

Research Paper

Integrating explainable machine learning models with geospatial features to enhance the prediction of late-spring frost risks in apple production

Jie Zhou^{a,b}, Xiaoya Ru^{b,f}, Zhihao He^{a,b}, Kaiyuan Gong^{a,b}, Jinghong Wang^{c,d}, Zhao Wang^{c,d}, Meirong Li^{c,d}, Hao Feng^e, Qiang Yu^e, Jianqiang He^{a,b,d,*}

^a Key Laboratory for Agricultural Soil and Water Engineering in Arid Area of Ministry of Education, Northwest A&F University, Yangling 712100, Shaanxi, China

^b Institute of Water-Saving Agriculture in Arid Areas of China, Northwest A&F University, Yangling 712100, Shaanxi, China

^c Shaanxi Meteorological Service Center of Agricultural Remote Sensing and Economic Crops, Xi'an, Shaanxi 710014, China

^d Key Laboratory of Eco-Environment and Meteorology for the Qinling Mountains and Loess Plateau, Shaanxi Provincial Meteorological Bureau, Xi'an, Shaanxi 710015, China

^e State Key Laboratory of Soil and Water Conservation and Desertification Control, Northwest A&F University, Yangling 712100, China

^f College of Horticulture, Zhejiang Agricultural and Forestry University, Lin'an 311300, Hangzhou, China

ARTICLE INFO

Article history:

Received 28 January 2026

Received in revised form 8 July 2026

Accepted 24 August 2026

Available online xxxx

Keywords:

Apple

Late-spring frost

Explainable machine learning model (EML)

Shapley additive explanations (SHAP)

Channel attention mechanism (CAM)

ABSTRACT

Late-spring frost during the flowering period is one of the most severe agro-meteorological disasters affecting apple production in China. Traditional low-temperature experiments and frost-risk index based on apple cultivars and did not fully consider the impacts of other important meteorological factors and the geographical features. Machine learning models could more comprehensively capture these impacts, although with relatively low interpretability.

In this study, we developed an explainable frost risk prediction framework by integrating daily meteorological observations with geographic variables across the four major apple-growing regions of China (Bohai Bay, Loess Plateau, Southwest Cool Highland, and Xinjiang). Ten candidate predictors were screened using Pearson correlation and variance inflation factors to remove highly collinear variables. Two experimental datasets were constructed: a combination of only meteorological variables and a combination of meteorological and geographic variables. Five machine learning models (Random forest, RF; Support vector machine, SVM; Multilayer perceptron, MLP; Extreme gradient boosting, XGBoost; Long short-term memory neural network, LSTM) and an attention-enhanced deep learning architecture (LSTM-CAM) were then trained to predict late-spring frost risk during apple flowering. Model interpretability was achieved using Shapley Additive Explanations (SHAP). Results showed that RF achieved the best performance (overall accuracy = 0.90, AUC = 0.96). SHAP analysis identified minimum temperature as the dominant meteorological predictor and latitude as the most influential geographic factor. LSTM-CAM model effectively captured the interactions between meteorological and geographical variables by adaptively weighting feature importance through its attention mechanism.

This study demonstrates that an explainable machine learning approach, which integrates multi-source meteorological and geospatial features, is effective in forecasting late-spring frost risks, thereby providing a robust decision-support tool for climate-resilient apple production and orchard management.

© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Apple (*Malus domestica* Borkh.) is a fruit crop widely cultivated in temperate regions in the world. Well-known for its high nutritional value, richness in dietary fiber and various vitamins, apple has become one of the most popular fruits worldwide. China ranks the largest

producer and consumer of apples globally, with a total cultivation area of 1.955×10^6 hm² in 2023, accounting for more than 40% of the world's apple cultivation area and production (FAOSTAT, 2013). However, spring frost injury during apple flowering stage have become one of the most devastating meteorological disasters that affect apple production in China (Qu and Zhou, 2016).

Frost damage refers to a type of short-term low temperature disaster in which plants are suddenly exposed to cold air intrusions or rapidly cooled by intensive ground radiation during a relatively warm period. The body temperature of apple trees could drop below 0 °C, damaging

* Corresponding author at: Key Laboratory for Agricultural Soil and Water Engineering in Arid Area of Ministry of Education, Northwest A&F University, Yangling 712100, Shaanxi, China.

E-mail address: jianqiang_he@nwsuaf.edu.cn (J. He).

<https://doi.org/10.1016/j.aia.2026.08.017>

2589-7217/© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Nomenclature

Symbol	Definition (Unit)
Alt	Altitude (km)
Asp	Aspect (°)
AUC	The area under the receiver operating characteristic curve (%)
CAM	Channel attention mechanism (-)
eXML	Explainable machine learning (-)
FN	False negatives (-)
FP	False positives (-)
FPR	False positive rate (%)
Hum	Humidity (%)
Lat	Latitude (°)
LSTM	Long-short term memory neural network (-)
Lon	Longitude (°)
MLP	Multilayer perceptron (-)
OA	Overall accuracy (%)
Pre	Precipitation (mm)
r	Pearson correlation coefficient (-)
RF	Random forest (-)
ROC	Receiver operating characteristic curve (-)
SHAP	Shapley Additive exPlanations (-)
Slo	Slope (°)
Sr	Solar radiation (MJ m^{-2})
SVM	Support vector machine (-)
Tav	Average temperature (°C)
TP	True positives (-)
TPR	True positive rate (%)
TN	True negatives (-)
$T_{\min}$	Minimum temperature (°C)
VIF	Variance inflation factor (-)
Ws	Wind speed (m s^{-1})
XGBoost	Extreme gradient boosting (-)

their photosynthetic tissues, resulting in reduced yield and lower quality, or even leading to a complete crop failure (Frederiks et al., 2012; Inouye, 2000). The greatest risk of frost injury to woody plants in temperate climates usually occurs during the period of vegetation onset in spring (Vitasse et al., 2018). For fruit trees, late-spring frost primarily occurs during the budding and flowering stages. The newly grown organs such as buds, leaves, and flowers are more susceptible to frost injury than woody tissues (Cebulj et al., 2021; Rodrigo, 2000). The impacts could be long-lasting and probably persist for several years after the disaster (Inouye, 2000; Vanoni et al., 2016), thus causing a decline in apple yield and quality, as well as serious economic losses for local apple growers.

Recent reports indicated that the major apple-growing regions of China often experienced at least one mild frost damage every

2–3 years and one severe frost damage every 3–5 years during apple flowering stage (Yang et al., 2023). For example, the four major apple-growing regions in China, especially the Loess Plateau, suffered an extraordinarily severe late-spring frost damage in early April of 2018 (Fig. 1), which had terribly impacted local apple production. Approximately 53.7% of the national apple-growing counties were severely affected, where apple fruit-setting rates dropped by about 46% compared to previous years, resulting in an overall reduction of 17.9%–19.4% for national apple production (Wang et al., 2020; Qiu et al., 2020). In the April of 2020, apple orchards in China were hit again by frost damage due to the repetitive intrusions of strong cold airs. In the Yan'an City of the Loess Plateau, various degrees of frost damages were caused in April 10th–13th, where mild frost damage affected about 2313 hm^2 or 37.5%; moderate frost damage affected about 1860 hm^2 or 30.2%; and severe frost damage affected about 1993 hm^2 or 32.3% of total cultivation area in the city (Guo et al., 2022; Liu et al., 2021a, 2021b; Yang et al., 2022; Yang et al., 2020a, 2020b).

The main reasons for apple late-spring frost, where air temperatures drop below 0 °C, could be attributed to radiative frost and advective frost (Kalma et al., 1992). Radiative frost usually occurs under the weather conditions of clear sky, calm wind, and high-pressure, where strong long-wave radiation causes rapid heat loss from the ground at night, leading to a sharp temperature drop and causing water vapor in the air to condense directly into frost. Advective frost occurs when robust intrusions of cold airs make air temperatures falling below the frost point and thereby induce frost formation. This phenomenon frequently occurs under windy and overcast weather conditions in early spring and late autumn.

Due to the impacts of land surface heterogeneity on the spatial variation of near surface temperatures, the occurrence of frost is closely related to land surface characteristics (Kotikot et al., 2020). The frequency and intensity of frost could vary dramatically over time and space according to the spatial geographic factors such as latitude, altitude, and topography (Allevato et al., 2019; Alvares et al., 2018; Diniz et al., 2021; Elli et al., 2020). For instance, Diniz et al. (2021) reported that geographical factors such as latitude, relative altitude, and longitude were strong determinants for frost risks in the artificial forest regions of south-central Brazil. Vitasse et al. (2018) showed that risks of frost exposure to fruit and forest trees increased at elevations above 800 m but kept unchanged at lower elevations over the period of 1975–2016 in Switzerland. The increased frost risks to fruit and forest trees in high-altitude regions were mainly because the advancement of tree spring phenology outpaced the occurrence of early late-spring frosts. In contrast, the fundamental reason for the unchanged frost risks in low-altitude regions was because both the rate and magnitude of apple spring phenological advancement was constrained by insufficient chilling accumulation and photoperiod limitation.

Frost events tend to occur more frequently in narrow valleys, then in low-lying and flat areas, while higher and more convex areas rarely experience frost (Lindkvist et al., 2000). This kind of pattern can be



Fig. 1. Apple flowers with different levels of late-spring frost injury in the four major apple-growing regions of China on April 6th in 2018.

explained by the concept of cold air pooling, where cold, dense air flows downslope, settles beneath warmer air, and causes stagnant conditions (Yoshino, 1984). Additionally, slope orientation also plays a crucial role in influencing air and soil temperatures (Monteith and Unsworth, 2013), since different slopes receive varying amounts of solar radiation. In the Northern Hemisphere, south-facing slopes receive more solar radiation during the day, leading to higher daytime air and soil temperatures. In contrast, north-facing slopes are more shaded and receive less solar radiation, resulting in cooler temperatures. Thus, frost is more likely to affect the cooler north-facing slopes than to affect the warmer south-facing slopes (Allevato et al., 2019).

So far, different methods, such as low-temperature control experiments on different tree species, construction of frost risk indices, and machine learning models, have been used to assess the risks of late-spring frost (Chen et al., 2024; Diniz et al., 2021; Eccel et al., 2007; Morin and Chuine, 2014; Omazić et al., 2024; Ru et al., 2023a; Ru et al., 2023b). However, low-temperature control experiments on different tree species were usually constrained by various factors, such as the need to recreate daily temperature variations under a climate warming scenario in the field and the frost damage to tree tissues (Morin and Chuine, 2014). Traditional temperature-based statistical methods generally estimate frost occurrence using predefined air-temperature thresholds or frost indices. Nevertheless, temperature-threshold-based methods primarily characterize the climatic conditions under which frost is likely to occur, whereas they often have limited ability to represent the complex interactions among meteorological conditions, topography, and local geographic characteristics that jointly determine the occurrence of actual frost damage (Omazić et al., 2024). At the same time, the current frost risk indices often did not fully account for all meteorological factors except for temperature and the spatial geographical features that influence the occurrence of spring frost (Ru et al., 2023a).

Currently, machine learning models have become a powerful tool to detect frost risks since they could provide useful information about frost risk in complex terrains without requiring a physical explanation for the development of cold air dynamics (Eccel et al., 2007; Ghielmi and Eccel, 2006). For instance, Diniz et al. (2021) used locally generated terrain data to predict frost risks in artificial forests in Brazil. They found that the random forest (RF) classifier was the best classifier for the occurrence of frost events. Omazić et al. (2024) effectively captured the conditions under which frost events would most likely to occur in Croatia based on machine learning models. Generally, machine learning models can reliably classify the occurrence of extreme climatic events (e.g., late-spring frost in this study) using historical record data and provide robust likelihood predictions of such events through supervised learning. However, most previous relevant studies usually relied on limited meteorological data or spatial geographical information in their predictions and did not include all possible predictive variables that could affect the occurrence of frost events, which made these studies unable to fully capture the mechanisms of frost risk occurrence. As pointed out by Omazić et al. (2024), the predictive capability could be increased if wind speed and net radiation were accounted for since they could provide more information about the nature of frost event occurrence. Thus, it is necessary to comprehensively assess the meteorological and geographic factors that influence the formation and occurrence of frost risks.

Compared to the process-based mechanical models, machine learning models are often considered as “black boxes” since there is typically a limited understanding of how the outputs are produced from the model inputs. This resulted in a low transparency and interpretability of the machine learning models (Lundberg and Lee, 2017). However, the explainable machine learning (eXML) methods have garnered great attentions since they could improve prediction interpretability and provide reliable insights into the otherwise opaque machine learning frameworks (Gaur and Drewry, 2024; Gou et al., 2023; Jiang et al., 2024; Liu et al., 2024; Lundberg et al., 2020). For instance, the SHAP

(SHapley Additive exPlanations) theory, based on a unified framework from game theory, constructs an additive explanation model to enhance the interpretability of machine learning models (Shapley, 1997). The SHAP theory calculates the linear weighted contribution value of each feature variable for each sample, aiming to explain the model output and ensure the causal inference within the model. Specifically, SHAP quantifies the contribution of each feature variable to model output by calculating the average marginal contribution of all features to the prediction, thereby attributing the final model output to individual input features (Gunning et al., 2019; Lundberg and Lee, 2017). However, there was no reported study that combined the machine learning models and the SHAP to explain the relationship between the meteorological & geographical features and the occurrence of late-spring frost events in the major apple-growing regions in China.

Hence, the objectives of this study were (1) to assess the efficiency of different machine learning models in predicting apple frost risks; (2) to assess the impacts of different combinations of input variables on the predictive performance of the machine learning models for apple frost risk prediction; and (3) to use the SHAP method to interpret the importance of each input variable in the “black-box” machine learning models and to identify the key influential factors for the occurrence of late-spring frost risks in the four major apple-growing regions in China. The prediction framework for apple late-spring frost risks based on interpretable machine learning models in this study can provide a valuable tool to quantitatively analyze the impacts of multiple meteorological and geographic features on the occurrence of late-spring frost risks of fruit trees in the world.

