-S3MADR	54	54	85	83	114	111
	AIP-S4MADR	54	57	82	83	122	116
	SAL-S1MADR	55	54	85	81	115	112
	SAL-S2COL	61	55	86	83	111	112
	SAL-S4COL	59	54	81	80	112	110
	MON-S2COL	52	56	77	83	116	116
CT21375	AIP-S1MADR	51	55	95	90	120	120
	AIP-S2MADR	54	56	83	83	109	114
	AIP-S3MADR	52	53	83	81	114	113
	AIP-S4MADR	52	56	81	84	122	117
	SAL-S1MADR	48	52	81	81	114	114
	SAL-S2COL	61	55	86	83	111	114
	SAL-S4COL	59	55	81	79	112	111
	MON-S2COL	52	56	72	84	126	122
Fed60	AIP-S1MADR	51	55	94	90	121	121
	AIP-S2MADR	55	54	88	84	109	114
	AIP-S3MADR	54	51	85	82	113	114
	AIP-S4MADR	54	54	85	82	122	118
	SAL-S1MADR	49	51	82	79	111	114
Fed733	AIP-S1MADR	49	53	95	90	120	120
	AIP-S2MADR	54	55	85	82	109	114
	AIP-S3MADR	53	52	83	80	115	113
	AIP-S4MADR	54	55	84	83	116	115
	SAL-S1MADR	48	53	83	80	115	114

SAL – Experimentos realizados en Saldaña (Tolima); AIP – Aipe (Huila); MON – Experimentos realizados en Montería (Córdoba); MADR – Experimentos realizados en el proyecto con el Ministerio de Agricultura; COL – Experimentos realizados en el marco del proyecto con Colciencias.

Tabla 11. Evaluación general del desempeño del modelo Oryza2000 para simular la fenología de las variedades Fedearroz 2000, Fedearroz 60, Fedearroz 733 y CT21375 - Datos de calibración.

Variable (Días)	Num	X <sub>mean</sub>	X <sub>sd</sub>	Y <sub>mean</sub>	Y <sub>sd</sub>	$\alpha$	$\beta$	R <sup>2</sup>	Pt	RMSE <sub>a</sub>	RMSE <sub>n</sub>
Eme - Ini. Pan	21	53	5	53	2	0.35	33.9	0.67	0.96	3	6.1
Eme - Flo	21	82	5	81	2	0.24	60.8	0.53	0.31	4	5.2
Eme - Mad. Fis	21	114	3	112	2	0.26	82.4	0.45	0.06	4	3.1

Tabla 12. Evaluación general del desempeño del modelo Oryza 2000 para simular la fenología de las variedades Fedearroz 2000, Fedearroz 60, Fedearroz 733 y CT21375 - Datos de validación.

Variabes (Días)	Num	X <sub>mean</sub>	X <sub>sd</sub>	Y <sub>mean</sub>	Y <sub>sd</sub>	$\alpha$	$\beta$	R <sup>2</sup>	Pt	RMSE <sub>a</sub>	RMSE <sub>n</sub>
Eme - Ini. Pan	26	53	4	54	2	0.17	45.27	0.14	0.37	3	6.5
Eme - Flo	26	85	5	83	3	0.45	45.33	0.54	0.23	4	4.7
Eme - Mad. Fis	26	116	6	115	4	0.49	58.55	0.72	0.84	4	3.3

Num – Número de par de datos; X<sub>mean</sub> – Promedio de datos observados; X<sub>sd</sub> – Desviación estándar de datos observados; Y<sub>mean</sub> – Promedio de datos simulados; Y<sub>sd</sub> – Desviación estándar de datos simulados;  $\alpha$  – Pendiente de la regresión lineal entre los datos simulados y datos observados;  $\beta$  - Intercepto de la regresión lineal entre los datos simulados y datos observados; R<sup>2</sup> – Coeficiente de determinación de la regresión; Pt - Prueba de medias t-Student asumiendo varianza desigual (95 % de nivel de significancia); RMSE<sub>a</sub> – Error medio cuadrático absoluto; RMSE<sub>n</sub> – Error medio cuadrático normalizado.

La Figura. 20 muestra la fenología observada vs la fenología simulada para los datos de calibración, en las localidades de Saldaña y Montería. De igual forma, la Figura. 21 muestra la relación entre la fenología observada vs la fenología simulada, para el grupo de datos de validación, en las localidades de Saldaña y Aipe.

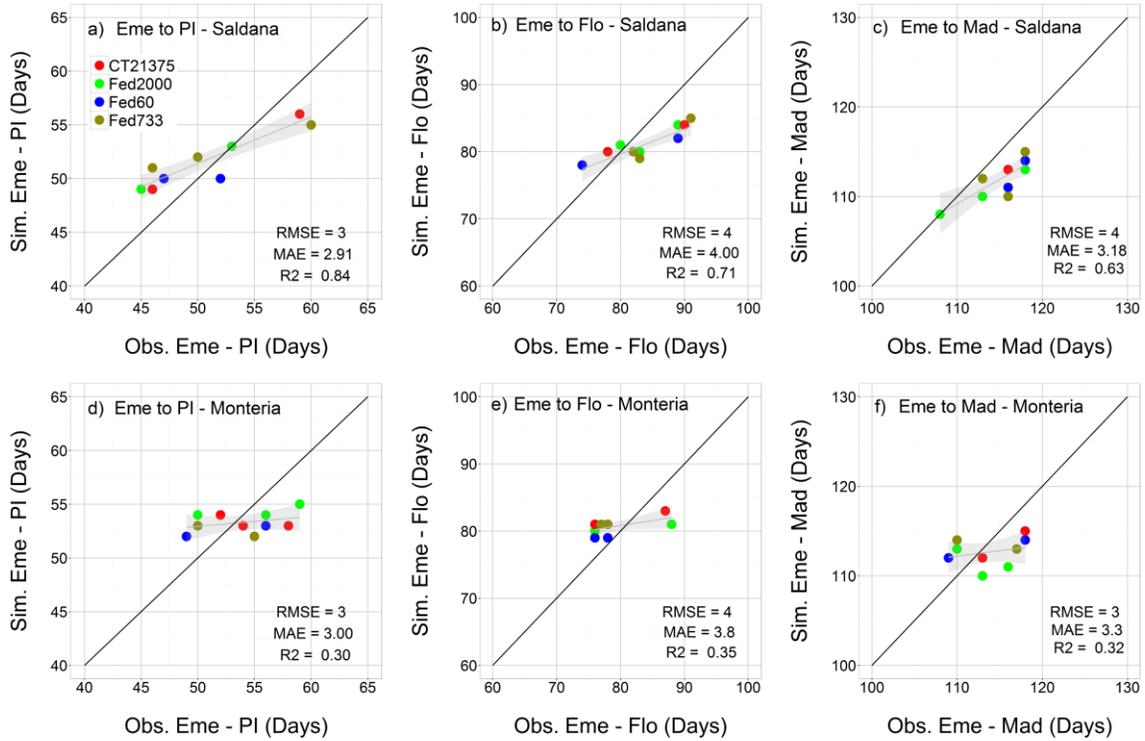


Figura. 20. Fenología: Observado vs simulado: a-c) Saldaña; d-f) Montería – Calibración

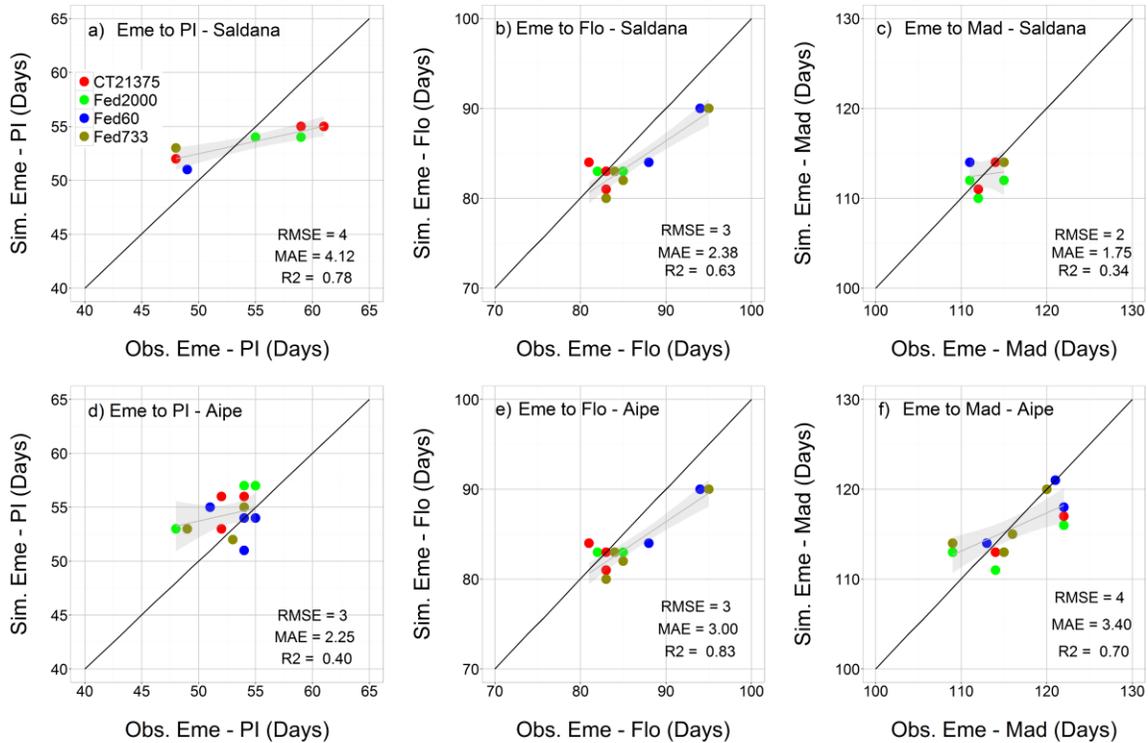


Figura. 21. Fenología: Observado vs simulado: a-c) Saldaña; d-f) Aipe - Validación  
Linea solida representa la relación 1:1.

### 3.1.2 Parámetros de producción de biomasa y desarrollo foliar

Los parámetros de distribución de biomasa, para cada variedad, se definieron con base en la fracción de materia seca disponible en cada órgano de la planta, a partir del conjunto de datos de calibración, como se muestra en la Tabla 13.

Tabla 13. Fracción de materia seca distribuida a las hojas verdes (FLV), hojas muertas (DRLV), tallos (FST) y panículas (FSO), en función del estado de desarrollo del cultivo (DVS).

Cultivar	DVS	FLV	DRLV	FST	FSO
Fedearroz 2000	0.00	0.76	0.00	0.24	0.00
	0.50	0.51	0.00	0.49	0.00
	0.64	0.18	0.00	0.82	0.00
	0.84	0.24	0.02	0.60	0.16
	1.00	0.02	0.03	0.13	0.86
	1.30	0.00	0.02	0.01	0.99
	2.50	0.31	0.06	0.00	0.69
Fedearroz 733	0.00	0.51	0.00	0.49	0.00
	0.50	0.53	0.00	0.47	0.00
	0.64	0.18	0.00	0.82	0.00
	0.84	0.24	0.02	0.60	0.16
	1.00	0.18	0.02	0.13	0.69
	1.30	0.08	0.02	0.01	0.91
	2.50	0.48	0.04	0.00	0.52
Fedearroz 60	0.00	0.76	0.00	0.24	0.00
	0.50	0.51	0.00	0.49	0.00
	0.64	0.18	0.00	0.82	0.00
	0.84	0.24	0.02	0.60	0.16
	1.00	0.39	0.02	0.13	0.48
	1.30	0.12	0.02	0.01	0.87
	2.50	0.50	0.06	0.00	0.50
CT2175	0.00	0.53	0.00	0.47	0.00
	0.50	0.62	0.00	0.38	0.00
	0.64	0.18	0.00	0.82	0.00
	0.84	0.24	0.02	0.60	0.16
	1.00	0.01	0.03	0.12	0.87
	1.30	0.00	0.02	0.01	0.99
	2.50	0.16	0.08	0.00	0.84

El estado de desarrollo del cultivo define su edad fisiológica y está relacionado con el proceso de formación de varios órganos y con el momento en que ocurre. Como muchos procesos fisiológicos y morfológicos se llevan a cabo con los cambios en las etapas fenológicas de la planta, la cuantificación exacta del estado fenológico es esencial en cualquier modelo de simulación de cultivos. En el modelo ORYZA2000, la fenología del cultivo de arroz se define a través del parámetro DVS, el cual toma un valor de (0) para indicar el estado de emergencia, (0.65) para la iniciación de la panícula, (1) para la floración y (2) para la madurez fisiológica (B. a M. Bouman et al., 2001).

De acuerdo a los resultados obtenidos, se observa que al inicio del ciclo de cultivo (DVS= 0 – 0.5) la mayor parte de asimilados son direccionados hacia las hojas de los distintos cultivares, con una fracción un poco mayor para el cultivar Fedearroz 60 (Figura. 22a). Posteriormente se observa una continua disminución hasta la etapa de floración (DVS=1). Por su parte, los asimilados que se encuentran en los tallos (cerca del 50% del total en la planta) se mantienen relativamente constantes hasta el momento previo a la iniciación de la panícula (DVS= 0 - 0.61), etapa en la cual se aprecia un marcado aumento que alcanza cerca del 80% del total de asimilados, después se evidencia una continua disminución hasta el punto de floración de cada una de las variedades (DVS=1) (Figura. 22b). Hasta este punto, el total de materia seca en cada una de las variedades es distribuida hacia las hojas y tallos, para luego ser traslocada hacia las panículas después de la floración.

Entre DVS= 1 - 1.5 se puede observar como la mayor parte de asimilados se transfieren hacia las panículas de la planta, siendo notablemente superior en las variedades Fedearroz 2000 y CT21375 (100%), lo cual contrasta con la cantidad traslocada en la variedad Fedearroz 60, que presenta los menores valores (80%) (Figura. 22c). Entre DVS = 1.5 – 2, se aprecia que todas las variedades disminuyen el contenido de materia seca en las panículas, siendo la variedad CT21375 la que termina todo el ciclo de cultivo con la mayor cantidad (84%), seguida por Fedearroz 2000 (69%), Fedearroz 733 (52%) y Fedearroz 60 (50%).

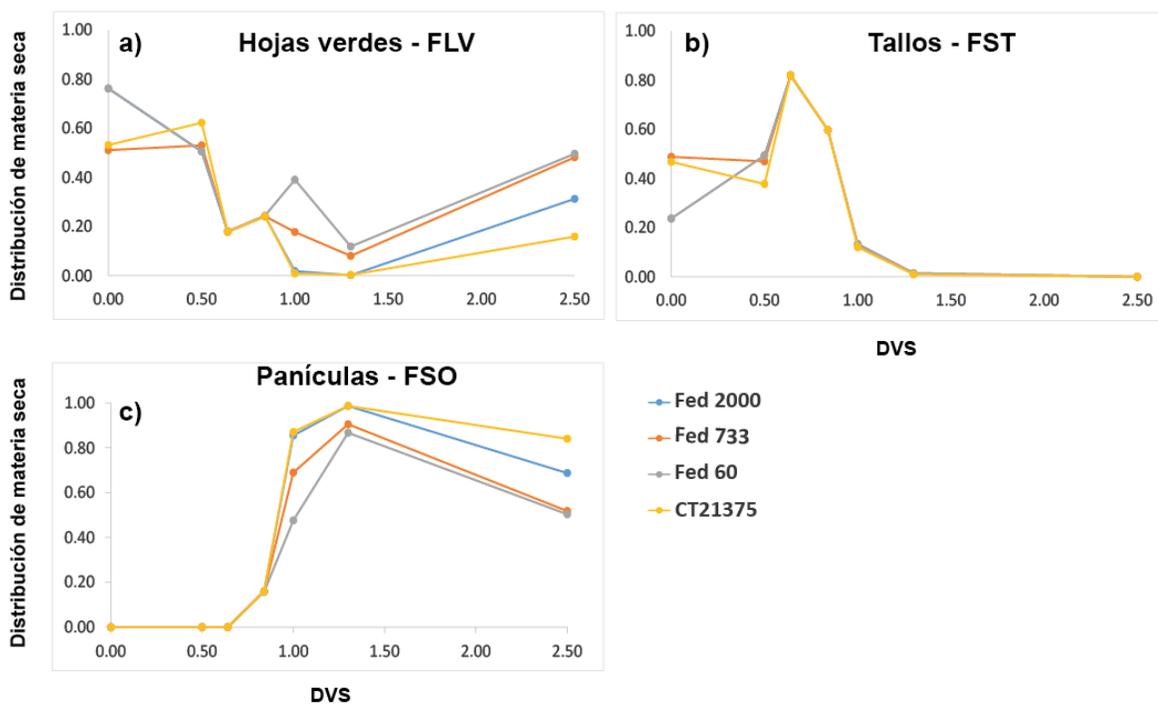


Figura. 22. Dinámica de distribución de asimilados en: a) Hojas (FLV); b) Tallos (FST); y Panículas (FSO), en función del desarrollo fenológico (DVS), para las variedades Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21375.

El desarrollo foliar de la planta se parametrizó a partir del coeficiente RGRL (tasa de crecimiento por unidad de área), el cual modula el desarrollo de las hojas en la fase de crecimiento exponencial del cultivo (desde la emergencia hasta el momento en que las hojas alcanzan un valor de LAI = 1). El área de las hojas determina la cantidad de luz interceptada para el proceso de fotosíntesis, y por tanto, permite determinar la cantidad de biomasa que puede ser acumulada por la planta (B. a M. Bouman et al., 2001).

En ORYZA2000, el parámetro RGRL se determina a partir de los valores mínimos (RGRLMN) y máximos (RGRLMX) del crecimiento relativo del área de las hojas. Estudios realizados por Kropff, H.H. Laar, & Matthews (1994) y B. a M. Bouman et al. (2001), reportaron los mayores valores de RGRL en ensayos experimentales en los que la planta no tuvo limitaciones de Nitrógeno (RGRLMX ( $^{\circ}\text{C día}^{-1}$ ) 0.0080 – 0.0085); mientras que los menores valores de RGRL se obtuvieron en ensayos con limitaciones de Nitrógeno (RGRLMN ( $^{\circ}\text{C día}^{-1}$ ) 0.0045 – 0.0050). El valor mínimo de RGRL se produjo cuando el contenido de N en las hojas ( $\text{kg N kg}^{-1}$  de materia seca) se redujo al 90% de su valor máximo.

En la Tabla 14 se presentan los valores de RGRLMN y RGRLMX obtenidos para los cultivares en estudio, los cuales se obtuvieron a partir de un proceso iterativo que tomó como parámetros iniciales los valores registrados para la variedad IR72 (RGRLMN = 0.0040 (°C día<sup>-1</sup>) y RGRLMX = 0.0085 (°C día<sup>-1</sup>)). En general, todas las variedades tuvieron valores mínimos y máximos de RGRL muy cercanos a los registrados en la literatura (0.004 °C día<sup>-1</sup> y 0.008 °C día<sup>-1</sup>, respectivamente), sin presentar diferencias significativas entre variedades. La variedad que obtuvo el mayor RGRL fue Fedearroz 2000 (0.00801 °C día<sup>-1</sup>), seguida por CT21375 (0.00799 °C día<sup>-1</sup>).

Tabla 14. Tasas de crecimiento relativo de hojas, mínimo (RGRLMN (°C día<sup>-1</sup>)) y máximo (RGRLMX (°C día<sup>-1</sup>)), para la variedades Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21375.

Cultivar	RGRLMN	RGRLMX
Fedearroz 2000	0.00373	0.00801
Fedearroz 733	0.00394	0.00799
Fedearroz 60	0.00404	0.00786
CT21375	0.00410	0.00799

Después de las primeras etapas de crecimiento del cultivo, el espesor de las hojas es otra característica importante que determina su tamaño. El espesor o grosor de la hoja está definido como el Área Foliar Específica (SLA), el cual hace referencia al área de las hojas por unidad de peso (ha kg<sup>-1</sup>). Las hojas de un cierto peso tienen mayor área superficial cuando su área foliar específica es mayor, por lo tanto, interceptan mayor luz solar, y muestran mayor actividad fotosintética. El SLA generalmente cambia a medida que avanza el crecimiento del cultivo, tomando valores más bajos a medida que la planta se acerca a la etapa de madurez fisiológica (B. a M. Bouman et al., 2001).

En la Tabla 15 se muestran los valores calibrados de SLA para cada una de las variedades, y en la Figura. 23 se observa su dinámica en función de la fenología del cultivo (DVS). Los valores calibrados de SLA se obtuvieron a partir de la relación de los registros secuenciales de materia seca en hojas verdes y su respectiva área foliar. Todos los datos se tomaron del grupo de experimentos considerados para la parametrización del modelo. Para cada

variedad, el SLA fue optimizado a partir de un proceso iterativo, que se inició a partir de los valores establecidos para la variedad de referencia IR72.

Tabla 15. Valores de Área Foliar Específica (SLA), parametrizados para las variedades Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21375.

Área Foliar Específica (ha kg <sup>-1</sup> )				
DVS	Fed 2000	Fed 733	Fed 60	CT21375
0.00	0.00279	0.00319	0.00279	0.00323
0.25	0.00263	0.00152	0.00263	0.00196
0.33	0.00184	0.00168	0.00181	0.00176
0.64	0.00173	0.00176	0.00176	0.00169
0.93	0.00174	0.00171	0.00171	0.00171
1.30	0.00148	0.00094	0.00105	0.00152
1.75	0.00071	0.00067	0.00070	0.00128
2.10	0.00066	0.00095	0.00066	0.00066
2.50	0.00066	0.00085	0.00085	0.00093

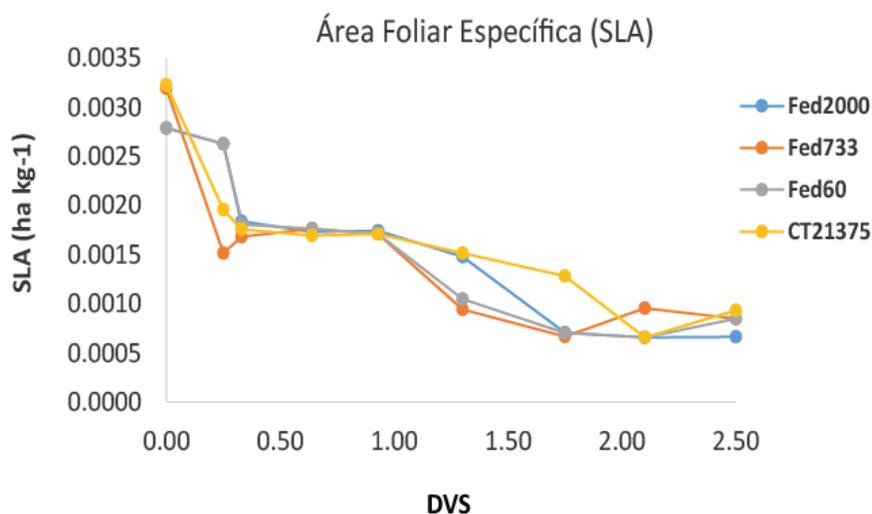


Figura. 23. Dinámica del Área Foliar Específica (SLA), parametrizada para las variedades Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21375.

### 3.1.3 Dinámica de Biomasa y Área Foliar (Datos de Calibración)

Las dinámicas de producción de materia seca total, en hojas, tallos, panículas; y el área foliar específica del cultivo, fueron bien simuladas para todas las variedades del grupo de ensayos de calibración, establecidos en las localidades de Saldaña (Figura. 24) y Montería (Figura. 27). No obstante, se observó un mejor ajuste en la dinámica de materia seca simulada para los experimentos de Saldaña, en relación a la materia seca simulada para las diferentes variedades en la localidad de Montería.

La comparación entre los datos observados y simulados de las variables de crecimiento del cultivo, mostró que los mejores resultados de simulación se obtuvieron para la biomasa total. La mayoría de los puntos se ubicaron dentro de los intervalos de confianza establecidos a lo largo de la recta 1:1 (Figura. 26a y Figura. 29a). De igual forma, el LAI fue otra de las variables para las cuales el modelo mostró un gran desempeño (Figura. 26e y Figura. 29e).

