

Article

NYUS.2: an automated machine learning prediction model for the large-scale real-time simulation of grapevine freezing tolerance in North America

Hongrui Wang^{1,*}, Gaurav D. Moghe², Al P. Kovaleski³, Markus Keller⁴, Timothy E. Martinson¹, A. Harrison Wright⁵, Jeffrey L. Franklin⁵, Andréanne Hébert-Haché⁶, Caroline Provost⁶, Michael Reinke⁷, Amaya Atucha³, Michael G. North³, Jennifer P. Russo¹, Pierre Helwi⁸, Michela Centinari⁹ and Jason P. Londo^{1,*}

¹School of Integrative Plant Science, Horticulture Section, Cornell AgriTech, Cornell University, Geneva, NY 14456, USA

²School of Integrative Plant Science, Plant Biology Section, Cornell University, Ithaca, NY 14850, USA

³Plant and Agroecosystem Sciences Department, University of Wisconsin–Madison, Madison, WI 53706, USA

⁴Department of Viticulture and Enology, Irrigated Agriculture Research and Extension Center, Washington State University, Prosser, WA 99350, USA

⁵Kentville Research and Development Centre, Agriculture and Agri-Food Canada, Kentville, Nova Scotia, B4N 1J5, Canada

⁶Centre de Recherche Agroalimentaire de Mirabel, Mirabel, Québec, J7N 2X8, Canada

⁷Southwest Michigan Research and Extension Center, Michigan State University, Benton Harbor, MI 49022, USA

⁸Martell & Co., 7 place Edouard Martell, Cognac 16100, France

⁹Department of Plant Science, The Pennsylvania State University, University Park, PA 16802, USA

*Corresponding authors. E-mail: Hongrui Wang, hw692@cornell.edu; Jason P. Londo, jpl275@cornell.edu

Abstract

Accurate and real-time monitoring of grapevine freezing tolerance is crucial for the sustainability of the grape industry in cool climate viticultural regions. However, on-site data are limited due to the complexity of measurement. Current prediction models underperform under diverse climate conditions, which limits the large-scale deployment of these methods. We combined grapevine freezing tolerance data from multiple regions in North America and generated a predictive model based on hourly temperature-derived features and cultivar features using AutoGluon, an automated machine learning engine. Feature importance was quantified by AutoGluon and SHAP (SHapley Additive exPlanations) value. The final model was evaluated and compared with previous models for its performance under different climate conditions. The final model achieved an overall 1.36°C root-mean-square error during model testing and outperformed two previous models using three test cultivars at all testing regions. Two feature importance quantification methods identified five shared essential features. Detailed analysis of the features indicates that the model has adequately extracted some biological mechanisms during training. The final model, named NYUS.2, was deployed along with two previous models as an R shiny-based application in the 2022–23 dormancy season, enabling large-scale and real-time simulation of grapevine freezing tolerance in North America for the first time.

Introduction

The global and regional distribution of perennial plants is primarily constrained by abiotic stresses associated with regional climate. For example, the cultivation of European grapevines (*Vitis vinifera*) in mid-winter cold regions in North America presents a significant challenge, as the minimum temperatures in these regions sometimes exceed the plants' maximum freezing tolerance. Cold-related damage is therefore a major limiting factor for the grape and wine industries in these regions [1–4]. Along with the development of preventative cultural practices and the introduction of cultivars with improved freezing tolerance, the monitoring of grapevine freezing tolerance has been a major focus of research groups in cool climate viticultural regions in North America to support the sustainability of the grape and wine industries.

Currently, the monitoring programs for grapevine freezing tolerance mainly rely on measuring the dormant bud low temperature exotherm (LTE), a burst of heat released when intracellular ice formation occurs, using a method called differential thermal analysis (DTA) [5]. DTA is conducted in programmable freezers, with buds exposed to a gradual decrease in temperature from 0°C to a lethal temperature (–40 to –50°C) at specific rates (e.g. –4°C·h^{–1}), and LTEs are recorded as a voltage change by a thermoelectric module placed under sample plates [6, 7]. Although this method facilitates a rapid assessment of bud freezing tolerance compared to a visual assessment of oxidative browning [4, 6, 8], the whole procedure, along with sample collection and preparation typically takes days, which eliminates its large-scale application with high temporal resolution. As a result, most grapevine freezing tolerance monitoring systems established in cool climate viticultural regions in North America only report weekly or

Received: 28 August 2023; Accepted: 17 December 2023; Published: 29 December 2023; Corrected and Typeset: 1 February 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nanjing Agricultural University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

bi-weekly updated freezing tolerance of a few cultivars collected from a few sites. Due to the logistical constraints of this system, current DTA-based monitoring systems may not be sufficient for growers when climate change-associated weather extremes occur more frequently [9]. An advanced grapevine freezing tolerance prediction system that accommodates multi-site and multi-cultivar freezing tolerance with high temporal resolution could serve not only as a practical tool for vineyard management, but also as a guideline for the adaptation of viticulture in the era of climate change.

To this end, there has been an increased interest in developing prediction models to mathematically estimate grapevine freezing tolerance under field conditions. The most widely used models, WAUS.1 and its successor WAUS.2 (developed with more cultivars, refined initial and minimum freezing tolerance, altered deacclimation calculation, and generalized chilling degree days), are discrete-dynamic models, where daily changes in the lethal temperature for 50% of a bud population (bud LT_{50}) are phased with incremental time steps [10, 11]. These models estimate daily freezing tolerance change based on daily maximum and minimum temperatures and cultivar-specific parameters [10, 11]. A derivative of WAUS.2 was also developed for cold climate inter-specific cultivars in Wisconsin using local freezing tolerance data [12]. A new model, NYUS.1, is a biological model recently developed using phased integration of cold acclimation and deacclimation responses based on recent findings of dormancy-related grapevine freezing tolerance dynamics [13–15]. Two issues limit the widespread application of these models. First, the parameters used in the WAUS.2 and NYUS.1 models were optimized using local freezing tolerance data in Washington and New York, respectively, which may contribute to overfitting under local climate conditions and underperformance under other climate conditions. For example, the underperformance of WAUS.2 has been observed in New York, Wisconsin, and in subtropical viticultural regions in Chile [12, 14, 16]. Second, since these models are the mathematical realizations of the biological understanding of dormancy and freezing tolerance, the parameters in the models such as the dynamics of cold acclimation and deacclimation need to be updated regularly according to the advances in dormancy and freezing tolerance biology. The version update of these models can involve substantial manual parameter generation, selection, and tuning, which in turn compromises the deployment of the most updated version.

Machine learning (ML) and its subfield deep learning (DL) have significantly advanced in the last decade [17, 18]. ML algorithms are typically used for structured data, such as numerical data in tables, which usually involves manual feature extraction. DL algorithms are capable of processing unstructured data, such as images, audio, text, and sequence, which often require higher-level abstractions and more training data to be fully functional [19]. Regarding the modeling of weather-related plant physiological responses, large-scale applications of ML and DL are used for growth and yield estimation of major crops, such as corn, wheat, and soybean [20–24]. In the field of viticulture, DL tools, particularly computer vision tools such as convolutional neural networks, have facilitated the development of precision viticulture in yield estimation, water potential quantification, berry number estimation, variety recognition, disease detection, disease quantification, fungicide efficacy, and phenotyping for grape breeding [25–33]. Researchers recently used a DL tool, RNN (recurrent neural network), as a backbone to model grapevine freezing tolerance. Being trained with Washington state data, the RNN model performed equivalently or better for some cultivars than

WAUS.2 [34]. The model has yet to be evaluated in other regions due to ongoing development. Nevertheless, the modeling of plant physiology using ML is still at very early stages. A leading reason is that the rapid development of ML can make it difficult even for ML experts to efficiently incorporate novel practices and timely deployed ML models. In addition, training datasets are usually small due to the complex and time-consuming methods to measure plant physiological responses, such as measuring freezing tolerance with DTA [35, 36]. The individual prediction models generated from such small datasets are subjected to overfitting, which limits their applications in real-world scenarios [37]. The development of automated machine learning (Auto-ML) frameworks addresses some of these issues. These frameworks not only fuse all the tedious steps of ML model development, including data preprocessing, feature engineering, model training, and hyperparameter tuning into very few lines of code, but also incorporate novel techniques such as model stacking and weighting to minimize overfitting. These tools enable domain experts, such as plant physiologists, to build ML applications through an automated pipeline without much requirement for coding and ML knowledge [38].

In this study, we combined the on-site measurement of freezing tolerance data (LT_{50}) from multiple regions in North America and labeled each datapoint with 117 features generated from hourly temperatures and cultivar information. The dataset was used for the training of a prediction model through an Auto-ML framework, AutoGluon [39]. AutoGluon is hosted by Amazon Web Services and is known for its ease of use and high performance through built-in multi-layer stacking and weighting techniques [39, 40]. The first objective of this study was to build a generalizable prediction model that accurately estimates the grapevine freezing tolerance under different climate conditions. The second objective was to quantify the importance of each feature in the final model to understand the physiological control of freezing tolerance in grapevine. The third objective was to apply the model to cool climate viticultural regions to achieve a large-scale and real-time monitoring of grapevine freezing tolerance with high temporal resolution.

