

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Agronomy & Horticulture -- Faculty Publications

Agronomy and Horticulture Department

2016

Estimating yield potential in temperate high-yielding, direct-seeded US rice production systems

Matthew B. Espe

University of California, Davis, mespe@ucdavis.edu

Haishun Yang

University of Nebraska-Lincoln, hyang2@unl.edu

Kenneth Cassman

University of Nebraska-Lincoln, kcassman1@unl.edu

Nicolas Guilpart

University of Nebraska-Lincoln

Hussain Sharifi

University of California, Davis

See next page for additional authors

Follow this and additional works at: <https://digitalcommons.unl.edu/agronomyfacpub>



Part of the [Agricultural Science Commons](#), [Agriculture Commons](#), [Agronomy and Crop Sciences Commons](#), [Botany Commons](#), [Horticulture Commons](#), [Other Plant Sciences Commons](#), and the [Plant Biology Commons](#)

Espe, Matthew B.; Yang, Haishun; Cassman, Kenneth; Guilpart, Nicolas; Sharifi, Hussain; and Linquist, Bruce A., "Estimating yield potential in temperate high-yielding, direct-seeded US rice production systems" (2016). *Agronomy & Horticulture -- Faculty Publications*. 975.

<https://digitalcommons.unl.edu/agronomyfacpub/975>

This Article is brought to you for free and open access by the Agronomy and Horticulture Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Agronomy & Horticulture -- Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

Matthew B. Espe, Haishun Yang, Kenneth Cassman, Nicolas Guilpart, Hussain Sharifi, and Bruce A. Linquist



Estimating yield potential in temperate high-yielding, direct-seeded US rice production systems



Matthew B. Espe^{a,*}, Haishun Yang^b, Kenneth G. Cassman^b, Nicolas Guilpart^b, Hussain Sharifi^a, Bruce A. Linquist^a

^a Department of Plant Sciences, University of California—Davis, One Shields Ave., Davis, CA 95616, USA

^b Department of Agronomy and Horticulture, University of Nebraska—Lincoln, 202 Keim Hall, Lincoln, NE 68583, USA

ARTICLE INFO

Article history:

Received 23 January 2016

Received in revised form 1 April 2016

Accepted 3 April 2016

Available online 6 May 2016

Keywords:

Crop model

Rice

Yield potential

Oryza(v3)/Oryza2000

ABSTRACT

Accurate estimation of a crop's yield potential (Y_p) is critical to addressing long-term food security via identification of the exploitable yield gap. Due to lack of field data, efforts to quantify crop yield potential typically rely on crop models. Using the ORYZA rice crop model, we sought to estimate Y_p of irrigated rice for two widely used rice varieties (M-206 and CXL745) in three major US rice-producing regions that together represent some of the highest yielding rice regions of the world. Three major issues with the crop model had to be addressed to achieve acceptable simulation of Y_p ; first, the model simulated leaf area index (LAI) and biomass agreed poorly for all direct-seeded systems using default settings; second, cold-induced sterility and associated yield losses were poorly simulated for environments with a large diurnal temperature variation; lastly, simulated Y_p was sensitive to the specified definition of physiological maturity. Except for the simulation of cold-induced sterility, all issues could be remedied within the existing model structure. In contrast, simulation of cold-induced sterility posed a continuing challenge to accurate simulation—one that will likely require changes to ORYZA's formulation. Estimates of Y_p from the modified model were validated against large multi-year data sets of experimental yields covering the majority of US rice production areas. Validation showed the adjusted model simulated Y_p well, with most top yields falling within 85% of Y_p for both varieties (77% and 78% observed yields within 15% of Y_p for CXL745 and M-206 respectively). Maximum estimated Y_p was 14.3 (range of 8.2–14.5) and 14.5 (range of 8.7–15.3) $t\ ha^{-1}$ for the Southern US and CA, respectively.

© 2016 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

By definition, yield potential (Y_p) is the yield obtained when an adapted crop variety is grown under optimal conditions without limitations from water supply, nutrients, weeds, insect pests, or disease (Evans and Fischer, 1993). Under these conditions, crop yield is limited only by solar radiation and temperature during the growing season. In trying to produce crops at Y_p levels, however, it is difficult to know whether all biophysical limitations on growth have been eliminated. Hence, it is a challenge to identify field studies in which yields approach the yield potential ceiling. Yields from high-yielding field studies managed explicitly to eliminate all biophysical stresses, winning yields from sanctioned yield contests,

and the top yielding fields amongst a population of farmers have been proposed as proxies for estimating Y_p (Duvick and Cassman, 1999; Lobell et al., 2009; van Ittersum et al., 2013). Previous literature has suggested that few farmers are ever able to achieve Y_p , with most well managed fields achieving between 70–85% of Y_p (Cassman, 1999; Cassman et al., 2003; Fischer, 2015). The difference between actual yields and 85% of Y_p has been termed the exploitable yield gap (van Ittersum et al., 2013; Connor et al., 2011).

Physiological crop models have been suggested as a method to estimate Y_p while avoiding some of the above challenges with estimating Y_p from field studies. Broadly, physiological crop models are seeing increased use for a wide variety of applications, including predicting climate change impacts (Wang et al., 2014; Angulo et al., 2012), assessing effects and implications of management and policy changes (Amiri et al., 2014; Artacho et al., 2011), as well as for yield gap assessments (van Ittersum et al., 2013; Van Wart et al., 2013).

* Corresponding author.

E-mail address: mespe@ucdavis.edu (M.B. Espe).

This increased use is largely due to the ability of these models to conduct broad quantitative assessments which integrate genotype, environment, and management effects that are otherwise difficult to capture (Loomis et al., 1979).

One such model is ORYZA (2000/v3) (hereafter ORYZA), a physiological rice model which was originally developed to simulate rice yield potential in tropical, transplanted rice systems (Bouman et al., 2001). The current version represents the third major version which includes simulation of water and nutrient-limited situations. ORYZA has been widely used in Asia, where it has been calibrated and validated for many common rice varieties (IRRI; <https://sites.google.com/a/irri.org/oryza2000/calibration-and-validation/calibrated-varieties>). Additionally, ORYZA has been used in more diverse environments, including rice production systems in Africa (van Oort et al., 2015), South America (Artacho et al., 2011), Australia (Gaydon et al., 2012), and Europe (Casanova et al., 2000). However, to our knowledge, ORYZA has not been calibrated or validated for rice production systems in the US, which are temperate direct-seeded systems that grow both tropical and temperate *japonica* rice varieties.

US rice production systems achieve some of the highest rice yields in the world (FAOSTAT, 2015), and the US is currently the world's fourth largest exporter of rice, making up over ten percent of annual global rice trade volume (USDA Economic Research Service, 2015). Due to population growth and increasing ethnic diversity, the US's global role as an exporter will be threatened unless rice production can be increased substantially. Given the yield plateaus seen in many high intensity rice production systems (Grassini et al., 2013), it is essential to be able to estimate Yp to answer whether this increase in production will require additional irrigated land, or whether the increased production can come from higher yields on existing land. Yet, to date, there has been no concerted effort to estimate US rice yield potential, partially because there are currently few well-validated models which can do so. Therefore, there is substantial interest in calibrating ORYZA for US production systems to evaluate and quantify Yp.

US rice production can be split into three ecological zones, the Upper Sacramento Valley (CA), the Gulf Coast (TX, LA, and parts of MS), and the Mississippi River Valley (AR, MO, and parts of MS) (Livezey and Foreman, 2004). Although each of these regions has unique climate characteristics, the climate in the two Southern regions (generally humid with a small range of diurnal temperature variation) is more similar to the environment for which ORYZA was originally developed (the Philippines) compared to the climate in CA (generally arid/semi-arid with a large range of diurnal temperature variation). Diurnal temperature variation in CA rice growing areas is driven by cool night winds originating from the San Francisco Bay (termed the "Delta breeze"). Furthermore, CA growers use primarily medium-grain temperate *japonica* rice varieties, whereas growers in the Southern regions primarily use long-grain tropical *japonica* and hybrid rice varieties. By comparison, most varieties that ORYZA has been calibrated for are long-grain *indica* and hybrid types.

In the course of calibration and validation of ORYZA for simulation of Yp in US rice systems, we encountered three challenges with model performance in terms of agreement between measured and simulated Yp: (1) LAI simulations in direct-seeded rice exhibited systematic error correlated with planting density, (2) poor agreement at locations where cool temperatures can cause reduction of fertile spikelets, and (3) high sensitivity of Yp to the determination of physiological maturity. The objective of this study is to investigate via sensitivity analyses of model responses the causes and potential solutions to these three issues. We then assess the suitability of ORYZA for modeling US direct-seeded, temperate rice production systems using large, multi-year multi-site data sets of yields from well managed research plots.

