UC Davis

UC Davis Electronic Theses and Dissertations

Title

Floodwater-driven microclimate alters thermal accumulation from gridded temperature data in rice systems

Permalink

https://escholarship.org/uc/item/6sr0x03r

ISBN

9798189228278

Author

Daniel, Lewis Edgar

Publication Date

2026-06-25

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Floodwater-driven microclimate alters thermal accumulation from gridded temperature data in rice systems

By

LEWIS EDGAR DANIEL III

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

Horticulture and agronomy

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Bruce A. Linquist, Chair

_____

Mark E. Lundy

_____

Kassim Al-Khatib

Committee in Charge

2026

Abstract

Accurate temperature is critical for thermal time–based phenological simulation modeling. In flooded rice systems, the presence of standing water modifies surface energy balance and canopy temperature. Temperature data from meteorological stations or gridded temperature data derived from these measurements may not capture the canopy-level microclimate of flooded agroecosystems. Our objective was to quantify differences between these two temperatures and how they impact modeled growing degree day (GDD) accumulation through the season. We evaluated temperature-source effects using nine years of observations from eight rice production sites (62 site-years) in California's Sacramento Valley, comparing PRISM gridded temperature to canopy-level and floodwater In-field sensor measurements. A threshold-bound (cardinal temperature) GDD model was developed for PRISM-based phenological prediction of days to heading. In-field water temperatures differed systematically from canopy-level air temperatures across the season. Relative to air temperature, water temperature was warmer early in the season and cooler later in the season as the canopy closed. Reflecting these differences, compared to PRISM air temperatures, In-field daily average air temperatures were higher (by 0.65 °C) early in the season until about 30 days after planting and lower (by 1.17 °C) for the rest of the season. This resulted in predictable stage-dependent differences in GDD accumulation between In-field and PRISM measurements throughout the season. Our results indicated that with the use of In-field temperature data, at the phenological stages of panicle initiation, heading, and R7, GDD accumulation was +11.5 °C·d, 0 °C·d, and -27 °C·d, respectively, different from when PRISM temperature data was used. These results indicate that the floodwater impacts the field microclimate and should be accounted for in GDD models to improve predictive capacity.

TABLE OF CONTENTS

ABSTRACT ii
LIST OF FIGURES iv
LIST OF TABLES v

1. INTRODUCTION 1
2. MATERIALS AND METHODS 4
   2.1 Variety trials 4
   2.2 Temperature data: collection and comparison 5
   2.3 Growing Degree Day Model (GDD) 7
      2.3.1 Crop phenology data 7
      2.3.2 Model development 8
3. RESULTS 11
   3.1 Temperature data 11
      3.1.1 Seasonal air and water temperatures 11
   3.2 GDD-Based Prediction of Days to Heading (DTH) 15
   3.3 GDD accumulation variation 18
4. DISCUSSION 19
   4.1 GDD Model 19
   4.2 Water temperature versus air temperature 20
      4.2.1 Seasonal changes in In-field water and air temperature ... 20
      4.2.2 Impact of floodwater on air temperature source 21
   4.3 Impact of temperature sources for GDD accumulation 22
      4.3.1 GDD predictions using the model 22
   4.4 Considerations 24
CONCLUSION 26
REFERENCES 27

LIST OF FIGURES

Figure 1. The main rice growing region in Sacramento Valley, California, with counties, approximate GPS coordinates of the eight general site locations, and PRISM 4-kilometer satellite temperature grid represented. 7

Figure 2. Daily In-field air and water temperature averaged across nine seasons (2016–2024). Panels show minimum, maximum, and average temperatures by days after planting (DAP). 12

Figure 3. Differences between PRISM and In-field daily air temperature averages across nine seasons (2016–2024) and 62 site-years, calculated as PRISM − In-field. 14

Figure 4. Observed versus predicted days to heading (DTH) in the validation dataset for In-field and PRISM temperature sources. 17

Figure 5. Average difference in modeled cumulative Growing Degree Days (GDD, °C·d) calculated as PRISM − In-field across the growing season. 19

Figure S1. Distribution of planting dates across site-years (2016–2024) after exclusion of environments removed from the analysis (see Methods). 32

Figure S2. Workflow used for model development and evaluation. 33

Figure S3. Example distribution of observations across counties for one representative random seed, showing the 70% training, 20% calibration, and 10% validation partitioning used for model development and evaluation. 34

LIST OF TABLES

Table 1. Mean and range of observed days to heading (DTH) for the five main Calrose medium-grain varieties across environments. 15

Table 2. Summary statistics of cumulative Growing Degree Days (GDD, °C·d) for days to heading (DTH) for five medium-grain varieties using In-field and PRISM temperature datasets. 17

Table S1. Summary statistics of cumulative Growing Degree Days (GDD, °C·d) to heading for the five main Calrose medium-grain varieties using the uncapped base-only formulation (Tb = 7 °C). 34

1. Introduction

Accurate representation of thermal regimes is central to agroecosystem modeling due to the influence of temperature on the rate of plant developmental processes and phenology. Rice (*Oryza sativa* L.) varieties that are not photoperiod-sensitive do not differ in their physiological responses to temperature (Yin et al., 1997; Mackill and Lei, 1997). Crop phenology can be determined through calendar-based approaches, visual staging systems (Counce et al., 2000), or remote sensing indicators such as NDVI and LAI derived from satellite platforms (Liao et al., 2022). While these approaches provide practical detection of developmental stages, temperature-driven thermal time models are the dominant framework for predictive applications. Growing degree day (GDD) formulations are the basis of phenological development in process-based crop models such as ORYZA2000 and CERES-Rice (Bouman et al., 2001; Timsina and Humphreys, 2006; van Oort et al., 2011).

Many GDD models assume that developmental rate increases linearly with temperature between biologically defined cardinal thresholds and ceases outside these bounds (Chuine et al., 2003). While conceptually simple, cumulative thermal time is highly sensitive to small biases in daily minimum and maximum temperatures. Small differences between temperature datasets (i.e. ±0.5–1 °C in daily minimum or maximum temperature), when integrated over a growing season, can result in multi-day differences in predicted phenological events (van Oort et al., 2011). Thus, phenological prediction accuracy depends not only on threshold parameterization but also on the quality of the temperature dataset used to drive the model.

Temperature data are commonly derived from either meteorological station networks or spatially interpolated gridded products. The Parameter-elevation Regressions on Independent Slopes

Model (PRISM) dataset (Daly et al., 2008) is widely used in the United States and serves as a foundational ground-based gridded climate product for agricultural and hydrological applications (Mourtzinis et al., 2017; Mankin et al., 2025). PRISM integrates meteorological station observations using physiographically informed regression techniques to produce spatially continuous daily temperature surfaces. Although gridded datasets provide spatial completeness and long-term consistency, interpolation procedures can reduce local variability and smooth temperature extremes relative to actual field conditions (Daly et al., 2002).

Because rice phenology is highly sensitive to temperature, the ability of interpolated temperature products to accurately represent field conditions is critical for reliable phenological prediction. Meteorological stations are typically situated on land and over mowed, irrigated grass. While this may provide adequate temperature estimates for upland crops, rice is grown in a flooded environment for much of the growing season. The presence of floodwater modifies the surface energy balance by increasing the effective heat capacity of the surface layer and altering the partitioning of sensible and latent heat fluxes (Snyder et al., 2001). Water possesses substantially greater thermal inertia than air, and flooded rice systems can therefore exhibit substantial divergence between water, near-water, and standard air temperatures (Confalonieri et al., 2005). This distinction is biologically important because the temperature experienced by the rice crop and driving phenological development is not the air temperature from a nearby station but the temperature of the near-canopy microclimate. For much of pre-floral rice development, the shoot apex is influenced by water temperature and near-canopy microclimate (Dingkuhn et al., 1995; Sharifi et al., 2018). In a rice field, the relationship between water and air temperature changes because paddy water temperature depends upon various factors, including floodwater depth and canopy density, which impact the amount of solar radiation and wind speed at the water surface

(Kuwagata et al., 2008). Empirical measurements have demonstrated divergence between floodwater temperature, near-canopy air temperature, and regional meteorological observations in rice production systems (Confalonieri et al., 2005; Sharifi et al., 2018).

Although prior studies have shown that rice phenology models are sensitive to temperature inputs and that water and air temperatures can diverge in flooded systems (Confalonieri et al., 2005; Kuwagata et al., 2008; Sharifi et al., 2018); no studies to our knowledge, have evaluated whether commonly used gridded air temperature products capture canopy-level thermal conditions in flooded rice fields and how these differences impact cumulative GDD trajectories. In Mediterranean climates characterized by large diurnal temperature ranges, such as California's Sacramento Valley, differences in nighttime cooling and daytime heating between flooded fields and surrounding land surfaces may amplify differences between gridded and canopy level air temperatures. Because thermal time models integrate temperature differences over time, this may result in significant differences in cumulative GDD and inaccurate prediction accuracy (van Oort et al., 2011). Therefore, evaluating temperature-source effects at biologically meaningful growth stages will provide an assessment of whether dataset-induced differences are amplified, diminished, or offset as development progresses.

The objectives of this study were to: (i) quantify seasonal differences in daily minimum, maximum, and average air temperatures between PRISM and canopy-level In-field sensors; (ii) quantify differences between In-field water and air temperature; and (iii) evaluate how these differences impact GDD accumulation over time using a model. By isolating the temperature source impact within a flooded agroecosystem, this study provides insight into how surface–atmosphere interactions and dataset selection jointly influence thermal time modeling outcomes.

To accomplish our objectives, we analyzed nine years of observations from eight sites in California's Sacramento Valley.

2. Materials and Methods

All temperature and field data were taken from the Sacramento Valley of California where roughly 95% of the rice is produced in California (Hill et al., 2006), and which produces about 20% of the rice grown in the USA (USDA NASS, 2023). The region is characterized by a Mediterranean climate with cool, wet winters and hot, dry summers. There is also substantial diurnal temperature variation with daily minimum and maximum temperatures differing by 15-20 °C during the growing season based on PRISM data in this study. In this region, rice is typically established by water-seeding and maintained under continuous, shallow flood conditions (approximately 8-15 cm depth) throughout the growing season until about 4 weeks before harvest, when fields are drained (Linquist et al., 2025). Rice is planted in late April through May and is harvested in late September through early November.

