

# Multifactor Models for Improved Prediction of Phenological Timing in Cold-climate Wine Grapes

James A. Schrader, Paul A. Domoto, and Gail R. Nonnecke

*Department of Horticulture, Iowa State University, Ames, IA 50011*

Diana R. Cochran

*Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD 20742*

*Additional index words.* multiple regression, mathematical modeling, environmental monitoring, weather data utilization, viticulture, interspecific hybrids, crop management

**Abstract.** An accurate predictive model for estimating the timing of seasonal phenological stages of grape (*Vitis* L.) would be a valuable tool for crop management. Currently the most used index for predicting the phenological timing of fruit crops is growing degree days (GDD), but the predictive accuracy of the GDD index varies from season-to-season and is considered unsatisfactory for grapevines grown in the midwestern United States. We used the methods of multiple regression to analyze and model the effects of multiple factors on the number of days remaining until each of four phenological stages (budbreak, bloom, veraison, and harvest maturity) for five cold-climate wine grape cultivars (Frontenac, La Crescent, Marquette, Petit Ami, and St. Croix) grown in central Iowa. The factors (predictor variables) evaluated in models included cultivar, numerical day of the year (DOY), DOY of soil thaw or the previous phenological stage, photoperiod, GDD with a base temperature of 10 °C (GDD 10), soil degree days with a base temperature of 5 °C (SDD 5), and solar accumulation. Models were evaluated for predictive accuracy and goodness of fit by calculating the coefficient of determination ( $R^2$ ), the corrected Akaike information criterion (AICc), and the Bayesian information criterion (BIC); testing for normal distribution of residuals; and comparing the actual number of days remaining until a phenological stage with the number of days predicted by models. The top-performing models from the training set were also tested for predictive accuracy on a validation dataset (a set of data not used to build the model), which consisted of environmental and phenological data recorded for one popular Midwest cultivar (Marquette) in 2019. At all four phenological stages, inclusion of multiple factors (cultivar and four to six additional factors) resulted in predictive models that were more accurate and consistent than models using cultivar and GDD 10 alone. Multifactor models generated from data of all five cultivars had high  $R^2$  values of 0.996, 0.985, 0.985, and 0.869 for budbreak, bloom, veraison, and harvest, respectively, whereas  $R^2$  values for models using only cultivar and GDD 10 were substantially lower (0.787, 0.904, 0.960, and 0.828, respectively). The average errors (differences from actual) for the top multifactor models were 0.70, 0.84, 1.77, and 3.80 days for budbreak, bloom, veraison, and harvest, respectively, and average errors for models that included only cultivar and GDD 10 were much larger (5.27, 2.24, 2.79, and 4.29 days, respectively). In the validation tests, average errors for budbreak, bloom, veraison, and harvest were 1.92, 1.31, 0.94, and 1.67 days, respectively, for the top multifactor models and 10.05, 2.54, 4.23, and 4.96 days, respectively, for models that included cultivar and GDD 10 only. Our results demonstrate the improved accuracy and utility of multifactor models for predicting the timing of phenological stages of cold-climate grape cultivars in the midwestern United States. Used together in succession, the models for budbreak, bloom, veraison, and harvest form a four-stage, multifactor calculator for improved prediction of phenological timing. Multifactor models of this type could be tailored for specific cultivars and growing regions to provide the most accurate predictions possible.

A mechanism that accurately estimates the timing of annual phenological stages of wine grapes or other fruit crops would be a valuable tool for crop management. In viticulture, the ability to predict phenological stages would be beneficial in planning and preparation for pruning, pesticide application, canopy management of shoot positioning and leaf or shoot removal, cluster thinning, and harvest. Past efforts to develop a predictive index for phenological timing

have focused mainly on measures of seasonal heat accumulation, and the most commonly investigated index for use with fruit crops is GDD (García de Cortázar-Atauri et al., 2009; Gentilucci and Burt, 2018; Zapata et al., 2017). Although the GDD index has been used successfully to estimate phenological timing in moderate climates (Fraga et al., 2016; Verdugo-Vásquez et al., 2017), its accuracy can vary greatly from season-to-season and it is considered inadequate as a

stand-alone index for predicting the phenological timing of cold-climate wine grapes grown in the midwestern United States (Fernández-González et al., 2013; Schrader et al., 2019). Nearly all research on phenological timing of grapevines has focused on describing the effects of environmental factors on select phenological stages and prescribing a threshold at which each stage will likely be accomplished. Reports of true predictive models that use real-time environmental data to estimate the arrival of a future phenological stage within the same season do not yet exist in the literature. The current report describes the development and evaluation of a prediction system that uses values from multiple factors to provide real-time predictions for the phenological timing of cold-climate wine grapes. The cultivars evaluated in this study are of strong interest to viticulturists in the midwestern and north-eastern United States, and information about their characteristics and culture is widely available (Minnesota Grape Growers Assn., 2016; Schrader et al., 2019, 2020; Smiley et al., 2016).

Although phenological models based on only air temperature and heat accumulation are predominant in the literature (Fraga et al., 2016; García de Cortázar-Atauri et al., 2009; Verdugo-Vásquez et al., 2017), researchers have demonstrated that other environmental factors can affect phenological timing (Basler and Körner, 2014; Greer et al., 2006; Kliever, 1975; Rezazadeh and Stafne, 2018; Schaber and Badeck, 2003; Way and Montgomery, 2015; Williams et al., 1985). The influence of daylength (photoperiod) as an environmental cue in directing seasonal phenology has been demonstrated with grapevines and other woody plants (Basler and Körner, 2014; Rezazadeh and Stafne, 2018; Schaber and Badeck, 2003; Way and Montgomery, 2015). Schrader et al. (2019) proposed that photoperiod may be more important in directing annual phenology than common theory suggests, but that the importance of photoperiod is not easily observed except during years when heat accumulation and photoperiod are poorly correlated. Root-zone temperatures and solar radiation also have been shown to affect phenological timing of grapevines and other fruit crops (Greer et al., 2006; Kliever, 1975; Williams et al., 1985). With the increase in availability of environmental data from both private and public weather stations over the past few decades, it is now feasible to include data such as root-zone temperature and solar accumulation in models along with air temperature data and photoperiod to improve predictive accuracy. We hypothesize that multifactor models could be especially beneficial for use in areas with cold and variable climate, areas such as the midwestern United States where air temperatures and heat accumulation rates are often volatile and vary from year-to-year.

Dunkler et al. (2014) describe statistical models as “simple mathematical rules derived from empirical data describing the

association between an outcome and several explanatory variables.” There are three main purposes for developing and using statistical models: 1) prediction, 2) explanation, and 3) description (Heinze et al., 2018; Shmueli, 2010). Descriptive models are meant to capture the association between dependent and independent variables (Shmueli, 2010), but do not consider causality in a formal manner (Heinze et al., 2018). Explanatory models are used to estimate causal effects and summarize the impacts of independent variables, with particular emphasis placed on minimizing bias (Heinze et al., 2018). The primary goal of predictive models is to accurately predict an outcome value from a set of independent variables (predictors) (Heinze et al., 2018). The aim of our research was to develop predictive models that would be suitable for estimating the timing of phenological stages within a reasonable range of error. Although we believe that causal relationships exist among the variables in our project, the goals of our research were not to provide proof of causality but to develop accurate and useful predictive models based on empirical research data.

Multiple regression is used routinely in research studies for modeling processes in plant biology (Bock et al., 2011; Matas et al., 2005; Webb et al., 2012), and has been used in many studies to develop predictive models for plant phenology (Anderson et al., 1978; Constable and Rose, 1988; Weikai and Wallace, 1998). When using multiple regression to build predictive models, one of the goals is to balance accuracy with efficiency, to develop and choose the most accurate model possible by including all useful factors, and excluding factors that are not useful (Heinze et al., 2018). If a model is excessively complex, it becomes impractical for use; however, if too much emphasis is placed on simplicity (minimal number of factors), the model is likely to become inaccurate as values of predictors move away from norms (Burnham and Anderson, 2002). Statisticians have developed several metrics that aid in variable selection and help to balance model accuracy with simplicity. The coefficient of determination ( $R^2$ ), which is the proportion of

variation in the outcome that is explained by the predictor variables, is useful for judging the goodness of fit, but  $R^2$  almost always indicates a better fit as the number of predictors increases (Kassambara, 2018). Akaike’s Information Criterion (AIC) is a metric that penalizes the inclusion of additional variables to a model, providing a measure for balancing accuracy and complexity (Akaike, 1974). BIC is a variant of AIC with a stronger penalty for including additional variables to the model (Kassambara, 2018). The capacity for calculating these metrics and others is common to most modern statistical software platforms, facilitating the creation and selection of quality models through multiple regression.

Correlation (collinearity) of predictor variables can be a potential issue when building multiple regression models, a situation that can make it difficult to separate the effects of independent variables and to detect statistical significance of predictors (Dormann et al., 2013; Frost, 2020; Meloun et al., 2002). Although it is common in studies that include multiple predictors, the existence of collinearity does not influence the predictions, the accuracy of the predictions, or the goodness-of-fit statistics of a model, and therefore it is not a concern when developing and using models for prediction only (Frost, 2020; Neter et al., 1996; Shmueli, 2010). Although often overlooked during criticisms of regression methods, specific requirements for the three types of models (explanatory, descriptive, and predictive) are different. Although collinearity can lead to inflated standard errors that may interfere with inference in explanatory and descriptive modeling, moderate collinearity is not an important issue for predictive modeling where inference is not an objective (Makridakis et al., 1998; Shmueli, 2010; Vaughan and Berry, 2005). The focus of predictive modeling is the accuracy (predictive power) of the model when it is applied to new data, and removal of a useful predictor variable only because it is correlated with another predictor can cause an unnecessary reduction in predictive accuracy (Frost, 2020; Hyndman and Athanasopoulos, 2018; Shmueli, 2010). Inclusion of a predictor variable that is highly correlated with another (near complete collinearity) should be avoided because it can increase the possibility of overfitting the model, allowing it to fit well on the training set but to perform poorly when used for predictions based on new samples (Frost, 2020; Martens and Naes, 1989).

During the creation of predictive models, measures of  $R^2$  and F statistics for the original data (the training set) are used as indicators for the level of association of variables but not as a gauge for causation (Frost, 2020; Shmueli, 2010). Metrics such as AIC aid in the selection of models for best predictive accuracy based on the training set (Berk, 2008; Konishi and Kitagawa, 2007; Shmueli, 2010), and utilization of the AIC metric can prevent overfitting (Dettling, 2015). AICc is more effective for models based on small

sample sizes (N), and when the N increases well beyond the number of variables, AICc converges to AIC, a feature that makes AICc effective regardless of sample size (Burnham and Anderson, 2002). Methods of elastic net regression are effective for selecting variables and building strong predictive models even when collinearity is present in a group of predictors (Boehmke and Greenwell, 2019; Kelly, 2014; Zou and Hastie, 2005). After model creation, evaluation of model accuracy on a validation set (a set of data not used to build the model) is the best indicator of predictive power (Geisser, 1975; Picard and Cook, 1984; Stone, 1974).

We used the methods of multiple regression to analyze and model the effects of multiple factors on the number of days remaining until each of four phenological stages (budbreak, bloom, veraison, and harvest maturity) for five cold-climate wine grape cultivars evaluated in central Iowa. The objectives of our research were as follows: 1) to create a series of multifactor predictive models that use inputs of local, real-time data to accurately and consistently estimate the timing for the seasonal arrival of four key phenological stages of cold-climate wine grapes; 2) compare and demonstrate the predictive accuracies of these models by testing them on the original data set and a validation dataset; and 3) provide regression equations from the top-performing mathematical models and describe basic methods for using the equations in a spreadsheet program to predict the annual timing of the four phenological stages. The resulting mathematical models provide a predictive mechanism for real-time estimation of the arrival of phenological stages based on seasonal thresholds, photoperiod, and cumulative environmental metrics.

## Materials and Methods

Nursery-grown plants of ‘Frontenac’, ‘La Crescent’, ‘Marquette’, ‘Petit Ami’, and ‘St. Croix’ were received from Double A Vineyards, Inc. (Fredonia, NY) as part of the NE-1020 project titled “Multi-state Evaluation of Winegrape Cultivars and Clones.” Vines were planted on 20 May 2008 at the Iowa State University Horticulture Research Station near Gilbert, IA [lat. 42°6′27″ N, long. 93°35′24″ W; USDA hardiness zone 5a (USDA, 2019)]. Plants were arranged in a randomized complete block design with three-vine panels replicated six times (18 vines per cultivar) and bordered by guard rows and end vines. Soil at the research plot was a well-drained Clarion loam (fine-loamy, mixed, superactive, mesic Type Hapludoll), and no fertilizer was added. Vines of experimental units were trained to a high cordon (single curtain, bilateral cordon), with the trellis wire 1.83 m above the ground and vine spacing of 2.44 × 3.05 m. Pests and diseases were managed according to established protocols of integrated pest management (Hoover et al., 2011). Vines were pruned and managed according to the protocols

Received for publication 12 Aug. 2020. Accepted for publication 21 Sept. 2020.

Published online 26 October 2020.

This research was supported in part by the U.S. Department of Agriculture/National Institute of Food and Agriculture NE-1020 project “Multi-State Evaluation of Winegrape Cultivars and Clones” and by Iowa State University. We thank Nicholas Howell and the staff of the Iowa State University Horticulture Research Station for their assistance with management of the research plot, and we thank Pete Lawlor and Kenneth McCabe for their helpful critiques of the manuscript.