2. Methods

2.1. Weather data

Following protocols established for the Global Yield Gap Atlas (GYGA; www.yieldgap.org), weather stations were selected to obtain the greatest coverage of rice production area within a 100 km radius of selected stations (van Bussel et al., 2015) (Figs. 1 and 2). We selected weather stations first from COOP stations (i.e., stations operated by state agricultural agencies or grower co-operative networks; e.g. agricultural research stations) where available. In cases where such stations were not available, we selected stations from the National Oceanic and Atmospheric Administration National Centers for Environmental Information database (NOAA-NCEI; <https://www.ncei.noaa.gov>). Local agronomists were consulted during the selection stage to ensure chosen weather stations were representative of rice production areas in each region. Locations and elevations of selected weather stations are provided in supplementary information (Table S1). For consistency, incident solar radiation data from the Prediction of Worldwide Energy Resource data set from the US National Aeronautics and Space Administration (NASA-POWER; <http://power.larc.nasa.gov>) were used for all simulations since incident solar radiation was not collected for most stations. Previous work has shown that these data are suitable for Yp modeling purposes (Van Wart et al., 2013, 2015).

2.2. Crop data

Two rice varieties, M-206 and Clearfield XL745 (CXL745) were chosen to represent the two widely-planted rice types grown in CA (M-206) and the Southern US (CXL745). M-206 is a widely planted medium grain *japonica* rice currently grown on approximately 50% of CA rice area (CA Cooperative Rice Research Foundation, 2014). CXL745 is a non-transgenic, herbicide-resistant hybrid rice variety widely planted in most Southern states (the most widely planted rice variety in the USA) (RiceTec, 2015). US rice production systems are direct seeded; CA rice is primarily water-seeded (i.e.—pre-soaked seed is applied to flooded fields via airplane), whereas Southern US rice is primarily drill-seeded. Although data was not collect on site-specific management, management was assumed to follow recommended best practices at each site. Likewise, we assume both varieties are planted at recommended rates; 350 seeds m⁻² for M-206 (Hill, 2013) and 50 seeds m⁻² for CXL745 (Runsick and Wilson, 2009).

For each variety, calibration data were collected that included observed dates of key growth stages; emergence, panicle initiation, 50% heading, and R7 (defined as when the first grain on the panicle transitions from green to yellow color; Counce et al., 2000). For M-206, the calibration data set included plot-level yields from nine sites over two years. For CXL745 the calibration data set included three or four plots per year at a single location in AR over three years. All experimental plots were well managed with the goal of avoiding yield loss from nutrient deficiencies, weeds, insect pests and disease.

For this project, physiological maturity was defined as 225 heat units (approximately 12–17 d) after observed R7 for M-206 and 300 heat units (approximately 15–20 d) for CXL745 (Counce et al., 2015). Heat units were determined using ORYZA's hourly approximation technique (see Bouman et al., 2001 for further detail). Since R7 typically occurs roughly 15 d following 50% heading, this definition of physiological maturity coincides with the general guideline that physiological maturity falls 35–40 d after 50% heading for most rice varieties (IRRI, 2013). Sensitivity of simulated yields to this definition of physiological maturity is addressed in a following section. For the southern-most CA site, R7 was not observed and the date

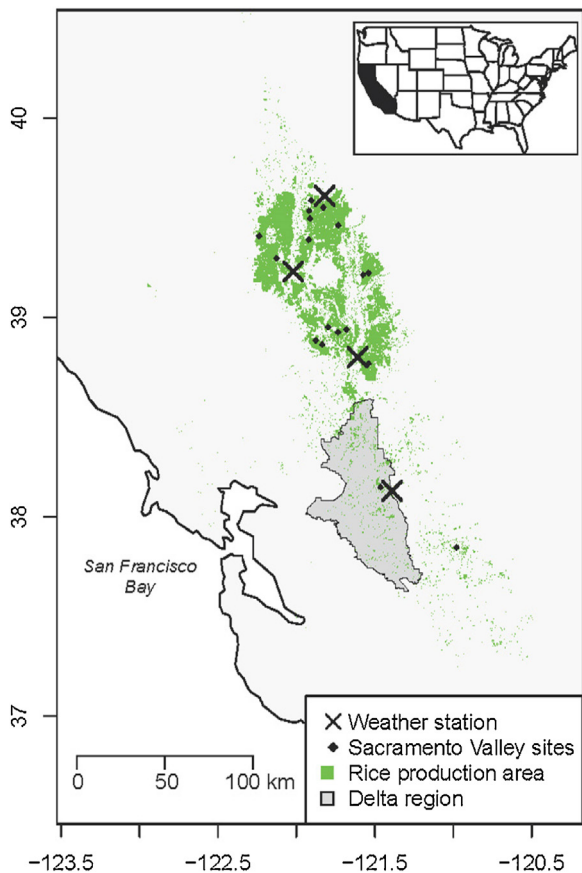


Fig. 1. Locations of weather stations and validation data collection sites for the Sacramento Valley (CA) used for the calibration and validation of ORYZA for rice variety M-206. Cool night-time winds originating from the San Francisco Bay create a temperature gradient from the Delta region (coolest) to the Upper Sacramento Valley (warmer).

of R7 was estimated based on the relationship between R7 for this site and the Northern sites in 2015.

To validate the model, we used a larger data set consisting of experimental plots (M-206 in CA; Fig. 1) or strip-trials (CXL745 in the Southern US; Fig. 2). Both CA and Southern US validation data were from variety testing trials conducted in farmers' fields and under commercial management. For CA, this data set included plot-level data from the Statewide Variety Trials over 15 years (1999–2014). Each year, M-206 was planted at six to eight locations across the Upper Sacramento Valley (Fig. 1). At each location, plots were 3 m × 6 m replicated three or four times in a completely randomized design. Replicate plot yields ($n = 533$) were pooled by trial prior to comparison to simulated Y_p ($n = 115$). Since CA varieties are pre-soaked prior to seeding, emergence date was assumed to be the day after sowing. For the Southern US, the validation data set consisted of strip trials conducted by RiceTec (www.ricetec.com; Houston, TX) from 2007 to 2014 ($n = 223$) across locations in the Southern US (AR, MO, MS, LA, and TX) (Fig. 2). Observed emergence date was recorded for all Southern locations. These data were the mean yield from two replicate 0.4 ha long strips across a field.

To compare simulated Y_p to cases in the validation data set most likely to be non-limited, we split the data by year and geographic region (Sacramento Valley, Gulf Coast, and Mississippi Valley) and the highest yielding 30% of observations in each subset was compared to simulated Y_p . This was deemed necessary because the validation data was collected from farmers' fields and not explicitly managed to avoid nutrient or pest stresses. The threshold of 30% was used for two reasons: (1) to avoid sites where crop and

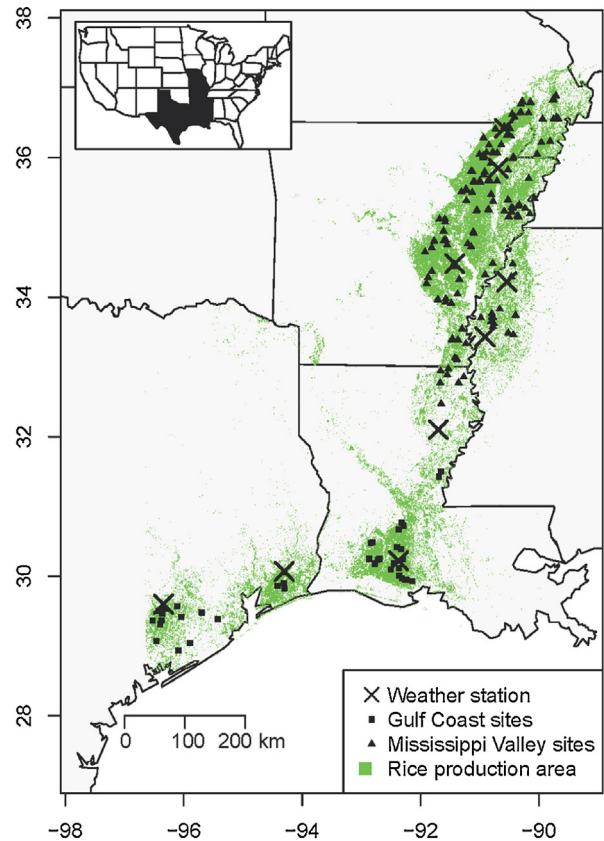


Fig. 2. Locations of weather stations and validation data collection sites for the Gulf Coast and Mississippi Valley (Southern US) rice production areas used for the calibration and validation of ORYZA for rice variety CXL745.

soil management may not have been optimal, and (2) to ensure at least three observations per year per region. All observed yields were corrected to 14% grain moisture.

2.3. Calibration

Crop and soil management at test sites used for model calibration were based on recommended “best management” practices and, in the case of on-farm trials, on farmer experience. Such management regimes seek to maximize profit by achieving highest cost-effective yield levels and are not designed to eliminate all biophysical stresses. Therefore, it was expected that a well-calibrated model would predict Y_p values greater than observed yields in most cases, consistent with work on other crops attempting to quantify Y_p (Cassman, 1999; Lobell et al., 2009). On the other hand, due to sampling variability and measurement error in both yield and weather data, it can be expected that simulated Y_p will be less than observed yields at some sites, but these cases should be a minority in a model that simulates Y_p well. Unlike many calibration efforts where the goal is to minimize the RMSE or absolute error, we expected the calibrated model would not perform well by these measures. In effect, the model was intentionally “biased” towards over-prediction. In this case the “bias” represents the real-world difficulty to achieve Y_p . For example, irrigated maize producers in central Nebraska consistently achieve yields that are 85% of Y_p (Grassini et al., 2011). Hence, in calibrating the model, we used the threshold of no more than 10% of observed yields could be larger than the associated simulated Y_p . Ten percent is the expectation when mean observed yields are 85% of Y_p with a coefficient of variation of 15%.