Results

Data collection, method suitability test, and site-transferability test

The dataset combines the LT_{50} measurements from the nine grapevine freezing tolerance programs in New York (NY), Michigan (MI), Wisconsin (WI), Washington (WA), Pennsylvania (PA), and Texas (TX) in the USA and British Columbia (BC), Québec (QC), and Nova Scotia (NS) in Canada between 2002 and 2023 (Fig. 1a, Table S1). These regions cover the major types of climate conditions in the cool climate viticultural regions in North America. The dataset contains 3458 LT_{50} measurements from 24 *Vitis* sp. hybrid cultivars and 6699 LT_{50} measurements from 21 *V. vinifera* cultivars, resulting in a total of 10157 datapoints (Fig. 1b and c). The 45 cultivars in this dataset represent a large proportion of the cultivars grown in North America. The full model dataset contains 117 features. The features used for modeling in this study are within two categories: cultivar features (Boolean-type features) and hourly temperature-based features, except for days in season, which is the number of days after 1 September.

The Python package AutoGluon (version 0.7.0) [39] was used to conduct model training. Since this study represents the first approach to grapevine freezing tolerance modeling using Auto-ML, the first step was to test for method suitability and

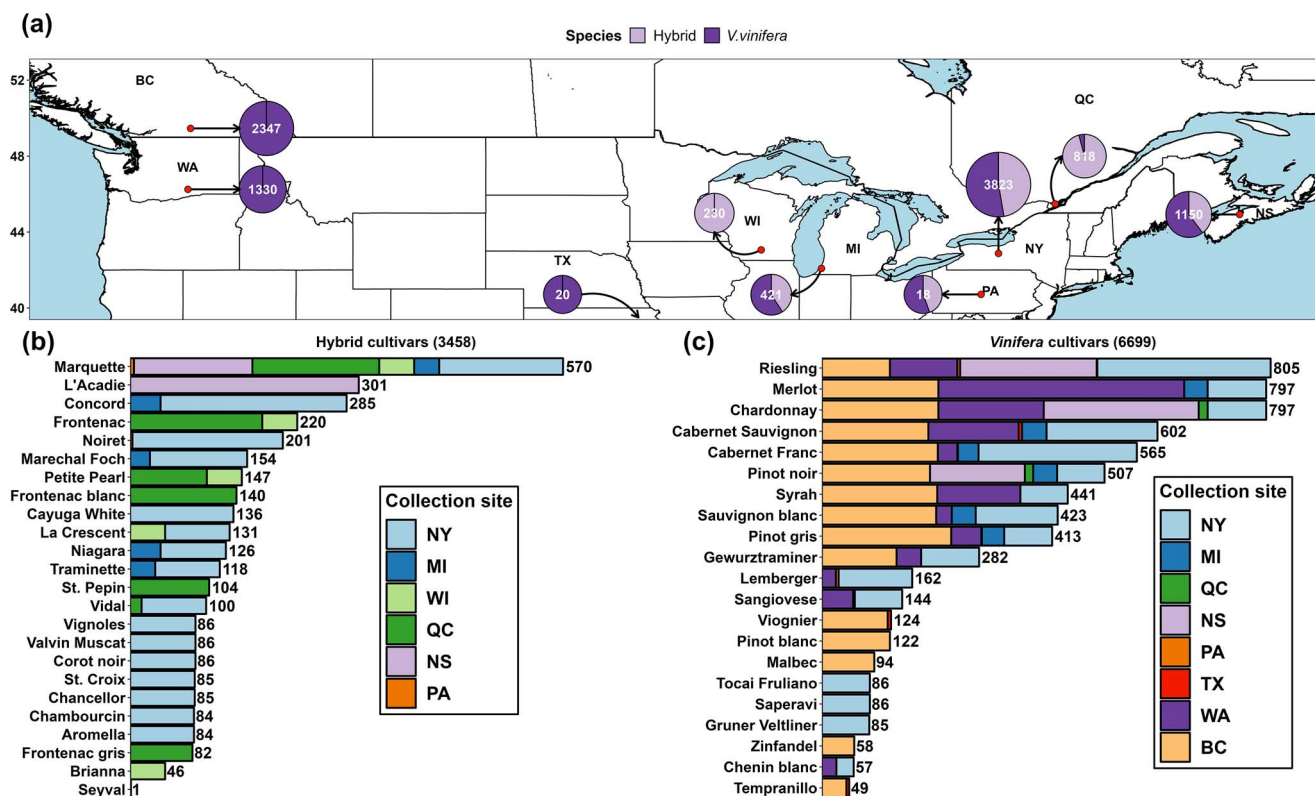


Figure 1. Grapevine freezing tolerance (LT_{50}) dataset composition. (a) Distribution of freezing tolerance data in the study. Numbers in each circle denote the number of data points and composition of each region's cultivar type denoted for hybrid and *Vitis vinifera*. (b) Data composition of hybrid cultivars. (c) Data composition of *V. vinifera* cultivars.

site-transferability. We therefore performed three types of validation: (i) cross-validation during model training, (ii) validation using an internal testing dataset ("internal"), and (iii) validation using an external testing dataset ("external"). Root-mean-square error (RMSE) was used as the evaluation metric for model performance. Training results, model performances in internal testing and external testing for the method suitability test, and the site-transferability test are shown in Fig. S1. During training, five levels of stacking and weighting were accomplished, resulting in five weighted ensemble models and 47 stacker models at different levels (Fig. S1a). Internal testing indicated that multiple models (from "WeightedEnsemble_L2" and the models thereafter) reveal the best performance with $RMSE = 1.34^{\circ}C$ to $RMSE = 1.37^{\circ}C$ (Fig. S1a). Thus, "WeightedEnsemble_L2" ($RMSE = 1.35^{\circ}C$) was selected as the alpha model for external testing for its least complexity. In the external testing, the alpha model achieved $RMSE = 1.59^{\circ}C$, $RMSE = 1.44^{\circ}C$, and $RMSE = 1.86^{\circ}C$ in the MI sub-dataset ($n = 421$), the PA sub-dataset ($n = 18$), and the TX sub-dataset ($n = 20$), respectively (Fig. S1b). The RMSE of individual cultivars with more than 10 LT_{50} values ranges from 1.14 to $2.02^{\circ}C$ and is greater in several hybrid cultivars (e.g. 'Concord' and 'Marquette') (Fig. S1b).

Final model performance and feature importance

After these tests, the complete dataset ($n = 10157$) was used to train and test the final grapevine LT_{50} prediction model. The complete dataset, along with date and site information, is provided in Data S1. Training data and internal testing data are provided in Data S2 and Data S3, respectively. Training results, model performance, and feature importance quantification using AutoGluon and SHAP value are shown in Fig. 2. During training, multiple

models (from "WeightedEnsemble_L2" and the models thereafter) reveal the best performance (Fig. 2a), and "WeightedEnsemble_L2" ($RMSE = 1.36^{\circ}C$) was selected as the final grapevine LT_{50} prediction model for its least complexity. The performance of the final model was visualized within each region (Fig. 2b). PA and TX data were omitted due to the limited amount of data categorized as internal testing data. The final model performance varies between $RMSE = 0.82^{\circ}C$ (BC sub-dataset) and $RMSE = 1.68^{\circ}C$ (NY sub-dataset) (Fig. 2b). The final model is an ensemble model at stacking level two and is composed of eight base models generated at stacking level one. The prediction of the ensemble model is a weighted outcome of the predictions in the eight base models, and the weight of each base model is determined on their performance during training. The final model prediction on the LT_{50} of all internal testing data is provided in Data S4. Specific computational requirements for running the model and the speed of the model predictions are determined automatically due to internal parallelization of the AutoGluon program. However, as a base value, the average time needed for each prediction (prediction of a single freezing tolerance value) is less than 0.001 s using 8-core i9-11900H CPU (Intel, Santa Clara, CA); thus, a prediction for a single site in a single winter season of about 200 days would require less than 0.2 s. Hyperparameter information and model requirements for running memory ratio, base model types, and base model weights have been provided as a JSON file as Supplemental Material 3.