D.R.C. is the corresponding author. E-mail: cochrand@umnd.edu.

This is an open access article distributed under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

described by Domoto (2014) and Minnesota Grape Growers Assn. (2016), including compensatory pruning following winters with significant bud injury. Dormant pruning was performed in mid-March for all seasons evaluated in the project.

Timing of four phenological stages [budbreak (at 50% budbreak), bloom (at 50% bloom), veraison (at 50% veraison), and harvest maturity] was recorded as the numerical DOY when each vine reached the measurement threshold for the specified stage. Vines were considered to have reached 50% budbreak when 50% of buds per vine had reached stage 4 of the Modified Eichhorn-Lorenz (Modified E-L) phenological scale (Dry and Coombe, 2004). Vines were considered to have reached 50% bloom when 50% of inflorescences per vine had reached Modified E-L stage 23, and vines were considered to have reached 50% veraison when 50% of berries per vine had reached Modified E-L stage 35 (Eichhorn and Lorenz, 1977). Harvest maturity for each cultivar was determined by assays of soluble solids content (SSC), pH, and titratable acidity (TA). Beginning at 100% veraison, berries were sampled weekly [morning collections of 100 berries (10 berries per vine) from 10 vines of each cultivar and selected randomly and proportionally from the top, middle, and bottom of clusters]. Sampling frequency was increased to daily as readings neared the desired values for harvest. Samples were juiced with a bench-top juicer (Model J8006; Omega, Harrisburg, PA) and pressed through cheesecloth. The SSC of grapes was determined by using a temperature-compensating refractometer (ATAGO, Bellevue, WA). Juice pH was measured with a pH meter (Orion 2-Star; Thermo Fisher Scientific, Waltham, MA), and a mini-titrator (Model HI84532U-01; Hanna Instruments, Woonsocket, RI) was used to quantify TA. Harvest maturity for each cultivar was judged based on SSC, pH, and TA standards set by Dhar-madhikari and Wilker (2001) for white (21% to 22% SSC, 3.2–3.4 pH, 7–9 g·L<sup>-1</sup> TA) and red table wines (22% to 24% SSC, 3.3–3.5 pH, 6–8 g·L<sup>-1</sup> TA) with adjustment for ‘St. Croix’ (SSC ≈18), which did not reach desired SSC before pH exceeded acceptable values. The primary data for models were collected during three growth seasons (2011, 2013, and 2014).

A comprehensive software package (JMP Pro 14; SAS Institute, Cary, NC) was used to analyze, create, and compare models representing the effects of multiple factors on the number of days remaining until fulfillment of the four phenological stages. The factors evaluated in models included cultivar, DOY, DOY of soil thaw or the previous phenological stage, photoperiod, GDD with a base temperature of 10 °C (GDD 10), soil degree days with a base temperature of 5 °C (SDD 5), and solar accumulation measured in MJ·m<sup>-2</sup> (solar acc.). Full datasets for the number of days remaining until each phenological stage were developed from primary data by incorporating environmental data

from each of the days preceding arrival of the specified stage. In preliminary research, we found that one of the key methods for improving model accuracy was the anchoring of environmental predictors to a suitable threshold for each of the stages. Therefore, GDD 10, SDD 5, and Solar acc. were included in models as cumulative amounts measured from soil thaw (for budbreak) or from accomplishment of the previous phenological stage (for bloom, veraison, and harvest). Raw data used to calculate DOY of soil thaw, GDD 10, SDD 5, and Solar acc. were collected continuously onsite by using a weather station with air and soil temperature probes (CS215-L and CS231-L; Campbell Scientific, Inc., Logan, UT), a solar radiation sensor (CS301 Pyranometer; Campbell Scientific, Inc.), and a data logger (CR 1000; Campbell Scientific, Inc.) that logged conditions every 15 min. The air temperature and solar radiation sensors were set at 2 m above the ground and the soil temperature sensor was located at 10.2 cm below the soil surface. Data for these measurements were accumulated and accessed through the Iowa Environmental Mesonet (Iowa State University, 2019). The DOY of soil thaw was defined as the earliest DOY in which the soil temperature (10.2 cm below surface) was above 0 °C and remained above 0 °C for the rest of the growing season. Photoperiod was calculated for each date by using data for sunrise and sunset that were specific for the location of the vineyard plot. These data were obtained from Sunrise-sunset.org (2020).

The GDD 10 metric was calculated in two ways (from hourly temperatures and average daily temperatures) for comparison of accuracy and utility. The GDD 10 based on average daily temperatures, GDD 10 (Avg), is the most popular method used for agriculture and is often included in local weather station datasets that are available to the public, such as those of the Iowa Environmental Mesonet (Iowa State University, 2019). The GDD 10 (Avg) metric is calculated by subtracting the base temperature of 10 °C from the average daily air temperature [(daily maximum + daily minimum) ÷ 2] with a cap of 30 °C (Iowa State University, 2020). The GDD 10 metric based on hourly temperatures, GDD 10 (Hrly), has been shown to be more precise than GDD 10 (Avg) for phenological purposes (Gu, 2016), therefore we included it in this study for comparison and use. The GDD 10 (Hrly) metric was calculated by subtracting the base temperature of 10 °C from the mean hourly air temperature in °C (with no temperature cap), dividing each hourly value by 24, then adding the 24 hourly values together to obtain the daily value.

The generalized regression platform of JMP Pro 14 was used to calculate metrics for model and variable selection, generate model equations, and test the equations on training and cross-validation sets to validate and quantify in-sample predictive accuracy (SAS Institute Inc., 2020). Models based on one or two factors were built using ordinary

least squares regression, which provides the best possible coefficient estimates when the model satisfies the assumptions for linear regression and collinearity does not exist (predictors are not correlated) (Berk, 2008). Models that included more than two predictor variables were built and evaluated using elastic net regression techniques, which include parameters that aid in variable selection and effectively manage collinearity if it is present (Boehmke and Greenwell, 2019; Kelly, 2014; Zou and Hastie, 2005). Two types of cross validation (AICc and KFold) were applied and compared for all elastic net regressions. All models were evaluated for normal distribution of residuals by using two JMP diagnostic functions (histogram of residuals and normal quartile plots).

Metrics used for selecting variables and judging model accuracy based on the training set were  $R^2$ , AICc, BIC, the average model error, and the largest model error. In our study, error was defined as the absolute difference between the predicted and actual number of days remaining until arrival of the phenological stage. The two error metrics were calculated on each observational unit by applying the model equation to the data to predict the number of days remaining, subtracting the predicted value from the actual value, then taking the square root of the squared difference to receive the absolute value. Using this method for calculating the average model error (mean difference between actual and predicted) enabled the calculation of mean separation statistics for this metric. Means separation analyses were conducted by using Tukey-Kramer honestly significant difference model ( $P \leq 0.05$ ) in JMP Pro 14. The “largest model error” was the largest error of prediction for any single experimental unit. With four of the five metrics (AICc, BIC, average model error, and largest model error), the lower the metric value, the better the model (Burnham and Anderson, 2002; Kassambara, 2018). With  $R^2$ , the higher the metric value, the better the fit of the model. Values for AICc and BIC are relative to the specific dataset, and therefore cannot be used to compare models built on different datasets. For each phenological stage, models were generated from the full dataset (all five cultivars) and separately using data from one cultivar only (Marquette) to gauge the importance of cultivar specificity for building models of this type and to confirm the accuracy of the multicultivar models. Models that were determined to be superior based on the five metrics were selected for evaluation on the out-of-sample validation dataset.

Out-of-sample validation performance of models was evaluated by using data from 144 ‘Marquette’ vines from a separate vineyard measured in 2019. The vineyard used for validation was located 400 m from the original vineyard at the same research station. Vines in the validation group were planted in 2012, grown on the same training system (single curtain, bilateral cordon with the trellis wire at 1.83 m), and managed by using

Table 1. Performance of basic (single-factor or cultivar plus one additional factor) and multifactor models for estimating the timing of budbreak (days remaining until 50% budbreak) for five cold-climate grape cultivars in central Iowa based on 3 years of data (2011, 2013, and 2014). The generalized models (for all five cultivars) were built based on data (N = 4403) from 'Frontenac', 'La Crescent', 'Marquette', 'Petit Ami', and 'St. Croix' over a range of 1 to 56 d remaining until 50% budbreak. A subset of models specific to one widely grown cultivar (Marquette) was generated to confirm the accuracy of the multicultivar model. Models specific to 'Marquette' were built based on data (N = 861) spanning the range from 1 to 53 d remaining until 50% budbreak. Models based on only one or two predictors (factors) were created by using ordinary least squares regression, and models based on more than two factors were created by using elastic net regression that included tuning parameters for variable selection and control of potential collinearity.

Factors included <sup>2</sup>	R <sup>2</sup>	Avg model error (d) <sup>3</sup>	Largest model error (d) <sup>3</sup>	AICc	BIC
All five cultivars					
Cultivar, DOY	0.949	1.80 e <sup>x</sup>	6.98	16,312	16,357
Cultivar, Photoperiod	0.946	1.97 e	6.57	17,378	17,422
Cultivar, GDD 10 (Hrly)	0.787	5.27 b	18.89	29,217	29,262
Cultivar, GDD 10 (Avg)	0.683	6.40 a	27.11	30,974	31,018
Cultivar, SDD 5	0.829	4.72 c	21.61	28,265	28,309
Cultivar, Solar acc. <sup>w</sup>	0.935	3.23 d	8.02	23,973	24,018
Cultivar, DOY, Soil thaw DOY, GDD 10 (Avg), SDD 5, Solar acc.	0.994	0.93 f	4.47	13,832	13,902
Cultivar, DOY, Soil thaw DOY, GDD 10 (Hrly), SDD 5, Solar acc.	0.996	0.75 g	3.81	11,932	12,002
Cultivar, DOY, Soil thaw DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	0.996	0.70 g	3.98	11,388	11,464
'Marquette' only					
GDD 10 (Hrly)	0.794	5.09 a	16.93	5,651	5,665
DOY, Soil thaw DOY, GDD 10 (Avg), SDD 5, Solar acc.	0.994	0.90 b	3.5	2,648	2,682
DOY, Soil thaw DOY, GDD 10 (Hrly), SDD 5, Solar acc.	0.997	0.65 c	2.73	2,091	2,124
DOY, Soil thaw DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	0.997	0.54 d	2.49	1,880	1,918

<sup>2</sup>Factor abbreviations and definitions: AICc = corrected Akaike information criterion; BIC = Bayesian information criterion; DOY = numerical day of the year; Photoperiod = day length on DOY measured in hours; GDD 10 (Hrly) = growing degree days with a base of 10 °C and calculated from hourly data; GDD 10 (Avg) = growing degree days with a base of 10 °C and calculated from daily average; SDD 5 = soil degree days with a base of 5 °C and based on hourly data measured at a soil depth of 10.2 cm; Solar acc. = accumulated solar radiation measured in MJ·m<sup>-2</sup>; Soil thaw DOY = the DOY when the soil temperature at a depth of 10.2 cm became >0 °C and remained >0 °C for the rest of the growing season. Factors with cumulative quantities [GDD 10 (Hrly), GDD 10 (Avg), SDD 5, Solar acc.] were measured from Soil thaw.

<sup>3</sup>Model error (difference between predicted and actual days remaining until budbreak) is reported as the absolute difference (not positive or negative) because residuals were normally distributed for nearly all models evaluated and therefore had relatively even distribution of positive and negative differences from actual.

<sup>x</sup>Within each analysis group, means followed by the same letter are not different according to Tukey-Kramer honestly significant difference test at  $P \leq 0.05$  (n = 4403 for models based on all five cultivars and n = 861 for models based on 'Marquette' only).

<sup>w</sup>Models with non-normal distribution of residuals.

Table 2. Performance of basic (single-factor or cultivar plus one additional factor) and multifactor models for estimating the timing of bloom (days remaining until 50% bloom) for select cold-climate grape cultivars in central Iowa based on 3 years of data (2011, 2013, and 2014). The generalized models (for all five cultivars) were built based on data (N = 2667) from 'Frontenac', 'La Crescent', 'Marquette', 'Petit Ami', and 'St. Croix' over a range of 1 to 32 d remaining until 50% bloom. A subset of models specific to one widely grown cultivar (Marquette) was generated to confirm the accuracy of the multicultivar model. Models specific to 'Marquette' were built based on data (N = 553) spanning the range from 1 to 32 d remaining until 50% bloom. Models based on only one or two predictors (factors) were created by using ordinary least squares regression, and models based on more than two factors were created by using elastic net regression that included tuning parameters for variable selection and control of potential collinearity.