Previous work calibrating ORYZA for a range of environments identified phenological development rates (Table 2) as the most sensitive factors governing model performance (van Oort et al., 2011, 2015). The base crop file used for both calibrations was the “standard crop” file included with ORYZA(v3). Two other calibrated crop files of medium-grain varieties (“Nipponbare” and “Takanari”) were also used as starting points for calibration for M-206 (also a medium-grain variety), but resulted in poorer predictions compared to the standard crop file. Development parameters were first tuned using the DRATES(v2) program included with ORYZA(v3). DRATES(v2) calculates “optimal” values for each of the four development rate parameters given the calibration dataset.

After calibration via DRATES(v2), more than 10% of observed yields were greater than their associated simulated Y_p and further calibration was required. First, we verified that the model simulated the timing of flowering and heading accurately. Development rates for panicle initiation and flowering were further adjusted to decrease the absolute prediction error of each of these events against the calibration set. Then the final development rate for grain filling was adjusted such that: (1) simulated yields were greater than observed yields in 90% or greater cases, (2) the timing of panicle initiation and 50% heading were well predicted in each case (within 4 d of observed), and (3) simulated heat units from planting to physiological maturity fell within 50 heat units of our observed physiological maturity (i.e. $R7 + 225$ GDDs) for each variety.

2.3.1. Relative growth rate of leaves

Using default settings, the model simulated harvest index outside of the expected 50–55% (typical of modern high yielding cultivars in experimental plots; Hill, 2013). Harvest index simulations were linked to simulated LAI, which was only weakly responsive to changes in development rates (data not presented). ORYZA uses planting density only at initialization of the model, where the LAI at time step 1 is calculated as the product of the state variable “initial LAI” (LAPE) and the user provided plant density. After this point, LAI is part of a feedback loop, where (1) increased LAI leads to greater carbon assimilation, (2) greater carbon assimilation increases biomass, including leaves, which in turn (3) increases LAI at the next time step (Bouman et al., 2001). The increase in biomass and LAI is moderated by the increasing respiration demands for assimilate by biomass and diminishing assimilate gains from additional LAI. However, under high planting densities, LAI and biomass increase too quickly due to the larger starting LAI. This leads to greater simulated biomass and an associated higher maintenance demand for assimilate before simulated grain formation in the model. Therefore, planting density only enters into the model once, but can lead to larger biomass, lower net-assimilate available for grain formation, and hence a lower than expected harvest index. To adjust LAI and thereby harvest index values, the maximum relative growth rate of leaves (RGRLMX) was calibrated for each variety. Changes were made iteratively in 1.0×10^{-4} steps until harvest index values for both varieties fell within the expected range of 50–55%. Unfortunately, data was not collected that allowed independent validation of the RGRLMX calibrations.

2.3.2. Additional calibration (M-206)

Recent work has shown that in arid and semi-arid environments such as CA, ORYZA’s phenology simulation can have systematic errors that are correlated with mean temperatures from emergence to flowering (van Oort et al., 2011, 2015). To mitigate this potential error, we followed van Oort et al. (2011) and adjusted the cardinal temperatures (the temperature thresholds used to calculate heat units) to minimize both the RMSE of predictions from emergence to 50% flowering and the slope of this correlation for the calibration of M-206. The base (TBD), optimum (TOD), and maximum (TMD) temperatures for phenological development were calibrated using the

‘pheno_opt_rice3’ program and the “bilinear1” routine. Two search patterns were used: (1) large steps with TMD up to 999 °C, and then (2) small steps to narrow in on optimal values. ORYZA was then calibrated as above using the values for TBD, TOD, and TMD that resulted in the lowest correlations between temperature and phenology prediction error from emergence to flowering. This calibration was then compared to the default cardinal temperatures.

In CA, a temperature gradient exists running north to south that is driven by the aforementioned “Delta breeze.” Historically, rice production in the Delta areas of CA (Fig. 1) has experienced lower yields due to cold-induced spikelet sterility. Parameters controlling cold-induced sterility were adjusted for M-206 and simulated yields compared to observed yields from the sites most likely to be cold-affected to produce a ‘cold-calibration’ data set (CAL-cold). In ORYZA, cold-induced sterility is controlled by adjustments to the threshold (COLDREP, °C T_{mean}), which is the temperature threshold at which cold-induced spikelet sterility occurs. ORYZA’s cold induced sterility routine accumulates cooling degree-days during the critical time just after simulated panicle initiation until after simulated 50% heading and then applies an adjustment to grain number at the end of the accumulation period based on the total cooling degree-days. Adjustments to the threshold were made incrementally in 0.1 °C steps, checking at each step that simulated yields for the non-cold calibration data set were unaffected by the changes and the absolute yield prediction error was reduced for the CAL-cold data set.

2.4. Sensitivity analyses

During calibration, several anomalies in the simulations prompted investigation of model performance using sensitivity analyses. These anomalies fell in two general categories: (1) continued poor agreement between observed and simulated yields at sites where weather data indicated low temperatures below the COLDREP threshold and (2) sensitivity to the designation of physiological maturity.

2.4.1. Cold induced yield losses

To investigate how the model simulated cold stress in responses to changes in the COLDREP threshold (independent of calibration), we created a series of altered weather data with adjusted T_{min} between panicle initiation and 50% heading. For this, we used an artificially constructed data file consisting of the 15-year average daily values for the Southern-most site in CA, which typically experiences cold induced yield losses. Deviations ranging from -5 to $+5$ °C were added to the daily observed T_{min} during the period from panicle initiation to 50% heading (as predicted by the model using the unaltered weather file). Then, five different values of the cold-induced sterility threshold (COLDREP) from the default 21–29 °C were used to simulate spikelet sterility and yield. All other crop parameters and input data were kept constant. This analysis provided an illustration of model performance over a range of conditions, some more extreme than might be observed. By doing this, model irregularities can be observed that might not be seen when simulating more moderate conditions.

2.4.2. Physiological maturity

In calibration of crop development rates in ORYZA, the user specifies the dates of panicle initiation, 50% heading, and physiological maturity (i.e. the time at which there is no further increase in grain yield). These dates define the duration of vegetative and reproductive stages, and accurate simulation requires robust assessment of each these events. Of the three, 50% heading is the most clearly defined and easiest to observe. Given an accurate date of 50% heading, panicle initiation can be estimated and ultimately has relatively small impact on the model results. Physiological maturity, on the

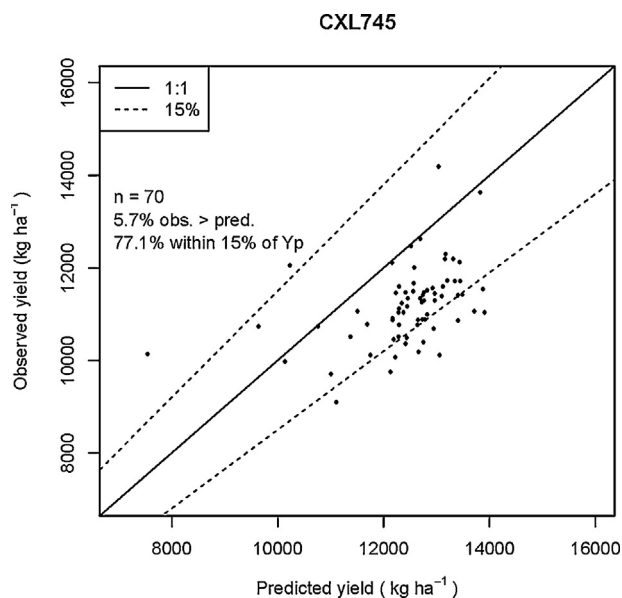


Fig. 3. Simulation results for the validation set for ORYZA calibrated for rice variety CXL745. Only the top performing sites are shown (top 30%). Each point represents the average yield from two replicate strip trials conducted by RiceTec from 2007 to 2014. Out of 70 observations in the validation data set, 77.1% were within 15% of yield potential (Y_p), with only 5.7% of observed values being greater than estimated yield potential. For this project, it was expected that most observed values should fall within 15% of estimated Y_p .

other hand, is difficult to estimate accurately from other events but has a large impact on simulated yields. Physiological maturity marks the end point of the simulations and is most accurately determined by daily measurements of grain carbohydrate accumulation. Most researchers use a proxy measurement which is easier to observe but may not reflect true physiological maturity.

Amongst the many different proxy indicators of physiological maturity used in the literature are: (1) grain moisture thresholds (e.g., 28% grain moisture; Espe et al., 2015), (2) number of days since planting or heading (e.g., 35 days following 50% heading; IRRI, 2013), and (3) the firmness of the grain (e.g., IRRI, 2013; Moldenhauer and Slaton, 2015). The definition listed in the ORYZA(v3) training manuals (IRRI) is: “Physiological maturity is visually identified when the grains on the lower portion of secondary and tertiary panicles harden and begin to lose their green color.” However, even this measure is uncertain as the development of grain on secondary and tertiary panicles will depend on tillering density and rate, which is closely linked to planting density (Hill, 2013). Due to this uncertainty, the date designated as physiological maturity could differ by 5–10 d, which gives rise to a large difference in grain filling time and Y_p .