2. Materials and methods

2.1. Study area

The study area covers the four main apple-growing regions in China (Fig. 2), including the Bohai Bay (BB) region, the Loess Plateau (LP) region, the Southwestern Cool Highland (SCH) region, and the Xinjiang (XJ) region (Qu and Zhou, 2016). Specifically, the Loess Plateau (LP) region is the largest apple production base of China, accounting for about 57.2% of the national apple planting area and 53.0% of the national apple production (Tang et al., 2021). The LP region, including the provinces of Shaanxi, Gansu, Shanxi, Henan, and Ningxia, is characterized by superior ecological conditions, deep soil layer, high altitude, and abundant sunlight, making it the largest and most concentrated high-quality apple production area in China. The Bohai Bay (BB) region, which primarily includes the provinces of Shandong, Liaoning, and Hebei and is known for its high level of industrialization in apple cultivation, is the second-largest apple-growing area, with about 26.9% of the national planting area and 36.0% of national production (Tang et al., 2021). The Southwestern Cool Highland region, which includes the provinces of Yunnan, Sichuan, and Guizhou and is characterized by unique microclimates and ample precipitation, mainly produces early and mid-season apples (Liu et al., 2021). The apple production in Xinjiang region is mainly concentrated in the Aksu Prefecture of Xinjiang Province, through which the Tarim River flows. This mid-latitude region is westerly shielded by the Pamir Plateau and northerly shielded by the Tianshan Mountains. These unique topographic conditions make it a unique and advantageous apple-growing region in China (Ilmiyaz et al., 2022; Qiu et al., 2020). The Southwestern Cool Highland (SCH) region and the Xinjiang (XJ) region totally account for about 15.9% of national apple planting area and 11.0% of national production (Liu et al., 2021; Ru et al., 2023b).

2.2. Datasets

2.2.1. Meteorological data and geographic data

Daily meteorological data for the representative apple-growing sites were obtained from the China Meteorological Data Network

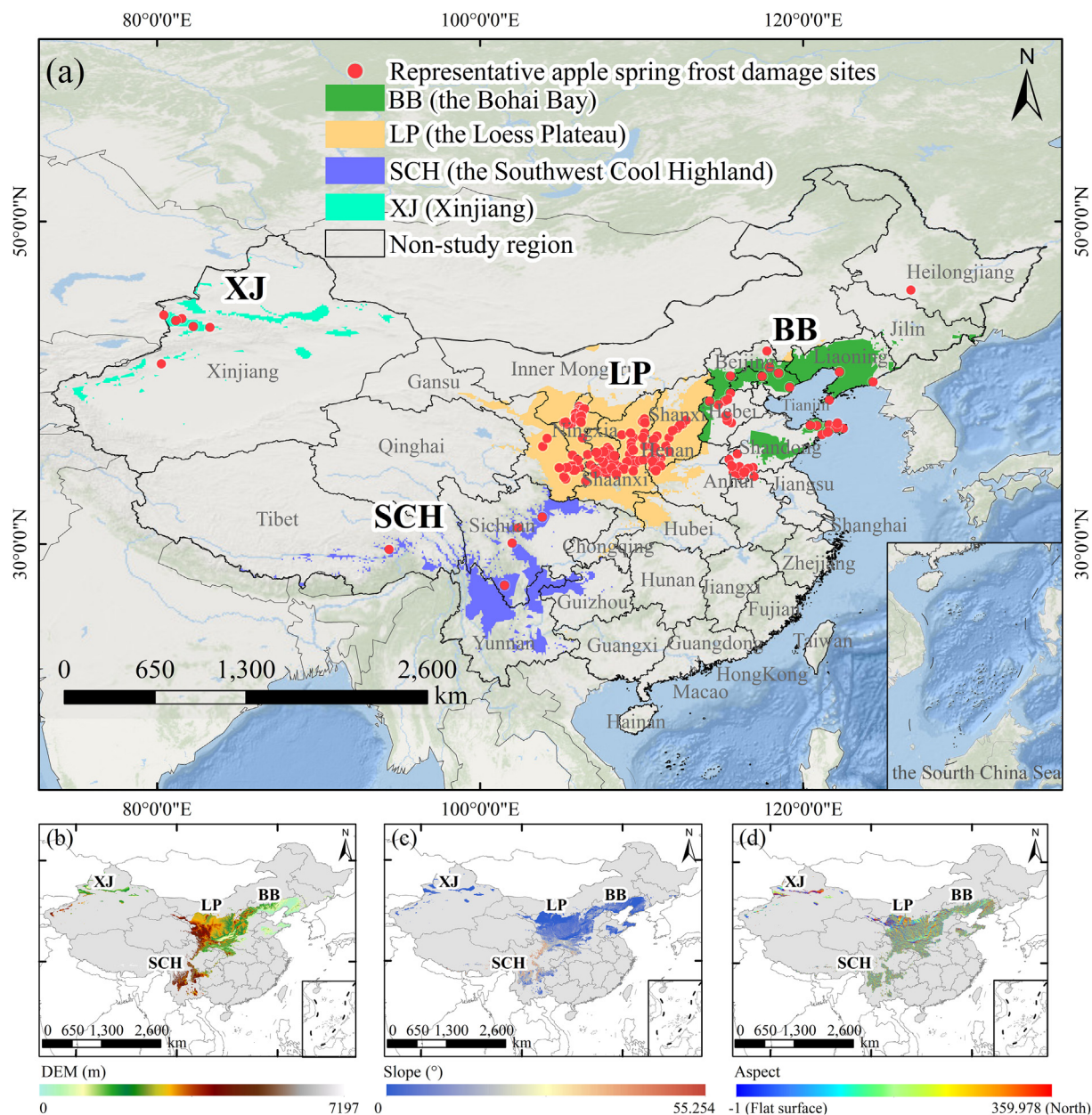


Fig. 2. (a) Distributions of the four main apple-growing regions (the Bohai Bay, BB; the Loess Plateau, LP; the Southwest Cool Highland, SCH; and Xinjiang, XJ) in China. The red solid dots represent the representative apple-growing sites where late-spring frost damage records were collected in this study. Panels (b), (c), (d) are the topography, slopes, and aspects of the four main apple-growing regions in China, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(<https://data.cma.cn/>), including average temperature ($^{\circ}\text{C}$), the minimum temperature ($^{\circ}\text{C}$), relative humidity (%), average wind speed (m s^{-1}), solar radiation (MJ m^{-2}), and precipitation (mm). Geographical data included slope ($^{\circ}$), aspect ($^{\circ}$), longitude ($^{\circ}$), latitude ($^{\circ}$), and altitude (km). Altitude data were sourced from the SRTM (Shuttle Radar Topography Mission) 30-m digital elevation dataset of NASA (Farr et al., 2007); slope and aspect data were calculated on the Google Earth Engine (GEE) platform. These variables are generally recognized as influential factors for frost events (Diniz et al., 2021; Kotikot et al., 2018). Meteorological and geographical feature variables were used as predictor variables in the machine learning models to predict the binary dependent variable, where 0 and 1 separately represents no frost occurrence and frost occurrence. The range and distribution of the predictor variables were shown with several descriptive statistical parameters (Table 1). No data transformation was conducted before the modeling.

2.2.2. Records of apple late-frost damage

In this study, we collected the historical data about late-spring frost damages to apples at 179 sites in 2018 and at 86 sites in 2020 (Table A.1) across the four main apple-growing regions in China. The years of 2018 and 2020 were two typical recent years with serious late-spring frost damage for apple production in China, providing comprehensive and spatially extensive frost observation records from representative sites covering the four major apple-growing regions. These data were mainly obtained from the China Meteorological Disaster Management System (<https://10.164.146/disaster/index>), the Shaanxi Meteorological Bureau (<http://sn.cma.gov.cn/>), the Encyclopedia of China's Meteorological Disasters (Comprehensive Edition) (Wen and Ding, 2008), the China Meteorological Disasters Yearbook (China Meteorology Administration, 2020; 2022), relevant reports in the Apple Pest Control Information Network (<https://apple-ipm.hebau.edu.cn/index.htm>), and the academic papers about apple late-spring frost risks and

Table 1

Statistics of the influential variables during apple flowering period in April for the predictions of late-spring frost risks across the four major apple-growing regions in China.

Statistics	Lat (°)	Lon (°)	Alt (km)	Asp (°)	Slo (°)	Hum (%)	T_{av} (°C)	T_{min} (°C)	Pre (mm)	Ws (m s ⁻¹)	Sr (MJ m ⁻²)
Minimum	26.867	80.267	4.6	0	0	7	-3.6	-10.7	0	0.2	4.504
1st Quartile	34.950	105.896	780.9	90	2	27	8.4	1.9	0	2.0	16.123
Median	35.717	109.067	1072.0	180	4	38	11.4	4.8	0	2.6	22.015
Mean	36.167	108.884	1024.4	178	7	41	11.2	4.5	0.6	2.8	19.830
3rd Quartile	37.433	110.333	1269.0	270	10	53	14.4	7.1	0.1	3.3	23.830
Maximum	45.748	126.669	3394.6	357	38	93	21.6	17.9	19.4	23.8	27.624

Note: Lat = latitude (°), Lon = longitude (°), Alt = Altitude (km), Asp = Aspect (°), Slo = Slope (°), Hum = Humidity (%), T_{av} = Average temperature (°C), T_{min} = Minimum temperature (°C), Pre = Precipitation (mm), Ws = Wind speed (m s⁻¹), Sr = Solar radiation (MJ m⁻²).

damages (Guo et al., 2020; Guo et al., 2022; Li, 2018; Liu et al., 2021a, 2021b; Liu et al., 2020; Ma et al., 2018; Shang, 2019; Xu et al., 2019; Yang et al., 2022; Yang et al., 2020a, 2020b; Yin, 2019). These records included the dates of frost occurrence, longitude and latitude of the affected sites, and the severity of late-spring frost disaster (Table A.1).

To ensure consistency among different data sources, frost-damage observations were harmonized according to the Chinese Meteorological Industry Standard (QX/T 392-2017): Grade of Flowering-Stage Frost Damage for Fuji Apple (China Meteorology Administration, 2017), frost damage is classified into three severity levels (mild, moderate, and severe) based on visible injury symptoms of apple flowers, corresponding frost-injury rates, and yield-loss rates. Specifically, mild frost damage is characterized by yellow-brown petals while the anthers, filaments, and ovaries of the central flower remain intact, corresponding to a frost-injury rate of less than 30%. Moderate frost damage is characterized by slight yellowing of pistils, yellow-brown anthers, reddened and deformed filaments, and partial blackening of the ovary, corresponding to a frost-injury rate of 30–80%. Severe frost damage is characterized by severe injury to pistils, completely blackened stigmas, blackened and deformed filaments, blackened internal tissues of the anthers, and complete blackening of the ovary, corresponding to a frost-injury rate greater than 80%. For binary classification modeling, all frost-damage severity levels (mild, moderate, and severe) were merged into a single frost-occurrence category and assigned a label of 1, whereas observations without frost damage were assigned a label of 0. The non-frost samples (label = 0) were strictly defined based on the same observational records as the frost-damage events. They were defined as the synchronous records from the same observation sites and dates where NO frost damage was sustained, which ensured that both frost and non-frost samples originated from the same spatial geographic features and meteorological background during the flowering period. Furthermore, all predictor variables were normalized using the Min-Max scaling method prior to model training to ensure comparability among variables with different units and magnitudes. These data were used for later machine learning modeling.

The suitable input variables were then screened to avoid the influence of collinearity on model output before modeling. First, all of the candidate meteorological and geographical variables were standardized to be dimensionless. Then, the Pearson correlation coefficients (r) between the candidate variables were calculated. It indicated there were no strong collinearity between the variables when $|r| < 0.7$, but highly correlated variables ($|r| > 0.7$) were removed from the dataset to avoid collinearity (Dormann et al., 2013). Finally, the variance inflation factor (VIF) was calculated to explore the multicollinearity of the variables. There was no multicollinearity among the variables when $VIF < 10$ (Wang et al., 2023). The finally selected input variables for the machine learning models were further detailed in Section 2.4.

2.3. Performance evaluation for the machine learning models

In this study, the receiver operating characteristic (ROC) curve was used to evaluate the predictive performance of the five machine learning models in identifying the occurrence and non-occurrence of apple

late-spring frost events. The ROC curve is a graphical tool used to assess the diagnostic ability of binary classifiers and has been widely used in research fields such as medicine, agriculture, and natural disasters (Baker, 2003; Kotikot et al., 2018; Tang et al., 2023). ROC curve is constructed by plotting pairs of true positive rate (TPR, sensitivity) and false positive rate (FPR, 1-specificity) with FPR as the x-axis and TPR as the y-axis. TPR is the proportion of positive samples correctly predicted as positive (Eq. (1)), while FPR is the proportion of negative samples incorrectly predicted as positive (Eq. (2)). By changing the classification threshold, a trade-off between these two values is obtained, which in turn allows for the ROC curve plotting (Kotikot et al., 2018). The closer the ROC curve is to the upper left corner (high TPR and low FPR), the better the model's performance.