Por otro lado, la relación entre biomasa observada y simulada en hojas verdes, fue la que mostró mayor dispersión alrededor de la recta 1:1, a pesar de que el modelo capturó muy bien la dinámica de producción de biomasa, hubo un par de experimentos en la localidad de Montería en los que se sobre estimó la biomasa producida, especialmente para las variedades Fedearroz 60 y Fedearroz 733, después de la floración del cultivo. En Saldaña, también se apreció un poco de sobre estimación en esta variable, principalmente en la etapa vegetativa.

En cuanto a las panículas, hubo dos experimentos en Montería (para las variedades Fedearroz 2000 y CT21375) en los cuales el modelo sub estimó la biomasa final alcanzada; mientras que en Saldaña, se observó una mayor dispersión en la biomasa simulada de este órgano de la planta (Figura. 26d y Figura. 29d).

Las mayores diferencias entre los datos observados y simulados se observó en la biomasa de tallos, especialmente en el periodo entre floración y madurez fisiológica, en el cual el modelo mostró una constante sub estimación de la biomasa seca producida. Este comportamiento se evidenció con mayor claridad en el grupo de datos de calibración seleccionados en la localidad de Montería (Figura. 26c y Figura. 29c).

Respecto al rendimiento simulado, en Saldaña todos los valores se ubicaron por encima de la recta 1:1, indicando una sobre estimación respecto a los valores observados (Figura. 26f). Por su parte, el rendimiento simulado en Montería tuvo un mejor ajuste respecto a los rendimientos observados (Figura. 29f).

En la Tabla 16 se muestran los parámetros de bondad de ajuste para todo el conjunto de datos de calibración (Dinámicas de producción de biomasa, índice de área foliar y rendimiento). Hubo algunas variaciones en las métricas estadísticas, en cuanto a experimentos y variedades, pero en términos generales el patrón de ajuste fue consistente en todos los casos.

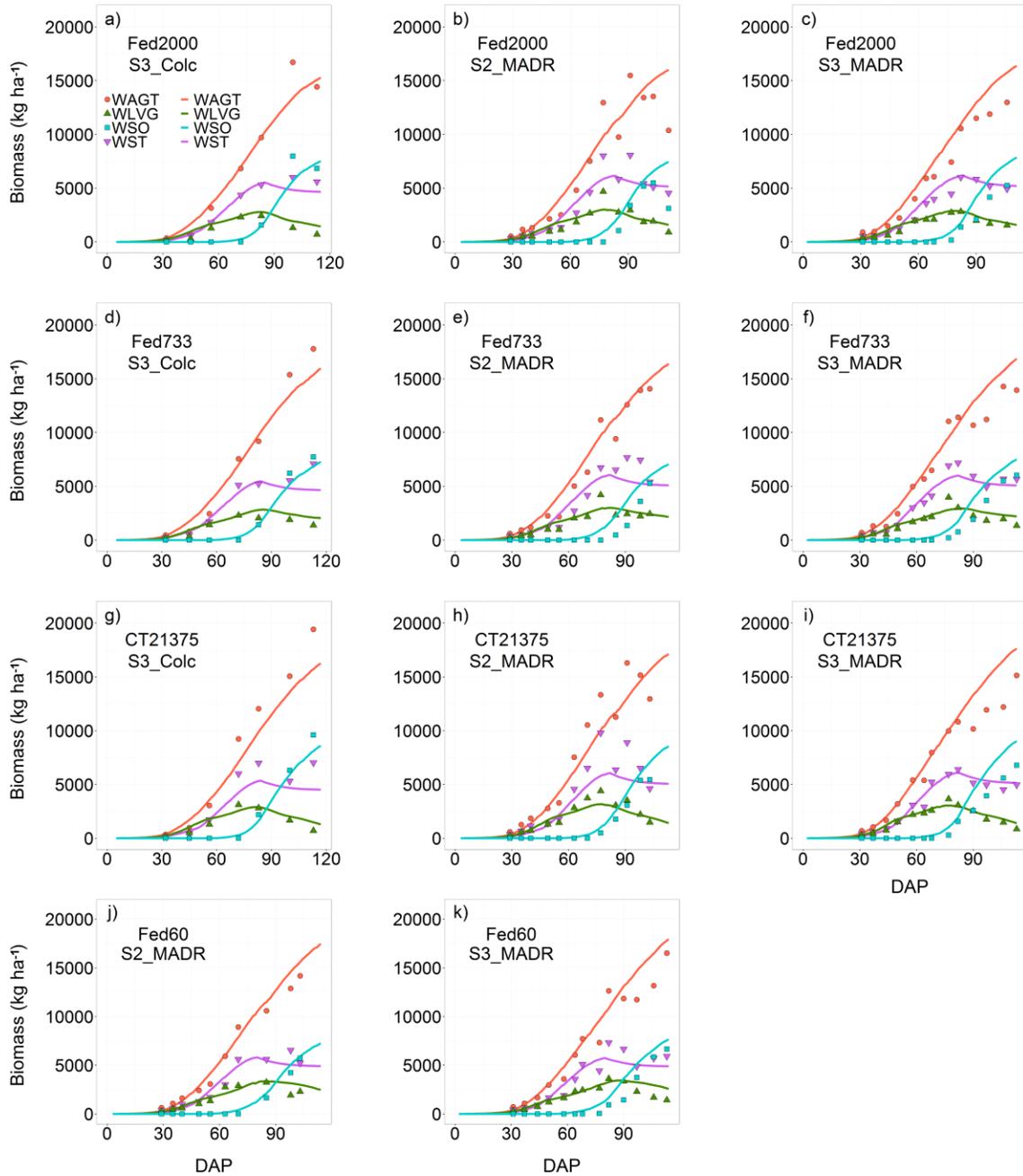


Figura. 24. Biomasa observada (puntos) y simulada (líneas) de materia seca total (WAGT), hojas verdes (WLVG), panículas (WSO), y tallos (WST), en función de los días después de emergencia (DAE); de los ensayos experimentales realizados durante los años 2013 – 2015 (Datos de calibración), en Saldaña – Tolima; para las variedades a-c) Fedearroz 2000; d-f) Fedearroz 733; g-i) CT21375 y j-k) Fedearroz 60.

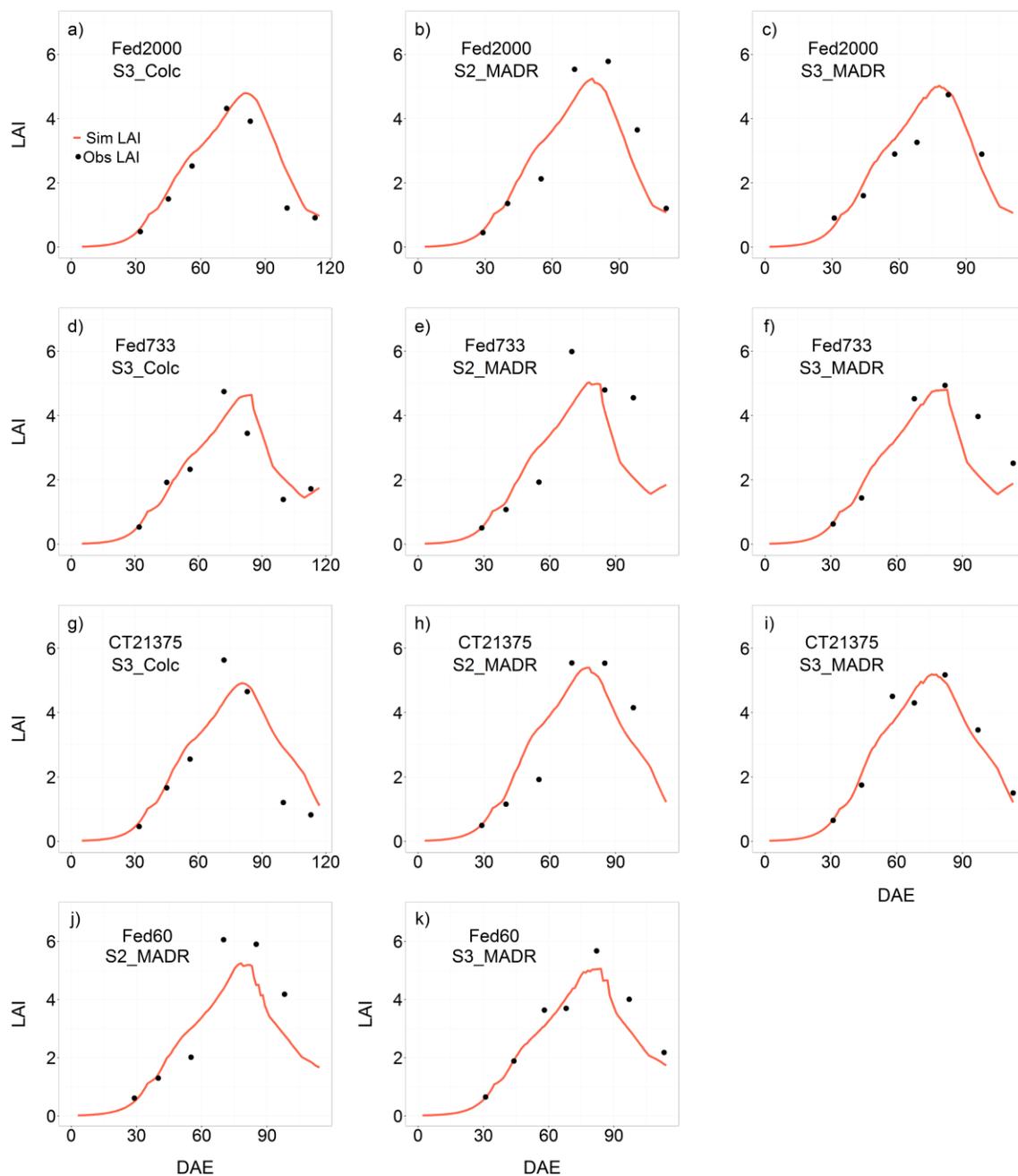


Figura. 25. Índice de Área Foliar (LAI) observado (puntos) y simulado (líneas), en función de los días después de emergencia (DAE); de los ensayos experimentales realizados durante los años 2013 – 2015 (Datos de calibración), en Saldaña – Tolima; para las variedades a-c) Fedearroz 2000; d-f) Fedearroz 733; g-i) CT21375 y j-k) Fedearroz 60.

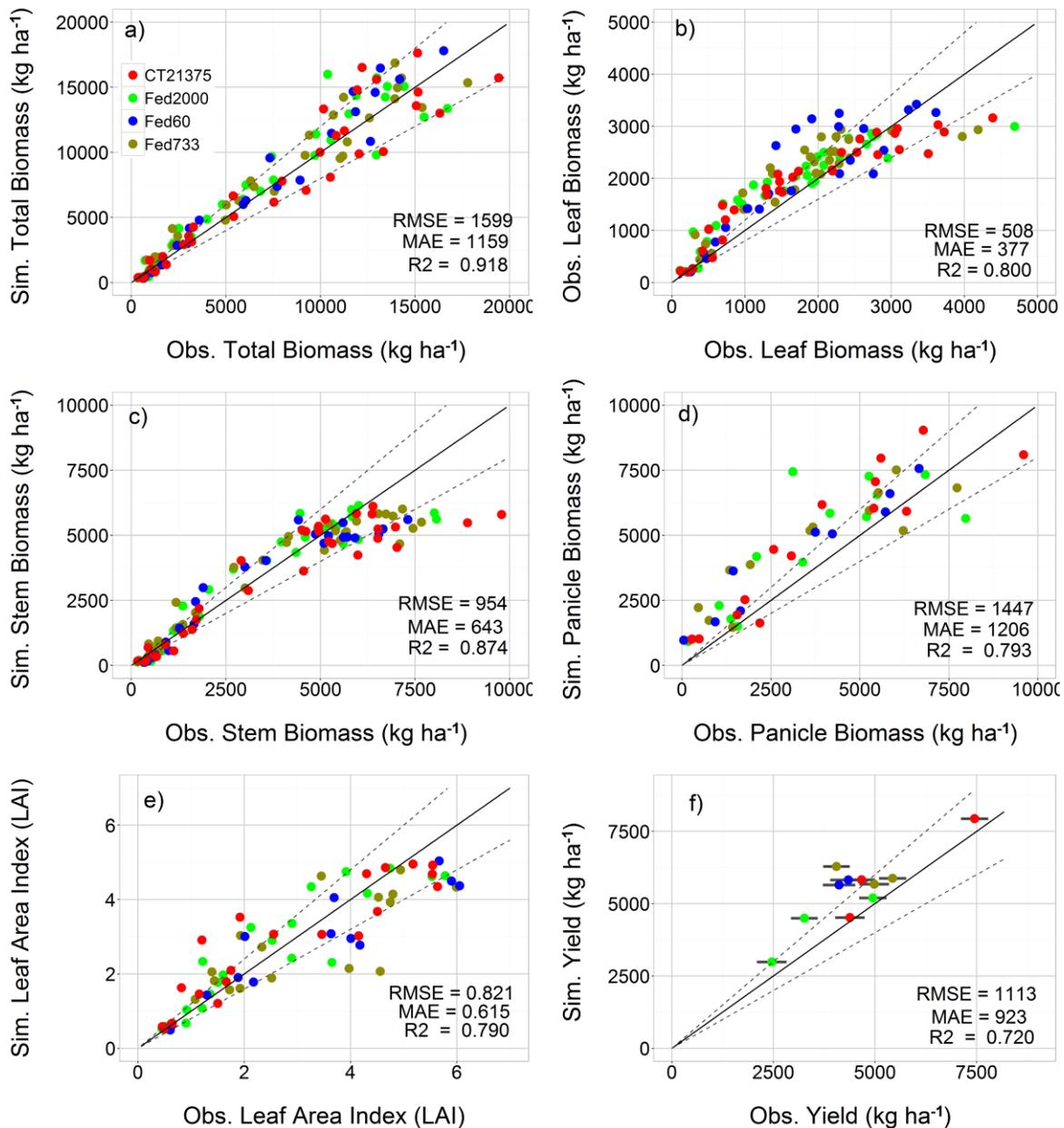


Figura. 26. Datos observados vs simulados de a) biomasa total; b) biomasa en hojas; c) biomasa en tallos; d) biomasa en panículas; e) Índice de Área Foliar (LAI); y f) rendimiento del cultivo, para el grupo de datos de calibración en Saldaña-Tolima; para las variedades CT21375 (puntos rojos), Fedearroz 2000 (puntos verdes), Fedearroz 60 (puntos azules) y Fedearroz 733 (puntos cafés). La línea sólida es la recta de relación 1:1; Las líneas punteadas corresponden a la desviación estándar de los datos observados, alrededor de la recta 1:1. RMSE – Error medio cuadrado; MAE – Error medio absoluto; R<sup>2</sup> – Coeficiente de determinación.

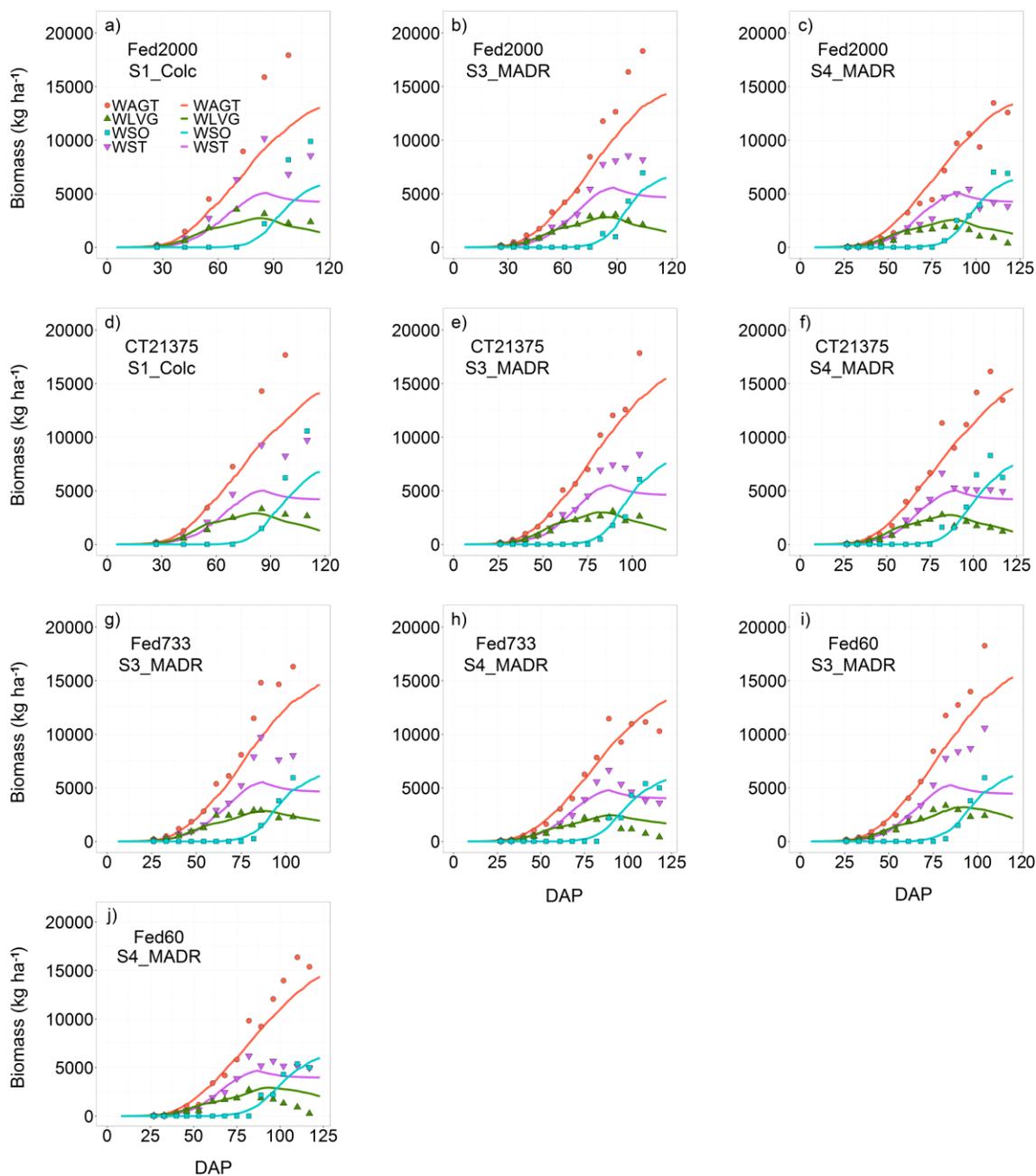


Figura. 27. Biomasa observada (puntos) y simulada (líneas) de materia seca total (WAGT), hojas verdes (WLVG), panículas (WSO), y tallos (WST), en función de los días después de emergencia (DAE); de los ensayos experimentales realizados durante los años 2013 – 2015 (Datos de calibración), en Montería – Córdoba; para las variedades a-c) Fedearroz 2000; d-f) CT21375; g-h) Fedearroz 733 e i-j) Fedearroz 60.

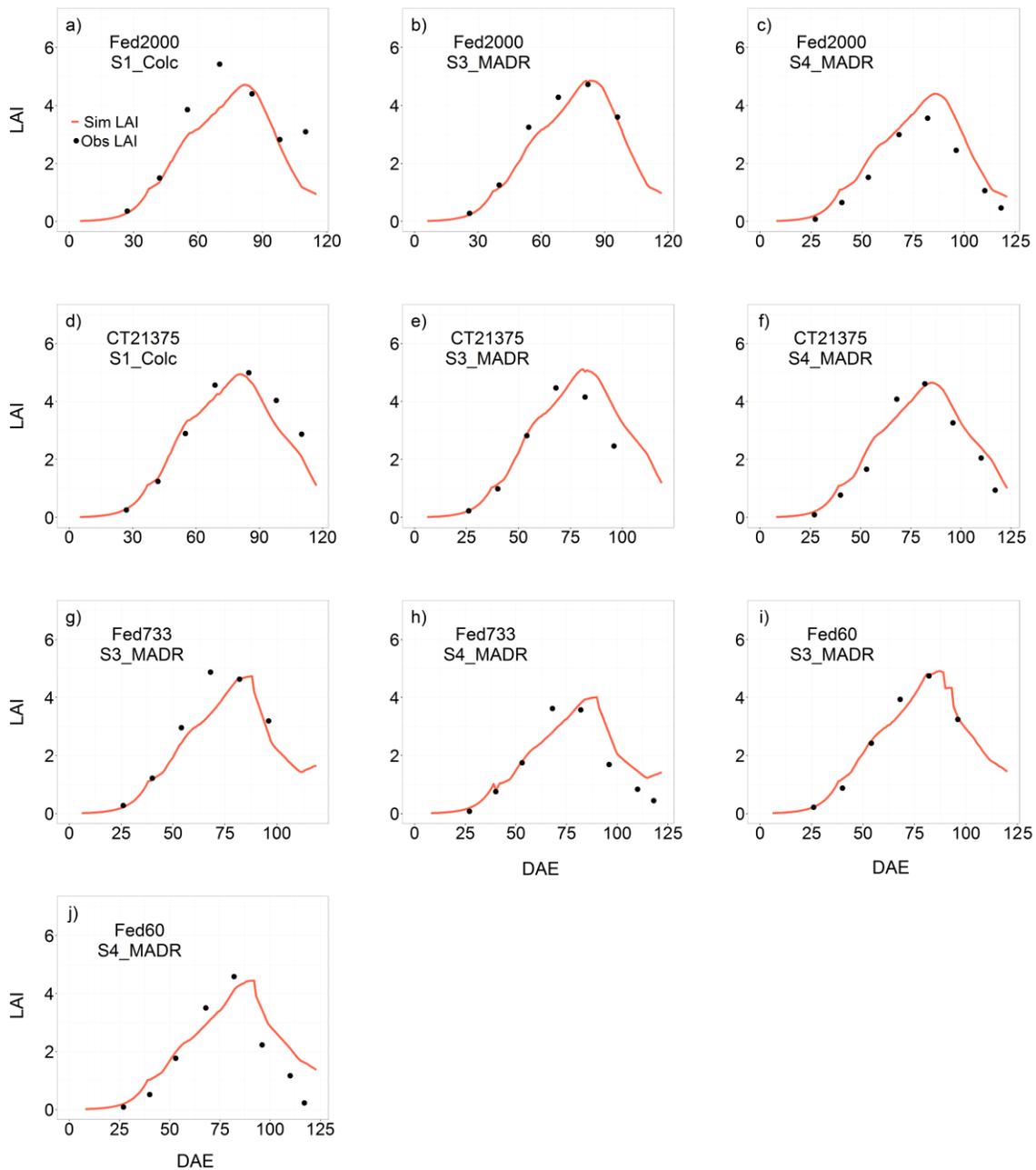


Figura. 28. Índice de Área Foliar (LAI) observado (puntos) y simulado (líneas), en función de los días después de emergencia (DAE); de los ensayos experimentales realizados durante los años 2013 – 2015 (Datos de calibración), en Montería - Córdoba; para las variedades a-c) Fedearroz 2000; d-f) CT21375; g-h) Fedearroz 733 e i-j) Fedearroz 60.