We evaluated the ramifications of this uncertainty by adjusting the development rate for the final crop stage (DVRR), for a single site/year in CA (‘Canal’ 2013). This site was chosen for its high-yield potential and close proximity to a reliable weather station (<10 km). To allow for different grain-filling lengths, DVRR was adjusted from 1.0×10^{-3} to 4.0×10^{-3} . This altered the length of the simulated grain filling time over a range of roughly 40 d.

2.5. Data formatting and analysis

All model output files were analyzed using the R statistics program (v. 3.2, R Core Team, 2015). Likewise, weather files were programmatically constructed from source files or online data sources using R (v. 3.2, R Core Team, 2015) and the RCurl package (v. 1.95, Temple Lang, 2015). Weather files were assessed for outlier values and extreme values (e.g. -40°C T_{\min}), and suspect values

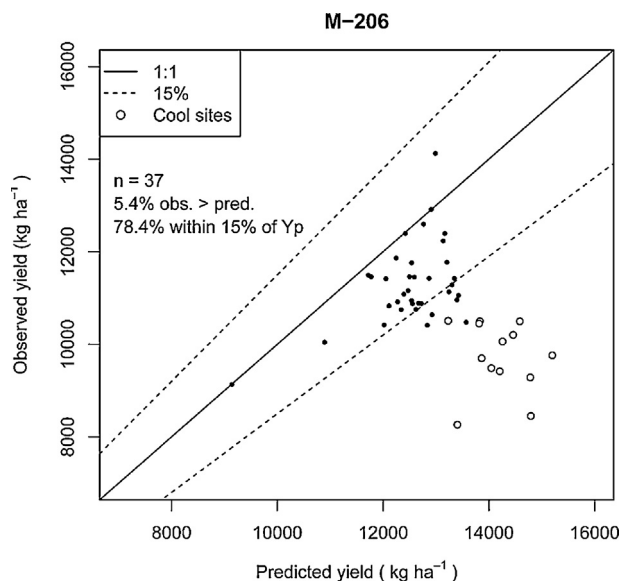


Fig. 4. Simulation results for the validation set for ORYZA calibrated for rice variety M-206. Only the top performing sites within each year are shown (top 30%). Each point represents the mean of three or four experimental plots in the California Statewide Variety Trials from 1999 to 2014. Open points are the sites that regularly experience cool temperature induced sterility. Out of 37 observations in the validation data set, 78.4% were within 15% of yield potential (Y_p), with only 5.4% of observed values being greater than estimated yield potential. For this project, it was expected that most observed values should fall within 15% of estimated Y_p .

and missing values were replaced with values from the NASA-POWER database. In all cases, outlier and missing data consisted of less than 5% of seasonal measurements.

3. Results

3.1. Validation

For CXL745 in the Southern US, the model simulated yields well, with 77% of observed yields falling $\pm 15\%$ around simulated Y_p (Fig. 3). Fewer than 6% of observed yields were greater than simulated. Despite only calibrating phenological parameters, the model captured variations in observed yields quite well, although several simulations were conspicuously under-predicted by the model. Average estimated Y_p was 12.3 t ha^{-1} for the Southern US, with a range of $8.2\text{--}14.5 \text{ t ha}^{-1}$. Observed yields across all sites averaged 9.8 t ha^{-1} , with a range of $4.3\text{--}14.2 \text{ t ha}^{-1}$.

For M-206, however, comparable simulation results were obtained only after calibration of COLDREP (Fig. 4). The number of observed yields greater than simulated Y_p (5.4%) and percentage of observed yields within 15% of Y_p (78%) were similar to CXL745. Although the range of observed yields was relatively narrow due to selecting the top 30% of observed yields by year, the model was able to capture differences between yields at several sites with low Y_p . Adjustments to cardinal temperatures as recommended by van Oort et al. (2011) were able to reduce the correlation between temperature and prediction error, but at the cost of increasing prediction error from emergence to 50% heading (Supplemental Fig. 1). Best model performance was achieved by using default cardinal temperatures, despite some persistent error in phenology prediction (Fig. 5). Therefore, the final calibration for M-206 used the default cardinal temperatures (TBD = 8°C , TOD = 30°C , and TMD = 42°C). The persistent phenology error resulted in over-prediction of days to heading in cases where heading was accelerated and under-prediction in cases where heading was delayed (Fig. 5). Average estimated Y_p for CA was 12.9 t ha^{-1}

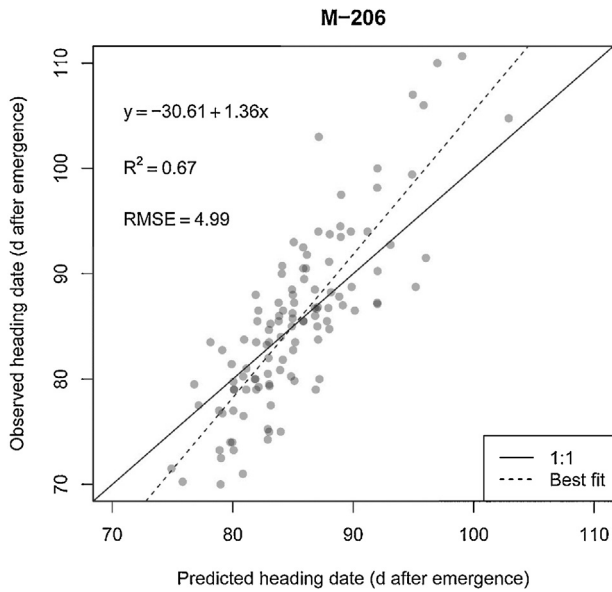


Fig. 5. Observed vs predicted days to 50% heading for rice variety M-206. Cases where heading was delayed or accelerated compared to average (85 d) are less accurately predicted by ORYZA.

with a range of 8.7–15.3 t ha⁻¹ (cool sites included) or 14.3 t ha⁻¹ (cool sites excluded). Observed yields across all sites averaged 10.4 t ha⁻¹, with a range of 7.8–14.1 t ha⁻¹.

3.2. Relative growth rate of leaves

The adjustments to RGRLMX required to bring the simulated harvest index within the range of 50–55% were opposite the planting densities for each variety. For M-206, which is planted at high density (350–470 plants m⁻², UCCE, 2015), the relative growth rate had to be decreased from the default of 0.0085–0.0060 (Table 1). For CXL745, which is planted at low density (40–150 plants m⁻², Runsick and Wilson, 2009), the relative growth rate was increased to 0.0110 (Table 1). (For the response of harvest index to changes in both RGRLMX and planting density, please refer Supplemental Fig. 2).

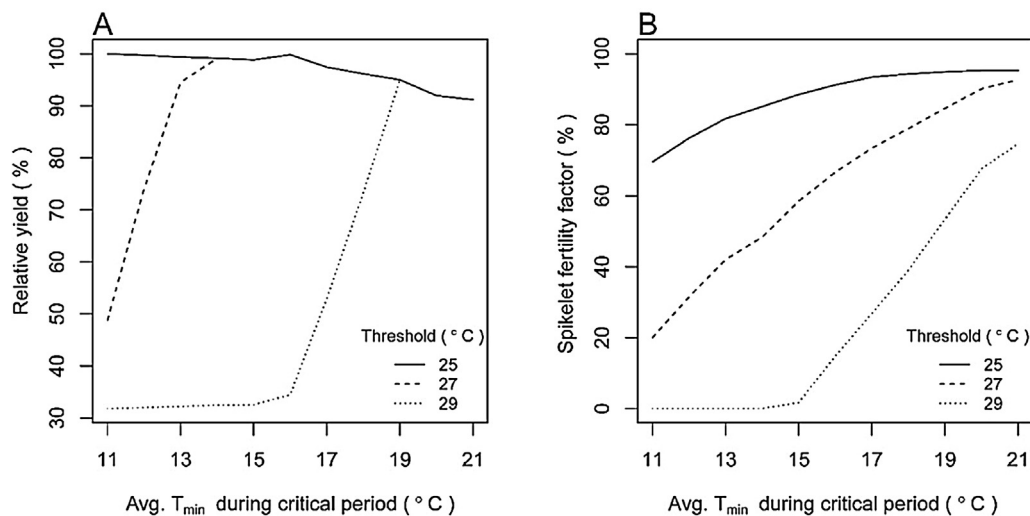


Fig. 6. Sensitivity of ORYZA to changes of cold sterility threshold (i.e., COLDREP parameter) and daily minimum temperature in weather data. The threshold is the T_{mean} below which ORYZA accumulates cold sterility. A higher threshold signifies greater sensitivity to cold. The critical period is defined as the period from just after panicle initiation to 50% heading (crop development stage 0.75–1.2). Simulations of yield and spikelet fertility factor were not different for thresholds between 21–25 °C (not displayed). Due to the fact that ORYZA applies a reduction in spikelet number at development stage 1.2, but yield formation begins at development stage 1.0, ORYZA simulates some yield gain even at complete spikelet sterility. Increasing yields despite increased spikelet sterility suggests that simulated yields are not sink limited and that decreased temperatures are predicted to benefit yield formation, likely due to decreases in respiration.