The area under the receiver operating characteristic curve (AUC) quantifies the ROC curve's performance. AUC values range between 0 and 1 and measure the model's effectiveness in distinguishing between positive and negative samples. A higher AUC value indicates stronger diagnostic ability, with the corresponding ROC curve closer to the upper left corner (Carter et al., 2016). The evaluation criteria are as follows: $0.5 < AUC \leq 0.6$ indicating no ability to distinguish between positive and negative samples; $0.6 < AUC \leq 0.7$ indicating poor performance; $0.7 < AUC \leq 0.8$ indicating average performance; $0.8 < AUC \leq 0.9$ indicating good performance; and $0.9 < AUC \leq 1$ indicating excellent performance in distinguishing between positive and negative samples (Tang et al., 2023). Overall accuracy indicates a comprehensive evaluation metric of the model, representing the proportions of both frost occurrence and non-occurrence events (Eq. (3)). It considers all categories of predictions and reflects the model's classification accuracy across the entire dataset.

$$TPR(sensitivity) = \frac{TP}{TP + FN} \quad (1)$$

$$FPR(1 - specificity) = \frac{FP}{FP + TN} \quad (2)$$

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

where TP and TN are the numbers of correctly retrieved frost events (true positive) and non-frost events (true negative); FP and FN are the numbers of incorrectly retrieved frost events (false positive) and non-frost events (false negative), respectively.

2.4. Machine learning models for apple late-spring frost prediction

2.4.1. Five different machine learning models

In this study, five different machine learning models were established to predict the occurrence of apple late-spring frost events, including the Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Multilayer Perceptron (MLP), and Long Short-Term Memory Neural Network (LSTM). These models are recognized for their abilities to handle complex and nonlinear data generated by the physical processes of climatic conditions and have been widely

used for frost risk predictions (Cao et al., 2009; Diniz et al., 2021; Naing and Htike, 2015; Omazić et al., 2024; Shi et al., 2015).

The Random Forest (RF) algorithm is an ensemble learning method that aggregates the predictions of multiple decision trees to obtain the final prediction. It is mainly used for classification and regression tasks (Breiman, 2001). Each internal node performs a feature-based condition split, each branch represents the result of that condition (True or False), and each leaf node represents a class label (for classification problems) or the mean prediction value (for regression problems). The RF algorithm is a very powerful machine learning algorithm for its efficiency and robustness and has been widely used across various domains. In this study, the RF model was constructed through the “*randomForest*” package in R language.

The Support Vector Machine (SVM) algorithm was introduced by Vapnik (1995) based on statistical learning theory, aiming to solve classification and regression problems. The core idea of SVM is to find an optimal separating hyperplane that maximizes the margin between different classes of data points in a high-dimensional feature space. This hyperplane can be linear; or through kernel functions (such as radial basis function, linear kernel, polynomial kernel, and etc.), nonlinear problems can be transformed into linearly separable problems. The advantage of SVM lies in its performance in high-dimensional spaces and its ability to handle nonlinear boundaries with good generalization capabilities. Moreover, SVM defines the hyperplane using support vectors, meaning that only a small number of representative sample points are considered during the training process to enhance the computational efficiency. In this study, the SVM model was constructed through the “*e1701*” package in R language.

As a relatively new decision-tree-based model, the Extreme Gradient Boosting (XGBoost) has demonstrated good performance due to its ability to handle missing data effectively. This algorithm incorporates L1 and L2 regularization techniques in its cost function, which effectively reduces overfitting and controls model complexity, making the learned model simpler and faster to compute (Natekin and Knoll, 2013). This is also a feature that distinguishes XGBoost from traditional Gradient Boosting Decision Tree (GBDT) algorithms. In this study, the XGBoost model for apple frost event prediction was constructed through the “*xgboost*” package in R language.

The Multilayer Perceptron (MLP) algorithm is a fully connected feedforward neural network model composed of multiple neurons arranged in layers, including at least three layers (i.e., input layer, output layer, and hidden layers). Each neuron undergoes nonlinear transformations through multiple hidden layers, connected with weight coefficients affecting the network architecture, ultimately reaching the output layer to complete classification or regression tasks (Haykin, 2001). In this study, the MLP model was constructed based on the TensorFlow v2.9 framework (through Keras) in Python v3.9.

The Long Short-Term Memory Neural Network (LSTM) algorithm is a type of Recurrent Neural Network (RNN) model developed by Hochreiter and Schmidhuber (1997). LSTM introduces a memory cell and gating mechanisms, including an input gate, output gate, and forget gate, to address the gradient vanishing and explosion problems often encountered by traditional RNNs when processing long sequences (Greff et al., 2016). The internal state of the LSTM cell acts as a dynamic storage mechanism, enabling the accumulation of key historical information. Through the input gate, LSTM selectively accepts new data to facilitate learning of key predictors, while the forget gate assesses and adjusts the influence of previous historical data. Finally, the output gate determines the predictions generated by the model.

2.4.2. The LSTM-CAM model framework

Additionally, we proposed a LSTM-CAM model framework that combined the LSTM algorithm with the Channel Attention Mechanism (CAM) to further enhance the performance of the LSTM model in predicting complex time-series data (Fig. 3). Channel-wise attention mechanisms have been widely demonstrated to strengthen feature

representation by adaptively reweighting input channels, enabling the model to emphasize the most informative predictors while suppressing less relevant signals (Niu et al., 2021). Recent advances combining attention modules with LSTM architectures have shown improved performance in multi-source environmental time-series prediction, validating the benefits of attention-guided feature selection over vanilla LSTM models (Kumar et al., 2023; Lagap et al., 2025; Tian et al., 2024).

Coupled with the CAM, the LSTM model can more effectively focus on the key features that influence the predictions of apple late-spring frost events. The attention mechanism was initially proposed to mimic the selective perception of human vision, where it essentially selected more important information and suppressed useless information (Vaswani, 2017). In the case of multi-source input variables (such as the meteorological and geographic variables in this study), simply concatenating the variables may not fully leverage the advantages of each feature, while the CAM can adaptively adjust the weights of different feature channels through assigning a weight to each type of feature channel. Thus, the CAM can enhance the focus on important features, provide more granular explanations, and thereby achieve better feature fusion from multi-source data (Niu et al., 2021).

Specifically, a CAM function was defined first, where global average pooling and global max pooling were applied to each channel. Then, the pooled results were processed through two shared fully-connected layers, which could learn the relationships between different features and generate attention weights. These attention weights were then multiplied by the input features to obtain the weighted features, further enhancing the relevance of key features. Finally, the outputs of the two feature channels were combined, passed through a fully connected layer, and regularized using a dropout layer to generate the final output. In this study, the baseline LSTM model was applied to the single-source predictor set containing only meteorological variables, where capturing temporal dependencies was the primary objective. The combined predictor set incorporated both meteorological and geographic variables, forming a heterogeneous multi-channel input space. To more effectively exploit this richer information structure, the LSTM-CAM model was evaluated only under the combined predictor set. The CAM mechanism enabled the network to adaptively reweight feature channels, highlight predictors most strongly associated with late-spring frost formation, and improve multi-source feature integration relative to the baseline LSTM. Both models were all constructed based on the TensorFlow v2.9 in Python v3.9, with LSTM-CAM used as the primary deep-learning architecture in subsequent analyses involving combined variables.

2.4.3. Construction of the prediction models for apple late-spring frost events

The combination of screened meteorological variables and the combination of meteorological and geographic variables were separately input into the five machine learning models (i.e., RF, SVM, XGBoost, MLP, and LSTM) to generate the output of the occurrence of apple late-spring frost events (0 = no frost occurrence and 1 = frost occurrence; Fig. 4). For each model, 70% of the whole dataset was first split into training data and 30% into testing data, with a random seed to ensure the result reproducibility. To further evaluate the robustness of the developed machine-learning models to class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was applied to generate synthetic samples for the minority class. The SMOTE procedure was performed only on the training dataset after the training-testing split, while the independent testing dataset remained unchanged to ensure an unbiased evaluation of model performance. Subsequently, all machine-learning models were retrained and their hyperparameters were re-optimized using the balanced training dataset. The training data were used to evaluate the model's ability to correctly predict frost events and determine the best configuration for each model, which was then applied to determine the model's hyperparameters. Accurate hyperparameters are crucial for model performance, as they can help effectively control model complexity, avoid overfitting in the training

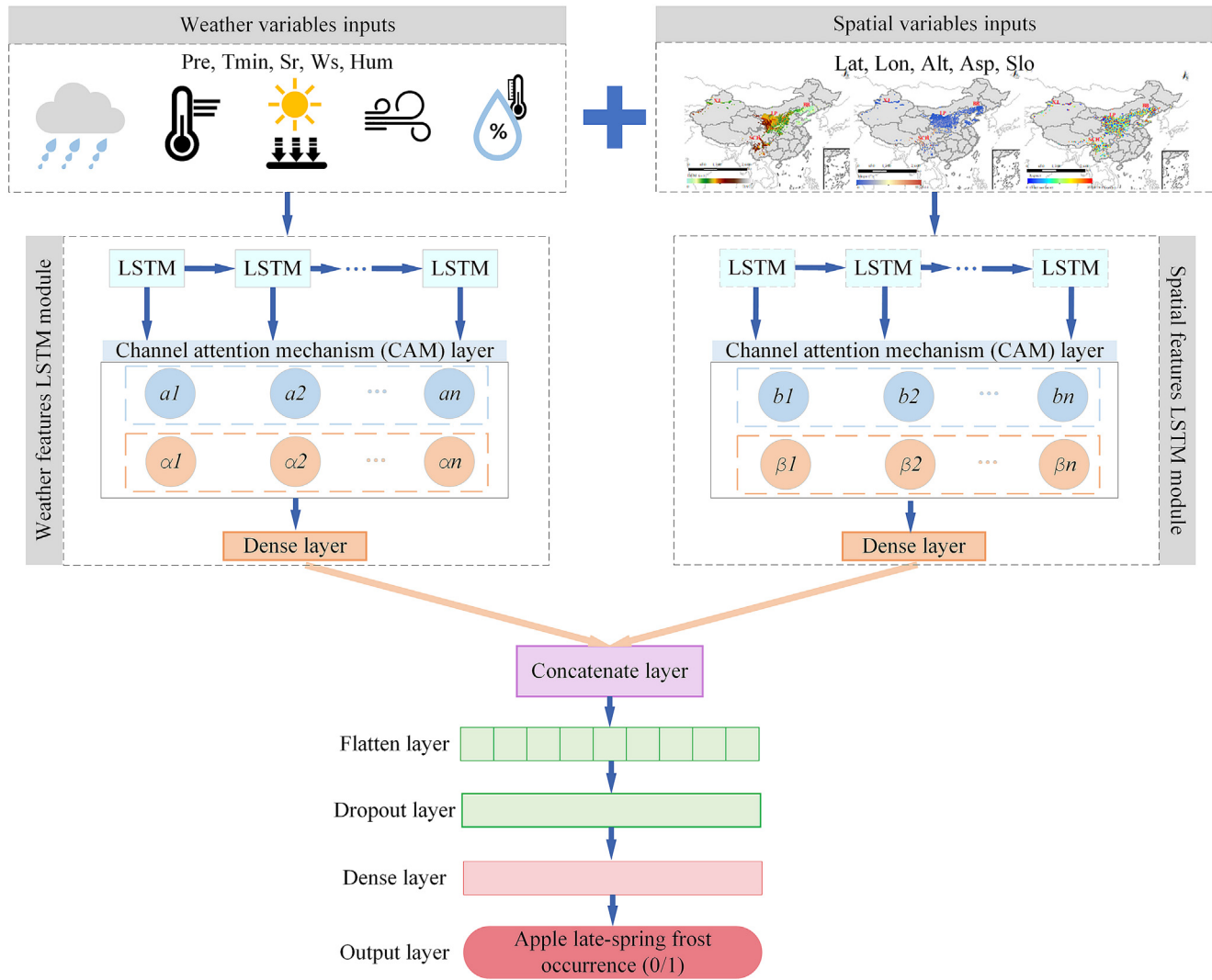


Fig. 3. Architecture of the LSTM-CAM model framework developed in this study for the predictions of apple late-spring frost events. LSTM is the Long Short-Term Memory Neural Network and CAM is the Channel Attention Mechanism.

process, and improve the generalization in the testing process. Hyperparameter tuning must be determined before model training, typically performed using the Grid Search or Random Search methods (Bergstra et al., 2011). This study used the Grid Search method to optimize the parameters of each machine learning model since it could exhaustively search all possible parameter combinations to determine the optimal parameter values (Bergstra and Bengio, 2012). Generally, it is essential to consider the model's generalization ability on unseen data in the processes of determining the optimal configuration of machine learning model and optimizing the hyperparameter.

To more comprehensively assess the stability and generalization of the model's performance, this study used a 10-fold cross-validation to obtain the sample with the highest training accuracy in model training. Cross-validation effectively evaluates the model's performance on different data subsets by repeatedly splitting the dataset for training, ensuring that all data are involved in both training and prediction processes, thereby providing a more reliable performance assessment and hyperparameter adjustment basis (King et al., 2021). Table A.2 listed the ranges of hyperparameters used for the five machine learning models in this study and information about their physical significance. The Receiver Operating Characteristic (ROC) curve and Overall Accuracy (OA) values were used to evaluate the overall performance of the models in distinguishing between the occurrence and non-occurrence

of apple late-spring frost events. More details about the construction of the prediction models of apple late-spring frost events based on the five machine learning models can be found in Fig. 4.