Factors included <sup>2</sup>	R <sup>2</sup>	Avg model error (d) <sup>3</sup>	Largest model error (d) <sup>3</sup>	AICc	BIC
All five cultivars					
Cultivar, DOY	0.921	2.19 bc <sup>x</sup>	5.21	12,317	12,358
Cultivar, Photoperiod	0.900	2.28 b	5.85	12,923	12,964
Cultivar, GDD 10 (Hrly)	0.904	2.24 bc	6.71	12,837	12,878
Cultivar, GDD 10 (Avg)	0.866	2.73 a	7.46	13,720	13,761
Cultivar, SDD 5 <sup>w</sup>	0.920	2.10 c	5.88	12,342	12,383
Cultivar, Solar acc. <sup>w</sup>	0.954	1.49 d	5.37	10,885	10,926
Cultivar, DOY, Budbreak DOY, GDD 10 (Avg), SDD 5, Solar acc.	0.985	0.85 e	2.80	7,835	7,900
Cultivar, DOY, Budbreak DOY, GDD 10 (Hrly), SDD 5, Solar acc.	0.985	0.84 e	2.83	7,822	7,887
Cultivar, DOY, Budbreak DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	0.985	0.84 e	2.78	7,816	7,886
'Marquette' only					
GDD 10 (Hrly)	0.907	2.23 a	6.86	2,683	2,696
DOY, Budbreak DOY, GDD 10 (Avg), SDD 5, Solar acc.	0.991	0.68 b	2.44	1,396	1,430
DOY, Budbreak DOY, GDD 10 (Hrly), SDD 5, Solar acc.	0.991	0.68 b	2.51	1,410	1,440
DOY, Budbreak DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	0.991	0.65 b	2.41	1,379	1,414

<sup>2</sup>Factor abbreviations and definitions: AICc = corrected Akaike information criterion; BIC = Bayesian information criterion; DOY = numerical day of the year; Photoperiod = daylength on DOY measured in hours; GDD 10 (Hrly) = growing degree days with a base of 10 °C and calculated from hourly data; GDD 10 (Avg) = growing degree days with a base of 10 °C and calculated from daily average; SDD 5 = soil degree days with a base of 5 °C and based on hourly data measured at a soil depth of 10.2 cm; Solar acc. = accumulated solar radiation measured in MJ·m<sup>-2</sup>; Budbreak DOY = the DOY when the experimental unit reached 50% budbreak. Factors with cumulative quantities [GDD 10 (Hrly), GDD 10 (Avg), SDD 5, Solar acc.] were measured from 50% budbreak.

<sup>3</sup>Model error (difference between predicted and actual days remaining until bloom) is reported as the absolute difference (not positive or negative) because residuals were normally distributed for nearly all models evaluated and therefore had relatively even distribution of positive and negative differences from actual.

<sup>x</sup>Within each analysis group, means followed by the same letter are not different according to Tukey-Kramer honestly significant difference test at  $P \leq 0.05$  (n = 2667 for models based on all five cultivars and n = 553 for models based on 'Marquette' only).

<sup>w</sup>Models with non-normal distribution of residuals.

the same methods as the original vineyard. Methods for sampling, data collection, and data preparation for the validation group were the same as those for the original train-

ing group. The two metrics used for judging the performance of models on the validation dataset were the average model error and largest model error. Metrics such as R<sup>2</sup>, AICc,

and BIC are not applicable for out-of-sample validation of the type used in our methods because the models were tested on a completely separate dataset than the one used to

Table 3. Performance of basic (single-factor or cultivar plus one additional factor) and multifactor models for estimating the timing of veraison (days remaining until 50% veraison) for select cold-climate grape cultivars in central Iowa based on 3 years of data (2011, 2013, and 2014). The generalized models (for all five cultivars) were built based on data (N = 5298) from ‘Frontenac’, ‘La Crescent’, ‘Marquette’, ‘Petit Ami’, and ‘St. Croix’ over a range of 1 to 67 d remaining until 50% veraison. A subset of models specific to one widely grown cultivar (Marquette) was generated to confirm the accuracy of the multicultivar model. Models specific to ‘Marquette’ were built based on data (N = 952) spanning the range from 1 to 57 d remaining until 50% veraison. Models based on only one or two predictors (factors) were created by using ordinary least squares regression, and models based on more than two factors were created by using elastic net regression that included tuning parameters for variable selection and control of potential collinearity.

Factors included <sup>2</sup>	R <sup>2</sup>	Avg model error (d) <sup>y</sup>	Largest model error (d) <sup>y</sup>	AICc	BIC
All five cultivars					
Cultivar, DOY	0.979	1.94 h <sup>*</sup>	6.55	23,711	23,757
Cultivar, Photoperiod <sup>w</sup>	0.811	6.01 a	27.06	36,446	36,492
Cultivar, [Photoperiod] <sup>w,v</sup>	0.863	4.68 b	27.90	34,724	34,776
Cultivar, GDD 10 (Hrly)	0.960	2.79 c	10.00	26,255	26,301
Cultivar, GDD 10 (Avg)	0.971	2.39 de	8.83	26,537	26,583
Cultivar, SDD 5	0.973	2.32 ef	7.94	26,198	26,244
Cultivar, Solar acc.	0.975	2.23 g	8.18	25,707	25,753
Cultivar, DOY, Bloom DOY, GDD 10 (Avg), Solar acc.	0.985	1.77 i	5.87	23,070	23,136
Cultivar, DOY, Bloom DOY, GDD 10 (Hrly), Solar acc.	0.985	1.77 i	5.64	23,040	23,106
‘Marquette’ only					
GDD 10 (Hrly)	0.974	2.10 a	6.34	4,337	4,351
DOY, Bloom DOY, GDD 10 (Avg), SDD 5, Solar acc.	0.994	1.02 b	3.08	3,122	3,156
DOY, Bloom DOY, GDD 10 (Hrly), SDD 5, Solar acc.	0.994	1.01 b	2.93	3,117	3,151

<sup>2</sup>Factor abbreviations and definitions: AICc = corrected Akaike information criterion; BIC = Bayesian information criterion; DOY = numerical day of the year; Photoperiod = daylength on DOY measured in hours; GDD 10 (Hrly) = growing degree days with a base of 10 °C and calculated from hourly data; GDD 10 (Avg) = growing degree days with a base of 10 °C and calculated from daily average; SDD 5 = soil degree days with a base of 5 °C and based on hourly data measured at a soil depth of 10.2 cm; Solar acc. = accumulated solar radiation measured in MJ·m<sup>-2</sup>; Bloom DOY = the DOY when the experimental unit reached 50% bloom. Factors with cumulative quantities [GDD 10 (Hrly), GDD 10 (Avg), SDD 5, Solar acc.] were measured from 50% bloom.

<sup>y</sup>Model error (difference between predicted and actual days remaining until veraison) is reported as the absolute difference (not positive or negative) because residuals were normally distributed for nearly all models evaluated and therefore had relatively even distribution of positive and negative differences from actual.

<sup>\*</sup>Within each analysis group, means followed by the same letter are not different according to Tukey-Kramer honestly significant difference test at  $P \leq 0.05$  (n = 5298 for models based on all five cultivars and n = 952 for models based on ‘Marquette’ only).

<sup>w</sup>Models with non-normal distribution of residuals.

<sup>v</sup>Factor designations within [brackets] indicate that the factor was included as nonlinear (second degree polynomial) for that specific model.

create them. Values for average model error and largest model error for the validation set were calculated by the same methods as those of the training set.

## Results and Discussion

Models with strong predictive accuracy were created for each of the four phenological stages (budbreak, bloom, veraison, and harvest maturity) by using the methods of multiple regression, and the five metrics used for gauging the fit and accuracy of models provided definitive criteria for selecting the best models. Initial screening of models and predictor variables showed that nearly all of the models evaluated had normally distributed residuals. The few models with non-normal distribution of residuals were rejected and were identified as non-normal in tabular results. Initial comparisons of the two types of GDD variables evaluated [GDD 10 (Hrly) and GDD 10 (Avg)] showed that GDD 10 (Hrly) was more effective as a predictor than GDD 10 (Avg). Therefore, our results focus on the use of GDD 10 (Hrly), but include representative models with GDD 10 (Avg) for potential use if or when data are unavailable for calculating GDD 10 (Hrly). Our finding that GDD 10 (Hrly) performed better than GDD 10 (Avg) in prediction models supports the conclusions of Gu (2016). Of the two types of elastic net validation evaluated (AICc and KFold), models constructed by using AICc validation performed better on the ‘Marquette’ validation dataset (data not shown), therefore models built using AICc

validation were chosen over those built using KFold and are the only elastic net models included in tables. The “cultivar” variable was found to be essential for models built on data from all five cultivars to provide adjustment for innate differences among the cultivars (data not shown). Therefore, “cultivar” was included in all models that were built on data from more than one cultivar.

### Model selection and performance based on the training set

At all four phenological stages, the fit and accuracy of models was improved by inclusion of multiple factors. The elastic net regression functions optimized the selection of predictor variables and provided the best multifactor model for each stage based on the training dataset. Along with performance metrics for basic models built on individual variables (cultivar and one more variable) and those of the top multifactor model, we included results for other select models that could be useful depending on availability of data in certain geographical areas. For example, inclusion of photoperiod and GDD 10 (Hrly) delivered the best model in nearly all contexts, but photoperiod data may be difficult to obtain for some areas and GDD 10 (Avg) data are more commonly available than GDD 10 (Hrly) data. Therefore, models built using GDD 10 (Avg) and models built without the photoperiod variable were included in tabular results.

**Budbreak.** For prediction of budbreak, models created from data of all five cultivars showed a range of  $R^2$  values from 0.683 for

the poorest performing basic model [built using cultivar and GDD 10 (Avg)] to 0.996 for the highest performing multifactor model built by using cultivar and six other predictor variables [DOY, soil thaw DOY, photoperiod, GDD 10 (Hrly), SDD 5, solar acc.] (Table 1). Based on all five metrics of model performance, all of the multifactor models outperformed the basic models, showing higher  $R^2$  values and lower values for average model error, largest model error, AICc, and BIC (Table 1). The top-performing multifactor model contained all seven predictor variables (cultivar plus six additional variables), had the lowest average model error (0.70 d), the second lowest value for largest model error (3.98 d), and the lowest values for AICc and BIC. Although the model with all seven variables was selected as the best model based on all five metrics, the second best multifactor model (built without the photoperiod variable) had the same  $R^2$  as the top model, an average model error that was not significantly different from the top model, and a value for largest model error that was less than that of the top model (Table 1). The lower values for AICc and BIC for the seven-factor model confirm the benefit of including the photoperiod variable when these data are available. Budbreak models based on data from one cultivar (‘Marquette’ only) had similar results, with the best model (built with all six predictor variables) showing an  $R^2$  of 0.997, average model error of 0.54 d, largest model error of 2.49 d, and the lowest values for AICc and BIC (Table 1). These results based on the training set indicate that

Table 4. Performance of basic (single-factor or cultivar plus one additional factor) and multifactor models for estimating the timing of harvest maturity (days remaining until harvest) for select cold-climate grape cultivars in central Iowa based on 3 years of data (2011, 2013, and 2014). The generalized models (for all five cultivars) were built based on data (N = 3381) from 'Frontenac', 'La Crescent', 'Marquette', 'Petit Ami', and 'St. Croix' over a range of 1 to 51 d remaining until harvest. A subset of models specific to one widely grown cultivar (Marquette) was generated to confirm the accuracy of the multicultivar model. Models specific to 'Marquette' were built based on data (N = 760) spanning the range from 1 to 48 d remaining until harvest. Models based on only one or two predictors (factors) were created by using ordinary least squares regression, and models based on more than two factors were created by using elastic net regression that included tuning parameters for variable selection and control of potential collinearity.

Factors included <sup>a</sup>	R <sup>2</sup>	Avg model error (d) <sup>y</sup>	Largest model error (d) <sup>y</sup>	AICc	BIC
All five cultivars					
Cultivar, DOY <sup>x</sup>	0.837	4.35 a <sup>w</sup>	11.19	20,614	20,657
Cultivar, Photoperiod <sup>a</sup>	0.835	4.36 a	11.20	20,650	20,693
Cultivar, GDD 10 (Hrly)	0.828	4.29 ab	14.55	20,786	20,829
Cultivar, GDD 10 (Avg)	0.828	4.30 a	14.97	20,791	20,835
Cultivar, SDD 5	0.841	4.12 c	11.78	20,414	20,456
Cultivar, Solar acc.	0.840	4.17 bc	12.09	20,535	20,577
Cultivar, DOY, Veraison DOY, GDD 10 (Avg), SDD 5, Solar acc.	0.860	3.91 d	10.42	20,094	20,161
Cultivar, DOY, Veraison DOY, GDD 10 (Hrly), SDD 5, Solar acc.	0.861	3.90 d	10.65	20,064	20,132
Cultivar, DOY, Veraison DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	0.869	3.80 e	10.13	19,870	19,944
'Marquette' only					
GDD 10 (Hrly)	0.814	4.54 a	9.72	4,747	4,761
DOY, Veraison DOY, GDD 10 (Avg), SDD 5, Solar acc.	0.978	1.43 b	6.02	3,123	3,155
DOY, Veraison DOY, GDD 10 (Hrly), SDD 5, Solar acc.	0.979	1.33 c	5.92	3,093	3,125
DOY, Veraison DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	0.979	1.32 c	5.91	3,093	3,125

<sup>a</sup>Factor abbreviations and definitions: AICc = corrected Akaike information criterion; BIC = Bayesian information criterion; DOY = numerical day of the year; Photoperiod = daylength on DOY measured in hours; GDD 10 (Hrly) = growing degree days with a base of 10 °C and calculated from hourly data; GDD 10 (Avg) = growing degree days with a base of 10 °C and calculated from daily average; SDD 5 = soil degree days with a base of 5 °C and based on hourly data measured at a soil depth of 10.2 cm; Solar acc. = accumulated solar radiation measured in MJ·m<sup>-2</sup>; Veraison DOY = the DOY when the experimental unit reached 50% veraison. Factors with cumulative quantities [GDD 10 (Hrly), GDD 10 (Avg), SDD 5, Solar acc.] were measured from 50% veraison.