3.3. Cold responsiveness

Model response to cold temperature was improved with adjustments to COLDREP without influencing simulations of non-cold affected sites, but the influence of cool temperatures on yield was still not fully captured by the model (Fig. 4). Sensitivity analyses revealed simulated yields were only responsive to changes in COLDREP greater than 25 °C (Fig. 6). Despite greater than 60% simulated spikelet fertility factor (Fig. 6B), yields increased as T_{min} decreased (Fig. 6A) when COLDREP equaled 25 °C or lower. The adjustment of -5 °C brought the average T_{min} below temperatures known to induce sterility (12–15 °C) in CA rice production systems (Board and Peterson, 1982; UCCE 2015) for most of the critical period between panicle initiation and 50% heading. Complete spikelet sterility would be expected in this extreme case, yet ORYZA simulates 4 t ha⁻¹ grain yield. This discrepancy is due to cold-induced sterility being applied in the model later than the start of grain accumulation (development stage 1.2 and 1.0, respectively). Thus for the most cold-affected sites in our data, our calibration of ORYZA still predicts a difference between observed yields and simulated Y_p of up to 5 t ha⁻¹ (Fig. 4).

3.4. Physiological maturity

Simulated yields of M-206 at a high-yielding site increased by roughly 230 kg ha⁻¹ for each extra day increase in simulation for maturity dates earlier in the season (25–35 d after 50% heading, Fig. 7), though this relationship tapered off as maturity was pushed later in the season. Our determination of physiological maturity (based on a number of heat units following the observed date of R7) resulted in simulations that met our expectations, with observed yields roughly 85% of simulated Y_p in most cases.

4. Discussion

4.1. Relative growth rate of leaves

Rice has a well-documented plasticity in response to planting density (UCCE, 2015; Connor et al., 2011; Yoshida, 1981). Rice plants can aggressively produce tillers at low planting densities

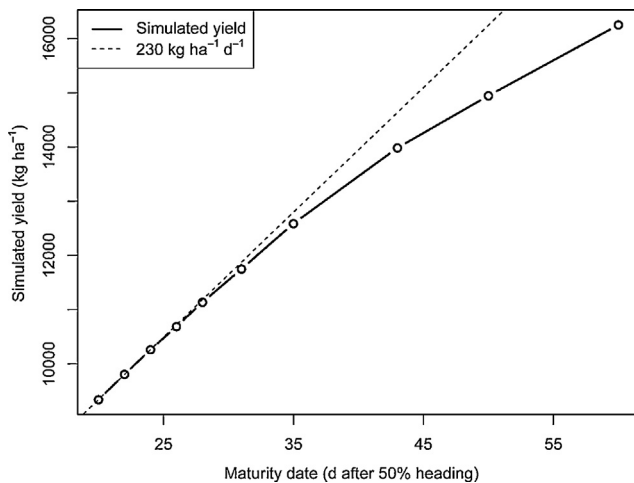


Fig. 7. Simulated yield resulting from extending the date of physiological maturity for rice variety M-206. The yield increase per day extension of the grain filling period is strongest early in the simulation. For M-206, 50% heading typically occurs at 85 d after emergence.

to have equivalent yield components as higher density plantings. Hence it is possible to observe similar grain yields and harvest index values across a wide range of planting densities (Hill, 2013). In this study we had both extremely low and high planting densities, yet similar observed yields, suggesting differences in relative growth rates. Changing the RGRLMX partially captured this plasticity of rice plants. Our values suggest that the hybrid variety (planted at low densities) has the capacity for much faster growth compared to the non-hybrid (planted at high densities) which supports previous reports of increased rates of growth in hybrid compared to non-hybrid rice varieties (Bueno and Lafarge, 2009). It has also been observed that the same rice variety sown at higher densities has lower relative growth rates compared to that same rice variety sown at low densities (San-oh et al., 2004), further supporting the results found here.

However, our solution has some notable disadvantages. First, we did not possess the required data to validate the RGRLMX parameter, therefore the values of RGRLMX for these two varieties should be treated as an untested assumption until more work can properly validate these values. Second, our implementation requires manual adjustment of RGRLMX in response to changes in planting density. A better implementation would make simulated LAI responsive to planting density and thereby would reflect the real-world plasticity of rice plants that is currently missing from the ORYZA model.

It is worth noting that in early calibration attempts, we found it possible to calibrate the model using the default values such that simulated yields approximated the calibration sets, yet the model performed poorly for the validation set in these cases (data not presented). It is possible this issue with LAI would have gone unnoticed had the only output of concern been yield, or had we not had a multi-year, multi-site data set to validate against. Although this specific issue was relatively easy to resolve via adjustment of a single parameter, it highlights one of the pitfalls of applying a crop model to new production systems. Researchers attempting to use such a model in production systems outside its system of origin should be aware of these possibilities. Best practices such as utilizing both a calibration and validation step and inspecting the complete model output rather than just the output of interest are critical.

4.2. Cold responsiveness

ORYZA's subroutine for cold induced sterility during flowering is empirically derived from the accumulation of cooling degree days

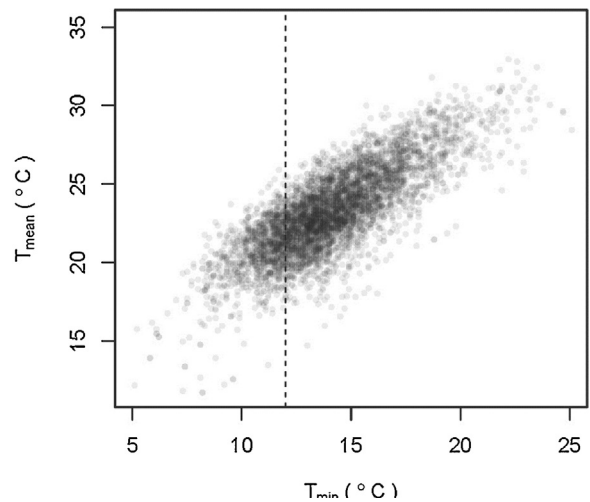


Fig. 8. Comparison of T_{\min} and T_{mean} for CA sites between June 25th and Aug. 28th (the range of simulated panicle initiation to 50% heading in CA rice for all planting dates) from 1999 to 2014. The vertical dashed line (12 °C) is widely regarded as the threshold at which cold induced sterility is experienced by CA rice varieties.

(Bouman et al., 2001). Ultimately, we show here that this subroutine is not adequate, especially for environments such as CA where there are large range of diurnal temperature variation. Cooling degree days are calculated as a function of the number of days with T_{mean} below a certain threshold (21 °C by default). This fails to capture cold induced sterility in environments where the diurnal swing of temperatures is large. For example, a site with T_{max} of 30 °C and T_{min} of 20 °C will have the same cooling degree days as one with T_{max} of 40 °C and T_{min} of 10 °C.

Research supports the importance of T_{min} rather than T_{mean} during the sensitive period between panicle initiation and 50% heading for the determination of cold-induced sterility (Farrell et al., 2006). For CA rice varieties (widely regarded as cold-tolerant) T_{min} of 12–15 °C during this critical period is generally acknowledged as the threshold for cold-induced sterility (Board et al., 1980; Board and Peterson, 1982). In our comparison of T_{min} against T_{mean} for CA sites, it is clear that if rice experiences spikelet sterility below 12 °C T_{min} , this corresponds to a 10 °C range of observed T_{mean} values (Fig. 8). To fully capture every occurrence of cool temperatures at or below 12 °C, a T_{mean} -based threshold would need to be increased to roughly 25 °C. However, at this high threshold many occurrences of T_{min} above 12 °C would also be considered as cool enough to induce sterility, creating over-sensitivity in the model. Thus it is clear that for environments with a large range of diurnal temperature fluctuations, ORYZA cannot hope to capture cold sterility appropriately using the existing T_{mean} -based model structure.

Unfortunately, there are several other sources of error that can also lead to gross inaccuracies in the modeling of cold induced yield reductions. First, since ORYZA's crop model is phenology driven, and since cold induced sterility impacts the crop during a relatively narrow window between panicle initiation and flowering, the accuracy of ORYZA's cold sterility routine is directly tied to the accuracy of the phenology sub-model. Second, sterility is highly dependent on the micro-climate directly surrounding the panicle (Julia and Dingkuhn, 2013) which might not be accurately reflected in the weather data. Third, the model assumes that cold induced sterility can be modeled as an accumulation of cold-stress over the period from panicle initiation, yet the pattern of flowering is non-linear (Yoshida, 1981) and cool nights during certain points of flowering may have greater impact than others. Lastly, ORYZA reduces spikelet number all at once at a relatively late stage (development

stage 1.2), but simulates yield beginning at development stage 1.0. Due to this, ORYZA simulates some yield gain even at complete spikelet sterility (Fig. 6).

Hence there are broadly four scenarios where model simulations of cold induced sterility will be inaccurate: 1) The site experiences a large range of diurnal temperature fluctuation and any threshold based on T_{mean} cannot accurately capture cool temperatures experienced by the crop, 2) phenology is not accurately modeled, so the simulated crop experiences a different set of temperatures than the observed crop, 3) phenology is accurately modeled, but climate data does not reflect the micro-climate experienced by the developing spikelets, and 4) a short period of cool temperatures coincides with peak spikelet sensitivity, which will have little impact on the total accumulated cooling degree days in the model but a large impact on observed sterility. The structure of ORYZA's sub-routine for cold-induced sterility is poorly equipped to counter these sources of error. Taken together, these issues suggest that the sub-routine ORYZA uses to simulate cold effects on sterility requires substantive revision.