2.5. Prediction contribution quantification based on the Shapley additive exPlanations (SHAP) theory

To enhance the interpretability of the machine learning models constructed in this study, the Shapley Additive exPlanations (SHAP) theory was used to quantify the contribution of different meteorological and geographic variables to the predictions of apple late-spring frost events and to evaluate the relative importance of each variable to the model performance. Compared with the traditional mechanistic models, the complexity of machine learning models presents the challenges in interpretability and understanding of model system dynamics. To address these issues, Lundberg and Lee (2017) developed a common exML algorithm, or the Shapley Additive exPlanations (SHAP), to explain the outputs of machine learning model. The concept of SHAP is derived from the Shapley value in game theory (Shapley, 1997), which defines the average marginal contribution of a player across all coalitions they belong to. SHAP was first used to fairly distribute the gains in cooperative games and later applied to feature attribution problems (Biecek and Burzykowski, 2021). Generally, SHAP is an additive feature attribution

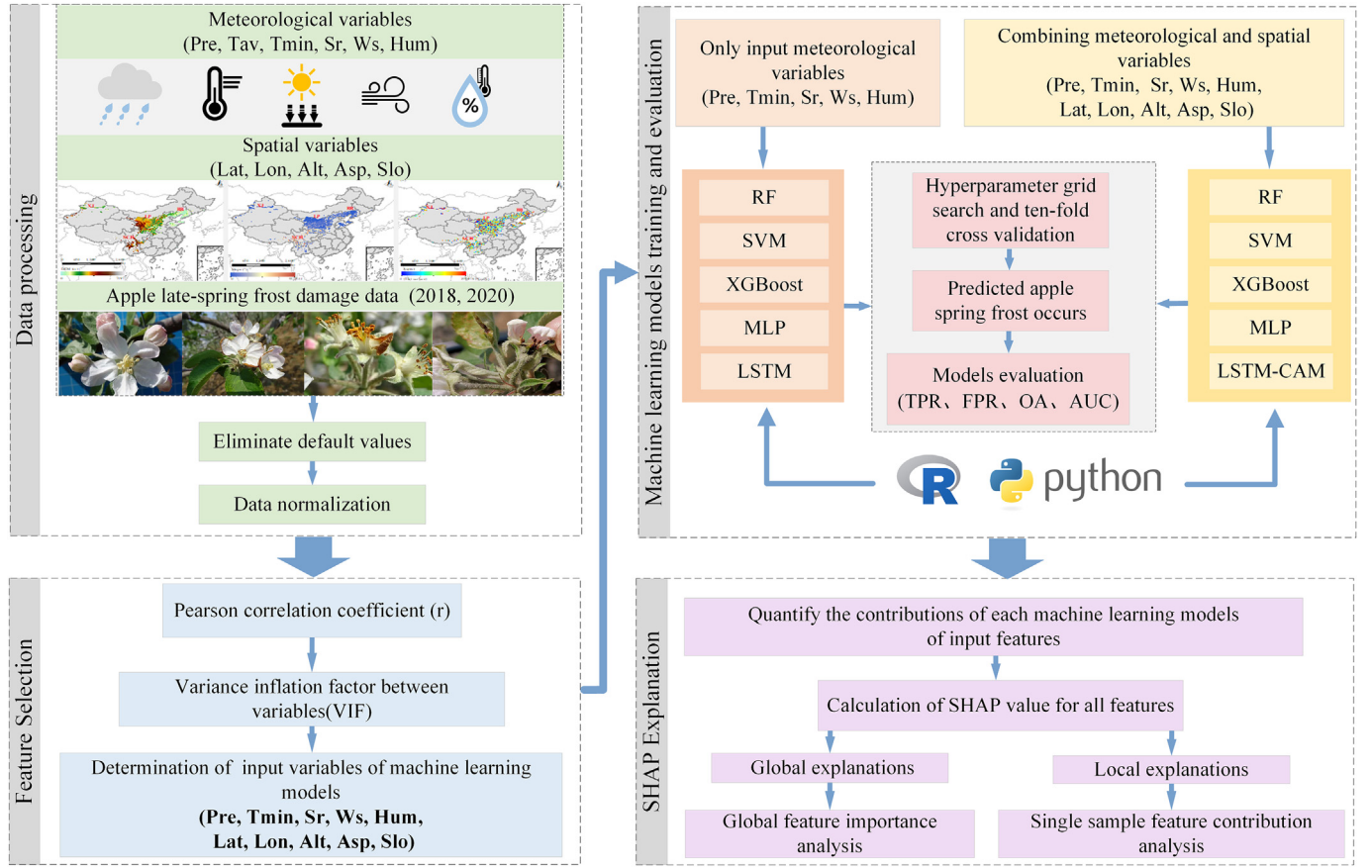


Fig. 4. Flowchart for the construction of prediction models for apple late-spring frost events based on different machine learning algorithms (RF, SVM, XGBoost, MLP, and LSTM) in this study.

method that calculates the weighted average of each feature's marginal contribution by considering all possible feature combinations (Lundberg et al., 2020).

For a model with n -dimensional input features, the set of features is $N = (X_1, X_2, X_3, \dots, X_n)$, namely, meteorological and spatial geographic variables in this study. Then, the contribution of feature subset S to the model output can be calculated through the feature value function $v(S)$ (Eq. (4)).

$$v(S) = \int \hat{f}(x_1, x_2, \dots, x_n) dP_{x \notin S} - E_X(\hat{f}(X)) \quad (4)$$

where S is a subset of the feature set N ; $v(S)$ is the feature value function for feature subset S ; $\hat{f}(x_1, x_2, \dots, x_n)$ is the function to calculate the model's prediction value for the given feature subset S ; P is the cumulative distribution function used to describe the distribution of features outside the feature subset; the conditional expectation E_X is used to estimate the expected value of the model prediction $\hat{f}(X)$ when the feature subset S is fixed. This expected value is the average of $\hat{f}(X)$ when keeping $X_S = x_s$ unchanged under all other possible features X .

Then, the Shapley value (ϕ_i) for each feature i is defined (Eq. (5)).

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} [v(S \cup \{i\}) - v(S)] \quad (5)$$

where $S \subseteq N \setminus \{i\}$ is all possible feature subsets; $\frac{|S|!(n - |S| - 1)!}{n!}$ is the weight of feature i , which calculates the average marginal contribution of feature i across all possible feature subsets, ensuring that the contribution of each feature is fairly considered under all possible feature combinations; $v(S \cup \{i\}) - v(S)$ is the marginal contribution of feature i to subset

S , indicating the change in model prediction when feature i is added to subset S .

If the Shapley value is positive, it means that feature i has a positive impact on model outputs and vice versa. Larger absolute Shapley value indicates the greater impact of feature i on model outputs (Štrumbelj and Kononenko, 2014). Through coupling the five machine learning models with the SHAP method (Fig. 4), it could help us understand the impacts of different meteorological and geographic factors on the occurrence predictions of apple late-spring frost events and gain more insights into the dynamic process about how each input factor affects the model outputs.

Unlike the other eXML techniques, SHAP can provide both global and local explanations for model outputs (Lundberg and Lee, 2017). For global explanation, the SHAP value represents the weighted averaged contribution across all samples to quantify the overall importance of the input feature variables, thereby summarizing how strongly and in which direction each variable influences the model as a whole. For local explanation, on the other hand, the SHAP value represents the contribution of each input variable to a single sample, focusing on explaining the unique features of each individual sample (Lundberg et al., 2020).

In this study, the 'beeswarm' plot was used to depict the local impact of input variables on model outputs. The 'beeswarm' plot shows the impact of each data point, and looks denser when multiple points are concentrated around the same location. A bar chart describes the overall impact of each feature on model output across all prediction samples, which allows us to clearly evaluate the average contribution of each input variable to model output from a global perspective. The length of each bar represents the importance of the feature to model predictions. Through combining the 'beeswarm' plot and bar charts, a more

comprehensive understanding could be gained for the interpretability of SHAP values from both local and global perspectives.

3. Results

3.1. Model input variables screening

To avoid the effects of multicollinearity on model prediction results, the input variables were screened in advance through comprehensively analyzing the relationships between the eleven candidate variables (i.e., *Lat*, *Lon*, *Alt*, *Asp*, *Slo*, *Hum*, *T_{av}*, *T_{min}*, *Pre*, *Ws*, and *Sr*). The results of correlation analyses were summarized in the matrix in Fig. 5.

The correlation between *T_{av}* and *T_{min}* was the highest ($r = 0.75$), which indicated a high level of collinearity between these two variables. The scatter plot also confirmed the linear relationship between these two variables. This strong positive correlation was likely due to the synchrony between these two variables during apple

growing seasons. Thus, the variable *T_{av}* was removed and *T_{min}* was retained to ensure the independence of model input variables. Although there was a certain positive correlation between *Pre* and *Hum* ($r = 0.48$), as well as between *Ws* and *Lon* ($r = 0.22$), these correlations were relatively weak. In addition, the negative correlations between *Alt* and *Lon* ($r = -0.65$), *Sr* and *Hum* ($r = -0.57$), and *Alt* and *Lat* (-0.47) essentially reflect the impacts of climatic and geographical heterogeneity on the distribution patterns of model input variables across the four major apple growing regions in China. Importantly, the absolute values of these negative correlation coefficients were all below 0.7, which did not reach the commonly used threshold for problematic collinearity.

To further investigate the multicollinearity among the remaining 10 variables (i.e., *Lat*, *Lon*, *Alt*, *Asp*, *Slo*, *Hum*, *T_{min}*, *Pre*, *Ws*, and *Sr*), the values of variance inflation factor (VIF) were calculated (Fig. 6). The results showed that all variables had VIF values far less than 3.0, indicating that there was no significant multicollinearity between the variables.

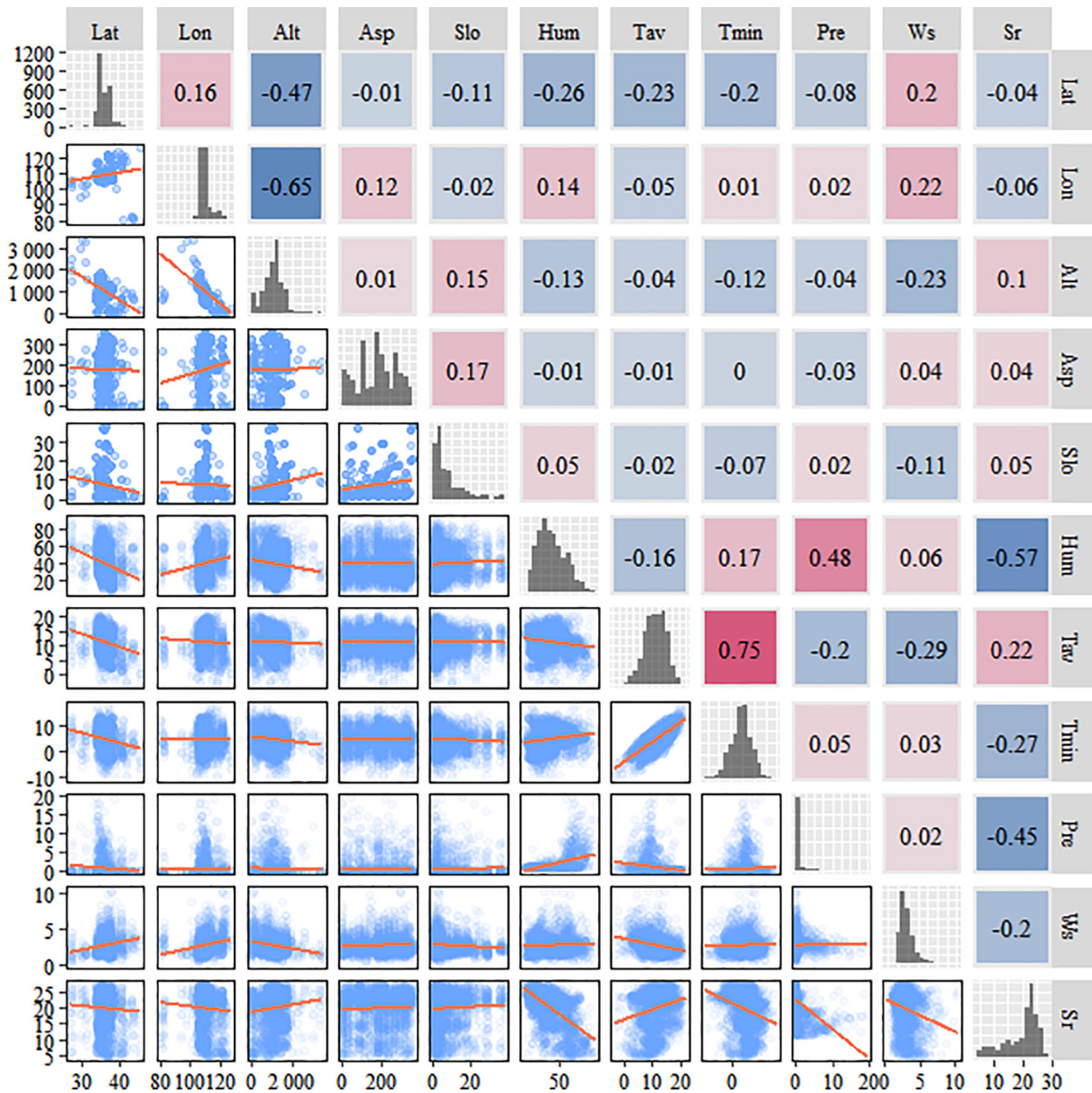


Fig. 5. Pearson correlation coefficients of the eleven candidate input variables for the construction of machine learning models for the predictions of apple late-spring frost events. The upper triangle represents the Pearson correlation coefficients (r) between different input variables; the lower triangle shows the scatter plots for each pair of variables; and the diagonal displays the histograms of the variables. Redder color indicates the stronger positive correlation between the variables; and bluer color the stronger negative correlation.

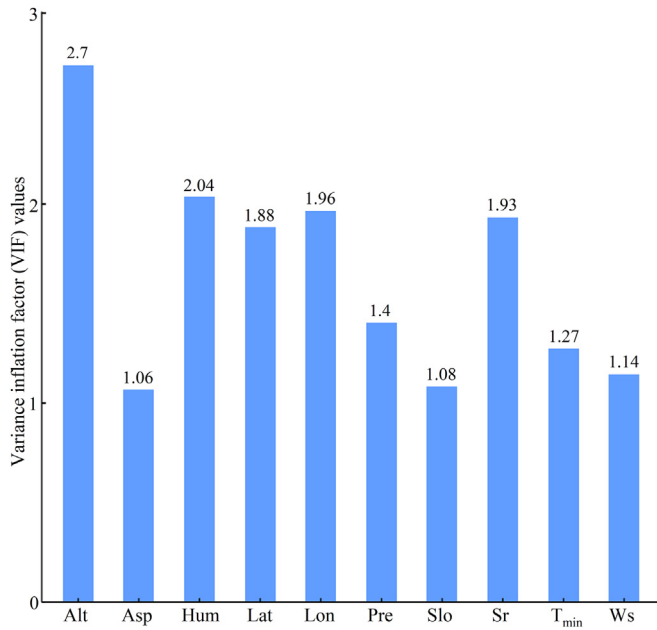


Fig. 6. Variance inflation factor (VIF) between the ten selected input variables (i.e., *Lat*, *Lon*, *Alt*, *Asp*, *Slo*, *Hum*, *T_{min}*, *Pre*, *Ws*, and *Sr*) for the construction of machine learning models for the predictions of apple late-spring frost events.

