

# Incorporating dynamic gene-based process modules into a crop simulation model



Fabio A. A. Oliveira<sup>1</sup>, James W. Jones<sup>1</sup>, Carlos. E. Vallejos<sup>2</sup>, Melanie J. Correll<sup>1</sup>, José M. C. Fernandes<sup>3</sup>, Willingthon Pavan<sup>4</sup>, and Gerrit Hoogenboom<sup>1,5</sup>



<sup>1</sup>Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL, USA - <sup>2</sup>Horticultural Sciences Department, University of Florida, Gainesville, FL, USA - <sup>3</sup>EMBRAPA Trigo, Passo Fundo, RS, Brazil - <sup>4</sup>International Fertilizer Development center, Muscle Shoals, AL, USA - <sup>5</sup>Food Systems Institute, University of Florida, Gainesville, FL, USA

## Introduction

- Scientific advances in understanding plant genes, combined with advances in technologies for rapidly and inexpensively identifying the genetic makeup of plants, have fueled considerable interest in using genetic information to predict plant phenotypes.
- Crop models are tools that can predict plant phenotype from the genotype based on parameters to represent features of specific cultivars, often referred to as “genetic coefficients” (GC).
- In the current crop models, genetic coefficients are estimated through phenotypic observations, not representing the genetic information, which requires new experiments to estimate genetic coefficients when cultivars are released.

## Goals and objectives

- The goal of this study was, therefore, to address the questions of how this type of integration can be done in a comprehensive dynamic crop model and what complications and limitations are likely to occur.
- The objective was to develop and integrate and evaluate a dynamic statistical gene-based module into the CSM-CROPGRO-Drybean model to predict the time of first flower appearance using data obtained from the MET bean studies.

## Materials and Methods

### Genotype population

- The bean MET was conducted to collect the time-to-flowering phenotypes giving rise of to 187 recombinant inbred lines (RILs) population from a cross between the Andean bean cultivar, Calima, and a Mesoamerican cultivar, Jamapa to measure the effects of quantitative trait loci (QTLs), environment (E) and QTL × E interactions (Bhakta et al. (2017).

### Experimental sites

- Three of the five sites are located in the USA: Prosper, North Dakota (ND); Citra, Florida (FL); and Isabela, Puerto Rico (PR), while the other two sites are located in Colombia: Palmira (PA) and Popayan (PO)

### Dynamic mixed linear model

- Vallejos et al. (2021) described procedures used to develop a dynamic mixed linear model to determine the rate of progress towards first flowering.

Table 2. Description of model abbreviations.

Module	Dynamic module description
DMLM	Dynamic mixed linear model developed by Vallejos et al. (2021).
DMLM-DL	Dynamic mixed linear module by Vallejos et al. (2021) using day length calculated with the CSM model
DPLM	Dynamic piecewise linear module (integrated into CSM-CROPGRO-Drybean model)
Full crop model	Full crop model description
DPLM-CSM	DPLM gene-based module integrated into the CSM-CROPGRO-Drybean model using QTL inputs
CSMG	CSM-CROPGRO-Drybean model using genetic specific coefficients (GSPs)

$$FR_{s,g}(t) = \mu + a_1 \cdot (DayL_s(t) - DayL_m) + a_2 \cdot (DayL_s(t) - DayL_m) \cdot QTL_{3,g} + a_3 \cdot (DayL_s(t) - DayL_m) \cdot QTL_{7,g} + a_4 \cdot (DayL_s(t) - DayL_m) \cdot QTL_{12,g} + a_5 \cdot (Srad_s(t) - Srad_m) + a_6 \cdot (Srad_s(t) - Srad_m) \cdot QTL_{12,g} + a_7 \cdot (Tmax_s(t) - Tmax_m) + a_8 \cdot (Tmax_s(t) - Tmax_m) \cdot QTL_{5,g} + a_9 \cdot (Tmin_s(t) - Tmin_m) + a_{10} \cdot (Tmin_s(t) - Tmin_m) \cdot QTL_{2,g} + a_{11} \cdot (Tmin_s(t) - Tmin_m) \cdot QTL_{3,g} + a_{12} \cdot (QTL_{1,g} \cdot QTL_{12,g}) + \sum_{q=1}^{12} \beta_q \cdot (QTL_{q,g}) + \varepsilon_{s,g,t}$$