Feature importance was quantified using internal testing data ($n = 1016$) through SHapley Additive exPlanations (SHAP) value and AutoGluon. The 15 most influential features identified in SHAP include three chilling models, two growing degree hours (GDH) models, nine exponential weighted moving average

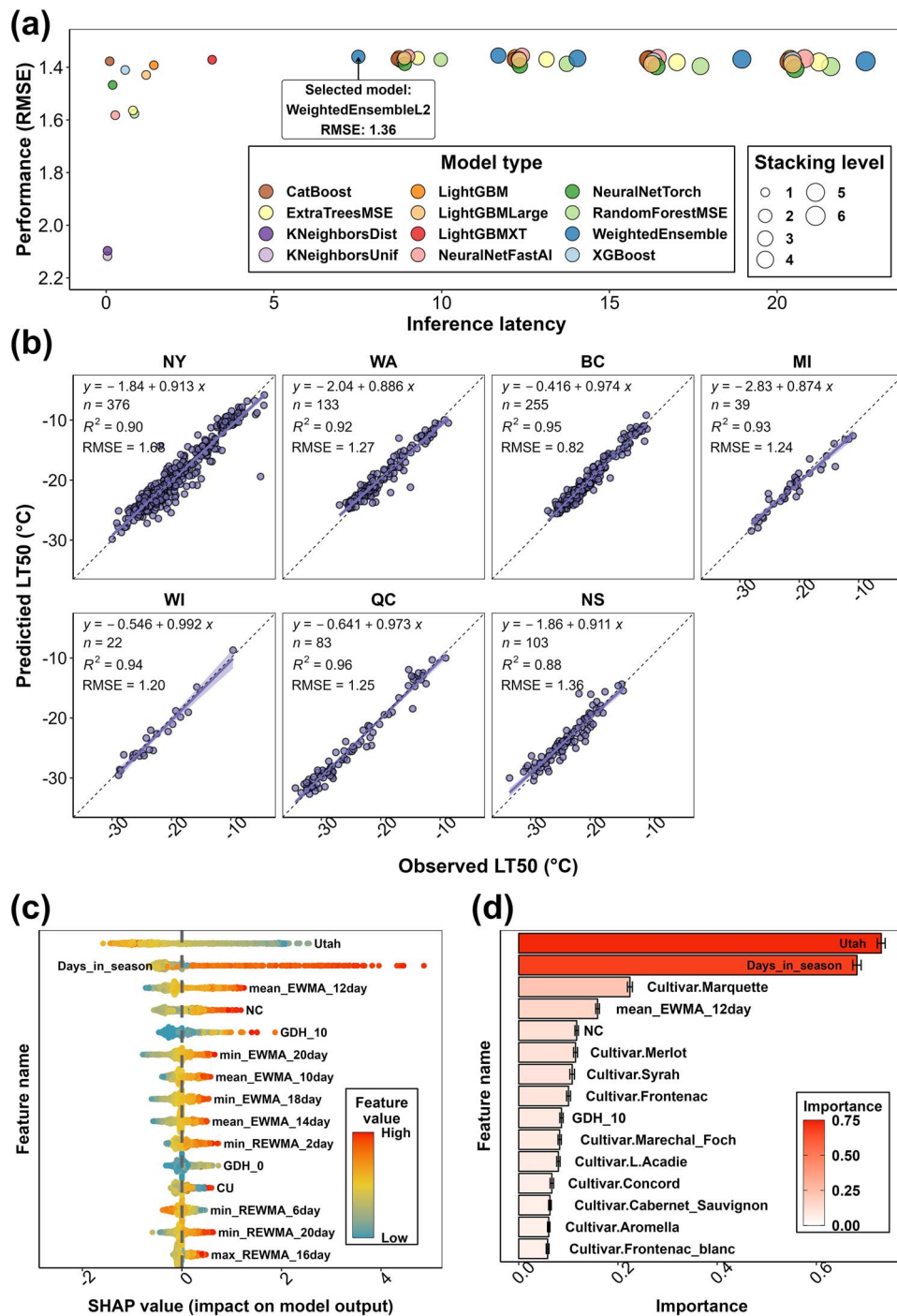


Figure 2. The final grapevine freezing tolerance model training result, performance, and feature importance. (a) The performance (RMSE) of all the models generated from training. (b) The performance of the final grapevine LT_{50} prediction model by individual sub-datasets. The data used to compute model performance are from internal testing data. (c) SHAP value distribution of the top 15 features of the final model. Feature importance is ranked based on the mean of absolute SHAP values of all the samples per feature. (d) AutoGluon feature importance of the top 15 features of the final model. The importance indicates the drop in model performance when the values of the feature are randomly shuffled across rows.

(EWMA)/reverse exponential weighted moving average (REWMA) temperature features, and days in season (Fig. 2c). The 15 most influential features identified by AutoGluon include two chilling models, one GDH model, one EWMA temperature, ten Boolean-type cultivar features, and days in season (Fig. 2d). Utah chilling hour model and days in season are the leading features that have far more impact on model performance (Fig. 2d). Randomizing the values of these two features across rows would result in 0.73 and 0.68 decrease of model performance (0.73°C and 0.68°C increase

of RMSE). Complete SHAP values of each internal testing data and each feature are provided in Data S5. A complete AutoGluon feature importance list is provided in Data S6.

Model performance comparison

The final model is thereafter referred to as “Auto-ML model” in this study. The Auto-ML model was subjected to model performance comparison with NYUS.1 and WAUS.2, and the results of this comparison are shown in Fig. 3a. The prediction error is

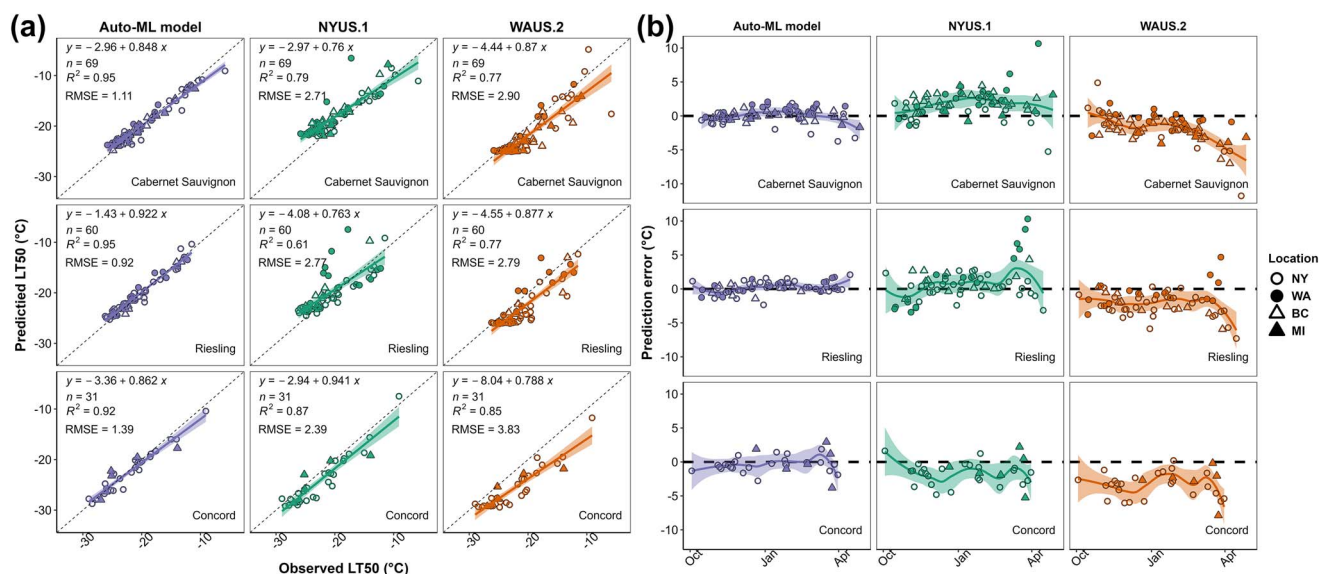


Figure 3. Grapevine freezing tolerance model comparison result and prediction error distribution. **(a)** The performance (RMSE) of Auto-ML model, NYUS.1, and WAUS.2 in ‘Cabernet Sauvignon’, ‘Riesling’, and ‘Concord’ in internal testing data. **(b)** Prediction error distribution by model and cultivar. Prediction error is calculated by $LT_{50}^{\text{predicted}} - LT_{50}^{\text{observed}}$. Negative prediction error indicates an overprediction, and positive prediction error indicates an underprediction. Different shapes and shading of datapoints indicate different regional data collection.

calculated by $LT_{50}^{\text{predicted}} - LT_{50}^{\text{observed}}$. As LT_{50} refers to subzero temperatures and thus is a negative value, negative prediction error indicates an overestimation, and positive prediction error indicates an underestimation of LT_{50} . The shade of the prediction error distribution represents the 95% confidence interval of the trendline fitted using “loess” and can serve as a rough measure of error-proneness. The temporal distribution of prediction error in the testing data is combined across all the dormant seasons and shown in Fig. 3b. The Auto-ML model prediction error is always relatively stable at around 0°C in these three cultivars, but the confidence interval tends to be wider in late season (i.e. early spring), indicating relatively lower model stability as the buds approaching budbreak (Fig. 3b). NYUS.1 usually underestimates the freezing tolerance of ‘Cabernet Sauvignon’ and ‘Riesling’ but overestimates the freezing tolerance of ‘Concord’ (Fig. 3b). Significant prediction errors were identified in late season, and these errors primarily occurred in the predictions of WA, BC, and MI data, suggesting a regional aspect that is being poorly predicted (Fig. 3b). WAUS.2 always overestimated the freezing tolerance of the three cultivars, and the overestimation became more severe in late season (Fig. 3b).