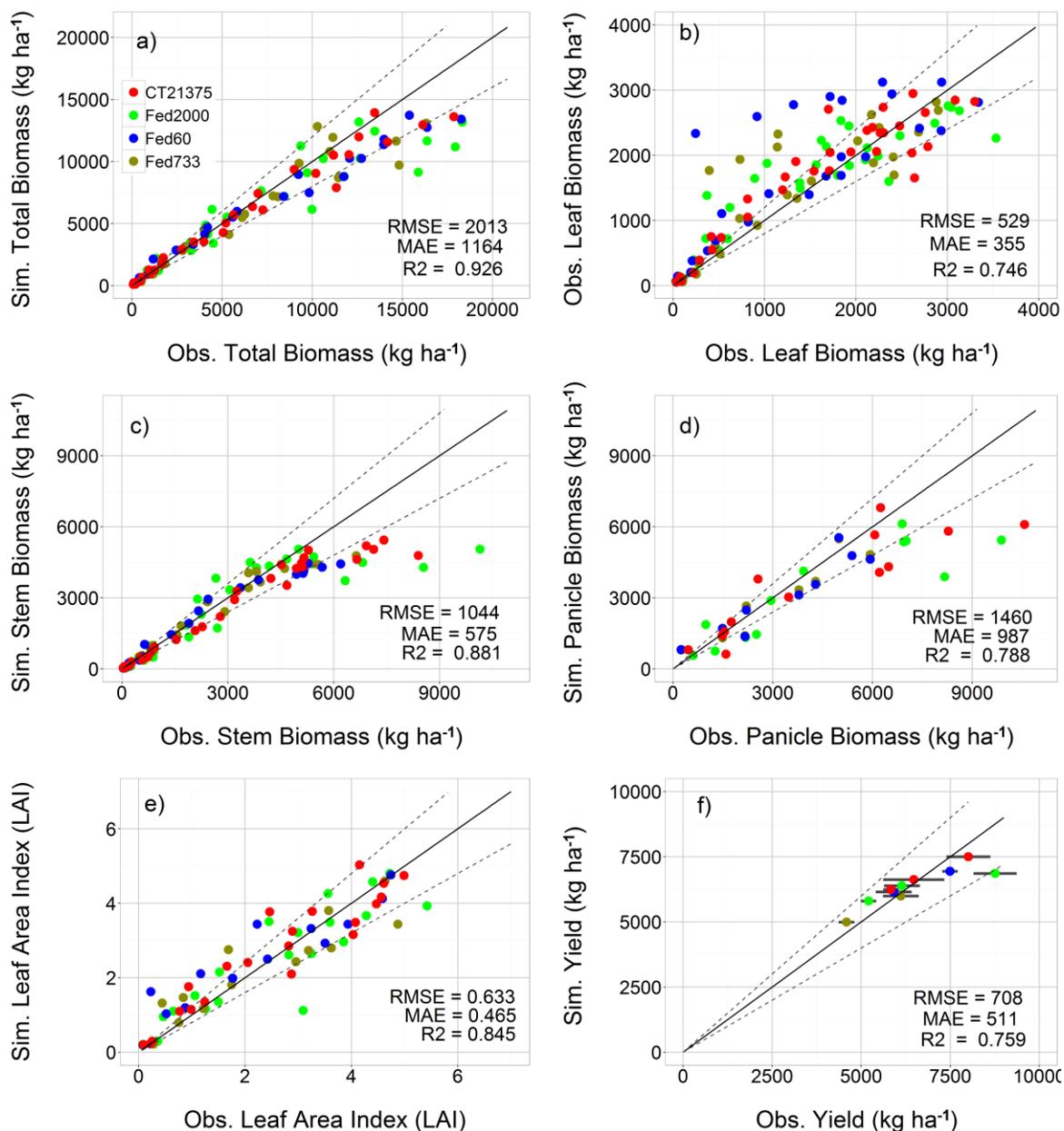


Figura. 29. Datos observados vs simulados de a) biomasa total; b) biomasa en hojas; c) biomasa en tallos; d) biomasa en panículas; e) Índice de Área Foliar (LAI); y f) rendimiento del cultivo, para el grupo de datos de calibración en Montería - Córdoba; para las variedades CT21375 (puntos rojos), Fedearroz 2000 (puntos verdes), Fedearroz 60 (puntos azules) y Fedearroz 733 (puntos cafés). La línea sólida es la recta de relación 1:1; Las líneas punteadas corresponden a la desviación estándar de los datos observados, alrededor de la recta 1:1. RMSE – Error medio cuadrado; MAE – Error medio absoluto; R<sup>2</sup> – Coeficiente de determinación.

Tabla 16. Parámetros cuantitativos de bondad de ajuste para evaluar el desempeño del modelo Oryza2000, considerando el grupo de datos de calibración.

Variable del cultivo	Num	X <sub>mean</sub>	X <sub>sd</sub>	$\alpha$	$\beta$	R <sup>2</sup>	Pt	RMSE <sub>a</sub>	RMSE <sub>n</sub>
<i>Fedearroz 2000</i>									
WAGT	65	7018	5783	0.80	1108	0.84	0.275	2352	33.5
WSO	26	4050	2703	0.61	1520	0.54	0.833	1826	45.1
WST	65	3520	2747	0.69	691	0.79	0.014	1346	38.2
WLVG	65	1513	1033	0.74	537	0.81	0.013	481	31.8
LAI	41	2.5	1.6	0.82	0.44	0.81	0.973	0.69	27.7
Yield	6	5131	2230	0.58	2297	0.85	0.440	991	19.0
<i>Fedearroz 733</i>									
WAGT	58	6718	5294	0.92	504	0.92	0.838	1483	22.1
WSO	22	3393	2269	0.80	1128	0.78	0.055	1135	33.4
WST	58	3582	2733	0.71	573	0.88	0.002	1136	31.7
WLVG	58	1516	998	0.80	500	0.77	0.003	516	34.1
LAI	33	2.5	1.71	0.70	0.58	0.77	0.260	0.84	33.3
Yield	5	5035	789	0.50	5494	0.45	0.277	1075	15.3
<i>Fedearroz 60</i>									
WAGT	49	6769	5550	0.91	434	0.92	0.390	1561	23.1
WSO	18	3376	2134	0.88	741	0.84	0.133	905	26.8
WST	49	3490	2774	0.65	648	0.82	0.004	1380	39.5
WLVG	49	1555	1042	0.87	512	0.71	0.001	661	42.5
LAI	27	2.6	1.8	0.71	0.68	0.85	0.548	0.76	28.6
Yield	4	5467	1569	0.36	4179	0.96	0.230	1105	20.0
<i>CT21375</i>									
WAGT	65	7562	6034	0.81	788	0.87	0.023	2270	30.0
WSO	27	4137	2866	0.73	1148	0.69	0.875	1570	37.9
WST	65	3831	2886	0.66	523	0.83	0.050	1491	38.9
WLVG	65	1704	1100	0.77	449	0.86	0.255	428	25.1
LAI	41	2.7	1.8	0.80	0.59	0.86	0.531	0.67	24.8
Yield	6	6138	1460	0.78	1637	0.87	0.350	582	9.5

WAGT - Materia seca total; WLVG - Hojas verdes; WSO - Panículas; WST - Tallos; Yield - Rendimiento; Num – Número de par de datos; X<sub>mean</sub> – Promedio de datos observados; X<sub>sd</sub> – Desviación estándar de datos observados;  $\alpha$  – Pendiente de la regresión lineal entre los datos simulados y datos observados;  $\beta$  - Intercepto de la regresión lineal entre los datos simulados y datos observados; R<sup>2</sup> – Coeficiente de determinación de la regresión; Pt - Prueba de medias t-Student

---

asumiendo varianza desigual (95 % de nivel de significancia);  $RMSE_a$  – Error medio cuadrático absoluto;  $RMSE_n$  – Error medio cuadrático normalizado.

### 3.1.4 Dinámica de Biomasa y Área foliar (Datos de Validación)

Con el objetivo de evaluar la capacidad del modelo ORYZA2000 para reproducir el crecimiento y desarrollo del cultivo de arroz, en condiciones agroambientales distintas a las establecidas durante el proceso de parametrización; en la etapa de validación se incluyeron ensayos experimentales realizados en la localidad de Aipe – Huila, además de nuevos experimentos establecidos en Saldaña y Montería, los cuales no fueron tenidos en cuenta para la identificación de los coeficientes genéticos de cada una de las variedades en estudio.

En Aipe se consideraron 16 experimentos para la validación del modelo, cuatro para cada una de las variedades. Además de la fenología, se evaluó el desempeño del modelo para simular las dinámicas de producción de materia seca en hojas, tallos, panículas y materia seca total, así como el rendimiento final de los cuatro cultivares (ver Tabla 6).

En la Figura. 30 se compara los datos observados y simulados de las variables de crecimiento y su rendimiento del cultivo, para los ensayos de validación en la localidad de Aipe. En general se obtuvo un muy buen desempeño del modelo para todas las variables simuladas, siendo notablemente mejor que los resultados obtenidos en el proceso de calibración. La biomasa total de cada una de las variedades fue muy bien simulada, la mayoría de los datos simulados se ubicaron dentro de los intervalos de desviación estándar establecidos a lo largo de la recta 1:1. Una prueba de ello, es el alto valor del coeficiente de determinación ( $R^2 = 0.95$ ), el cual refleja la buena habilidad del modelo para simular la baja dispersión en los datos observados (Figura. 30a).

Al igual que la biomasa total, la biomasa simulada de los tallos mostró un buen ajuste respecto a los datos observados ( $R^2 = 0.87$ ), la gran mayoría de los datos se ubicaron dentro de los rangos de desviación, sin embargo, se pudo observar una sub estimación de la biomasa de tallos simulada durante la floración del cultivo, lo cual se puede evidenciar en la Figura. 30c. Las variables que mostraron mayor dispersión en esta localidad fueron la materia seca simulada de hojas y panículas, en ambos casos se observó una sobreestimación respecto a lo observado. No obstante, las métricas estadísticas mostraron

una buena capacidad del modelo para simular el comportamiento, tanto de hojas (Figura. 30b) como panículas (Figura. 30c), bajo las condiciones agroambientales de Aipe.

En cuanto al rendimiento simulado, todos los datos se ubicaron dentro de los intervalos de desviación estándar, la mayoría de ellos por encima de la recta 1:1, indicando que a pesar de haber similitud entre los datos observados y simulados, el modelo sobre estimó en  $\pm 800 \text{ kg ha}^{-1}$  el rendimiento global de las variedades simuladas en esta localidad (Figura. 30e).

En Saldaña se utilizaron ocho experimentos para la validación del modelo, las variables de crecimiento y desarrollo del cultivo que se evaluaron en esta localidad fueron: La materia seca en hojas, tallos, panículas, materia seca total y LAI. De igual forma, se consideró el rendimiento final obtenido en cada uno de los ensayos (ver Tabla 6). En la Figura. 31 se comparan los datos observados y simulados de las variables de crecimiento del cultivo y se evalúa el nivel de ajuste respecto a la línea de referencia 1:1.

El modelo reprodujo bastante bien la dinámica de biomasa total de cada una de las variedades, la mayoría de los puntos de datos se ubicaron dentro de los intervalos de desviación estándar establecidos a lo largo de la recta 1:1, mostrando una baja dispersión ( $R^2= 0.91$ ), que confirma el buen ajuste logrado entre la información medida en campo y lo simulado por el modelo (Figura. 31a). No obstante, cabe resaltar la existencia de un grupo de datos por fuera de los intervalos de desviación, los cuales evidencian una subestimación de la materia seca total por parte del modelo de cultivo, especialmente para la variedad CT21375. Este comportamiento se debió a que el modelo subestimó notablemente la materia seca de tallos, especialmente en la etapa de llenado de granos de esta variedad, para el ensayo realizado el 29 de abril de 2013 (S1MADR).

Respecto a la materia seca producida en tallos, hojas verdes y panículas, gran cantidad de los datos simulados se ubicaron dentro de los intervalos de desviación, pero por debajo de la recta 1:1, lo que demuestra una subestimación generalizada por el modelo de cultivo para estas variables. Dichas discrepancias fueron reflejadas por los altos valores obtenidos de RMSE, en relación a los resultados encontrados para la localidad de Aipe.

Por su parte, el LAI (sólo para CT21375 y Fedearroz 2000), y el rendimiento de las diferentes variedades fueron simulados relativamente bien en esta localidad. En general, el LAI simulado se subestimó en una proporción muy baja (RMSE = 0.7) con respecto al

---

LAI observado (Figura. 31e). Mientras tanto, los rendimientos simulados para las cuatro variedades se ubicaron dentro de los intervalos de desviación establecidos. El bajo valor del coeficiente de determinación ( $R^2 = 0.65$ ), pudo ser un reflejo de la dispersión de los rendimientos medidos en campo. En promedio, el modelo estimó el rendimiento simulado con un error de  $900 \text{ kg ha}^{-1}$ , respecto a los datos observados (Figura. 31f).

En la localidad de Montería se consideraron sólo dos experimentos para llevar a cabo la evaluación del modelo, divididos entre las variedades CT21375 y Fedearroz 2000. Al igual que en Saldaña, se evaluó el desempeño de las simulaciones de la materia seca en hojas, tallos, panículas, materia seca total, LAI y rendimiento del cultivo (ver Tabla 6). A diferencia de las otras dos localidades, para Montería no se incluyó una gráfica de dispersión para evaluar el rendimiento simulado por el modelo, debido a que sólo se contó con dos datos de rendimiento.

En general, la simulación de la biomasa en los diferentes órganos y del área foliar específica, tuvo un menor desempeño en relación a los resultados obtenidos en la calibración, y a las validaciones realizadas en Aipe y Saldaña. Los mejores resultados se observaron en las simulaciones de la biomasa total ( $R^2 = 0.96$ ), sin embargo, se observó una subestimación generalizada respecto a los datos observados (Figura. 32a), la cual estuvo asociada a la subestimación de materia seca de tallos (Figura. 32c) y panículas (Figura. 32d) para las dos variedades. A pesar de ello, gran parte de los datos simulados estuvieron muy cerca de los intervalos de desviación considerados. El LAI simulado generalmente fue superior a lo observado, sobre todo durante la etapa de floración del cultivo ( $R^2 = 0.59$ ). Por su parte, el rendimiento simulado estuvo muy cerca a los rendimientos promedios observados, con RMSE de  $700 \text{ kg ha}^{-1}$ , siendo menor al error encontrado en las localidades de Aipe y Saldaña, con el grupo de datos de validación.

En la Tabla 17 se muestran los parámetros estadísticos que evalúan el desempeño de ORYZA2000, para simular la dinámica de las diferentes variables de crecimiento, y el rendimiento de cada variedad, considerando todos los experimentos de validación, realizados en Aipe, Saldaña y Montería.

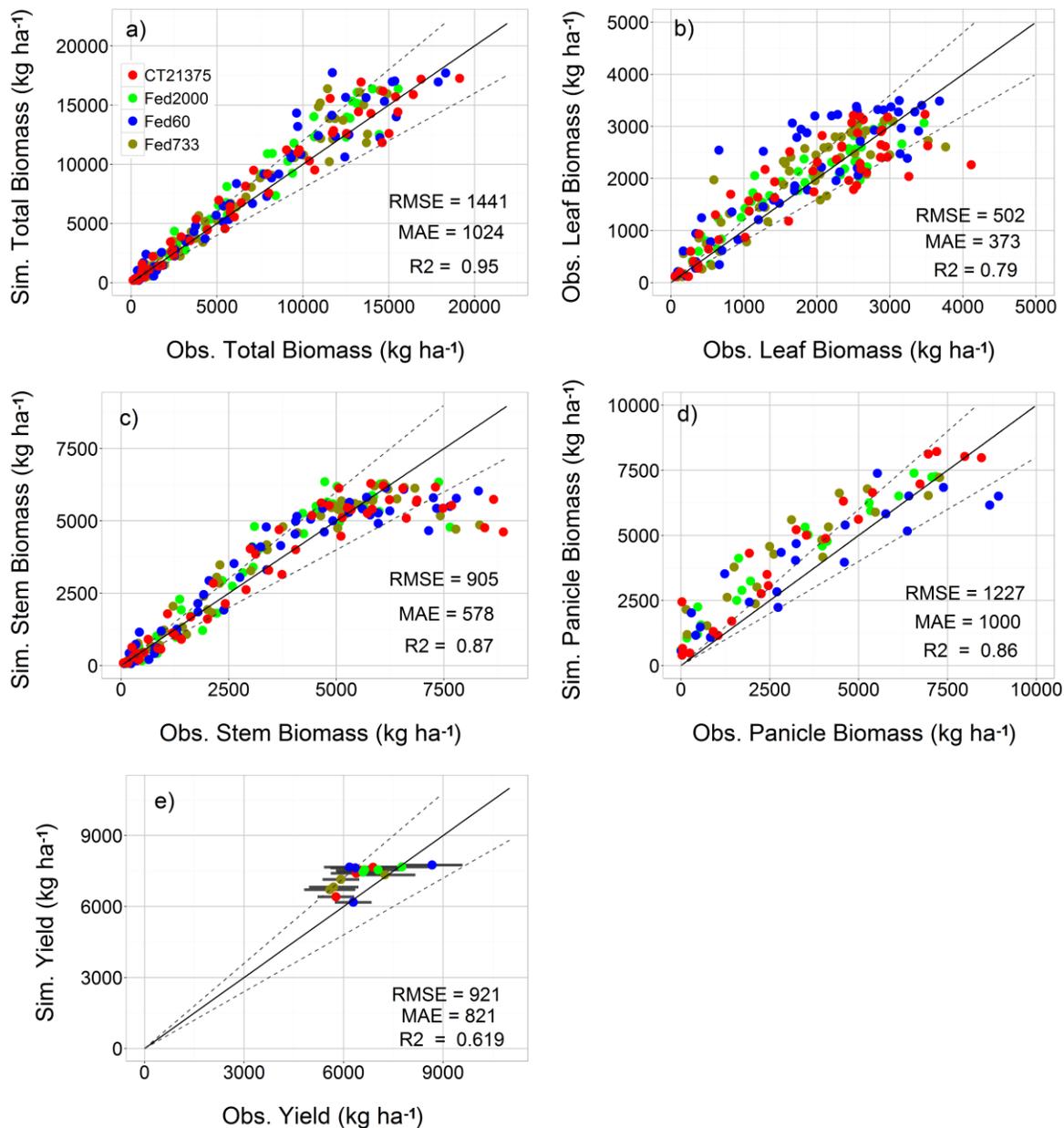


Figura. 30. Datos observados vs simulados de a) biomasa total; b) biomasa en hojas; c) biomasa en tallos; d) biomasa en panículas; y e) rendimiento del cultivo, para el grupo de datos de validación en Aipe - Huila; para las variedades CT21375 (puntos rojos), Fedearroz 2000 (puntos verdes), Fedearroz 60 (puntos azules) y Fedearroz 733 (puntos cafés). La línea sólida es la recta de relación 1:1; Las líneas punteadas corresponden a la desviación estándar de los datos observados, alrededor de la recta 1:1. RMSE – Error medio cuadrado; MAE – Error medio absoluto; R<sup>2</sup> – Coeficiente de determinación.

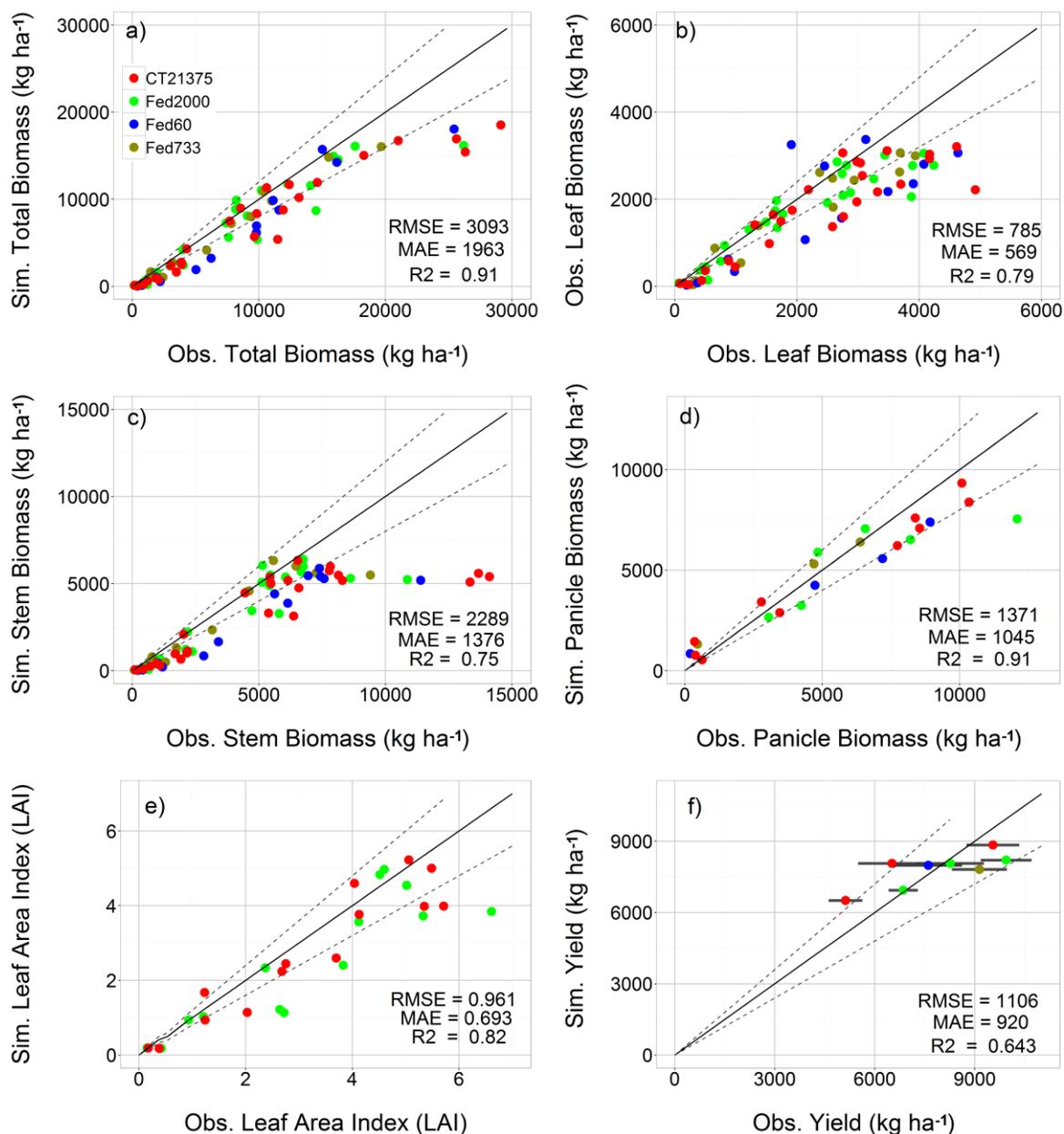


Figura. 31. Datos observados vs simulados de a) biomasa total; b) biomasa en hojas; c) biomasa en tallos; d) biomasa en panículas; e) LAI; y f) rendimiento del cultivo, para el grupo de datos de validación en Saldaña - Tolima; para las variedades CT21375 (puntos rojos), Fedearroz 2000 (puntos verdes), Fedearroz 60 (puntos azules) y Fedearroz 733 (puntos cafés). La línea sólida es la recta de relación 1:1; Las líneas punteadas corresponden a la desviación estándar de los datos observados, alrededor de la recta 1:1. RMSE – Error medio cuadrado; MAE – Error medio absoluto; R<sup>2</sup> – Coeficiente de determinación.

