

# A comparison of phenology models for harvest readiness in walnuts

*Elise Hellwig<sup>1</sup>, Katherine Pope<sup>2</sup>, Robert J. Hijmans<sup>1</sup>*

*1. University of California, Davis*

*2. Cooperative Extension, University of California*

## Introduction

Phenology models are used in research and farm operations to predict the timing of plant life stages, such as emergence, flowering, and maturity. For example, phenology models estimate development rates for crop growth models that simulate the effect of varieties, management, and the environment on growth and production. In addition to their use in research, farmers use in-season phenological model predictions in tactical decision-making to prepare for the logistics associated with flowering and harvest, such as pesticide application and equipment rental for harvest.

While both temperature and photoperiod can drive plant development, here we focus on temperature driven phenology models. Temperature based models use thermal time, instead of chronological time, to measure plant development. Thermal time is also commonly referred to as a “heat sum”. Thermal time summarizes how plants’ rates of development change based on the ambient temperature surrounding the plant.

In these models, thresholds are used to predict when a plant reaches a particular phenological stage, like flowering. We distinguish between two broad types of models based on the unit of this threshold value: heat threshold (HT) and time threshold (TT) models.

HT models use thresholds in thermal units. The thresholds are used to determine the thermal time required to complete the developmental stage. When the heatsum threshold is reached, the model reports the number of days it took to reach that threshold. In the base model, this day count is exactly the stage length. While this type of model is useful in research, it provides little in the way of predictive value for farm management. We can develop an extended version of the model if we use the number of days it takes to reach the heat threshold is used to predict the stage length, but do not require it to be the stage length itself. Because the extended model can produce predictions earlier in the season it may be more useful for in season decision-making.

The threshold for TT models is in time units, most commonly days. The base time threshold model is quite simple. Since the threshold is the length of the stage, and the threshold is in days, the base TT model simply constitutes the mean stage length value. This ‘model’ is almost never used in on a year to year basis. However, farmers use mean season length estimates from agricultural companies and extension offices to ensure crosspollination between tree cultivars and varieties that will mature within the local growing season. Because of this, we will use it as a worst-case scenario for stage length prediction.

If we extend the TT model, the resulting model is a mirror image of the extended HT model. The model counts the number of thermal units accumulated before the number of days, set by the threshold, have past. Next, the model uses the number of thermal units accumulated this way to predict the length of the stage, in days. This approach is commonly used in fruit and nut phenology models to make in season predictions for farm orchard management. The use of TT models is supported by findings that temperature regime during the first 30 to 60 days after flowering has much larger effects on fruit development than conditions during the rest of the maturation period [Mimoun and DeJong, 1998, Day et al., 2007, DeBuse et al., 2008, Tombesi et al., 2010, Ruml et al., 2011]

One key component of both model types is the method of converting temperature data into thermal time. This is typically done by using temperature thresholds called cardinal temperatures. The models in this paper use base ( $T_b$ ), optimal ( $T_o$ ), and critical temperatures ( $T_c$ ). Thermal time is zero below the base temperature and above the critical temperature. For Mediterranean fruit and nut trees, a typical value for  $T_b$  is 4°C and

for  $T_c$  is 36°C. Thermal time is highest at the optimum temperature, which may be around 25°C for tree crops [Richardson et al., 1975]. Growth chamber experiments with different temperature treatments can determine cardinal temperature values. However, this work is expensive and is rarely done.

While some effort has been made to compare different TT phenological models [e.g. Marra et al., 2001, Tombesi et al., 2010], the performance of TT and HT models have not been compared. In this paper, compare the predictive power of 8 TT and 14 HT phenology models for predicting harvest readiness in walnuts. We use optimization methods to estimate both the amount of thermal time accumulation and the cardinal temperatures used to calculate thermal time. The resulting models can be used for in season walnut harvest date prediction.

## Phenology model types

### Heatsum Threshold (HT) Model

The base HT model assumes that the model threshold represents the total amount of thermal time, the heatsum, the plant in question requires to complete a particular developmental stage. We call this value  $\Lambda_j$ , where  $j$  is the development stage. Then the number of days required to accumulate  $\Lambda_j$  thermal time is  $x_i(\Lambda_j)$ , for a given year or growing season  $i$ . The stage length, in days, is  $y_i$ , this means that

$$y_{ij} = x_i(\Lambda_j) + \varepsilon_i \quad (1)$$

where  $\varepsilon_i$  is the associated error term. We refer to equation (1) as the base model, because it contains the fewest possible elements while still remaining an HT model. This is due to the assumption that the heatsum represents the total amount of thermal time the plant needs to accumulate to complete a given growth stage. If we relax that assumption and allow the plant to reach the threshold prior to the end of the development stage, we observe two complications.

First, because the heatsum threshold is not necessarily the total amount of thermal time necessary for the plant to complete the stage, the number of days to accumulate the thermal time will be lower than the stage length. Let  $\lambda_j$  be the heatsum threshold in the case where the heatsum threshold does not necessarily equal the total thermal time necessary for the plant to complete the stage. Then  $x_i(\lambda_j)$ , is the number of days necessary to accumulate  $\lambda_j$  thermal time units. If

$$\frac{\lambda_j}{\Lambda_j} < 1$$

then  $x_i(\lambda_j)$  will be less than the total stage length,  $y_{ij}$ . To compensate for this we will need to include a conversion factor,  $\beta_{j\lambda}$ , to convert days taken to reach the model threshold to days in the stage. In principle,

$$\beta_{j\lambda}$$

should equal  $\frac{\Lambda_j}{\lambda_j}$ . However, because of the difficulty in estimating  $\Lambda_j$ ,  $\beta_j$  is estimated numerically using ordinary least squares regression.

Second, if the heatsum threshold is zero, the plant will reach the threshold immediately, causing  $x_i$  to be zero as well. This would make the stage length zero, regardless of how large the conversion factor  $\beta_{j\lambda}$  is. To compensate for this we include a second parameter,  $\alpha_{j\lambda}$ , the stage length if the plant reaches the threshold before completing a day of the stage. While this is biologically impossible, it is necessary to consider mathematically because of the way the model is

If we add these two parameters to equation (1) we get the ordinary linear model (equation 2), giving the extended HT model its name.

$$y_{ij} = \alpha_{j\lambda} + \beta_{j\lambda} \cdot x_i(\lambda_j) + \varepsilon_i \quad (2)$$

In fact, we can consider equation (1) one of the most basic types of linear model, where  $\alpha_j = 0$  and  $\beta_j = 1$ . Consequently, we refer to the HT model that uses equation (1) as the base HT model.

## Time Threshold (TT) Model

In the TT model, the threshold is measured in time, not thermal time. If we assume that the time threshold is the total number of days needed for the plant to complete the stage, temperature and thus thermal time becomes irrelevant. In the base TT model, the stage length is exactly the time threshold (equation 3).

$$y_{ij} = \Lambda_j + \varepsilon_i \quad (3)$$

where  $y_{ij}$  is the stage length,  $\Lambda_j$  is time threshold, and  $\varepsilon_i$  is the error term. This is equivalent to the most basic linear model with an intercept but no predictor variables, which is also equivalent to the grand mean. This means we can write the base TT model as

$$y_{ij} = \bar{y}_j + \varepsilon_i$$

The extended TT model has essentially the same structure as equation (2). The only thing that differs is the interpretations of the variables. The variable  $\lambda_j$  is the threshold in days, instead of thermal time units, and  $x_i(\lambda_j)$  is the amount of thermal time accumulated in  $\lambda_j$  days. Like before,  $y_{ij}$  is the length, in days, of stage  $j$ . Then,  $\alpha_{j\lambda}$  is the predicted stage length if  $x_i$  is zero. In this case that would mean the trees did not experience any thermal time accumulation during the first  $\lambda_j$  days of stage  $j$ . and  $\beta_{j\lambda}$  is the change in stage length due to accumulating one additional thermal time unit during the first  $\lambda_j$  days of the stage.

## Materials and Methods

There are four steps in specifying a phenology model for in season prediction of harvest readiness or another growth stage: (1) selection of the type of model including whether the base or extended form is used; (2) selection of the function used to derive thermal time from temperature (see below); (3) estimating the cardinal temperature and thermal time accumulation parameters; and (4) fitting of a statistical model that predicts the length in days ( $y$ ) of growth stage of interest from time or thermal time accumulation.

### Thermal time computation functions

We evaluated five functions for calculating thermal time. While there are additional functions [Marra et al., 2001], these five capture a variety of assumptions about tree responses to temperature and are computationally viable options for model fitting on most computers. The simplest “Linear” thermal time function has just one parameter [Yang et al., 1995].

$$\text{Linear TT} = \begin{cases} 0 & T \leq T_b \\ T - T_b & T_b \leq T \end{cases}$$

where  $TT$  is the thermal time accumulated,  $T$  is the average temperature for the chosen time period (day or hour), and  $T_b$  is the base temperature. We evaluated the Linear function for both hourly and daily temperatures (LinearHour and LinearDay).

The “Plateau” function is a slightly more complicated function that includes two parameters.

$$\text{Plateau TT} = \begin{cases} 0 & T \leq T_b \\ T - T_b & T_b \leq T \leq T_o \\ T_o & T_o \leq T \end{cases}$$

where  $T_o$  is the optimal temperature. The Plateau model caps thermal time at  $T_o$ , but still neglects a potential slowdown in development at very high temperatures.

One function that accounts for this is the three parameter triangle function from, where  $T_b$ ,  $T_o$ , and  $T_c$  are the base, optimal and critical temperatures respectively.

$$\text{Triangle TT} = \begin{cases} 0 & T \leq T_b \\ T - T_b & T_b \leq T \leq T_o \\ \frac{(T_c - T)(T_o - T_b)}{T_c - T_o} & T_o \leq T \leq T_c \\ 0 & T_c \leq T \end{cases}$$

A widely used function in tree crop phenology is another three parameter model, called the Asymcur model, from [Anderson et al., 1985].

$$\text{Asymcur TT} = \begin{cases} 0 & T \leq T_b \\ \frac{T_o - T_a}{2} \left[ 1 + \cos \left( \pi + \pi \cdot \frac{T - T_b}{T_o - T_b} \right) \right] & T_b \leq T \leq T_o \\ (T_o - T_a) \left[ 1 + \cos \left( \frac{\pi}{2} + \frac{\pi}{2} \cdot \frac{T - T_o}{T_c - T_o} \right) \right] & T_o \leq T \leq T_c \\ 0 & T_c \leq T \end{cases}$$

As this function is almost always used with the cardinal temperatures reported by Anderson et al. [1985], we use these as well to have a basis of comparison to other studies ( $T_b = 4$ ,  $T_o = 25$ ,  $T_c = 36$ ). We refer to the Asymur function with cardinal temperatures from Anderson et al. [1985] as the Anderson function.

One thermal time function uses daily minimum and maximum temperatures data to calculate thermal time accumulation. The widely used Growing Degree Day (GDD) function has only one parameter, the base temperature  $T_b$  [Zalom and Goodell, 1983, McMaster and Wilhelm, 1997].

$$\text{MinMax TT} = \begin{cases} 0 & T_{max} \leq T_b \\ \frac{(T_{max} - T_b)^2}{2(T_{max} - T_{min})} & T_{min} \leq T_b \leq T_{max} \\ T - T_{avg} & T_{min} \geq T_b \end{cases}$$

Where  $T_{avg}$  is the mean of the daily minimum and maximum daily temperature and  $T_b$  is the base cardinal temperature.  $T_{avg}$  is sometimes replaced with the minimum daily temperature ( $T_{min}$ ) or the maximum daily temperature ( $T_{max}$ ) [Ruml et al., 2011]. A simpler GDD function is sometimes used as well [Snyder et al., 1999]. While there are other models used in the literature [Marra et al., 2001], these five capture a variety of assumptions about tree responses to temperature, and are all computationally feasible.

