Predicting 'Kerman' Pistachio Yield Using Previous-Yield and Hourly Air-Temperature Accumulations

Craig Kallsen, U.C. Cooperative Extension – Kern County

Version 3, May 2017

After 22 years in working with pistachios, I have not found any of the methods available for estimating whether or not the fall and winter chill requirement had been met to be of much help in predicting how the pistachios in the San Joaquin Valley of California would fare in the upcoming year. Recently, I had what I am calling a minor epiphany of sorts related to the winter 'rest period' or 'winter chilling' that I wanted to share. It dawned on me that what many in the industry really want to know is not if fall and winter chilling were adequate or surpassed some threshold. What many really want to know is what their yield will be like at harvest time. To begin the statistical process that might help answer the latter question, I asked a well-known farming operation to share some of their yield data with me and they graciously consented. I obtained CPC yield data from three large orchards, some of which dated back to the mid-1980s.

In short, the objective of this research was to correlate the yield of these orchards with air-temperature records available from some nearby CIMIS stations. For those that want the complete report on this project: the methods used, the scientific citations upon which this work is built and a discussion of the results (in metric units), the full article is available in the April 2017 issue of the following journal:

Kallsen, C.E. 2017. Temperature-related variables associated with yield of 'Kerman' pistachio in the San Joaquin Valley of California. HortScience: 52(4):598-605.

Fortunately, although useful in understanding the procedures used and some of the implications, reading the full published report is not necessary. I have worked to condense the results of this study for this newsletter and developed an Excel spreadsheet that will allow you to estimate your future yields, based on the results of this study. *This research examined Kerman and Peters orchards, only, and the results will be most applicable for Kerman and Peters orchards that are 15 years and older and which are located in the southern San Joaquin Valley.* The model that was developed assumes that crop nutrition, irrigation scheduling and pest control were above average, and that water quality was adequate.

Generally, pistachio is an alternate bearing crop, so it follows that if you want to predict yield for pistachio, the best input variable available is the yield you harvested the previous year. With an alternate bearing crop, almost by definition, if your yield was low last year you can predict it will be higher this year and vice versa. The objective of this research was to improve the prediction by adding various air temperature variables into the predictive model. A large number of air temperature-related variables were examined that summed or accumulated the number of hours that air temperature was above, below or inbetween set limit values during discreet calendar periods of time. For example, one of the many variables entered into the model was the accumulation of the number of hours less than 45 degrees F. from the time period from November 15 through February 15. Most of the variables calculated were rejected in favor of a few that did a better job of predicting yield.

The statistical procedure used was forward, stepwise multiple regression. The idea was to use this procedure to determine the top three or four variables, in addition to previous yield, that increased our ability to predict yield at harvest time. This objective was achieved by using stringent F thresholds for variable to enter and exit the model, and by using the Durbin-Watson statistic and related statistics to find variables that minimized collinearity in the variables selected for the model.

When the results were in, the regression procedure found the most useful variable for predicting subsequent yield in these older Kerman orchards was previous yield. As expected, if the preceding season yield was low, subsequent yield was high and vice versa. Interestingly, and not very surprisingly to me based on my previous experience, variables such as accumulated chilling hours and chill portions, as measured during the fall and winter preceding harvest were not very useful in predicting subsequent yield or were negatively correlated with yield in the pooled data from the three Kerman orchards in the San Joaquin Valley that I studied. The lack of significance of variables that measured chill related accumulations, in combination with the significance of variables that measured warmer temperatures during the fall and winter, suggest that we usually get enough winter chilling but the interruptions of warmer temperatures during the winter and fall, reverse or interrupt the fall and winter 'rest period' of the trees.

Of the 30 or more variable entered into this statistical procedure the following variables were the most powerful in estimating CPC yield (i.e. also called 'total inshell weight'):

1. The square of previous season CPC yield (i.e. total inshell weight) in lbs. /acre
2. The square of accumulated hours greater than or equal to 45 °F <u>and</u> less than or equal to 60 °F as summed from Nov. 15 through Feb. 15.
3. The accumulated hours greater the 65 °F as summed from Nov. 15 through Feb. 15
4. The accumulated hours greater than 80 °F as summed from March 20 through April 25.

Previous yield, hours greater that 65 degrees F. from Nov. 15 through Feb. 15, and hours greater than 80 °F. from March 20 through April 25 (roughly the bloom and early nuts set period) were negatively correlated with future yield, while the variable that summed the number of hours between and equal to 45 and 60 °F was positively correlated to yield. So, in short, if you want a big yield at the upcoming harvest, your best chance is as follows: to have had a small yield the previous harvest, to accumulate as many hours as possible between 45 and 60 °F from Nov. 15 through Feb. 15, have no hours greater than 65 °F from Nov. 15 through Feb. 15 and to have had no hours with temperatures greater than 80 °F from March 20 through April 25. In the three orchards upon which this model is based, the mean values over the course of the 31 years were as follows: for previous yield the mean was <u>3541 lbs. /acre</u>, for hourly accumulations from 45 to 60 °F from Nov. 15 through Feb. 15 was <u>1172 hours</u>, for hourly accumulations greater than 65 °F for the same period was <u>75 hours</u> and for hourly accumulations greater than 80 °F from March 20 to April 25 was <u>39 hours</u>. If you are going to use this model and track hourly accumulations during the fall, winter and bloom periods, actual hourly accumulations measured in your orchard can be compared to these mean values to give a rough indication how the upcoming harvest yield will be impacted. However, these are means and are not 'threshold' hourly accumulations. The results of this model suggests that, in general, thresholds do not exist but that yield will increase or decrease linearly or curvilinearly, depending on previous yield and the most significant hourly air temperature accumulations identified in this study.