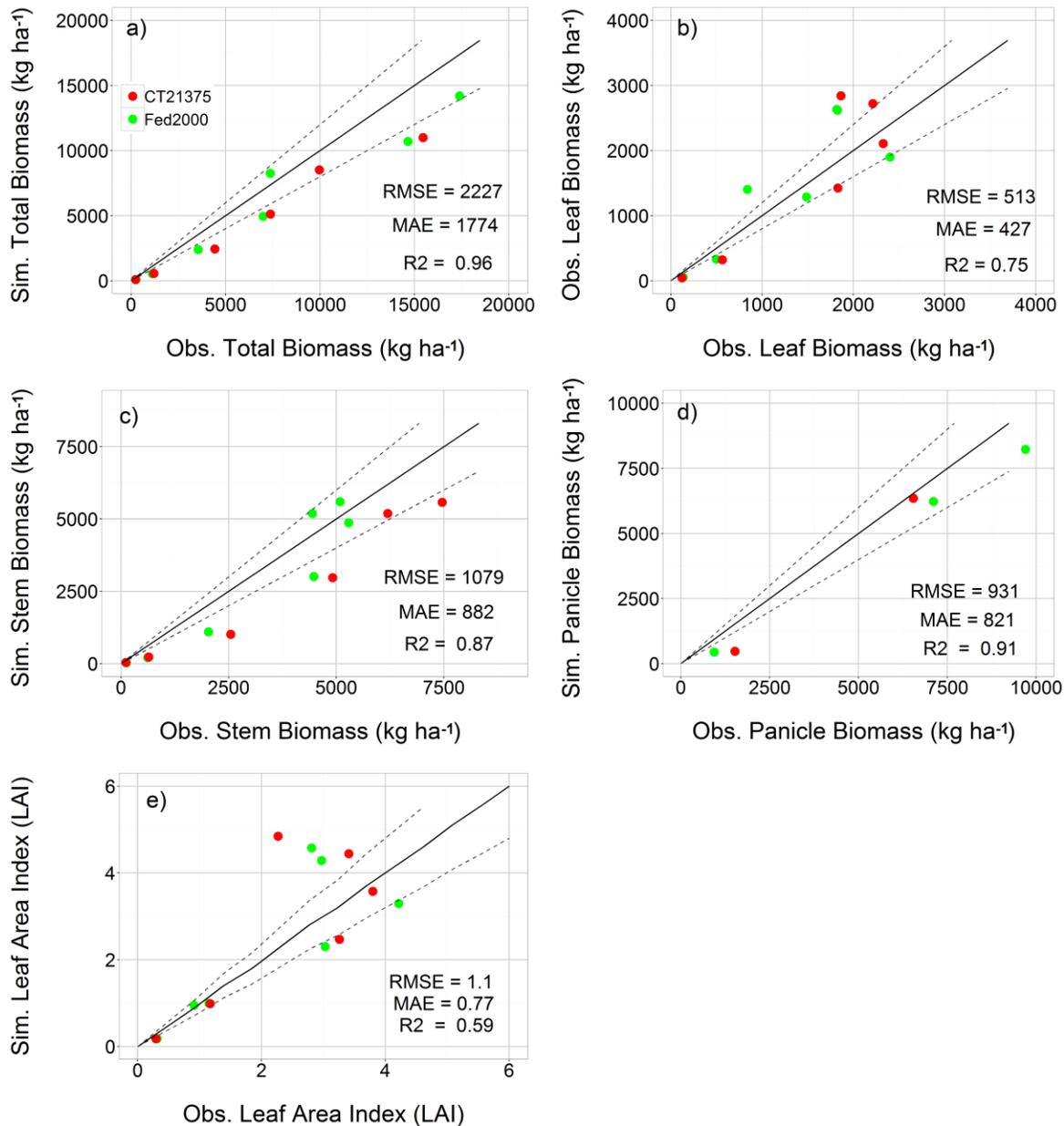


Figura. 32. Datos observados vs simulados de a) biomasa total; b) biomasa en hojas; c) biomasa en tallos; d) biomasa en panículas; y e) LAI, para el grupo de datos de validación en Montería - Córdoba; para las variedades CT21375 (puntos rojos) y Fedearroz 2000 (puntos verdes). La línea sólida es la recta de relación 1:1; Las líneas punteadas corresponden a la desviación estándar de los datos observados, alrededor de la recta 1:1. RMSE – Error medio cuadrado; MAE – Error medio absoluto; R<sup>2</sup> – Coeficiente de determinación.

Tabla 17. Parámetros cuantitativos de bondad de ajuste para evaluar el desempeño del modelo Oryza 2000, considerando el grupo de datos de validación.

Variable del cultivo	Num	$X_{\text{mean}}$	$X_{\text{sd}}$	$\alpha$	$\beta$	$R^2$	Pt	RMSE <sub>a</sub>	RMSE <sub>n</sub>
<i>Fedearroz 2000</i>									
WAGT	84	6741	5583	0.91	620	0.88	0.943	1938	28.8
WSO	29	4085	3153	0.62	1707	0.70	0.646	1736	42.5
WST	84	3369	2500	0.86	308	0.83	0.124	1056	31.3
WLVG	84	1730	1097	0.79	337	0.82	0.641	459	26.5
LAI	21	2.9	1.8	0.77	0.26	0.70	0.086	1.07	37.5
Yield	7	7582	1209	0.28	5503	0.66	0.460	838	11.9
<i>Fedearroz 733</i>									
WAGT	64	6560	5048	0.97	158	0.94	0.013	1457	22.2
WSO	23	3054	2266	0.87	1378	0.87	0.050	1279	41.9
WST	64	3449	2561	0.83	448	0.87	0.267	944	27.4
WLVG	64	1765	1071	0.83	357	0.80	0.298	487	27.6
LAI	----	----	----	----	----	----	----	----	----
Yield	5	6718	1509	0.28	5283	0.92	0.277	1075	15.3
<i>Fedearroz 60</i>									
WAGT	66	7280	5806	0.96	398	0.89	0.741	1934	26.6
WSO	25	3973	2938	0.68	1389	0.84	0.658	1260	31.7
WST	66	3670	2764	0.73	559	0.82	0.006	1279	34.9
WLVG	66	1908	1190	0.80	499	0.67	0.219	714	37.4
LAI	----	----	----	----	----	----	----	----	----
Yield	5	7018	1088	0.30	5305	0.21	0.247	982	14.0
<i>CT21375</i>									
WAGT	85	7749	6716	0.80	902	0.87	0.012	2531	32.7
WSO	34	4021	3245	0.81	1001	0.84	0.272	1314	32.7
WST	85	4026	3316	0.60	751	0.73	0.040	1987	49.4
WLVG	85	1947	1217	0.70	416	0.72	0.026	662	34.0
LAI	20	2.9	1.78	0.84	0.28	0.75	0.373	0.91	31.3
Yield	8	6701	1292	0.54	3839	0.79	0.089	1004	14.3

WAGT - Materia seca total; WLVG - Hojas verdes; WSO - Panículas; WST - Tallos; Yield - Rendimiento; Num - Número de par de datos;  $X_{\text{mean}}$  - Promedio de datos observados;  $X_{\text{sd}}$  - Desviación estándar de datos observados;  $\alpha$  - Pendiente de la regresión lineal entre los datos simulados y datos observados;  $\beta$  - Intercepto de la regresión lineal entre los datos simulados y datos observados;  $R^2$  - Coeficiente de determinación de la regresión; Pt - Prueba de medias t-Student asumiendo varianza desigual (95 % de nivel de significancia); RMSE<sub>a</sub> - Error medio cuadrático absoluto; RMSE<sub>n</sub> - Error medio cuadrático normalizado.

## 3.2 Identificación y Caracterización de ambientes

Con base en los rendimientos históricos simulados de las variedades Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21375; obtenidos en 38 estaciones meteorológicas (28 años de registro diario para cada estación) distribuidas dentro de la región de estudio, bajo seis diferentes tipos de suelos y 18 fechas de siembra (distribuidas entre los periodos 15 de Marzo – 3 Junio y 15 de Septiembre - 4 de Diciembre); se realizó el análisis numérico propuesto por Charrad et al (2014) para determinar el número óptimo de grupos ambientales y su mejor esquema de agrupamiento.

De acuerdo a los resultados obtenidos, se clasificaron los datos de rendimientos simulados en tres grupos, después de conocer que ocho de los 30 índices estadísticos evaluados sugirieron la existencia de tres Clústers contrastantes (Figura. 33a). Posteriormente, se implementó el análisis de clustering jerárquico, y se trazó una línea horizontal sobre el tercer nivel del dendrograma para especificar la agrupación final (Figura. 33b).

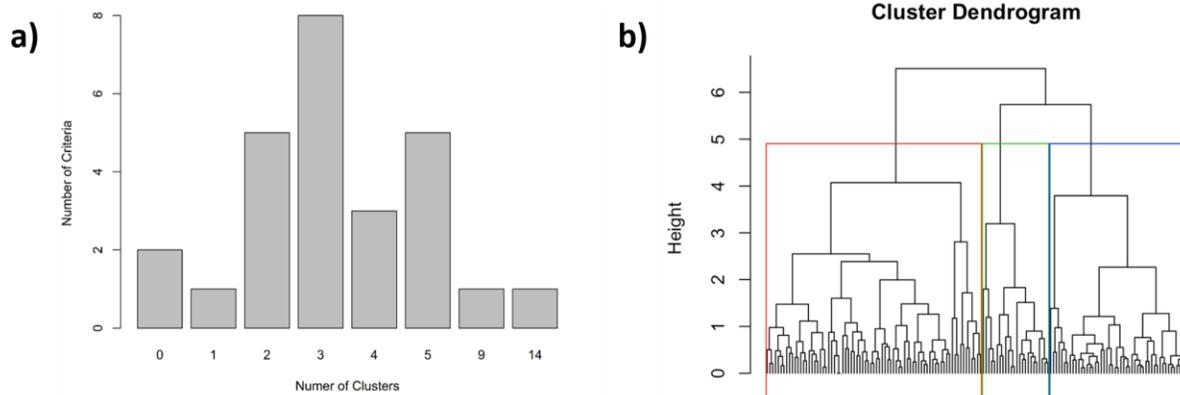


Figura. 33. Identificación de grupos ambientales a partir de los rendimientos simulados de las variedades Fedearroz 2000, Fedearroz 733, Fedearroz 60 y CT21375. a) Número de grupos sugeridos con base en el criterio de los 30 indicadores estadísticos (Charrad et al., 2014); b) Clústering jerárquico de los rendimientos simulados a partir del dendrograma.

De acuerdo a los rendimientos obtenidos en cada uno de los grupos, los ambientes de producción se clasificaron de la siguiente manera: Altamente Favorable (HFE), Favorable (FE) y Menos Favorable (LFE). El HFE fue el ambiente con los mayores rendimientos promedios ( $9,140 \text{ kg ha}^{-1}$ ) y representó el 18.5% del total de área de estudio. Por su parte, el FE tuvo rendimientos promedios de  $7,578 \text{ kg ha}^{-1}$ , ocupando el 50.2% de toda el área productora. Por último, el LFE se caracterizó por ser el ambiente con los menores rendimientos (en promedio  $6,000 \text{ kg ha}^{-1}$ ), y fue el segundo ambiente con la mayor probabilidad de ocurrencia (31.3%) dentro de la región de estudio (Figura. 34).

A nivel espacial, el HFE se distribuyó en dos pequeñas regiones ubicadas al centro-norte del departamento del Tolima y al extremo sur del Huila. Municipios como Ibagué (Tolima), Gigante, Tesalia, Agrado y Pital (Huila), se distinguieron por tener los mayores rendimientos simulados y por presentar una alta estabilidad ambiental (87% – 100%). Por otro lado, el FE ocupó una extensa región que comprendió los municipios de las zonas centro y norte de los departamentos de Tolima y Huila, dentro de los cuales ubican dos de los tres sitios experimentales (Saldaña y Aipe) seleccionados para la calibración y evaluación del modelo de cultivo. Esta región se distinguió por presentar una mediana estabilidad ambiental (54% - 86%) (Ver Figura. 35).

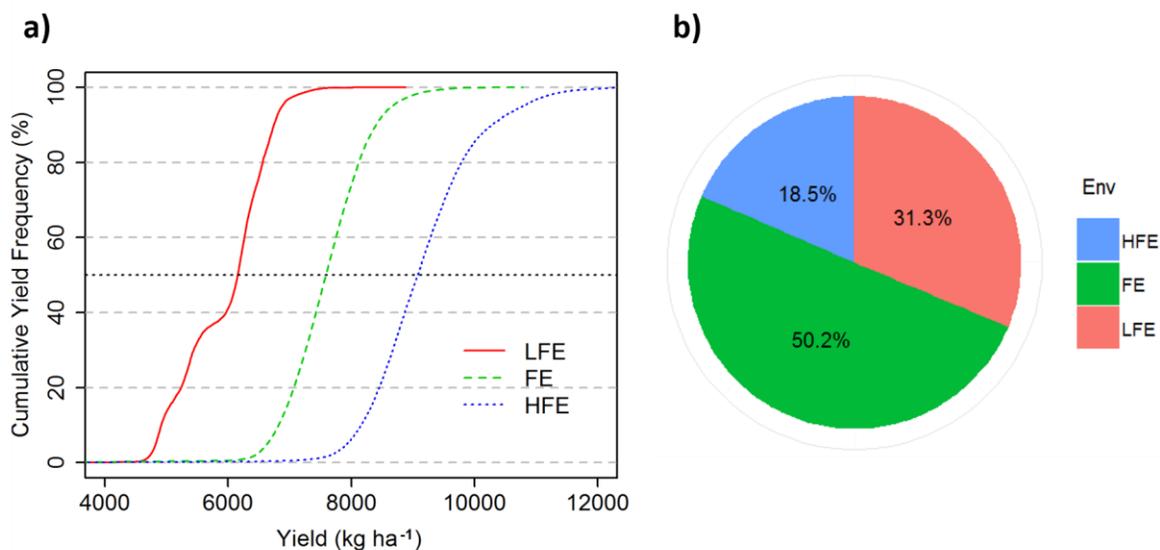


Figura. 34. Rendimientos simulados en la región de estudio. a) Frecuencia acumulada del rendimiento simulado en los diferentes grupos ambientales; b) Diagrama de tortas - Frecuencia de ocurrencia de los grupos ambientales.

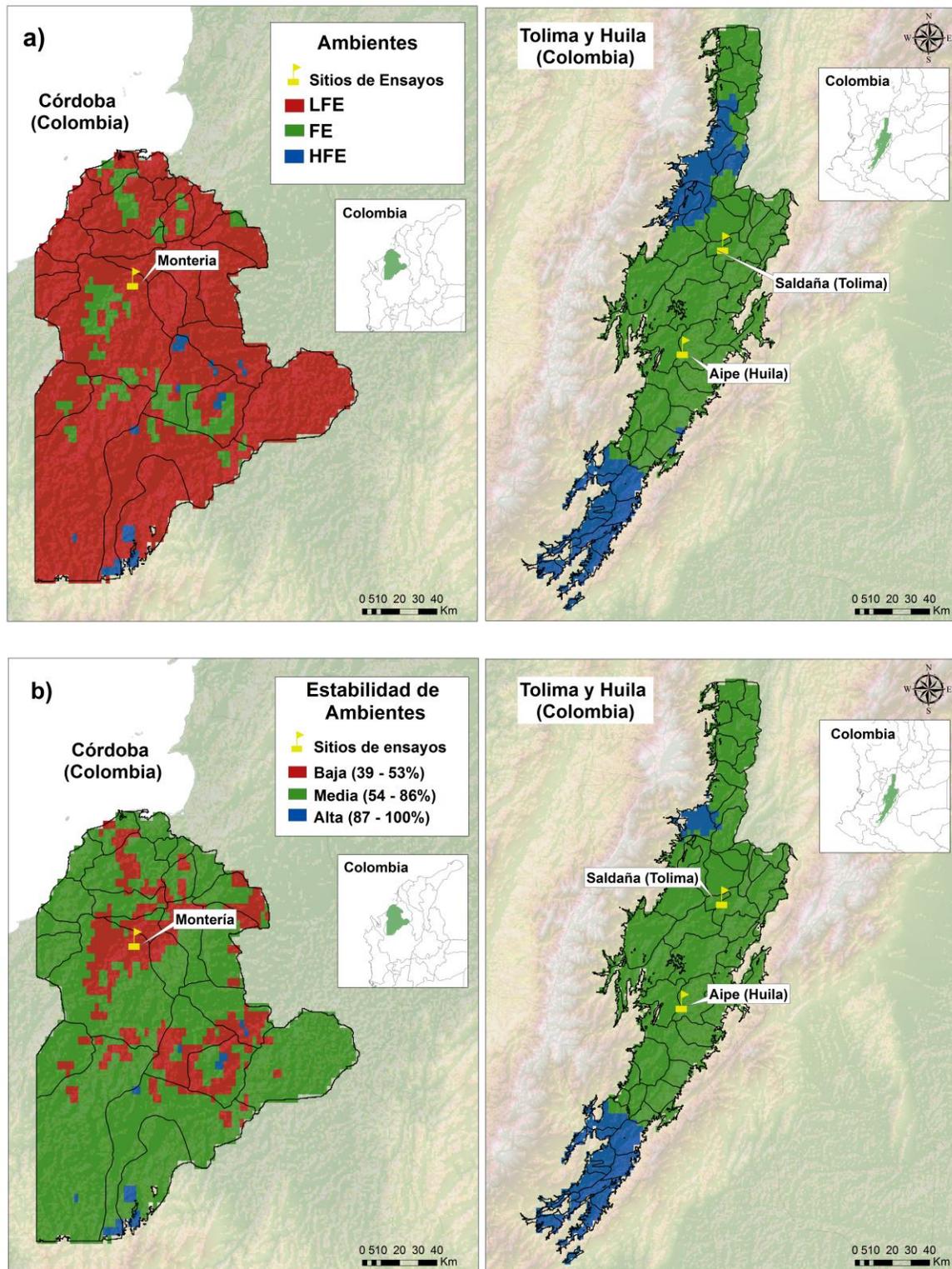


Figura. 35. Mapas de la distribución espacial de los grupos ambientales HFE, FE y LFE (a) y la estabilidad de cada uno de los ambientes (b). La estabilidad ambiental hace referencia a la frecuencia con la cual cada ambiente se presentó durante el periodo simulado (1984 - 2013).

Algunas áreas en el departamento de Córdoba también hicieron parte del ambiente FE, especialmente en los municipios de Montería, Planeta Rica, Loricá y San Bernardo del Viento; sin embargo, se caracterizaron por presentar una baja estabilidad ambiental (39% - 53%). Respecto al ambiente LFE, este se distribuyó totalmente en las regiones productoras del departamento de Córdoba, las cuales en su gran mayoría, presentaron una mediana estabilidad ambiental (54% - 86%) (Figura. 35).

Respecto a los rendimientos obtenidos por variedad, se observó una notable diferencia entre los rendimientos simulados de cada uno de los cultivares en los diferentes ambientes. En el HFE todos los cultivares expresaron los mayores rendimientos con relación a los otros dos ambientes, el cultivar CT21375 se distinguió por alcanzar los mejores rendimientos, la mayor frecuencia de ocurrencia (75%) y la mayor variabilidad (5,000 – 12,000 kg ha<sup>-1</sup>); mientras que las otras tres variedades no superaron el 20% en probabilidad de ocurrencia (Figura. 36). Los altos rendimientos simulados para CT21375 son una respuesta a los coeficientes genéticos establecidos durante el proceso de calibración, dado a que esta fue la variedad que tuvo una notable superioridad en la tasa de distribución de asimilados hacia las panículas (80% del total de asimilados en la planta después de la floración), además de tener una alta tasa máxima de crecimiento relativo en hojas (0.00799 °C día<sup>-1</sup>), y el mayor índice de área foliar específico durante toda la etapa del cultivo.

En el FE, el rendimiento de los cultivares estuvo centrado en 7,500 kg ha<sup>-1</sup>, variando entre 5,000 y 10,000 kg ha<sup>-1</sup>. En este ambiente, la variedad Fedearroz 2000 fue la que tuvo mayor frecuencia de ocurrencia (80%). Después de CT21375, el cultivar Fedearroz 2000 fue el que tuvo los mayores coeficientes genéticos durante la calibración del modelo. Se destaca que, del total de asimilados que acumula esta variedad después de floración, 70% son traslocados a las panículas y el 30% restante lo distribuye en hojas, siendo este último aspecto la gran diferencia con CT21375, la cual distribuye menos asimilados hacia las hojas después de floración, para concentrar una mayor cantidad en las panículas. Por su parte, los cultivares Fedearroz 60 y Fedearroz 733, tienen una frecuencia de ocurrencia en el FE, de 65% y 60%, respectivamente.

En el LFE fue evidente los bajos rendimientos alcanzados por las variedades, en relación a los otros dos ambientes (4,500 kg ha<sup>-1</sup> – 7000 kg ha<sup>-1</sup>). Los cultivares Fedearroz 60 y Fedearroz 733 fueron los que tuvieron mayor participación, con una frecuencia de

ocurrencia entre 20% y 30%. Por su parte, Fedearroz 2000 tuvo una baja participación, que no superó el 10%; mientras que CT21375 no clasificó dentro de este grupo ambiental (Figura. 36).

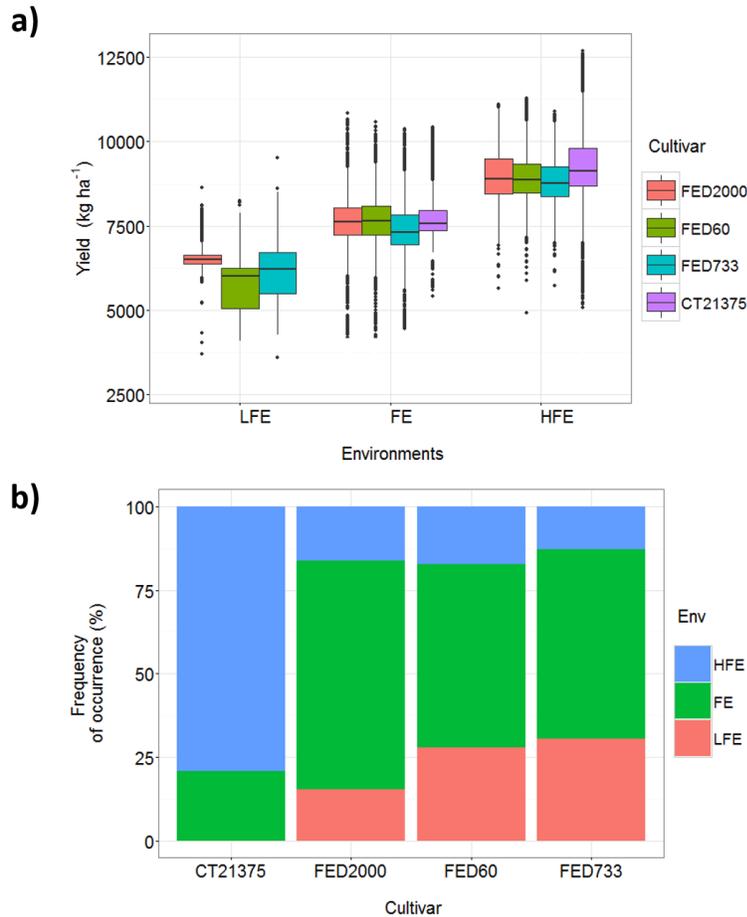


Figura. 36. Rendimiento de las variedades (a) y su frecuencia de ocurrencia a través de los diferentes grupos ambientales (b). Para los diagramas de caja, las líneas horizontales a través de las cajas representan la mediana, la altura representa el rango intercuartílico (desde el percentil 25 hasta el percentil 75) y los bigotes, las líneas que se extienden desde las cajas, se extienden hasta los valores mínimo (5%) y máximo (95%) de los datos.

Dado a que la gama de rendimientos simulados se mantuvo bastante amplia, en algunos casos, varios cultivares se comportaron de la misma forma en los diferentes ambientes (ej. Fedearroz 2000 y Fedearroz 60), generando cierta dificultad al momento de clasificarlos

en un determinado grupo. Por tanto, es importante considerar otros factores agroambientales (Temperatura, radiación solar, tasa de evapotranspiración, etc.) al momento de realizar el agrupamiento, lo cual permita explorar la posibilidad de encontrar un mayor número de condiciones ambientales con características bien diferenciadas entre ellas. En términos generales, los grupos ambientales identificados se pueden caracterizar de la siguiente manera:

***Ambiente Altamente Favorable (HFE):***

Este ambiente siempre mostró una consistencia en tener los más altos rendimientos a lo largo de todo el periodo histórico simulado (1984 - 2013), presentó un rendimiento de arroz promedio de 9,140 kg ha<sup>-1</sup>. Durante la primera época de siembra en el año (15 de Marzo – 3 de Junio), los mejores rendimientos se alcanzaron para las siembras realizadas entre el 14 de Mayo y el 3 de Junio, mientras que durante el segundo semestre (15 de Septiembre y 4 de Diciembre), el cultivo tuvo su mejor desarrollo bajo las siembras realizadas entre el 14 de Noviembre y 4 Diciembre. Durante estas épocas, se tuvo un ciclo promedio de 118 días, distribuidos de la siguiente manera: 55 días entre Emergencia - Iniciación de Panícula, 28 días entre Iniciación de Panícula – Floración, y 35 días entre Floración - Madurez Fisiológica.