## Data

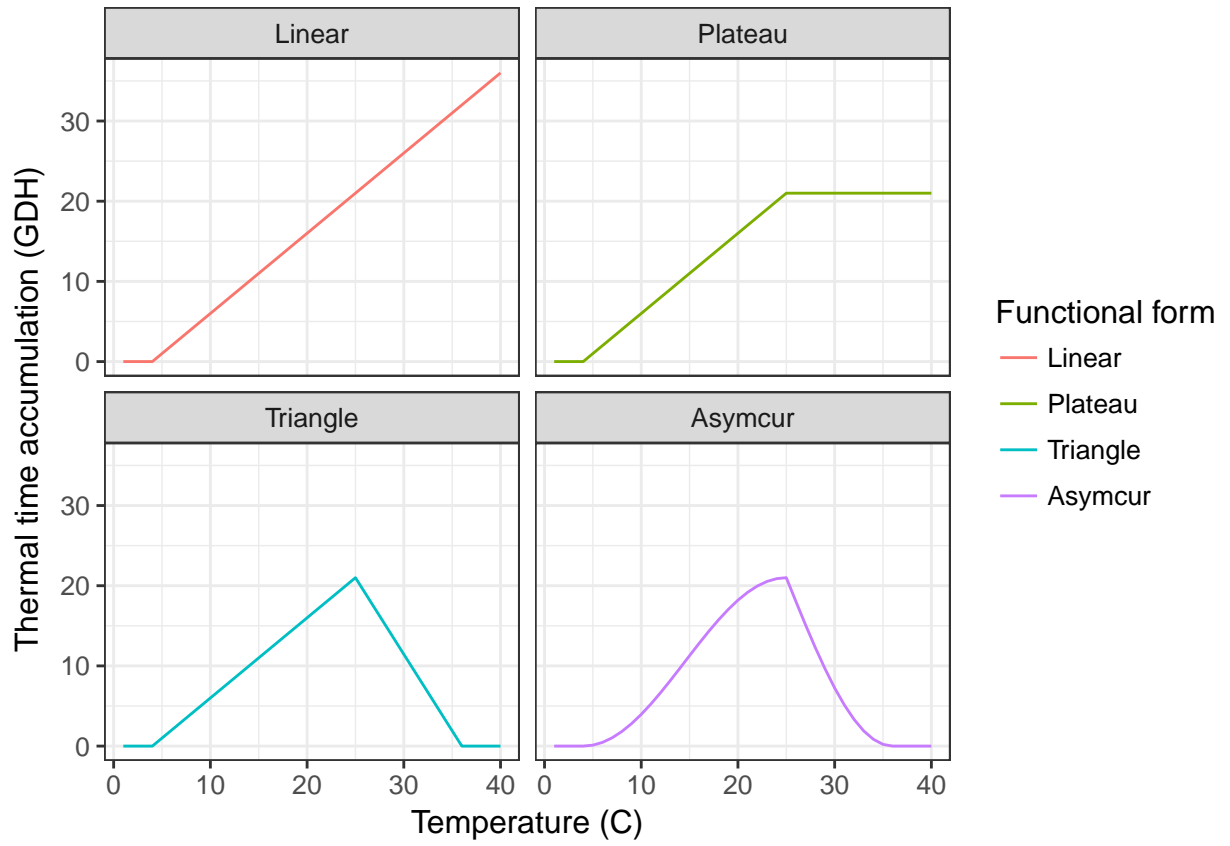


Figure 1: Visual representations of thermal time functions. For all of the functions  $T_b = 4^\circ\text{C}$ ,  $T_o = 25^\circ\text{C}$ , and  $T_c = 36^\circ\text{C}$ .

Table 1: Walnut phenology data summary. Leaf-out and harvest are given in day of the year, and season length is given in days. The mean value for each measure is displayed, with the first and third quartile values in parentheses to the right. All data is from the UC Davis Walnut Breeding Program.

Cultivar	Years with data	Season Length	Leaf Out	Harvest
Amigo	35	173 (167, 178)	89 (85, 92)	262 (258, 268)
Ashley	42	185 (180, 192)	77 (73, 80)	262 (257, 269)
Chandler	43	184 (178, 190)	94 (91, 98)	278 (274, 282)
Chico	44	183 (177, 190)	80 (76, 83)	263 (256, 270)
Franquette	45	181 (176, 185)	106 (101, 112)	287 (283, 291)
Hartley	50	184 (179, 190)	93 (91, 95)	278 (273, 283)
Howard	43	178 (174, 182)	93 (90, 96)	270 (266, 275)
Payne	52	184 (179, 190)	77 (74, 80)	261 (258, 267)
Serr	47	185 (180, 192)	78 (74, 80)	263 (258, 268)
Tehama	41	181 (176, 186)	88 (83, 92)	269 (265, 274)
Tulare	37	180 (175, 186)	92 (88, 95)	271 (267, 276)
Vina	59	181 (174, 187)	86 (82, 90)	267 (262, 272)

We used walnut phenology data from the University of California at Davis Walnut Breeding Program. Leaf out dates (LD) and Harvest readiness dates (HRD) were collected by Charles Leslie, Gale McGranahan and the members of the Walnut Improvement Program for a group of twelve walnut varieties over 35 to 59 years depending on variety (Table 1).

We used temperature data from NCDC and CIMIS weather stations in Davis, and nearby Winters and Woodland [Menne et al., 2015, California Department of Water Resources, 2015] Whenever possible, Davis temperature data was used. However, there were several gaps in the combined NCDC CIMIS time series. To fill these, linear models that related temperatures in Woodland and Winters to Davis were used to create predictions for the missing days, which were then averaged.

## Statistical methods

All statistical analysis was conducted using the *R* environment for statistical computing [R Core Team, 2017]. We used the following model and thermal time function combinations:

Table 2: Model and thermal time function combinations investigated in this study. TT - Day Threshold, HT - Thermal Time Threshold.

Model	Type	Thermal Time Function
HT	Full	LinearHour
HT	Full	Plateau
HT	Full	Triangle
HT	Full	Asymcur
HT	Full	LinearDay
HT	Full	MinMax
HT	Simplified	LinearHour
HT	Simplified	Plateau
HT	Simplified	Triangle
HT	Simplified	Asymcur
HT	Simplified	LinearDay
HT	Simplified	MinMax

Model	Type	Thermal Time Function
TT	Full	LinearHour
TT	Full	Plateau
TT	Full	Triangle
TT	Full	Asymcur
TT	Full	LinearDay
TT	Full	MinMax
TT	Simplified	NA

Parameter optimizations of both cardinal temperatures and threshold values were conducted using the DEoptim implementation [Ardia et al., 2015] of a Differential Evolution optimization method detailed in [Price et al., 2006]. The value minimized was the root mean squared error (RMSE) of the ordinary linear model relating thermal time or days accumulated to season length.

We chose to use the RMSE as our measure of error for several reasons. First, unlike the Mean Squared Error (MSE), the units of the RMSE are the same as the units of the dependent variable, which is useful for interpretation. Additionally the RMSE weights large errors greater than smaller ones. This is important because, when predicting harvest dates, models become useless if they can produce very large errors. Finally, the RMSE is continuous, which is important for optimizing parameter values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (4)$$

This optimization was conducted for each model, thermal time function, and cultivar combination. Five-fold cross-validation was used to estimate the accuracy of phenology models for out of sample data. Errors from these optimized models were compared based on the type of model and the function used to estimate thermal time. The model with the lowest cross-validated RMSE for each cultivar was selected and used to create a model to predict walnut season length in the Davis area.

In order to understand these models more fully, we created optimization landscapes of the model parameters. First, we selected a range of thermal time and day thresholds ranging from one to the maximum possible number allowed by the model. We then optimized cardinal temperatures for each of these thresholds and compared the accompanying RMSE values. This was done for all possible combinations of model, thermal time function, and cultivar.

Next, both the TT and HT models were evaluated using the LinearDay thermal time function, and error values recorded for every possible value of threshold and base temperature. This was repeated with the Flat and Triangle thermal time functions as well. All of these used the extended model type when evaluating the linear regression. We also calculated the variance in the predicted season lengths for all possible time and heatsum threshold values to further explore why the optimal threshold parameter values took on the values they did.

## Results

Walnut trees in Davis rarely experience temperatures below 0 °C during periods of fruit set and maturation. However temperatures in excess of 40 °C are not uncommon (Figure 2). Additionally, while maximum temperature vary widely between months (18°C in March and 34°C in July), the minimum temperatures stay more stable (6°C in March and 13°C in July) This means that the trees are exposed to a much greater range of temperatures in the summer than they do in the spring or fall.

Average leaf-out dates ranged from March 18 to April 16, and walnuts reached maturity between September 12th and October 18th. Mean season lengths ran from 173 days to 185 days (Table 1).

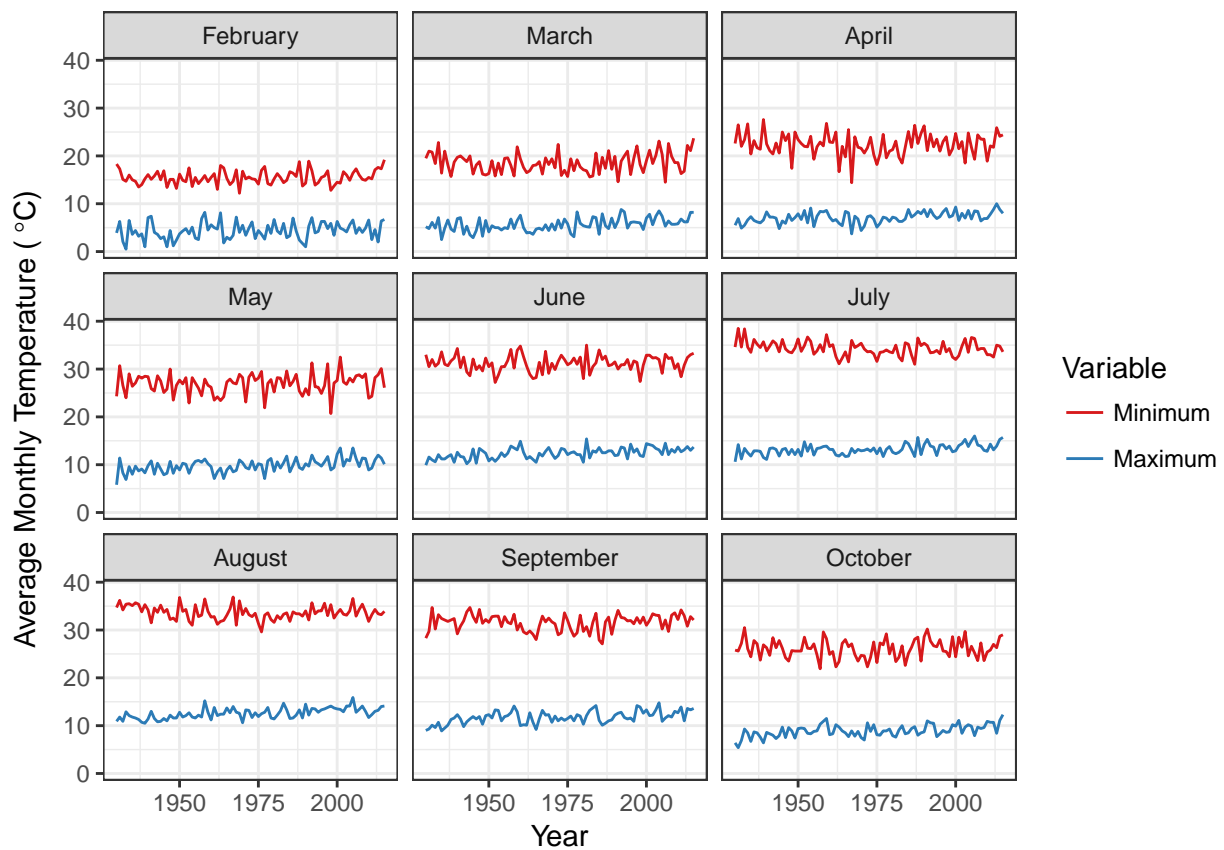


Figure 2: Average minimum and maximum monthly temperatures in Davis during the months relevant to walnut maturation, and the years included in the UC Davis Walnut Breeding Program dataset. The monthly averages were calculated using the temperature time series developed for this paper.



## Model Comparison

The extended models were more accurate than base models (RMSE 1.45 days fewer), regardless of model type (Figure 3, Table 3). Specifically, they better captured the variance in the season lengths from the original data than the base models did (Figure 4). For both the TT and HT models, variance in predicted season length was much lower than variance in actual season length. Additionally, that variance decreased the later in the season the predictions were made. Because base models predict season length at the end of the season, the predicted season variance was generally at its lowest. This meant the discrepancy between the variance in predicted and observed season length was at its greatest.