Model deployment

During the 2022–2023 dormant season, the Auto-ML model was deployed in the major cool climate viticultural regions in the USA through a new freezing tolerance website (https://grapecoldhardiness.shinyapps.io/grape_freezing_tolerance/). Predicted LT_{50} values throughout the dormancy season were made available for 16 cultivars at each weather station. Multiple sites in NY contributed on-site measurements of grapevine LT_{50} during the season through the data upload portal. Among these sites, Portland, NY (station ID: “USC00306747”) and Geneva, NY (station ID: “USC00303184”) contributed whole season LT_{50} measurements. The predictions from Auto-ML model, NYUS.1, and WAUS.2 were compared to these measurements (Fig. 4). We note that these measurements were not used at all in the model development and were completely unseen by the model.

During the season, the predicted LT_{50} values from the Auto-ML model prediction were typically intermediate to those predicted by WAUS.2 and NYUS.1 models and were more responsive to the change of minimum temperature (Fig. 4a). The Auto-ML model outperformed the previous models for all three cultivars at both sites (Fig. 4a). Prediction error distribution for the Auto-ML model across the season indicates that the model performed well in early and mid-season (Fig. 4b). In late season, consistent overestimation was observed for all three cultivars and at both sites (Fig. 4b). The comparison of the predictions and the on-site measurements for the other cultivars are shown in Fig. 4c. Overall, the performance of the Auto-ML model ranged between $RMSE = 1.19^{\circ}\text{C}$ and $RMSE = 2.20^{\circ}\text{C}$, except for ‘Marquette’ whose $RMSE = 3.20^{\circ}\text{C}$ (Fig. 4c). Since the observed LT_{50} of ‘Marquette’ from mid to late season are much higher than the observed LT_{50} of other hybrid cultivars, which is abnormal, we assume the underperformance of the Auto-ML model on ‘Marquette’ might be from repeated measurement errors or sample collection errors. Nevertheless, late season overestimation was also observed in most other cultivars, though it was more severe in some hybrid cultivars and *V. vinifera* cultivars with faster deacclimation response (e.g. ‘Gewurztraminer’ and ‘Lemberger’) (Fig. 4c).

Freezing damage potential was estimated based on the temperature gap between predicted LT_{50} and daily minimum temperature. The freezing damage potential over the dormant season is estimated for the Northeastern and Midwestern United States and updated daily using gridded weather data. Maximum freezing damage potential and maximum freezing tolerance (minimum predicted LT_{50}) in the 2022–23 dormant season of 16 cultivars are shown in Fig. S2 and Fig. S3, respectively.

Discussion

In this study, we used AutoGluon, an Auto-ML platform, to assist in the development of a multi-site grapevine freezing tolerance model. Results from method suitability and site-transferability tests indicated the feasibility of the feature extraction and modeling method. The resulting Auto-ML model outperformed two previous models in three test cultivars and accurately predicted

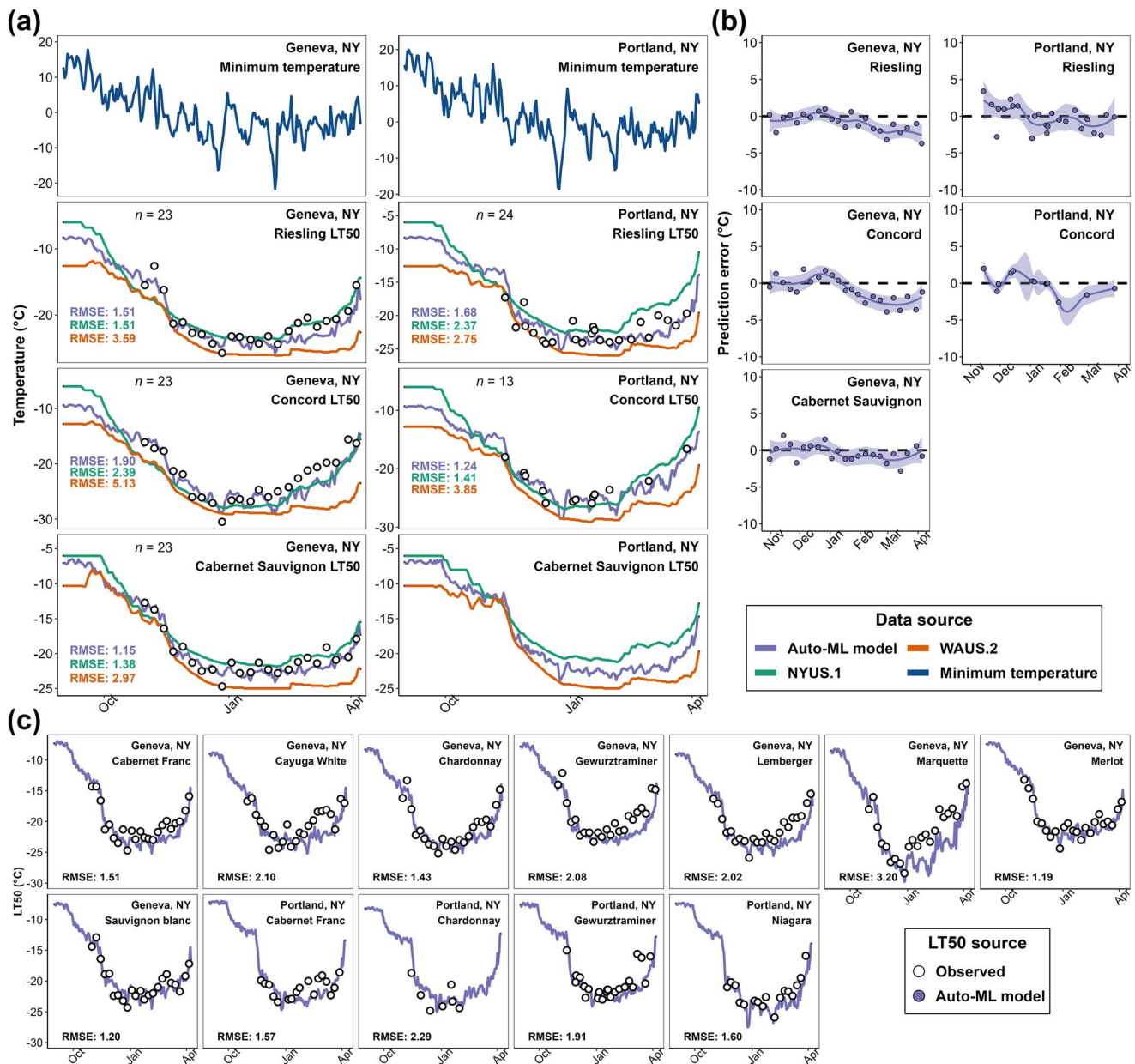


Figure 4. Model deployment results in Geneva, NY and Portland, NY in 2022–23 dormant season. (a) Prediction comparison of Auto-ML model, NYUS.1, and WAUS.2 and on-site measurements in 'Cabernet Sauvignon', 'Riesling', and 'Concord'. (b) Prediction error distribution of Auto-ML model by cultivar. (c) Predictions from Auto-ML model and on-site measurements in other cultivars. Open black circles indicate the LT₅₀ measurements.

the freezing tolerance in the 2022–23 dormant season. The Auto-ML model was deployed in the major cool climate viticultural regions in the USA at new grapevine freezing tolerance website, enabling a large-scale real-time simulation of grapevine freezing tolerance in North America for the first time. The development of such a predictive system is crucial to estimating how cultivated plants can and will adapt to changing climatic conditions. In addition, freezing tolerance prediction will be advantageous for farmers to determine management strategies and select suitable cultivars under existing and changing climatic conditions [41, 42]. It is also advantageous for the rest of the supply chain and the consumers since appropriate and timely mitigation strategies may reduce fluctuations in produce and product supplies.

Diverse modeling approaches, such as empirical modeling and mechanistic modeling, have been intensively used in the history of plant biology. ML-empowered or DL-empowered data-driven

modeling has been gaining favor in recent years for its efficiency and high accuracy, especially in "omics" data that exhibit high complexity [43]. However, modeling of complex physiological changes in plants using local or regional data often results in models that tend to overfit. In addition, mechanistic and DL-empowered models can be complicated, reducing the user-friendliness and ultimately adoption [44]. A detailed discussion of the reasoning for these limitations is included in Note S1. As a novel approach designed to address both issues (tendency of overfitting and easiness of modeling), we collected training data from multiple regions and used Auto-ML to model grapevine LT₅₀. The accumulation of grapevine LT₅₀ data across North America presented in this study was likely to generate a generalizable model for different climate conditions, and Auto-ML would in turn ease the modeling. As the aim is to generate a site-transferable grapevine LT₅₀ prediction model, the modeling

engine (AutoGluon) and the features were first evaluated for their suitability to model grapevine LT_{50} and their capacity to generate a stable prediction model with high site-transferability. During the method suitability and site-transferability tests, the alpha model generated using partial data resulted in a similar performance in internal testing ($RMSE = 1.35^{\circ}C$) and external testing ($RMSE = 1.59^{\circ}C$) (Fig. S1). These results demonstrated an improvement over predictions produced by the NYUS.1, WAUS.2, and the RNN model during the testing on unseen data [14, 34], indicating that the alpha model passed the method suitability test and the site-transferability test. As the internal and external testing data are from different sites, this result also suggests that the AutoGluon method did not overfit the training data and exhibited high site-transferability.

