

# 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. Phenology models estimate development rates for crop growth simulations that assess the effects of different varieties, management decision, and the environment on productions. Farmers use phenological predictions from these models to prepare for the logistics associated with flowering and harvest, such as pesticide application and equipment rental.

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 summarizes how plants' rates of development change based on the ambient temperature surrounding the plant. These phenology models also use a threshold value to define the most temperature sensitive period of development for the plant in question. We distinguish between two broad types of models based on the unit of this threshold value: thermal time threshold (TTT) and day threshold (DT) models.

TTT models set a threshold in thermal time units. Then the model counts the number of days it takes to reach that threshold. In TTT models of annual crops, this day count is exactly the stage length. However, this can be viewed as a simplified version of a more general model we develop in this paper, where the day count predicts the stage length via a linear regression model. We refer to this more general model as the full model. TTT models are generally used to study plant responses to different weather conditions. However, the full TTT models could also be used for within season predictions.

Models of perennial crops, specifically trees, generally use DT phenology models. The threshold for DT models is given in days. The model counts the amount of thermal time accumulated during that time. Next, the model uses the thermal time accumulated to predict the length of the stage, in days. This approach is commonly used in fruit and nut phenology models, including peaches, almonds, and prunes, to make in season predictions for farm orchard management. DT modeling is supported by finding that temperature variation within the first 30 to 60 days of flowering have disproportionately large effects on fruit development [mimoun1998; day2007; debuse2008; tombesi2010; ruml2011]

One key component of both model types is the method for converting daily or hourly temperatures 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 ( $T_c$ ). Above Thermal time is zero below the base temperature and above the critical temperature. Typical values of  $T_b$  and  $T_c$  for mediterranean fruit and nut trees are around 4°C and 36°C respectively). Thermal time is highest at the optimum temperature, which may be around 25°C for tree crops [richardson1975]. Growth chamber experiments can determine cardinal temperature values. However, while this is relatively straightforward for annuals, it is impractical for tree crops.

While some effort has been made to compare different DT phenological models [e.g. marra2001; tombesi2010], the performance of DT and TTT models have not been compared. In this paper, we develop the full TTT model and compare the predictive power of 8 DT and 14 TTT 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 models

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 full or simplified form is used; (2) selection of the function used to derive thermal time from temperature; (3) estimating the cardinal temperature and thermal time accumulation parameters; and (4) fitting of a statistical model that predicts the length of growth stage of interest from time or thermal time accumulation.

### Thermal Time Threshold (TTT) Model

For all the phenology models  $\lambda$  will represent the model threshold. In TTT models,  $\lambda_j$  is the accumulated thermal time threshold for stage  $j$ . The independent variable in the TTT model is the number of days necessary to accumulate  $\lambda_j$  thermal time units, which we represent as  $x_i(\lambda_j)$  for a given year  $i$ . For the simplified TTT model the thermal time threshold for the stage ( $\lambda_j$ ) is equal to the amount of thermal time required for the plant to complete the stage developmentally ( $\Lambda_j$ ). Thus in the simplified TTT model, we can define the length of stage  $j$ , in days, as the number of days it takes to accumulate  $\Lambda_j$  thermal time units,  $x_i(\Lambda_j)$  (1).

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

Here,  $y_{ij}$  is the length (in days) of stage  $j$  during the  $i$ th year of observation and  $\varepsilon_i$  is the associated error term. However, this model is just a simplification of the full TTT model. Specifically, in the full model the thermal time threshold does not have to equal the total thermal time required for the plant to complete the  $j$ th stage of development. Since  $\lambda_j$  may be less than  $\Lambda_j$ , a change in one day to  $x_i(\lambda_j)$  can mean a change in stage length of more than one day. To accommodate this we multiply  $x_i(\lambda_j)$  by  $\beta_j$  in our model equation. In addition, because  $\lambda_j$  can be quite variable, it is possible for the plant to meet the thermal time threshold during the first day of the phenological stage. If this is the case,

$$\begin{aligned} y_{ij} &= \beta_j \cdot x_i(\lambda_j) \\ &= 0 \end{aligned}$$

regardless of how large  $\beta_j$  is. To compensate for this, we include the parameter  $\alpha_j$  to represent the mean stage length if the plant reaches the thermal time threshold during the first day of the stage. If we add these two parameters into equation (1) we get

$$y_{ij} = \alpha_j(\lambda_j) + \beta_j(\lambda_j) \cdot x_i(\lambda_j) + \varepsilon_i \quad (2)$$

which is exactly the equation for univariate linear regression. In equation (2),  $y_{ij}$  is the length (in days) of stage  $j$  during the  $i$ th year of observation. This means that  $\alpha_j(\lambda_j)$  is the stage length if the plant reaches the thermal time threshold before the first day of  $j$  stage is over. Then  $\beta_j(\lambda_j)$  is the change in the predicted stage length, in days, if the plant takes one day longer to reach the thermal time threshold ( $\lambda_j$ ). The  $\alpha_j(\lambda_j)$  and  $\beta_j(\lambda_j)$  values depend on the values of  $x_i$ . Since  $x_i$  itself depends on  $\lambda_j$ , so do  $\alpha_j(\lambda_j)$  and  $\beta_j(\lambda_j)$ , if only indirectly. As the above phenological model incorporates the full linear regression, we refer to this model as the Full TTT Model.