Equation 1 – Dynamic mixed linear model

$$FRMAX_g = \gamma_g + \sum_{q=1}^{12} \rho_q \cdot (QTL_{q,g})$$

Equation 2 – Maximum rate of development towards first flowering

$$FRD_{s,g}(t) = \min(FR_{s,g}(t), FRMAX_g)$$

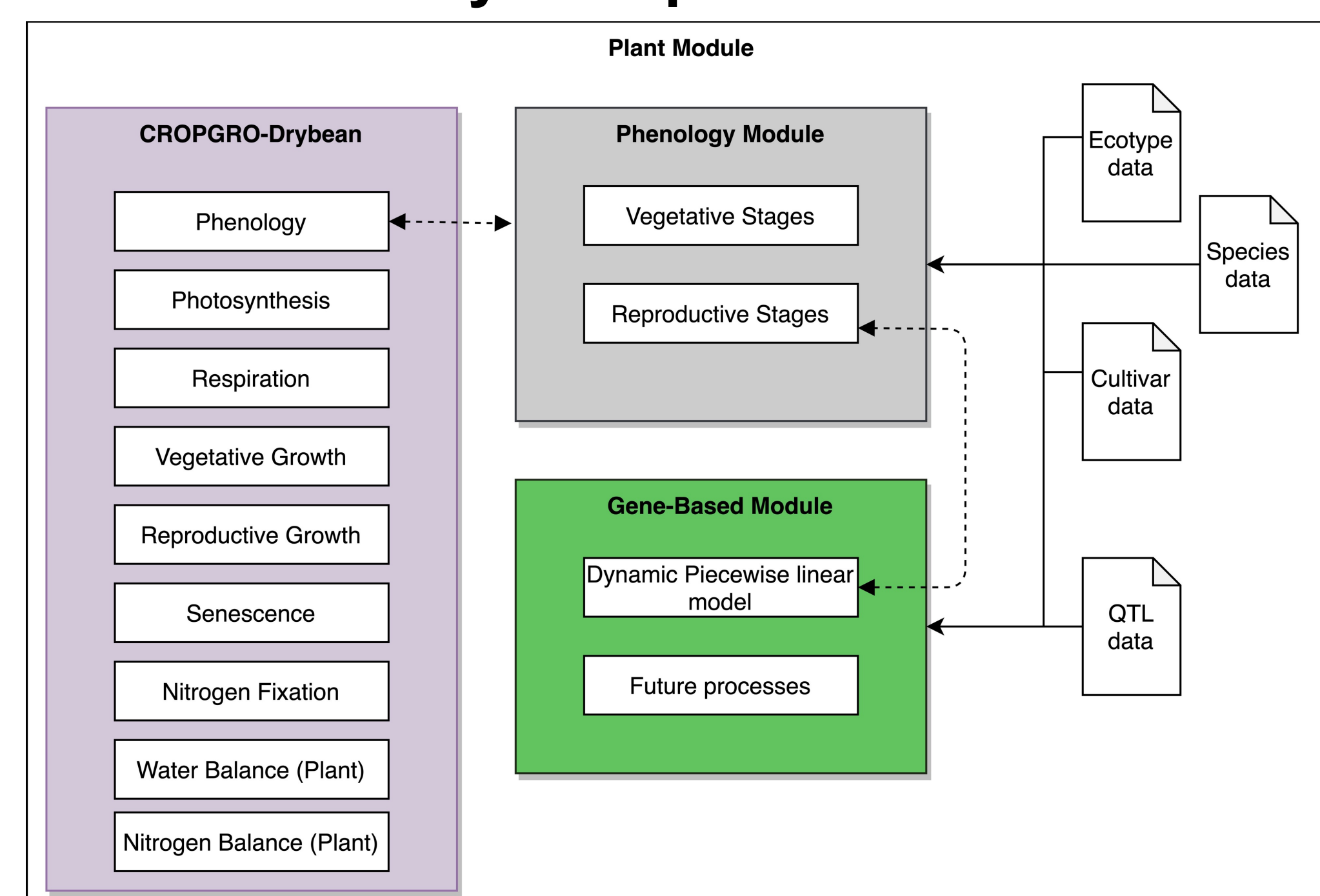
Equation 3 – Limits the maximum rate of flowering set by FRMAX

$$SUMFRD_{s,g}(t) = SUMFRD_{s,g}(t - 1) + FRD_{s,g}(t) \cdot dt$$

Equation 4 – Integrates the flowering rate at time t (in days) in the model

## Results

### CROPGRO-Drybean plant module



- The DPLM module was designed to be flexible, operating in a hybrid mode with the original CSMG using GSPs or QTL input data to simulate the development of first flowering.