Among the different potential models trained on the entire dataset, “WeightedEnsemble_L2” exhibited high accuracy and low complexity (thus low inference latency) and was chosen as the final model to deploy. Internal testing showed that the performance of the final model (Auto-ML model) varied in different regions (Fig. 2b); however, the regions that show greater RMSE have more complex data composition (e.g. a mix of hybrid cultivars and *V. vinifera* cultivars in NY dataset), indicating potential model underperformance in some cultivars. The Auto-ML model, in general, outperformed previous models (Fig. 3a). Prediction error distribution showed that NYUS.1 and WAUS.2 lost prediction power when tested in disparate regions. For example, WAUS.2 performs well in WA sub-dataset; however, the predictions in other sub-datasets exhibit frequent overestimation (Fig. 3b). Similarly, NYUS.1 accurately predicts the LT_{50} in NY sub-dataset, yet the prediction of WA and BC sub-datasets is less accurate (Fig. 3b). The underperformance of these two models is more severe in late season as the buds are deacclimating and approaching budbreak (Fig. 3b). To compare, the Auto-ML model leverages the multi-region training dataset and generalizes the predicted response across different regions, which resulted in a more evenly distributed prediction error in different regions (Fig. 3b). This result again indicates that the final model performs as a generalizable model that exhibited high site-transferability.

However, the Auto-ML model also has several limitations. First, the number of parameters in the Auto-ML model is considerably more than the mechanistic models, which may explain its better performance as the same process was formulated with more factors. However, utilizing more parameters may also impair the efficiency of deployment as more computational resources are needed, limiting the potential adoption of this model to the users with sufficient computational support. Unlike the mechanistic models, in which each prediction is explainable by the parameters, the Auto-ML method utilizes features in a complicated way through stacking and weighting of multiple base models (the black box effect), which impairs the ability to attribute model changes to biological or physiological factors. The black box effect of these models brings limited insight toward understanding grapevine dormant season biology [45]. Nevertheless, the impact of each feature can be determined using AutoGluon and SHAP analysis, thus partially unboxing the black box of the model. Although the result differs due to different computational approaches, five features (the Utah chilling accumulation, days in season, EWMA of mean temperature with window size of 12, the NC chilling accumulation, and the GDH with $10^{\circ}C$ as base temperature) were identified for their importance by both algorithms (Fig. 2c and d). These features have much higher statistical importance for the model than the other features; however, the model prediction is not fully explainable using these features only.

In fact, the Auto-ML model prediction incorporates complicated interactions of feature groups, and each group contributes to the final prediction in different ways. Detailed analysis and reasoning are included in Note S2 and Fig. S4.

Among these five features, Utah and NC are two chilling models that calculate the accumulation of chilling in the season. Chilling accumulation has been determined for its importance in the physiology of grapevine in the dormant season [4, 13, 46, 47]. Different chilling models also help quantify the dynamics of cold acclimation and deacclimation in the NYUS.1 and WAUS.2 models [11, 14], and therefore errors from chilling models can hinder predictions. The Auto-ML model utilized individual chilling models in a different manner than the previous prediction models (Fig. 2c and Fig. S5a). The SHAP value of the Utah chilling model exhibits an “S” shape as Utah chilling increases, while NC and CU chilling models exhibit “U” shapes (Fig. S5a). A potential explanation for the importance of both Utah and NC chilling models in Auto-ML is that the Auto-ML model combined the effect of the chilling models and generated a unique algorithm to dynamically describe the effect of chilling accumulation. As the physiological mechanism driving chilling accumulation is unknown, we are reliant on historical patterns of chilling to explain changes in dormancy status. Until this mechanism is clarified, our current chilling models are at best estimations of the process. Although the black box effect of the Auto-ML model impairs its explainability, our analysis suggests that neither model alone is a better fit to the data than when combining aspects of both models across the season (Note S3 and Fig. S5).

Another issue of the Auto-ML model is the inconsistency of its performance for different cultivars and different times of the season. Site-transferability tests and model comparison results suggest that the Auto-ML model underperforms for some hybrid cultivars (Fig. S1 and Fig. 3). This underperformance compounds with a prediction that is more error prone in late season (Fig. 4), which was also found in the predictions in NYUS.1 and WAUS.2. Grapevines enter ecodormancy during late season and rely on cold temperatures to prevent growth and development. During this phase, the buds progressively increase in temperature sensitivity, specifically responding to warm temperatures, which can be physiologically reflected in more variable freezing tolerance when measured with DTA [13, 48]. Modeling freezing tolerance during this period likely requires a more detailed and higher temporal resolution (e.g. hourly temperature), while all the features used for the prediction models are based on daily temperature [11, 14]. Another potential aspect impacting the accuracy of predictions during this phase is the impact of solar radiation. During periods of sun exposure, the dark bud tissue can absorb infrared radiation, warming to levels higher than ambient air temperatures [49]. As a result, predictions based on daily or even hourly temperature likely underestimate bud temperatures. Currently, there is not a comprehensive measure or validated conversion for solar radiation (langleys) to bud temperature, and many weather station datasets lack tracking of solar radiation. Future studies documenting the relationship between langleys and bud perceived temperatures may improve predictions during ecodormancy. Finally, as fast deacclimating grapevine buds exhibit less freezing tolerance, the LTE of these buds are sometimes less detectable, which may result in less accurate measurements of freezing tolerance with DTA [7]. Due to the nature of Auto-ML, the exact mechanism resulting in late season lower predictability cannot be delineated in this study.

To our knowledge, this study represents the first attempt at modeling weather-related perennial crop physiology using only an

Auto-ML modeling engine. The current model is named NYUS.2 based on the previously proposed model naming system [14]. The method was validated, the final model revealed high site-transferability, and outperformed two previous mechanistic models in all testing regions. Feature importance analysis indicates that the model may have adequately extracted some biological mechanisms during training. For example, the strong effect of chilling models in the Auto-ML model indicate the role of chilling fulfillment of grapevine, and the effect one-hot encoded cultivar features indicate the proper differentiation of cultivar-specific acclimation and deacclimation dynamics [13, 15, 47, 48]. Deployment of NYUS.2 in the 2022–23 dormant season provided daily updated model estimates of grapevine freezing tolerance of 16 cultivars at 2035 independent locations in 23 states in the USA using the weather station database ACIS (<https://www.rcc-acis.org/>). NYUS.2 and the Shiny-based web application represent the launch of the first large-scale real-time grapevine freezing simulation system with high temporal resolution in North America. In ongoing work, we plan to further expand the deployment of the NYUS.2 model to more cool climate viticultural regions in North America, such as Quebec, Ontario, and British Columbia in Canada. Furthermore, then incorporation of more worldwide on-site measurements of freezing tolerance and more climate descriptors would enable further training of the model, likely yielding a more generalizable model to help understand the biology of grapevine freezing tolerance and quantify the threat of freezing under climate change. In addition, recent research has demonstrated that dormancy progression and cold acclimation and deacclimation seem to be conserved among woody perennials [50], thus the biological drivers of winter hardiness (although not fully elucidated), captured by the NYUS.2 model, may also be transferable. The use of the NYUS.2 model, along with the feature extraction methods used here, may prove useful for generating a universal freezing tolerance model for other perennial plant species. Currently, the primary limitation of applying these methods to other plant species is the need for large datasets collected across a variety of years and regions.

The current NYUS.2 model could be deployed on other available weather databases to generate specific grapevine freezing tolerance monitoring systems for specific needs. Moreover, the model could be used to trace the grapevine freezing tolerance in the past, thus facilitating the quantification of climate change's impact on grapevine freezing tolerance/damage, which could help estimate site suitability for grape cultivars and wine industries in the future. In addition, supercooling species where freezing tolerance data are collected, such as cherries [51] and peaches [52], would be good candidates for extension of this model framework if sufficient on-site measurement of freezing tolerance is available for model re-training. It should be noted that the model framework and approach is not limited to the prediction of freezing tolerance and could be extended to any plant phenotypes where sufficient data has been collected. The detailed scripts for the above steps and model deployment along with a ready-to-use NYUS.2 model are available in GitHub at <https://github.com/imbattery11/NYUS.2>.