### Day Threshold (DT) Model

The Day Threshold 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,  $\alpha_j(\lambda_j)$  is the predicted stage length if the trees did not experience any thermal time accumulation during the first  $\lambda_j$  days of stage  $j$  and  $\beta_j(\lambda_j)$  is the change in stage length due to accumulating one additional thermal time unit during the first  $\lambda_j$  days of the stage. Like before, we refer to this model as the Full Day Threshold Model.

When simplifying the DT model, we set the model threshold to be  $\Lambda_j$ , total the number of days the plant will accumulate thermal time in stage  $j$ . This is necessarily the length of stage  $j$ , in days. Now,  $\beta_j(\Lambda_j)$  is the change in stage length if the plant accumulates one more unit of thermal time in  $\Lambda_j$  days. However, since  $\Lambda_j$  is, by definition, the total number of days that the plant will accumulate thermal time,  $\beta_j(\Lambda_j)$  must be zero. This removes the  $x_i$  term from the linear regression, leaving the equation

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

Since the stage length,  $y_{ij}$ , no longer depends on  $x_i$ , the model will predict the same stage length for all values of thermal time accumulation, namely the mean stage length,  $\bar{y}$ .

## Thermal time computation functions

We evaluated five functions for calculating thermal time. While there are additional functions (ex. @marra2001), 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 [yang1995].

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

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 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 (citation?), 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 [anderson1985].

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

As this function is almost always used with the cardinal temperatures reported by @anderson1985, 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 @anderson1985 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  $T_b$  [@zalom1983; @mcmaster1997].

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

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 ( $T_{max}$ ) [@ruml2011]. A simpler GDD function is sometimes used as well [@snyder1999]. While there are other models used in the literature [@marra2001], these five capture a variety of assumptions about tree responses to temperature, and are all computationally feasible.

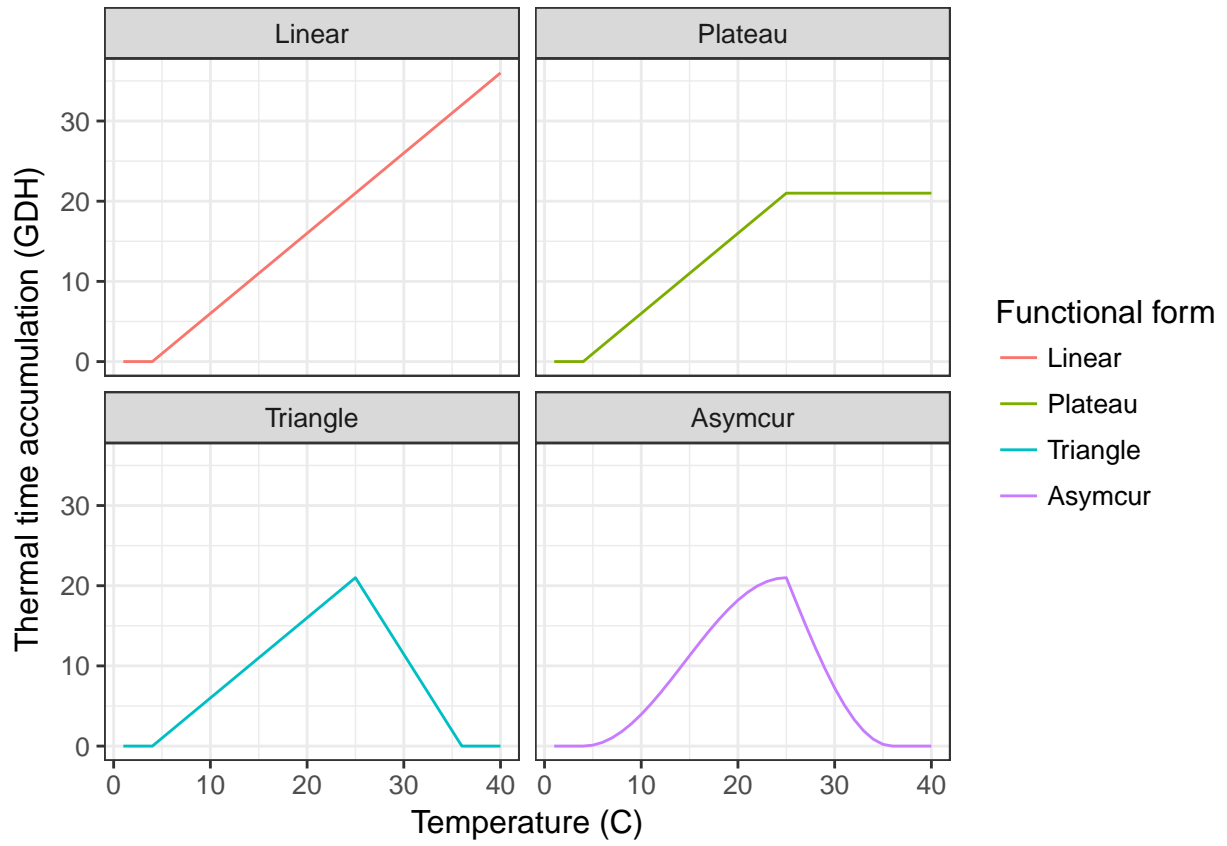


Figure 1: Visual representations of thermal time functions. For all of the functions  $T_b = 4^{\circ}\text{C}$ ,  $T_o = 25^{\circ}\text{C}$ , and  $T_c = 36^{\circ}\text{C}$ .