2.1 Variety trials

Data were compiled across nine growing seasons (2016-2024) from the University of California Cooperative Extension statewide variety trials conducted in the Sacramento Valley each year (Fig. 1). There are typically seven to eight trials in any given year, providing a total of 62 site-years for analysis. Trial plots are replicated using randomized complete block designs (RCBD), each plot is 3 x 6 m, and replicated multiple times. These trials are located within farmers' fields and are farmer managed. Generally, trial location remains on the same farm each year, but the field location may vary. Data collected from these trials include early season vigor, 50% heading

date, lodging and yield. Days to heading (DTH), defined as the number of days from planting until 50% heading, was used for subsequent GDD model development and evaluation.

2.2 Temperature data: collection and comparison

Temperature data were collected from two sources: PRISM and In-field sensors. Throughout this manuscript, "In-field" refers to temperature measurements collected directly within production rice fields using on-site sensors. For each temperature source, daily minimum (T_{min}) and maximum (T_{max}) air temperatures were extracted for each site and year.

In-field air and water temperatures were measured using HOBO U23 Pro v2 data loggers (Onset Computer Corporation, Bourne, MA, USA) with two sensors recording temperatures on an hourly basis. These sensors were located inside the field within the variety trial described above. Sensors were placed in the field on the day of planting and removed when the trial was harvested. The air temperature sensor was housed in a passive radiation shield and positioned 20-40 cm above the canopy (or water, as is the case before rice canopy develops about water). The other sensor was placed on the soil surface and underneath the floodwater, effectively measuring the temperature of the floodwater at the soil surface for the majority of the season. Early in the season the field may be drained for short periods (i.e. for seedling establishment purposes) and later in the season (at about 110 days), fields are drained in preparation for harvest. During these times, the sensor measures air temperature at the soil level. Daily T_{min} and T_{max} values were collected and screened for missing data or outliers. Problematic observations occurred primarily during the initial deployment period on the first or second day. In the case of outliers, which were usually a single data point and represented only 2.5% of observations, the gap was filled using linear interpolation between the two adjacent days. If the sensors were deployed one or two

days after planting, we adjusted values for the missing days based on the PRISM T_{max} and T_{min} data for that location, adjusted for the difference between In-field and PRISM T_{max} and T_{min} daily data (discussed later in this manuscript).

PRISM integrates meteorological station observations using informed regression models to generate spatially continuous temperature surfaces with a 4-km grid (Daly et al., 2008). We used the site-specific data based on the latitude and longitude coordinates of each variety trial to collect PRISM T_{min} and T_{max} data for each location. The data collected from PRISM and In-field sensors were standardized so that all daily temperatures were relative to the day of planting. Actual day of planting ranged from April 23rd to June 3rd, with the average planting date occurring on May 13 (median May 14), which is representative for California (Linquist et al., 2025). However, most planting dates were concentrated within a narrower window, with 50% of site-years planted between May 6 and May 19 and approximately 75% planted by May 19 (Fig. S1). Because subsequent analyses are expressed relative to days after planting (DAP), this distribution provides calendar-date context for interpreting seasonal temperature and GDD trends. Average daily temperature (T_{avg}), was calculated from daily T_{min} and T_{max}. We compared PRISM to In-field sensor data for T_{min}, T_{max}, and T_{avg} temperature for each day throughout the season by examining the absolute temperature differences.

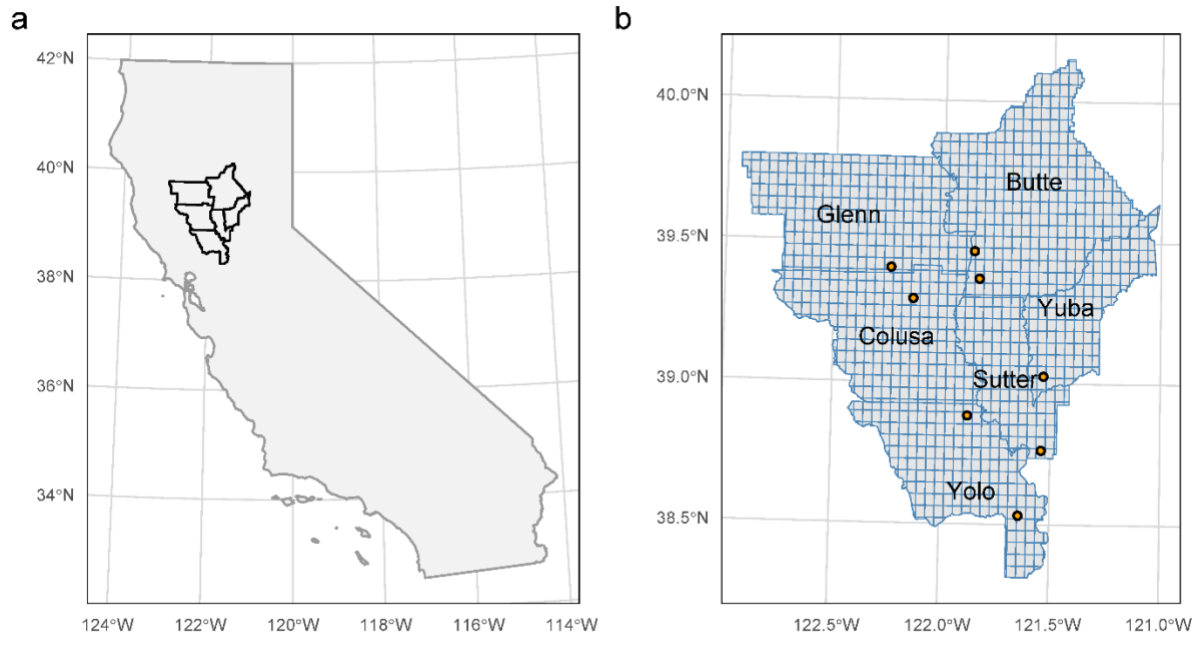


Figure 1. The main rice growing region in Sacramento Valley, California, with counties, approximate GPS coordinates of the eight general site locations, and PRISM 4-kilometer satellite temperature grid represented.

2.3 Growing Degree Day Model

2.3.1 Crop phenology data

From each of the variety trials, five commercially grown medium-grain Calrose-type cultivars (M-105, M-206, M-209, M-210, and M-211) were evaluated. To evaluate the impact of temperature source on GDD accumulation, we used days to heading (DTH - number of days from planting to 50% heading for each variety) as the phenological response variable for model calibration and validation because it is a clearly defined developmental stage commonly used in rice phenology studies. DTH was defined as the date at which 50% of the main tillers reached R3 (exhibited panicle exertion) according to Counce et al. (2000).

2.3.2 Model development

To evaluate how the temperature data source influences growing degree day (GDD) accumulation and phenological prediction, a GDD model was developed to estimate DTH. Daily GDD was calculated using a threshold-bound formulation:

$$GDD = \left(\frac{\min(T_{\max}, T_u) + \max(T_{\min}, T_l)}{2}\right) - T_b$$

where T_{max} and T_{min} represent daily maximum and minimum air temperature, T_u and T_l represent upper and lower temperature thresholds, and T_b represent the base temperature for development. Daily T_{max} exceeding $T_u$were set equal to T_u, and daily T_{min}below T_l were set equal to T_l prior to calculating GDD. This threshold-bound formulation is related to other thermal-time approaches used in crop phenology, ranging from simple base-temperature degree-day models such as DD10 (Counce et al., 2015) to rice phenology models that incorporate cardinal temperature thresholds (referred to here as a threshold-bound model), such as ORYZA2000 (Bouman et al., 2001).

Candidate threshold (cardinal) temperatures were informed based on published data for rice development (Gao et al., 1987; Yoshida, 1981; Yin et al., 1997) and systematically evaluated across biologically plausible ranges using PRISM temperature data. T_l, T_u, and T_b temperature values were evaluated within biologically plausible ranges (T_l: 10–18 °C, T_u: 25-40 °C, T_b: 7–10 °C) reported for rice development in previous studies (Yin et al., 1997; Gao et al., 1992; Confalonieri et al., 2009).

The dataset was stratified by site and quartiles of observed DTH and partitioned into three subsets: training (70% of data), calibration (20%), and validation (10%). This partitioning was repeated across five random seeds, where each seed generated a different random split of the data, to attain robust model selection and performance. Five repetitions were used as a practical compromise to evaluate robustness across multiple independent splits while maintaining computational tractability. The training data were used for model development, including internal five-fold cross-validation of candidate temperature parameter sets, resulting in 144 candidate threshold combinations evaluated within each fold. The output for each model was evaluated using RMSE, mean bias, and R^2 within the training cross-validation procedure. The calibration data were then used to compare candidate parameterizations with the same metrics and to identify the best-performing shared-threshold model. The validation data were not used for model selection and were reserved for final independent assessment of prediction accuracy using RMSE, mean bias, and R^2.

For each variety × location × year environment, cumulative GDD was summed from day of planting onward. For a given combination of T_l, T_u, and T_b, a variety-specific thermal requirement for heading was estimated from the training data as the cumulative GDD associated with the observed DTH across training environments for that variety. Predicted DTH was then defined as the first day after planting on which cumulative GDD reached that variety-specific thermal requirement. This procedure was repeated for each candidate T_l, T_u, and T_b combination at 1 °C increments across ranges of 10–18 °C, 25–40 °C, and 7–10 °C. Model performance within each training fold was quantified by comparing observed and predicted DTH using RMSE, mean bias, and R^2.

For most GDD models, temperatures from a meteorological station or gridded database (like PRISM) are used. Candidate threshold-bound models were developed using PRISM daily T_{min} and T_{max} temperatures. The selected PRISM-based model was then run using In-field temperature data for a direct comparison of the temperature sources under a common parameterization rather than conflating temperature-source differences with differences introduced by separate optimization. For this reason, the candidate parameterization, mean RMSE, mean bias, and mean R^2 were summarized across the five training folds. The independent calibration dataset was then used to identify the final shared threshold set by jointly considering RMSE, mean bias, and R^2.