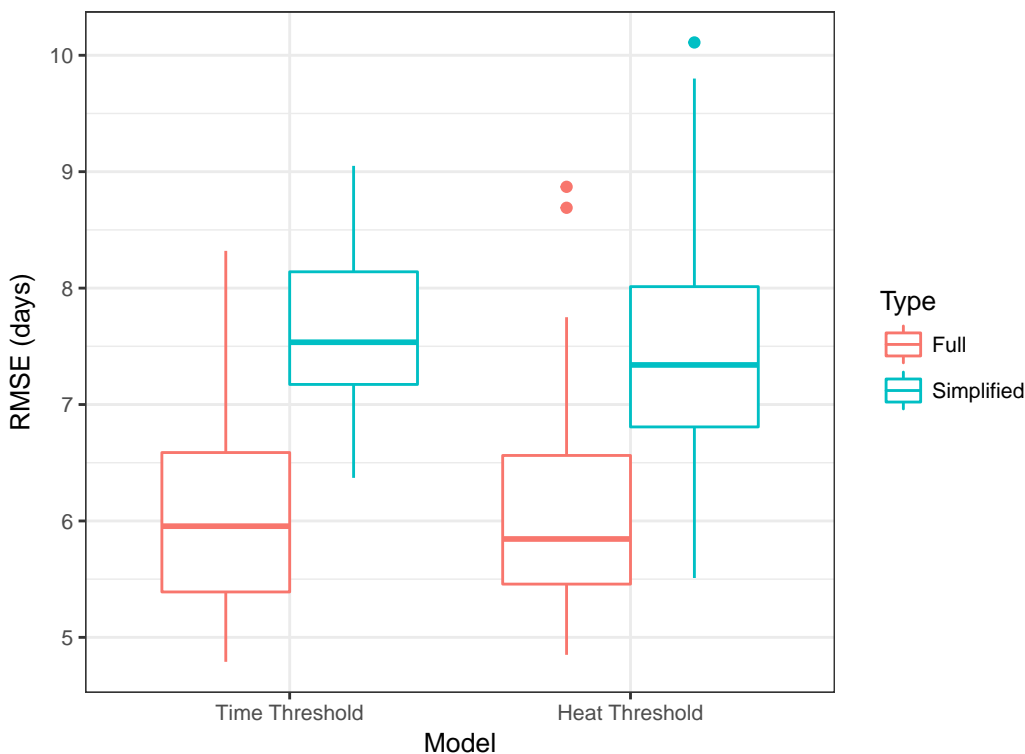


Figure 3: A comparison of model accuracy of by type. RMSE distributions reflect differences in error values between cultivars.

Many extended HT and TT models had RMSE values below 6, though that performance was cultivar dependent (Figure 5). The HT model paired with the LinearHour thermal time function had a very compact distribution of RMSE values, and the LinearDay and Asymcur functions also had error distributions with low variance. However, the LinearHour function only had an average performance when predicting season length for ‘Chandler’, the most commonly planted walnut cultivar. Chandler was the only cultivar best predicted by a base model. However, the difference between that and the extended model was minimal, only 0.06 days (Figure 5, Table 3, S1).

Table 3: Optimal model by cultivar. RMSE determined by 5-fold crossvalidation.

Cultivar	Model	Type	Function	RMSE
Amigo	Heat Threshold	Full	LinearHour	6.31
Ashley	Heat Threshold	Full	LinearDay	5.34
Chandler	Heat Threshold	Simplified	Plateau	7.05
Chico	Time Threshold	Full	Plateau	4.81
Franquette	Time Threshold	Full	Anderson	5.04

Cultivar	Model	Type	Function	RMSE
Hartley	Heat Threshold	Full	Plateau	6.34
Howard	Time Threshold	Full	LinearDay	5.13
Payne	Time Threshold	Full	LinearDay	5.11
Serr	Heat Threshold	Full	LinearHour	5.02
Tehama	Time Threshold	Full	LinearDay	4.79
Tulare	Heat Threshold	Full	Anderson	5.79
Vina	Heat Threshold	Full	LinearHour	5.87

## Optimization evaluation

### Model threshold and base temperature

Of all the parameters, the RMSE responded most to changes in threshold value. This is reflected in the fact that the base models, which essentially set the accumulation length or threshold to be constant, tended to have much higher RMSE values than the extended models (Figure 3). The majority of cultivars had a clear optimum time thresholds that was not affected by thermal time function (Figure 6). Most time threshold values were around 50 days after leaf out. Additionally, the more flexible functions, like the triangle function, tended to have lower, but more variable, RMSE values. Many of the HT models did have an optimum threshold or threshold range early on in the season. However, these error minima were generally wider, and less marked than their TT counterparts (Figure 7).

Of the cardinal temperatures, the base temperature influenced model fit the most, regardless of cultivar or thermal time function (Figures 8-9). Additionally, comparing parameter optimization across functions showed that the threshold parameter as well as the base temperature were frequently quite close together across functional forms. On the other hand, the optimal and critical cardinal temperatures tended to vary widely (Table S1).

The relationship between threshold value, base temperature, and RMSE value is similar between the TT and HT models (Figure 8-9). However, for HT models, not all combinations of model threshold and base temperature produce models where a tree can reach the heatsum threshold. Because of this, changes in RMSE values for HT models based on changes to model threshold or base temperature are never entirely independent of each other (Figure 9).

### Optimal and critical temperatures

For models with optimal and critical temperatures, it is clear that as long as those cardinal temperatures are above some “reasonable value”, approximately 15°C to 25°C for the optimal temperatures and 40°C for critical temperatures, they had very little effect on RMSE (Figures S1-S2). This is in stark contrast to the threshold and base temperature parameters which very frequently had a narrow range of optimal values.

These cardinal temperatures varied widely between cultivars, as well as between models of the same cultivar (Figures 10-11, Table S1). Additionally, many cultivars had cardinal temperatures that were estimated to be within a degree of each other making the visual depiction of them flat (For example, Chandler Asymcur). This is not biologically plausible, though it may tell us something about tree development. This problem is also less pronounced in the HT models (Figure 11).

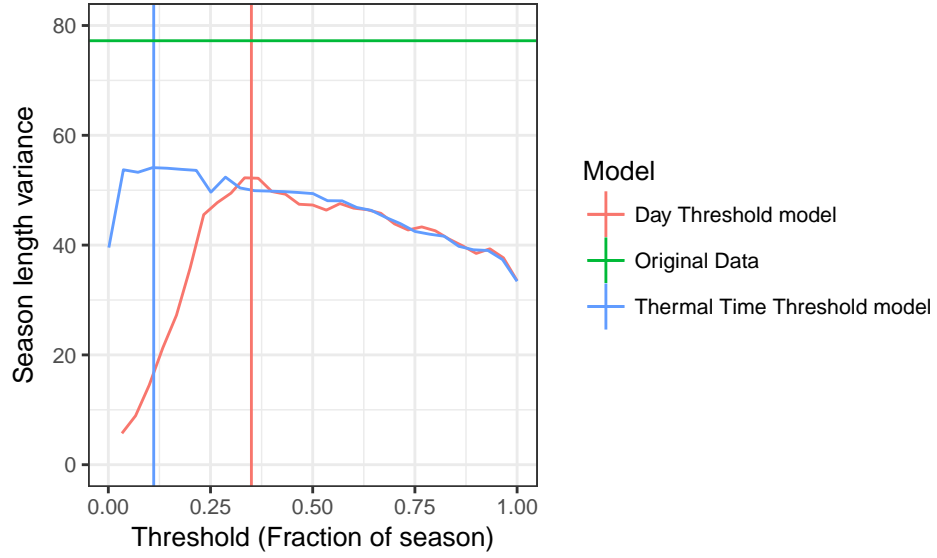


Figure 4: Variances in predicted season length values with both day and thermal threshold models using linear thermal time accumulation for Payne walnuts. The vertical lines mark optimal day threshold value of 63 days or 35% through the season, and the optimal thermal time threshold value of 9335 degree hours or 11% through the season. The horizontal line marks the variance (77.2) in the season lengths of the original data.

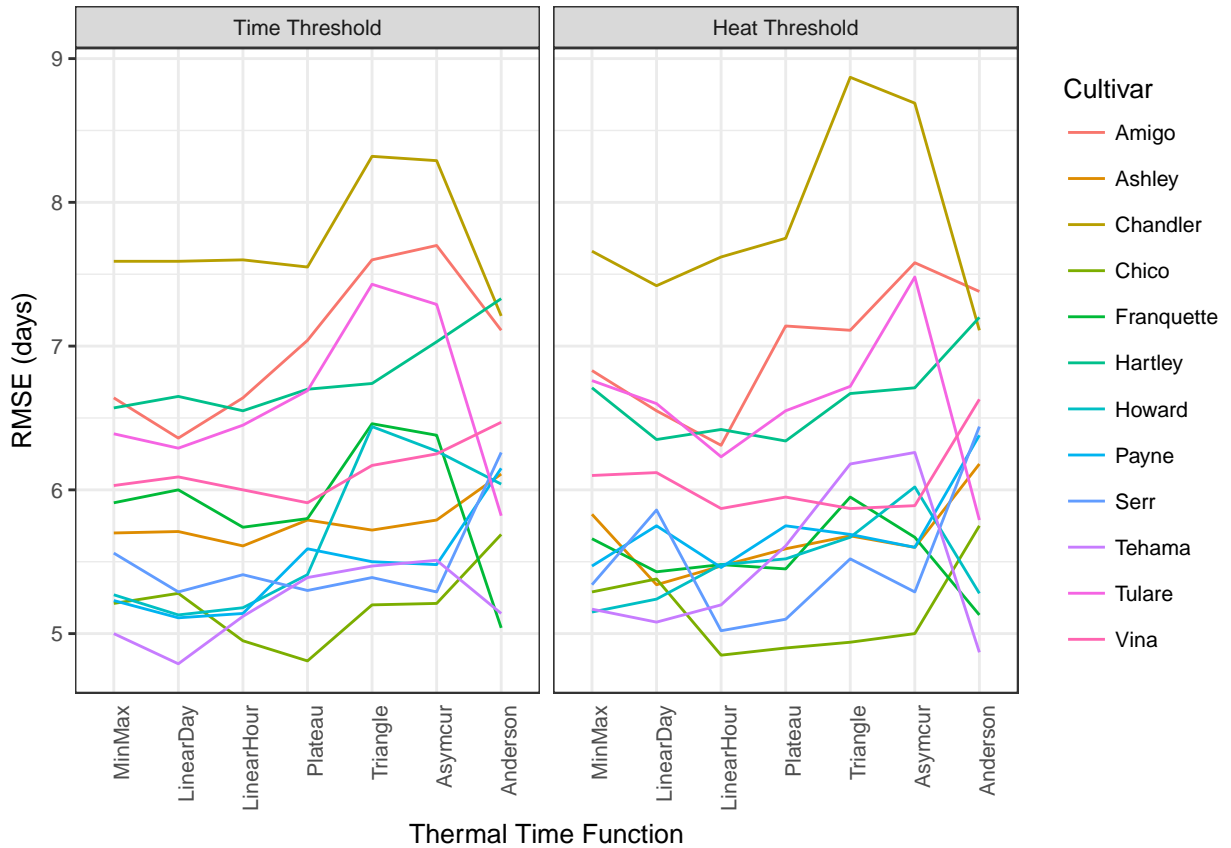


Figure 5: A comparison of RMSE values for full models by thermal time function and cultivar.