<sup>y</sup>Model error (difference between predicted and actual days remaining until harvest) is reported as the absolute difference (not positive or negative) because residuals were normally distributed for nearly all models evaluated and therefore had relatively even distribution of positive and negative differences from actual.

<sup>w</sup>Models with non-normal distribution of residuals.

<sup>x</sup>Within each analysis group, means followed by the same letter are not different according to Tukey-Kramer honestly significant difference test at  $P \leq 0.05$  (n = 3381 for models based on all five cultivars and n = 760 for models based on 'Marquette' only).

the best multifactor models for the prediction of budbreak explained more than 99% of the variance ( $R^2 > 0.99$ ) and predicted the arrival of budbreak within an average model error of <1 d and a largest model error of <4 d over a range from 1 to 56 d before the arrival of budbreak.

**Bloom.** For prediction of bloom, models created from data of all five cultivars showed a range of  $R^2$  values from 0.866 for the poorest performing basic model [built using cultivar and GDD 10 (Avg)] to 0.985 for the top-performing multifactor model built by using cultivar and six other predictor variables [DOY, soil thaw DOY, photoperiod, GDD 10 (Hrly), SDD 5, solar acc.] (Table 2). The top three multifactor models for bloom [one with all seven variables, one without photoperiod, and one with GDD 10 (Avg) and without photoperiod] showed very little difference in performance based on the five metrics. All three had the same  $R^2$ , their average model errors were not statistically different from each other, and their largest model errors were within only a few hundredths of a day of each other (Table 2). The multifactor models performed substantially better than the basic models, with two of the basic models showing non-normal distribution of residuals (cultivar with SDD 5 and cultivar with solar acc.) and the other four models performing poorly (lower  $R^2$  and higher values for average model error, largest model error, AICc, and BIC) in comparison with the multifactor models (Table 2). Multifactor models for bloom that were based on data from one cultivar ('Marquette' only) had

$R^2$  values of 0.991, average model errors of  $\leq 0.68$  d, largest model errors of  $\leq 2.51$  d, AICc of  $\leq 1410$ , and BIC of  $\leq 1440$ , whereas the basic model built with only GDD 10 (Hrly) did not perform as well (0.907, 2.23 d, 6.86 d, 2683, and 2696, respectively) (Table 2). Therefore, the best multifactor models for the prediction of bloom explained more than 98% of the variance ( $R^2 > 0.98$ ) in the training set and predicted the arrival of bloom within an average model error of <0.9 d and a largest model error of <3 d over a range from 1 to 32 d before the arrival of bloom.

**Veraison.** Modeling for the prediction of veraison had one important difference compared with modeling of the other three phenological stages. For the other three phenological stages, the numeric value for photoperiod increased (preceding budbreak and bloom) or decreased (preceding harvest) with time in a predominately linear fashion. In the days leading up to veraison the incremental values for photoperiod increase until summer solstice (DOY 172) then decrease for the time remaining until veraison (generally taking place between DOY 207 and 215 in central Iowa), causing the values for photoperiod to be nonlinear through the time period leading up to this phenological stage. Because of this innate nonlinear relationship, we evaluated both the linear and nonlinear (polynomial) effects of photoperiod in basic models (cultivar and photoperiod only) and found non-normal distribution of residuals for both models (Table 3). When included as potential predictors for multifactor models,

photoperiod variables (both linear and polynomial) were eliminated from models by elastic net regression. Therefore, based on results from the training set, it was determined that photoperiod was not useful as a predictor for estimating the timing of veraison.

For models of veraison based on all five cultivars, predictive power was improved by inclusion of multiple factors but the amount of improvement was not as large as it was for the other phenological stages. For basic models (cultivar plus one other variable) with normal residuals,  $R^2$  values ranged from 0.960 to 0.979, average model errors ranged from 1.94 to 2.79 d, largest model errors ranged from 6.55 to 10.00 d, AICc ranged from 23,711 to 26,537, and BIC ranged from 23,757 to 26,583 (Table 3). For the two selected multifactor models, both of which had photoperiod and SDD 5 eliminated by elastic net regression, the  $R^2$  values were 0.985, average model errors were 1.77 d, largest model errors were  $\leq 5.87$  d, AICc values were  $\leq 23,070$ , and BIC values were  $\leq 23,136$ . Multifactor models for veraison that were based on data from one cultivar ('Marquette' only) had  $R^2$  values of 0.994, average model errors of  $\leq 1.02$  d, largest model errors of  $\leq 3.08$  d, AICc values of  $\leq 3122$ , and BIC values of  $\leq 3156$ , whereas the basic model built with only GDD 10 (Hrly) had values for these metrics of 0.974, 2.10, 6.34, 4337, and 4351, respectively (Table 3). Therefore, the top multifactor models for the prediction of veraison explained over 98% of the variance ( $R^2 >$

Table 5. Predictive accuracy of models based on the validation set (set of data not used to build the model), which consisted of environmental and phenological data recorded for one widely grown Midwest cultivar (Marquette) in 2019. The vineyard used for validation contained 144 vines and was located 400 m from the original vineyard at the same research station. Models based on only one predictor were created by using ordinary least squares regression, and models based on multiple predictors were created by using elastic net regression that included tuning parameters for variable selection and control of potential collinearity. Models that include “cultivar” as a predictor variable (factor) were created from data of all five cultivars, and the cultivar adjustment coefficient for Marquette was used for these models. Models listed without “cultivar” as a predictor variable were created from data of Marquette only.

Factors included <sup>a</sup>	Avg model error (d) <sup>b</sup>	Largest model error (d) <sup>b</sup>
<b>Budbreak</b>		
GDD 10 (Hrly)	10.05 a <sup>x</sup>	20.99
DOY, Soil thaw DOY, GDD 10 (Avg), SDD 5, Solar acc.	3.79 b	8.51
DOY, Soil thaw DOY, GDD 10 (Hrly), SDD 5, Solar acc.	2.25 c	6.03
DOY, Soil thaw DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	1.96 d	5.97
Cultivar, DOY, Soil thaw DOY, GDD 10 (Hrly), SDD 5, Solar acc.	2.32 c	5.98
Cultivar, DOY, Soil thaw DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	1.92 d	5.94
<b>Bloom</b>		
GDD 10 (Hrly)	2.54 a	7.47
DOY, Budbreak DOY, GDD 10 (Avg), SDD 5, Solar acc.	1.39 b	5.75
DOY, Budbreak DOY, GDD 10 (Hrly), SDD 5, Solar acc.	1.34 bc	5.41
DOY, Budbreak DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	1.36 bc	5.64
Cultivar, DOY, Budbreak DOY, GDD 10 (Hrly), SDD 5, Solar acc.	1.31 c	4.94
Cultivar, DOY, Budbreak DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	1.32 c	5.02
<b>Veraison</b>		
GDD 10 (Hrly)	4.23 a	9.13
DOY, Bloom DOY, GDD 10 (Avg), SDD 5, Solar acc.	1.07 b	5.89
DOY, Bloom DOY, GDD 10 (Hrly), SDD 5, Solar acc.	1.04 b	5.66
Cultivar, DOY, Bloom DOY, GDD 10 (Hrly), Solar acc.	0.94 c	5.73
<b>Harvest</b>		
GDD 10 (Hrly)	4.96 a	10.06
DOY, Veraison DOY, GDD 10 (Avg), SDD 5, Solar acc.	2.09 b	7.10
DOY, Veraison DOY, GDD 10 (Hrly), SDD 5, Solar acc.	1.99 c	7.23
DOY, Veraison DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	2.01 c	6.83
Cultivar, DOY, Veraison DOY, GDD 10 (Hrly), SDD 5, Solar acc.	1.96 c	5.77
Cultivar, DOY, Veraison DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	1.67 d	6.27

<sup>a</sup>Factor abbreviations and definitions: DOY = numerical day of the year; Photoperiod = daylength on DOY measured in hours; GDD 10 (Hrly) = growing degree days with a base of 10 °C and calculated from hourly data; GDD 10 (Avg) = growing degree days with a base of 10 °C and calculated from daily average; SDD 5 = soil degree days with a base of 5 °C and based on hourly data measured at a soil depth of 10.2 cm; Solar acc. = accumulated solar radiation measured in MJ·m<sup>-2</sup>; Soil thaw DOY, Budbreak DOY, Bloom DOY, and Veraison DOY = the DOY when the experimental unit reached soil thaw at depth of 10.2 cm, 50% budbreak, 50% bloom, and 50% veraison, respectively. Factors with cumulative quantities [GDD 10 (Hrly), GDD 10 (Avg), SDD 5, Solar acc.] for each phenological stage were measured from the accomplishment of the applicable threshold (soil thaw for budbreak and the previous phenological stage for bloom, veraison, and harvest).

<sup>b</sup>Model error (difference between predicted and actual days remaining until the phenological stage) is reported as the absolute difference (not positive or negative) because residuals were normally distributed for nearly all models evaluated and therefore had relatively even distribution of positive and negative differences from actual. N = 8226, 4570, 6861, and 6014 for budbreak, bloom, veraison, and harvest, respectively.

<sup>x</sup>Within each phenological stage, means followed by the same letter are not different according to Tukey-Kramer honestly significant difference test at  $P \leq 0.05$  (n = 8226, 4570, 6861, and 6014 for budbreak, bloom, veraison, and harvest, respectively).

0.98) in the training set and predicted the arrival of veraison within an average model error of <1.80 d and a largest model error of <5.9 d over a range from 1 to 67 d before the arrival of veraison.

**Harvest.** For prediction of harvest maturity, models created from data of all five cultivars showed a range of  $R^2$  values from 0.828 for the poorest performing basic models [built using cultivar with GDD 10 (Avg) or GDD 10 (Hrly)] to 0.869 for the highest performing multifactor model built by using cultivar and six other predictor variables [DOY, soil thaw DOY, photoperiod, GDD 10 (Hrly), SDD 5, solar acc.] (Table 4). The three selected multifactor models performed better than the basic models, showing higher  $R^2$  values and lower values for average model error, largest model error, AICc, and BIC than the basic models with normal residuals. Multifactor models for harvest that were based on data from one cultivar (‘Marquette’ only) had  $R^2$  values  $\geq 0.978$ , average model errors  $\leq 1.43$  d, largest model errors  $\leq 6.02$  d, AICc values  $\leq 3123$ , and BIC values  $\leq 3155$ , whereas the basic model built with GDD 10 (Hrly) had values for these metrics of 0.814, 4.54, 9.72, 4747, and 4761, respectively (Table 4).

Compared with the models for harvest created from data of all five cultivars, models created from data of ‘Marquette’ performed much better (Table 4), a result that indicated a higher level of variability in the timing of harvest maturity for one or more of the other cultivars used for building the five-cultivar models. Examination of the raw data showed that the standard deviation for harvest DOY of ‘Marquette’ was 3.63 d and was 1.63, 7.83, 6.14, and 6.24 d for ‘Frontenac’, ‘La Crescent’, ‘Petit Ami’, and ‘St. Croix’, respectively. These results suggest that model accuracy may be lower when used for prediction of harvest for three of the cultivars (La Crescent, Petit Ami, and St. Croix) and that creating a consistently strong model for these three cultivars may be difficult. Another possible contributor to the high variance for timing of harvest may be the need for growers to harvest within a range of maturity that works for them. For example, growers (including those of our research station) may choose to harvest earlier or later than the optimum to avoid unfavorable weather conditions or to ensure that they have an adequate harvest crew in place. Regardless of the exact cause of the higher variation in our harvest results for the three cultivars, the multifactor models still

performed substantially better than basic models for predicting the timing of harvest, and use of multiple factors strongly improved predictive power compared with models based on GDD only (Table 4). Based on the training datasets, the top multifactor models for the prediction of harvest explained more than 86% of the variance for five-cultivar models and more than 97% for ‘Marquette’ models ( $R^2 > 0.86$  and  $> 0.97$ , respectively), and predicted the arrival of harvest maturity within an average model error of <4 d or <1.5 d, respectively, and a largest model error of <10.7 d or <6.05 d, respectively, over a range from 1 to 51 d before the arrival of harvest.