We also compared PRISM and In-field temperature sources in terms of GDD accumulation through a season. Cumulative GDD differences between temperature sources were first calculated using the selected threshold-bound model to evaluate how the final operational parameterization impacted temperature-source differences through a season. Secondly, to evaluate how cardinal temperature constraints influenced GDD accumulation between temperature sources, we examined a non-threshold bound base-only degree-day model. In this model, daily thermal time was calculated similarly as before, using the same T_b (7 °C), but without T_l and T_u thresholds:

$$GDD = \left(\frac{\min(T_{\max}) + \max(T_{\min})}{2}\right) - T_b$$

Comparison of these two thermal-time formulations allowed assessment of temperature-source differences independent of the temperature thresholds, as other GDD models use (i.e. Counce et

al., 2015). Particular attention was given to phenologically important developmental stages of panicle initiation (PI), DTH, and R7, with PI and R7 timing estimated from Sharifi et al. (2017).

3. Results

3.1 Temperature data

3.1.1 Seasonal air and water temperatures

Overall, the seasonal patterns in air temperature from the In-field canopy-level sensor reflect typical Mediterranean summer conditions characterized by warm daytime temperatures and cooler nighttime conditions (Fig. 2). Because temperatures are presented relative to days after planting (DAP) rather than calendar date, early-season conditions correspond approximately to May (average planting date occurred on May 13). Based on the In-field sensors, the highest mid-season temperatures occurred primarily during July and August. Following planting, both air T_{min} and T_{max} increased steadily as the season progressed. Air T_{min} averaged 13-14 °C during the early vegetative period and rose to 16-17 °C by mid-season. Air T_{max} also increased over this interval, from 28–29 °C early in season to 32–33 °C during mid-season. Similarly, T_{avg} rose from 20–21 °C early in the season to 24–25 °C in the mid-season. T_{max} increased more rapidly during the early season and remained elevated through the reproductive stages of the crop, whereas T_{min} peaked at mid-season then gradually declined as the season progressed.

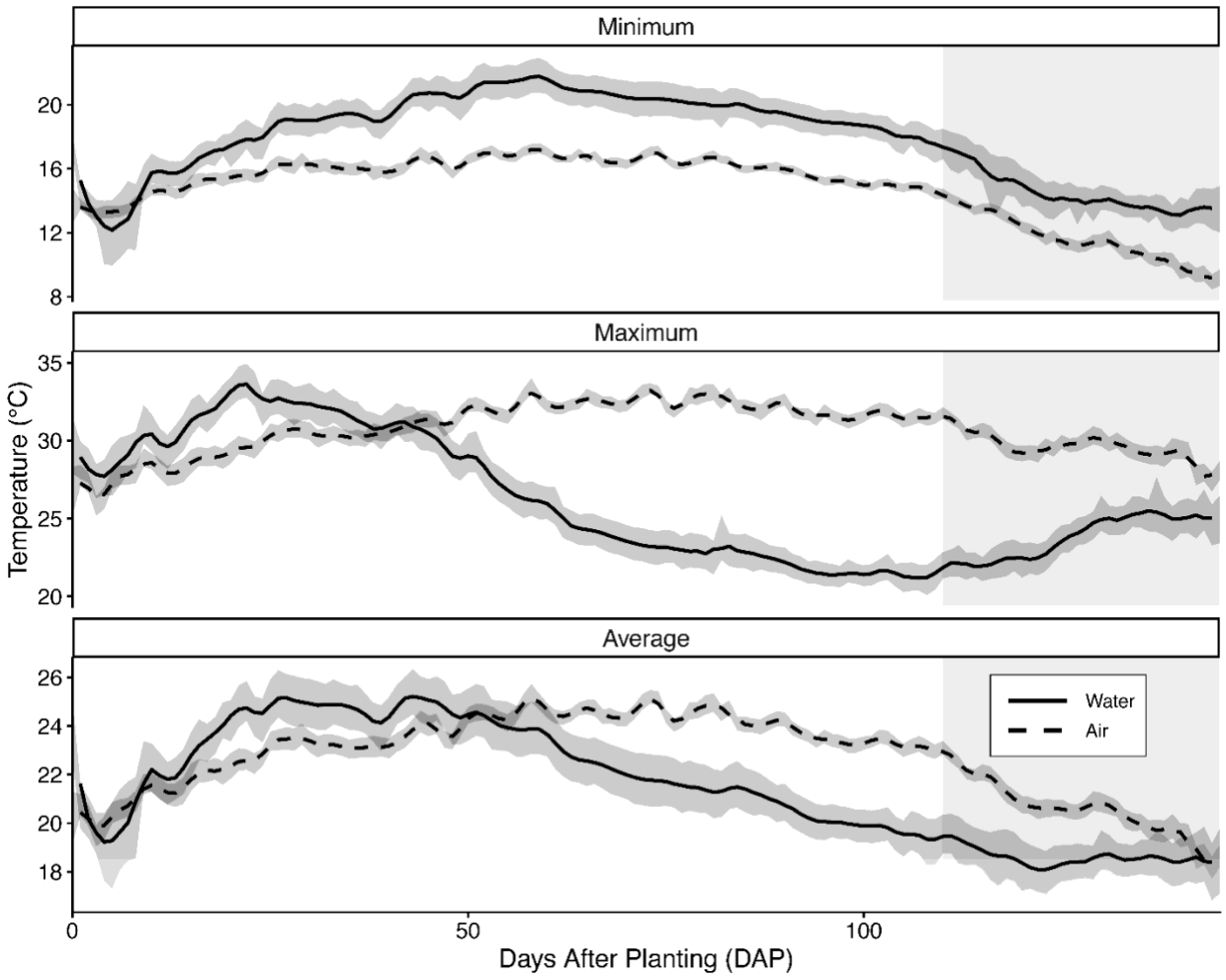


Figure 2. Daily In-field air and water temperature averaged across nine seasons (2016–2024). Panels show minimum, maximum, and average temperatures by days after planting (DAP). The shaded area around each line represents ± standard error across sites and years. The shaded region beginning at 110 DAP represents the approximate drainage period used for visualization. Actual drainage timing varied among site-years but generally occurred approximately two weeks before harvest. During this period, the "water" temperature sensor no longer measures floodwater temperature but instead the temperatures at the soil surface beneath the canopy.

The difference between In-field and PRISM temperature is shown in Figure 3. PRISM T_{min} was lower than In-field T_{min} from planting until 74 DAP, with the largest early-season differences being 1.8 °C lower during the first 18 days of the season. By contrast, PRISM T_{max} was higher than In-field T_{max} throughout the season. At the onset of the season, T_{max} differences were small to negligible; however, PRISM T_{max} averaged 2.5 °C higher than In-field T_{max} from 50 to 110

DAP. PRISM T_{avg} was lower than In-field early in the season until 29 DAP, after which, it was higher for the remainder of the season. PRISM T_{avg} averaged 1.2 °C higher than In-field T_{avg} from 30 to 110 DAP and reached a maximum difference of 1.8 °C at 113 DAP.

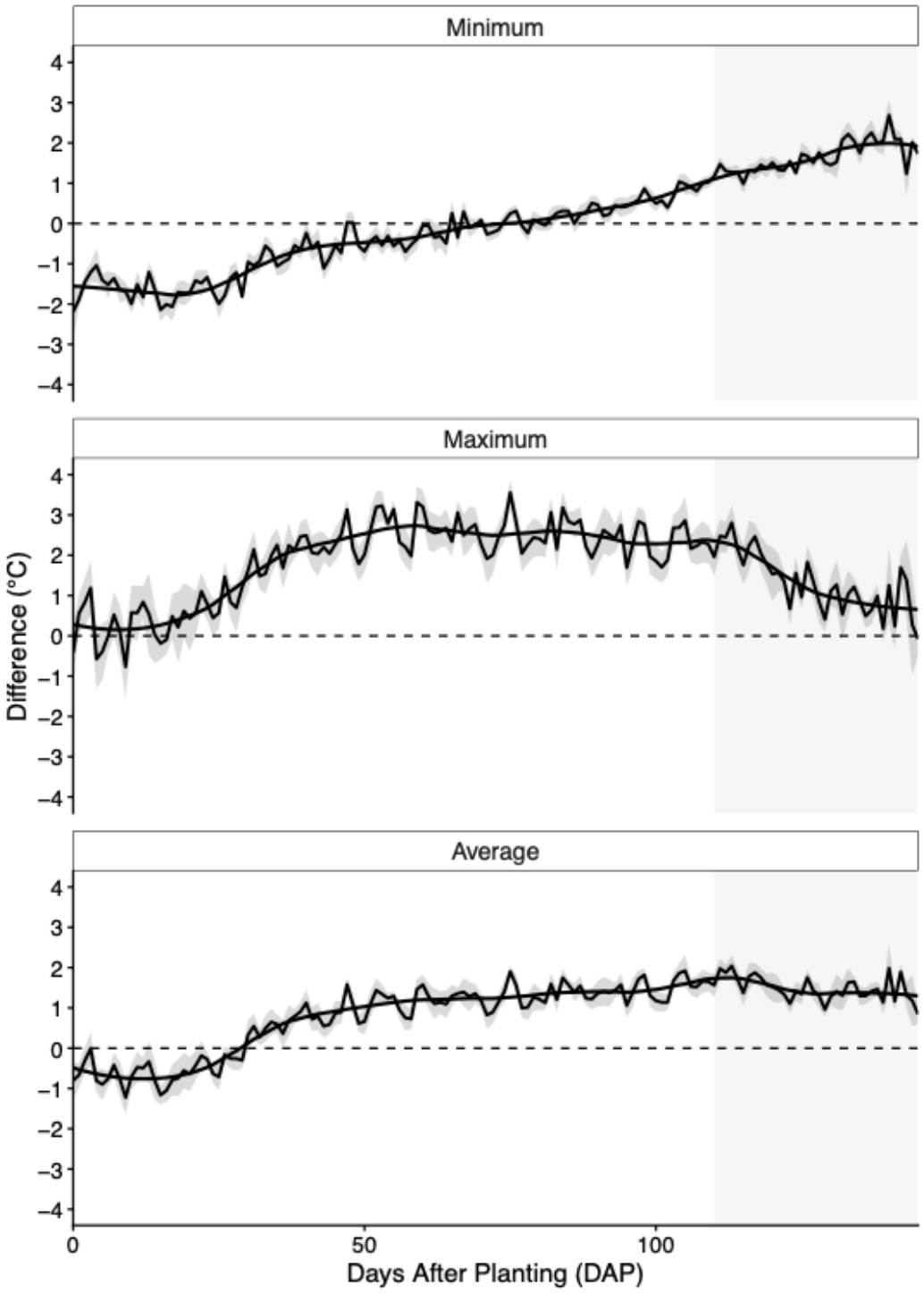


Figure 3. Differences between PRISM and In-field daily air temperature averages across nine seasons (2016–2024) and 62 site years, calculated as PRISM − In-field. Panels show minimum, maximum, and average air temperature differences by days after planting (DAP). Values above the zero dashed line indicate higher PRISM temperatures, while values below zero indicate higher In-field temperatures. Grey ribbons represent ± standard error across years. The grey shaded region indicates the approximate drainage period beginning at 110 DAP.