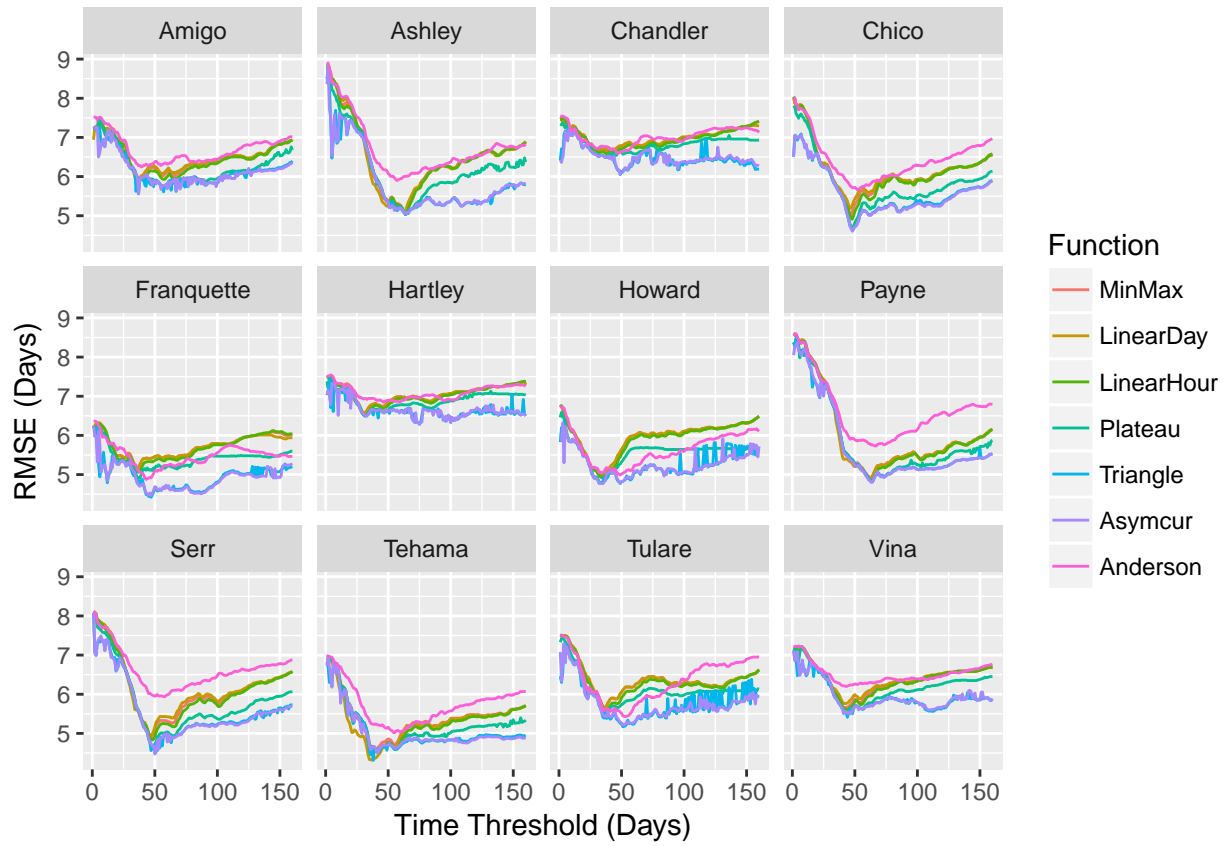


Figure 6: A comparison of the accuracy, assessed using RMSE, of all possible day threshold parameter values for the Day Threshold Model by cultivar and function.

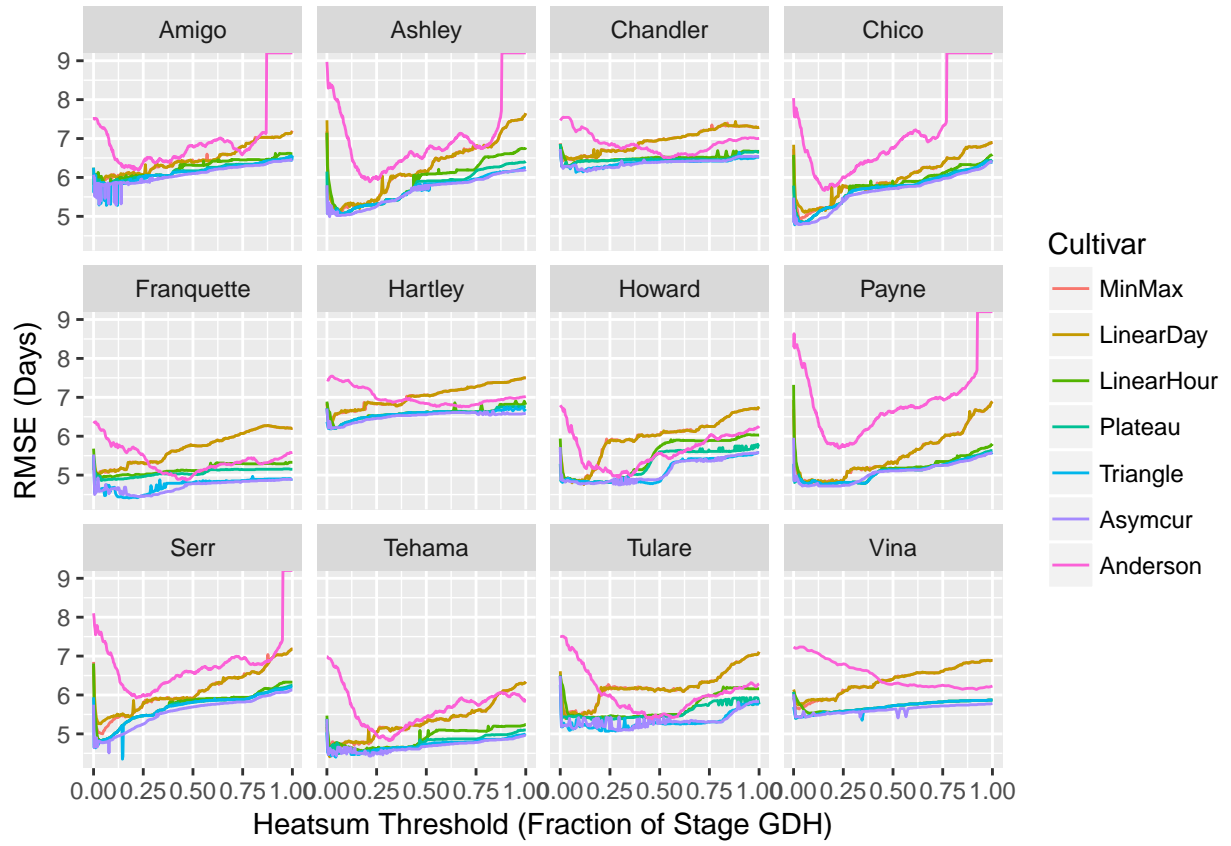


Figure 7: Comparison of accuracy of thermal time threshold parameter values by cultivar and function.

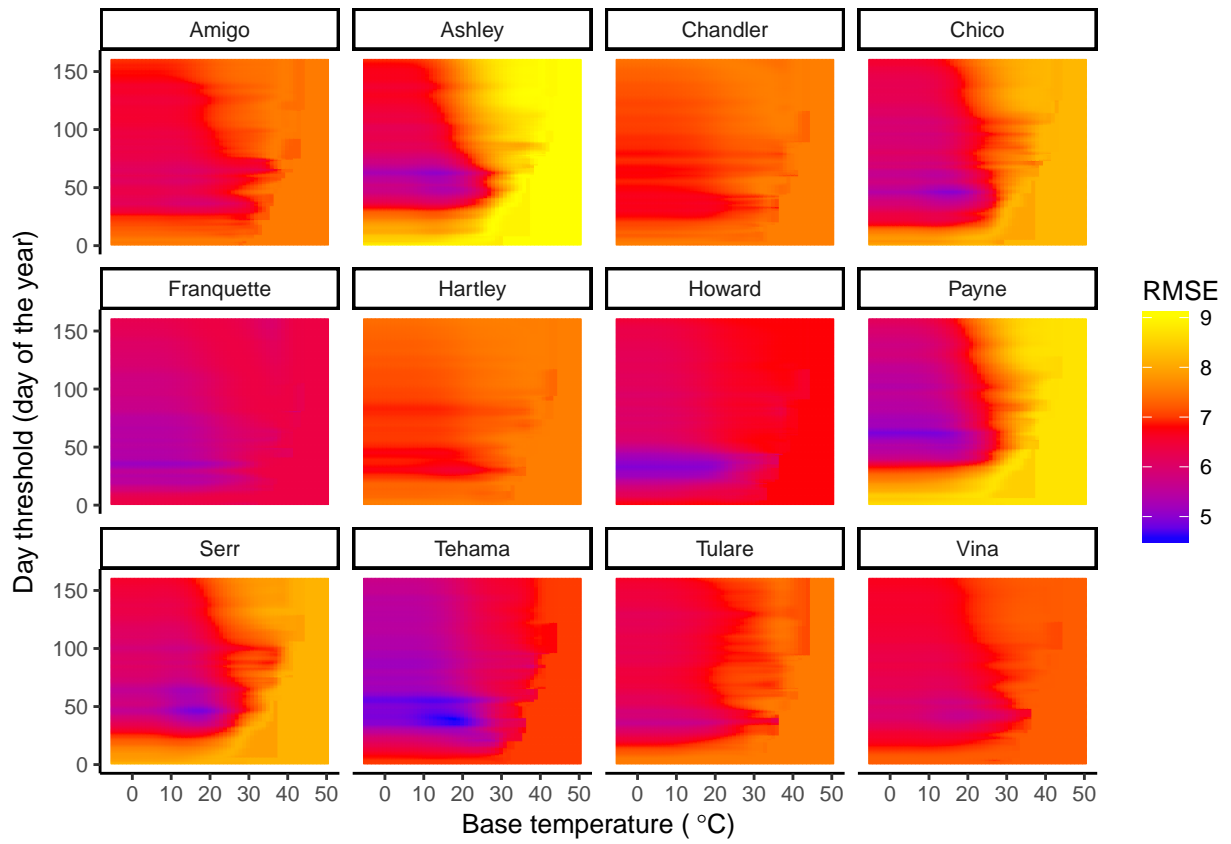


Figure 8: The RMSE optimization surface for the day threshold and base temperature parameters of Day Threshold model using the MinMax function. Results are separated by cultivar.

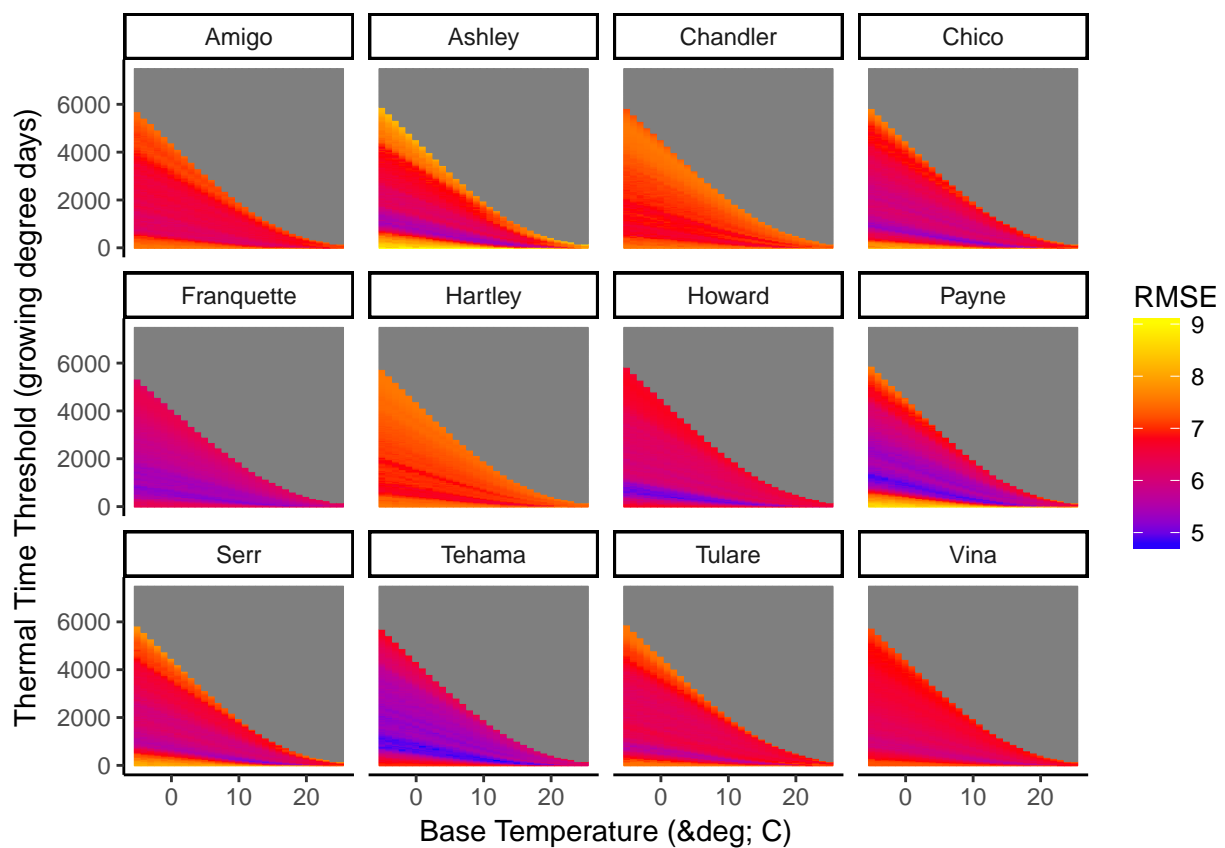


Figure 9: The RMSE optimization surface of the thermal time threshold and base temperature parameters in the Thermal Time Threshold model using the MinMax function. Results separated by cultivar. Grey indicates an invalid model.