The variable independence could improve the predictive accuracy and stability of the models, as well as enhance our confidence in the selected model input variables.

3.2. Overall performance of different machine learning models

3.2.1. Influences of input variable combinations on model prediction

Based on the performance of five different machine learning models (i.e., RF, SVM, XGBoost, MLP, and LSTM-CAM) under different combinations of input variables (namely, the combination of only meteorological variables, the combination of meteorological and geographic variables (i.e., *Lat*, *Lon*, *Alt*, *Asp*, and *Slo*) with the commonly selected meteorological variables (i.e., *Hum*, *T_{min}*, *Pre*, *Ws*, and *Sr*) could obviously improve the prediction accuracy of apple late-spring frost events (Table 2). The performance of all five models were sensitive to the geographic input variables. Compared with the model performance under the combination of only meteorological variables, the overall accuracies of RF (OA = 0.900), SVM (OA = 0.882), XGBoost (OA = 0.900), MLP (OA = 0.834), and LSTM-CAM (OA = 0.895) models all remarkably increased with the introduction of geographic input variables.

Table 2

Values of TP (true positives), TN (true negatives), FP (false positives), FN (false negatives), TPR (true positive rate), FPR (false positive rate), OA (overall accuracy), and AUC (the area under the receiver operating characteristic curve) of the machine learning models (i.e., RF, SVM, XGBoost, MLP, LSTM and LSTM-CAM) for the predictions of apple late-spring frost events under different combinations of input variables. The best results are marked in bold.

Input variable combination	Model	TP	TN	FP	FN	TPR (%)	FPR (%)	OA (%)	AUC (%)
Combination of only meteorological factors	RF	152	622	72	34	0.817	0.104	0.880	0.930
	SVM	143	623	81	33	0.813	0.115	0.870	0.900
	XGBoost	154	619	37	70	0.688	0.056	0.878	0.910
	MLP	93	626	25	136	0.406	0.038	0.817	0.850
	LSTM	141	617	34	88	0.616	0.052	0.861	0.900
Combination of both meteorological and geographic factors	RF	149	642	75	14	0.914	0.105	0.900	0.960
	SVM	141	635	83	21	0.870	0.116	0.882	0.930
	XGBoost	152	639	17	72	0.679	0.026	0.900	0.950
	MLP	113	621	30	116	0.493	0.046	0.834	0.860
	LSTM-CAM	175	613	38	54	0.764	0.058	0.895	0.950

According to the values of TPR (True positive rate; the ratio of correctly predicted frost occurrences; Eq. (1)) and FPR (False positive rate; the ratio of incorrectly predicted non-frost occurrences; Eq. (2)), it was found that the combination of meteorological and geographic variables could enhance the models' ability to predict actual frost events (Table 2). The TPR values under the combination of meteorological and geographic factors reached 0.914 for the RF model, 0.870 for the SVM model, and 0.764 for the LSTM-CAM model. Additionally, the introduction of geographic variables also reduced the rate of incorrect predictions of frost events, since the FPR value decreased to 0.026 for the XGBoost model. In general, the geographic variables could dramatically influence the predictions of apple late-spring frost events.

To further evaluate the robustness and generalization ability of the developed machine learning models under the final input variable combination (i.e., the integration of meteorological and geographic factors), their predictive performances on both the training and testing datasets were further compared (Table A.3). As expected, all models achieved better predictive performance on the training dataset than on the testing dataset. Nevertheless, the relatively high performance maintained on the independent testing dataset indicates good generalization ability and robust predictive capability of the developed models.

In addition, the Synthetic Minority Over-sampling Technique (SMOTE) was implemented during the training phase to mitigate data imbalance problem, and prevent classifiers from collapsing toward the majority non-frost class. Notably, the SMOTE-balanced results are highly consistent with the original analysis, with the relative performance rankings of different machine learning models remaining unchanged (Table A.4). The sustained high TPR values across optimized tree ensembles and attention-based deep models.

3.2.2. Performance evaluation of the five machine learning models

Either under the combination of only meteorological variables or under the combination of both meteorological and geographic variables, the RF model all performed the best, with the highest OA and AUC values of 0.90 and 0.96, respectively (Fig. 7). Each point on the ROC curve represents a pair of sensitivity (true positive rate) and specificity (true negative rate) corresponding to a given decision threshold. Generally, the ROC curves closer to the left and top edges indicated higher TPR (True positive rates) and lower FPR (False positive rates) values, signifying better model performance. In this study, a value of 0.5 was selected as the specific decision probability threshold. Thus, a probability greater than 0.5 indicated the occurrence late-spring frost during apple flowering, otherwise no frost event occurrence. An AUC value closer to 1.0 indicated better model performance and the model's ability to distinguish between positive and negative samples. Under the combination of both meteorological and geographic variables, the high TPR (0.914) value and relatively low FPR (0.105) value of the RF model

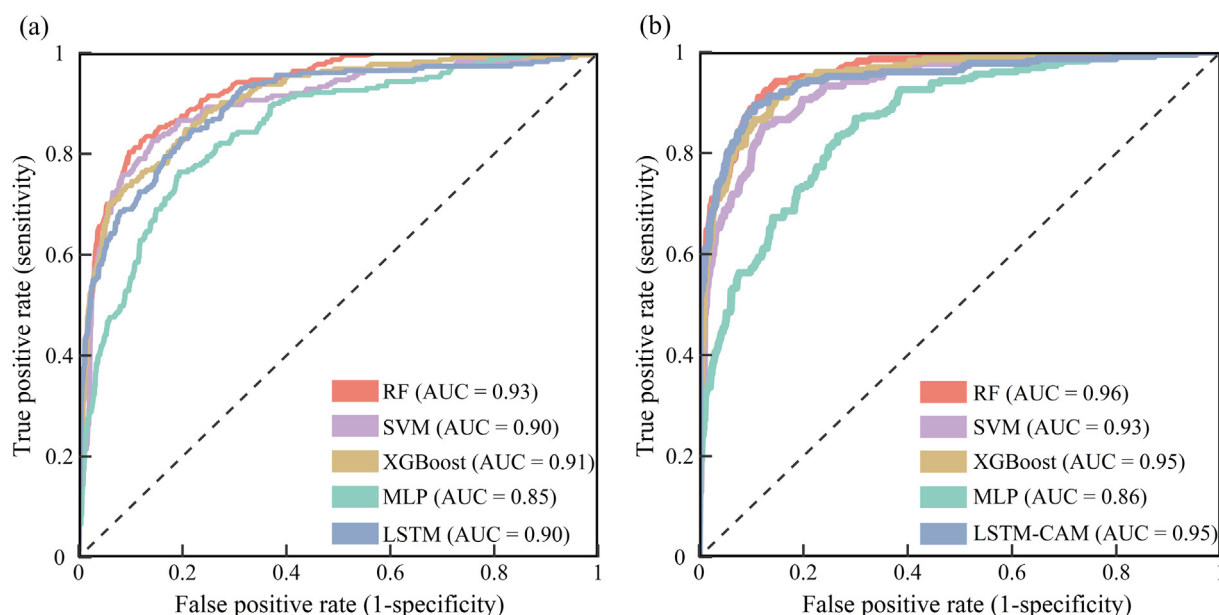


Fig. 7. ROC curves and AUC values of different machine learning models (i.e., RF, SVM, XGBoost, MLP, LSTM and LSTM-CAM) for the predictions of apple late-spring frost events under the combinations of input variables of only meteorological variables (a) and both meteorological and geographical variables (b).

(Table 2 and Fig. 7) demonstrated its superiority in accurately predicting the apple late-spring frost events, with fewer instances of incorrect predictions. The combination of both meteorological and geographic variables could obviously increase the TPR from 0.813 to 0.870 for the SVM model, indicating that the introduction of geographic variables could help the model better distinguish between frost and non-frost events. Similarly, although the XGBoost model (TPR = 0.679) was not as good as the RF model (TPR = 0.914) in identifying true frost events, it had the lowest error rate for non-frost event prediction (FPR = 0.026), resulting in a final AUC of 0.95 and better overall performance.

Either under the combination of only meteorological variables or under the combination of both meteorological and geographic variables, the MLP model was only slightly inferior to the other four models in distinguishing frost and non-frost events, with an AUC of 0.86. The LSTM-CAM (LSTM model with the Channel Attention Mechanism) model had the highest TP (True positives) value of 175, showing its excellent performance in identifying true frost events. The performance of LSTM-CAM model was notably improved compared to the original LSTM model, as its OA value increased from 0.861 to 0.895, TPR from 0.616 to 0.764, and AUC to 0.95 (Fig. 7). This illustrated showed that the attention mechanism could help better capture the occurrence of frost events. In general, under the combination of both meteorological and geographic variables, the general performance ranking of the five machine learning models for the predictions of apple late-spring frost events was: RF > XGBoost = LSTM-CAM > SVM > MLP.

Furthermore, this study further introduced two historical late-spring frost events in 2012 and 2013 as an out-of-sample validation dataset that was completely independent of the model training period, in order to evaluate the extrapolation capability of the pretrained models across different years (Leave-Unseen-Years-Out Validation). During this validation process, no model retraining or parameter tuning was performed. In contrast to the nationwide large-scale frost events in 2018 and 2020, the frost processes in 2012 and 2013 were mainly characterized by localized distributions, primarily occurring in major apple-producing counties such as Linyi and Wanrong in Shanxi Province, Baishui and Luochuan in Shaanxi Province, and Jingning in Gansu Province.

The results show that all models were able to identify the occurrence of frost events in 2012 and 2013, although the overall performance decreased compared with the main test datasets (Table A.5). In the 2012 dataset, the RF model still achieved the highest overall accuracy (OA = 0.93), indicating strong capability in correctly identifying non-frost samples. However, its relatively low AUC (0.71) and TPR (0.33) indicate limited ability to detect frost events under conditions of localized late frost events. In contrast, the LSTM-CAM model exhibited more balanced performance, with an AUC of 0.81 and an increased TPR of 0.67, demonstrating stronger sensitivity to localized frost events. Similar patterns were also observed in the 2013 validation. The RF model still maintained a relatively high OA (0.90), while the LSTM-CAM model achieved the highest AUC (0.84) and TPR (0.70), showing the best overall discriminative performance. The SVM and XGBoost models performed at an intermediate level, whereas the MLP model showed comparatively weaker performance in both years of out-of-sample validation.

Overall, the lower TPR values compared with OA indicate that model performance was affected by the scarcity of positive frost samples. Nevertheless, the relatively stable performance across fully independent years still suggests that the models possess a certain degree of cross-year generalization capability. In particular, the LSTM-CAM model showed stronger stability in identifying localized frost events, indicating that attention mechanisms may have potential advantages in modeling sparse extreme disaster events under temporal transfer conditions.

3.3. Global and local SHAP explanations for model performance

The SHAP technique was applied to conduct interpretability analysis on the prediction results of the five machine learning models (i.e., RF, SVM, XGBoost, MLP, and LSTM-CAM) under the input combination of both meteorological and geographic variables. The SHAP values were calculated in the test process to assess the contribution of each input variable to the model prediction results. The magnitude, prevalence, and direction of the effects of input variables were succinctly illustrated in the SHAP summary plots (Fig. 8). The standard bar charts of average absolute SHAP values reflected the global importance of each prediction factor (Fig. 8). For the RF model, the global feature importance