Materials and methods

Grapevine freezing tolerance data and weather data

On-site measurements of grapevine bud freezing tolerance from multiple viticultural regions in North America were used for model training and evaluation. In all data collection regions, the freezing tolerance of a cultivar on a single date was determined

using multiple buds through standard DTA with slight variations between collaboration groups [6, 7]. Briefly, grapevine buds were excised from canes collected from the field at specified time throughout the dormancy season and subjected to a decrease of temperature from 0°C to a lethal temperature (−40 to −50°C) in programmable freezers, and thermoelectric modules detect LTE that corresponds to the freezing tolerance of the buds. Bud LT₅₀ was determined by either taking the mean or median of all LTEs, or through probit analysis. Slight variations in protocol do not significantly affect measurements in grapevine [7]. Hourly or daily temperature data were obtained from various sources (Table S1) to represent local weather conditions at the sites of field data sampling. At the sites where hourly temperature was unavailable, hourly temperature was estimated based on daily maximum and minimum temperature using “stack_hourly_temps” of the R package “chillR” [53]. A detailed glossary of technical terms used in this paper is detailed in Table S2. The NY dataset was obtained from Cornell University Cornell AgriTech, Geneva, NY, USA (42.88° N, 77.03° W) and composed of 3823 LT₅₀ data. Among these data, 1807 (47%) were collected from hybrid cultivars, and 2016 (53%) were collected from *V. vinifera* cultivars. MI dataset was obtained from Michigan State University Southwest Michigan Research and Extension Center, Benton Harbor, MI, USA (42.08° N, 86.36° W) and composed of 435 LT₅₀ data. Among these data, 183 (42%) were collected from hybrid cultivars, and 252 (58%) were collected from *V. vinifera* cultivars. The WI dataset was obtained from University of Wisconsin Madison, West Madison Agricultural Research Station, Verona, Wisconsin, USA (43.06° N, 89.53° W) and composed of 230 LT₅₀ data collected from hybrid cultivars. WA dataset was obtained from Washington State University, Irrigated Agriculture Research and Extension Center, Prosser, WA, USA (46.25° N, 119.74° W) and composed of 1330 data collected from *V. vinifera* cultivars. PA and TX datasets are two minor datasets contributed by The Pennsylvania State University and Texas A&M University AgriLife Extension Service, respectively. BC dataset was obtained from Agriculture and Agri-Food Canada (AAFC), Summerland Research and Development Center, Summerland, BC, Canada (49.46° N, 119.60° W) and composed of 2347 LT₅₀ data collected between 2012 and 2022 from 15 *V. vinifera* cultivars. QC dataset was obtained from Centre de Recherche Agroalimentaire de Mirabel, Mirabel, QC, Canada (45.49° N, 74.05° W) and composed of 818 LT₅₀ data. Among these data, 784 (95%) were collected from hybrid cultivars, and 34 (5%) were collected from *V. vinifera* cultivars. NS dataset was obtained from AAFC, Kentville Research and Development Centre, Kentville, NS, Canada (44.93° N, 64.32° W) and composed of 1150 LT₅₀ data. Among these data, 457 (40%) were collected from hybrid cultivars, and 693 (60%) were collected from *V. vinifera* cultivars.

Feature generation and extraction

The full model dataset contains 117 features. Cultivar features were generated by transforming single-column categorical cultivar names to multiple-column binary cultivar names through one-hot encoding [18]. This approach addresses the difficulties of managing mixed-type variables in machine learning and facilitates the application of model interpretation techniques [19]. Forty-five Boolean-type features were obtained from this transformation. The remaining features are hourly temperature-based features composed of daily temperature descriptors, cumulative temperature descriptors, EWMA temperatures, and REWMA temperatures. Daily temperature descriptors and cumulative temperature descriptors are detailed in Note S4. EWMA and REWMA temperatures are unique temperature descriptors computed based

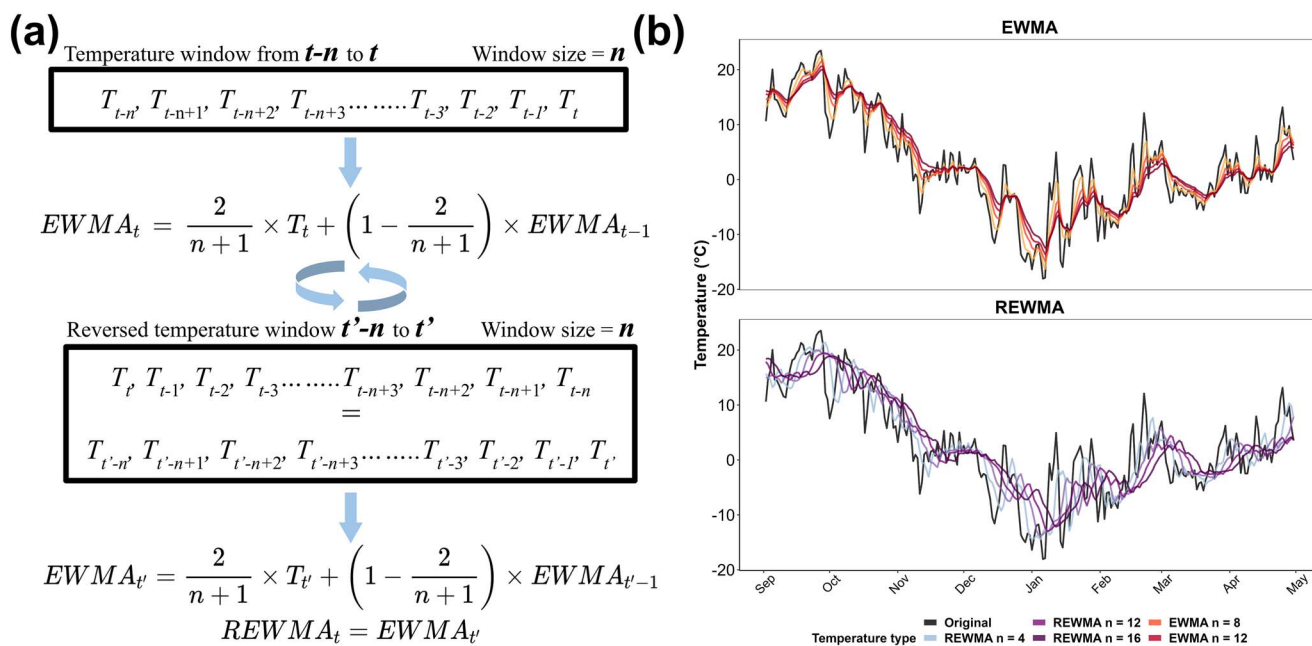


Figure 5. Computation of exponential weighted moving average (EWMA) and reverse exponential weighted moving average (REWMA) temperatures and their impacts on daily temperatures. **(a)** Schematic representation of EWMA and REWMA temperatures in a window from $t-n$ to t . T is the temperature of a day. To compute the EWMA of a date t with the window size of n , weight is exponentially ascendingly added from $t-n$ to t . To compute the REWMA of a date t with the window size of n , the window is first reversed to $t'-n$ to t' (where $T_{t'} = T_{t-n}$ and $T_{t'-1} = T_{t-n+1}$, etc.), and weight is exponentially descendingly added from $t-n$ to t . **(b)** Effect of EWMA and REWMA with four window sizes (4, 8, 12, and 16 days) on mean and daily temperatures in a dormant season. Original temperatures represent the recorded temperature in the field during the season.

on temperature windows [54]. The mathematical expression of EWMA and REWMA, along with their impacts on daily mean, maximum, and minimum temperatures, are described in Fig. 5. EWMA is a smoothing method used for normalization during data preprocessing and to decrease variation in time series data in ML [19]. In this study, the EWMA temperature of a date t is computed based on a temperature window of a size n (in days) with exponentially ascending weight added to each date from $t-n$ to t (Fig. 5a). This method allows all the temperatures in the window to contribute to the final EWMA temperature, while the closer the date is to date t , the more impact the temperature of the date has on the final EWMA temperature. As the window size n increases, the EWMA temperatures follow the same trend as the original temperatures, but they are increasingly smoothed as more data are included in the weighted average (Fig. 5b). Including EWMA temperatures as features allows Auto-ML to factor temperature as a cumulative descriptor from a temperature window rather than only relying on daily temperatures. This approach may result in higher accuracy since the change of freezing tolerance is more likely a dynamic consequence of continuous exposure to a temperature window rather than a transient response to current temperature [55].

REWMA temperature is a new expression of temperature to incorporate the theory of “cold priming” and the “cold shock effect” in the modeling of grapevine freezing tolerance. The biological reasoning for including REWMA temperatures is available in Note S5. The major difference between REWMA and EWMA is that the assignment of weight is reversed in REWMA, thus allowing earlier temperature to have more impact. To compute REWMA, the temperature window (size = n) from date $t-n$ to date t (T_{t-n} to T_t , where T is the temperature of a day) is reversed. EWMA is applied on the reversed window ($T_{t'-n}$ to $T_{t'}$, where $T_{t'} = T_{t-n}$ and $T_{t'-1} = T_{t-n+1}$, etc.) (Fig. 5a). In this method, weight exponentially descends from $T_{t'}$ to $T_{t'-n}$ (T_{t-n} to T_t), whereas all the tempera-

tures in the window contribute to the final REWMA $_t$. Similar to EWMA, as window size n increases, the REWMA temperatures are smoothed as more data are used for computation. However, in REWMA, all the temperatures also better correlate with earlier dates (visually horizontally shifted) as compared to the original temperatures (Fig. 5b).

We extracted EWMA and REWMA temperatures of daily maximum, minimum, and mean temperatures with window sizes in 2-day increments (2:20) for a total of 60 features. A correlation matrix of the EWMA and REWMA of daily minimum temperature is shown in Fig. S6. As the window size increases, the correlations between all these temperature features decrease (Fig. S6). The correlations between REWMA temperatures are generally weaker than those between EWMA temperatures, and the correlations between EWMA temperatures and REWMA temperatures are even weaker (Fig. S6).