Figure 2 illustrates the seasonal differences between In-field water and air temperatures. Water T_{min} was higher than air T_{min} throughout the season. In contrast, water T_{max} was higher than air T_{max} early in the season, but after 40 DAP (with increased canopy shading), water T_{max} became increasingly lower than air T_{max}. As a result, water T_{avg} was higher than the air temperature early in the season, but after about 50 DAP, it became lower and remained so throughout the rest of the season. During the first 10 days of the season, the differences between water and air temperature were not consistent. This is likely due to differences in water management practices at the onset of the season between fields due to varying establishment practices and weed control, which can result in the removal of floodwater from the field for short periods of time. Also, after the end of the season, the field is drained to prepare for harvest (about 110 DAP), so the water sensor measurements no longer represent standing floodwater temperature but instead reflect temperatures at the soil surface beneath the canopy.

3.2 GDD-Based Prediction of Days to Heading (DTH)

Observed DTH varied among cultivars and across environments. Average days to heading ranged from 84 to 92 days after planting, with individual observations spanning 73-104 days depending on variety year and location (Table 1). M105 was the earliest variety (average: 84 DTH) while M209 and M211 were the latest (average: 92 DTH).

Table 1. Mean and range of observed days to heading (DTH) for the five main Calrose medium-grain varieties across environments.

Variety	Mean DTH	SD	Range
M105	84	5.2	73-97
M206	86	5.3	74-100
M209	92	5.5	78-104
M210	86	5.1	73-99
M211	92	5.5	79-104

Candidate combinations of T_l, T_u, and T_b were evaluated at 1 °C increments across their specified ranges, resulting in 144 candidate threshold combinations evaluated within each cross-validation fold. The joint consideration of RMSE, mean bias, and R^2 in the calibration dataset identified T_l = 12 °C, T_u = 33 °C, and T_b = 7 °C as the best-performing set of cardinal temperatures. These thresholds were subsequently applied to the independent validation dataset, where the model performance metrics showed an RMSE of 4.02 days, R^2 of 0.51, and mean bias of 0.04 days (Fig. 4). The same set of cardinal temperatures was also applied using In-field temperature data to permit direct comparison of temperature sources under a common parameterization. Using In-field temperature data resulted in similar performance to that obtained with PRISM temperatures, with an RMSE of 3.88 days, R^2 = 0.54, and a mean bias of 0.08 days. Furthermore, both models had very similar GDD requirements for DTH (Table 2). Averaged across the five varieties, differences in cumulative GDD at heading between PRISM and In-field temperature sources were 8.2 GDD. Given that roughly 16-17 GDD corresponds to one day in this dataset, an 8.2 GDD difference is approximately on the order of half a day. Using the PRISM data, M105 required the least GDD to reach heading (1358 °C·d) and M211 required the most (1497 °C·d), a difference of 139 °C·d (about 8 to 9 days).

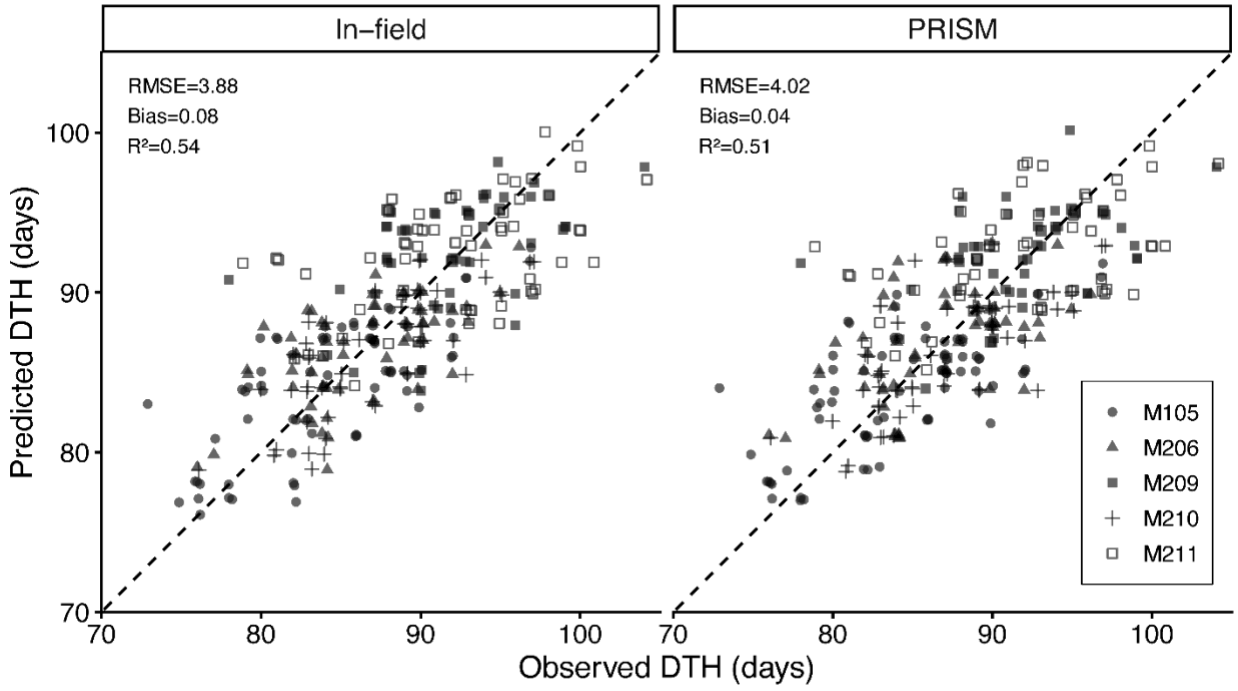


Figure 4. Observed versus predicted days to heading (DTH) in the validation dataset for In-field and PRISM temperature sources using the following thresholds: T_l = 12 °C, T_u = 33 °C, and Tb = 7 °C. Validation results are pooled across five independent train–calibration–validation partitions generated using different random seeds. Each point represents a variety × location × year observation.

Table 2. Summary statistics of Minimum, Mean, Median, Maximum, and Standard Deviation of cumulative Growing Degree Days (GDD, °C·d) for days to heading (DTH) for five medium grain varieties M105, M206, M209, M210, and M211. Data show statistics of modeled output for In-field and PRISM temperature datasets. n = number of variety-specific heading observations across environments. All GDD values (Min, Mean, Median, Max, and SD) are reported in °C·d.

Variety	Source	*n*	Min	Mean	Median	Max	SD
M105	In-field	62	1184	1352	1354	1470	61
	PRISM	62	1165	1358	1373	1511	68
M206	In-field	62	1199	1389	1390	1502	61
	PRISM	62	1184	1396	1404	1544	69
M209	In-field	62	1264	1483	1473	1619	69
	PRISM	62	1248	1494	1502	1633	75
M210	In-field	62	1184	1392	1392	1518	60
	PRISM	62	1165	1399	1407	1561	67
M211	In-field	60	1282	1487	1479	1636	73
	PRISM	60	1266	1497	1502	1670	79

3.3 GDD accumulation variation

We quantified how temperature source influenced cumulative GDD accumulation through the growing season (Fig. 5). Early in the season (up to about 80 DAP), PRISM cumulative GDD was lower than In-field accumulation. After 80 DAP, PRISM cumulative GDD exceeded In-field accumulation. Specifically, through the PI growth stage (~45–55 DAP), PRISM cumulative GDD was up to 10–15 °C·d lower than In-field accumulation. Differences narrowed approaching heading (~70–80 DAP), where cumulative GDD from both sources converged near 0 °C·d difference. Later in the season, during R7 (105–120 DAP), PRISM cumulative GDD was approximately 27 °C·d higher than In-field accumulation. By the end of the season, this difference had increased to approximately 54 °C·d.

Using the model without temperature thresholds, PRISM cumulative GDD differed more strongly from In-field accumulation than when the threshold-bound model was used (Fig. 5). Early in the season, there was little difference between PRISM and In-field GDD accumulation; however, after about 40 DAP, the non-threshold bound PRISM GDD began accumulating GDD more rapidly, and by the end of the season it exceeded the In-field GDD accumulation by approximately 144 °C·d. This greater divergence relative to the bound model is primarily attributable to the upper temperature threshold (T_u) of 33 °C in the threshold-bound model, which limited the influence of the warmer PRISM temperatures. To illustrate how often the upper-temperature threshold altered heat accumulation, across all site-years, daily maximum temperature exceeded 33 °C on 4,415 days in PRISM data and 2,544 days in In-field data.