Figure 1 – Overview of the DPLM-CSM model developed to integrate the CSM-CROPGRO-Drybean model (CSMG) with the dynamic piecewise linear module (DPLM) using a new gene-based module. The DPLM simulates the first time of first flowering module developed from the dynamic mixed linear model first developed by Vallejos et al. (2021).

### Simulating time to first flower and yield

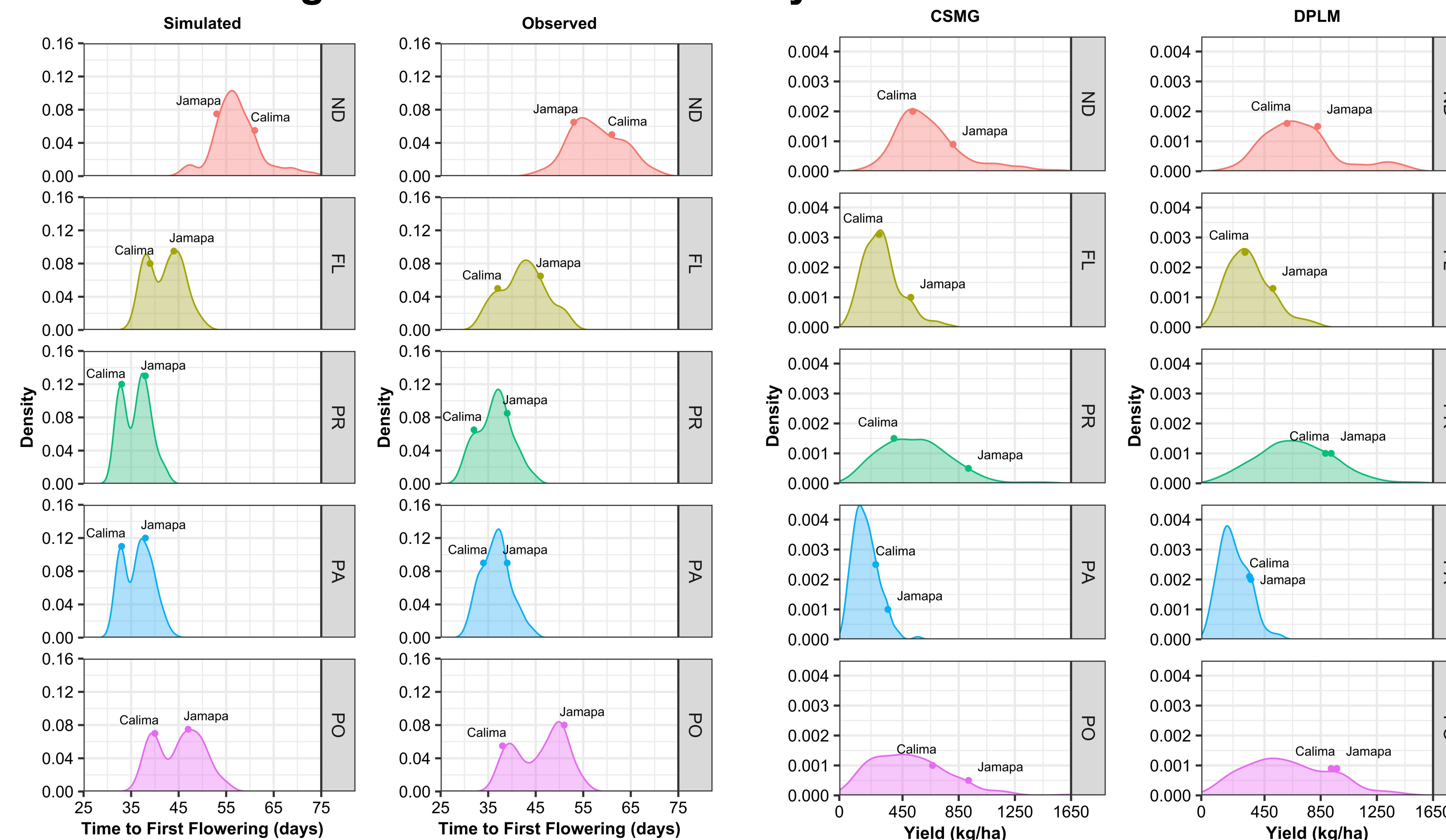