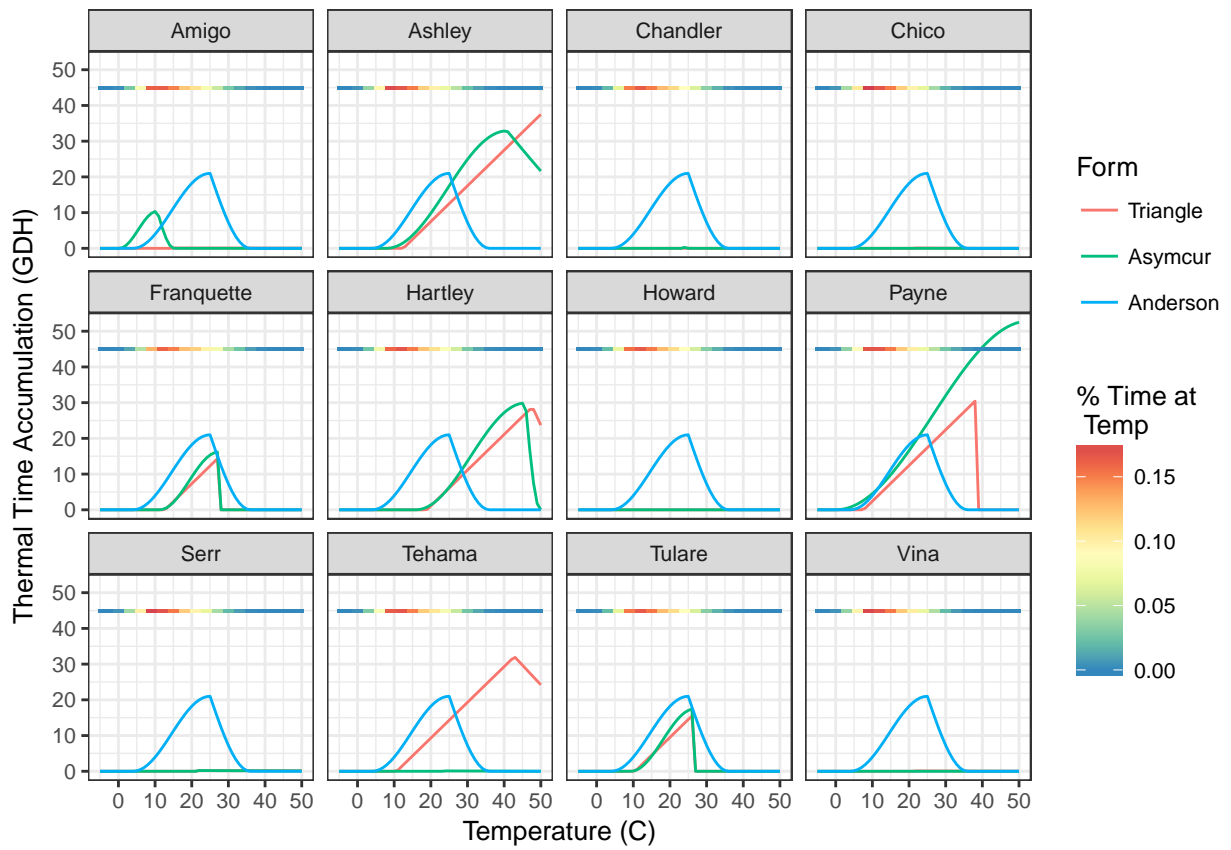


Figure 10: Comparison of different thermal time functional forms along with what temperatures the trees experience most, Day Threshold model.



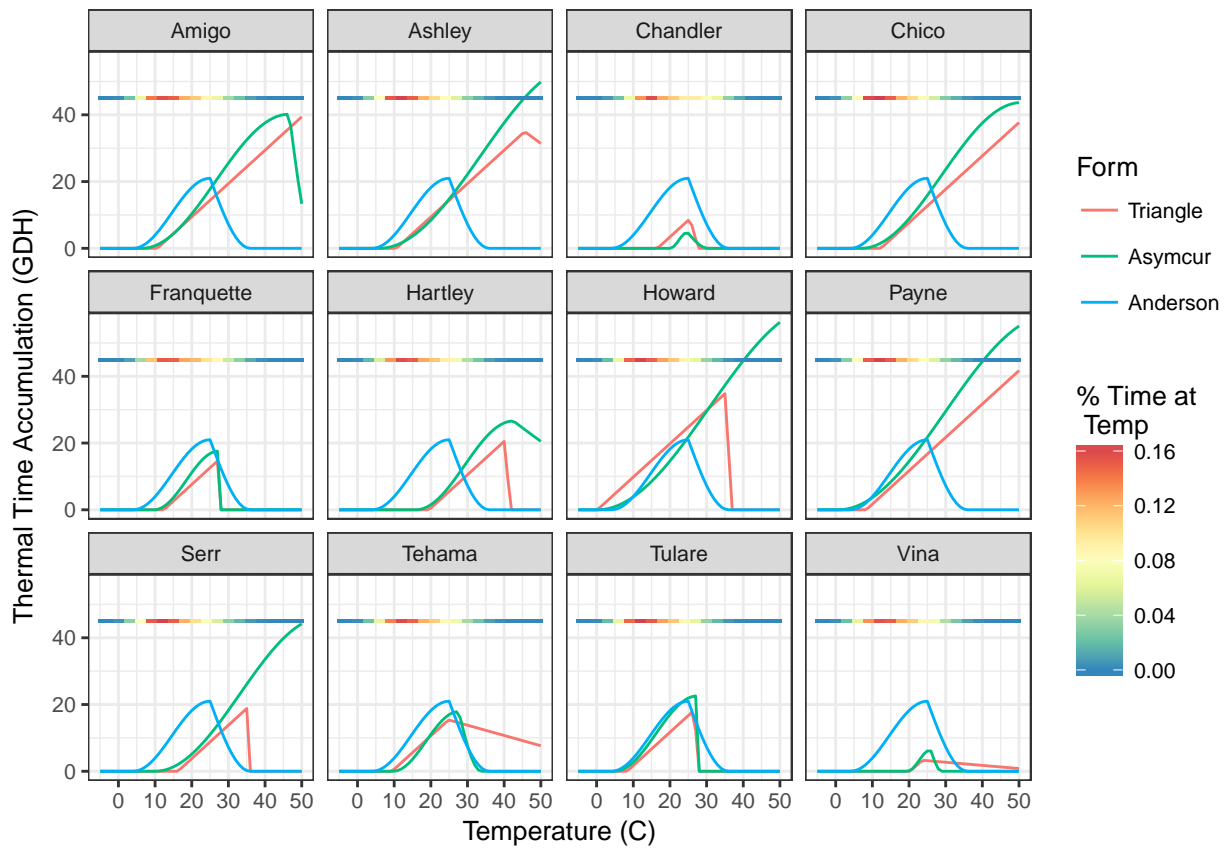


Figure 11: Comparison of different functional forms of thermal time accumulation along with what temperatures the trees experience most, Full model.

# Discussion

## Model complexity

Using the extended model was more accurate than using the base model in almost all cases. This is because the variation in predicted season length values is greatest early on in crop development. Predicted season lengths from a linear regression will always have lower variance than observed season lengths, because of a statistical phenomenon called regression towards the mean. However, some models may capture the season length variation more effectively than others.

This is the case with the base and extended models. At the beginning of the season, variation in predicted season lengths increases as the amount of thermal time or days accumulated increases from zero to a non-zero value. At some point the variance in predictions hits a peak. This corresponds to the optimal threshold value. Then, as the season advances, the amount of variation in the predictions decreases. It is only in cases where the change in variance is very minimal that, the base models performed well. Biologically speaking, this seems to indicate that walnut development is especially sensitive to temperatures at the beginning of the season but less so later in the summer, at least under the conditions of our experiments.

The idea that predicting the season length early on in the season would yield better predictions than predicting it later in the season may be counterintuitive. However, this approach is informed by our understanding of walnut biology. Walnuts have two main stages of development: cell division and differentiation, and cell enlargement. During the cell division stage, the walnut creates all the cells it will need for the rest of its development. Then during the cell enlargement stage the tree pumps carbohydrates, fats, and other nutrients into the nut [Ramos, 1997]. It is generally thought that fruit and nut growth rates respond more to temperature during the cell division stage than during the cell enlargement stage [Warrington et al., 1999, Bertin, 2005, Zhang et al., 2006]. So, predicting season length early in the season may enable us to focus our maximize the proportion of our data that comes from the highly temperature sensitive cell division stage.

Ideally, we would be able to identify the end of the cell division stage and only count thermal time during that period. Unfortunately, it is very difficult, as well as labor intensive to collect this data, so as of yet there is no time series of cell division end dates available for walnuts.

## Model Optimization

Overall, the simplest thermal time functions performed the best. In particular, the LinearHour function used in conjunction with the HT model produced reasonable results for all cultivars and is computationally simple to estimate. However, it does require hourly temperature estimates. If these are impossible to obtain, they can be interpolated using formulas from [Cesaraccio et al., 2001]. The more flexible functions, Triangle and Asymcur, were sensitive to overfitting and should be used with caution. The Anderson function performed worse than some of the simpler models for many of the cultivars. However, in situations where cardinal temperature optimization is not possible, it can present a viable option.

Optimization clearly did not identify the biologically meaningful cardinal temperatures, especially in the TT models. One notable example was the estimation of base and optimal temperatures within 1°C of each other in the TT models. This meant thermal time accumulation was extremely low. However, as the thresholds were also very low, this did not seem to unduly affect prediction, and the models with these parameters performed about as well as models with more reasonable cardinal temperatures.

Another noteworthy issue was the instability of estimated optimal and critical cardinal temperatures. We suspect this is due to the fact that this dataset does not include data from trees experiencing warm enough temperatures early enough in the season to accurately estimate these parameters using empirical methods. Since we found that the temperatures early in the walnuts development seemed to have a disproportional effect on their development, it is likely that even though these trees may have experienced 40°C weather on occasion, it was late enough in the season, not to slow down nut development. Incorporating walnut data from areas with different climates could help remedy this problem.

## Model Type

The TT and HT models are roughly equivalent in terms of accuracy. However, their biological underpinnings are quite different. The TT model assumes that the determining factor in when various phenological stages begin or end is the day of the year, determined by photoperiod. The HT model assumes that the ambient temperature the tree experiences has a greater effect in determining when phenological stages start and end than photoperiod does. Clearly, which assumption is appropriate will depend on what crop is being modeled.

## Model Accuracy

The UC Davis Walnut Breeding Program uses hull-split as their indicator of harvest readiness. It is easy to assess and accurate enough for the purposes of analyzing variety characteristics. However, it is also affected by water and nitrogen availability, which can vary significantly from year to year [Goldhamer and Viveros, 2000]. Growers most commonly use an indicator called packing tissue brown (PTB), which correlates very strongly with nut quality and color [Sibbett et al., 1974]. Unfortunately packing tissue brown is much more labor intensive to assess and so consequently not normally used in breeding programs, where exact harvest dates are not as critical. Because of this, model errors are higher than would be expected for data sets with packing tissue brown as the measure of harvest date.

Model error values vary widely by cultivar. Unfortunately, the most commonly grown cultivar, Chandler, is also the hardest to predict. It is unclear why this is the case. One possibility is that Chandler walnuts are more affected by factors like humidity, that vary somewhat independently of harvest readiness but do affect hull split.

## Conclusions

For most cultivars, predicting season length (and harvest date) early in the season yields the most accurate results. This points towards a model of walnut development that is more responsive to temperature early on in the process. However, physiological studies are needed to confirm this. TT and HT models performed similarly for walnuts, so any decision would need to be made on a basis of underlying biological knowledge. The more accurate models of thermal time accumulation did not perform better than the less accurate ones, and often suffered from overfitting. The key to realizing most of the reduction in model error is to fit either an extended TT or HT model with a relatively simple thermal time function.