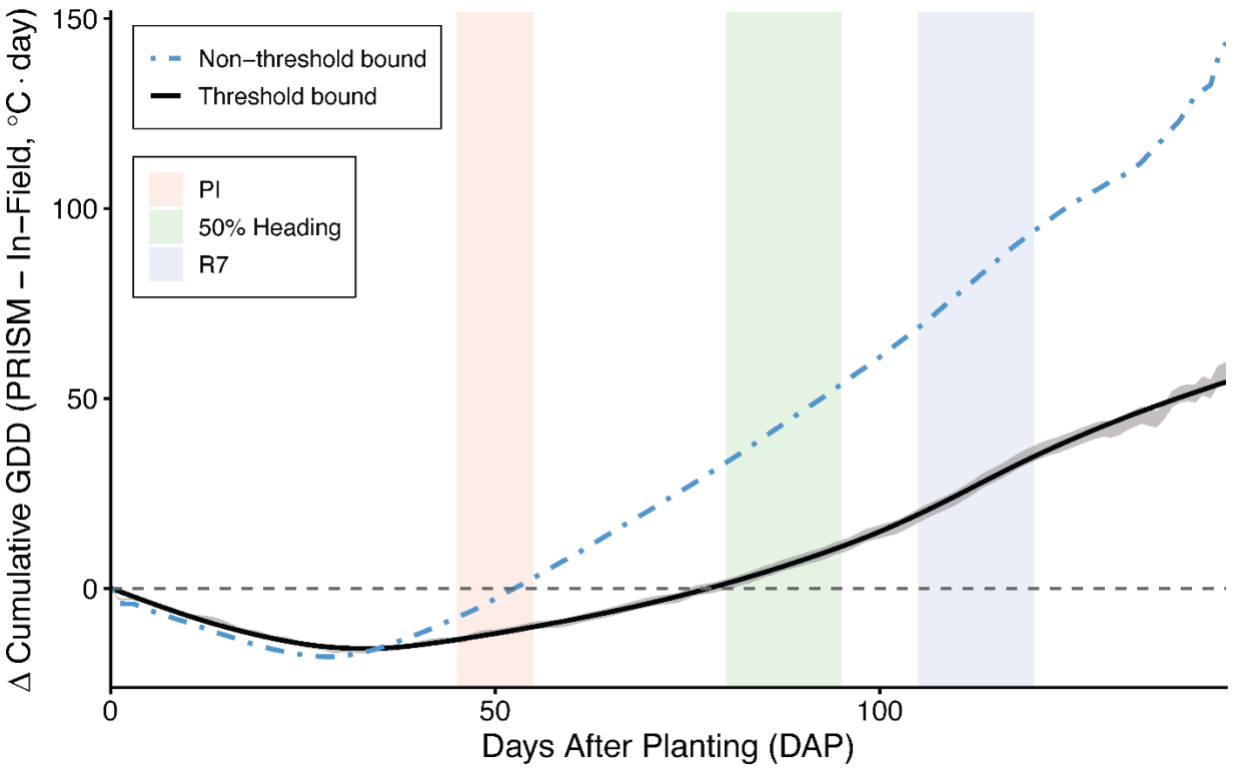


Figure 5. Average difference in modeled cumulative Growing Degree Days (GDD, °C·d) calculated as PRISM − In-field across the growing season. Values above the zero dashed line indicate that PRISM temperatures accumulate GDD faster than In-field temperatures, while values below zero indicate faster GDD accumulation using In-field temperatures. The solid black line represents the mean difference across site-years, and the grey ribbon represents ±1 standard error. The horizontal dashed line marks zero difference. Shaded regions indicate the approximate timing of panicle initiation (PI), heading, and R7.

4. Discussion

4.1 GDD Model

The GDD model used to compare temperature sources in this study had the following cardinal temperatures: T_l = 12 °C, T_u = 33 °C, and T_b = 7 °C. These emerged as the best-performing PRISM-based parameterization and provided reasonably accurate predictions of DTH across environments (Fig. 4). The base temperature (T_b) of 7 °C is lower than the 8–10 °C range commonly used in rice phenology models but falls within the range of base temperatures previously reported for rice development (Gao et al., 1987; Yin et al., 1997). In our analysis,

using a T_b of 10 °C as the accepted model without altering T_l and T_u resulted in an RMSE of 4.86, R^2 of 0.28, and a mean bias of 0.16 in the validation dataset (data not shown), which was poorer than the accepted T_b = 7 °C model (RMSE = 4.02, R^2 = 0.51, mean bias = 0.04). The lower T_b relative to other studies may reflect cooler early-season conditions typical of California rice production environments and the use of temperate japonica varieties adapted to cooler temperate climates (Mackill & McKenzie, 2003). In the mid-south of the USA, a T_b of 10 °C is used for rice models (Gao et al., 1987); however, the rice varieties grown there are typically tropical japonica varieties adapted to warmer climates (Mackill & McKenzie, 2003). The lower threshold (T_l = 12 °C) aligns closely with cardinal temperatures reported for rice developmental processes in water-seeded and continuous flooded systems (Gao et al., 1987; Yin et al., 1997). While the upper threshold (Tu = 33 °C) falls within the range typically used in crop simulation models to represent reduced developmental response at high temperatures (Yoshida, 1981; Bouman et al., 2001), it is also similar to that reported by Sharifi et al. (2017) for a California-based GDD model.

4.2 Water temperature versus air temperature

4.2.1 Seasonal changes in In-field water and air temperature

The seasonal differences in water and air temperature (Fig. 2) observed in this study are consistent with previous work in similar environments (Sharifi et al., 2018). Rice field floodwater temperature is strongly impacted by canopy development (affecting solar radiation and wind speed on water surface), as well as floodwater depth (Confalonieri et al., 2005; Kuwagata et al., 2008; Dingkuhn et al., 1995). In a water-seeded system, the field is primarily open water for the first two to three weeks after planting allowing the water to receive the full

amount of full solar radiation. Full canopy cover is not reached until 40 to 50 days which occurs around PI; however, this can also vary depending on plant density (Miller et al., 1991). The impact of canopy cover is clearly seen by comparing In-field water and air T_{max} (Fig. 2). Early in the season when there is open water and limited canopy development, water T_{max} is greater than In-field air T_{max}. However, after about three weeks when the rice is emerging above the floodwater, the difference between water T_{max} and air T_{max} diminishes. At 45 DAP and for the rest of the season when there is complete canopy cover, T_{max} is less than air T_{max} (Fig. 2). While this difference is most striking in T_{max}, it is also reflected in T_{avg} because water and air T_{min} had a relatively consistent difference across the season, with water T_{min} being higher than air T_{min}.

4.2.2 Impact of floodwater on air temperature source

In flooded rice systems, the presence of a floodwater layer modifies surface–atmosphere energy exchange because floodwater has greater thermal inertia and heat capacity than air, which buffers nocturnal cooling and moderates daytime heating (Snyder et al., 2001; Confalonieri et al., 2005; Dingkuhn et al., 1995). For example, Confalonieri et al. (2005) reported that increasing floodwater depth affected air temperature 20 cm above water. In this study, the impact of the floodwater on air temperature is evident by examining the differences between canopy level air temperature (In-field) and PRISM air temperature (Fig. 3). PRISM represents gridded air temperature derived from regional meteorological station observations, whereas the In-field sensors measured canopy-level air temperature within flooded production rice fields and were directly impacted by a floodwater layer. Similar to other reports (Confalonieri et al., 2005; Kuwagata et al., 2008; Sharifi et al., 2018), we found that early in the season, when fields are flooded, and canopy cover is minimal, open water is exposed to direct solar radiation, and In-

field T_{min} (and T_{avg}) is warmer than PRISM T_{min}. This can be attributed to the warmer floodwater during this period (Fig. 2). As the season progressed and the rice canopy covers the water, the water temperatures cool relative to the air. This results in the air In-field T_{min}, T_{max}, and T_{avg} being cooler than the corresponding PRISM temperatures. While the observed differences are not large (T_{avg} differed by roughly -1 to +2 °C throughout the season), these differences compound throughout the season and lead to errors in phenological prediction when integrated in thermal time models (van Oort et al., 2011), discussed below.

4.3 Impact of temperature sources for GDD accumulation

4.3.1 GDD predictions using the model

Using the same GDD model developed using PRISM temperatures, but with In-field air temperature, the model gave similar results in terms of predictability (Fig. 4; Table 2). This is despite the systematic differences in daily T_{min} and T_{max} between these two temperature sources (Fig. 3). The reason model performance for DTH was similar was that early-season differences in temperature (In-field > PRISM) were offset by later-season differences where PRISM temperatures were higher than In-field temperatures (Fig. 3). This is clearly illustrated in Figure 5, where at heading (about 80 DAP), the difference in cumulative GDD accumulation between PRISM- and In-field GDD accumulation is near zero. This type of convergence in cumulative thermal time despite earlier temperature differences has also been reported in previous rice modeling studies (van Oort et al., 2011).

While at the heading stage, the GDD difference between the two temperature data sources was close to zero, there were noted differences at other growth stages (Fig. 5). This stage dependence

is important because while PRISM may produce acceptable estimates of heading, it introduces errors at other developmental stages. For example, PRISM-based accumulation delays the estimated time to PI and advances the estimate of later growth stages such as R7. Using a rough conversion of about 16 GDD accumulation per day, the observed divergence indicates that PI would be shifted one day earlier and R7 about 2 days later when PRISM temperature is used instead of In-field temperature.

Given that the differences in temperature and subsequent GDD accumulation between temperature sources are consistent and predictable, there are ways to improve phenological prediction models based on these findings. First, the relationships observed in this study could be used to model water temperature or canopy-level air temperature from gridded temperature data sources. Water temperature is important, especially in the first half of the season, as the growing apex for rice is below the water surface. Many have recognized its importance for inclusion into phenological models (Kuwagata et al., 2008; Confalonieri et al, 2005; Shimono et al., 2002). Sharifi et al. (2018) compared models in which water or air temperature was used and found that up until PI, using water temperature instead of air temperature improved model accuracy. In-field canopy level air temperature is also critical, especially for the latter stages of development, and our findings suggest that the In-field temperatures could be modeled from gridded temperature data. Second, one could develop stage-dependent thresholds for rice systems as Sharifi et al. (2017) did and find that overall prediction accuracy improved. Doing so indirectly accounts for differences in temperature source as well as potential differences in optimal temperature thresholds between growth stages. Finally, it is recognized that gridded data sets such as PRISM do not fully capture local microclimates (Mankin et al., 2025). They could take into account rice systems and their impacts on microclimate. While possible, there are challenges

to this, such as annual land use changes where rice may be rotated with other upland crops (Salvato et al., 2025) that have a different impact on the microclimate.

While the introduced error is relatively small when predicting DTH, error or variability may be especially relevant for applications focused on early-season rice development, including germination and establishment (Lundy et al., 2014; Driver et al., 2024), as well as for models of insect pest development (Zou et al., 2004). Also, for phenological predictions of later growth stages, such as harvest maturity, there is the potential for much larger errors.

Importantly, many models do not use temperature-bound thresholds as we did here (Yoshida, 1981; Bouman et al., 2001). Under conditions without temperature thresholds (blue line; Fig. 5), the divergence between temperature source was greater. In this study, the differences were largely attributed to T_{max} values being higher than the T_u (33 °C) Many phenological models apply T_u limits to represent reduced developmental response at high temperatures (Bouman et al., 2001; Yin et al., 1997). Thus, when using models without a T_u, thermal time accumulates at a greater rate during hot periods.