Figure 2 – Density plots of time to first flower in days across five sites. Dynamic piecewise linear module (DPLM) (left panel) and the distribution of observed time to first flower (right panel).

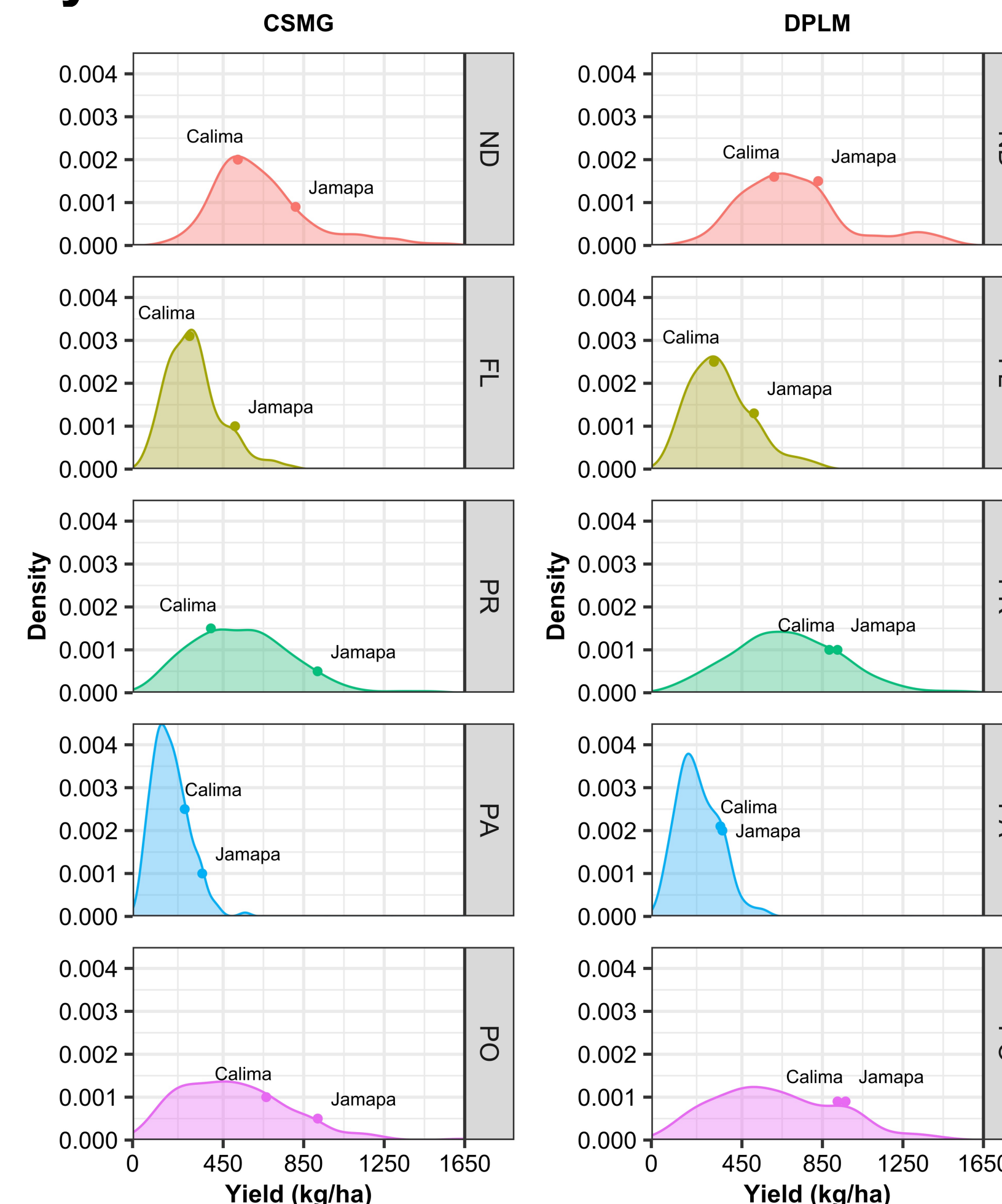


Figure 3 – Density plots for simulated yield using the original CSM-CROPGRO-Drybean model and genetic specific coefficients (CSMG) (left panel) and the dynamic piecewise linear module with CROPGRO (DPLM-CSM) (right panel).