Durante la etapa de desarrollo del cultivo, la temperatura mínima promedio estuvo relativamente estable, variando solo entre 21°C y 21.5 °C; mientras que la temperatura máxima promedio siempre aumentó a medida transcurría la etapa del cultivo, pasando de 31.5 °C, entre emergencia e iniciación panicular, hasta 33.1° C entre Floración y madurez fisiológica. De igual manera, la radiación solar promedio, mostró un incremento paulatino durante el ciclo de cultivo, pasando de 17.0 MJ m<sup>-2</sup> día<sup>-1</sup> al inicio del ciclo, hasta alcanzar 19.5 MJ m<sup>-2</sup> día<sup>-1</sup> en la madurez fisiológica. Los mayores cambios de la oscilación de temperatura en el ambiente, se presentaron entre Floración y Madurez Fisiológica, llegando a ser de hasta 12°C.

Por otro lado, en el HFE se evidenció un alto porcentaje en la probabilidad de ocurrencia (90%) de tener suelos de texturas pesadas, principalmente suelos Arcillosos y Arcillo Arenosos (Figura. 37). Finalmente, este ambiente se distinguió por tener una alta estabilidad (87% - 100%) a través del tiempo.

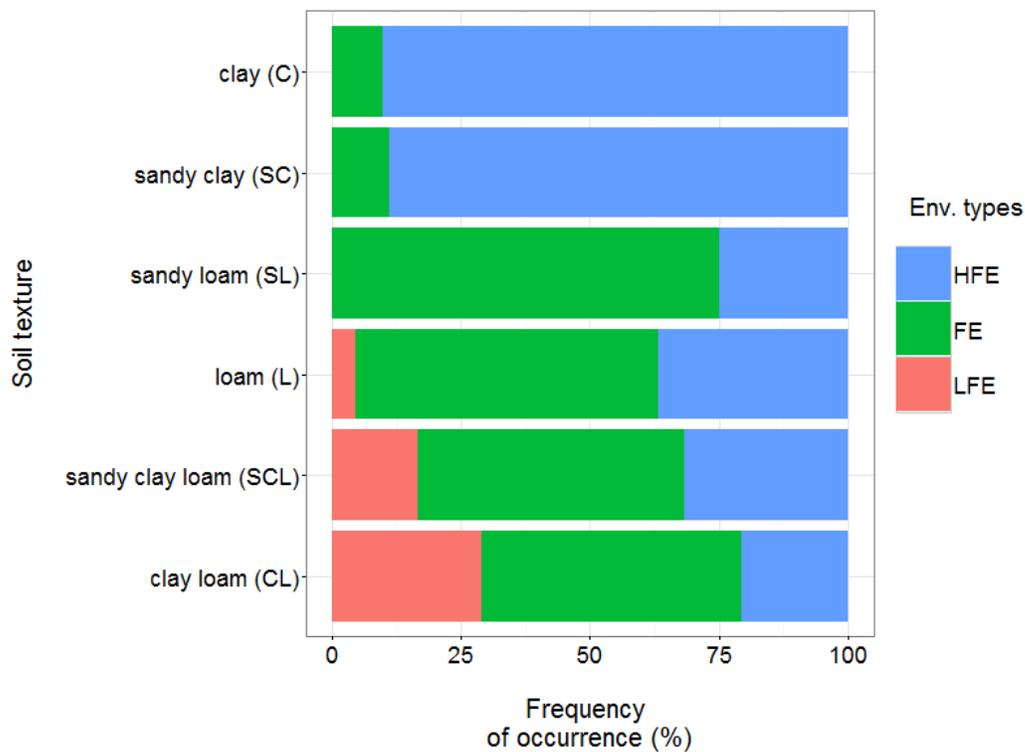


Figura. 37. Frecuencia de ocurrencia de los diferentes tipos de suelos, en los diferentes grupos ambientales. Clay – Arcilloso; Sandy Clay – Arcillo Arenoso; Sandy Loam – Franco Arenoso; Loam – Franco; Sandy Clay Loam – Franco Arcilloso Arenoso; Clay Loam – Franco Arcilloso.

**Ambiente Favorable (FE):**

El ambiente FE se caracterizó por tener rendimientos promedios de 7,500 kg ha<sup>-1</sup>, especialmente tuvo mayor presencia en las regiones Centro y Norte de los departamentos de Tolima y Huila; teniendo una mayor frecuencia de ocurrencia de los rendimientos simulados para las variedades Fedearroz 2000, Fedearroz 60 y Fedearroz 733. Las épocas de siembra en las cuales se obtuvieron los mayores rendimientos, coincidieron con la presencia de las temporadas típicas de lluvia en la región (Mayo - Junio) y (Noviembre – Diciembre). Al igual que en el HFE, los mejores rendimientos se alcanzaron en las siembras realizadas entre los periodos 14 de Mayo - 3 de Junio y 14 de Noviembre – 4 de Diciembre. Por su parte, las épocas en las cuales el cultivo tuvo un menor desempeño, fueron las comprendidas entre el 15 de Marzo – 4 de Abril y el 15 de Septiembre – 5 de Octubre. En promedio, el total del ciclo de cultivo fue más corto que lo registrado en el

HFE; se encontró una duración promedio de 114 días, de los cuales, 54 estuvieron entre Emergencia – Iniciación de Panícula, 28 días entre Iniciación de Panícula – Floración, y 32 días entre Floración – Madurez fisiológica.

Respecto a las características agroclimáticas, se encontró que las temperaturas mínima y máxima, a través del ciclo de cultivo, siguieron el mismo patrón encontrado en el HFE; no obstante, los valores promedios fueron mayores. En el caso de la temperatura en horas de la noche, esta osciló entre 22.5°C – 23.5°C; mientras que la temperatura diurna se incrementó a lo largo del ciclo de cultivo, pasando de 32°C durante las primeras etapas, hasta alcanzar los 33.5 °C al momento de la madurez fisiológica. Los valores de radiación solar promedio encontrados a lo largo del ciclo, fueron más bajos que los registrados en el HFE, en general estuvieron entre 16.0 MJ m<sup>-2</sup> día<sup>-1</sup> al inicio del ciclo, incrementándose progresivamente hasta alcanzar ±18.5 MJ m<sup>-2</sup> día<sup>-1</sup> durante la etapa de llenado de grano.

Los suelos más predominantes en este grupo ambiental se caracterizaron por ser de textura media, hubo una mayor proporción de suelos Franco Arenosos (75%), seguidos por suelos de textura Franca (60%), y en menor proporción predominaron suelos Franco Arcillo Arenosos y Franco Arcillosos. La estabilidad de este ambiente osciló entre 54% - 86%.

***Ambiente Menos Favorable (LFE):***

El LFE fue el ambiente que presentó los rendimientos más bajos durante todo el ciclo histórico simulado (Figura. 38). Los rendimientos de este ambiente oscilaron entre 4,500 kg ha<sup>-1</sup> y 6,500 kg ha<sup>-1</sup>, con valores promedios de 5,500 kg ha<sup>-1</sup>. Especialmente, el LFE se distribuyó en el departamento de Córdoba, y en su gran mayoría presentó una mediana estabilidad ambiental (54% - 86%). Los rendimientos alcanzados por las variedades Fedearroz 733 y Fedearroz 60 fueron los que tuvieron mayor presencia.

El crecimiento y desarrollo de las variedades se caracterizó por tener ciclos de crecimiento de 112 días en promedio, siendo notablemente inferior a lo encontrado en los otros dos ambientes. La etapa comprendida entre Emergencia – Iniciación de panícula, tuvo una duración promedio de 52 días, mientras tanto, el periodo entre Iniciación de Panícula – Floración duró alrededor de 27 días, y por último, la etapa entre Floración – Madurez Fisiológica comprendió de 33 días.

A diferencia de los ambientes con mejores rendimientos, en el LFE el cultivo tuvo el mejor desempeño en las siembras realizadas entre el 15 de Marzo – 15 de Abril, y 24 de Noviembre – 4 de Diciembre. Durante estas épocas, la temperatura mínima del ciclo de cultivo osciló entre 23.0°C y 23.5°C, siendo así el ambiente con las temperaturas nocturnas más altas. Por su parte, a temperatura máxima se mantuvo estable durante el ciclo de cultivo (32.5 °C), siendo inferior a lo reportado en el FE. A consecuencia de este comportamiento, fue característico de este ambiente tener bajas tasas de oscilación de temperatura durante el ciclo de cultivo, las cuales no superaron los 10°C día<sup>-1</sup>. Por otro lado, la radiación solar promedio presentó la mayor tasa de variación, oscilando entre 18 MJ m<sup>-2</sup> día<sup>-1</sup> - 21 MJ m<sup>-2</sup> día<sup>-1</sup>.

Respecto a las condiciones edafológicas, no se observó un dominio marcado de algún tipo de suelo, las texturas Franco Arcillosa y Franco Arcillo Arenosa, fueron las que tuvieron mayor participación.

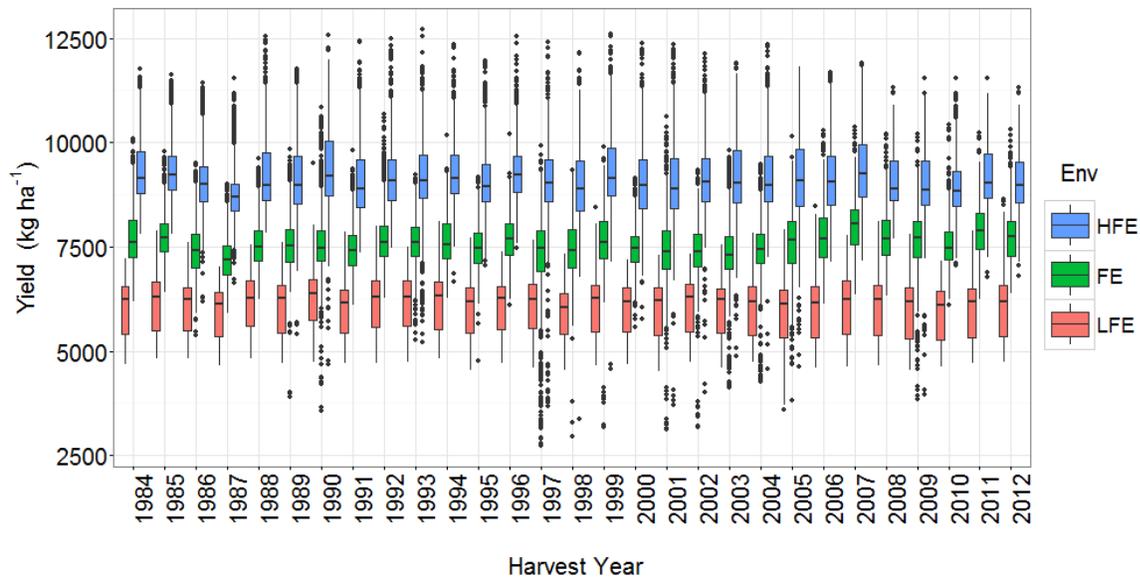


Figura. 38. Rendimientos históricos simulados (1984 - 2012) a través de los diferentes grupos ambientales. Para los diagramas de caja, las líneas horizontales a través de las cajas representan la mediana, la altura representa el rango intercuartílico (desde el percentil 25 hasta el percentil 75) y los bigotes, las líneas que se extienden desde las cajas, se extienden hasta los valores mínimo (5%) y máximo (95%) de los datos.

### **3.3 Determinación del mayor estrés abiótico y su influencia sobre el rendimiento del cultivo**

#### **Desempeño de los modelos de Red Neuronal Artificial**

La implementación de modelos de regresión a partir de redes neuronales artificiales se llevó a cabo con el fin de identificar la relación existente entre el rendimiento del cultivo de arroz en cada uno de los grupos ambientales, y factores climáticos tales como temperaturas (mínima y máxima), precipitación y radiación solar incidente, en distintas fases del ciclo vegetativo del cultivo.

Para seleccionar la mejor arquitectura de cada modelo, se evaluó el Error medio cuadrático normalizado (RMSEn) y el coeficiente de determinación  $R^2$  en 30 diferentes arquitecturas, seleccionando las que registraron menor RMSEn, mayor  $R^2$  y menor número de neuronas en la capa oculta de los modelos de cada ambiente. Cada arquitectura se entrenó 100 veces, en la capa de entrada se utilizaron 16 variables predictoras Tabla 7, y el rendimiento como variable de salida. Debido a que los modelos de redes neuronales no se implementaron con el objetivo de realizar predicciones o pronósticos del rendimiento, sino encontrar la importancia relativa de las variables predictoras respecto a la variable de respuesta, durante la fase de entrenamiento de cada modelo se utilizó la totalidad de rendimientos simulados obtenidos del modelo de cultivo (LFE  $\approx$  23,490 datos; FE  $\approx$  51,156 datos; HFE  $\approx$  9,396 datos), sin considerar nuevos grupos de datos para la validación de los modelos obtenidos.

En la Figura. 39 se presenta el comportamiento promedio del RMSEn para cada arquitectura, el número de neuronas seleccionadas en la capa oculta, y el diagrama de red neuronal obtenido para los ambientes LFE, FE y HFE. De acuerdo a los resultados obtenidos, en los tres set de datos hay un comportamiento decreciente asintótico del RMSEn a medida que aumenta el número de neuronas en la capa oculta. Es decir, a mayor número de neuronas hay una mejor capacidad predictiva del modelo, no obstante, después de cierto número de neuronas el error se vuelve prácticamente constante.

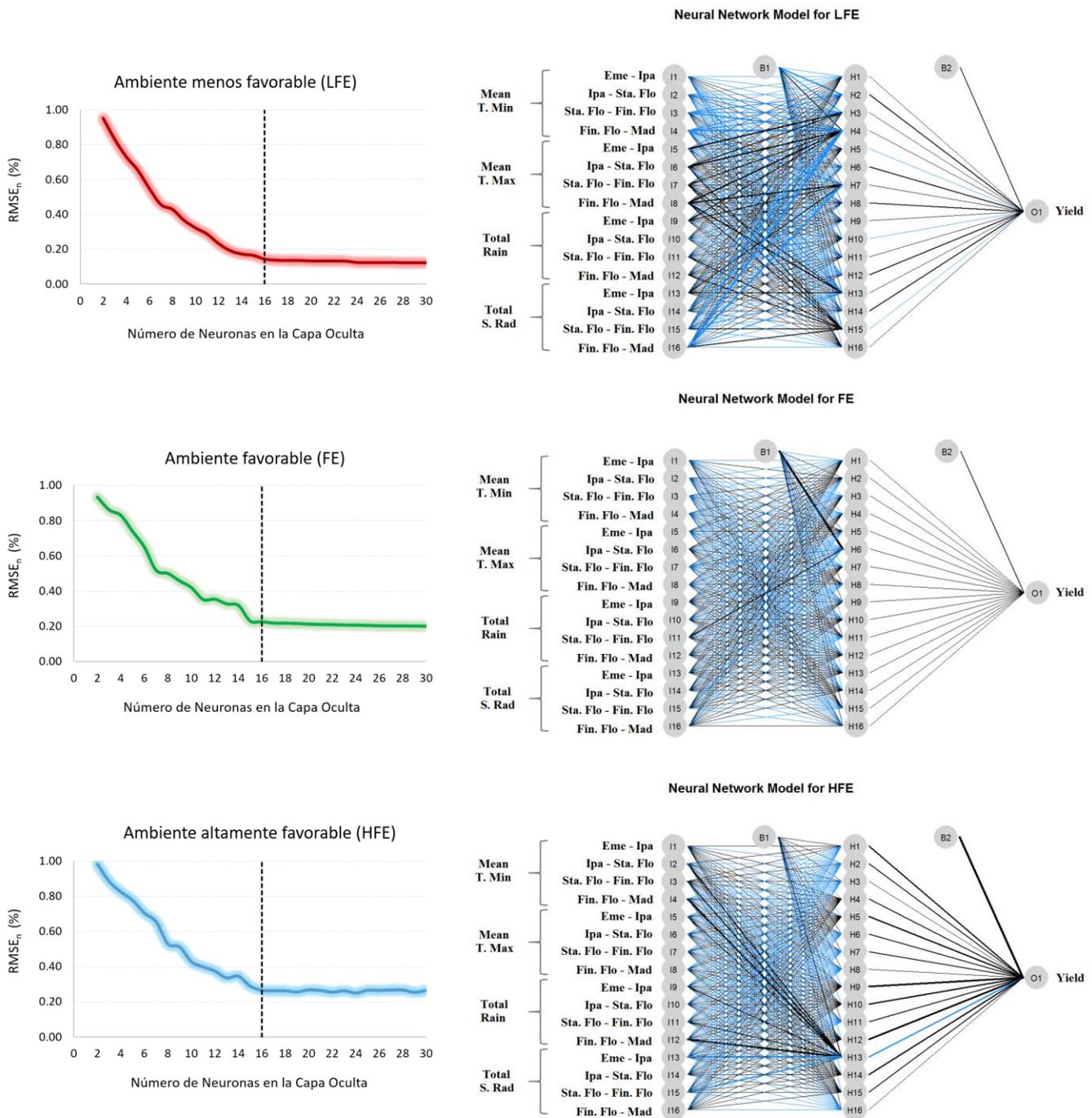


Figura. 39. Arquitecturas de Red Neuronal Artificial. Al lado izquierdo se observa la variación promedio del RMSE<sub>n</sub> para las arquitecturas evaluadas en el LFE (Línea roja), FE (Línea verde) y el HFE (Línea azul). La línea vertical (Color negro) representa el número de neuronas en la capa oculta para el cual se obtiene el menor RMSE<sub>n</sub>. Al lado derecho se muestran las estructuras de las redes neuronales identificadas en este estudio. Cada estructura de red neuronal está compuesta por una capa de entrada, la cual contiene tantas neuronas como variables que entran al sistema; una capa

---

oculta con un número de neuronas determinadas empíricamente; y una capa de salida con una neurona, correspondiente a la única variable de respuesta.

Para los tres modelos establecidos, el número de óptimo de neuronas en la capa oculta fueron 16, generando tres modelos con la siguiente arquitectura: Una capa de entrada con un total de 16 neuronas, correspondientes a las 16 variables de entrada al modelo, una capa oculta con 16 neuronas, y la capa de salida con una sola neurona, correspondiente a la variable de respuesta del modelo [16 16 1].

Con la anterior arquitectura, para el LFE se obtuvo un modelo con un error de entrenamiento del 14%, mientras que para el FE el error fue del 23%, y para el HFE se obtuvo un error promedio del 27%. Respecto a los resultados obtenidos de los coeficientes de determinación, la comparación entre los rendimientos observados y simulados de los diferentes modelos, mostraron para el LFE un  $R^2 \approx 0.85$  ( $P < 0.01$ ), para FE un  $R^2 \approx 0.76$  ( $P < 0.01$ ) y para HFE un  $R^2 \approx 0.73$  ( $P < 0.01$ ). En general se observó que la capacidad de predicción y ajuste de los modelo de redes neuronales artificiales son buenos, con errores cuadráticos medios estandarizados inferiores al 30%, y con coeficientes de determinación superiores en todos los casos a 0.70.

Una vez establecidas las estructuras de redes neuronales para cada ambiente, se utilizaron Diagramas de Interpretación Neuronal, con el fin de realizar un primer análisis cualitativo de la importancia de las variables explicativas, dada su influencia relativa sobre la variable de respuesta. Esta técnica, propuesta por Özesmi & Özesmi (1999), permite realizar una interpretación visual de los pesos de las conexiones entre las neuronas, la magnitud relativa de cada peso de la conexión está representada por el espesor de la línea (es decir, líneas más gruesas representan mayores pesos), y el color de las líneas de conexión representan la dirección o el sentido de cada uno de los pesos (por ejemplo, líneas de color negro representan señales con pesos positivos, mientras que líneas azules representan señales negativas). El seguimiento de la magnitud y la dirección de los pesos entre las neuronas, permite identificar los efectos individuales y la interacción de las variables de entrada sobre la variable de salida (Olden & Jackson, 2002).

En la Figura. 40 se muestra el NID obtenido para la estructura de red neuronal del LFE. Se puede observar en mayor detalle la distribución de los pesos de las variables de entrada, y el sentido (positivo o negativo) de cada uno de ellos, a través de las diferentes capas de la red.

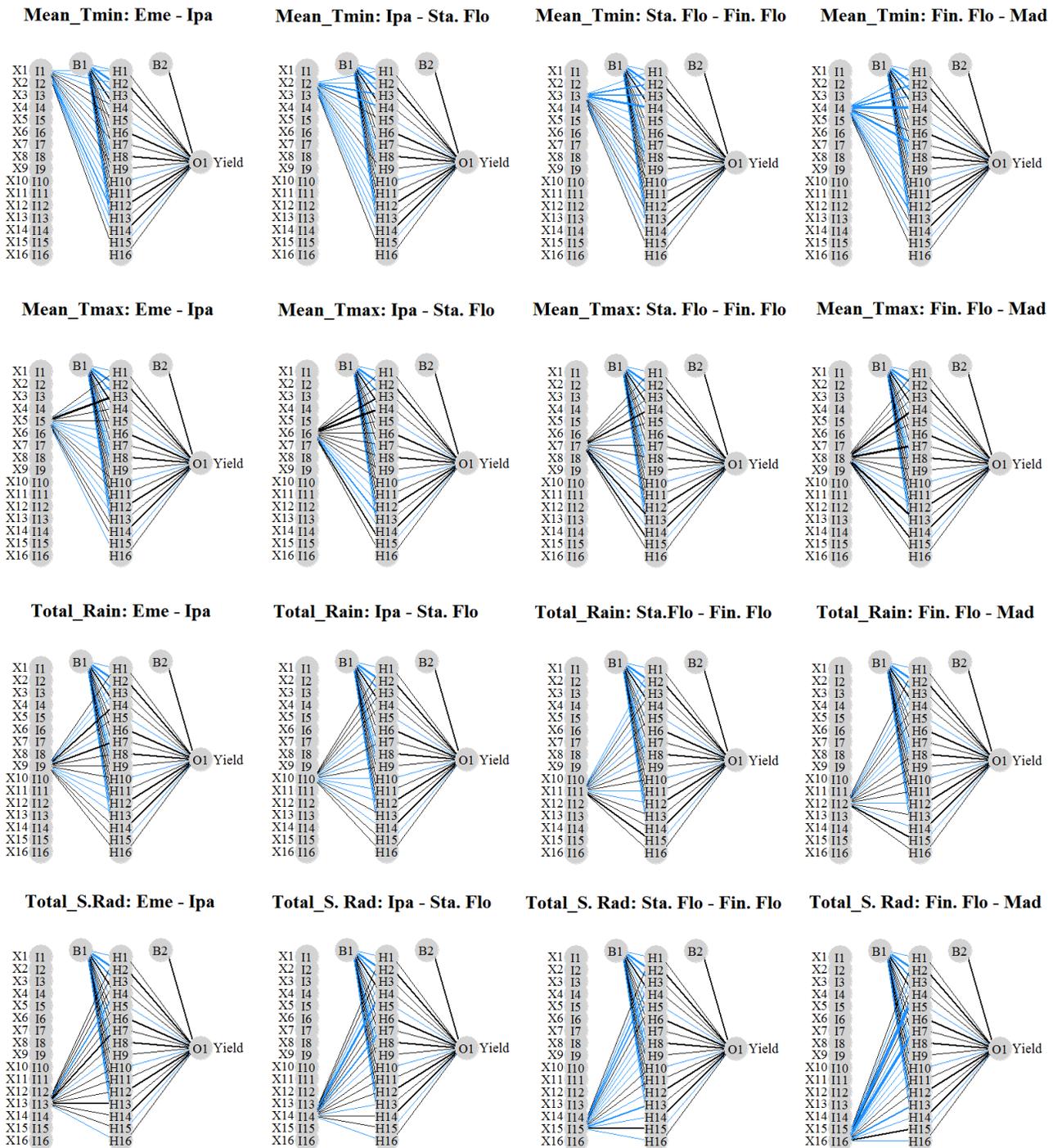


Figura. 40. Diagramas de Interpretación Neuronal para la descomposición del modelo RNA obtenido en el grupo ambiental LFE. El grosor de las líneas que unen neuronas es proporcional a la magnitud del peso de la conexión, y el color de cada línea indica la dirección de la interacción entre las neuronas: Las conexiones negras representan pesos positivos y las conexiones azules indican pesos negativos.