#### Out-of-sample validation of models

Evaluation of model accuracy on an out-of-sample validation dataset is the best way to confirm the predictive power of regression models (Geisser, 1975; Picard and Cook, 1984; Stone, 1974). With our validation dataset recorded from ‘Marquette’ vines in 2019, the multifactor models performed much better than models based on only GDD 10 (Hrly) at all four phenological stages (Table 5). Average model errors for the basic model created using only GDD 10 (Hrly) were 10.05, 2.54, 4.23, and 4.96 d for budbreak,



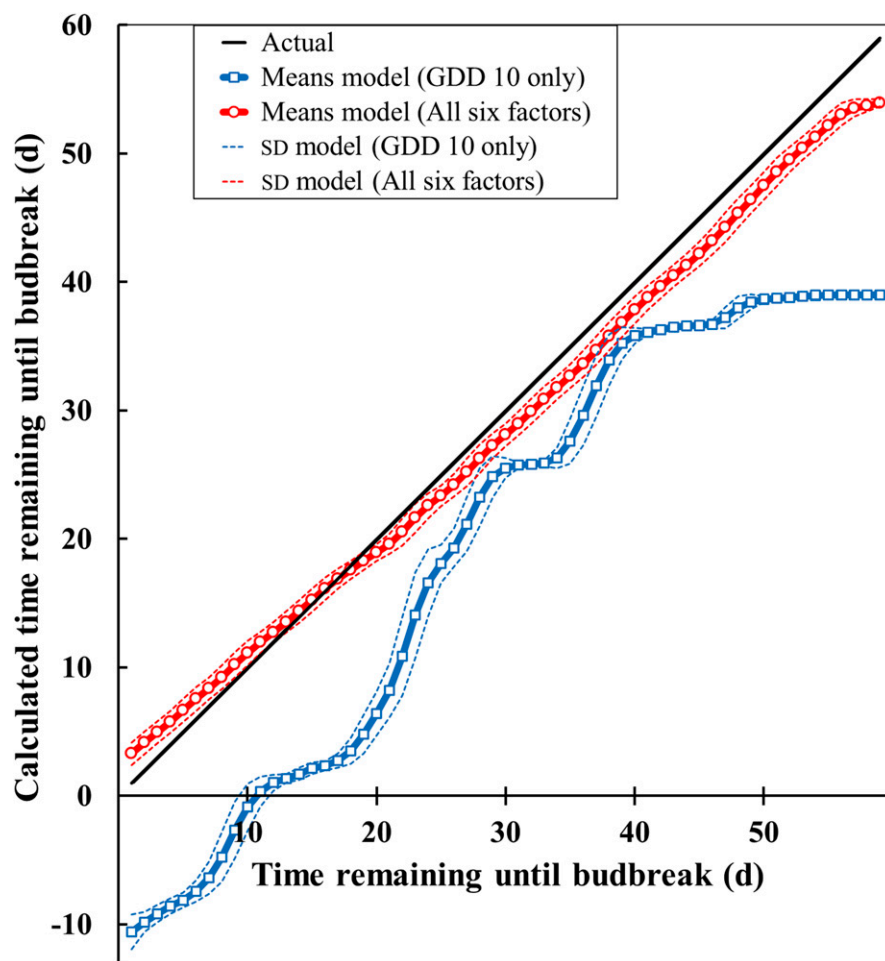


Fig. 1. Actual number of days remaining until budbreak compared with the mean number of days remaining as predicted by two of the models [the model based on only cultivar and GDD 10 (Hrly), and the model based on cultivar and all six additional variables] applied to the validation dataset recorded for ‘Marquette’ vines in 2019. Model equations were used to calculate predictions for the number of days remaining until budbreak based on values for variables from each day preceding budbreak, beginning with the day after soil thaw (a range from 1 to 59 d before budbreak). Means and standard deviations were calculated from values recorded for individual vines ( $n = 144$ ). The model labeled “GDD 10 only” contained predictor components for cultivar and GDD 10 (Hrly). The model labeled “All six factors” contained predictor components for cultivar and six additional independent variables [DOY, Soil thaw DOY, Photoperiod, GDD 10 (Hrly), SDD 5, and Solar acc.]. DOY = numerical day of the year; Soil thaw DOY = DOY on which soil thawed for the season; Photoperiod = daylength on DOY; GDD 10 (Hrly) = growing degree days with a base of 10 °C accumulated since soil thaw and calculated from hourly data; SDD 5 = soil degree days with a base of 5 °C accumulated since soil thaw; Solar acc. = solar radiation accumulated since soil thaw.

bloom, veraison, and harvest, respectively, whereas the average model errors for the selected multifactor models were much lower ( $\leq 3.79$ ,  $\leq 1.39$ ,  $\leq 1.07$ , and  $\leq 2.09$  d, respectively). Largest model errors for the basic model built on only GDD 10 (Hrly) were 20.99, 7.47, 9.13, and 10.06 d for budbreak, bloom, veraison, and harvest, respectively, and the largest model errors for the selected multifactor models were much lower ( $\leq 8.51$ ,  $\leq 5.75$ ,  $\leq 5.89$ , and  $\leq 7.10$  d, respectively). Of the multifactor models selected for validation, the models created using GDD 10 (Avg) instead of GDD 10 (Hrly) were the least accurate at all four phenological stages (Table 5). Although the average model error for the models using GDD 10 (Avg) was not significantly different from some of the other multifactor models at each phenological

stage, it was significantly greater than the average model error for the best multifactor model at all of the phenological stages. Among the selected multifactor models, those created from data of all five cultivars performed very well on the ‘Marquette’ validation set, having among the lowest values for both average model error and largest model error at all four phenological stages (Table 5).

**Budbreak.** For budbreak, the two top-performing models (the five-cultivar model with cultivar and all six predictor variables, and the ‘Marquette’ model with all six predictor variables) were nearly identical in effectiveness based on the validation dataset (Table 5). The values for average model error for these two models were within a few hundredths of a day of each other, as were

the values for largest model error. The best of these two models (the five-cultivar model with cultivar and all six predictor variables) predicted budbreak of the validation set within an average error of 1.92 d and a largest error of 5.94 d over a range from 1 to 59 d before the arrival of budbreak (Table 5). Plotting the actual number of days remaining until budbreak along with the mean number of days remaining as predicted by two of the five-cultivar models [the basic model with cultivar and GDD 10 (Hrly), and the multifactor model with cultivar and all six variables] provided an effective comparison of the performance of the models over the range of 1 to 59 d before budbreak (Fig. 1). Beginning at the top right corner (59 d before budbreak) and moving toward the bottom left corner of the plot (1 d before budbreak), it is evident that the multifactor model outperformed the basic model over the entire range of days preceding budbreak. The basic model underpredicted the timing of budbreak across the entire range of 58 d, and its poorest performance was during the early days of prediction (40 to 59 d before budbreak) and again as budbreak was approaching (<24 d before budbreak), where it commonly underpredicted the arrival of budbreak by more than 10 d. The multifactor model slightly underpredicted the days remaining until budbreak during the early days of prediction (20 to 59 d before budbreak) then slightly overpredicted the days remaining as budbreak was approaching (<15 d before budbreak), but the prediction mean for the multifactor model was within 3 d of the actual remaining number of days across the entire range of days <56 d preceding budbreak (Fig. 1).

**Bloom.** In evaluations of models for bloom, all five of the multifactor models selected for validation performed well, with only one multifactor model [the model with GDD 10 (Avg)] having an average model error that was significantly greater than that of the top-performing model (Table 5). Based on the values for average model error and largest model error, the five-cultivar model built without the photoperiod variable showed a slightly better performance than the other models, predicting timing of bloom for the validation set within an average error of 1.31 d and a largest error of 4.94 d over a range from 1 to 34 d before the arrival of bloom. In the plot comparing the actual number of days remaining until bloom with the daily mean number of days predicted by two of the five-cultivar models [the basic model with cultivar and GDD 10 (Hrly), and the multifactor model with cultivar and all six variables], the multifactor model outperformed the basic model over the entire range of days preceding bloom except for three of the days (21, 22, and 23 d before bloom) where the basic model was slightly more accurate (a few hundredths of a day) (Fig. 2). The models for bloom were generally more accurate than the models for budbreak, but the shape of the plots was similar. The basic model underpredicted the timing of bloom across the entire range of 33 d, and its



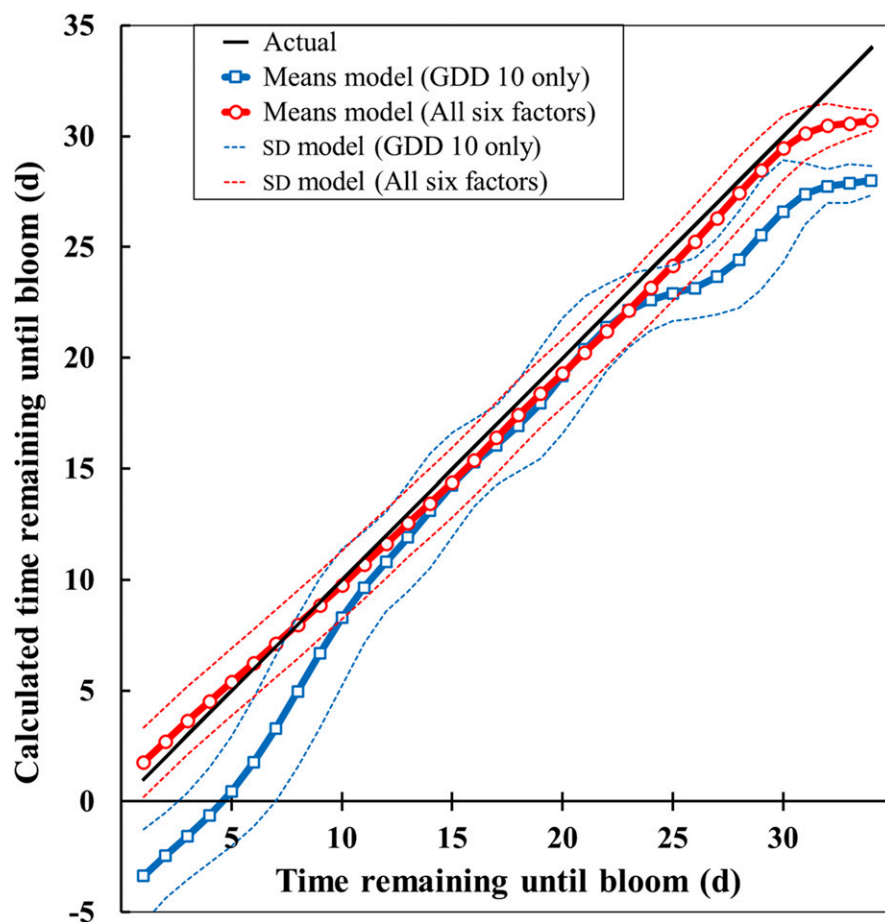


Fig. 2. Actual number of days remaining until bloom compared with the mean number of days remaining as predicted by two of the models [the model based on only cultivar and GDD 10 (Hrly), and the model based on cultivar and all six additional variables] applied to the validation dataset recorded for ‘Marquette’ vines in 2019. Model equations were used to calculate predictions for the number of days remaining until bloom based on values for variables from each day preceding bloom, beginning with the day after budbreak (a range from 1 to 34 d before bloom). Means and standard deviations were calculated from values recorded for individual vines ( $n = 144$ ). The model labeled “GDD 10 only” contained predictor components for cultivar and GDD 10 (Hrly). The model labeled “All six factors” contained predictor components for cultivar and six additional independent variables [DOY, Budbreak DOY, Photoperiod, GDD 10 (Hrly), SDD 5, and Solar acc.]. DOY = numerical day of the year; Budbreak DOY = DOY on which budbreak occurred for the season; Photoperiod = daylength on DOY; GDD 10 (Hrly) = growing degree days with a base of 10 °C accumulated since budbreak and calculated from hourly data; SDD 5 = soil degree days with a base of 5 °C accumulated since budbreak; Solar acc. = solar radiation accumulated since budbreak.

poorest performance was during the early days of prediction (25 to 34 d before bloom) and again as bloom was approaching (<12 d before bloom). The multifactor model slightly underpredicted the days remaining until bloom during the early days of prediction (>12 d before bloom), then slightly overpredicted the days remaining as budbreak was approaching (<6 d before bloom), but the mean predictions from the multifactor model were within 1.2 d of the actual days remaining across the entire range except for the earliest days of prediction (32, 33, 34 d before bloom) (Fig. 2).

**Veraison.** For veraison, the top-performing model was the five-cultivar model built on cultivar and four other predictor variables [DOY, bloom DOY, GDD 10 (Hrly), solar acc.], which predicted the timing of veraison for the validation set within an average error of 0.94 d and a largest error of 5.73 d over a range

from 1 to 48 d before the arrival of veraison (Table 5). The plot comparing the actual number of days remaining until veraison with the mean number of days predicted by two of the five-cultivar models [the basic model with cultivar and GDD 10 (Hrly), and the multifactor model with cultivar and four variables] showed very accurate predictions provided by the multifactor model across the entire range of 47 d (Fig. 3). The multifactor model slightly underpredicted the actual number of days remaining across the 47 d, with mean daily predictions that were all within 1 d of the actual number remaining except for those from the earliest two days of prediction (47 and 48 d before veraison). The basic model built from only cultivar and GDD 10 (Hrly) overpredicted the number of days remaining until budbreak across the entire 47 d, with mean daily predictions that overestimated by >5 d for the time period from

29 to 43 d before veraison, then gradually improved to overestimates of  $\approx 2.2$  d for the time period from 1 to 11 d before veraison (Fig. 3).

**Harvest.** For prediction of harvest, the best performing model evaluated on the validation dataset was the five-cultivar model built with cultivar and all six predictor variables. Even though the in-sample performance of this model was hindered by high variation in the timing of harvest for three of the five cultivars in the training dataset, the variation of ‘Marquette’ was relatively low, and therefore had little impact on the accuracy of the model when used on the ‘Marquette’ validation set. The five-cultivar multifactor model predicted the timing of harvest for the validation set within an average error of 1.67 d and a largest error of 6.27 d over a range from 1 to 41 d before the arrival of harvest (Table 5). In the plot comparing the actual number of days remaining until harvest with the daily mean number of days predicted by two of the five-cultivar models [the basic model with cultivar and GDD 10 (Hrly), and the multifactor model with cultivar and all six variables], the multifactor model performed much more consistently than the basic model (Fig. 4). The multifactor model slightly underestimated the time remaining until harvest (by <2 d) across the entire range of 40 d. The basic model performed well for the early days of prediction (20 to 41 d before harvest), then grew increasingly inaccurate as harvest approached, with its poorest performance (overestimation of  $\approx 6.6$  d) occurring 1 d before harvest (Fig. 4).