4.3. Physiological maturity

We show that simulated yields are highly influenced by adjusting the date of physiological maturity. Yields increased by roughly the theoretical maximum carbohydrate accumulation for a C3 crop under non-limited situations (200–300 kg ha⁻¹; Connor et al., 2011). Although this behavior is not unexpected, it is important to highlight how critical the determination of physiological maturity is for accurate modeling of crop yield performance. As mentioned previously, there are many definitions of physiological maturity currently in use which all imply different dates of physiological maturity. As shown here, in non-limiting environments relatively small differences in physiological maturity can result in large differences in simulated Yp. Users of crop models need to be explicit about how physiological maturity was determined and the sensitivity of their conclusions to this measure. Unfortunately, this is rarely the case (e.g., Zhang and Tao, 2013; Amiri et al., 2014; Artacho et al., 2011). This is especially important in cases where maturity is estimated from inaccurate proxies such as average harvest date, approximate days since planting, etc. It should be noted that there are many determinants of harvest maturity (i.e., when the crop is deemed ready to harvest) that are distinct from the determinants of physiological maturity. Determining harvest maturity often involves factors other than the point at which carbohydrate accumulation ceases (e.g., grain quality, drying, and avoidance of losses during harvest), and therefore indicators of harvest timing often are not suitable for determining physiological maturity. In cases where these inaccurate measures are necessitated, we recommend a sensitivity analysis to help readers understand the implications of error in the estimate of physiological maturity.

4.4. Final calibration and validation

To our knowledge, this is the first large-scale effort to calibrate and validate ORYZA for US rice production systems. Furthermore, it is unique for modeling studies to have a validation data set of the quality and temporal and spatial scope such as ours here. We found that for the Southern US, the model performed well after calibration of only the four development rate parameters and calibration of RGRLMX (Table 1). This is possibly due to similarities between the Southern US and the environment that ORYZA was originally developed under. It is noteworthy that the ORYZA model can perform well outside its original domain with such a simple calibration.

This stands in contrast to the calibration of ORYZA for CA. Additional calibration of cold-sterility thresholds was needed before

Table 1

Calibrated ORYZA parameter values for two varieties representative of US rice varieties.

| Variety | DVRJ | DVRI | DVRP | DVRR | RGRLMX | COLDREP |
|---------|---------|---------|---------|---------|--------|---------|
| Default | 0.00077 | 0.00076 | 0.00078 | 0.00178 | 0.0085 | 21.0 |
| CXL745 | 0.00088 | 0.00076 | 0.00080 | 0.00151 | 0.0110 | 21.0 |
| M-206 | 0.00159 | 0.00076 | 0.00088 | 0.00235 | 0.0060 | 23.3 |

Table 2

ORYZA parameter abbreviations and definitions.

| Abb. | Description |
|---------|---------------------------------------------------|
| DVRI | Development rate for early juvenile phase |
| DVRJ | Development rate for photo-period sensitive phase |
| DVRP | Development rate for flowering phase |
| DVRR | Development rate for grain filling phase |
| RGRLMX | Maximum relative growth rate for leaves |
| TBD | Base temperature for phenological development |
| TOD | Optimum temperature for phenological development |
| TMD | Maximum temperature for phenological development |
| COLDREP | Threshold for cold induced sterility |

the model could capture yield trends as well as in the Southern US (Table 1). Persistent poor simulation of Yp for the most cold-affected site (Fig. 4) suggests model inaccuracies in simulation of cold-induced sterility need to be better addressed for better fit. While it is possible that the use of a modified version of ORYZA with a different cold-sterility routine (e.g.—ORYZA2000v2n13sXX; van Oort et al., 2015) may have been able to simulate CA rice production more accurately, we sought to calibrate and validate the most widely used branch of ORYZA (v. 3). In general, this fragmentation of the model is problematic; ideally improvements to the model structure would be integrated into the main branch when possible. That said, current rice production in the most cold-affected region (the Delta region) of CA is limited to less than 3000 ha (less than 10% of CA rice production area; Fig. 1). Hence this deficiency in the model may have little impact on region-wide estimates of current Yp. However, given that all CA sites have some potential to experience sterility inducing night-time temperatures, the impact may extend beyond just the Delta region in certain years. The development of a more accurate cold-sterility routine would improve our understanding of mechanism leading to cold-induced sterility, and, while beyond the scope of this effort, is needed for more accurate simulation of this system and systems like it.

Since ORYZA relies on phenology to trigger various events throughout the simulation, it is also possible that the persistent phenology error we encountered (Fig. 5) has contributed to the inaccuracy of the model simulations. Contrary to van Oort et al. (2011), calibration of cardinal temperatures did not improve model performance (Supplemental Fig. 1) and therefore default cardinal temperatures were used in our final calibration for both varieties. Similar results are reported and more thoroughly analyzed by Sharifi et al. in CA rice production systems. A potential cause of this discrepancy with van Oort et al. (2011) is the bias-variance tradeoff in predictive models (James et al., 2015). By removing temperature bias for the calibration data set, the model was potentially over-fit which would account for increased variance for the validation data set. Unfortunately, van Oort et al. (2011) did not utilize separate calibration and validation steps, so we are unable to assess whether they would have seen similar results with out-of-sample predictions. Likewise, we lack the required data for the Southern US to investigate if this error is unique to CA or present throughout US rice simulations. Lastly, it is difficult to assess the specific cause of this error in the face of other differences from the system that ORYZA was developed for (e.g., arid to semi-arid climate with large range of diurnal variation, direct-seeded rice at extremely high seeding rates via water-seeding, varieties developed with quality as a pri-

ority, little disease or pest pressure, etc.). Further investigation is needed to determine the cause and scope of this persistent error and the mechanism behind this discrepancy from van Oort et al. (2011).

Estimates of Yp in this study differ substantially from previous estimates. For example, all estimates in this study (Figs. 3 and 4) were much less than those of Sheehy and Mitchell (2015) for subtropical semi-dwarf rice Yp (20.1 t ha⁻¹). However, their estimate assumes a longer growing season (168 d) than is feasible in temperate regions. Therefore, our estimates are likely more applicable to US rice production systems. On the other hand, estimates of Yp from our study are generally higher than estimates of rice Yp based on the maximum average regional yields in similar climates (e.g., Mueller et al., 2012; Foley et al., 2011). Our results suggest, contrary to these previous studies, there may be an exploitable yield gap in these highly intensified rice production systems. Further work is needed to quantify the yield gap for these systems using these revised estimates.

Lastly, while we were able to achieve an acceptable calibration of ORYZA for US rice production, a full calibration of the model using more extensive data on carbohydrate partitioning throughout the season could further improve model performance and provide validation for our values for relative growth rates. However, more extensive calibration would not address the fundamental issues with the CA calibration (phenology error and poor simulation of cold-sterility). Modification of the model structure is required to address these issues.

5. Conclusions

ORYZA can adequately simulate Yp for Southern US environments with straightforward calibration of variety-specific phenological parameters. Simulation of Yp for CA required more extensive calibration with attention to representation of cold tolerance and physiological maturity. More extensive calibration, however, will not address structural deficiencies, such as the modeling of plasticity in tillering or cold-induced sterility. Accurately capturing these complex phenomena will likely require updating several structural components of the model. Despite these issues, we show that ORYZA can be acceptably calibrated and validated for the majority of US rice production environments, both in the Southern US and CA.

Acknowledgements

We would like to thank the CA Rice Research Board, RiceTec, Merle Anders, the UC Davis Dept. of Plant Sciences graduate student fellowship, and the members of the Agro-ecosystems lab for support and assistance.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.fcr.2016.04.003>.