4.4 Considerations

The impact of temperature data source on thermal time modeling is likely system- and climate-dependent. Our findings may be unique to the climate and rice production system used in California. However, the general impact of floodwater on the canopy-level microclimate will apply to all rice systems to varying degrees. Given the importance of floodwater for moderating the canopy level microclimate, consideration needs to be given to different water management practices used in rice systems. Water-seeded systems, such as those reported here, employ

different water management practices than other establishment systems used globally, such as drill/dry-seeded rice and transplanted rice systems. In drill/dry-seeded systems, the crop is planted and established similarly to upland crops and does not receive a permanent flood until about a month after planting (Linquist et al., 2015). Thus, the early-season temperature differences observed in this study that are linked to the presence of a flood layer may not be seen in dry-seeded systems. Also, in transplanted rice systems, common to Asia, seedlings are transplanted into floodwater (Ahmed et al., 2025). At the time of transplanting, the rice seedling leaves are above the water surface providing some canopy cover. Also, one must also consider the growing adoption of non-continuous flooding practices such as alternate-wetting-and-drying (AWD), which are practiced to conserve water (Bo et al., 2024), reduce GHG emissions (Qian et al., 2023), and reduce heavy metal concentration in grain (Taaime et al., 2026). These practices vary widely in terms of how long and when the floodwater is on a rice field. Finally, water depth was not considered in this study; however, increasing water depth may strengthen the influence of floodwater on canopy-level air temperatures because deeper water has a greater capacity to buffer temperature fluctuations (Confalonieri et al., 2005).

California's Mediterranean climate is characterized by relatively large diurnal temperature ranges. In this study, the difference in day and night temperatures averaged 15.8 °C for In-field measurements and 17.6 °C for PRISM, which may accentuate temperature-source differences. These large diurnal temperature ranges are comparable rice grown at higher latitudes such as Japan, Korea, northern China, Europe, and Australia; or the Sahel region, (Dingkuhn, 1995; Confalonieri et al., 2005). In other production environments common to rice (i.e. tropical), diurnal temperature ranges are smaller, and the interaction between water temperature, air temperature, and canopy temperature will differ.

Conclusion

Phenological prediction is at the foundation of most crop models, with temperature being a primary driver of phenology. In most models, temperature data come from a nearby meteorological station or a gridded temperature data source. While temperature data from these sources may accurately represent temperature for upland crops, the floodwater in flooded rice systems introduce microclimatic effects that are not fully captured by gridded air temperature datasets. This impact led to predictable differences between In-field canopy level temperature and temperature from a gridded database (PRISM). Using a GDD model, we illustrated how these two temperature sources impacted GDD accumulation through a rice growing season. Due to the presence of open floodwater, GDD accumulated faster early in the season using In-field sensors; however, as the season progressed and floodwater cooled due to the canopy closing, PRISM temperatures were warmer, resulting in faster GDD accumulation. Importantly, these temperature differences should be considered when developing phenological models, particularly for developmental stages before and after heading where divergence in cumulative GDD was greatest. While these findings are specific to California's water-seeded rice system, the principles apply to other rice-growing regions to varying degrees. Future work should evaluate how these relationships vary across contrasting climates, establishment methods, and water management regimes.

References

Ahmed, S., V. Kumar, A.U. Zaman, M.R. Dewan, A. Khatun, K. Hossain, S. Singh, J. Timsina, T.J. Krupnik (2025). Dry direct-seeded and broadcast rice: A profitable and climate-smart alternative to puddled transplanted *aus* rice in Bangladesh. Field Crops Research. https://doi.org/10.1016/j.fcr.2025.109739.

Araghi, A., Martinez, C. J., Olesen, J. E., & Hoogenboom, G. (2022). Assessment of nine gridded temperature data for modeling of wheat production systems. *Computers and Electronics in Agriculture*, *199*, 107189–107189. https://doi.org/10.1016/j.compag.2022.107189

Aslam, M. A., Ahmed, M., Stöckle, C. O., Higgins, S. S., Hassan, F. ul, & Hayat, R. (2017). Can Growing Degree Days and Photoperiod Predict Spring Wheat Phenology? *Frontiers in Environmental Science*, *5*. https://doi.org/10.3389/fenvs.2017.00057

Bo, Y., Wang, X., van Groenigen, K.J., Linquist, B.A., Müller, C., Li, T., Yang, J., Jägermeyr, J., Qin, Y., and Zhou, F. (2024). Improved alternate wetting and drying irrigation increases global water productivity. *Nature Food,* doi.org/10.1038/s43016-024-01081-z

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*. IRRI.

Blankenau, Philip A., et al. "An Evaluation of Gridded Weather Data Sets for the Purpose of Estimating Reference Evapotranspiration in the United States." *Agricultural Water Management*, vol. 242, Dec. 2020, p. 106376, https://doi.org/10.1016/j.agwat.2020.106376.

Buban, M. S., Lee, T. R., & Baker, C. B. (2020). A Comparison of the U.S. Climate Reference Network Precipitation Data to the Parameter-Elevation Regressions on Independent Slopes Model (PRISM). *Journal of Hydrometeorology*, *21*(10), 2391–2400. https://doi.org/10.1175/jhm-d-19-0232.1

C. Wikarmpapraharn, & E. Kositsakulchai. (2010). Evaluation of ORYZA2000 and CERES-rice models under potential growth condition in the central plain of Thailand. *Thai Journal of Agricultural Science*, *43*(1), 17–29.

Confalonieri, R., et al. (2005). "Analysis and Modelling of Water and near Water Temperatures in Flooded Rice (Oryza Sativa L.)." *Ecological Modelling*, vol. 183, no. 2-3, Apr. 2005, pp. 269–280, https://doi.org/10.1016/j.ecolmodel.2004.07.031. Accessed 22 Apr. 2026.

Counce, P. A., Keisling, T. C., & Mitchell, A. J. (2000). A Uniform, Objective, and Adaptive System for Expressing Rice Development. *Crop Science*, *40*(2), 436–443. https://doi.org/10.2135/cropsci2000.402436x

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. *Annals of Applied Biology*, *167*(2), 262–276. https://doi.org/10.1111/aab.12226

CSM-CERES-Rice. (2020, April 16). DSSAT.net. https://dssat.net/csm-ceres-rice/

Daly, C., & Bryant, K. (2013). *The PRISM Climate and Weather System -An Introduction*.

Daly, C., Gibson, W., Taylor, G., Johnson, G., & Pasteris, P. (2002). A knowledge-based approach to the statistical mapping of climate. *Climate Research*, *22*, 99–113. https://doi.org/10.3354/cr022099

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., Curtis, J., & Pasteris, P. P. (2008). Physiographically sensitive mapping of climatological temperature

and precipitation across the conterminous United States. *International Journal of Climatology*, *28*(15), 2031–2064. https://doi.org/10.1002/joc.1688

Dingkuhn, M., et al. (1995). Climatic Determinants of Irrigated Rice Performance in the Sahel — I. Photothermal and Micro-Climatic Responses of Flowering. *Agricultural Systems*, vol. 48, no. 4, pp. 385–410, https://doi.org/10.1016/0308-521x(94)00027-i.

Dingkuhn, M. (1995). Climatic Determinants of Irrigated Rice Performance in the Sahel — III. Characterizing Environments by simulating crop phenology. *Agricultural Systems*: 48:435-456.

Gao, L. Z., Jin, Z. Q., & Li, L. (1987). Photo-thermal models of rice growth duration for various varietal types in China. *Agricultural and Forest Meteorology*, *39*(2-3), 205–213. https://doi.org/10.1016/0168-1923(87)90038-4

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.

Hatfield, J. L., & Prueger, J. H. (2015). Temperature extremes: Effect on plant growth and development. *Weather and Climate Extremes, 10*, 4–10. https://doi.org/10.1016/j.wace.2015.08.001

Islam, M., & Sikder, S. (1970). Phenology and degree days of rice cultivars under organic culture. *Bangladesh Journal of Botany*, *40*(2), 149–153. https://doi.org/10.3329/bjb.v40i2.9770

Jagadish, S. V. K., Murty, M. V. R., & Quick, W. P. (2015). Rice responses to rising temperatures: Challenges, perspectives and future directions. *Plant, Cell & Environment, 38*(9), 1686–1698. https://doi.org/10.1111/pce.12430

Ke, Driver, et al. (2024). "Fascicularis in California Water-Seeded Rice." *Weed Technol*, vol. 38, no. e8, 3 Jan. 2024, pp. 1–5, https://doi.org/10.1017/wet.2024.2.

Kuwagata, Tsuneo, et al. (2008). "Modeling Water Temperature in a Rice Paddy for Agro-Environmental Research." *Agricultural and Forest Meteorology*, vol. 148, no. 11, Oct. 2008, pp. 1754–1766, https://doi.org/10.1016/j.agrformet.2008.06.011.

Liao, S., Xu, X., Xie, H., Chen, P., Wang, C., Jin, Y., Tong, X., & Xiao, C. (2022). A Modified Shape Model Incorporating Continuous Accumulated Growing Degree Days for Phenology Detection of Early Rice. *Remote Sensing*, *14*(21), 5337–5337. https://doi.org/10.3390/rs14215337

Linquist, B.A., R. Snyder, F. Anderson, L. Espino, G. Inglese, S. Marras, R. Moratiel, R. Mutters, P. Nicolosi, H. Rejmanek, A. Russo, T. Shapland, Z. Song, A. Swelam, G. Tindula, and J. Hill. (2015). Water balances and evapotranspiration in water- and dry-seeded rice systems. *Irrigation Science* 33:375-385.

Linquist, B., Al-Khatib, K., Brim-DeForest, W., Espe, M. B., Espino, L., Leinfelder-Miles, M., & Stogsdill, J. R. (2025). Predictors of high rice yields in a high-yielding environment: Lessons from a yield contest. Field Crops Research, 322, 109693. https://doi.org/10.1016/j.fcr.2024.109693

Liu, L.-W., Lu, C.-T., Wang, Y.-M., Lin, K.-H., Ma, X., & Lin, W.-S. (2022). Rice (Oryza sativa L.) Growth Modeling Based on Growth Degree Day (GDD) and Artificial Intelligence Algorithms. *Agriculture*, *12*(1), 59. https://doi.org/10.3390/agriculture12010059

Lu, H., Redus, M. A., Coburn, J. R., Rutger, J. N., McCouch, S. R., & Tai, T. H. (2005). Population Structure and Breeding Patterns of 145 U.S. Rice Cultivars Based on SSR Marker Analysis. *Crop Science*, *45*, 66–76.