### Equations for the top-performing models

Equations for the top-performing models at each phenological stage are provided in Table 6, along with equations for models that may be useful if data for photoperiod and/or GDD 10 (Hrly) are unavailable. The prediction value for each model is the sum of three components: the cultivar adjustment (a value added or subtracted for the specific cultivar), variables with coefficients, and a constant. The cultivar adjustment component is of a type required for categorical variables, variables that can take on one of a limited, and usually fixed, number of possible values (Mukunthu et al., 2019; Statistics Knowledge Portal, 2020). The cultivar adjustment corrects for the intrinsic phenological differences among the cultivars in the five-cultivar models. The cultivar adjustment was shown to be very effective with the ‘Marquette’ validation set, where the five-cultivar model with the ‘Marquette’ cultivar adjustment performed as well or better at all four phenological stages than the model created specifically from ‘Marquette’ data (Table 5). The other two components (variables with coefficients and the constant) are typical for regression models. The number of significant digits included in model equations is deliberately high to maximize the accuracy of predictions (Table 6). The models are not intended to summarize the effects of the predictors, they are intended to predict the arrival of the phenological stages as accurately as possible based on the

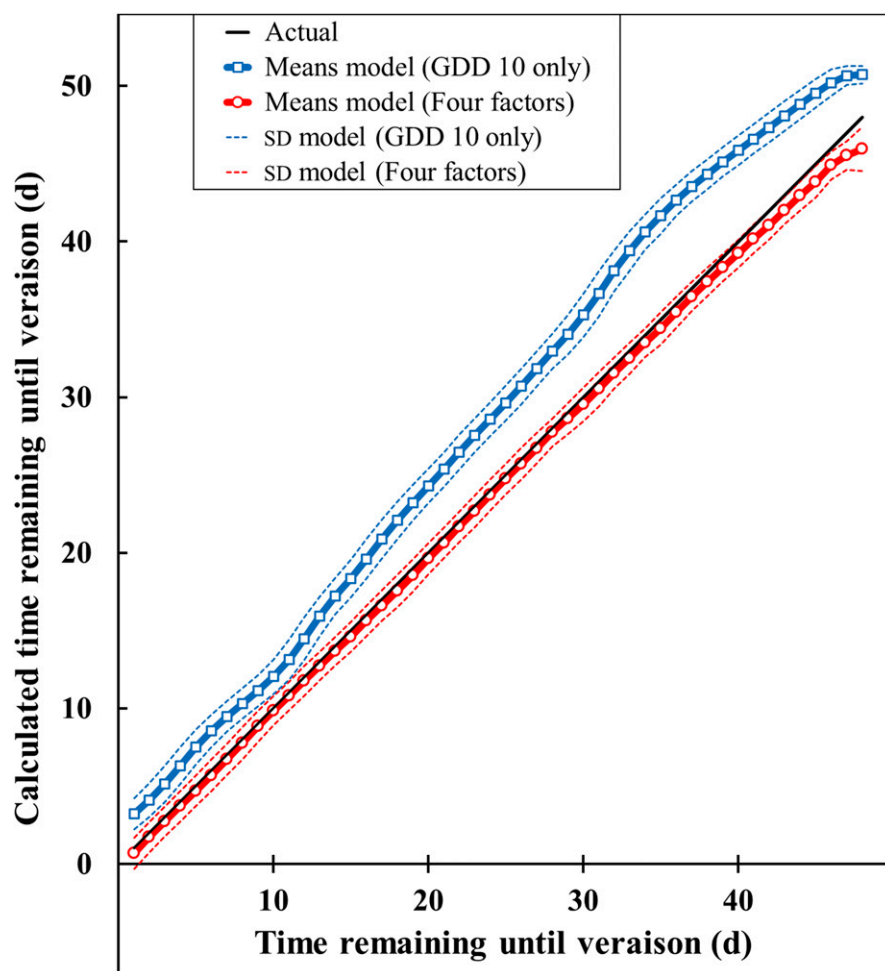


Fig. 3. Actual number of days remaining until veraison compared with the mean number of days remaining as predicted by two of the models [the model based on only cultivar and GDD 10 (Hrly), and the model based on cultivar and four additional variables] applied to the validation dataset recorded for 'Marquette' vines in 2019. Model equations were used to calculate predictions for the number of days remaining until veraison based on values for variables from each day preceding veraison, beginning with the day after bloom (a range from 1 to 48 d before veraison). Means and standard deviations were calculated from values recorded for individual vines ( $n = 144$ ). The model labeled "GDD 10 only" contained predictor components for cultivar and GDD 10 (Hrly). The model labeled "Four factors" contained predictor components for cultivar and four additional independent variables [DOY, Bloom DOY, GDD 10 (Hrly), and Solar acc.]. DOY = numerical day of the year; Bloom DOY = DOY on which bloom occurred for the season; GDD 10 (Hrly) = growing degree days with a base of 10 °C accumulated since bloom and calculated from hourly data; Solar acc. = solar radiation accumulated since bloom.

available data. The high number of significant digits is warranted, and they cause no difficulty when the equations are used with a spreadsheet program for automated calculation of time remaining until the phenological stage.

#### Automated calculation of models using spreadsheet programs

When used together in succession, the models for budbreak, bloom, veraison, and harvest can function as a four-stage, multifactor calculator for improved prediction of phenological timing. Spreadsheet programs such as Microsoft Excel (Microsoft Corp., Redmond, WA) contain tools that can be used to automate the calculation of predictions from model equations and provide a predicted date of arrival for each phenological stage. A simplified (non-Macro) example of

automated calculation using a spreadsheet program is provided to demonstrate how models can be used by growers and technicians who have a moderate level of expertise with spreadsheet software (Fig. 5). Using the top-performing five-cultivar model, the figure illustrates the spreadsheet components and mathematical functions required for calculation of days remaining until budbreak and the two-step conversion of this value to receive the predicted date on which budbreak will likely take place (Fig. 5). From left to right, the columns contain the name of the cultivar followed by the input for each of the factors as measured during the specified DOY. When using the model to predict budbreak for more than one cultivar at the same time, a separate column for cultivar adjustment is required. If the model is used for only

one cultivar, the value for cultivar adjustment for that specific cultivar can be included in the function for the calculation column, and the extra column for cultivar adjustment can be removed. The "calculation" column is the most complex. It contains the entire model equation in the form of a spreadsheet function that acts on the values contained in the input columns to calculate the predicted number of days until budbreak. The final two columns are used to convert the number of days until budbreak into an actual forecasted date for budbreak. The function in the first of those two columns adds the number of days remaining until budbreak to the current value for DOY, then rounds it off to receive the predicted DOY for budbreak rounded to the nearest full day. The function in the last column (far right) converts the predicted DOY to the calendar date on which budbreak is predicted to take place (Fig. 5).

#### Impact and application

Our results demonstrate the effectiveness and utility of multifactor models for predicting the timing of phenological stages for cold-climate wine grapes. Comparison of basic (cultivar plus one predictor variable) and multifactor models (cultivar plus four to six predictor variables) based on the same large dataset reveals the improvement in predictive power that can be attained when multiple predictor variables are used and scaled by the methods of multiple regression. At all four phenological stages evaluated in our research (budbreak, bloom, veraison, and harvest maturity), multifactor models were selected as the top performers according to elastic net regression metrics that penalize for added variables. In agreement with results from the metrics for variable selection, the multifactor models had lower mean model errors than basic models at all four phenological stages for both in-sample and out-of-sample evaluations.

Using multiple regression to predict the "number of days remaining" until a phenological stage arrives, rather than regressing the effects of variables on the DOY that the phenological stage arrived, facilitated the development of a real-time predictive mechanism for estimating the arrival of an upcoming phenological stage. This mechanism is a strong improvement over techniques based on a variable reaching a certain threshold (for example, monitoring GDD 10 and expecting budbreak at a threshold of 125). Compared with recent models reported by others, our multifactor models performed very well. Models created and evaluated by Zapata et al. (2017) that used GDD with base temperatures that were adjusted for each cultivar achieved their best predictions for budbreak, bloom, and veraison with mean errors for the calibration (training) set of 5.4, 3.0, and 6.6 d, respectively, across 16 cultivars. Best predictions for budbreak, bloom, and veraison based on our models had mean errors for the training set of 0.70, 0.84, and 1.77 d, respectively, across five cultivars. For the evaluation (validation) sets, the models created by Zapata et al. (2017) achieved their

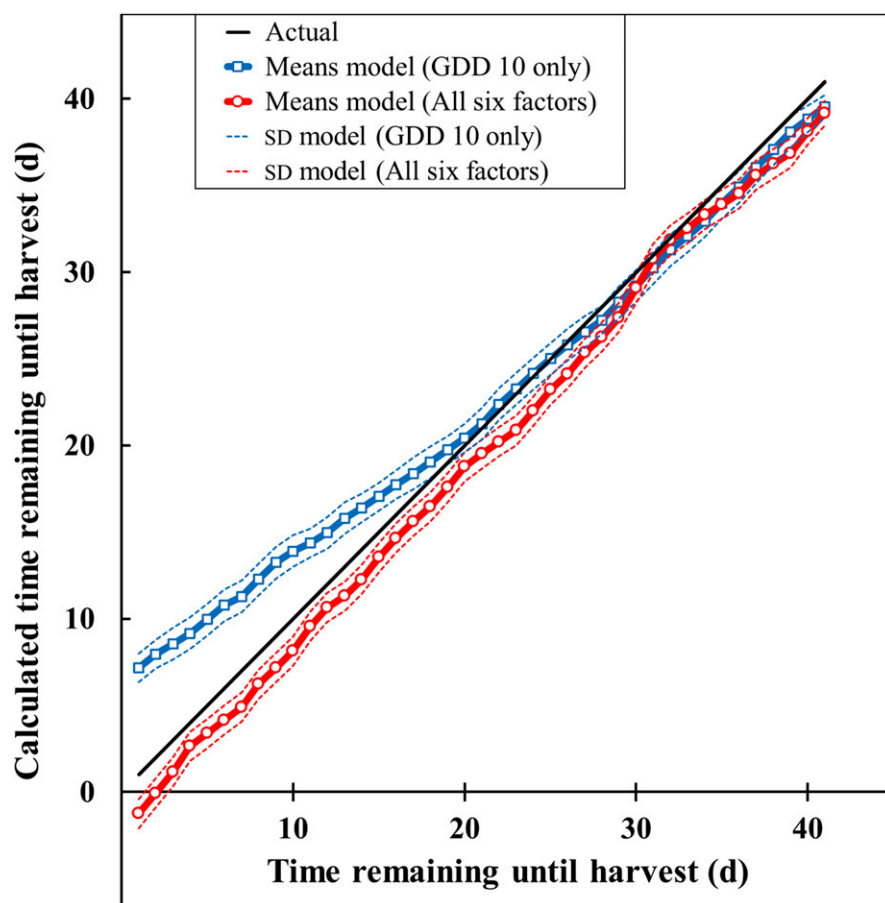


Fig. 4. Actual number of days remaining until harvest compared with the mean number of days remaining as predicted by two of the models [the model based on only cultivar and GDD 10 (Hrly), and the model based on cultivar and all six additional variables] applied to the validation dataset recorded for 'Marquette' vines in 2019. Model equations were used to calculate predictions for the number of days remaining until harvest based on values for variables from each day preceding harvest, beginning with the day after veraison (a range from 1 to 41 d before harvest). Means and standard deviations were calculated from values recorded for individual vines ( $n = 144$ ). The model labeled "GDD 10 only" contained predictor components for cultivar and GDD 10 (Hrly). The model labeled "All six factors" contained predictor components for cultivar and six additional independent variables [DOY, Veraison DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.]. DOY = numerical day of the year; Veraison DOY = DOY on which veraison occurred for the season; Photoperiod = daylength on DOY; GDD 10 (Hrly) = growing degree days with a base of 10 °C accumulated since veraison and calculated from hourly data; SDD 5 = soil degree days with a base of 5 °C accumulated since veraison; Solar acc. = solar radiation accumulated since veraison.

best predictions for budbreak, bloom, and veraison with mean errors of 5.6, 3.0, and 5.9 d, respectively, and our models achieved best predictions for budbreak, bloom, and veraison with mean errors of 1.92, 1.31, and 0.94 d, respectively, on the validation set. Although the two types of models evaluated different cultivars in different climates, and therefore should not be compared formally, this informal comparison of results demonstrates the high level of predictive accuracy achieved by our multifactor models.

As with all types of prediction or forecasting of future events, there are uncertainties inherent in attempting to make predictions of phenological timing by using the described models. Predictions based on phenological models or any other models are estimates that must be interpreted and used with a degree of caution (Box, 1979; Hyndman and Athanasopoulos, 2018; Rebba

et al., 2006). Strong models perform well when values for the predictor variables are not substantially different from those of the training set, but can perform poorly when values deviate widely from those of the training set (Burnham and Anderson, 2002; Heinze et al., 2018; Shmueli, 2010). Therefore, extra caution should be exercised with models during years when environmental variables are far from their norms.