---

Las líneas de color azul que conectan los nodos de la capa de entrada y la capa oculta, muestran un efecto negativo sobre el rendimiento. Mientras que las líneas de color negro, representan las variables que tienen un efecto positivo sobre la variable de respuesta en la red neuronal.

De acuerdo a la Figura. 40, los cuatro primeros nodos (I1 a I4) en la capa de entrada, correspondientes a la temperatura mínima media registrada en diferentes etapas del cultivo (I1 - Emergencia a Iniciación de panícula, I2 – Iniciación de panícula a inicio de la floración, I3 - Inicio de floración a fin de floración, I4 – Fin de floración a madurez fisiológica), tienen una alta frecuencia de conexiones con pesos negativos (líneas azules) que se propagan hasta la capa oculta; siendo más fuerte el nivel de asociación entre las variables de los nodos I3 e I4 (Tem. mínima entre Inicio de floración a fin de floración y Tem. mínima entre Fin de floración a madurez fisiológica) con la variable de respuesta, ya que presentan líneas azules de mayor grosor.

Un comportamiento similar se observa con la radiación solar acumulada para las mismas etapas del cultivo (I15 e I16), en las cuales se observa una alta frecuencia de interconexiones con pesos negativos entre la capa de entrada y la capa oculta. Sin embargo, solo en las conexiones que se despliegan desde el nodo I16 se aprecia una mayor intensidad en la relación de la variable predictora, correspondiente a I16, y la variable de respuesta.

Respecto a las neuronas asociadas a la temperatura máxima media (I5 – I8) y a la precipitación acumulada (I9 – I12), en las diferentes etapas del cultivo, se observa una mayor frecuencia de interconexiones con pesos positivos entre las diferentes capas de la red, y muy pocas conexiones que evidencien una fuerte relación entre alguna variable predictora y la variable de respuesta. Este es el caso de la temperatura máxima media entre fin de floración y madurez fisiológica (I8), que muestra interconexiones positivas de mayor intensidad.

En los anexos A y B Se muestran los Diagramas de Interpretación Neuronal obtenidos para los grupos ambientales FE y HFE.

## Importancia relativa de las variables predictoras

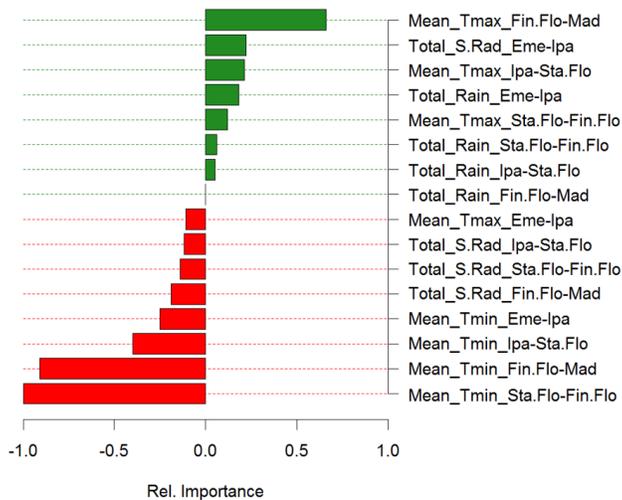
La importancia relativa de las variables predictoras, se define como la contribución de cada una de las variables de entrada en una red neuronal para predecir la variable dependiente o variable de respuesta en la capa de salida (Ibrahim, 2013).

En este estudio se utilizaron los métodos propuesto por Garson & David (1991) y Goh (1995), para dividir los pesos de las conexiones de la red neuronal, con el fin de determinar la importancia relativa de cada variable de entrada en la red. Dichas metodologías fueron seleccionadas, porque además de que han sido aplicadas ampliamente en estudios de ecología y agricultura (Ibrahim, 2013; Olden & Jackson, 2002; Uli, Olden, Ld, Ackson, & Olden, 2002; Zhou et al., 2015), ofrecen una fácil interpretación de los resultados.

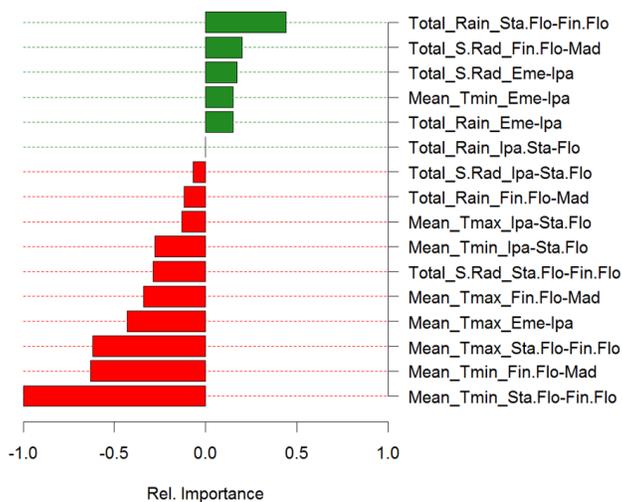
En la Figura. 41 se muestra los resultados de la importancia relativa de las variables predictoras sobre el rendimiento simulado en los ambientes LFE, FE y HFE. Las barras de color verde hacen referencia a las variables que tuvieron una importancia relativa positiva (entre 0 y 1) sobre el rendimiento, mientras que las barras de color rojo indican las variables con importancia relativa negativa (entre -1 y 0) sobre la variable de respuesta. Dado que el objetivo de este análisis es identificar, dentro de las variables consideradas, el factor ambiental que afecta el rendimiento del cultivo de arroz en los diferentes ambientes de producción, los comentarios a continuación están enfocados sobre las variables que ejercen la mayor importancia relativa de forma negativa sobre la variable de respuesta.

Los resultados obtenidos revelaron que existe una fuerte relación entre el rendimiento y la temperatura mínima media, registrada entre el inicio y final de la floración del cultivo de arroz. En los tres ambientes se encontró que esta variable es la que ejerce el mayor efecto negativo sobre el rendimiento, teniendo una importancia relativa (negativa) que supera el 80%; seguido por la temperatura mínima media registrada entre el final de la floración y la madurez fisiológica, con una importancia relativa superior al 60%. Este resultado evidencia que en los tres ambientes, el rendimiento del cultivo de arroz disminuye significativamente a medida aumenta la temperatura mínima durante las fases fenológicas mencionadas anteriormente (fase reproductiva).

Low Favorable Environment (LFE)



Favorable Environment (FE)



High Favorable Environment (HFE)

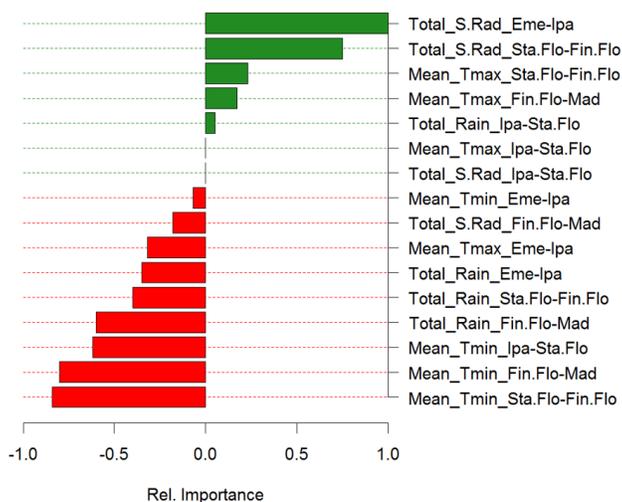


Figura. 41. Importancia relativa de las variables predictoras en los modelos neuronales de los diferentes ambientes, basados en los algoritmos de Garson & David (1991) y Goh (1995).

Otras variables como la temperatura mínima media y máxima media, registradas entre la iniciación de panícula y el inicio de la floración del cultivo, aparecen como factores que también afectan negativamente el rendimiento, pero con una menor importancia relativa.

# Capítulo 4.

## Discusión

“¿Cómo que si puedo? Para el que cree, todo es posible”

**Marcos 9:23 (NVI)**

### 4.1 Consistencia entre los ensayos experimentales y las simulaciones de ORYZA2000.

Este estudio demostró que el modelo ORYZA2000 tuvo un muy buen desempeño al representar, bajo diferentes condiciones ambientales, los patrones de fenología, producción de biomasa, y rendimiento de las variedades consideradas durante esta investigación.

#### ***Parametrización de la fenología del cultivo***

Gracias a que las tasas de desarrollo que se obtuvieron para las distintas etapas del cultivo (DVRJ, DVRP, y DVRR) fueron bastante precisas, no se encontraron diferencias estadísticamente significativa entre la fenología simulada por el modelo y las observaciones reales en los experimentos de campo (ver Tabla 8, Tabla 11, y Tabla 12). De acuerdo a los análisis estadísticos realizados, hubo una variación no superior a 4 días en los valores de RMSE absoluto para cada etapa fenológica. La prueba t-Student (95%

de confianza) indicó que los valores simulados de la fenología fueron muy similares a los observados en campo.

### ***Dinámica de Biomasa, Área Foliar y Rendimiento (Datos de Calibración)***

De acuerdo a la Tabla 16, para el grupo de datos considerado en el proceso de calibración del modelo de cultivo, para todas las variedades, la regresión lineal entre los valores observados y simulados tuvo una pendiente  $\alpha$  relativamente cercana a 1 ( $> 0.7$ ), los mayores valores se obtuvieron para la biomasa total ( $\alpha \pm 0.80$ ), mientras que para el rendimiento se reportaron los valores más bajos ( $\alpha \pm 0.55$ ).

Para todos los casos, el intercepto  $\beta$  fue más pequeño respecto al promedio de los datos observados, excepto para el rendimiento de la variedad Fedearroz 733, indicando una sobre estimación en los rendimientos simulados para los ensayos de esta variedad. El coeficiente de determinación calculado para todos los parámetros fue relativamente alto ( $R^2 \pm 0.80$ ), lo que demuestra una baja dispersión de los datos. Para todas las variedades, la prueba t-Student a un nivel de confianza del 95%, indicó que las variables simuladas fueron similar a los valores observados, excepto para la biomasa de hojas verdes en las variedades Fedearroz 733 y Fedearroz 60, que reportaron valores de  $P_t < 0.05$ .

Por su parte, el RMSE absoluto fue consistentemente menor a la desviación estándar de los datos observados, tanto en los parámetros de crecimiento, como en LAI y rendimiento. Por otro lado, el RMSE estandarizado promedio fue relativamente bajo para el rendimiento y la biomasa total del cultivo, tomando valores de 15% y 25%, respectivamente, mientras que la biomasa simulada en panículas y tallos mostraron los mayores RMSE estandarizado, oscilando alrededor del 35%.

### ***Dinámica de Biomasa, Área Foliar y Rendimiento (Datos de Validación)***

Para todas las variedades, la variable que mostró la mayor relación entre datos observados y simulados, fue la materia seca total (WAGT) del cultivo. Los parámetros  $\alpha$ ,  $\beta$  y  $R^2$  de la regresión lineal, el RMSE absoluto y el estadístico t-Student, entre otros, estuvieron dentro de los rangos deseados para justificar un buen ajuste de los valores simulados. En promedio, la pendiente de la regresión estuvo muy cerca de 1 ( $\alpha > 0.80$ ), el intercepto  $\beta$  fue notablemente inferior a los valores promedios de materia seca total observada en campo. El alto valor del coeficiente de determinación ( $R^2 > 0.85$ ) evidencia un buen ajuste

de los valores simulados respecto a los datos observados, a partir del modelo de regresión lineal implementado. Para todas las variedades, EL RMSEa de la materia seca total fue bastante pequeño en relación a la desviación estándar de los datos observados, y el RMSEn varió entre 20% - 30%.

La biomasa de hojas verdes (WLVG), tallos (WST) y panículas (WSO), también fueron bien simuladas por el modelo, a pesar de que en algunos casos hubo sobreestimación o subestimación de estas variables, los dinámicas de producción de materia seca en estos órganos fue bien representada por el modelo. Los valores de la prueba t-Student obtenidos para estas variables fueron superior a 0.05, indicando que las dinámicas de biomasa simulada y observada fueron relativamente similares, a un nivel de confianza del 95%. Respecto al RMSEn, todas estas variables presentaron un error entre 30% – 37%, y al igual que la materia seca total simulada, presentaron un RMEa notablemente inferior a la desviación estándar de los datos observados.

El índice de área foliar para las variedades CT21375 y Fedearroz 2000, fue simulado con un error absoluto de  $\pm 1.0$ , siendo subestimado en la mayoría de las veces durante la etapa de floración del cultivo. No obstante, los valores obtenidos de  $\alpha$ ,  $\beta$  y  $R^2$ , confirman el buen desempeño del modelo para simular la dinámica del LAI en las diferentes etapas fenológicas del cultivo.

Por último, los parámetros de bondad de ajuste relacionados con el rendimiento simulado de las cuatro variedades, muestran que esta fue la variable con menor RMSEn (< 15%). A pesar de que en la mayoría de los casos se observó una sobre estimación por parte del modelo de cultivo, los valores simulados siempre estuvieron dentro de los intervalos de desviación estándar establecidos alrededor de la recta 1:1, con un RMSEa promedio de 900 kg ha<sup>-1</sup>. Para todas las variedades siempre se obtuvo un estadístico t-Student > 0.05, y un coeficiente de determinación promedio > 0.65.

## 4.2 Efecto de la alta temperatura mínima sobre el cultivo de arroz, en los diferentes grupos ambientales.

Los resultados encontrados en esta investigación mostraron que la temperatura mínima es el principal factor ambiental que afecta de manera negativa el rendimiento del cultivo de arroz en los tres ambientes identificados (Figura. 41), especialmente durante el periodo comprendido entre las etapas de iniciación de panícula (DVS = 0.65) y la floración al 50% (DVS = 1.0). Estos resultados concuerdan con los resultados obtenidos por Garcés Varón (2013), quien observó a través de ensayos experimentales, un aumento en el porcentaje de vaneamiento y reducción de la tasa de llenado de grano de las variedades Fedearroz 733 y Fedearroz 50, al ser sometidas a estrés térmico en la etapa de floración.

En la Figura. 42 muestra el cambio en el rendimiento (a - c) y en el número de espigas (d - f) del cultivo de arroz, en los ambientes LFE, FE y HFE; como respuesta al incremento en la temperatura mínima media (g - i) durante la etapa comprendida entre el inicio y el final de la floración del cultivo. De acuerdo a la Figura. 42 (a - c), el incremento relativo de la temperatura mínima media, en la etapa de floración, genera una pérdida gradual en el rendimiento del cultivo. En el LFE las temperaturas nocturnas oscilan entre 20°C – 26°C (23.5°C en promedio), y los valores más bajos de rendimiento ( $< 5,000 \text{ kg ha}^{-1}$ ), y número de espigas ( $50,000 \text{ m}^{-2}$ ) se alcanzan cuando la temperatura mínima excede los 24°C. Por su parte, el FE tiene un rango de variación de temperatura mínima mucho más amplio en relación LFE, con una mayor frecuencia de temperaturas nocturnas más bajas, las cuales varían entre 15°C y 27°C (21.5°C en promedio), ocasionando menor desempeño tanto en rendimiento ( $< 7,000 \text{ kg ha}^{-1}$ ) como en número de espigas ( $47,000 \text{ m}^{-2}$ ) cuando se exceden los 23.5°C. Por último en el HFE, donde se presentan las más bajas temperaturas mínimas (20°C en promedio) de toda la región de estudio, se alcanzan las altas disminuciones, tanto de rendimientos como de número de espigas, cuando las temperaturas nocturnas superan los 20°C (ver Figura. 43).

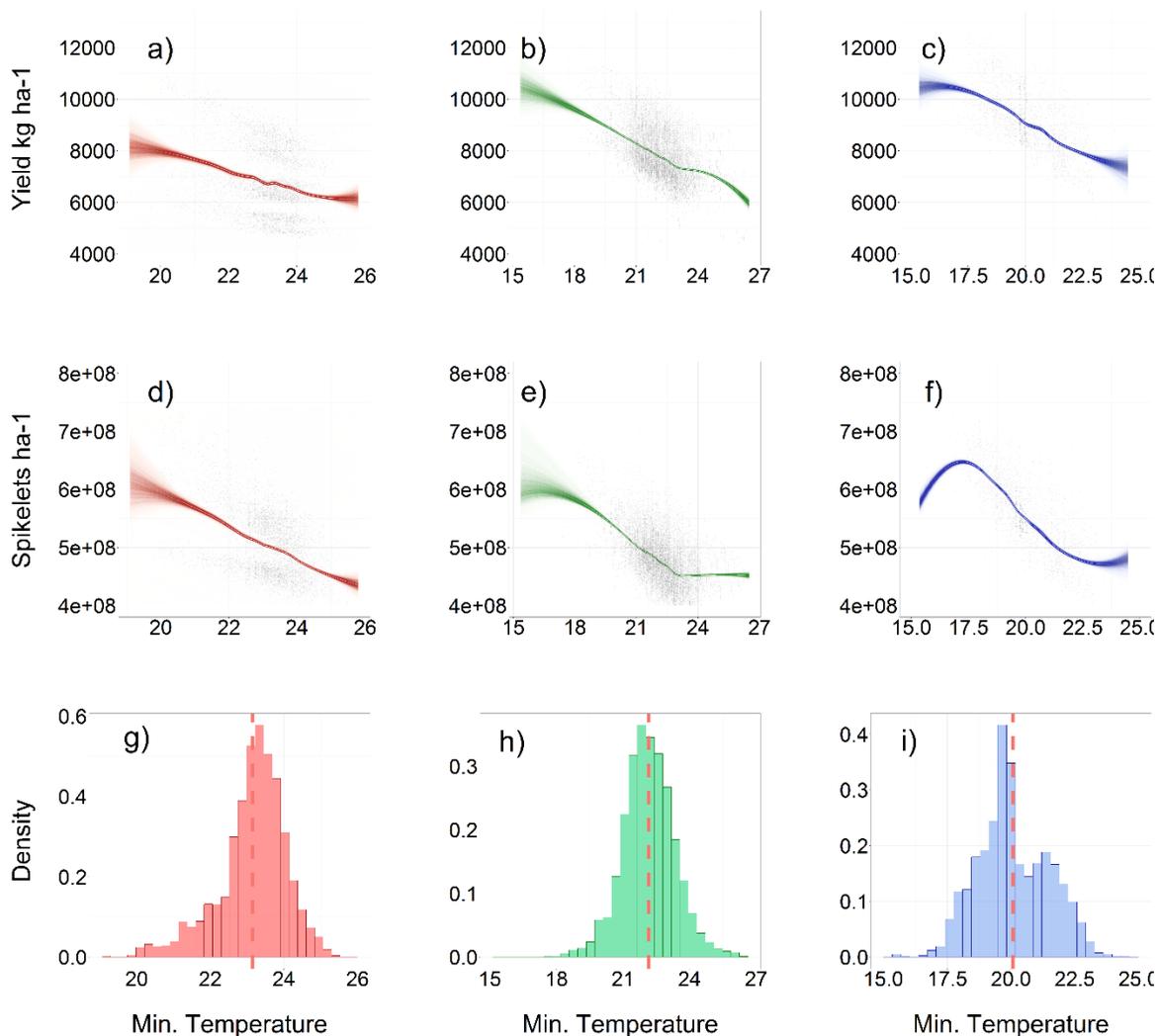


Figura. 42. Respuesta del rendimiento simulado y del número de espigas, al cambio de temperatura mínima media entre el inicio y el final de la floración del cultivo, en los ambientes menos favorables (LFE), Favorable (FE) y Altamente Favorable (HFE). a – c) Perfil promedio de la variación en el rendimiento ( $\text{kg ha}^{-1}$ ) en LFE, FE y HFE. d – f) Perfil promedio de la variación en el número de espigas por hectárea ( $\text{Num ha}^{-1}$ ). g – i) Distribución de probabilidad de la temperatura mínima media entre el inicio y el final de la floración del cultivo, para cada uno de los escenarios de simulación que forman parte de los ambientes LFE, FE y HFE. La línea punteada roja representa la temperatura mínima promedio de cada uno de los histogramas.

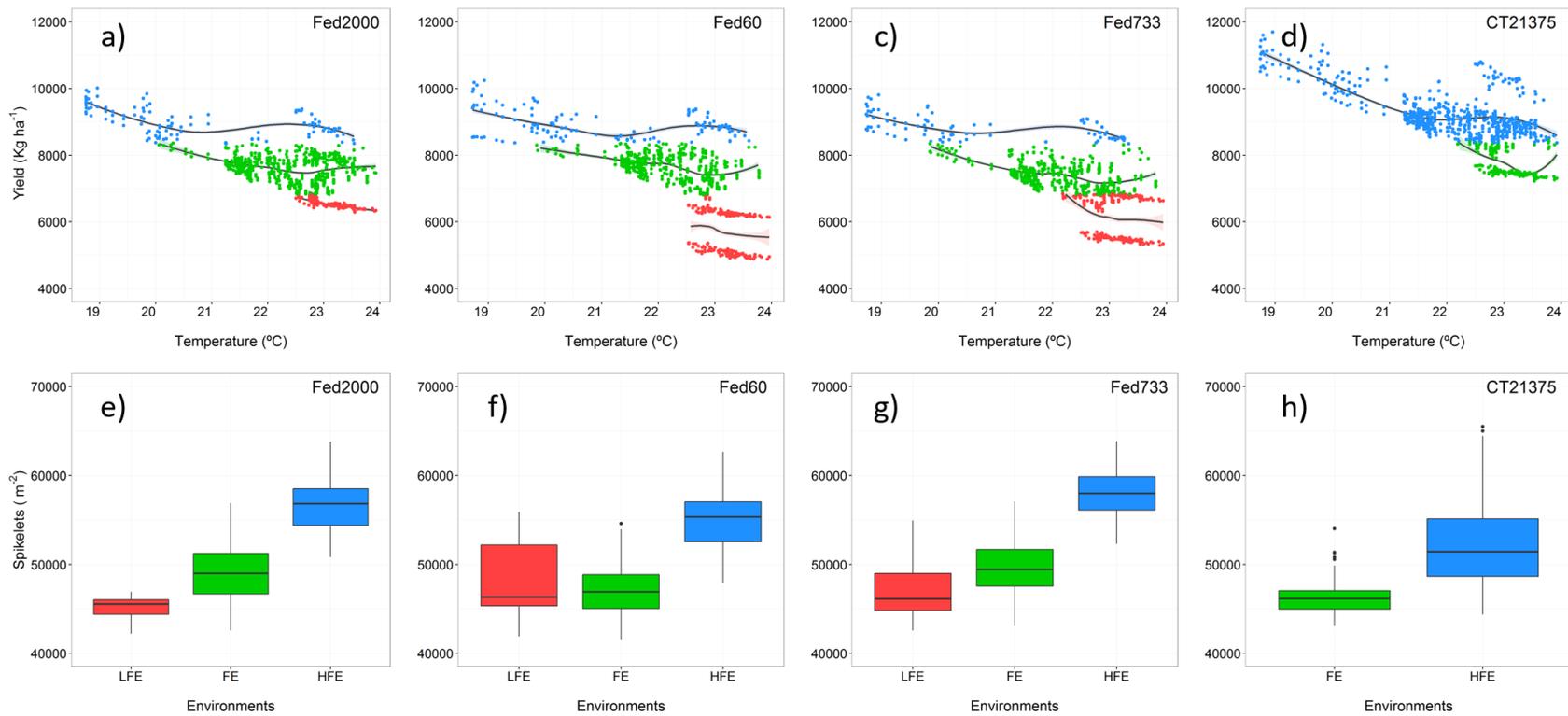


Figura. 43. a – d) Perfil de distribución del rendimiento promedio (kg h<sup>-1</sup>) de las variedades Fedearroz 2000, Fedearroz 60, Fedearroz 733 y CT21375, en función del cambio de la temperatura mínima media durante el inicio y el final de la floración, para los ambientes LFE (puntos rojos), FE (puntos verdes) y HFE (puntos azules). e – f) Número de espigas (Num m<sup>-2</sup>) de cada una de las variedades en los diferentes ambientes de producción (LFE – Cajas rojas; FE – Cajas verdes; HFE – Cajas azules). Para los diagramas de caja, las líneas horizontales a través de las cajas representan la mediana, la altura representa el rango intercuartílico (desde el percentil 25 hasta el percentil 75) y los bigotes, las líneas que se extienden desde las cajas, se extienden hasta los valores mínimo (5.0 %) y máximo (95.0 %) de los datos.