Lundy, M.E., et al. (2014). "Site-Specific, Real-Time Temperatures Improve the Accuracy of Weed Emergence Predictions in Direct-Seeded Rice Systems." *Agricultural Systems*, vol. 123, Jan. 2014, pp. 12–21, https://doi.org/10.1016/j.agsy.2013.08.006.

Mackill, D. J., & Lei, X. (1997). Genetic Variation for Traits Related to Temperate Adaptation of Rice Cultivars. *Crop Science*, *37*(4), 1340–1346. https://doi.org/10.2135/cropsci1997.0011183x003700040051x

Mackill, D. J., & McKenzie, K. S. (2003). Origin and characteristics of U.S. rice cultivars. In C. W. Smith & R. H. Dilday (Eds.), *Rice: Origin, history, technology, and production* (pp. 87-100). John Wiley & Sons.

Mankin, K. R., Mehan, S., Green, T. R., & Barnard, D. M. (2025). Review of gridded climate products and their use in hydrological analyses reveals overlaps, gaps, and the need for a more objective approach to selecting model forcing datasets. *Hydrology and Earth System Sciences*, *29*(1), 85–108. https://doi.org/10.5194/hess-29-85-2025

Michler, J. D., Josephson, A., Kilic, T., & Murray, S. (2021). Estimating the Impact of Weather on Agriculture. *The World Bank Policy Research Working Paper*, *9867*.

Miller, B.C., Hill, J.E., & Roberts, S.R. 1991. Plant population effects on growth and yield in water seeded rice. *Agronomy Journal* 83:291-297.

Mourtzinis, S., J.I.R. Edreira, S.P. Conley, P. Grassini. (2017). From grid to field: Assessing quality of gridded weather data for agricultural applications. European Journal of Agronomy 82:163-172. https://doi.org/10.1016/j.eja.2016.10.013.

Qian H., Zhu, X., Huang, S., Linquist, B., Kuzyakov, Y., Wassmann, R., Minamikawa, K., Martinez-Eixarch, M., Yan, X., Zhou, F., Ole Sander, B., Zhang, W., Shang, Z., Zou, J., Zheng, X., Li, G., Liu, Z., Wang, S., Ding, Y., van Groenigen, J.K., Jiang, Y. (2023). Greenhouse gas emissions and mitigation in rice agriculture. *Nature Reviews Earth & Environment*. doi.org/10.1038/s43017-023-00482-1

Parent, B., & Tardieu, F. (2012). Temperature responses of developmental processes have not been affected by breeding in different ecological areas for 17 crop species. *New Phytologist*, *194*(3), 760–774. https://doi.org/10.1111/j.1469-8137.2012.04086.x

R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org/

Roberts, E., Summerfield, R., Ellis, R., & Qi, A. (1993). Adaptation of Flowering in Crops to Climate. *Outlook on Agriculture*, *22*(2), 105–110. https://doi.org/10.1177/003072709302200207

Salvato, L., Pittelkow, C., and Linquist, B.A. (2025). Evaluating agricultural land use change in relation to environmental, biophysical, and economic drivers. *Field Crops Research* https://doi.org/10.1016/j.fcr.2025.110145

Schneider, J., & Ford, D. (2010). Evaluating PRISM Precipitation Grid Data As Possible Surrogates For Station Data At Four Sites In Oklahoma. *Proc. Okla. Acad. Sci*, *90*, 93–104.

Schwartz, M. (2014). *Phenology : an integrative environmental science.* Springer.

Shimono, H., Hasegawa, T., Iwama, K. (2002). Response of growth and grain yield in paddy rice to cool water at different growth stages. Field Crop Res. 73, 67–79.

Sharifi, H., Hijmans, R. J., Hill, J. E., & Linquist, B. A. (2017). Using Stage-Dependent Temperature Parameters to Improve Phenological Model Prediction Accuracy in Rice Models. *Crop Science*, *57*(1), 444–453. https://doi.org/10.2135/cropsci2016.01.0072

Sharifi, H., Hijmans, R. J., Hill, J. E., & Linquist, B. A. (2018). Water and air temperature impacts on rice (Oryza sativa) phenology. *Paddy and Water Environment*, *16*(3), 467–476. https://doi.org/10.1007/s10333-018-0640-4

Smith, E., Grant, N., Luo, X., & Diaz, D. (2025). Evaluating the ability of gridded climate datasets to capture temperature and precipitation trends and extremes. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-97570-7

Snyder, R. L., Spano, D., Duce, P., & C. Cesaraccio. (2001). Temperature data for phenological models. *International Journal of Biometeorology*, *45*(4), 178–183. https://doi.org/10.1007/s004840100103

Stern, M., Flint, L., Flint, A., Boynton, R., Stewart, J., Wright, J., & Thorne, J. (2022). *Selecting the Optimal Fine-Scale Historical Climate Data for Assessing Current and Future Hydrological Conditions*.

Strachan, Scotty, and Christopher Daly (2017). "Testing the Daily PRISM Air Temperature Model on Semiarid Mountain Slopes." *Journal of Geophysical Research: Atmospheres*, vol. 122, no. 11, 7 June 2017, pp. 5697–5715, https://doi.org/10.1002/2016jd025920.

Streck, A., Lago, N. I., Oliveira, B., None A. B. Heldwein, de, A., & Bosco, C. (2011). Modeling the Development of Cultivated Rice and Weedy Red Rice. *Transactions of the ASABE*, *54*(1), 371–384. https://doi.org/10.13031/2013.36234

Summerfield, R. J., Collinson, S. T., Ellis, R. H., Roberts, E. H., & De Vries, F. W. T. P. (1992). Photothermal Responses of Flowering in Rice (Oryza sativa). *Annals of Botany*, *69*(2), 101–112. https://doi.org/10.1093/oxfordjournals.aob.a088314

Surajit K De Datta. (1981). *Principles and practices of rice production.* Wiley.

Taaime, N., S.V.J. Asseko, D. Carrijo, L. Celi, K. Abhishek, M. Martin, M. Romani, S. Rothenberg, A. Seyfferth and B.A. Linquist. (2026). Water management impacts on potentially toxic elements in rice grain: a quantitative review. *Agricultural Water Management*.

Timsina, J., & Humphreys, E. (2006). Performance of CERES-Rice and CERES-Wheat models in rice–wheat systems: A review. *Agricultural Systems*, *90*(1-3), 5–31. https://doi.org/10.1016/j.agsy.2005.11.007

University of California, Davis Agronomy. (2023, August 18). *About California rice*. https://agronomy-rice.ucdavis.edu/about-california-rice

USDA National Agricultural Statistics Service (NASS). (2023). *Rice Acreage and Production Annual Summary.* United States Department of Agriculture, Washington, DC. http://www.nass.usda.gov/

P.A.J. van Oort, Zhang, T., M.E. de Vries, Heinemann, A. B., & Meinke, H. (2011). Correlation between temperature and phenology prediction error in rice (Oryza sativa L.). *Agricultural and Forest Meteorology*, *151*(12), 1545–1555. https://doi.org/10.1016/j.agrformet.2011.06.012

Yang, Zhao, et al. (2017). "Impact of Irrigation over the California Central Valley on Regional Climate." *Journal of Hydrometeorology*, vol. 18, no. 5, 12 Apr. 2017, pp. 1341–1357, https://doi.org/10.1175/jhm-d-16-0158.1.

Yin, X. (1996). Differential Effects of Day and Night Temperature on Development to Flowering in Rice. *Annals of Botany*, *77*(3), 203–213. https://doi.org/10.1006/anbo.1996.0024

Yin, X., Kropff, M. J., Horie, T., Nakagawa, H., Centeno, H. G. S., Zhu, D., & Goudriaan, J. (1997). A model for photothermal responses of flowering in rice I. Model description and

parameterization. *Field Crops Research*, *51*(3), 189–200. https://doi.org/10.1016/s0378-4290(96)03456-9
Yin, X., Kropff, M. J., Nakagawa, H., Horie, T., & Goudriaan, J. (1997). A model for photothermal responses of flowering in rice II. Model evaluation. *Field Crops Research*, *51*(3), 201–211. https://doi.org/10.1016/s0378-4290(96)03457-0
Yin, X., & Kropff, M. J. (1998). The effect of photoperiod on interval between panicle initiation and flowering in rice. *Field Crops Research*, *57*(3), 301–307. https://doi.org/10.1016/s0378-4290(98)00074-4
Yoshida, S. (1981). *Fundamentals of rice crop science*. International Rice Research Institute.
Zhang, S., & Tao, F. (2013). Modeling the response of rice phenology to climate change and variability in different climatic zones: Comparisons of five models. *European Journal of Agronomy*, *45*, 165–176. https://doi.org/10.1016/j.eja.2012.10.005
Zhang, T., Zhu, J., & Yang, X. (2008). Non-stationary thermal time accumulation reduces the predictability of climate change effects on agriculture. *Agricultural and Forest Meteorology*, *148*(10), 1412–1418. https://doi.org/10.1016/j.agrformet.2008.04.007
Zou, L., Stout, M., & Ring, D. (2004). Degree-day models for emergence and development of the rice water weevil (Coleoptera: Curculionidae) in southwestern Louisiana. Environmental Entomology, 33 (6), 1541-1548. https://doi.org/10.1603/0046-225X-33.6.1541