For prediction of budbreak specifically, it is important to acknowledge the impact that the winter chilling requirement has on dormancy release, and therefore on the timing of budbreak. It is widely understood that budbreak is delayed when the chilling requirement (measured as hours with temps 0 and 7.2 °C) has not been met, because buds continue to exhibit some degree of endodormancy. Cumulative chilling hours for our research vineyards were greater than 1000 for

each of the winters preceding our evaluations. This indicates that ample chilling hours for overcoming endodormancy were met for the vines used in preparation and validation of our models because northern hybrids are low-chill species that require <1000 h of chilling (Londo and Johnson, 2014). Along with other assumptions that must be met for optimal accuracy of our models (such as similar and consistent timing of winter pruning), the condition of vines with regard to endodormancy and chilling hours must be considered when judging the potential accuracy of the model for prediction of budbreak. If required chilling hours have not been achieved, model estimates for budbreak should be used with greater caution.

The models provided in this report should be sufficient, accurate, and useful for managing the cultivars included in our project (Frontenac, La Crescent, Marquette, Petit Ami, and St. Croix) when grown in the midwestern United States, but they should not be considered to be universal for all cultivars or all growing regions. However, the methods of multiple regression can be used to create similar models for specific cultivars and regions to provide the most accurate predictions possible, and models could be continuously improved by adding to the dataset year after year (Neter et al., 1996; Shmueli, 2010). We recommend at least 3 years of data for preparation of models specific to other cultivars and/or regions. The inputs required for the multiple factors used in our top-performing models should be available in most areas. Values for these variables are either known by definition (cultivar and DOY) or can be compiled (soil thaw DOY, photoperiod, GDD 10, SDD 5, solar acc.) from commonly available databases (such as environmental Mesonet Web sites or climate data networks) or from private, dedicated weather stations that can be installed at vineyards for a moderate cost.

The multifactor models demonstrated in this report (and/or other potential models created for specific cultivars or regions) can be easily adapted for automated calculation using common spreadsheet software (Fig. 5). Automated calculation such as this can facilitate the practical use of multifactor models as a tool for crop management. The capacity to receive a reasonably accurate prediction for arrival of an upcoming phenological stage on a daily basis could be especially valuable for the planning and preparation for pruning, pesticide application, shoot and leaf positioning or removal, cluster thinning, and harvest. Predictions of budbreak would help growers gauge the potential for damage from late spring freeze events and would help guide decisions about timing of pruning and preparation of cold-protection methods for at-risk cultivars. Accurate prediction of phenological stages could improve accuracy when scheduling work crews at all stages of vine management and harvest, and insect and disease management plans could be developed to optimize pesticide applications by targeting the stages at which pests are most



Table 6. Equations of multifactor models for predicting the arrival of four key phenological stages (days remaining until budbreak, bloom, veraison, and harvest) of cold-climate wine grapes. Equations are included for the top-performing models at each phenological stage and other models that may be useful if data for photoperiod and/or GDD 10 (Hrly) are not available. Models were created by multiple linear regression and were based on 3 years of data (2011, 2013, and 2014) from five cultivars (Frontenac, La Crescent, Marquette, Petit Ami, and St. Croix) growing in central Iowa. Each equation includes a cultivar adjustment (a value added or subtracted for the specific cultivar), four to six variables with coefficients, and a constant.

Factors included <sup>a</sup>	Model equation
<b>Budbreak</b>	
Cultivar, DOY, Soil thaw DOY, GDD 10 (Avg), SDD 5, Solar acc.	(CULTIVAR: Frontenac, -1.01993727708912; La Crescent, -2.24504473186155; Marquette, -1.96684342864818; Petit Ami, 0.85444005165998; St. Croix, 0) + (-1.40590454880985 × DOY) + (0.35738862982858 × Soil Thaw DOY) + (-0.052822832729278 × GDD 10 (Avg)) + (-0.00827929232295 × SDD 5) + (0.0336195442732471 × Solar acc.) + (134.270314274879)
Cultivar, DOY, Soil thaw DOY, GDD 10 (Hrly), SDD 5, Solar acc.	(CULTIVAR: Frontenac, -1.02573798255656; La Crescent, -2.23553254712436; Marquette, -1.97077116313873; Petit Ami, 0.850935184883518; St. Croix, 0) + (-1.2461462437735 × DOY) + (0.185741969780306 × Soil Thaw DOY) + (-0.113952130163524 × GDD 10 (Hrly)) + (0.046370732387399 × SDD 5) + (0.0222764575409836 × Solar acc.) + (135.995089341884)
Cultivar, DOY, Soil thaw DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc. <sup>y</sup>	(CULTIVAR: Frontenac, -1.02470430016399; La Crescent, -2.24916944016656; Marquette, -1.98211314185571; Petit Ami, 0.863714874697284; St. Croix, 0) + (-2.41136571255827 × DOY) + (0.152194171942409 × Soil Thaw DOY) + (24.37271708590005 × Photoperiod) + (-0.134957195561917 × GDD 10 (Hrly)) + (0.0804770431283008 × SDD 5) + (0.0238725444616138 × Solar acc.) + (-66.4256807364779)
<b>Bloom</b>	
Cultivar, DOY, Budbreak DOY, GDD 10 (Avg), SDD 5, Solar acc.	(CULTIVAR: Frontenac, -0.955628757917193; La Crescent, 1.58312562269949; Marquette, 1.60199841506135; Petit Ami, -1.60732939728465; St. Croix, 0) + (-1.23985103940057 × DOY) + (1.61437791422849 × Budbreak DOY) + (0.00433342283031756 × GDD 10 (Avg)) + (0.00346124460855962 × SDD 5) + (0.00929235882372386 × Solar acc.) + (-20.2556272832215)
Cultivar, DOY, Budbreak DOY, GDD 10 (Hrly), SDD 5, Solar acc. <sup>y</sup>	(CULTIVAR: Frontenac, -0.933074188152312; La Crescent, 1.60097766918085; Marquette, 1.64452266568747; Petit Ami, -1.65183118707479; St. Croix, 0) + (-1.27461620671621 × DOY) + (1.66492679153369 × Budbreak DOY) + (-0.0072140037673947 × GDD 10 (Hrly)) + (0.00851476873590943 × SDD 5) + (0.0126135991535855 × Solar acc.) + (-22.3594374863197)
Cultivar, DOY, Budbreak DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc.	(CULTIVAR: Frontenac, -0.915252011030945; La Crescent, 1.59404777653477; Marquette, 1.64561510948753; Petit Ami, -1.64996076927766; St. Croix, 0) + (-1.23429946336684 × DOY) + (1.68611055351719 × Budbreak DOY) + (-2.20601314900715 × Photoperiod) + (-0.00874722811462959 × GDD 10 (Hrly)) + (0.00743245585534968 × SDD 5) + (0.0148663785802552 × Solar acc.) + (1.57504694741139)
<b>Veraison</b>	
Cultivar, DOY, Bloom DOY, GDD 10 (Avg), Solar acc.	(CULTIVAR: Frontenac, -3.38072822762651; La Crescent, -2.21673341416367; Marquette, -8.70700269937228; Petit Ami, 0.657698266214539; St. Croix, 0) + (-0.766940160746902 × DOY) + (-0.258239121186141 × Bloom DOY) + (0.00568369300137577 × GDD 10 (Avg)) + (-0.0142236840331889 × Solar acc.) + (223.293761406298)
Cultivar, DOY, Bloom DOY, GDD 10 (Hrly), Solar acc. <sup>y</sup>	(CULTIVAR: Frontenac, -3.44645227233593; La Crescent, -2.25904489748378; Marquette, -8.73548634409278; Petit Ami, 0.645850390672121; St. Croix, 0) + (-0.824353445797593 × DOY) + (-0.225104349200609 × Bloom DOY) + (0.00767645689048162 × GDD 10 (Hrly)) + (-0.0127790089607248 × Solar acc.) + (227.226237639516)
<b>Harvest</b>	
Cultivar, DOY, Veraison DOY, GDD 10 (Avg), SDD 5, Solar acc.	(CULTIVAR: Frontenac, 18.0210770689642; La Crescent, 5.58193874595989; Marquette, 9.73789679277882; Petit Ami, 0.717029608818815; St. Croix, 0) + (-0.497684071428533 × DOY) + (0.296891610683816 × Veraison DOY) + (0.0748411579130007 × GDD 10 (Avg)) + (-0.113043857490862 × SDD 5) + (0.0461242884282124 × Solar acc.) + (72.3322428401132)
Cultivar, DOY, Veraison DOY, GDD 10 (Hrly), SDD 5, Solar acc.	(CULTIVAR: Frontenac, 18.0344629645652; La Crescent, 5.66214305547262; Marquette, 9.76657952622938; Petit Ami, 0.691441211072622; St. Croix, 0) + (-0.526052439186431 × DOY) + (0.352509847156545 × Veraison DOY) + (0.0817383783949407 × GDD 10 (Hrly)) + (-0.106553270072933 × SDD 5) + (0.0378586656926828 × Solar acc.) + (66.3709894704132)
Cultivar, DOY, Veraison DOY, Photoperiod, GDD 10 (Hrly), SDD 5, Solar acc. <sup>y</sup>	(CULTIVAR: Frontenac, 18.1682918739522; La Crescent, 6.2466756725603; Marquette, 9.76554863591796; Petit Ami, 0.891179572436585; St. Croix, 0) + (-3.72934717065363 × DOY) + (0.554451403254974 × Veraison DOY) + (-70.7097780390826 × Photoperiod) + (0.144932729339421 × GDD 10 (Hrly)) + (-0.172014188318456 × SDD 5) + (0.0725028240892792 × Solar acc.) + (1727.12210529748)

<sup>a</sup>Factor abbreviations and definitions: DOY = numerical day of the year; Photoperiod = daylength on DOY measured in hours; GDD 10 (Hrly) = growing degree days with a base of 10 °C and calculated from hourly data; GDD 10 (Avg) = growing degree days with a base of 10 °C and calculated from daily average; SDD 5 = soil degree days with a base of 5 °C and based on hourly data measured at a soil depth of 10.2 cm; Solar acc. = accumulated solar radiation measured in MJ·m<sup>-2</sup>; Soil thaw DOY, Budbreak DOY, Bloom DOY, and Veraison DOY = the DOY when the experimental unit reached soil thaw at depth of 10.2 cm, 50% budbreak, 50% bloom, and 50% veraison, respectively. Factors with cumulative quantities [GDD 10 (Hrly), GDD 10 (Avg), SDD 5, Solar acc.] for each phenological stage were measured from the accomplishment of the applicable threshold (soil thaw for budbreak and the previous phenological stage for bloom, veraison, and harvest).

<sup>y</sup>The top-performing model for each phenological stage based on in-sample performance and out-of-sample validation.

## Model equation

$$\text{Days until budbreak} = (\text{CULTIVAR: Frontenac, -1.02470430016399; La Crescent, -2.24916944016656; Marquette, -1.98211314185571; Petit Ami, 0.863714874697284; St. Croix, 0}) + (-2.41136571255827 \times \text{DOY}) + (0.152194171942409 \times \text{Soil Thaw DOY}) + (24.3727170859005 \times \text{Photoperiod}) + (-0.134957195561917 \times \text{GDD 10 (Hrly)}) + (0.0804770431283008 \times \text{SDD 5}) + (0.0238725444616138 \times \text{Solar acc.}) + (-66.4256807364779)$$

## Functions in specified cells

$$= H2 + (-2.41136571255827 * B2) + (0.152194171942409 * C2) + (24.3727170859005 * D2) + (-0.134957195561917 * E2) + (0.0804770431283008 * F2) + (0.0238725444616138 * G2) + (-66.4256807364779)$$

$$= \text{ROUND}((B2 + I2), 0)$$

$$= \text{DATE}(2021, 1, J2)$$

## Spreadsheet

	A	B	C	D	E	F	G	H	I	J	K
1	Cultivar	DOY	Soil Thaw DOY	Photoperiod	GDD 10 (Hrly)	SDD 5	Solar acc.	Cultivar adjustment	Predicted days until budbreak	Predicted DOY of budbreak	Predicted date of budbreak
2	Frontenac	98	76	13.071389	21.452083	16.0672	291.3038	-1.02470430016399	31.73988696	130	5/10/2021
3	La Crescent	98	76	13.071389	21.452083	16.0672	291.3038	-2.24916944016656	30.51542182	129	5/9/2021
4	Marquette	98	76	13.071389	21.452083	16.0672	291.3038	-1.98211314185571	30.78247812	129	5/9/2021
5	Petit Ami	98	76	13.071389	21.452083	16.0672	291.3038	0.863714874697284	33.62830613	132	5/12/2021
6	St. Croix	98	76	13.071389	21.452083	16.0672	291.3038	0.0	32.76459126	131	5/11/2021
7											

Fig. 5. An example of automated calculation using a spreadsheet program with a multifactor model to predict the arrival of a phenological stage. The example uses the top-performing, five-cultivar model for prediction of days remaining until budbreak. The full model equation is included at the top of the figure. The lower portion of the figure is an illustration of the spreadsheet components and mathematical functions required for calculation of days remaining until budbreak and the two-step conversion of this value to receive the predicted date on which budbreak will likely take place. When using the model to predict budbreak for more than one cultivar at the same time, as in this example, a separate column for cultivar adjustment is required. If the model is used for only one cultivar, the value for cultivar adjustment for that specific cultivar can be included in the function for the calculation column, and the extra column for cultivar adjustment can be removed. DOY = numerical day of the year; Soil thaw DOY = DOY on which soil thawed for the season; Photoperiod = daylength on DOY; GDD 10 (Hrly) = growing degree days with a base of 10 °C accumulated since soil thaw and calculated from hourly data; SDD 5 = soil degree days with a base of 5 °C accumulated since soil thaw; Solar acc. = solar radiation accumulated since soil thaw measured in MJ-m<sup>-2</sup>.

active based on phenology and environmental data. In these and other ways, prediction models could be used as part of a viticulture management system to help improve efficiency and sustainability, reduce waste, and increase profitability.