So how much error is associated with a yield prediction in this model? This model predicts yield more accurately when the input values for the four variables are near the mean values listed above. By utilizing the 'standard error of the predicted value', which is a statistic calculated by the multiple regression model, we can estimate how much error is associated with the yield value predicted by this regression equation. By multiplying the standard errors of the predicted values by 2, over the range of predictions upon which this model is based, the error associated with a given yield prediction for these three orchards varied from a maximum of ± 952 lbs. /acre for yield predictions based on input values near their extreme values, to a minimum of ± 266 lbs. /acre for cases with values near the middle of the range. On average, using the standard error of the predicted value, a yield prediction was accurate within ± 524 lbs. /acre.

This model 'explained' about 64% of the total adjusted variation in CPC yield measured over the past 31 years. One should not spend too much time in looking for physiological importance in these variables. These variables are merely correlations which do not explain cause and effect as to what may be going on biologically. Nevertheless, this model should allow growers to estimate what their fall yields will be, with greater accuracy, than trying to estimate whether or not 'chilling' or 'winter rest' or 'dormancy' was or was not sufficient. This model has shown that warm temperatures during the fall and winter 'rest period' appear to have reduced yields to a greater extent than previously thought or, perhaps, allowed for, in other chilling models used in the past. The use of fall and winter foliar sprays of kaolinite clay or calcium carbonate crystals, if effective in reducing orchard air temperature below 65 ° F, appears to be supported by this model.

In the past, poor chilling may have been blamed for poor yields when the problem may have been excessively warm air temperatures during the bloom period. Results from this research suggest that the industry may have focused too much attention on chilling, and not enough on the orchard's current point in the alternate bearing cycle. Extremely low yields may be the result of a "perfect storm" of being on the wrong 'side' of high previous yield and these few temperature variables, even if everything else was done right as far as irrigating, fertilizing and controlling pests. By using this model, growers should have the ability to obtain an estimate of their yield for the upcoming harvest by the end of April, assuming that that their irrigation scheduling and water quality, crop nutrition, and pest control are adequate.

As an example of how this model works, let's assume that the date is April 26, and a person has downloaded their orchard temperature recorder and summed the hourly air temperature accumulations of the variables required in this model as follows:

1. The number of hour greater than or equal to 45 degrees F. and less than or equal to 60 degrees F. as summed from Nov. 15 through Feb. 15. was **<u>1150 hours</u>**.
2. The accumulated hours greater the 65 degrees F. as summed from Nov. 15 through Feb. 15 was **<u>136 hours</u>**.
3. The accumulated hours greater than 80 degrees F. as summed from March 20 through April 25 was **<u>45 hours</u>**.
4. By checking their records from the previous year, last year's total inshell yield (previous yield) was **<u>2200 lbs. /acre.</u>**

When this data is entered in the model equation, the predicted yield for the upcoming harvest in the fall will be estimated as **3968 ± 524 lbs. /acre**, assuming the average error value discussed above. As described above this prediction also assumes that crop nutrition, irrigation scheduling and pest control were above average, and that water quality was adequate. Obviously, there is plenty of room that might allow for your actual yield to differ from that predicted by the model. Based on initial feedback of growers that have used the model, this model does a better job of predicting yield in more 'normal' years. The model fails to predict yields accurately during the crop failure year of 2015 and the subsequent recovery in 2016. Clearly something extraordinary occurred in 2015, most likely related to weather during the fall and winter rest period, that is not accounted for in this regression, and which the polynomial functions of the regression cannot model. When additional grower data from 2015 and 2016 is included in the model development, the variable which measures temperatures greater than 65 °F during the period from November 15 through Feb. 15 accounts for an increasing percentage of the total variation accounted for in the regression. It appears some threshold was surpassed, perhaps excessive heating of the buds through direct solar radiation for example, that is not reflected in the air temperature data. Ideally, this model would be used in conjunction with other yield-related information as it becomes available during the season such as apparent nut set, early and late season field measurements such as nut abortion, and blanking and nut quality observations such as percent nut split, to increase the information base available to further refine yield estimates.

For ease of calculation, I produced an Excel spreadsheet that will calculate your predicted yield using the equation of this model if you input your values as measured in your orchard(s). The intercept and coefficients of the variables appear in an equation in one of the hidden columns of the Excel spreadsheet. You can 'unhide' the column if you wish to see the equation. I 'hid' the column to reduce the chance of users accidentally deleting or changing the equation that estimates yield. The spreadsheet is available at http://cekern.ucanr.edu/files/260338.xlsx . If you have any questions on how to use the spreadsheet do not hesitate to contact me at cekallsen@ucanr.edu or by phone at 661-868-6221.