Desde el punto de vista fisiológico, el número potencial de espigas en el cultivo de arroz se determina durante el inicio de la fase reproductiva cuando se presenta la diferenciación de las primeras ramificaciones de la panoja (Kato, Kamoshita, & Yamagishi, 2008). Durante esta fase, el cultivo presenta una alta sensibilidad a la temperatura, especialmente cuando se produce la formación de la panícula. Un acelerado aumento de temperatura durante esta etapa fenológica puede tener un marcado efecto sobre la formación panicular, reduciendo el número final de espiguillas o granos potenciales en la madurez del cultivo (Ansari, Yamamoto, Yoshida, Miyazaki, & Wang, 2003). Como ya se ha mencionado, los resultados del presente estudio demostraron que el aumento de las temperaturas mínimas durante la etapa reproductiva del cultivo redujo significativamente el número de espiguillas por panícula, y que el nivel de afectación varió de acuerdo a las condiciones de temperatura de cada uno de los ambientes identificados.

Estudios realizados por Laza et al (2015) reportaron que la incidencia de altas temperaturas nocturnas ( $>24.8 \pm 0.2$  °C) al inicio de la fase reproductiva del cultivo (R0 – R<sup>2</sup>) genera un menor número de espiguillas por panícula, afectando cerca del 30% del total de espigas generadas por planta, indicando que la presencia de altas temperaturas mínimas en la etapa de diferenciación de la panoja es perjudicial para la formación de las espigas. De igual manera, Y. Zhang et al (2013) encontraron una reducción en el número de espiguillas por panícula, causada por un ligero aumento de la temperatura nocturna durante el crecimiento reproductivo del cultivo. Por otro lado, Cheng et al. (2009) mostraron que el número de espiguillas por panícula no se afectó significativamente por el aumento de temperaturas mínimas en la etapa reproductiva, sin embargo, el porcentaje de fertilidad si tuvo una notable reducción, indicando un fuerte efecto de las temperaturas nocturnas en la formación del grano. Previos estudios han mostrado que temperaturas mínimas  $>29.0$ °C aumentaron la esterilidad de la espiga de arroz, y en consecuencia, se redujo tanto el número de granos como el rendimiento del cultivo (Morita et al., 2005; Satake2 & Yoshida, 1978).

A pesar de que algunos estudios han documentado los efectos de las altas temperaturas en horas de la noche sobre los procesos fisiológicos del cultivo de arroz, se debe reconocer que se sabe poco sobre los mecanismos fisiológicos detrás de la disminución del rendimiento del arroz a elevadas temperaturas nocturnas. Al respecto, estudios realizados por Mohammed & Tarpley (2009) reportaron una relación negativa entre el rendimiento y la tasa de respiración del cultivo en horas de la noche. De igual forma, Laza et al (2015)

encontraron un aumento la tasa de respiración como respuesta al incremento de la temperatura nocturna, además de registrar una reducción significativa en el área foliar y en la biomasa seca total, pasados 15 días después de la iniciación de la panícula. Dicha reducción en la producción de biomasa seca total se le atribuyó al aumento de la tasa de respiración y a la disminución de la tasa fotosintética reportada durante altas temperaturas mínimas.

### **4.3 Efecto de las altas temperaturas y la alta humedad relativa sobre los rendimientos del ambiente menos favorable (LFE).**

El ambiente menos favorable se caracteriza por ser el ambiente de producción con los más bajos rendimientos de la región de estudio ( $< 6,500 \text{ k ha}^{-1}$ ). Con una probabilidad de ocurrencia del 32%, el LFE se distingue por distribuirse espacialmente en el departamento de Córdoba. Esta región, además de presentar las mayores temperaturas mínimas de todo el área de estudio, también se caracteriza por las altas tasas de humedad relativa ( $> 80\%$ ) que se presentan durante todo el año.

De acuerdo a estudios realizados por Yan et al (2010), la humedad relativa en el LFE también puede jugar un papel importante en el rendimiento del cultivo de arroz, ya que la combinación de una alta humedad relativa y altas tasas de temperatura mínima durante la fase de floración del cultivo, afecta negativamente la fertilidad de la espiga. Abeysiriwardena, Ohba, & Maruyama (2002) encontraron que después de la iniciación de la panícula, condiciones de humedad relativa entre el 85.0% - 90.0%, y temperaturas nocturnas  $> 25^{\circ}\text{C}$ , inducen la esterilidad en el grano de arroz. Por su parte, Weerakoon, Maruyama, & Ohba (2008) observaron que el derramamiento del polen sobre la estigma se ve reducido cuando se presenta alta humedad relativa y alta temperatura en el ambiente, fenómeno que no se observó mientras se tuvo alta humedad y bajas temperaturas. Estas observaciones sugieren que el derramamiento de polen sobre el estigma y la posterior esterilidad de la espiguilla, se ven afectados por una alta humedad relativa junto a un aumento en la temperatura.

En este mismo estudio, se encontró que la temperatura dentro de la espiga disminuye con una reducción de la humedad relativa, posiblemente debido a que mejora la tasa de transpiración con el descenso de humedad. Esta reducción de temperatura dentro de la espiga aumenta la viabilidad de los granos de polen. Los granos de polen viables, absorben la humedad y se hinchan bajo niveles moderados a altos de humedad relativa, creando la presión requerida para la ruptura del septum, lo que ayuda en la deposición del polen sobre el estigma, produciendo así una espiguilla fertilizada.

## Capítulo 5.

### Conclusiones

“En paz me acuesto y me duermo,  
porque sólo tú, SEÑOR, me haces vivir  
confiado”

**Salmos 4:8 (NVI)**

- ✓ El modelo de cultivo ORYZA2000 mostró un gran potencial para simular las diferentes etapas fenológicas (iniciación de panícula, floración y madurez fisiológica), la dinámica de producción de biomasa de los diferentes órganos de la planta, el Índice de Área Foliar, y el rendimiento de los granos, arrojando valores cercanos a los datos observados en campo para los cuatro cultivares considerados en este estudio.
- ✓ Los índices de evaluación RMSE, RMSE normalizado y  $R^2$ , confirmaron que no hubo diferencias significativas entre las predicciones del modelo de cultivo y las observaciones en los experimentos de campo, lo que indica que ORYZA2000 es capaz de proporcionar una estimación fiable del rendimiento del cultivo bajo diferentes condiciones ambientales.
- ✓ De acuerdo a los resultados obtenidos, las regiones Centro y Bajo Cauca podrían dividirse en tres grupos ambientales para la producción del cultivo de arroz: Ambiente Altamente Favorable (HFE), Ambiente Favorable (FE), y Ambiente Menos Favorable (LFE).

- 
- ✓ La probabilidad de ocurrencia de los tres grupos ambientales a través de la región de estudio, se distribuye de la siguiente manera: HFE (18,5%), FE (50,2%) y LFE (31,3%).
  - ✓ La propuesta metodológica implementada, que combina modelos de cultivos con Redes Neuronales Artificiales, puede ser considerada una herramienta potencial para agrónomos y mejoradores de cultivos, ya que permite caracterizar y evaluar el crecimiento y desarrollo de un cultivo en un gran número de sitios, minimizando costos económicos y a un menor tiempo.
  - ✓ Los resultados mostraron que la temperatura mínima, entre la iniciación de la panícula y el final de la floración, fue el principal factor de estrés abiótico que afectó negativamente el rendimiento de los cultivos en los tres ambientes.
  - ✓ En el ambiente menos favorable (LFE), donde existe una alta frecuencia de altas temperaturas mínimas (23°C – 26°C) durante los periodos regulares de siembra y cosecha del cultivo del arroz, es necesario aplicar una estrategia de adaptación, la cual pueda contrarrestar este tipo de estrés.

## Capítulo 6.

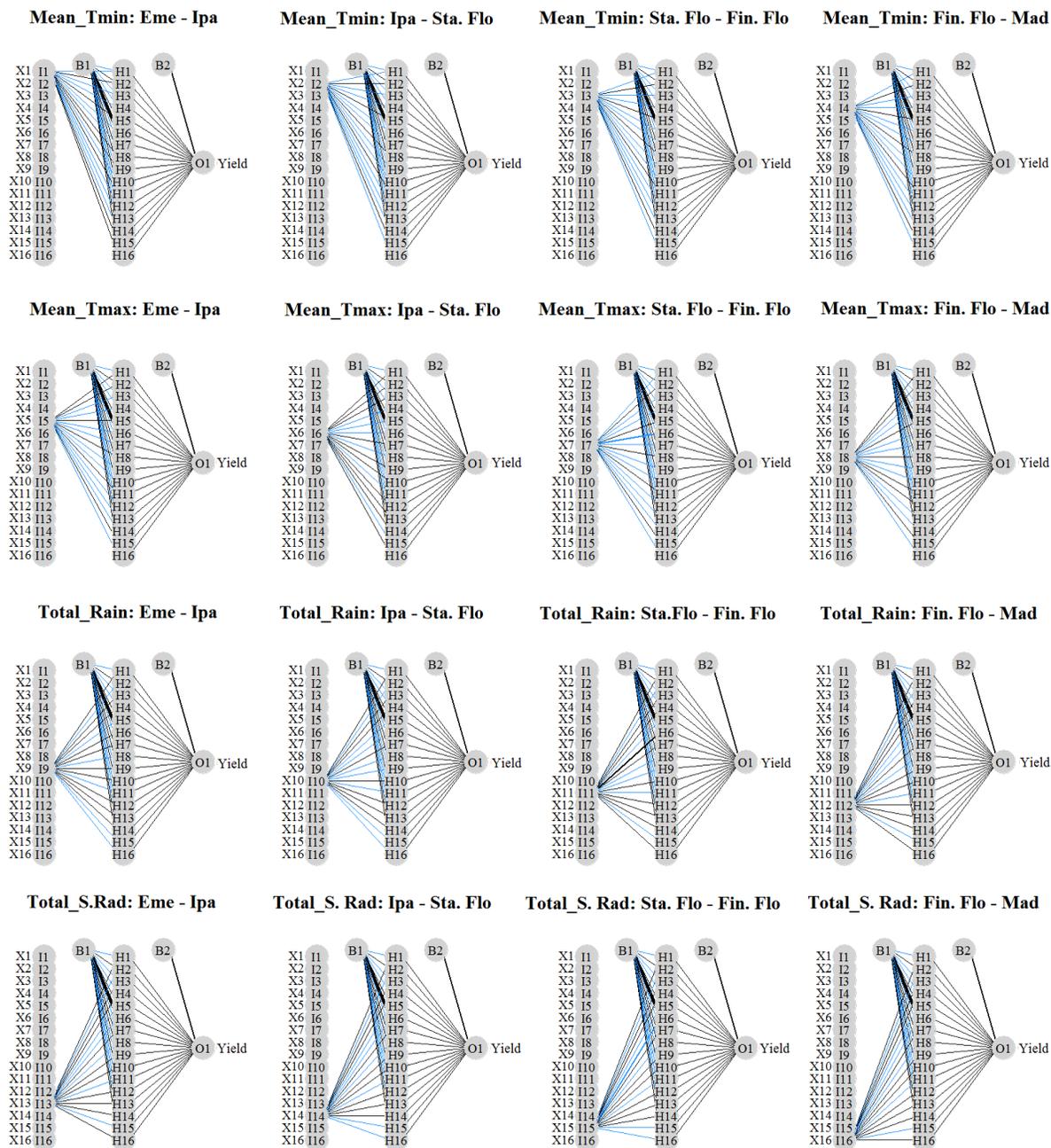
### Recomendaciones

“En paz me acuesto y me duermo,  
porque sólo tú, SEÑOR, me haces vivir  
confiado”

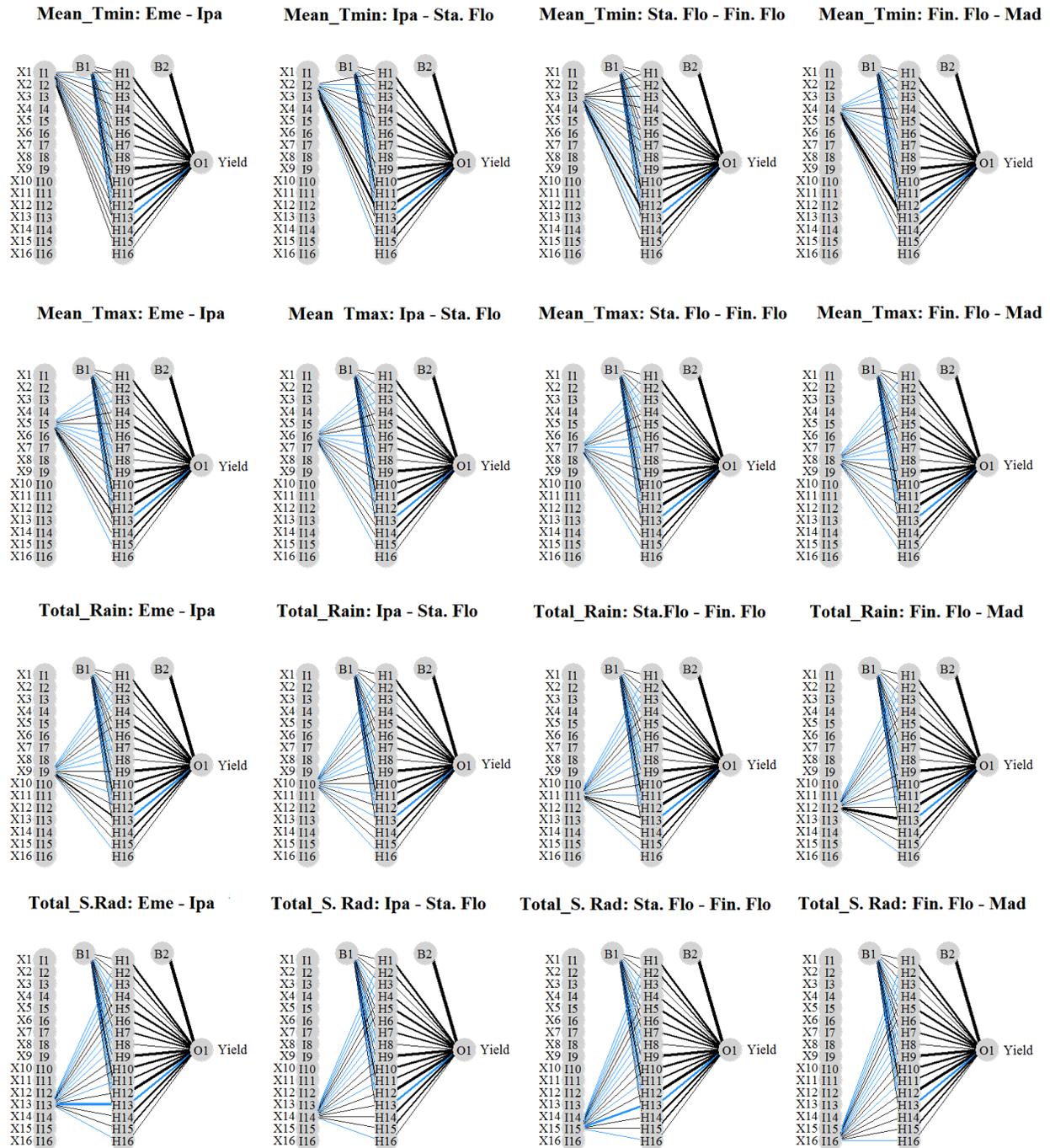
**Salmos 4:8 (NVI)**

- ✓ Es de gran importancia considerar este tipo de estudios en las futuras estrategias que los programas de mejoramiento de cultivos están implementando en el país. Una zonificación ambiental permite identificar las características agroclimáticas de una región productora y esto podría ser muy útil para aplicar de una manera localizada y eficiente las nuevas tecnologías.
- ✓ Se deben considerar estudios futuros que involucren las nuevas proyecciones de cambio climático, con el fin de dimensionar los cambios en la configuración espacial de los ambientes identificados, y las variaciones en la intensidad del factor limitante (estrés abiótico) al rendimiento de los cultivos.
- ✓ Se deben llevar a cabo nuevos estudios que combinen modelos de cultivos y ensayos multiambientales realizados por los programas de mejoramiento, con el fin de mejorar el nivel de entendimiento de la interacción existente entre el Genotipo, el Ambiente y el Manejo del cultivo.

# A: Anexo: Diagramas de Interpretación Neuronal para la descomposición del modelo RNA obtenido en el grupo ambiental FE

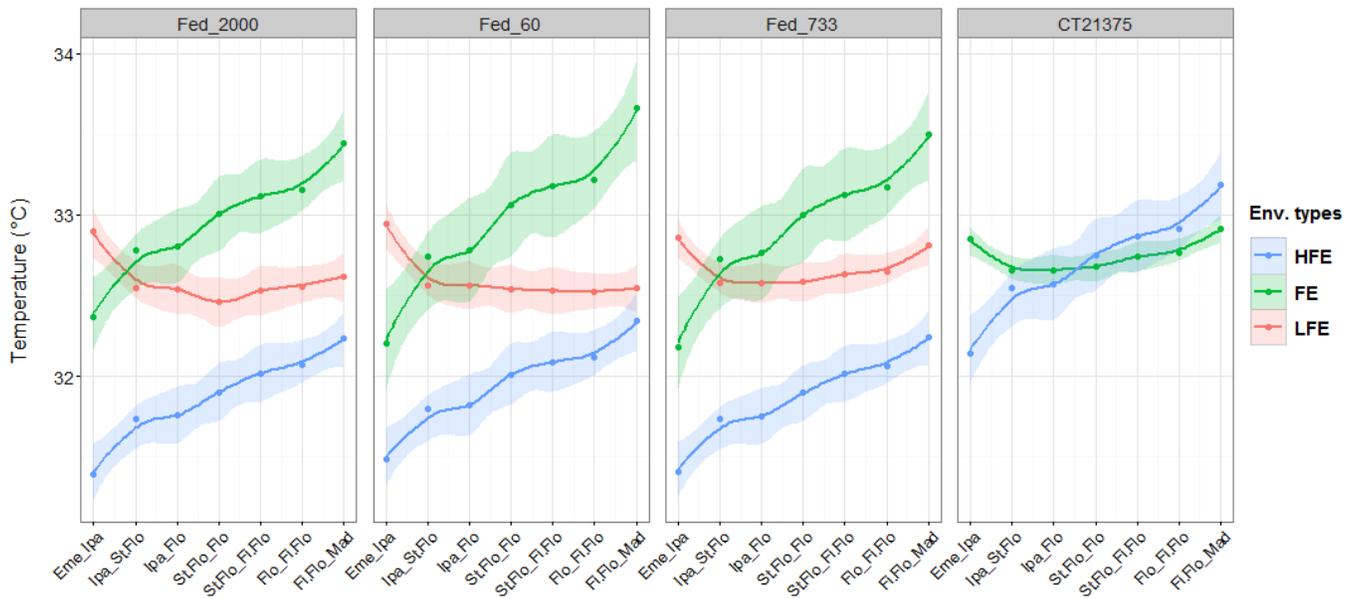


# B: Anexo: Diagramas de Interpretación Neuronal para la descomposición del modelo RNA obtenido en el grupo ambiental HFE

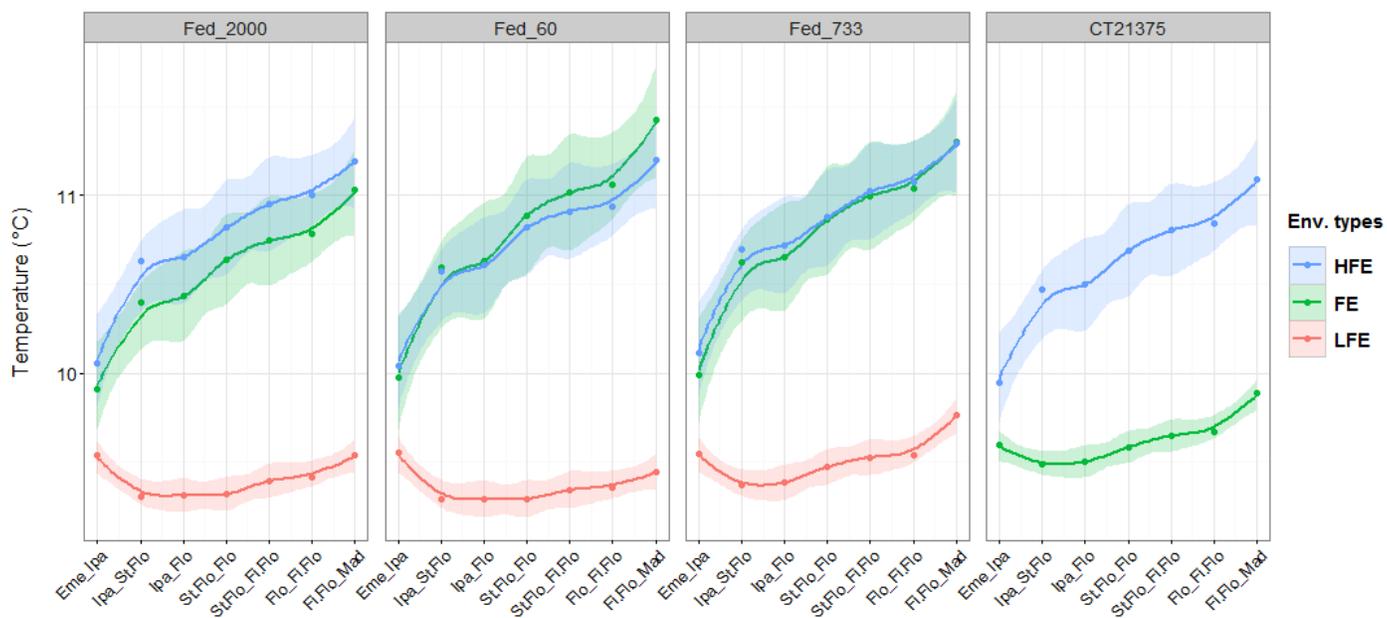




## D: Anexo: Perfil de variación de la temperatura máxima media, a través el ciclo fenológico de las diferentes variedades, en los ambientes LFE, FE y HFE



## E: Anexo: Perfil de variación de la Oscilación de temperatura, a través el ciclo fenológico de las diferentes variedades, en los ambientes LFE, FE y HFE



## Bibliografía

- Abey Siriwardena, D. S. D. Z., Ohba, K., & Maruyama, A. (2002). Influence of temperature and relative humidity on grain sterility in rice. *Journal of the National Science Foundation of Sri Lanka*, 30(1–2), 33–41.
- Aggarwal, P. K., Kropff, M. J., Cassman, K. G., & ten Berge, H. F. M. (1997). Simulating genotypic strategies for increasing rice yield potential in irrigated, tropical environments. *Field Crops Research*, 51(1–2), 5–17.
- Ansari, T. H., Yamamoto, Y., Yoshida, T., Miyazaki, A., & Wang, Y. (2003). Cultivar differences in the number of differentiated spikelets and percentage of degenerated spikelets as determinants of the spikelet number per panicle in relation to dry matter production and nitrogen absorption. *Soil Science and Plant Nutrition*, 49(3), 433–444.
- Areiza Segura, A. M. (2013). *Diagnóstico del mercado del arroz en Colombia (2000-2012)*. Retrieved from <http://www.sic.gov.co>
- Barnett, K., & Pearce, R. (1983). Source-sink ratio alteration and its effect on physiological parameters in maize. *Crop Science*.
- Barrios-Perez, C. (2012). *INCIDENCIA DE “EL NIÑO”- OSCILACIÓN DEL SUR (ENOS) SOBRE LA PRODUCTIVIDAD DEL CULTIVO DE LA CAÑA DE AZÚCAR EN EL VALLE DEL RÍO CAUCA (COLOMBIA)*. Universidad del Valle.
- Boling, A. A., Bouman, B. A. M., Tuong, T. P., Murty, M. V. R., & Jatmiko, S. Y. (2006). Modelling the effect of groundwater depth on yield-increasing interventions in rainfed lowland rice in Central Java, Indonesia ARTICLE IN PRESS. *Agricultural Systems*.
- Boote, K. J., Jones, J. W., White, J. W., Asseng, S., & Lizaso, J. I. (2013). Putting mechanisms into crop production models. *Plant, Cell and Environment*, 36(9), 1658–1672.
- Bouman, B. A. M., Van Keulen, H., Van Laar, H. H., & Rabbinge, R. (1996). The “School of de Wit” crop growth simulation models: A pedigree and historical overview. *Agricultural Systems*, 52(2–3), 171–198.