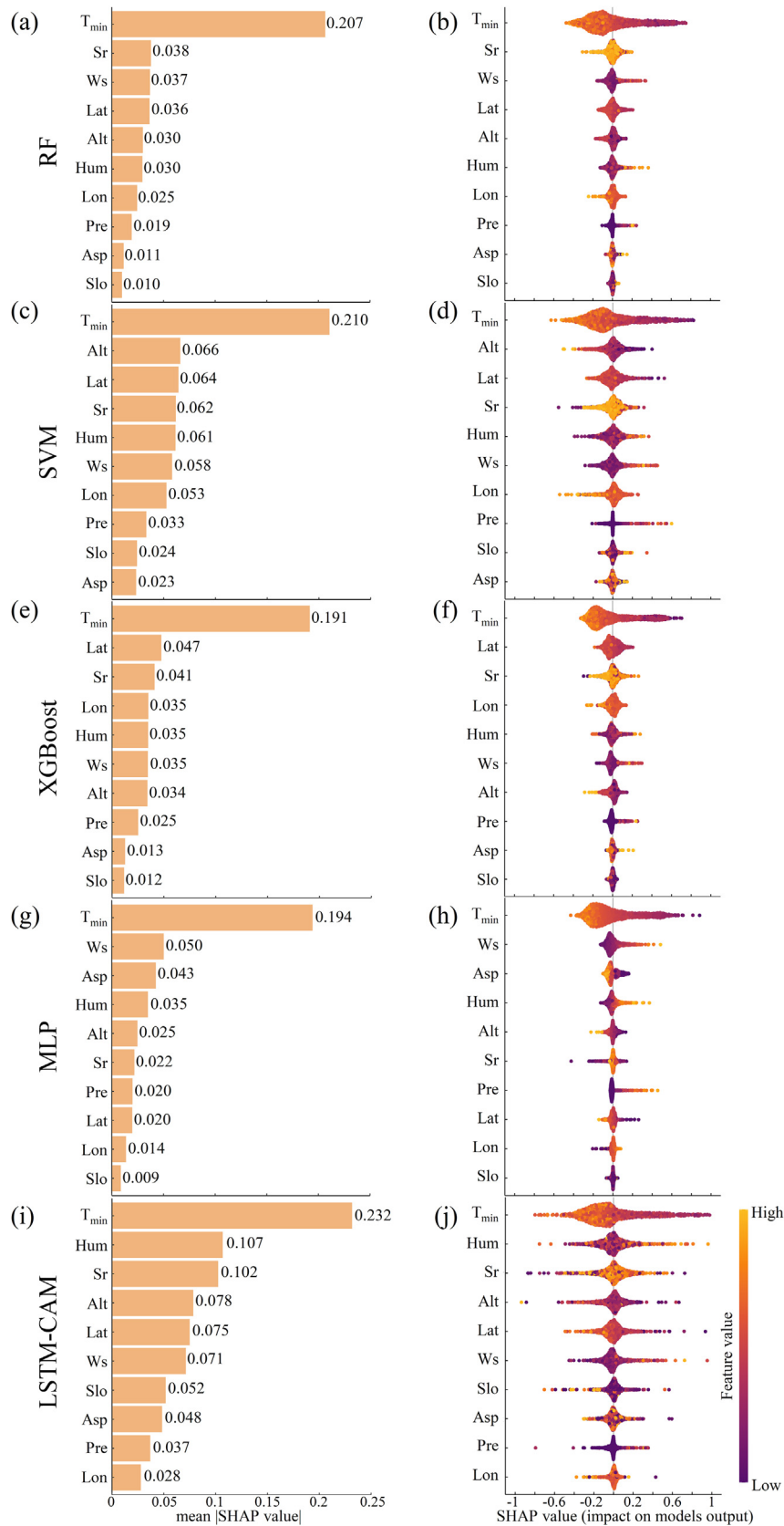


Fig. 8. SHAP summary plots for the machine learning models (i.e., RF, SVM, XGBoost, MLP, and LSTM-CAM) for the predictions of apple late-spring frost events under the input combination of both meteorological and geographic variables. Subplots (a), (c), (e), (g), (i) present the |SHAP value| for global feature importance; and (b), (d), (f), (h), (j) the SHAP value for local explanations summary plots. In subplots (b), (d), (f), (h), (j), the color gradient from purple to orange represents the increasing amplitude of the input variables. The X-axis represents the SHAP values, reflecting the impact of each input variable on the predictions of apple late-spring frost events. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ranking based on average SHAP values was: $T_{min} > Sr > Ws > Lat > Alt > Lon > Pre > Asp > Slo$ (Fig. 8a). For the SVM model, the global feature importance ranking was: $T_{min} > Alt > Lat > Sr > Hum > Ws > Lon > Pre > Slo > Asp$ (Fig. 8c). Similarly, for the XGBoost model, the ranking was: $T_{min} > Lat > Sr > Lon > Hum > Ws > Alt > Pre > Asp > Slo$ (Fig. 8e). For the MLP model, the ranking was: $T_{min} > Ws > Asp > Hum > Alt > Sr > Pre > Lat > Lon > Slo$ (Fig. 8g). And for the LSTM-CAM model, the ranking was: $T_{min} > Sr > Hum > Alt > Lat > Ws > Slo > Asp > Pre > Lon$ (Fig. 8i). The rankings revealed that the meteorological variable of T_{min} had the greatest contribution to the predictions of apple late-spring frost events despite of the models, while Slo and Asp had relatively lower contributions than the other factors. The meteorological variables generally contributed more than the geographic variables. In other words, the geographic variables had less influence on the occurrence of apple late-spring frost events, which was also illustrated by the concentrated SHAP value distributions (Fig. 8f, h).

Each point in the beeswarm plots corresponds to the SHAP value of a given variable for a single sample (Fig. 8), thereby reflecting its local contribution to the model output. The horizontal position of each point indicates the direction and magnitude of the SHAP value, while the vertical dispersion represents the distribution of variable effects across all samples. By aggregating SHAP values over the entire dataset, the overall width and tail length of each variable's distribution summarize its global importance, with longer tails indicating a greater overall influence on model predictions.

According to the beeswarm plots (Fig. 8b, d, f, h, j), larger SHAP values indicate stronger positive influences on the predictive probability of apple late-spring frost events, whereas smaller (negative) values represent suppressing effects. The color gradient from purple to orange reflects increasing magnitudes of the predictor values. In the RF model (Fig. 8b), lower T_{min} values clustered around higher frost probabilities, and the extended right tail of SHAP distribution illustrated how increasingly low temperatures substantially amplify frost risk. This long-tailed behavior highlights both the dominant global importance of T_{min} and its strong local impacts on individual samples. Other meteorological variables, such as Sr , Hum , and Ws , also exert positive effects on frost predictions, although their contributions were smaller than T_{min} .

The long tails also showed that predictors with relatively low global importance may still be important under specific conditions (Lundberg et al., 2020), namely the predictors with modest overall contributions may still be influential under particular environmental scenarios. For example, Hum and Ws had lower impacts on the predictions with the XGBoost model than the other variables, which was apparent in the beeswarm clustering around SHAP values of zero (Fig. 8f). But both variables had tails in two directions of their clusters, which indicated that they could be influential on frost-event predictions in some circumstances. A similar pattern appeared in the MLP model (Fig. 8h), where Hum and Pre exhibited low mean contributions but extended local tails, also indicating the importance under certain meteorological conditions.

More importantly, the SHAP results also highlighted the importance of specific predictors as well as particular combinations of meteorological and geographic factors in shaping frost-event probabilities (Fig. 8). Although geographic factors generally ranked lower in global importance across models, several exhibited meaningful local effects. For instance, lower Lat values exerted stronger positive impacts on frost predictions than Alt in the SVM model (Fig. 8b). In the LSTM-CAM model (Fig. 8j), variations in Slo and Asp occasionally generated substantial local SHAP responses. Across all models, T_{min} consistently emerged as the most influential predictor for frost-event occurrence predictions, especially under the input combination of both meteorological and geographic variables. Geographic predictors played a smaller global role but became important in combination with the meteorological variables, especially under the conditions prone to cold-air pooling. These findings

elucidated that the joint effects of meteorological and geographic factors were essential for accurately characterizing the occurrence of frost events across the four major apple-growing regions in China.

3.4. Reproduction of the severe apple late-spring frost disaster in 2018

We further evaluated the abilities of the machine learning models developed in this study to simulate the severe late-spring frost disaster during apple flowering in the year of 2018 in China, with a particular focus on the spatial consistency between the predicted and observed late-spring frost events across the four major apple-growing regions. At the national scale, all machine learning models could successfully reproduce the general spatial pattern of late-spring frost in 2018. The regions characterized by observed severe frost damages, particularly across the LP (Loess Plateau) and parts of the BB (Bohai Bay) region, generally coincided with the areas with high predicted frost probabilities.

Specifically, the RF model exhibited the highest spatial coincidence with the observed rates of frost damage on apple central-flowers (Fig. 9a). High-probability frost zones simulated by the RF model overlapped spatially with the sites with observed severe frost damages in the LP and BB regions. At the same time, low probabilities were predicted across most areas of the XJ and SCH regions, where the late-spring frost impacts were limited in the year of 2018. This strong spatial agreement suggested that the RF model effectively integrated the meteorological drivers with spatial geographical constraints to reproduce the actual pattern of the late-spring frost in 2018. The SVM and XGBoost models could capture the general extent of the regions affected by the severe frost damage (Fig. 9b, c). However, their simulations of the high-risk areas appeared more spatially fragmented. High frost probabilities were also simulated by these two models in the regions where the observed late-spring frost occurrences were sparse (e.g., parts of Liaoning Province), which suggested that these two models overpredicted the late-spring frost in some parts of China. The MLP model provided a more refined spatial pattern for the late-spring frost in 2018 than the SVM and XGBoost models (Fig. 9d). In particular, simulated high-probability zones in the LP region aligned more closely with the sites with observed apple frost damages. Nevertheless, despite relatively mild frost damage observed in the SCH region, the MLP model simulated comparatively high probabilities of frost occurrence, indicating a localized overestimation of frost risk in this region. Among all models, the LSTM-CAM model showed a notably enhanced spatial coincidence between the simulated frost probabilities and the observed late-spring frost patterns in 2018 (Fig. 9e). By introducing the channel attention mechanism, the LSTM-CAM model was able to effectively identify the severe frost-affected areas in the LP and BB regions, as well as the mild frost conditions in the XJ region. In contrast, the other models consistently overestimated the occurrence probabilities of late-spring frost in XJ in 2018. This comparison highlighted the advantage of the attention mechanisms in emphasizing the dominant signals and suppressing the less relevant information in the predictions of late-spring frost with the machine learning models across China in 2018.

In addition, quantitative evaluation metrics were employed to objectively benchmark the spatial prediction performance of different classifiers (Table A.6). The tree-based ensemble models demonstrated outstanding overall spatial alignment. The XGBoost model achieved the highest spatial overall accuracy (OA = 0.912), F1 (0.879), and Kappa coefficient (0.810). The RF model yielding the highest geographic discrimination capacity with an AUC of 0.969 and a balanced Kappa of 0.780. In comparison, while the neural network architectures (MLP and LSTM-CAM) excelled in visual noise suppression across certain macro-regions, they exhibited relatively lower F1 and Kappa coefficients at the specific point-station scale. These quantitative validation results further support the spatial agreement shown in Fig. 9.

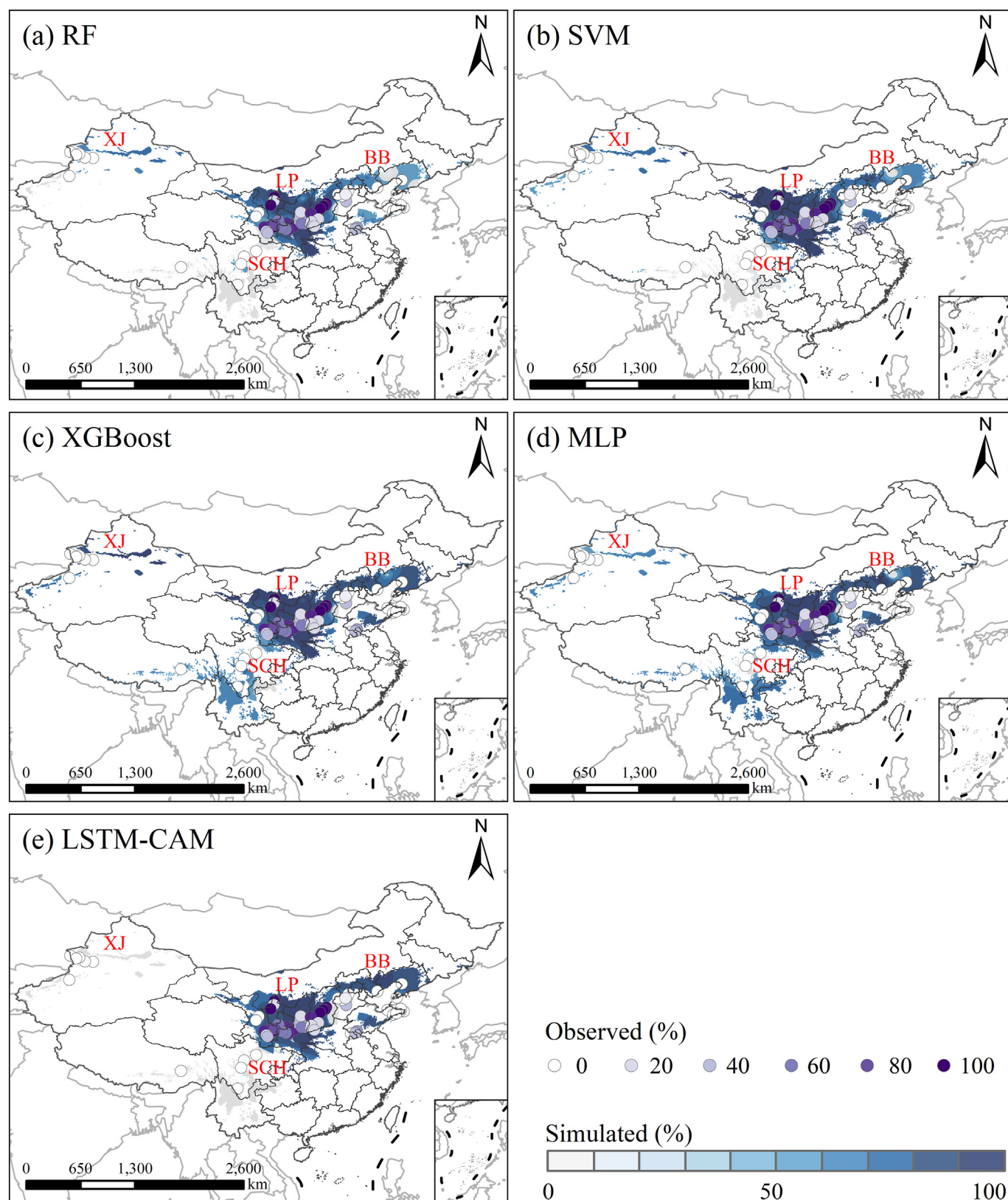


Fig. 9. Spatial distributions of late-spring frost disaster events during apple flowering period in 2018 across the four main apple-growing regions in China. The blue shaded areas indicate the simulated probabilities of late-spring frost occurrence (%) through machine learning models (RF, a; SVM, b; XGBoost, c; MLP, d; LSTM-CAM, e). The dots (from white to purple) denote the sites where different rates (0%–100%) of frost damage on apple central-flowers were observed during the apple late-spring frost disaster in 2018. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

4.1. Performance of different machine learning models in predicting apple late-spring frost events

Performance of machine learning models may vary greatly, depending on which input variables they considered most explanatory in predicting the likelihood of apple flowering frost risks. In other words, whether the input variables are useful or not depends on the model application purpose. The overall accuracy and TPR (True positive rate) were dramatically improved with the inclusion of spatial geographic variables, compared with the models under the input combination of only meteorological variables. The RF and XGBoost models achieved overall accuracy of 90%, and the accuracy of the LSTM-CAM model also rose to 89.5%. This boosted the models' ability to identify apple flowering frost events and improved the model prediction performance.