Supplemental Material

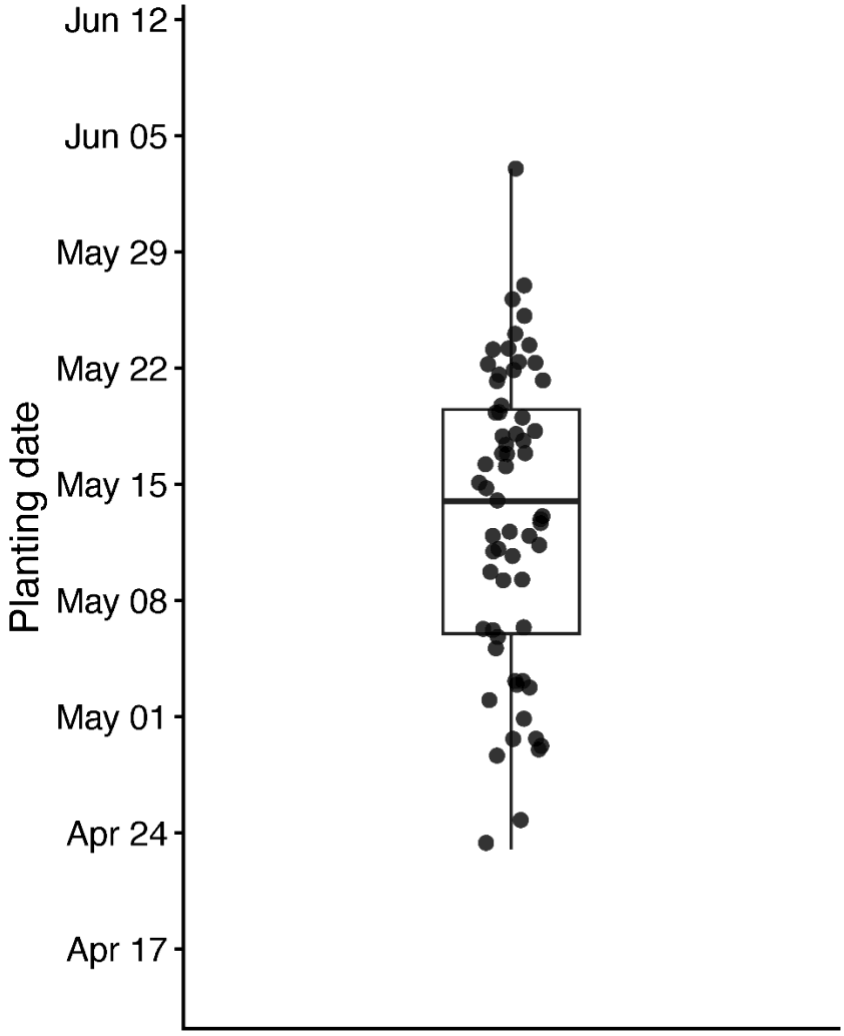


Figure S1. Distribution of planting dates across site-years (2016–2024) after exclusion of environments removed from the analysis (see Methods). Points represent individual site-year observations, and the boxplot shows the median and interquartile range. This figure provides calendar-date context for interpreting days-after-planting (DAP)-based temperature and growing degree day (GDD) analyses.

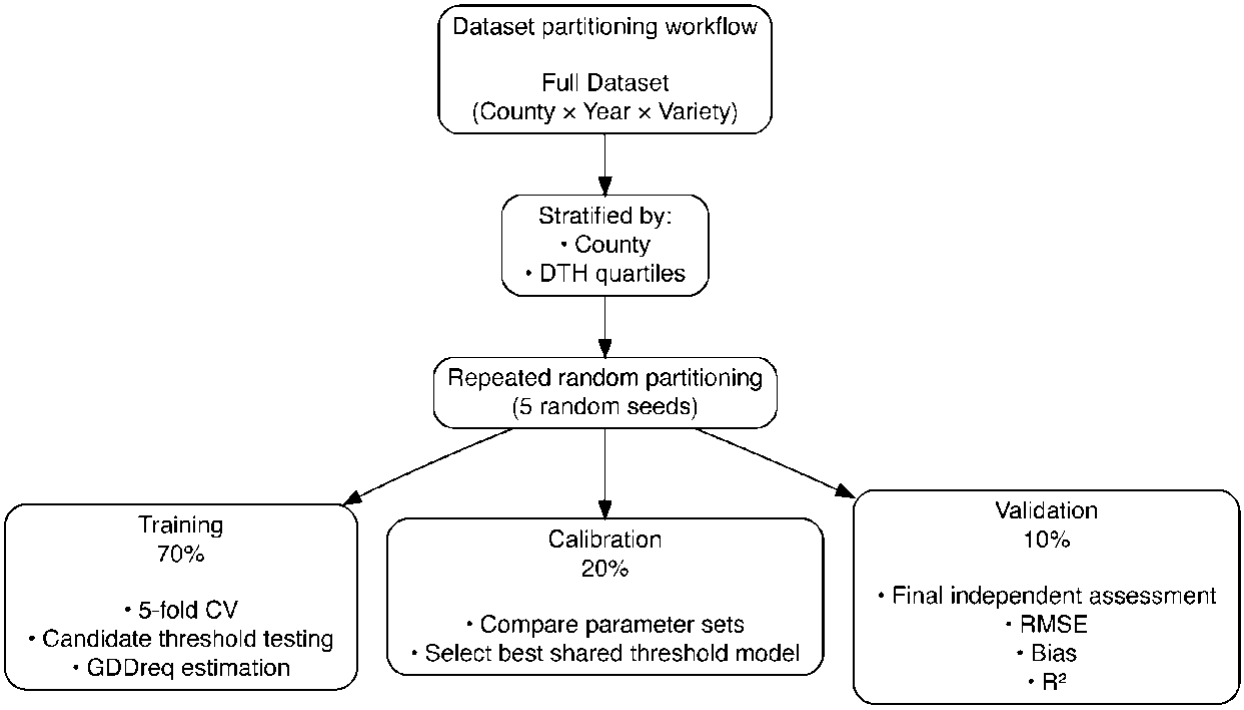


Figure S2. Workflow used for model development and evaluation. The full dataset was stratified by county and quartiles of observed days to heading (DTH) and partitioned into 70% training, 20% calibration, and 10% validation subsets. This partitioning was repeated across five random seeds to evaluate the robustness of model selection and prediction performance. The training subset was used for model development, including internal five-fold cross-validation of candidate temperature threshold parameterizations and estimation of variety-specific growing degree day (GDD) requirements. The calibration subset was used to compare candidate parameterizations and identify the best-performing shared threshold model. The validation subset was held out from model development and used for the final independent assessment of prediction accuracy. Validation results are presented in Figure 4.

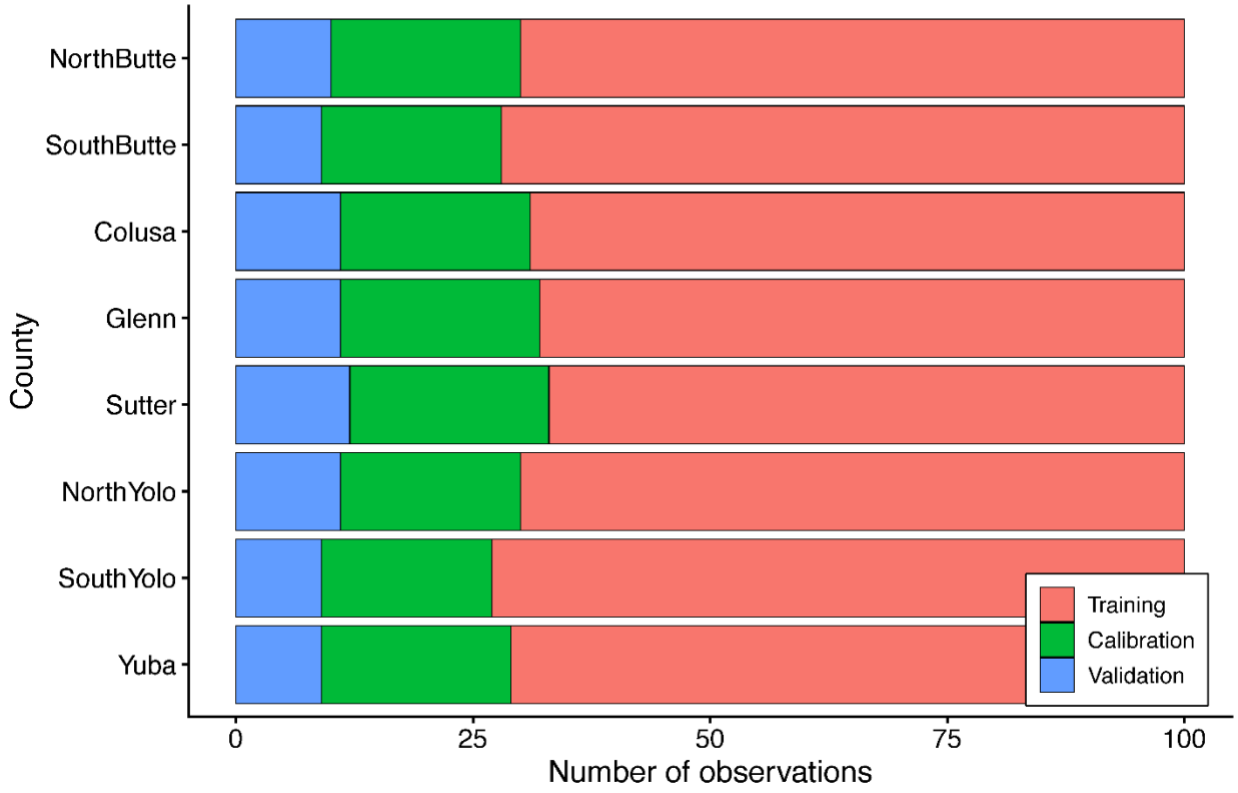


Figure S3. Example distribution of observations across counties for one representative random seed, showing the 70% training, 20% calibration, and 10% validation partitioning used for model development and evaluation. Stratification by county and quartiles of observed DTH ensured balanced representation of environments across subsets and maintained independence of the final validation dataset.

Table S1. Summary statistics of cumulative Growing Degree Days (GDD, °C·d) to heading for the five main Calrose medium-grain varieties using the uncapped base-only formulation (T_b = 7 °C). In-field uses sensor temperatures and PRISM uses gridded temperatures. n = number of variety-specific heading observations across environments. All GDD values (Min, Mean, Median, Max, and SD) are reported in °C·d.

Variety	Source	n	Min	Mean	Median	Max	SD
M105	In-field	62	1199	1377	1378	1529	67
M105	PRISM	62	1197	1417	1430	1627	77
M206	In-field	62	1214	1414	1408	1567	66
M206	PRISM	62	1217	1458	1460	1663	79
M209	In-field	62	1279	1511	1505	1695	76
M209	PRISM	62	1281	1561	1570	1767	86
M210	In-field	62	1199	1417	1421	1550	66
M210	PRISM	62	1197	1461	1472	1682	79
M211	In-field	60	1298	1514	1509	1712	80
M211	PRISM	60	1300	1565	1569	1813	90