Method suitability test, site-transferability test, model training, and model selection

We performed three types of validation to test the method and the model: (i) cross-validation during model training, (ii) validation using an internal testing dataset (“internal”), and (iii) validation using an external testing dataset (“external”). Specifically, the LT₅₀ data from three minor sub-datasets in MI, PA and TX ($n = 459$) were isolated and used as external testing data. The remaining data ($n = 9698$) were used to train and test an alpha grapevine LT₅₀ prediction model for the method suitability test and the site-transferability test. The partial dataset ($n = 9698$) was randomized before training to avoid structural bias and split into 90% and 10% subsets to be used for training ($n = 8728$) and internal testing ($n = 970$), respectively. Only the training data were used for model training and 10-fold cross-validation. In this study, the difference between internal testing data and external testing data

is that the internal testing data were from the same collection regions as the training data (thus named “internal”), and the external testing data were from the different collection regions (thus named “external”). The performance of the alpha model on the internal testing data informs if Auto-ML can generate a model that can accurately predict the LT_{50} of the data that were not used for training ($n = 970$), which indicates “method suitability”. In comparison, the performance of the alpha model on the external testing data informs whether the resulting model can accurately predict the LT_{50} of the data that were neither used for training nor originated from the same regions ($n = 459$), which indicates “site-transferability”.

The Python package AutoGluon (version 0.7.0) was used to conduct model training. Training was conducted through “TabularPredictor” with “fit(presets=‘best_quality’, num_bag_folds = 10, num_stack_levels = 5)”. All steps of data preprocessing were automatically managed by AutoGluon. All the models produced from training were evaluated with RMSE using internal testing data. As the training command specifies stacking and weighting at five levels (num_stack_levels=5), it is possible that model performance maximizes after certain stacking levels. In this case, the model with the least complexity (lowest level of stacking) was selected as the resulting model per training to enable highest prediction speed. After the method suitability test and the site-transferability test, the entire dataset was aggregated to include LT_{50} from every region ($n = 10157$) to generate a final grapevine LT_{50} prediction model. The training of the final model was the same as stated above, but the dataset was randomized again to be split into 90% training data and cross-validation data ($n = 9141$) and 10% internal testing data ($n = 1016$).

Feature importance quantification of the final model

The importance of each feature in the final model was quantified using internal testing data through two methods. First, AutoGluon internal feature importance quantification was conducted using “feature_importance” with “(num_shuffle_sets=100, subsample_size=1000)”. In this method, when testing data are perturbed by randomly shuffling the values of a feature across rows, the feature’s importance score reflects the resulting decrease in the model’s performance, a derivation of local interpretable model-agnostic explanations (LIME) [56]. Second, feature importance was quantified by computing the SHAP value of the feature using the Python package SHAP (version 0.41.0) [57]. The SHAP value of a feature represents the average marginal contribution of that feature across all possible feature combinations. The calculation of SHAP values involves computing the contributions of each feature subset, considering all possible permutations of features [58].

Model comparison

To further test model performance, the final model was compared to two previous grapevine LT_{50} prediction models, WAUS.2 and NYUS.1. These models were generated using Washington LT_{50} data and New York LT_{50} data, respectively [11, 14]. Since the current version of NYUS.1 is limited to the LT_{50} of cultivars ‘Cabernet Sauvignon’, ‘Riesling’, and ‘Concord’ [14], model performance comparison was conducted on a subset of internal testing data ($n = 1016$) of ‘Cabernet Sauvignon’ ($n = 69$), ‘Riesling’ ($n = 60$), and ‘Concord’ ($n = 31$) from NY, WA, BC, and MI sub-datasets. ‘Cabernet Sauvignon’ is a red-fruited *V. vinifera* grapevine cultivar whose maximum freezing tolerance typically does not exceed -24°C . ‘Riesling’ is a white-fruited *V. vinifera* grapevine cultivar whose maximum freezing tolerance typically does not exceed

-27°C . ‘Cabernet Sauvignon’ and ‘Riesling’ tend to acclimate and deacclimate slower compared to hybrid cultivars. ‘Concord’ is a hybrid cultivar derived from *Vitis labrusca* that typically develops greater freezing tolerance (e.g. -32°C) but also tends to acclimate and deacclimate faster than *V. vinifera* cultivars [1, 4, 11, 14, 47].

Model deployment

In the 2022–23 dormant season, the final model was applied on weather station temperature data and interpolated gridded temperature data in Applied Climate Information System (ACIS, <https://www.rcc-acis.org/>) to facilitate real-time grapevine freezing tolerance monitoring in cool climate viticultural regions in the USA. The model prediction was updated daily and interactively deployed using the R Shiny package [59]. The current monitoring system covers the 2035 weather stations in Global Historical Climatology Network (in ACIS) within the bounding box of Minnesota (NW), Maine (NE), Virginia (SE), and Missouri (SW). Gridded prediction is also available for Northeastern and Midwestern United States with a $5\text{ km} \times 5\text{ km}$ spatial resolution based on interpolated gridded temperature data from ACIS. Along with LT_{50} prediction, the potential that freezing damage has occurred (0 to 100) is estimated through a symmetric sigmoid function.

The sigmoid function is defined as

$$P = \frac{100}{1 + e^{-(\ln(1-P_h) - \ln(P_h)) \times g^{-1} \times (T - LT_{50\text{pred}})}} \quad (1)$$

where P is the potential of freezing damage for a day, T is the daily minimum temperature and $LT_{50\text{pred}}$ is the predicted LT_{50} from the model. P_h and g are two constants that represent a potential of freezing damage that is greater than 50% and the gap of T and $LT_{50\text{pred}}$ to reach that potential [60]. In our estimation, we used $P_h = 0.9$ and $g = 2$, assuming that 10% and 90% of the potential would have occurred when the ambient temperature is 2°C above and below the predicted LT_{50} , respectively.

Acknowledgments

The authors would like to thank Lynn Mills (WA), Beth Ann Workmaster (WI), Katherine Benedict (NS), Alexander Campbell and Jessee Tinsley (QC), Don Smith and Meredith Persico (PA), Kim Knappenberger (Portland, NY), and Hanna Martens, Felix Pike, and Bill Wilsey (Geneva, NY) for their help in LT_{50} data collection. This work was partially supported by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison with funding from the Wisconsin Alumni Research Foundation, and through the USDA ARS appropriated project 1910-21220-006-00D and New York Wine and Grape Foundation. The NY LT_{50} data collection was also supported by Cornell University Federal Capacity Funds Grant Program. The WA LT_{50} data collection was supported by the Washington Wine Industry Foundation. The PA LT_{50} data collection was supported by the USDA National Institute of Food and Agriculture (NIFA) Federal Appropriation under Projects PEN0 4794 (7003432). The NS LT_{50} data collection was supported by a Canadian Agricultural Partnership (CAP) project (ASC-12 Wine Grape Cluster Activity 7), the Canadian Grapevine Certification Network (CGCN), and the Grape Growers’ Association of Nova Scotia (GGANS). The QC LT_{50} data collection was supported by AgriScience program-cluster on behalf of Agriculture and Agri-Food Canada.

Author contributions

H.W. and J.P.L. assembled the dataset, conducted the modeling, did the model analysis, and designed the website. G.D.M. provided technical and biological guidance of machine learning and feature importance quantification. A.P.K. provided the code for WAUS.2 and NYUS.1. A.P.K., M.K., T.E.M., A.H.W., J.L.F., A.H.H., C.P., M.R., A.A., M.G.N., J.P.R., P.W., M.C., and J.P.L. collected LT₅₀ data and contributed the sub-datasets from different regions. H.W., J.P.L., and G.D.M. wrote most of the manuscript with contributions from all co-authors.

Data availability

The data underlying this article are available in GitHub at <https://github.com/imbattery11/NYUS.2>.

Conflict of interest statement

None declared.

Supplementary data

Supplementary data are available at *Horticulture Research* online.