## Supplemental Plots

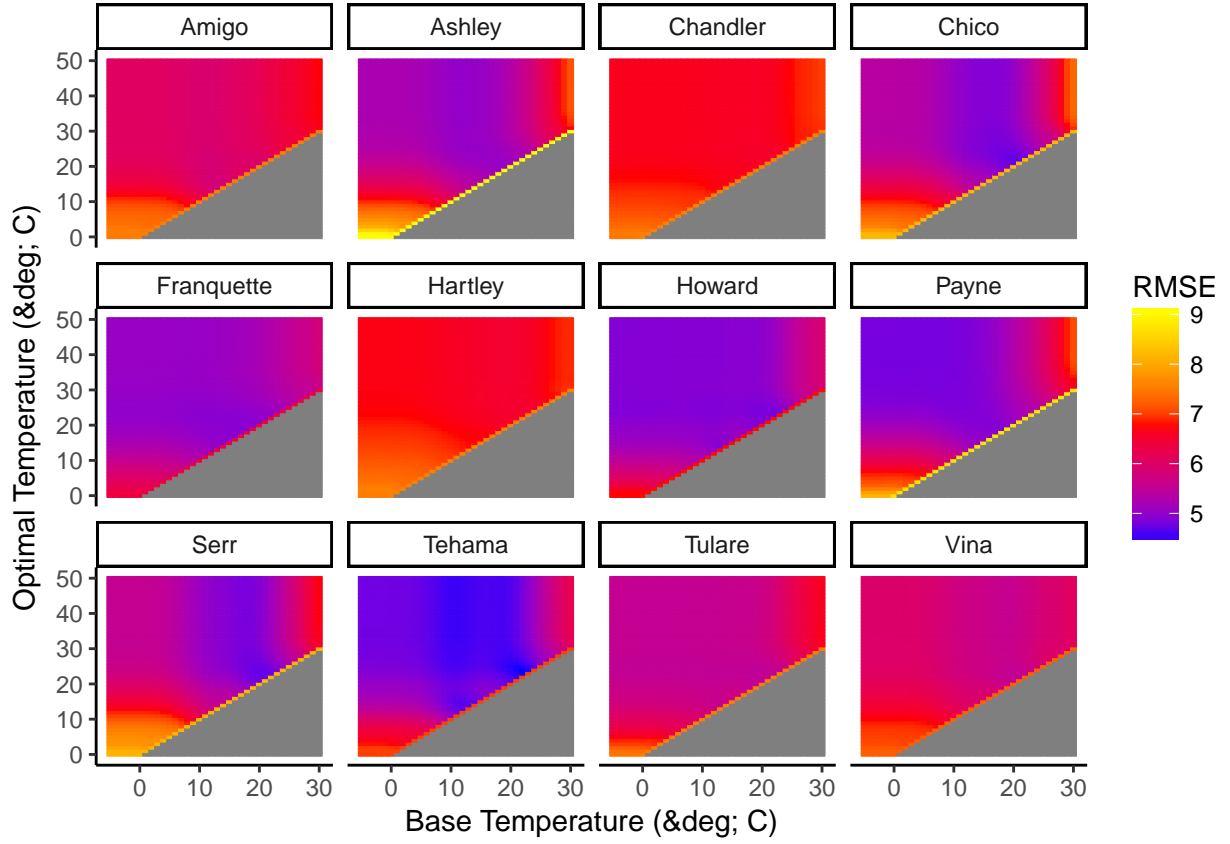


Figure 12: S1: Optimization surface for the Day Threshold Model model using Flat function by cultivar. Grey indicates an invalid model.

Table 4: S3: Cardinal temperatures for all models. RMSE values are from Five fold cross-validation.

Cultivar	Type	Model	Function	Length	Base	Opt	Crit	RMSE
Amigo	Full	Time Threshold	Anderson	39.00	4	25	36	7.11
Amigo	Full	Time Threshold	MinMax	38.00	18.3			6.64
Amigo	Full	Time Threshold	LinearDay	37.00	16.3			6.36
Amigo	Full	Time Threshold	LinearHour	37.00	10.7			6.64
Amigo	Full	Time Threshold	Plateau	57.00	14.9	15		7.04
Amigo	Full	Time Threshold	Triangle	57.00	14.9	15	100	7.60
Amigo	Full	Time Threshold	Asymcur	37.00	0.2	10.6	15.2	7.70
Ashley	Full	Time Threshold	Anderson	57.00	4	25	36	6.11
Ashley	Full	Time Threshold	MinMax	63.00	13.1			5.70
Ashley	Full	Time Threshold	LinearDay	63.00	12.3			5.71
Ashley	Full	Time Threshold	LinearHour	64.00	12.5			5.61
Ashley	Full	Time Threshold	Plateau	64.00	12.5	38.6		5.79
Ashley	Full	Time Threshold	Triangle	64.00	12.5	76.3	88.3	5.72
Ashley	Full	Time Threshold	Asymcur	64.00	7.9	40.8	82.2	5.79
Chandler	Full	Time Threshold	Anderson	47.00	4	25	36	7.21
Chandler	Full	Time Threshold	MinMax	38.00	17.4			7.59

Cultivar	Type	Model	Function	Length	Base	Opt	Crit	RMSE
Chandler	Full	Time Threshold	LinearDay	31.00	14.7			7.59
Chandler	Full	Time Threshold	LinearHour	38.00	18			7.60
Chandler	Full	Time Threshold	Plateau	37.00	21.9	22		7.55
Chandler	Full	Time Threshold	Triangle	49.00	23.3	23.4	24.6	8.32
Chandler	Full	Time Threshold	Asymcur	49.00	23.1	23.5	25.4	8.29
Chico	Full	Time Threshold	Anderson	52.00	4	25	36	5.69
Chico	Full	Time Threshold	MinMax	47.00	15.2			5.21
Chico	Full	Time Threshold	LinearDay	47.00	12.5			5.28
Chico	Full	Time Threshold	LinearHour	48.00	16.3			4.95
Chico	Full	Time Threshold	Plateau	48.00	21.7	21.7		4.81
Chico	Full	Time Threshold	Triangle	48.00	21.6	21.7	38.6	5.20
Chico	Full	Time Threshold	Asymcur	48.00	21.7	21.7	45.6	5.21
Franquette	Full	Time Threshold	Anderson	43.00	4	25	36	5.04
Franquette	Full	Time Threshold	MinMax	37.00	8.2			5.91
Franquette	Full	Time Threshold	LinearDay	36.00	12.2			6.00
Franquette	Full	Time Threshold	LinearHour	37.00	8.4			5.74
Franquette	Full	Time Threshold	Plateau	37.00	18.9	19.1		5.80
Franquette	Full	Time Threshold	Triangle	47.00	12.7	27.3	27.3	6.46
Franquette	Full	Time Threshold	Asymcur	47.00	11.2	27.4	27.4	6.38
Hartley	Full	Time Threshold	Anderson	46.00	4	25	36	7.33
Hartley	Full	Time Threshold	MinMax	31.00	18.2			6.57
Hartley	Full	Time Threshold	LinearDay	30.00	14.7			6.65
Hartley	Full	Time Threshold	LinearHour	32.00	18.9			6.55
Hartley	Full	Time Threshold	Plateau	32.00	18.9	93.8		6.70
Hartley	Full	Time Threshold	Triangle	32.00	18.9	47.7	60.7	6.74
Hartley	Full	Time Threshold	Asymcur	32.00	15.9	45.8	49.9	7.03
Howard	Full	Time Threshold	Anderson	49.00	4	25	36	6.04
Howard	Full	Time Threshold	MinMax	34.00	1.3			5.27
Howard	Full	Time Threshold	LinearDay	34.00	1.6			5.13
Howard	Full	Time Threshold	LinearHour	35.00	1.2			5.18
Howard	Full	Time Threshold	Plateau	35.00	22.3	22.3		5.41
Howard	Full	Time Threshold	Triangle	34.00	22.8	22.8	49.7	6.44
Howard	Full	Time Threshold	Asymcur	34.00	22.7	22.7	60.7	6.27
Payne	Full	Time Threshold	Anderson	67.00	4	25	36	6.15
Payne	Full	Time Threshold	MinMax	63.00	0			5.23
Payne	Full	Time Threshold	LinearDay	62.00	11.1			5.11
Payne	Full	Time Threshold	LinearHour	63.00	7.4			5.14
Payne	Full	Time Threshold	Plateau	63.00	0	38.2		5.59
Payne	Full	Time Threshold	Triangle	63.00	7.6	38.2	38.8	5.50
Payne	Full	Time Threshold	Asymcur	63.00	0	52.9	78.7	5.48
Serr	Full	Time Threshold	Anderson	57.00	4	25	36	6.26
Serr	Full	Time Threshold	MinMax	48.00	17.1			5.56
Serr	Full	Time Threshold	LinearDay	47.00	14.5			5.29
Serr	Full	Time Threshold	LinearHour	48.00	18.2			5.41
Serr	Full	Time Threshold	Plateau	49.00	21.6	21.8		5.30
Serr	Full	Time Threshold	Triangle	49.00	21.6	21.8	100	5.39
Serr	Full	Time Threshold	Asymcur	49.00	21.6	21.8	100	5.29
Tehama	Full	Time Threshold	Anderson	55.00	4	25	36	5.14
Tehama	Full	Time Threshold	MinMax	40.00	18.1			5.00
Tehama	Full	Time Threshold	LinearDay	37.00	15			4.79
Tehama	Full	Time Threshold	LinearHour	42.00	10.7			5.12
Tehama	Full	Time Threshold	Plateau	42.00	21.9	22		5.39

Cultivar	Type	Model	Function	Length	Base	Opt	Crit	RMSE
Tehama	Full	Time Threshold	Triangle	42.00	10.7	42.8	72.1	5.47
Tehama	Full	Time Threshold	Asymcur	42.00	23.6	23.7	68.7	5.51
Tulare	Full	Time Threshold	Anderson	53.00	4	25	36	5.82
Tulare	Full	Time Threshold	MinMax	37.00	0			6.39
Tulare	Full	Time Threshold	LinearDay	36.00	11.4			6.29
Tulare	Full	Time Threshold	LinearHour	37.00	3.5			6.45
Tulare	Full	Time Threshold	Plateau	37.00	20.5	20.6		6.69
Tulare	Full	Time Threshold	Triangle	36.00	10.5	26.8	26.9	7.43
Tulare	Full	Time Threshold	Asymcur	51.00	9.3	26.8	26.8	7.29
Vina	Full	Time Threshold	Anderson	43.00	4	25	36	6.47
Vina	Full	Time Threshold	MinMax	43.00	17.9			6.03
Vina	Full	Time Threshold	LinearDay	42.00	15.6			6.09
Vina	Full	Time Threshold	LinearHour	44.00	18.4			6.00
Vina	Full	Time Threshold	Plateau	44.00	21.9	22		5.91
Vina	Full	Time Threshold	Triangle	44.00	21.9	22	60.1	6.17
Vina	Full	Time Threshold	Asymcur	44.00	21.9	21.9	71.2	6.25
Amigo	Simplified	Time Threshold	None	173.00				7.52
Ashley	Simplified	Time Threshold	None	185.00				9.05
Chandler	Simplified	Time Threshold	None	184.00				7.55
Chico	Simplified	Time Threshold	None	183.00				8.17
Franquette	Simplified	Time Threshold	None	181.00				6.37
Hartley	Simplified	Time Threshold	None	184.00				7.55
Howard	Simplified	Time Threshold	None	178.00				6.80
Payne	Simplified	Time Threshold	None	184.00				8.70
Serr	Simplified	Time Threshold	None	185.00				8.13
Tehama	Simplified	Time Threshold	None	181.00				7.00
Tulare	Simplified	Time Threshold	None	180.00				7.51
Vina	Simplified	Time Threshold	None	181.00				7.23
Amigo	Full	Heat Threshold	Anderson	14580.00	4	25	36	7.38
Amigo	Full	Heat Threshold	MinMax	164.00	11.3			6.83
Amigo	Full	Heat Threshold	LinearDay	43.00	16.1			6.55
Amigo	Full	Heat Threshold	LinearHour	4262.00	10.6			6.31
Amigo	Full	Heat Threshold	Plateau	4268.00	10.6	86.1		7.14
Amigo	Full	Heat Threshold	Triangle	4283.00	10.6	61.6	82.1	7.11
Amigo	Full	Heat Threshold	Asymcur	4922.00	6.4	46.6	53.9	7.58
Ashley	Full	Heat Threshold	Anderson	14022.00	4	25	36	6.18
Ashley	Full	Heat Threshold	MinMax	259.00	9.7			5.83
Ashley	Full	Heat Threshold	LinearDay	211.00	12			5.34
Ashley	Full	Heat Threshold	LinearHour	5344.00	10.6			5.47
Ashley	Full	Heat Threshold	Plateau	5285.00	10.6	69.3		5.59
Ashley	Full	Heat Threshold	Triangle	5344.00	10.6	45.6	88.6	5.68
Ashley	Full	Heat Threshold	Asymcur	5132.00	5.9	61	72.8	5.60
Chandler	Full	Heat Threshold	Anderson	10999.00	4	25	36	7.11
Chandler	Full	Heat Threshold	MinMax	229.00	11			7.66
Chandler	Full	Heat Threshold	LinearDay	255.00	9.9			7.42
Chandler	Full	Heat Threshold	LinearHour	810.00	20.1			7.62
Chandler	Full	Heat Threshold	Plateau	820.00	20	32.3		7.75
Chandler	Full	Heat Threshold	Triangle	1827.00	16.6	25.5	27.7	8.87
Chandler	Full	Heat Threshold	Asymcur	2982.00	20	24.8	30.5	8.69
Chico	Full	Heat Threshold	Anderson	11449.00	4	25	36	5.75
Chico	Full	Heat Threshold	MinMax	155.00	12.8			5.29
Chico	Full	Heat Threshold	LinearDay	222.00	9.8			5.38