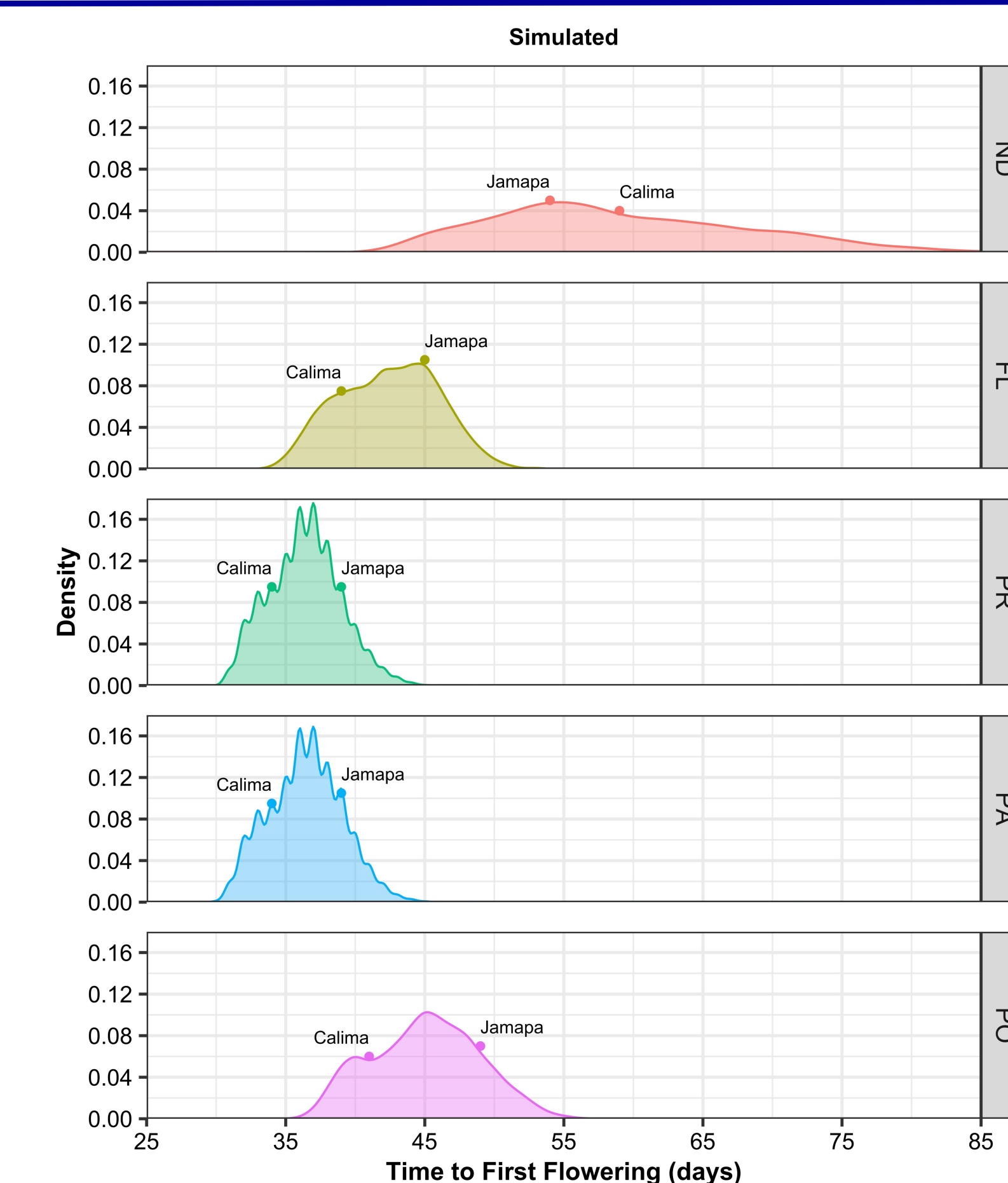


Figure 4 – Density plots for simulated time to first flower (days) across the five sites showing all possible genetic combinations.