References

- Zabadal TJ, Dami IE, Goffinet MC. *et al.* Winter injury to grapevines and methods of protection. Michigan State University Extension. 2007;
- Poling EB. Spring cold injury to winegrapes and protection strategies and methods. *HortScience*. 2008;**43**:1652–62
- Dami IE, Li S, Zhang Y. Evaluation of primary bud freezing tolerance of twenty-three winegrape cultivars new to the eastern United States. *Am J Enol Vitic*. 2016;**67**:139–45
- Londo JP, Kovaleski AP. Characterization of wild North American grapevine cold hardiness using differential thermal analysis. *Am J Enol Vitic*. 2017;**68**:203–12
- Pierquet P, Stushnoff C. Relationship of low temperature exotherms to cold injury in *Vitis riparia* Michx. *Am J Enol Vitic*. 1980;**31**:1–6
- Mills LJ, Ferguson JC, Keller M. Cold-hardiness evaluation of grapevine buds and cane tissues. *Am J Enol Vitic*. 2006;**57**:194–200
- Londo JP, Moyer MM, Mireles M. *et al.* Evaluation of sample preparation practices common with differential thermal analysis of grapevine bud cold hardiness. *Am J Enol Vitic*. 2023;**74**:0740002
- Wample RL, Reisenauer G, Bary A. *et al.* Microcomputer-controlled freezing, data acquisition and analysis system for cold hardiness evaluation. *HortScience*. 1990;**25**:973–6
- Cohen J, Agel L, Barlow M. *et al.* Linking Arctic variability and change with extreme winter weather in the United States. *Science*. 2021;**373**:1116–21
- Ferguson JC, Tarara JM, Mills LJ. *et al.* Dynamic thermal time model of cold hardiness for dormant grapevine buds. *Ann Bot*. 2011;**107**:389–96
- Ferguson JC, Moyer MM, Mills LJ. *et al.* Modeling dormant bud cold hardiness and budbreak in twenty-three *Vitis* genotypes reveals variation by region of origin. *Am J Enol Vitic*. 2014;**65**:59–71
- North M, Workmaster BA, Atucha A. Cold hardiness of cold climate interspecific hybrid grapevines grown in a cold climate region. *Am J Enol Vitic*. 2021;**72**:318–27
- Kovaleski AP, Reisch BI, Londo JP. Deacclimation kinetics as a quantitative phenotype for delineating the dormancy transition and thermal efficiency for budbreak in *Vitis* species. *AoB PLANTS*. 2018;**10**:ply066
- Kovaleski AP, North MG, Martinson TE. *et al.* Development of a new cold hardiness prediction model for grapevine using phased integration of acclimation and deacclimation responses. *Agric For Meteorol*. 2023;**331**:109324
- North M, Workmaster BA, Atucha A. Effects of chill unit accumulation and temperature on woody plant deacclimation kinetics. *Physiol Plant*. 2022;**174**:e13717
- Rubio S, Pérez FJ. Testing the Ferguson model for the cold-hardiness of dormant grapevine buds in a temperate and subtropical valley of Chile. *Int J Biometeorol*. 2020;**64**:1401–8
- Zhou Z-H. *Machine Learning*. 1st ed. Singapore: Springer Singapore; 2021:
- Ashenden SK, Bartosik A, Agapow P-M *et al.* Chapter 2 - Introduction to artificial intelligence and machine learning. In: Ashenden SK (eds.), *The Era of Artificial Intelligence, Machine Learning, and Data Science in the Pharmaceutical Industry*. Academic Press; 2021; 15–26
- Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. *Electron Mark*. 2021;**31**:685–95
- Khaki S, Wang L. Crop yield prediction using deep neural networks. *Front Plant Sci*. 2019;**10**:621
- Khaki S, Wang L, Archontoulis SV. A CNN-RNN framework for crop yield prediction. *Front Plant Sci*. 2020;**10**:1750
- Shahhosseini M, Hu G, Archontoulis SV. Forecasting corn yield with machine learning ensembles. *Front Plant Sci*. 2020;**11**:1120
- Gall GEC, Pereira TD, Jordan A. *et al.* Fast estimation of plant growth dynamics using deep neural networks. *Plant Methods*. 2022;**18**:21
- Ma C, Liu M, Ding F. *et al.* Wheat growth monitoring and yield estimation based on remote sensing data assimilation into the SAFY crop growth model. *Sci Rep*. 2022;**12**:5473
- Franczyk B, Hernes M, Koziarkiewicz A. *et al.* Deep learning for grape variety recognition. *Procedia Comput Sci*. 2020;**176**:1211–20
- Arab ST, Noguchi R, Matsushita S. *et al.* Prediction of grape yields from time-series vegetation indices using satellite remote sensing and a machine-learning approach. *Remote Sens Appl Soc Environ*. 2021;**22**:100485
- Liu E, Gold KM, Combs D. *et al.* Deep semantic segmentation for the quantification of grape foliar diseases in the vineyard. *Front Plant Sci*. 2022;**13**:978761
- Qiu T, Underhill A, Sapkota S. *et al.* High throughput saliency-based quantification of grape powdery mildew at the microscopic level for disease resistance breeding. *Hortic Res*. 2022;**9**:uhac187
- Palacios F, Melo-Pinto P, Diago MP. *et al.* Deep learning and computer vision for assessing the number of actual berries in commercial vineyards. *Biosyst Eng*. 2022;**218**:175–88
- Mohimont L, Alin F, Rondeau M. *et al.* Computer vision and deep learning for precision viticulture. *Agronomy*. 2022;**12**:2463
- Romero Galvan F, Pavlick R, Trolley GR. *et al.* Scalable early detection of grapevine virus infection with airborne imaging spectroscopy. *Phytopathology*. 2023;**113**:1439–46
- Gambhir N, Paul A, Qiu T. *et al.* Non-destructive monitoring of foliar fungicide efficacy with hyperspectral sensing in grapevine. *Phytopathology*. 2023
- Jenkins M, Mannsfeld A, Nikzad S. *et al.* Novel algorithms for high-resolution prediction of canopy evapotranspiration in grapevine. *OENO One*. 2023;**57**:3

34. Saxena A, Pesantez-Cabrera P, Ballapragada R et al. Grape cold hardiness prediction via multi-task learning. *AAAI*. 2023;**37**: 15717–23
35. Feng L, Zhang Z, Ma Y. et al. Alfalfa yield prediction using UAV-based hyperspectral imagery and ensemble learning. *Remote Sens*. 2020;**12**:2028
36. Yoosefzadeh-Najafabadi M, Earl HJ, Tulpan D. et al. Application of machine learning algorithms in plant breeding: predicting yield from hyperspectral reflectance in soybean. *Front Plant Sci*. 2021;**11**:624273
37. Ali I, Cawkwell F, Green S. et al. Application of statistical and machine learning models for grassland yield estimation based on a hypertemporal satellite remote sensing time series. 2014 *IEEE Geoscience and Remote Sensing Symposium*. Quebec City, QC, Canada, 2014,5060–3
38. He X, Zhao K, Chu X. AutoML: a survey of the state-of-the-art. *Knowl-Based Syst*. 2021;**212**:106622
39. Erickson N, Mueller J, Shirkov A. et al. AutoGluon-tabular: robust and accurate AutoML for structured data. 2020
40. Fakoor R, Mueller JW, Erickson N et al. Fast, accurate, and simple models for tabular data via augmented distillation. *Advances in Neural Information Processing Systems*. Vancouver: Curran Associates, Inc.; 2020; 8671–81
41. Raza A, Razzaq A, Mehmood SS. et al. Impact of climate change on crops adaptation and strategies to tackle its outcome: a review. *Plan Theory*. 2019;**8**:34
42. Körner C. The cold range limit of trees. *Trends Ecol Evol*. 2021;**36**: 979–89
43. Krantz M, Zimmer D, Adler SO. et al. Data management and modeling in plant biology. *Front Plant Sci*. 2021;**12**:717958
44. Ellis JL, Jacobs M, Dijkstra J. et al. Review: synergy between mechanistic modelling and data-driven models for modern animal production systems in the era of big data. *Animal*. 2020;**14**:s223–37
45. Azodi CB, Tang J, Shiu S-H. Opening the black box: interpretable machine learning for geneticists. *Trends Genet*. 2020;**36**:442–55
46. Dokoozlian NK. Chilling temperature and duration interact on the Budbreak of 'Perlette' grapevine cuttings. *HortScience*. 1999;**34**:1–3
47. Londo JP, Kovaleski AP. Deconstructing cold hardiness: variation in supercooling ability and chilling requirements in the wild grapevine *Vitis riparia*: cold hardiness in *Vitis riparia*. *Aust J Grape Wine Res*. 2019;**25**:276–85
48. Londo JP, Johnson LM. Variation in the chilling requirement and budburst rate of wild *Vitis* species. *Environ Exp Bot*. 2014;**106**: 138–47
49. Peña Quiñones AJ, Keller M, Salazar Gutierrez MR. et al. Comparison between grapevine tissue temperature and air temperature. *Sci Hortic*. 2019;**247**:407–20
50. Kovaleski AP. Woody species do not differ in dormancy progression: differences in time to budbreak due to forcing and cold hardiness. *Proc Natl Acad Sci*. 2022;**119**:e2112250119
51. Salazar-Gutiérrez MR, Chaves-Cordoba B. Modeling approach for cold hardiness estimation on cherries. *Agric For Meteorol*. 2020;**287**:107946
52. Liu J, Lindstrom OM, Chavez DJ. Differential thermal analysis of 'Elberta' and 'Flavorich' peach flower buds to predict cold hardiness in Georgia. *HortScience*. 2019;**54**:676–83
53. Luedeling E, Fernandez E. chillR: Statistical Methods for Phenology Analysis in Temperate Fruit Trees. 2022
54. Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. *Int J Forecast*. 2004;**20**:5–10
55. De Rosa V, Vizzotto G, Falchi R. Cold hardiness dynamics and spring phenology: climate-driven changes and new molecular insights into grapevine adaptive potential. *Front Plant Sci*. 2021;**12**:591
56. Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?": explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco California USA: ACM; 2016;1135–44
57. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates Inc.; 2017;4768–77
58. Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. *Knowl Inf Syst*. 2014;**41**:647–65
59. Chang W, Cheng J, Allaire JJ. et al. shiny: Web Application Framework for R. 2022
60. Badr G, Hoogenboom G, Abouali M. et al. Analysis of several bioclimatic indices for viticultural zoning in the Pacific northwest. *Clim Res*. 2018;**76**:203–23