Cultivar	Type	Model	Function	Length	Base	Opt	Crit	RMSE
Chico	Full	Heat Threshold	LinearHour	3506.00	13.3			4.85
Chico	Full	Heat Threshold	Plateau	3879.00	12.3	35.1		4.90
Chico	Full	Heat Threshold	Triangle	3876.00	12.3	65.1	88	4.94
Chico	Full	Heat Threshold	Asymcur	5098.00	7	50.7	66.5	5.00
Franquette	Full	Heat Threshold	Anderson	12637.00	4	25	36	5.13
Franquette	Full	Heat Threshold	MinMax	48.00	16.9			5.66
Franquette	Full	Heat Threshold	LinearDay	102.00	12.5			5.43
Franquette	Full	Heat Threshold	LinearHour	1579.00	15.6			5.48
Franquette	Full	Heat Threshold	Plateau	1149.00	15.9	20		5.45
Franquette	Full	Heat Threshold	Triangle	4765.00	12.4	27.3	27.4	5.95
Franquette	Full	Heat Threshold	Asymcur	6736.00	9.7	27.3	27.5	5.67
Hartley	Full	Heat Threshold	Anderson	14151.00	4	25	36	7.20
Hartley	Full	Heat Threshold	MinMax	35.00	18.8			6.71
Hartley	Full	Heat Threshold	LinearDay	77.00	13.6			6.35
Hartley	Full	Heat Threshold	LinearHour	934.00	19.5			6.42
Hartley	Full	Heat Threshold	Plateau	934.00	19.5	87.5		6.34
Hartley	Full	Heat Threshold	Triangle	927.00	19.5	40.5	41.4	6.67
Hartley	Full	Heat Threshold	Asymcur	1261.00	15.9	42.5	93.5	6.71
Howard	Full	Heat Threshold	Anderson	12168.00	4	25	36	5.28
Howard	Full	Heat Threshold	MinMax	408.00	1.4			5.15
Howard	Full	Heat Threshold	LinearDay	468.00	0.3			5.24
Howard	Full	Heat Threshold	LinearHour	6298.00	6.4			5.48
Howard	Full	Heat Threshold	Plateau	9097.00	2.5	72.5		5.52
Howard	Full	Heat Threshold	Triangle	10299.00	0.2	35.2	36.8	5.67
Howard	Full	Heat Threshold	Asymcur	6808.00	0.1	61.3	82.2	6.02
Payne	Full	Heat Threshold	Anderson	14100.00	4	25	36	6.38
Payne	Full	Heat Threshold	MinMax	942.00	0.1			5.47
Payne	Full	Heat Threshold	LinearDay	638.00	4.3			5.75
Payne	Full	Heat Threshold	LinearHour	9335.00	8.3			5.46
Payne	Full	Heat Threshold	Plateau	9336.00	8.3	84.3		5.75
Payne	Full	Heat Threshold	Triangle	9324.00	8.3	50.6	95.9	5.69
Payne	Full	Heat Threshold	Asymcur	12889.00	0.6	59	78	5.60
Serr	Full	Heat Threshold	Anderson	13091.00	4	25	36	6.44
Serr	Full	Heat Threshold	MinMax	140.00	13.9			5.34
Serr	Full	Heat Threshold	LinearDay	143.00	13.1			5.86
Serr	Full	Heat Threshold	LinearHour	2221.00	16.2			5.02
Serr	Full	Heat Threshold	Plateau	2204.00	16.2	32.9		5.10
Serr	Full	Heat Threshold	Triangle	2195.00	16.2	35.7	35.7	5.52
Serr	Full	Heat Threshold	Asymcur	3170.00	10	56.2	76	5.29
Tehama	Full	Heat Threshold	Anderson	14787.00	4	25	36	4.87
Tehama	Full	Heat Threshold	MinMax	269.00	10.9			5.17
Tehama	Full	Heat Threshold	LinearDay	52.00	14.6			5.08
Tehama	Full	Heat Threshold	LinearHour	8384.00	8.2			5.20
Tehama	Full	Heat Threshold	Plateau	7474.00	8.8	24.8		5.61
Tehama	Full	Heat Threshold	Triangle	6874.00	9.4	24.8	74.9	6.18
Tehama	Full	Heat Threshold	Asymcur	8201.00	9.7	27.6	33.8	6.26
Tulare	Full	Heat Threshold	Anderson	11941.00	4	25	36	5.79
Tulare	Full	Heat Threshold	MinMax	427.00	6.3			6.76
Tulare	Full	Heat Threshold	LinearDay	108.00	12.7			6.60
Tulare	Full	Heat Threshold	LinearHour	8695.00	7.4			6.23
Tulare	Full	Heat Threshold	Plateau	4904.00	11.1	22.4		6.55
Tulare	Full	Heat Threshold	Triangle	7054.00	8.3	26.9	27.2	6.72

Cultivar	Type	Model	Function	Length	Base	Opt	Crit	RMSE
Tulare	Full	Heat Threshold	Asymcur	10288.00	4.9	27.5	27.7	7.48
Vina	Full	Heat Threshold	Anderson	13750.00	4	25	36	6.63
Vina	Full	Heat Threshold	MinMax	106.00	16.2			6.10
Vina	Full	Heat Threshold	LinearDay	99.00	14.4			6.12
Vina	Full	Heat Threshold	LinearHour	1394.00	17.7			5.87
Vina	Full	Heat Threshold	Plateau	1433.00	18.6	24.4		5.95
Vina	Full	Heat Threshold	Triangle	674.00	20.3	23.6	58.8	5.87
Vina	Full	Heat Threshold	Asymcur	814.00	19.5	25.9	28.9	5.89
Amigo	Simplified	Heat Threshold	Anderson	45559.33	4	25	36	7.07
Amigo	Simplified	Heat Threshold	MinMax	3475.05	0.1			7.93
Amigo	Simplified	Heat Threshold	LinearDay	3342.00	1.1			8.04
Amigo	Simplified	Heat Threshold	LinearHour	82957.32	0			8.15
Amigo	Simplified	Heat Threshold	Plateau	49281.26	1.8	14.9		7.19
Amigo	Simplified	Heat Threshold	Triangle	67299.37	0.1	23.6	51.9	7.34
Amigo	Simplified	Heat Threshold	Asymcur	65349.75	0	21.7	62.8	7.70
Ashley	Simplified	Heat Threshold	Anderson	47797.47	4	25	36	6.85
Ashley	Simplified	Heat Threshold	MinMax	3362.02	1.5			7.68
Ashley	Simplified	Heat Threshold	LinearDay	3504.36	1.7			7.66
Ashley	Simplified	Heat Threshold	LinearHour	87113.16	0			7.73
Ashley	Simplified	Heat Threshold	Plateau	69221.16	1.3	21.2		6.85
Ashley	Simplified	Heat Threshold	Triangle	65912.04	1.9	25.4	48.1	6.80
Ashley	Simplified	Heat Threshold	Asymcur	73723.55	0.3	27.3	43.9	6.81
Chandler	Simplified	Heat Threshold	Anderson	49036.93	4	25	36	7.27
Chandler	Simplified	Heat Threshold	MinMax	3777.09	0.1			9.08
Chandler	Simplified	Heat Threshold	LinearDay	3581.65	1.6			9.02
Chandler	Simplified	Heat Threshold	LinearHour	89943.75	0			9.28
Chandler	Simplified	Heat Threshold	Plateau	50127.33	0.4	12.1		7.05
Chandler	Simplified	Heat Threshold	Triangle	21530.25	4.5	10.7	67.3	8.16
Chandler	Simplified	Heat Threshold	Asymcur	35758.11	3.1	13.7	78.3	7.12
Chico	Simplified	Heat Threshold	Anderson	47406.63	4	25	36	7.38
Chico	Simplified	Heat Threshold	MinMax	3628.36	0.1			8.03
Chico	Simplified	Heat Threshold	LinearDay	3595.11	0.2			7.35
Chico	Simplified	Heat Threshold	LinearHour	86578.19	0			7.45
Chico	Simplified	Heat Threshold	Plateau	74486.63	0.2	21.9		6.88
Chico	Simplified	Heat Threshold	Triangle	76029.99	0.3	26.7	53	6.67
Chico	Simplified	Heat Threshold	Asymcur	74678.48	0.1	24.3	69.2	6.79
Franquette	Simplified	Heat Threshold	Anderson	48482.93	4	25	36	6.40
Franquette	Simplified	Heat Threshold	MinMax	3742.90	0.2			9.48
Franquette	Simplified	Heat Threshold	LinearDay	3753.42	0			9.62
Franquette	Simplified	Heat Threshold	LinearHour	88920.55	0			10.11
Franquette	Simplified	Heat Threshold	Plateau	17136.83	5.7	9.7		6.35
Franquette	Simplified	Heat Threshold	Triangle	27992.35	5	14.6	43.8	6.32
Franquette	Simplified	Heat Threshold	Asymcur	28090.33	4.6	13.2	74.1	6.26
Hartley	Simplified	Heat Threshold	Anderson	48900.79	4	25	36	7.31
Hartley	Simplified	Heat Threshold	MinMax	3756.03	0.1			9.27
Hartley	Simplified	Heat Threshold	LinearDay	3767.62	0			9.40
Hartley	Simplified	Heat Threshold	LinearHour	89254.69	0			9.80
Hartley	Simplified	Heat Threshold	Plateau	41762.78	1.2	10.9		6.98
Hartley	Simplified	Heat Threshold	Triangle	21421.92	5.2	11.2	72.7	7.11
Hartley	Simplified	Heat Threshold	Asymcur	26008.81	5.2	12.8	95.1	7.50
Howard	Simplified	Heat Threshold	Anderson	47289.36	4	25	36	6.31
Howard	Simplified	Heat Threshold	MinMax	310.99	22			8.11