## Future work

### Predicting vegetative growth and development

- Number of days to emergence
- Number of days to V1
- Node addition rate
- Maximum number of nodes

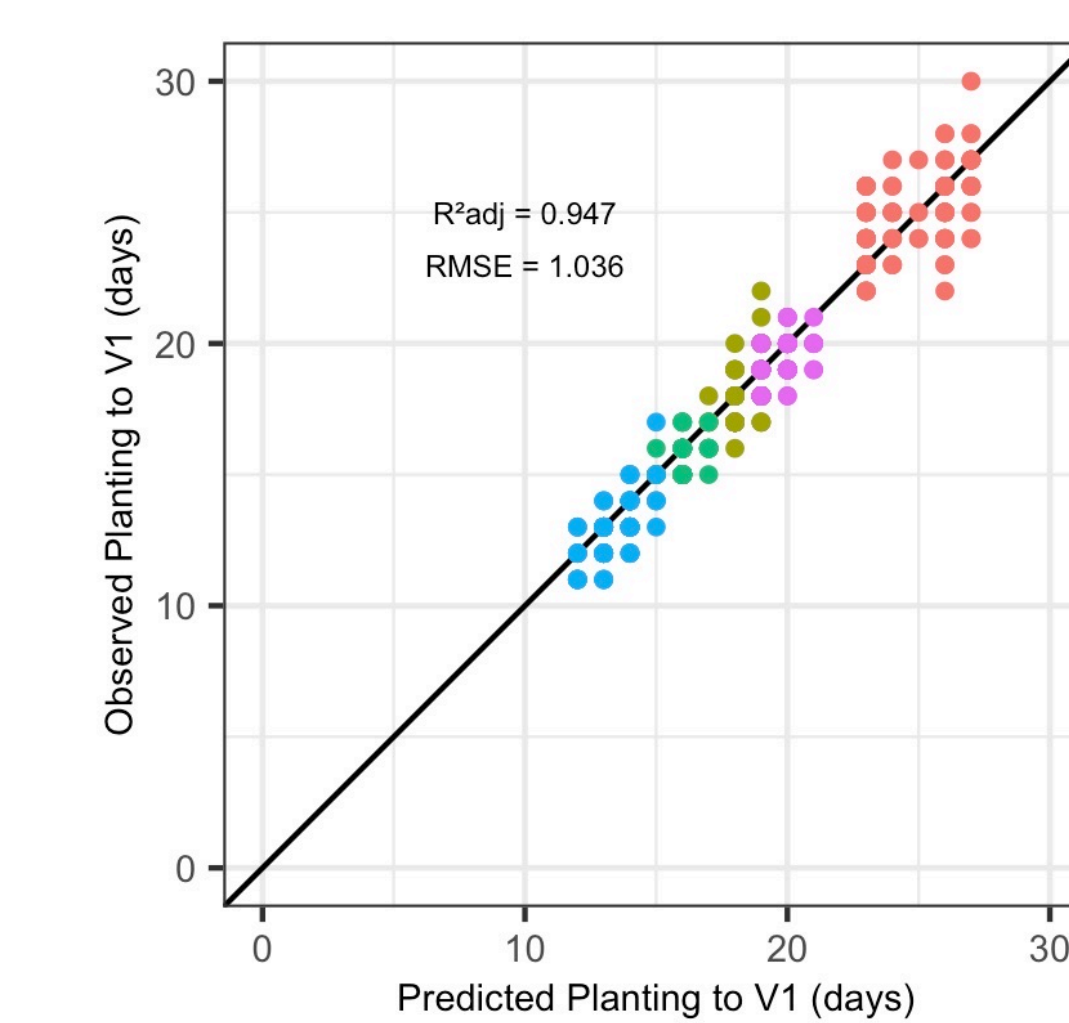


Figure 5 – Node addition rate model to predict planting to V1 stage

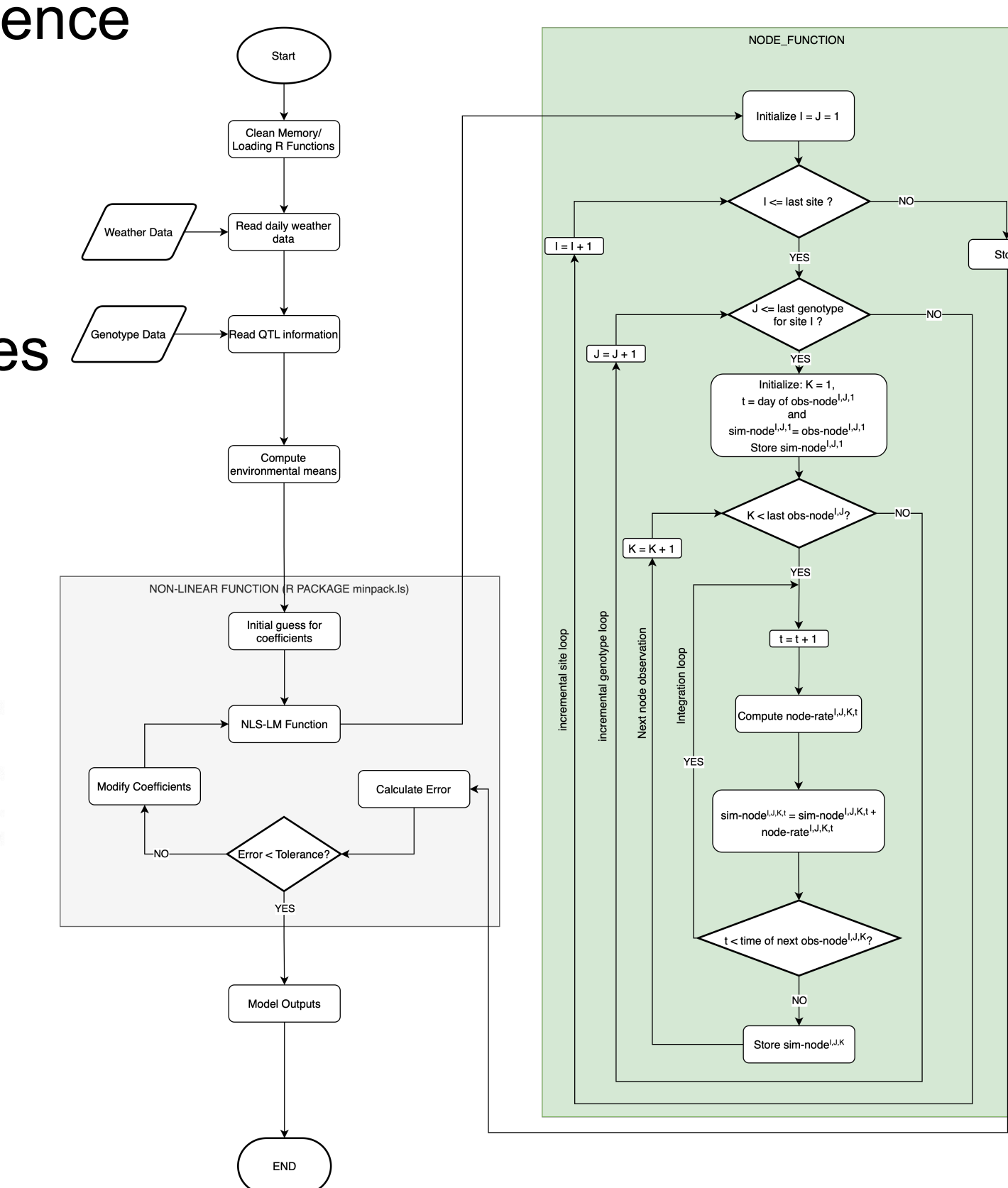


Figure 6 – Flowchart of the non-linear approach for node addition rate model

## Conclusion

- This study showed the potential for integrating a process-oriented gene-based module that only requires genetic input information into an existing comprehensive crop model. The CSM-CROPGRO-Drybean model with the integrated gene-based module was able to predict flowering date, using only QTL and weather information.

## Acknowledgements