The Random Forest (RF) model generally achieved the highest prediction accuracy and showed its potential in predicting frost events. This result was consistent with the study by Diniz et al. (2021), where the RF model proved to be the most accurate classifier to predict frost risks in artificial forests in southern Brazil compared with the SVM and MLP models. Eccel et al. (2007) found that the RF model was more effective in predicting the frost risks of the minimum temperatures in Italian Alps. Diedrichs et al. (2018) also reported the enhanced performance of the RF model in predicting frost events, with excellent prediction results in the plateau region of Mendoza Province in Argentina. The superior performance of the RF model among the machine learning models investigated in this study for apple late-spring frost prediction, could be attributed to its characteristics of ensemble learning algorithms based on decision trees and their ability to handle complex datasets (Breiman, 2001). The RF model builds multiple decision trees and aggregates their predictions through averaging or majority voting to obtain the final prediction. This approach dramatically improves the model's generalization ability by introducing randomness, reducing overfitting, and enhancing overall predictive performance.

The occurrence of frost events is influenced by various meteorological and spatial geographic factors. The RF model driven by the combination of meteorological and spatial geographic variables could effectively capture the complex relationships beneath the occurrence of frost events, especially the interactions among different features. The SVM model finds the optimal hyperplane that separates data points of different classes at the maximum possible margin. This hyperplane is identified through an optimization algorithm, ensuring that any new data point, when classified, is positioned as far as possible from the decision boundary, thereby improving the accuracy of classification (Hearst et al., 1998). Under the combination of meteorological and spatial geographic variables, the SVM model could better separate frost and non-frost events and thus improve the TPR (0.870) (Table 2). Spatial geographic variables provided additional information to allow the SVM model to more accurately capture the characteristics of frost event occurrence, although the improvement of TPR and the FPR values were relatively small if compared with the predictions when only meteorological variables were used as input variables. This indicated that the SVM model performed good in reducing the false alarm rate of frost events.

The XGBoost model is an ensemble method based on gradient boosting integrated approach that optimizes the model by progressively building weak classifiers (usually decision trees). It is robust to outliers and noise and can effectively reduce the false alarm rate. In this study, the XGBoost model had the lowest false alarm rate (FPR = 0.026) in predicting the non-frost events, but slightly inadequate in identifying the frost events (TPR = 0.679). This was probably due to the fact that the XGBoost model focuses more on reducing the overall error during the optimization process (Chen and Guestrin, 2016).

The MLP model is a static feedforward neural network that cannot directly handle the temporal dependencies in time series data, whereas frost event prediction often involves temporal occurrence dependencies. This limits the prediction capability of MLP model. Additionally, the performance of MLP model could be affected by the volume and complexity of data samples. Neural network operation depends on the complex weight relationships between inputs and outputs, which impact the classification performance (Diniz et al., 2021; Fuentes et al., 2018; Hosseini et al., 2019). Thus, both the output (frost event occurrence or non-occurrence) and the input variables could affect the classification performance of neural networks. In this study, compared with the other machine learning models, the MLP struggled to capture the complex relationships between the meteorological and spatial geographic variables and had poorer prediction performance.

In summary, the RF model performed the best due to its strong ability to capture the interactions between the input variables. The SVM and XGBoost models both had improved prediction accuracy when under the combination of meteorological and geographic variables. The SVM model dramatically improved the ability of frost event identification, while the XGBoost model excelled in reducing the false alarm rates. The MLP model generally performed poor due to inherent structural limitations.

4.2. Ability of channel attention mechanism (CAM) to enhance the prediction performance of the LSTM model

In this study, the channel attention mechanism (CAM) was introduced into the authentic LSTM model, which could dynamically adjust the importance of different features. By incorporating the spatial geographic and meteorological factors as two separate CAM layers into the hidden layer of LSTM model, the LSTM model was able to deeply focus on the key features of the input variables and thus more accurately capture the frost event characteristics. Interestingly, the LSTM-CAM model demonstrated strong capability in frost event predictions (OA = 0.895 and TPR = 0.764), although the FPR value increased. This prediction performance improvement could be attributed to the network structure of the CAM itself. By weighting the contribution of each channel, the model could focus on the key features. In this study, we utilized two separate CAM layers of input variables instead of simply "concatenating" the input variables. Thus, the LSTM-CAM model could automatically learn and adjust which feature channels were the most valuable for prediction. In general, by introducing the channel attention mechanism, the LSTM-CAM model could dramatically improve the prediction capability for apple late-spring frost events.

4.3. Analysis of the driving factors for apple late-spring frost events based on the SHAP theory

The results of SHAP (Shapley Additive exPlanations) analysis highlighted the importance of combining meteorological and geographic variables in the predictions of apple late-spring frost events. The meteorological variable of T_{min} was the most important predictor when under the input combination of meteorological and geographic variables regardless of the machine learning models (Fig. 8). Apple flowering mainly occurs in early spring from March to April. During this spring transition period, temperature fluctuations are usually serious. The temperatures were above 0 °C in daytime, but might drop below 0 °C at night or in early morning. Under low-temperature conditions, water vapor in the air was more likely to condense and form frost. When T_{min} reached a certain intensity and persisted for a certain period of time, frost event might occur. Levels of frost damage to apple flowers depend on the duration of low temperatures. The longer the low temperature lasts, the greater the apple frost injury will be. Similarly, Ding et al. (2019) also reported that the minimum temperature is the most important factor in the frost risk forecasting based on machine learning models.

The meteorological variable of *Sr* (Solar radiation) was also a major contributor in the prediction models of apple late-spring frost events, ranking second in importance in the RF and LSTM-CAM models. During daytime, solar radiation heats the surface, causing soil surface temperature to rise and store heat. However, solar radiation disappears after sunset and at night and soil surface temperature gradually decreases due to radiative heat loss. Frost event occurs more easily when the soil surface loses heat excessively and the atmospheric temperature drops rapidly. Gu et al. (2023) noted that solar radiation not only reflects the changes in daylight length (or photoperiod) but also affects complex weather conditions (such as clouds and snows), leading to more complex microclimate changes. Daylight length (or photoperiod) is an indicator of the total accumulated sunlight, which is directly linked to the active growing period of plants after leaf emergence and serves as a pivotal environmental trigger for early or delayed spring phenology. On the other hand, earlier leaf emergence due to climate warming and shorter photoperiods may reduce the temperature sensitivity of spring phenology, and thus prevent plants from emerging too early and avoid frost damage (Flynn and Wolkovich, 2018; Piao et al., 2019). This situation was confirmed by the results in this study (Fig. 8b, d), where high solar radiation had a negative contribution to frost events or could reduce frost risk, while low solar radiation had a positive contribution or could increase frost risk.

The meteorological variable of *Ws* (Wind speed) also had remarkable impacts on the RF values of the apple frost prediction model. Although wind speed ranked lower in importance in the SVM, XGBoost, and LSTM-CAM models under the input combination of meteorological and geographic variables, high wind speeds could help reduce frost risk (Fig. 8b, d). *Ws* influenced frost occurrence by affecting the heat exchange and mixing in the atmosphere. When *Ws* was high, there was increased heat exchange between atmospheric layers and warmer air aloft mixed with cooler air at the surface, which could suppress the rapid drop of soil surface temperatures. Strong winds could disrupt the stagnation of cold air and promote the downward transport of heat, thereby to some extent preventing soil surface temperatures from falling below 0 °C. Additionally, *Ws* affects the accumulation of cold air at soil surface. Under the conditions of low wind speeds, cold air tended to accumulate at soil surface, especially in valleys or low-lying areas where frost was more likely to form. Although surface radiative cooling occurred more rapidly on clear and calm nights, the absence of wind would cause cold air to pool together and greatly increase the likelihood of frost formation. Thus, farmers often used large fans or wind machines to enhance air circulation to prevent the formation of cold air pockets around fruit trees.

The variable of *Lat* (Latitude) was the most crucial geographic factor for the RF model as the overall accuracy of this model was the highest. Although the RF and XGBoost models achieved comparable overall accuracy, the relative importance of geographic factors differed among models. In particular, *Lat* played a less prominent role than *Alt* in the SVM and LSTM-CAM models, suggesting model-specific sensitivities to spatial controls on frost occurrence. Previous studies have similarly highlighted the role of large-scale geographic gradients in frost risk. For example, Simões et al. (2015) reported that latitude and continentality as key macro-environmental factors governing frost occurrence in southern Brazil, where vineyards located at higher latitudes experience lower temperatures and were therefore more prone to frost risks. And according to the study by Allevato et al. (2019), the severity of frost in beech tree crowns decreased from the southern to the northern Apennine Mountains.

The variable of *Alt* (Altitude) ranked second in importance among the geographic variables in the RF model and also the second in the SVM model. The timing of tree spring phenology varied across different altitudes, with most tree species exhibiting earlier spring phenology at lower altitudes. Čufar et al. (2012) found that the trend of leaf unfolding was closely related to altitude. Regions with a greater advance in spring phenology were more susceptible to apple frost risk (Ru et al., 2023a).

This was because an earlier onset of phenology was more likely to trigger frost events, which would coincide with the sensitive period from bud breaking to fruit setting, ultimately increasing frost risk (Chen et al., 2024; Hoffmann and Rath, 2013; Parker et al., 2021).

The SHAP value distributions for the geographic variables of *Slo* (Slope) and *Asp* (Aspect) were relatively concentrated, and their importance in predicting frost events was less significant than *Lat* (latitude) and *Alt* (altitude). Slope and aspect are local terrain features and their impacts on frost event occurrence are also local. Although they could influence the flow and accumulation of cold air (e.g., cold air lingering longer in depressions or downhill areas), their influences were limited in specific small areas. Thus, in large-scale frost event predictions, their effects may be overshadowed by other factors such as latitude and elevation.

4.4. Reproduction of the severe apple late-spring frost disaster in 2018

There was general spatial consistency between the simulated occurrence probabilities of late-spring frost damage and the observed rates of frost damage on central apple flowers in 2018, which indicated that the machine learning models developed in this study were able to reproduce the geographical extents of the areas vulnerable to late-spring frost disaster. As shown in Fig. 9, the model simulations clearly identified the main frost-prone zones in the LP and BB apple-growing regions, suggesting that the models effectively learned the key environmental conditions associated with late-spring frost occurrence during apple flowering period.

The RF model exhibited the highest spatial agreement between the simulated late-spring frost damage probabilities and the observed frost damage distributions. When the meteorological variables and the spatial geographic variables were combined, this tree-based ensemble model demonstrated greater reliability in delineating the frost risk zones. By integrating nonlinear meteorological controls with spatial constraints, the RF model maintained relatively high spatial stability and consequently good generalization performance across different apple-growing regions. This finding was highly consistent with the conclusions by Diniz et al. (2021), who reported that the RF model achieved the highest accuracy among the tested classifiers and outperformed the other machine learning models in capturing the spatial heterogeneity of the frost-environment relationships, particularly in regions with complex terrains.

In contrast, the SVM and XGBoost models tended to provide more fragmented and spatially dispersed predictions of high frost probability, with some predicted high-risk areas extending into the regions with only mild frost damage or no frost occurrence. Although these models could achieve acceptable performance according to the overall classification accuracy, their spatial verification remained limited when applied to the landscapes characterized by pronounced topographic variability and strong climate gradients. Similar problem was noticed in some previous studies, where the SVM and XGBoost models exhibited remarkable spatial generalization biases under complex landscape conditions due to the insufficient integration of topographic and regional climatic constraints (Diniz et al., 2021).

The MLP model was able to generate relatively fine spatial patterns in core frost-prone regions (e.g., the LP region), but local overestimations persisted along the regional margins (e.g., parts of the SCH region). This indicated that although the feed-forward nonlinear mappings could capture the complex interactions among predictors, they did not inherently guarantee spatial consistency in predicted risk patterns. The LSTM model incorporating a CAM (Channel attention mechanism) could remarkably improve the predictions of spatial realism of late-spring frost in 2018. The CAM enhanced the contributions of key feature channels most relevant to frost formation and suppressed weakly related or noisy signals, thereby effectively reduced the spurious high-probability facts, particularly in the regions with weak observed frost impacts (e.g., the XJ region). This result was mechanistically consistent

with the findings by Kirchhof et al. (2025), who emphasized that systematic biases in model-based frost risk assessments often arose from insufficient consideration of the microclimatic factors (such as humidity, wind speed, cloud cover, topography, and plant height), which directly affected the translatability between the empirically measured cold-hardiness thresholds and the critical field-level damage temperatures. By strengthening feature selection and weighting through the attention mechanisms, the LSTM-CAM model could partially mitigate the spatial prediction biases induced by microclimatic complexity and thereby reduce the false frost risk signals.

Despite the overall agreements, discrepancies between the simulated frost probabilities and the observed late-spring frost disaster patterns remained in the year of 2018 due to multiple reasons. First, station-based late-spring frost damage observations were spatially heterogeneous and the observed rates of frost damage on apple central-flowers represented site-specific point-scale samples of frost impacts, which limited their capacity to fully capture the broad regional distribution of late-spring frost disaster in 2018. Omazić et al. (2024) highlighted the inherent limitations of traditional point-based frost measurements, particularly with respect to the spatial resolution and coverage in heterogeneous landscapes, which constrained the accurate characterization of frost risk at regional scales. Next, regional differences in climatic conditions, phenological periods, and cultivar managements among the XJ, LP, BB, and SCH regions imply that the identical frost thresholds might correspond to substantially different levels of apple frost damage. Frost thresholds in field damage observations were often based on air temperatures, which might differ from the actual internal temperature at which tissues froze (Litschmann and Strěda, 2019). Kirchhof et al. (2025) further demonstrated that frost risk thresholds were not universally applicable and must be calibrated to regional phenological and climatic contexts to avoid misleading predictions under climate change. Thus, appropriate selection of frost thresholds is critical to the model predictive reliability.