Cultivar	Type	Model	Function	Length	Base	Opt	Crit	RMSE
Howard	Simplified	Heat Threshold	LinearDay	3648.95	0			8.15
Howard	Simplified	Heat Threshold	LinearHour	87125.20	0			8.24
Howard	Simplified	Heat Threshold	Plateau	41757.20	1	11.1		6.10
Howard	Simplified	Heat Threshold	Triangle	56617.40	3	28.5	33.3	6.32
Howard	Simplified	Heat Threshold	Asymcur	65468.70	0.3	26	37.1	6.33
Payne	Simplified	Heat Threshold	Anderson	47520.42	4	25	36	7.04
Payne	Simplified	Heat Threshold	MinMax	3643.94	0.1			6.82
Payne	Simplified	Heat Threshold	LinearDay	3637.14	0.1			6.80
Payne	Simplified	Heat Threshold	LinearHour	86661.85	0			6.97
Payne	Simplified	Heat Threshold	Plateau	75616.71	0.4	23.1		6.52
Payne	Simplified	Heat Threshold	Triangle	62647.63	3	26.3	47.6	6.38
Payne	Simplified	Heat Threshold	Asymcur	79691.84	0.1	28.8	50.2	6.49
Serr	Simplified	Heat Threshold	Anderson	47669.87	4	25	36	8.01
Serr	Simplified	Heat Threshold	MinMax	3580.45	0.6			7.96
Serr	Simplified	Heat Threshold	LinearDay	46720.89	89			8.02
Serr	Simplified	Heat Threshold	LinearHour	87227.68	0			8.10
Serr	Simplified	Heat Threshold	Plateau	60294.44	0.6	15.8		7.62
Serr	Simplified	Heat Threshold	Triangle	64322.30	1.9	23.1	60.8	7.61
Serr	Simplified	Heat Threshold	Asymcur	76506.20	0.1	24.9	68.8	7.89
Tehama	Simplified	Heat Threshold	Anderson	47421.72	4	25	36	5.90
Tehama	Simplified	Heat Threshold	MinMax	3661.18	0			7.15
Tehama	Simplified	Heat Threshold	LinearDay	3401.78	1.2			6.95
Tehama	Simplified	Heat Threshold	LinearHour	86520.39	0			6.98
Tehama	Simplified	Heat Threshold	Plateau	47266.81	3	15.1		5.82
Tehama	Simplified	Heat Threshold	Triangle	75597.00	0.1	25.6	57.7	5.51
Tehama	Simplified	Heat Threshold	Asymcur	74038.81	0.3	25.6	57.2	5.71
Tulare	Simplified	Heat Threshold	Anderson	48097.63	4	25	36	6.69
Tulare	Simplified	Heat Threshold	MinMax	3673.95	0.3			7.50
Tulare	Simplified	Heat Threshold	LinearDay	3706.73	0			7.42
Tulare	Simplified	Heat Threshold	LinearHour	87766.27	0			7.67
Tulare	Simplified	Heat Threshold	Plateau	62465.23	0.3	16.3		6.69
Tulare	Simplified	Heat Threshold	Triangle	26025.68	4.9	12.4	76.1	7.85
Tulare	Simplified	Heat Threshold	Asymcur	59086.57	2.4	24.9	42.9	7.64
Vina	Simplified	Heat Threshold	Anderson	47423.86	4	25	36	7.34
Vina	Simplified	Heat Threshold	MinMax	3672.47	0			8.11
Vina	Simplified	Heat Threshold	LinearDay	3678.44	0			8.12
Vina	Simplified	Heat Threshold	LinearHour	87337.07	0			8.46
Vina	Simplified	Heat Threshold	Plateau	37839.86	1.8	10.8		6.99
Vina	Simplified	Heat Threshold	Triangle	15006.44	6.3	10.7	79.1	6.98
Vina	Simplified	Heat Threshold	Asymcur	64823.29	0.3	22.4	49.6	7.52

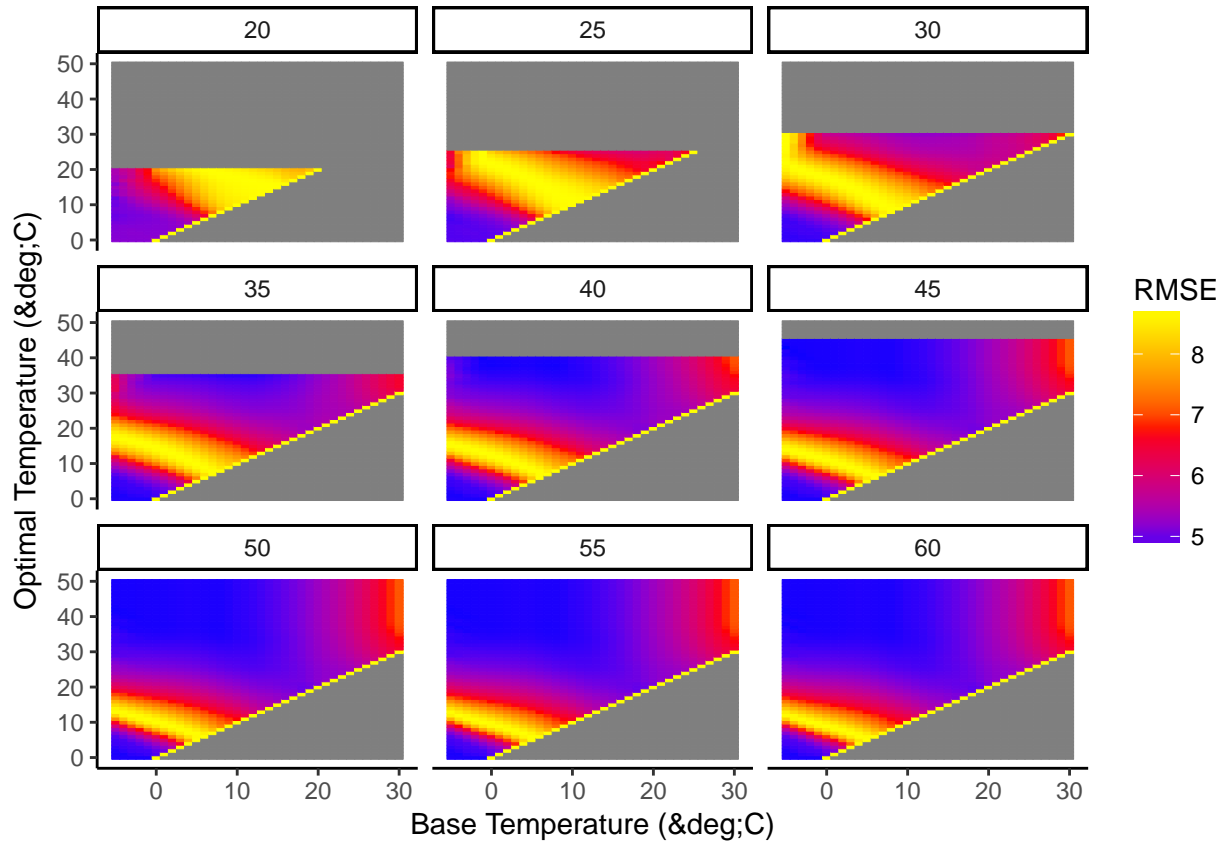


Figure 13: S2: Optimization surface for Day Threshold model with Triangle function applied to the 'Payne' walnut cultivar. Base vs optimal temperature heat maps are shown by critical temperature. All temperatures are given in °C. Grey indicates an invalid model.

---

## References

- J. L. Anderson, E. A. Richardson, and C. D. Kesner. Validation of chill unit and flower bud phenology models for 'Montmorency' sour cherry. In *I International Symposium on Computer Modelling in Fruit Research and Orchard Management 184*, pages 71–78, 1985. URL [http://www.actahort.org/books/184/184\\_7.htm](http://www.actahort.org/books/184/184_7.htm).
- David Ardia, Katharine M. Mullen, Brian G. Peterson, and Joshua Ulrich. *DEoptim: Differential Evolution in R*. 2015. URL <http://CRAN.R-project.org/package=DEoptim>. version 2.2-3.
- N. Bertin. Analysis of the tomato fruit growth response to temperature and plant fruit load in relation to cell division, cell expansion and DNA endoreduplication. *Annals of botany*, 95(3):439–447, 2005. URL <http://aob.oxfordjournals.org/content/95/3/439.short>.
- California Department of Water Resources. CIMIS Weather Observation Data, 2015. URL <http://www.cimis.water.ca.gov/>.
- C. Cesaraccio, D. Spano, P. Duce, and R. L. Snyder. An improved model for determining degree-day values from daily temperature data. *International Journal of Biometeorology*, 45(4):161–169, November 2001. ISSN 0020-7128, 1432-1254. doi: 10.1007/s004840100104. URL <https://link.springer.com/article/10.1007/s004840100104>.
- K. Day, G. Lopez, and T. DeJong. Using growing degree hours accumulated thirty days after bloom to predict peach and nectarine harvest date. In *VIII International Symposium on Modelling in Fruit Research and Orchard Management 803*, pages 163–167, 2007. URL [http://www.actahort.org/books/803/803\\_20.htm](http://www.actahort.org/books/803/803_20.htm).
- C. DeBuse, G. Lopez, and T. DeJong. Using Spring Weather Data to Predict Harvest Date for 'Improved French' Prune. In *IX International Symposium on Plum and Prune Genetics, Breeding and Pomology 874*, pages 107–112, 2008. URL [http://www.actahort.org/books/874/874\\_13.htm](http://www.actahort.org/books/874/874_13.htm).
- David A. Goldhamer and Mario Viveros. Effects of preharvest irrigation cutoff durations and postharvest water deprivation on almond tree performance. *Irrigation Science*, 19(3):125–131, 2000. doi: 10.1007/s002710000013.
- F. P. Marra, P. Inglese, T. M. DeJong, and R. S. Johnson. Thermal time requirement and harvest time forecast for peach cultivars with different fruit development periods. In *V International Peach Symposium 592*, pages 523–529, 2001. URL [http://www.actahort.org/books/592/592\\_70.htm](http://www.actahort.org/books/592/592_70.htm).
- Gregory S. McMaster and W. W. Wilhelm. Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, 87(4):291–300, 1997. URL <http://www.sciencedirect.com/science/article/pii/S0168192397000270>.
- MJ Menne, I Durre, B Korzeniewski, S McNeal, K Thomas, X Yin, S Anthony, R Ray, RS Vose, BE Gleason, and TG Houston. Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.22, 2015. URL <http://doi.org/10.7289/V5D21VHZ>.
- Ben M. Mimoun and T. M. DeJong. Using the relation between growing degree hours and harvest date to estimate run-times for peach: a tree growth and yield simulation model. In *V International Symposium on Computer Modelling in Fruit Research and Orchard Management 499*, pages 107–114, 1998. URL [http://www.actahort.org/books/499/499\\_10.htm](http://www.actahort.org/books/499/499_10.htm).
- Kenneth V. Price, Rainer M. Storn, and Jouni A. Lampinen. *Differential Evolution - A Practical Approach to Global Optimization*. Natural Computing. Springer-Verlag, January 2006. ISBN 540209506.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL <https://www.R-project.org/>.

- David E. Ramos. *Walnut production manual*, volume 3373. UCANR Publications, 1997. URL <https://books.google.com/books?hl=en&lr=&id=yKawaOt88A0C&oi=fnd&pg=PP7&dq=Walnut+production+manual&ots=bvzJNjZWkv&sig=PFaXdavMNYLWNcOoemC7g6HqFZY>.
- E. A. Richardson, S. D. Seeley, D. R. Walker, J. L. Anderson, and G. L. Ashcroft. Pheno-climatology of spring peach bud development [Cold damage]. *HortScience*, 1975. URL <http://agris.fao.org/agris-search/search.do?recordID=US19760051429>.
- Mirjana Ruml, Dragan Milatović, Todor Vulić, and Ana Vuković. Predicting apricot phenology using meteorological data. *International journal of biometeorology*, 55(5):723–732, 2011. URL <http://link.springer.com/article/10.1007/s00484-010-0387-0>.
- G. Sibbett, L. Hendricks, G. Carnill, W. Olson, R. Jeter, D. Ramos, G. Martin, C. Davis, and others. Walnut quality and value maximized by harvest management. *California Agriculture*, 28(7):15–17, 1974. URL <http://calag.ucanr.edu/archive/?article=ca.v028n07p15&sharebar=share>.
- R. L. Snyder, Donatella Spano, Carla Cesaraccio, and Pierpaolo Duce. Determining degree-day thresholds from field observations. *International Journal of Biometeorology*, 42(4):177–182, 1999. URL <http://link.springer.com/article/10.1007/s004840050102>.
- S. Tombesi, R. Scalia, J. Connell, B. Lampinen, and T. M. DeJong. Fruit development in almond is influenced by early Spring temperatures in California. *The Journal of Horticultural Science and Biotechnology*, 85(4):317–322, 2010. URL <http://www.tandfonline.com/doi/abs/10.1080/14620316.2010.11512674>.
- I. J. Warrington, T. A. Fulton, E. A. Halligan, and H. N. De Silva. Apple fruit growth and maturity are affected by early season temperatures. *Journal of the American Society for Horticultural Science*, 124(5):468–477, 1999. URL <http://journal.ashspublications.org/content/124/5/468.short>.
- Senshan Yang, Joanne Logan, and David L. Coffey. Mathematical formulae for calculating the base temperature for growing degree days. *Agricultural and Forest Meteorology*, 74(1):61–74, 1995. URL <http://www.sciencedirect.com/science/article/pii/016819239402185M>.
- Frank G. Zalom and Peter B. Goodell. *Degree days: the calculation and use of heat units in pest management*. University of California, Division of Agriculture and Natural Resources, 1983.
- Caixi Zhang, Kenji Tanabe, Shiping Wang, Fumio Tamura, Akira Yoshida, and Kazuhiro Matsumoto. The impact of cell division and cell enlargement on the evolution of fruit size in *Pyrus pyrifolia*. *Annals of Botany*, 98(3):537–543, 2006. URL <https://aob.oxfordjournals.org/content/98/3/537.full>.