References

- Amiri, E., Rezaei, M., Rezaei, E.E., Bannayan, M., 2014. Evaluation of Ceres-Rice, Aquacrop and Oryza2000 models in simulation of rice yield response to different irrigation and nitrogen management strategies. *J. Plant Nutr.* 37 (11), 1749–1769.
- Angulo, C., Becker, M., Wassmann, R., 2012. Yield gap analysis and assessment of climate-induced yield trends of irrigated rice in selected provinces of the Philippines. *J. Agric. Rural Dev. Trop. Subtrop.* 113 (1), 61–68.
- Artacho, P., Meza, F., Antonio Alcalde, J., 2011. Evaluation of the Oryza2000 rice growth model under nitrogen-limited conditions in an irrigated Mediterranean environment. *Chil. J. Agric. Res.* 71 (1), 23–33.
- Board, J.E., Peterson, M.L., 1982. Rice sterility varies with variety and area. *Calif. Agric.* (January–February), 6–7.
- Board, J.E., Peterson, M.L., Ng, E., 1980. Floret sterility in rice in a cool environment. *Argon. J.* 72, 483–487.
- Bouman, B.A.M., Kropff, M.J., Tuong, T.P., Wopereis, M.C.S., Ten Berge, H.F.M., Van Laar, H.H., 2001. ORYZA2000: modeling lowland rice. International Rice Research Institute, Los Baños, Philippines, and Wageningen University and Research Centre, Wageningen, Netherlands, 235 pp.
- Bueno, C.S., Lafarge, T., 2009. Higher crop performance of rice hybrids than of elite inbreds in the tropics: 1. Hybrids accumulate more biomass during each phenological phase. *Field Crops Res.* 112, 229–237.
- CA Cooperative Rice Research Foundation, 2014. Rice Production of CRRF RES Varieties, 2012–2013, <http://www.crrf.org/crrf.res-v40.028.htm> (accessed July 2015).
- Casanova, D., Goudriaan, J., Bosch, A.D., 2000. Testing the performance of ORYZA1 an explanatory model for rice growth simulation, for Mediterranean conditions. *Eur. J. Agron.* 12, 175–189.
- Cassman, K.G., 1999. Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5952–5959.
- Cassman, K.G., Dobermann, A., Walters, D.T., Yang, H.S., 2003. Meeting cereal demand while protecting natural resources and improving environmental quality. *Ann. Rev. Environ. Resour.* 28, 315–358.
- Connor, D.J., Loomis, R.S., Cassman, K.G., 2011. *Crop Ecology: Productivity and Management in Agricultural Systems*. Cambridge University Press, New York, 562 pp.
- Counce, P.A., Keisling, T.C., Mitchell, A.J., 2000. A uniform objective, and adaptive system for expressing rice development. *Crop Sci.* 40, 436–443.
- Counce, P.A., Siebenmorgen, T.J., Ambardekar, A.A., 2015. Rice reproductive development stage thermal time and calendar day intervals for six US rice cultivars in the Grand Prairie, Arkansas, over 4 years: rice reproductive development stage intervals. *Ann. Appl. Biol.* 167, 262–276.
- Duvick, D.N., Cassman, K.G., 1999. Post-green-revolution trends in yield potential of temperate maize in the north-central United States. *Crop Sci.* 39, 1622–1630.
- Espe, M.B., Kirk, E., van Kessel, C., Horwath, W.H., Linquist, B.A., 2015. Indigenous nitrogen supply of rice is predicted by soil organic carbon. *Soil Sci. Soc. Am. J.* 79 (2), 569–576.
- Evans, L.T., Fischer, R.A., 1993. Yield potential: Its definition, measurement, and significance. *Crop Sci.* 39 (36), 1544–1561, <http://dx.doi.org/10.2135/cropsci1999.3961544x>.
- FAOSTAT, 2015. Food and Agriculture Organization of the United Nations Statistics Division. <http://faostat3.fao.org/> (accessed June 2015).
- Farrell, T.C., Fox, K.M., Williams, R.L., Fukai, S., 2006. Genotypic variation for cold tolerance during reproductive development in rice: screening with cold air and cold water. *Field Crops Res.* 98, 178–194.
- Fischer, R.A., 2015. Definitions and determination of crop yield gaps, and of rates of change. *Field Crops Res.* 182, 9–18.
- Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockstrom, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., 2011. Solutions for a cultivated planet. *Nature* 478, 337–342.
- Gaydon, D.S., Probert, M.E., Buresh, R.J., Meinke, H., Suriadi, A., Dobermann, A., Bouman, B.A.M., Timsina, J., 2012. Rice in cropping systems—modelling transitions between flooded and non-flooded soil environments. *Eur. J. Agron.* 39, 9–24.
- Grassini, P., Thornburn, J., Burr, C., Cassman, K.G., 2011. High-yield irrigated maize in the Western U.S. Corn Belt: I. On-farm yield, yield potential, and impact of agronomic practices. *Field Crops Res.* 120, 144–152.
- Grassini, P., Eskridge, K.M., Cassman, K.G., 2013. Distinguishing between yield advances and yield plateaus in historical crop production trends. *Nat. Commun.* 4, 2918.
- Hill, J., 2013. Rice crop growth and development. In: Dickey, D. (Ed.), California Rice Production Workshop Manual. UC Davis Cooperative Extension, Davis, CA, http://ucanr.edu/sites/UCRiceProject/Publications/Rice_Production_Workshop_Manual (accessed June 2015).
- IRRI, Postharvest Unit, 2013. *Harvesting*. IRRI, Los Baños, Philippines.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2015. *An Introduction to Statistical Learning with Applications in R*, 6th ed. Springer, New York, 440 pp.
- Julia, C., Dingkuhn, M., 2013. Predicting temperature induced sterility of rice spikelets requires simulation of crop-generated microclimate. *Eur. J. Agron.* 49, 50–60.
- Livezey, J., Foreman, L., 2004. Characteristics and Production Costs of U.S. Rice Farms. US Dept. of Ag. Statistical Bulletin, pp. 974–977.
- Lobell, D.B., Cassman, K.G., Field, C.B., 2009. Crop yield gaps: their importance, magnitudes, and causes. *Annu. Rev. Environ. Resour.* 34, 179–204.
- Loomis, R.S., Ragginage, R., Ng, E., 1979. Explanatory models in crop physiology. *Ann. Rev. Plant Physiol.* 30, 339–367.
- Moldenhauer, K., Slaton, N., 2015. Rice growth and development. In: Hardke, J.T. (Ed.), Arkansas Rice Production Handbook. Arkansas State Cooperative Extension, <http://www.uaex.edu/publications/pdf/mp192/mp192.pdf> (accessed July 2015).
- Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N., Foley, J.A., 2012. Closing yield gaps through nutrient and water management. *Nature* 490, 254–257.

- National Oceanic and Atmospheric Administration National Centers for Environmental Information (formerly the National Climatic Data Center). <https://www.ncei.noaa.gov/> (accessed March 2015).
- Core Team, R., 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/>.
- RiceTec, 2015. Clearfield Hybrids, <http://www.ricetec.com/Products/Hybrids/clearfield> (accessed July 2015).
- Runsick, S., Runsick, C.E., 2009. Rice Seeding Rate Recommendations for Arkansas. University of Arkansas, United States Department of Agriculture, and County Governments Cooperating.
- San-oh, Y., Mano, Y., Ookawa, T., Hirasawa, T., 2004. Comparisons of dry matter production and associated characteristics between direct-sown and transplanted rice plants in a submerged paddy field and relationships to planting patterns. *Field Crops Res.* 87, 43–58.
- Sheehy, J.E., Mitchell, P.I., 2015. Calculating maximum theoretical yield in rice. *Field Crop Res.* 182, 68–75.
- Temple Lang, D., 2015. RCurl: General Network (HTTP/FTP/..) Client Interface for R. R Package Version 1.95-4.6. <http://CRAN.R-project.org/package=RCurl>.
- University of California Cooperative Extension (UCCE), 2015. Rice Production Manual, <http://rice.ucanr.edu/Reports-Publications/Rice.Production-Workshop.Manual/> (accessed July 2015).
- USDA Economic Research Service, 2015. Rice Yearbook, <http://www.ers.usda.gov/data-products/rice-yearbook-2015.aspx> (accessed Jan 2016).
- van Bussel, L.G.J., Grassini, P., van Wart, J., Wolf, J., Claessens, L., Yang, H.S., Boogaard, H., de Groot, H., Hugo, S., Saito, K., Cassman, K.G., van Ittersum, M.K., 2015. From field to atlas: upscaling of location-specific yield gap estimates. *Field Crops Res.* 177, 98–108.
- van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Yield gap analysis with local to global relevance—a review. *Field Crops Res.* 143, 4–17.
- van Oort, P.A.J., Zhang, Tianyi, de Vries, M.E., Heinemann, A.B., Meinke, H., 2011. Correlation between temperature and phenology prediction error in rice (*Oryza sativa* L.). *Agric. Forest Meteorol.* 151 (12), 1545–1555.
- 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), e0118114.
- Van Wart, J., Kersebaum, C.K., Peng, S., Milner, M., Cassman, K.G., 2013. Estimating crop yield potential at regional to national scales. *Field Crops Res.* 143, 34–43.
- Van Wart, J., Grassini, P., Yang, H.S., Claessens, L., Jarvis, A., Cassman, K.G., 2015. Creating long-term weather data from the thin air for crop simulation modeling. *Agric. For. Meteorol.* 209–210, 49–58.
- Wang, W.G., Yu, Z.B., Wei, Z., Shao, Q.X., Zhang, Y.W., Luo, Y.F., Jiao, X.Y., Xu, J.Z., 2014. Responses of rice yield, irrigation water requirement and water use efficiency to climate change in China: Historical simulation and future projections. *Agric. Water Manag.* 146, 249–261.
- Yoshida, S., 1981. *Fundamentals of Rice Crop Science*. IRRI, Las Baños, Philippines (265 pp).
- Zhang, S., Tao, F., 2013. Modeling the response of rice phenology to climate change and variability in different climatic zones: comparisons of five models. *Eur. J. Agron.* 45, 165–176.