## Literature Cited

- Akaike, H. 1974. A new look at the statistical model identification, p. 215–222. In: E. Parzen, K. Tanabe, and G. Kitagawa (eds.). *Selected papers of Hirotugu Akaike*. Springer Series in Statistics (Perspectives in Statistics). Springer, New York, NY.
- Anderson, W.K., R.C.G. Smith, and J.R. McWilliam. 1978. A systems approach to the adaptation of sunflower to new environments I: Phenology and development. *Field Crops Res.* 1:141–152.
- Basler, D. and C. Körner. 2014. Photoperiod and temperature responses of bud swelling and bud burst in four temperate forest tree species. *Tree Physiol.* 34:377–388.
- Berk, R.A. 2008. *Statistical learning from a regression perspective*. Springer, New York, NY.
- Bock, A., T. Sparks, N. Estrella, and A. Menzel. 2011. Changes in the phenology and composition of wine from Franconia, Germany. *Clim. Res.* 50:69–81.
- Boehmke, B. and B.M. Greenwell. 2019. *Hands-on machine learning with R*. CRC Press, Boca Raton, FL.
- Box, G.E.P. 1979. Robustness in the strategy of scientific model building, p. 201–236. In: R.L. Launer and G.N. Wilkinson. (eds.). *Robustness in statistics*. Academic Press, Cambridge, MA.
- Burnham, K.P. and D.R. Anderson. 2002. *Model selection and multimodel inference: A practical information-theoretic approach*. Springer, New York, NY.
- Constable, G.A. and I.A. Rose. 1988. Variability of soybean phenology response to temperature, daylength and rate of change in daylength. *Field Crops Res.* 18:57–69.
- Dettling, M. 2015. *Applied statistical regression*, AS 2015. Eidgenössische Technische Hochschule, Zürich. 3 June 2020. <https://stat.ethz.ch/education/semesters/as2015/asr/Script\_v151119.pdf>.
- Dhamadhikari, M.R. and K.L. Wilker. 2001. *Micro vinification: A practical guide to small scale wine production*. Midwest Viticult. Enol. Ctr., Southwest Missouri State Univ., Mountain Grove, MO.
- Domoto, P. 2014. Pruning grape vines - Evaluating and adjusting for cold injury. *Wine Growers News* #261. Iowa State Univ. Ext., Ames, IA.
- Dormann, C.F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J.R.G. Marquéz, B. Gruber, B. Lafourcade, P.J. Leitão, T. Münkemüller, C. McClean, P.E. Osborne, B. Reineking, B. Schröder, A.K. Skidmore, D. Zurell, and S. Lautenbach. 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36:27–46.
- Dry, P. and B. Coombe. 2004. *Grapevine growth stages - The modified E-L system*. Viticulture 1 - Resources. 2nd ed. Winetitles Media, Broadview, Australia.
- Dunkler, D., M. Plischke, K. Leffondré, and G. Heinze. 2014. Augmented backward elimination: A pragmatic and purposeful way to develop statistical models. *PLoS One* 9:e113677. doi: 10.1371/journal.pone.0113677.
- Eichhorn, K.W. and D.H. Lorenz. 1977. *Phänologische entwicklungsstadien der rebe*. Nachrichtenblatt des Deutschen Pflanzenschutzdienstes. Braunschweig 29:119–120.
- Fernández-González, M., F.J. Rodríguez-Rajo, O. Escuredo, and M.J. Aira. 2013. Influence of thermal requirement in the aerobiological and phenological behavior of two grapevine varieties. *Aerobiologia* 29:523–535.
- Fraga, H., J.A. Santos, J. Moutinho-Pereira, C. Carlos, J. Silvestre, J. Eiras-Dias, T. Mota, and A.C. Malheiro. 2016. Statistical modelling of grapevine phenology in Portuguese wine regions: Observed trends and climate change projections. *J. Agr. Sci., Cambridge* 154: 795–811.
- Frost, J. 2020. Multicollinearity in regression analysis: Problems, detection, and solutions. 28 May 2020. <https://statisticsbyjim.com/regression/multicollinearity-in-regression-analysis/>.
- García de Cortázar-Atauri, I., N. Brisson, and J.P. Gaudillere. 2009. Performance of several models for predicting budburst date of grapevine (*Vitis vinifera* L.). *Intl. J. Biometeorol.* 53:317–326.
- Geisser, S. 1975. The predictive sample reuse method with applications. *J. Amer. Stat. Assn.* 70:320–328.
- Gentilucci, M. and P. Burt. 2018. Using temperature to predict the end of flowering in the common grape (*Vitis vinifera*) in the Macerata wine region, Italy. *Euro-Mediterranean J. Environ. Intl.* 3:38, doi: 10.1007/s41207-018-0079-4.
- Greer, D.H., J.N. Wünsche, C.L. Norling, and H.N. Wiggins. 2006. Root-zone temperatures affect phenology of bud break, flower cluster development, shoot extension growth and gas

- exchange of 'Braeburn' (*Malus domestica*) apple trees. *Tree Physiol.* 26:105–111.
- Gu, S. 2016. Growing degree hours - a simple, accurate, and precise protocol to approximate growing heat summation for grapevines. *Intl. J. Biometeorol.* 60:1123–1134.
- Heinze, G., C. Wallisch, and D. Dunkler. 2018. Variable selection – a review and recommendations for the practicing statistician. *Biometrical J.* 60:431–449.
- Hoover, E., S. Wold-Burkness, J. Hilton, D. Mollov, E. Burkness, T. Galvan, P. Hemstad, and W. Hutchison. 2011. Grape IPM guide for Minnesota producers. Univ. Minnesota Ext. 8 Jan. 2020. <<https://conservancy.umn.edu/bitstream/handle/11299/166875/Grape%20IPM%20Guide.pdf?sequence=1&isAllowed=y>>.
- Hyndman, R.J. and G. Athanasopoulos. 2018. *Forecasting: Principles and practice*. 2nd ed. OTexts. 1 June 2020. <<https://otexts.org/fpp2/>>.
- Iowa State University. 2019. Iowa Environmental Mesonet. 18 Nov. 2019. <<https://mesonet.agron.iastate.edu/>>.
- Iowa State University. 2020. Growing degree days and applications. Department of Agronomy. 5 June 2020. <<http://agron-www.agron.iastate.edu/courses/Agron541/classes/541/lesson02b/2b.1.1.html>>.
- Kassambara, A. 2018. Regression model validation. *STHDA*. 27 May 2020. <<http://www.sthda.com/english/articles/38-regression-model-validation>>.
- Kelly, R. 2014. Linear model selection and regularization. 4 June 2020. <[https://rstudio-pubs-static.s3.amazonaws.com/22067\\_48fad02fb1a944e9a8fb1d56c55119ef.html](https://rstudio-pubs-static.s3.amazonaws.com/22067_48fad02fb1a944e9a8fb1d56c55119ef.html)>.
- Kliewer, W.M. 1975. Effect of root temperature on budbreak, shoot growth, and fruit-set of 'Cabernet Sauvignon' grapevines. *Amer. J. Enol. Viticult.* 26:82–89.
- Konishi, S. and G. Kitagawa. 2007. *Information criteria and statistical modeling*. Springer, New York, NY.
- Londo, J.P. and L.M. Johnson. 2014. Variation in the chilling requirement and bud burst rate of wild *Vitis* species. *Environ. Expt. Bot.* 160:138–147.
- Makridakis, S.G., S.C. Wheelwright, and R.J. Hyndman. 1998. *Forecasting: Methods and applications*. 3rd ed. Wiley, New York, NY.
- Martens, H. and T. Naes. 1989. *Multivariate calibration*. Wiley, New York, NY.
- Matas, A.J., G. López-Casado, J. Cuartero, and A. Heredia. 2005. Relative humidity and temperature modify the mechanical properties of isolated tomato fruit cuticles. *Amer. J. Bot.* 92:462–468.
- Meloun, M., J. Milítký, M. Hill, and R.G. Brereton. 2002. Crucial problems in regression modelling and their solutions. *Analyst* 127:433–450.
- Minnesota Grape Growers Assn. 2016. *Growing grapes in Minnesota*. 10th ed. revised by P. Domoto, C. Anderson, M. Clark, and I. Geary. 9 Dec. 2019. <<https://www.mngrapes.org/page/GrowingGrapes>>.
- Mukunthu, D., P. Shah, and W.H. Tok. 2019. Practical automated machine learning on Azure: Using Azure machine learning to build AI solutions. O'Reilly Media, Inc., Champaign, IL.
- Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman. 1996. *Applied linear statistical models*. 4th ed. WCB McGraw-Hill, New York, NY.
- Picard, R. and R. Cook. 1984. Cross-validation of regression models. *J. Amer. Stat. Assn.* 79: 575–583.
- Rebba, R., S. Mahadevan, and S. Huang. 2006. Validation and error estimation of computational models. *Reliab. Eng. Syst. Saf.* 91:1390–1397.
- Rezazadeh, A. and E.T. Stafne. 2018. Effect of chilling and photoperiod on budbreak in three hybrid grape cultivars. *HortTechnology* 28:737–742.
- SAS Institute Inc. 2020. JMP statistical discovery: Fitting linear models. 8 June 2020. <<https://www.jmp.com/support/help/en/15.1/index.shtml#page/jmp/model-specification.shtml#262668>>.
- Schaber, J. and F.W. Badeck. 2003. Physiology-based phenology models for forest tree species in Germany. *Intl. J. Biometeorol.* 47:193–201.
- Schrader, J.A., D.R. Cochran, P.A. Domoto, and G.R. Nonnecke. 2019. Phenology and winter hardiness of cold-climate grape cultivars and advanced selections in Iowa climate. *HortTechnology* 29:906–922.
- Schrader, J.A., D.R. Cochran, P.A. Domoto, and G.R. Nonnecke. 2020. Yield and berry composition of cold-climate grape cultivars and advanced selections in Iowa climate. *HortTechnology* 30:193–203.
- Shmueli, G. 2010. To explain or to predict? *Stat. Sci.* 25:289–310.
- Smiley, L.A., D. Cochran, P. Domoto, G. Nonnecke, and W.W. Miller. 2016. A review of cold climate grape cultivars. *Iowa State Univ. Ext. Publ. Hort.* 3040.
- Statistics Knowledge Portal. 2020. Multiple linear regression with categorical predictors. 30 June 2020. <[https://www.jmp.com/en\\_us/statistics-knowledge-portal/what-is-multiple-regression/mlr-with-categorical-predictors.html](https://www.jmp.com/en_us/statistics-knowledge-portal/what-is-multiple-regression/mlr-with-categorical-predictors.html)>.
- Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. [Ser. A]* 36:111–147.
- Sunrise-sunset.org. 2020. Sunset and sunrise times. 20 Apr. 2020. <<https://sunrise-sunset.org>>.
- USDA. 2019. USDA plant hardiness zone map. 8 June 2020. <<https://planthardiness.ars.usda.gov/PHZMWeb/InteractiveMap.aspx>>.
- Vaughan, T.S. and K.E. Berry. 2005. Using Monte Carlo techniques to demonstrate the meaning and implications of multicollinearity. *J. Stat. Educ.* 13:1, doi: 10.1080/10691898.2005.11910640.
- Verdugo-Vásquez, N., C. Pañitru-De la Fuente, and S. Ortega-Farías. 2017. Model development to predict phenological scale of table grapes (cvs. Thompson, Crimson and Superior Seedless, and Red Globe) using growing degree days. *OENO One* 51(3):277–288, doi: 10.20870/oeno-one.2017.51.2.1833.
- Way, D.A. and R.A. Montgomery. 2015. Photoperiod constraints on tree phenology, performance and migration in a warming world. *Plant Cell Environ.* 38:1725–1736.
- Webb, L., P. Whetton, J. Bhend, R. Darbyshire, P.R. Briggs, and E.W.R. Barlow. 2012. Earlier wine-grape ripening driven by climatic warming and drying and management practices. *Nat. Clim. Chang.* 2:259–264.
- Weikai, Y. and D.H. Wallace. 1998. Simulation and prediction of plant phenology for five crops based on photoperiod×temperature interaction. *Ann. Bot.* 81:705–716.
- Williams, D.W., H.L. Andris, R.H. Beede, D.A. Luvisi, M.V.K. Norton, and L.E. Williams. 1985. Validation of a model for the growth and development of the Thompson Seedless grapevine. II. Phenology. *Amer. J. Enol. Viticult.* 36:283–289.
- Zapata, D., M. Salazar-Gutierrez, B. Chaves, M. Keller, and G. Hoogenboom. 2017. Predicting key phenological stages for 17 grapevine cultivars (*Vitis vinifera* L.). *Amer. J. Enol. Viticult.* 68:60–72.
- Zou, H. and T. Hastie. 2005. Regularization and variable selection via the elastic net. *J. R. Stat. Soc.* 67:301–320.