- 
- Bouman, B. A. M., & Van Laar, H. H. (2006). Description and evaluation of the rice growth model ORYZA2000 under nitrogen-limited conditions. *Agricultural Systems*, 87(3), 249–273.
  - Bouman, B. a M., Kropff, M., Tuong, T., Wopereis, M., Ten Berge, H., & van Laar, H. (2001). *ORYZA2000: Modeling lowland rice*.
  - Briceño, J. L. (2012). *Regiones rurales “Propuesta de regionalización rural del país.”* Universidad Nacional de Colombia.
  - Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M., & Ruget, F. (1998). STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie*.
  - Casadebaig, P., Zheng, B., Chapman, S., Huth, N., Faivre, R., & Chenu, K. (2016). Assessment of the potential impacts of wheat plant traits across environments by combining crop modeling and global sensitivity analysis. *PLoS ONE*, 11(1), 1–27.
  - Castro, L. M., & Carvajal, Y. (2010). Análisis de tendencia y homogeneidad de series climatológicas. *Universidad Del Valle - Ingeniería de Recursos Naturales Y Del Ambiente*, 9, 15–25.
  - Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, 61(6), 1–36.
  - Chauhan, Y. S., & Rachaputi, R. C. N. (2014). Defining agro-ecological regions for field crops in variable target production environments: A case study on mungbean in the northern grains region of Australia. *Agricultural and Forest Meteorology*, 194, 207–217.
  - Chauhan, Y., Wright, G., Rachaputi, N., & Mccosker, K. (2008). Identifying chickpea homoclimes using the APSIM chickpea model. *Australian Journal of Agricultural Research*, 59, 260–269.
  - Cheng, W., Sakai, H., Yagi, K., & Hasegawa, T. (2009). Interactions of elevated [CO<sub>2</sub>] and night temperature on rice growth and yield. *Agricultural and Forest Meteorology*, 149(1), 51–58.

- Chenu, K. (2015). Characterizing the crop environment–nature, significance and applications. *Crop Physiology. Applications for Genetic Improvement and Agronomy*, (DECEMBER).
- Chenu, K., Cooper, M., Hammer, G. L., Mathews, K. L., Dreccer, M. F., & Chapman, S. C. (2011). Environment characterization as an aid to wheat improvement: Interpreting genotype-environment interactions by modelling water-deficit patterns in North-Eastern Australia. *Journal of Experimental Botany*, 62(6), 1743–1755.
- Chenu, K., Deihimfard, R., & Chapman, S. C. (2013). Large-scale characterization of drought pattern: A continent-wide modelling approach applied to the Australian wheatbelt - spatial and temporal trends. *New Phytologist*, 198(3), 801–820.
- Confalonieri, R., Acutis, M., Bellocchi, G., & Donatelli, M. (2009). Multi-metric evaluation of the models WARM, CropSyst, and WOFOST for rice. *Ecological Modelling*.
- Cuadras, C. M. (2014). NUEVOS MÉTODOS DE ANÁLISIS MULTIVARIANTE.
- D'Afonseca, D. S., Heinemann, A. B., da Silva, S. C., & da Cunha Moraes, A. (2013). Metodologia para o preenchimento de dados climáticos diários faltantes: uma aplicação para os municípios de Goiás a partir da base de dados de clima da Embrapa Arroz e Feijão. *Embrapa Arroz E Feijão. Documentos*, (296).
- DANE. (2014). *Boletín técnico I semestre - Encuesta Nacional de Arroz Mecanizado*.
- DANE. (2015). *Boletín técnico I semestre - Encuesta Nacional de Arroz Mecanizado* (Vol. 2015).
- Diepen, C. Van, Keulen, H. Van, & Vries, F. de. (1987). Simulated variability of wheat and rice yields in current weather conditions and in future weather when ambient CO<sub>2</sub> has doubled.
- Dingkuhn, M., Kumar, U., Laza, M., & Pasco, R. (2016). SAMARA: A crop model for simulating rice phenotypic plasticity.

- 
- Drenth, H., ten Berge, H. F. M., & Riethoven, J. J. M. (1994). ORYZA simulation modules for potential and nitrogen limited rice production. *SARP Research Proceedings, IRRI/AB-DLO, Wageningen, Netherlands*, 223.
  - EKANAYAKE, I. J., DATTA, S. K. DE, & STEPONKUS, P. L. (1989). Spikelet Sterility and Flowering Response of Rice to Water Stress at Anthesis. *Annals of Botany*, 63(2), 257–264.
  - Erazo-Meza, E. (2011). *EFFECTO DE LOS FACTORES LIMITANTES DE LA PRODUCTIVIDAD DEL CULTIVO DE CAÑA DE AZÚCAR A NIVEL INTRASUERTE*. Universidad Nacional de Colombia.
  - Espinosa, J. (2002). Rice Nutrition Management in Latin America, 16(May), 36–39.
  - FAOSTAT. (2016). Food and Agriculture Organization of the United Nations, FAOSTAT (Database).
  - Farquhar, G., Caemmerer, S., & Berry, J. (1980). A biochemical model of photosynthetic CO<sub>2</sub> assimilation in leaves of C<sub>3</sub> species. *Planta*.
  - FEDEARROZ. (2008a). III Censo Nacional arrocero Zona Bajo Cauca.
  - FEDEARROZ. (2008b). III Censo Nacional arrocero Zona Centro, 102.
  - FEDEARROZ. (2011). Adopción masiva de tecnología. *Revista Arroz*, 1–52.
  - FEDEARROZ. (2012). Arroz. Edición Especial, 60 Años de Historia. *Revista Arroz*, 60(496).
  - FEDEARROZ. (2016). Federación Nacional de Arroceros, FEDEARROZ (Database).
  - Feng, L., Bouman, B. A. M., Tuong, T. P., Cabangon, R. J., Li, Y., Lu, G., & Feng, Y. (2007). Exploring options to grow rice using less water in northern China using a modelling approach: I. Field experiments and model evaluation. *Agricultural Water Management*, 88(1), 1–13.
  - Fischer, R. (1985). Number of kernels in wheat crops and the influence of solar radiation and temperature. *The Journal of Agricultural Science*.

- Fukai, S., Basnayake, J., & Cooper, M. (2000). Modelling water availability, crop growth, and yield of rainfed lowland rice genotypes in northeast Thailand.
- Fukai, S., Rajatsasereekul, S., & Boonjung, H. (1995). Simulation modeling to quantify the effect of drought for rainfed lowland rice in Northeast Thailand.
- Fumoto, T., Kobayashi, K., Li, C., & Yagi, K. (2008). Revising a process-based biogeochemistry model (DNDC) to simulate methane emission from rice paddy fields under various residue management and fertilizer. *Global Change*.
- Garcés Varón, G. A. (2013). *influencia de las altas temperaturas, la baja luminosidad y la época de siembra sobre el comportamiento fisiológico de dos cultivadores de arroz*. Universidad Nacional de Colombia.
- Garson, & David, G. (1991). Interpreting neural-network connection weights. *AI Expert*, 6(4), 46–51.
- Gaydon, D., Probert, M., Buresh, R., & Meinke, H. (2012). Rice in cropping systems—Modelling transitions between flooded and non-flooded soil environments. *European Journal of*.
- Gaydon, D. S., Probert, M. E., Buresh, R. J., Meinke, H., & Timsina, J. (2012). Modelling the role of algae in rice crop nutrition and soil organic carbon maintenance. *European Journal of Agronomy*, 39, 35–43.
- Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143–151.
- Gonzáles Calderón, D. C. (2015). *CARACTERIZACIÓN DE LA ARQUITECTURA DE LA PANÍCULA Y CARACTERES AGRONÓMICOS EN UNA POBLACIÓN F2 ENTRE DOS TIPOS DE PLANTA DE ARROZ (*Oryza sativa L.*) CONTRASTANTE*. Universidad del Tolima.
- Hammer, G. L., & Jordan, D. R. (2007). an Integrated Systems Approach To Crop Improvement. *Scale and Complexity in Plant Systems Research: Gene-Plant-Crop Relations*, (November 2016), 45–61.

- 
- Han, E., Ines, A., & Koo, J. (2015). Global high-resolution soil profile database for crop modeling applications, (October), 37.
  - Hasegawa, T., & Horie, T. (1997). Modelling the effect of nitrogen on rice growth and development. *Applications of Systems Approaches at the Field*.
  - Heinemann, A. B., Barrios-Perez, C., Ramirez-Villegas, J., Arango-Londoño, D., Bonilla-Findji, O., Medeiros, J. C., & Jarvis, A. (2015). Variation and impact of drought-stress patterns across upland rice target population of environments in Brazil. *Journal of Experimental Botany*, 66(12), 3625–3638.
  - Heinemann, A. B., Ramirez-Villegas, J., Souza, T., Didonet, A., di Stefano, J., Boote, K. J., & Jarvis, A. (2016). Drought impact on rainfed common bean production areas in Brazil. *Agricultural and Forest Meteorology, in revisio*, 57–74.
  - Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., & Samuel-Rosa, A. (2014). SoilGrids1km -- Global Soil Information Based on Automated Mapping. *PLoS ONE*, 9(8).
  - Heredia, L. M. C., Escobar, Y. C., & Díaz, Á. J. Á. (2012). ANÁLISIS CLÚSTER COMO TÉCNICA DE ANÁLISIS EXPLORATORIO DE REGISTROS MÚLTIPLES EN DATOS METEOROLÓGICOS. *Ingeniería de Recursos Naturales Y Del Ambiente*, (11), 11–20.
  - Hijmans, R. J., Lens, I. M. G., & Diepen, C. A. van. (1994). WOFOST 6.0 : user's guide for the WOFOST 6.0 crop growth simulation model.
  - Hoang, L., Ngoc, T. A., & Maskey, S. (2016). A robust parameter approach for estimating CERES-Rice model parameters for the Vietnam Mekong Delta. *Field Crops Research*, 196, 98–111.
  - Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., ... Keating, B. A. (2014). APSIM – Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, 62, 327–350.
  - Horie, T., Nakagawa, H., & Centeno, H. (1995). The rice crop simulation model SIMRIW and its testing.

- Horie, T., Yajima, M., & Nakagawa, H. (1992). Yield forecasting. *Agricultural Systems*, 40(1–3), 211–236.
- Huang, J., Nie, L., Shah, T., Chen, C., & Wa, D. K. (2017). Impact of high-temperature stress on rice plant and its traits related to tolerance.
- Husson, F., Josse, J., Le, S., & Maitiner, J. M. (2016). *Multivariate Exploratory Data Analysis and Data Mining*.
- Ibrahim, O. M. (2013). A comparison of methods for assessing the relative importance of input variables in artificial neural networks. *Journal of Applied Sciences Research*, 9(11), 5692–5700.
- Islam, M. S., & Morison, J. I. L. (1992). Influence of solar radiation and temperature on irrigated rice grain yield in Bangladesh. *Field Crops Research*, 30(1–2), 13–28.
- Jame, Y. W., & Cutforth, H. W. (1996). Crop growth models for decision support systems.
- JIANG, M., & JIN, Z. (2009). A Method for Upscaling Genetic Parameters of CERES-Rice in Regional Applications. *Rice Science*, 16(4), 292–300.
- Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., ... Ritchie, J. T. (2003). *DSSAT Cropping System Model*. *European Journal of Agronomy* (Vol. 18).
- Kadiyala, M. D. M., Jones, J. W., Mylavarapu, R. S., Li, Y. C., & Reddy, M. D. (2015). Identifying irrigation and nitrogen best management practices for aerobic rice–maize cropping system for semi-arid tropics using CERES-rice and maize models. *Agricultural Water Management*, 149, 23–32.
- Kato, Y., Kamoshita, A., & Yamagishi, J. (2008). Preflowering Abortion Reduces Spikelet Number in Upland Rice ( *L.*) under Water Stress. *Crop Science*, 48(6), 2389.
- Kropff, M. J., H.H. Laar, van, & Matthews, R. B. (1994). ORYZA 1 : an ecophysiological model for irrigated rice production.

- 
- Larijani, B. A., Sarvestani, Z. T., Nematzadeh, G., Manschadi, A. M., & Amiri, E. (2011). Simulating Phenology, Growth and Yield of Transplanted Rice at Different Seedling Ages in Northern Iran Using ORYZA2000. *Rice Science*, 18(4), 321–334.
  - Lau, C., Jarvis, A., & Ramírez, J. (2010). Colombian agriculture: Adapting to climate change. *CIAT Policy Brief No. 1.*, 4.
  - Laza, M. R. C., Sakai, H., Cheng, W., Tokida, T., Peng, S., & Hasegawa, T. (2015). Differential response of rice plants to high night temperatures imposed at varying developmental phases. *Agricultural and Forest Meteorology*, 209–210, 69–77.
  - Li, T., Hasegawa, T., Yin, X., Zhu, Y., Boote, K., Adam, M., ... Bouman, B. (2015). Uncertainties in predicting rice yield by current crop models under a wide range of climatic conditions. *Global Change Biology*, 21(3), 1328–1341.
  - Li, T., Raman, A. K., Marcaida, M., Kumar, A., Angeles, O., & Radanielson, A. M. (2013). Simulation of genotype performances across a larger number of environments for rice breeding using ORYZA2000. *Field Crops Research*, 149, 312–321.
  - Lin, G.-F., & Chen, L.-H. (2006). Identification of homogeneous regions for regional frequency analysis using the self-organizing map. *Journal of Hydrology*, 324(1), 1–9.
  - Luis, J., & Villardón, V. (2007). INTRODUCCION AL ANALISIS DE CLUSTER (p. 22). Salamanca - España.
  - Matthews, R. B., & Hunt, L. A. (1994). GUMCAS: a model describing the growth of cassava (*Manihot esculenta* L. Crantz). *Field Crops Research*, 36(1), 69–84.
  - Mazuera Fernandez, C. A., & Neira Rodriguez, J. D. (2009). *Análisis de los Costos de Producción de Arroz, Oriza Sativa L., en el Municipio de Saldaña, Tolima. Método Pulver vs Método Tradicional de Manejo*. Universidad de la Salle.
  - McLean, J., Hardy, B., & Hettel, G. (2013). *Rice Almanac, 4th edition*. IRRI, Los Baños, Philippines.

- Mennamy, J. M., & O'toole, J. (1983). Rice mod: A physiologically based rice growth and yield model. *IRRI Research Paper Series*.
- Mohammed, A. R., & Tarpley, L. (2009). High nighttime temperatures affect rice productivity through altered pollen germination and spikelet fertility.
- Monteith, J., & Moss, C. (1977). Climate and the efficiency of crop production in Britain.
- Morita, S., Yonemaru, J.-I., & Takanashi, J.-I. (2005). Grain growth and endosperm cell size under high night temperatures in rice (*Oryza sativa L.*). *Annals of Botany*, 95(4), 695–701.
- Nebel, Á. L. C., Timm, L. C., Cornelis, W., Gabriels, D., Reichardt, K., Aquino, L. S., ... Reinert, D. J. (2010). Pedotransfer functions related to spatial variability of water retention attributes for lowland soils. *Revista Brasileira de Ciência Do Solo*, 34(3), 669–680.
- Olden, J. D., & Jackson, D. A. (2002). Illuminating the “black box”: a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, 154(1), 135–150.
- Oteng-Darko, P., Yeboah, S., Addy, S. N. T., Amponsah, S., & Danquah, E. O. (2013). Crop modeling: A tool for agricultural research – A review. *E3 Journal of Agricultural Research and Development*, 2(1), 1–6.
- Özesmi, S. L., & Özesmi, U. (1999). An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecological Modelling*, 116(1), 15–31.
- Penning de Vries, F. W. T. (1989). *Simulation of ecophysiological processes of growth in several annual crops*. IRRI.
- Peterson, L. E. (2002). Hierarchical cluster and principal-component analysis of microarray-based transcriptional profiles. *Genome Biology*, 3(7), SOFTWARE0002.
- R Development Core Team. (2008). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

- 
- Radha, V., & Murthy, K. (2004). CROP GROWTH MODELING AND ITS APPLICATIONS IN AGRICULTURAL METEOROLOGY.
  - Richardson, C. W., & Wright, D. A. (1984). WGEN : A Model for Generating Daily Weather Variables. *U. S. Department of Agriculture, Agricultural Research Service, ARS-8*, 83.
  - Roetter, R., Hoanh, C., & Teng, P. (1998). A systems approach to analyzing land use options for sustainable rural development in South and Southeast Asia; SysNet special project report.
  - Sadras, V., Calderini, D., Messina, C., Hammer, G., Dong, Z., Podlich, D., & Cooper, M. (2009). Modelling Crop Improvement in a G×E×M Framework via Gene–Trait–Phenotype Relationships. *Crop Physiology*, 235–581.
  - Sánchez, B., Rasmussen, A., & Porter, J. R. (2014). Temperatures and the growth and development of maize and rice: A review. *Global Change Biology*, 20(2), 408–417.
  - Satake, T., & Yoshida, S. (1978). High Temperature-Induced Sterility in Indica Rices at Flowering1). *Crop Sci*, 47(1), 6–17.
  - Satapathy, S. S., Swain, D. K., & Herath, S. (2014). *Field experiments and simulation to evaluate rice cultivar adaptation to elevated carbon dioxide and temperature in sub-tropical India. European Journal of Agronomy* (Vol. 54).
  - Satizábal, H. F., & Pérez-Urbe, A. (2007). Relevance Metrics to Reduce Input Dimensions in Artificial Neural Networks.
  - Singh, U., Ritchie, J., & Godwin, D. (1993). *A User's Guide to CERES Rice, V2*. 10.
  - Spitters, C., & Keulen, H. Van. (1989). A simple and universal crop growth simulator: SUCROS87. *Simulation and Systems*.
  - Sudhir-Yadav, S., Li, T., Humphreys, E., Gill, G., & Kukal, S. S. (2011). Evaluation and application of ORYZA2000 for irrigation scheduling of puddled transplanted rice in north west India. *Field Crops Research*, 122(2), 104–117.
  - Svirezhev, Y. M. (1992). *Simulation of ecophysiological process of growth in several annual crops. Field Crops Research* (Vol. 28).

- Tang, L., Zhu, Y., Hannaway, D., Meng, Y., & Liu, L. (2009). RiceGrow: A rice growth and productivity model.
- Tao, F., & Zhang, Z. (2013). Climate change, high-temperature stress, rice productivity, and water use in Eastern China: a new superensemble-based probabilistic projection.
- Tomasella, J., & Hodnett, M. (2004). Pedotransfer functions for tropical soils. *Developments in Soil Science*, 30(C), 415–429.
- Tomasella, J., Pachepsky, Y., Crestana, S., & Rawls, W. J. (2003). Comparison of Two Techniques to Develop Pedotransfer Functions for Water Retention. *Soil Science Society of America Journal*, 67(4), 1085.
- Turner, N. C., O’Toole, J. C., Cruz, R. T., Namuco, O. S., & Ahmad, S. (1986). Responses of seven diverse rice cultivars to water deficits I. Stress development, canopy temperature, leaf rolling and growth. *Field Crops Research*, 13, 257–271.
- UCHIJIMA, T. (1976). Some Aspects of the Relation Between Low Air Temperature and Sterile Spikelets Numbers in Rice Plants. *Journal of Agricultural Meteorology*, 31(4), 199–202.
- Uli, J., Olden, A. D., Ld, D., Ackson, A. J., & Olden, J. D. (2002). A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology*, 47.
- van Ittersum, M. ., Leffelaar, P. ., van Keulen, H., Kropff, M. ., Bastiaans, L., & Goudriaan, J. (2003). On approaches and applications of the Wageningen crop models. *European Journal of Agronomy*, 18(3), 201–234.
- Van Oort, P. A. J., De Vries, M. E., Yoshida, H., & Saito, K. (2015). Improved climate risk simulations for rice in arid environments. *PLoS ONE*, 10(3), 1–27.
- van Wart, J., van Bussel, L. G. J., Wolf, J., Licker, R., Grassini, P., Nelson, A., ... Claessens, L. (2013). Use of agro-climatic zones to upscale simulated crop yield potential. *Field Crops Research*, 143, 44–55.
- Wang, W., Yu, Z., Zhang, W., Shao, Q., Zhang, Y., Luo, Y., ... Xu, J. (2014). Responses of rice yield, irrigation water requirement and water use efficiency to climate change in China: Historical simulation and future projections. *Agricultural Water Management*, 146, 249–261.

- 
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*.
  - Weerakoon, W. M. W., Maruyama, A., & Ohba, K. (2008). Impact of Humidity on Temperature-Induced Grain Sterility in Rice (*Oryza sativa* L). *Journal of Agronomy and Crop Science*, 194(2), 135–140.
  - Woodhead, T. (1979). Simulation of assimilation, respiration and transpiration of crops. By D. T. de Wit et al. Pudoc (Wageningen), 1978. Pp. 140. D.F1.22.50. *Quarterly Journal of the Royal Meteorological Society*, 105(445), 728–729.
  - Wopereis, M., Bouman, B., Tuong, T., Berge, H., & Bafios, L. (1996). SARP Research Proceedings ORYZA\_W: Rice growth model for irrigated and rainfed environments.
  - Xinyou, Y. (1996). Quantifying the effects of temperature and photoperiod on phenological development to flowering in rice.
  - Xue, C., Yang, X., Bouman, B. A. M., Deng, W., Zhang, Q., Yan, W., ... Wang, H. (2008). Optimizing yield, water requirements, and water productivity of aerobic rice for the North China Plain. *Irrigation Science*, 26(6), 459–474.
  - Yan, C., Ding, Y., Wang, Q., Liu, Z., Li, G., Muhammad, I., & Wang, S. (2010). The impact of relative humidity, genotypes and fertilizer application rates on panicle, leaf temperature, fertility and seed setting of rice. *Journal of Agricultural Science*, 148, 329–339.
  - Yang, R.-C., Blade, S. F., Crossa, J., Stanton, D., & Bandara, M. S. (2005). Identifying Isoyield Environments for Field Pea Production. *Crop Sci*, 106–113.
  - Yin, X., & van Laar, H. H. (2005). *Crop Systems Dynamics*. The Netherlands: Wageningen Academic Publishers.
  - Yoshida, H., & Horie, T. (2010). A model for simulating plant N accumulation, growth and yield of diverse rice genotypes grown under different soil and climatic conditions. *Field Crops Research*.

- Yoshida, S. (1981). *Fundamentals of rice crop science*. (International Rice Research Institute, Ed.). Manila (Philippines).
- Yoshida, S., & Parao, F. (1976). Climatic influence on yield and yield components of lowland rice in the tropics. *Climate and Rice*.
- Zhang, J., Feng, L., Zou, H., & Liu, D. L. (2015). Using ORYZA2000 to model cold rice yield response to climate change in the Heilongjiang province, China. *The Crop Journal*, 3(4), 317–327.
- Zhang, Y., Tang, Q., Peng, S., Zou, Y., Chen, S., Shi, W., ... Laza, M. R. C. (2013). Effects of high night temperature on yield and agronomic traits of irrigated rice under field chamber system condition. *Australian Journal of Crop Science*, 7(1), 7–13.
- Zhou, B., Vogt, R. D., Lu, X., Xu, C., Zhu, L., Shao, X., ... Xing, M. (2015). Relative Importance Analysis of a Refined Multi-parameter Phosphorus Index Employed in a Strongly Agriculturally Influenced Watershed.
- Zorrilla, G., Martínez, C., Berrío, L., Corredor, E., Carmona, L., & Pulver, E. (2012). Chapter 10: Improving Rice Production Systems in Latin America and the Caribbean.