Supplemental Information

Supplemental Table 1: Weather stations used to simulate yields in US rice production systems

| State | Station Name | Latitude | Longitude | Elevation (m) |
|-------|------------------------------------|----------|-----------|---------------|
| LA | Sweet potato ^a | 32.100 | -91.700 | 21 |
| LA | Rice center (Crowley) ^a | 30.220 | -92.370 | 6 |
| MS | Stoneville ^b | 33.430 | -90.910 | 38 |
| MS | Lyon ^b | 34.220 | -90.540 | 53 |
| AR | Stuttgart ^c | 34.480 | -91.420 | 64 |
| AR | ASU ^d | 35.850 | -90.690 | 79 |
| AR | Corning ^d | 36.420 | -90.586 | 88 |
| CA | Durham ^e | 39.610 | -121.820 | 130 |
| CA | Colusa ^e | 39.230 | -122.020 | 55 |
| CA | Lodi West ^e | 38.130 | -121.390 | 25 |
| CA | Verona ^e | 38.800 | -121.610 | 24 |
| TX | Beaumont ^f | 30.070 | -94.302 | 5 |
| TX | Eagle Lake ^f | 29.600 | -96.345 | 53 |

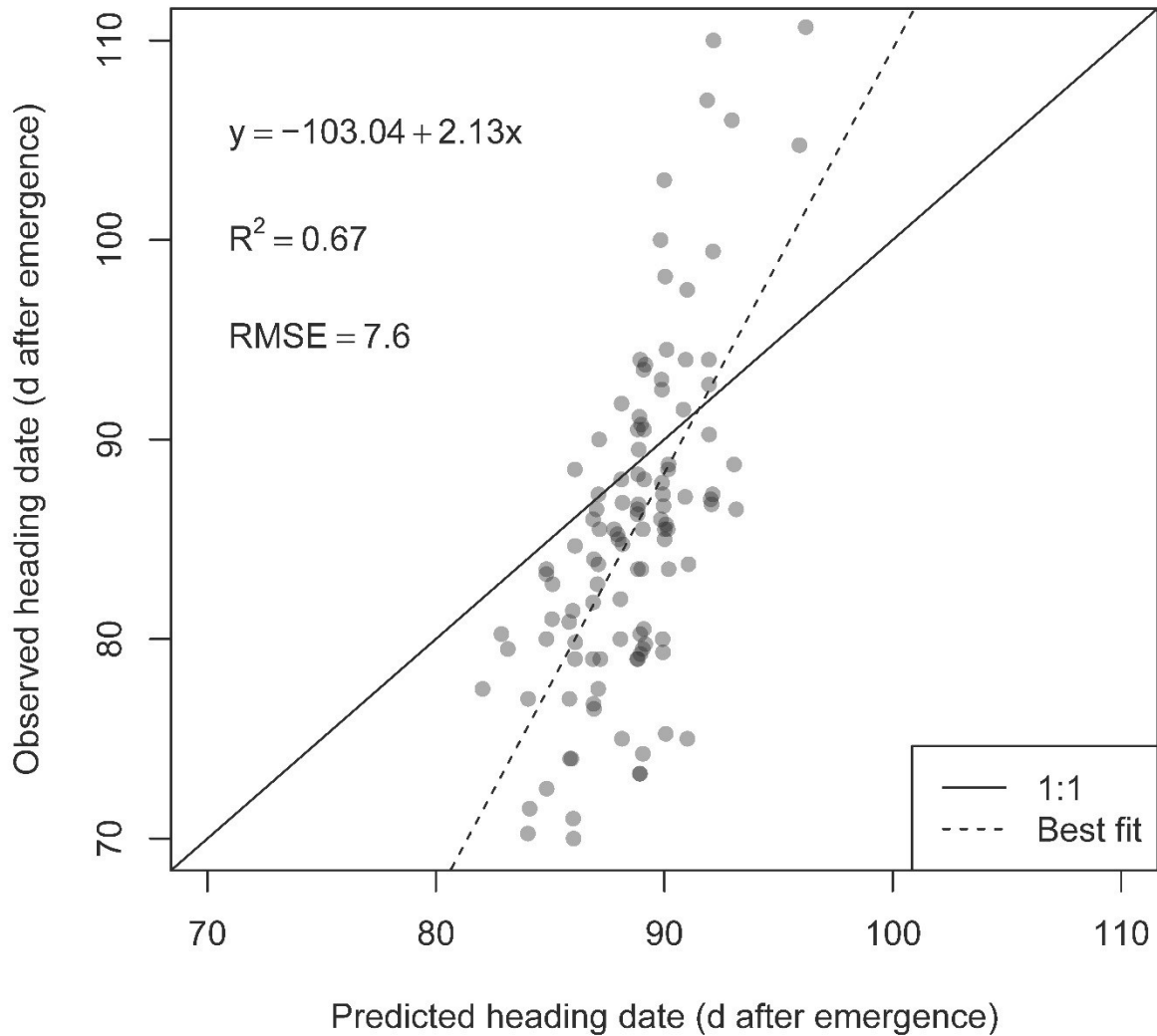
Sources: ^a <http://weather.lsuagcenter.com/>; ^b www.deltaweather.msstate.edu; ^c

<http://www.ars.usda.gov/Main/Docs.htm?docid=23623>; ^d [http://www.ncdc.noaa.gov/data-access/land-based-station-](http://www.ncdc.noaa.gov/data-access/land-based-station-data/)
[data](http://www.cimis.water.ca.gov/); ^e <http://www.cimis.water.ca.gov/>; ^f <https://beaumont.tamu.edu/climaticdata/>

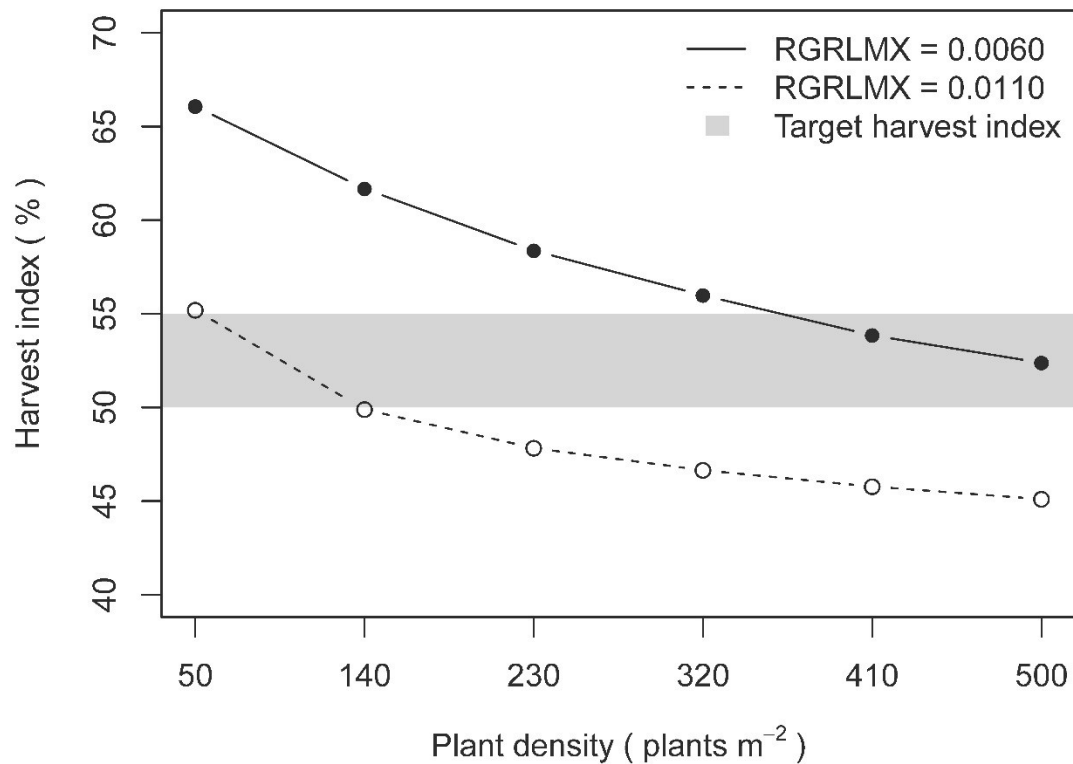
Supplemental Table 2. Values of calibrated cardinal temperatures. Calibration followed the steps outlined in van Oort et al. (2011) to calibrate the base (TBD), optimum (TOD), and maximum (TMD) temperatures for phenological development using the ‘pheno_opt_rice3’ program and the “bilinear1” routine. Two search routines were used: (1) large step with TMD up to 999°C, and then (2) small step to narrow in on optimal values. The values displayed are those that resulted in the lowest correlation between temperatures and phenology prediction error from emergence to flowering. Calibration reduced the correlation between temperature and phenology prediction error, but at the cost of increased error in phenology predictions (Supplemental Fig. 1).

| | Base temperature (TBD) | Optimum temperature (TOD) | Maximum temperature (TMD) |
|------------------|---------------------------|------------------------------|------------------------------|
| Calibrated value | 4.0 | 26.0 | 40.0 |

M-206



Supplemental Figure 1: Observed date of 50% heading versus the predicted date of 50% heading for model simulations in which the cardinal temperatures used to calculate heat units have been calibrated according to van Oort et al. (2011). Calibration of cardinal temperatures decreased the correlation between phenology error and temperature in the calibration data set, but resulted in poorer predictive performance in the validation data set compared to using default temperatures (RMSE of 7.6 verses 5).



Supplemental Figure 2: Sensitivity of simulated harvest index values in the ORYZA crop model to changes in the maximum relative growth rate of leaves (RGR_LMX) for two values of RGR_LMX and plant density values ranging from 50 to 500 plants m⁻². The target harvest index was between 50 and 55%, typical of modern high-yielding rice varieties. In order to achieve these harvest index values, RGR_LMX was assumed to be 0.0060 for variety M-206 (planted at 350 plants m⁻²) and 0.0110 for variety CXL745 (planted at 50 plants m⁻²).