Differences between simulated frost probabilities and site observed frost damage should reflect not only the intrinsic model limitations but also the mismatches between the observation processes and the physical mechanisms of frost formation, particularly under radiative frost conditions where near-surface or tissue-level temperatures could be substantially lower than standard air temperature measurements at 2 m height. Generally, traditional site-based observations and generalized frost thresholds were probably the major uncertainty sources in reproducing the actual patterns of apple late-spring frost disaster in 2018. At regional scale, single temperature threshold and sparse station observations were insufficient to fully characterize the actual processes of apple frost damage. Thus, it shows strong potential for spatial reproduction and early warning of severe late-spring frost disasters by integrating multi-source information into the machine learning models. Finally, region-specific temperature thresholds for frost risks should be established and calibrated for different apple-growing regions to improve the predictive reliability of apple late-spring frost risks with the machine learning models.

4.5. Limitations and future challenges

Despite the satisfactory performance the proposed explainable machine learning framework for spatial prediction of late-frost in apples, several limitations remain. Apple flowering phenology is considered an important biological factor influencing frost susceptibility, frost damage typically occurs when low-temperature events coincide with spring phenological sensitive developmental stages (flowering and fruit setting) (Ru et al., 2023b). In the study, all frost and non-frost samples were collected during the apple flowering period. The phenological stage was implicitly controlled through the sampling strategy, and the developed models were designed to predict frost damage under flowering-stage conditions rather than throughout the entire growing

season. Nevertheless, explicitly incorporating observed flowering dates or dynamically simulated phenology into the prediction model may better capture the biological timing of frost susceptibility and further improve model performance.

In addition, long-term time series and spatially extensive observational records of apple late-spring frost events are crucial for further improving the temporal generalization ability and spatial robustness of the model; however, such high-quality and continuous disaster observations are still relatively scarce in practice. Existing data are mainly concentrated in a few extreme frost years, which makes it difficult to fully capture the interannual variability of frost events under different climatic conditions. Therefore, future studies should explore the integration of phenological models with machine learning algorithms and incorporate multi-source long-term observational datasets to establish a more biologically informed and generalizable frost-risk prediction framework, thereby improving the robustness and stability of extreme meteorological disaster prediction models.

5. Conclusions

This study developed an explainable machine learning framework to predict the occurrence of late-spring frost events during the apple flowering period across the four major apple-growing regions in China. By integrating multiple meteorological variables with spatial geographic factors and incorporating the SHAP (Shapley Additive Explanations), we developed different machine learning models, systematically evaluated their performance, identified the key driving factors for frost events, and assessed the spatial consistency of model predictions.

First, the results demonstrated that the inclusion of spatial geographic factors could substantially improve the predictive accuracy and robustness of the machine learning models for the predictions of apple late-spring frost risks. Compared with models driven solely by the combination meteorological variables, the incorporation of geographic factors of longitude, latitude, elevation, slope, and aspect consistently enhanced model discrimination and stability, highlighting the importance of geographic variables in regional assessment of apple frost risks.

Next, among all of the models evaluated, the RF (Random forest) model exhibited superior overall prediction performance. When the meteorological variables and spatial geographic features were combined, the tree-based ensemble approach effectively captured the non-linear frost-environment relationships and maintained strong spatial stability and generalization across heterogeneous landscapes in China. This endorsed the RF model as a robust tool for regional-scale assessment of frost risks in complex terrains.

Third, SHAP interpretability analysis revealed that T_{min} was the dominant meteorological driver for late-spring frost occurrence, whereas *Lat* emerged as the most influential geographic factor. Importantly, the SHAP framework provided both global and local interpretations to elucidate how the meteorological and geographic factors jointly influenced the regional patterns of late-spring frost occurrence. These insights substantially enhanced the transparency and explainability of otherwise “black-box” machine learning models.

Fourth, spatial reconstruction of the severe apple late-spring frost disaster in 2018 across the four major apple-growing regions in China provided strong evidence for the reliability of the machine learning models developed in this study. All models could successfully reproduce the extensive frost-prone belts over the LP (Loess Plateau) and BB (Bohai Bay) regions. However, simulation differences were evident in spatial realism. The RF and LSTM-CAM (Long short-term memory neural network-channel attention mechanism) models showed the highest spatial consistency with the observed rates of frost damage on central apple flowers, whereas the SVM (Support vector machine) and XGBoost (Extreme gradient boosting) models tended to generate spatially fragmented patterns and overestimate frost risks in the areas only

affected by mild frost. By introducing a channel attention mechanism (CAM), the LSTM-CAM model could more effectively capture the key drivers of frost occurrence. It improved the overall prediction accuracy by about 3.4% and dramatically reduced the spurious high-risk facts, particularly in the XJ (Xinjiang) region where observed frost impacts were relatively weak. This highlighted the advantage of CAM in emphasizing informative features and suppressing weakly relevant or noisy signals in the predictions of apple late-spring frost risks with the machine learning models.

Overall, this study demonstrated that the explainable machine learning models, particularly the RF model and attention-enhanced deep learning framework, could reliably reproduce the spatial patterns of the severe apple late-spring frost disaster and provide interpretable insights into the key driving factors of frost event occurrence. The proposed framework offered a scalable and transparent tool for regional assessment and early warning of apple late-spring frost risks. It could also facilitate the identification of frost-vulnerable areas across the major apple-producing regions and support scientific decisions to enhance climate resilience in apple production under ongoing climate change in China.

CRedit authorship contribution statement

Jie Zhou: Writing – original draft, Validation, Software, Methodology, Conceptualization. **Xiaoya Ru:** Visualization, Validation. **Zhihao He:** Validation, Conceptualization. **Kaiyuan Gong:** Visualization, Methodology. **Jinghong Wang:** Supervision, Resources. **Zhao Wang:** Supervision, Resources. **Meirong Li:** Resources, Methodology. **Hao Feng:** Funding acquisition, Conceptualization. **Qiang Yu:** Methodology, Conceptualization. **Jianqiang He:** Writing – review & editing, Methodology, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China (No. 52579046), the National Key Research and Development Program of China (2021YFD190070402). We would like to thank Chinese Meteorological Administration (CMA) for providing the meteorological data and Shaanxi Meteorological Service Center of Agricultural Remote Sensing and Economic Crops for providing apple data. The authors acknowledge the anonymous referees for their valuable comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aiaa.2026.08.017>.

References

Allevato, E., Saulino, L., Cesarano, G., Chirico, G.B., D'Urso, G., Falanga Bolognesi, S., Rita, A., Rossi, S., Saracino, A., Bonanomi, G., 2019. Canopy damage by spring frost in European beech along the Apennines: effect of latitude, altitude and aspect. *Remote Sens. Environ.* 225, 431–440. <https://doi.org/10.1016/j.rse.2019.03.023>.

Alvares, C.A., Sentelhas, P.C., Stape, J.L., 2018. Modeling monthly meteorological and agro-nomic frost days, based on minimum air temperature, in Center-Southern Brazil. *Theor. Appl. Clim.* 134, 177–191.

Baker, S.G., 2003. The central role of receiver operating characteristic (ROC) curves in evaluating tests for the early detection of cancer. *JNCI J. Natl. Cancer Inst.* 95 (7), 511–515.

Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* 13, 281–305.

Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B., 2011. Algorithms for hyper-parameter optimization. *Adv. Neural Inf. Process. Syst.* 2546–2554.

Biecek, P., Burzykowski, T., 2021. *Explanatory Model Analysis: Explore, Explain, and Examine Predictive Models*. first ed. Chapman and Hall/CRC, New York.

Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.

Cao, Z., Han, H., Gu, B., Ren, N., 2009. A novel prediction model of frost growth on cold surface based on support vector machine. *Appl. Therm. Eng.* 29 (11–12), 2320–2326. <https://doi.org/10.1016/j.applthermaleng.2008.11.015>.

Carter, J.V., Pan, J., Rai, S.N., Galanduk, S., 2016. ROC-ing along: evaluation and interpretation of receiver operating characteristic curves. *Surgery* 159 (6), 1638–1645. <https://doi.org/10.1016/j.surg.2015.12.029>.

Cebulj, A., Mikulić-Petkovšek, M., Veberič, R., Jakopic, J., 2021. Effect of spring frost damage on apple fruit (*Malus domestica* Borkh.) inner quality at harvest. *Agric* 12 (1), 14. <https://doi.org/10.3390/agriculture12010014>.

Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. *Proceedings of the 22nd acm sigkdd International Conference on Knowledge Discovery and Data Mining*, pp. 785–794.

Chen, R., Wang, J., Li, Y., Bai, R., Huang, M., Zhang, Z., Zhao, L., Qu, Z., Liu, L., 2024. Higher risk of spring frost under future climate change across China's apple planting regions. *Eur. J. Agron.* 159, 127288. <https://doi.org/10.1016/j.eja.2024.127288>.

China Meteorology Administration, 2017. Grade of fluorescence freezing injury to Fuji apple (QX/T 392-2017). Meteorological Industry Standard of the People's Republic of China. https://www.cma.gov.cn/zfxgk/gknr/flfgbz/bz/202209/t20220921_5099023.html.

China Meteorology Administration, 2020. *Yearbook of Meteorological Disasters in China: 2019*. first ed. China Meteorological Press, Beijing.

China Meteorology Administration, 2022. *Yearbook of Meteorological Disasters in China: 2021*. first ed. China Meteorological Press, Beijing.

Čufar, K., De Luis, M., Saz, M.A., Črepinšek, Z., Kajfež-Bogataj, L., 2012. Temporal shifts in leaf phenology of beech (*Fagus sylvatica*) depend on elevation. *Trees* 26, 1091–1100. <https://doi.org/10.1007/s00468-012-0686-7>.

Diedrichs, A.L., Bromberg, F., Dujovne, D., Brun-Laguna, K., Watteyne, T., 2018. Prediction of frost events using Bayesian networks and random forest. *IEEE Internet Things J.* <https://doi.org/10.1109/JIOT.2018.2867333> hal-01867780.

Ding, L., Noborio, K., Shibuya, K., 2019. Frost forecast using machine learning – from assistance to causality. *Proced. Comput. Sci.* 159, 1001–1010. <https://doi.org/10.1016/j.procs.2019.09.267>.

Diniz, É.S., Lorenzon, A.S., de Castro, N.L.M., Marcatti, G.E., dos Santos, O.P., de Deus Júnior, J.C., Cavalcante, R.B.L., Fernandes-Filho, E.L., Hummelo Amaral, C., 2021. Forecasting frost risk in forest plantations by the combination of spatial data and machine learning algorithms. *Agric. For. Meteorol.* 306, 108450. <https://doi.org/10.1016/j.agrformet.2021.108450>.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leão, P.J., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36 (1), 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>.

Eccel, E., Ghielmi, L., Granitto, P., Barbiero, R., Grazzini, F., Cesari, D., 2007. Prediction of minimum temperatures in an alpine region by linear and non-linear post-processing of meteorological models. *Nonlin. Processes Geophys.* 14 (3), 211–222. <https://doi.org/10.5194/npg-14-211-2007>.

Elli, E.F., Sentelhas, P.C., Huth, N., Carneiro, R.L., Alvares, C.A., 2020. Gauging the effects of climate variability on Eucalyptus plantations productivity across Brazil: a process-based modelling approach. *Ecol. Indic.* 114, 106325. <https://doi.org/10.1016/j.ecolind.2020.106325>.

FAOSTAT, F., 2013. Food and Agriculture Organization of the United Nations. Statistical Database. <https://www.fao.org/faostat/en/#data/FBS.2023.5.20> (accessed on 20th May, 2023).

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The Shuttle Radar Topography Mission. *Rev. Geophys.* 45 (2), RG2004. <https://doi.org/10.1029/2005RG000183>.

Flynn, D.F.B., Wolkovich, E.M., 2018. Temperature and photoperiod drive spring phenology across all species in a temperate forest community. *New Phytol.* 219 (4), 1353–1362. <https://doi.org/10.1111/nph.15232>.

Frederiks, T.M., Christopher, J., Harvey, G.L., Sutherland, M.W., Borrell, A.K., 2012. Current and emerging screening methods to identify post-head-emergence frost adaptation in wheat and barley. *J. Exp. Bot.* 63 (15), 5405–5416. <https://doi.org/10.1093/jxb/ers215>.

Fuentes, M., Campos, C., García-Loyola, S., 2018. Application of artificial neural networks to frost detection in central Chile using the next day minimum air temperature forecast. *Chil. J. Agric. Res.* 78 (3), 327–338. <https://doi.org/10.4067/S0718-58392018000300327>.

Gaur, S., Drewry, D.T., 2024. Explainable machine learning for predicting stomatal conductance across multiple plant functional types. *Agric. For. Meteorol.* 350, 109955. <https://doi.org/10.1016/j.agrformet.2024.109955>.

Ghielmi, L., Eccel, E., 2006. Descriptive models and artificial neural networks for spring frost prediction in an agricultural mountain area. *Comput. Electron. Agric.* 54 (2), 101–114. <https://doi.org/10.1016/j.compag.2006.09.001>.

Gou, R., Buchmann, N., Chi, J., Luo, Y., Mo, L., Shekhar, A., Feigenwinter, I., Hörtnagl, L., Lu, W., Cui, X., Meng, Y., Song, S., Lin, G., Chen, Y., Liang, J., Guo, J., Peng, H., Lin, G., 2023. Temporal variations of carbon and water fluxes in a subtropical mangrove forest: insights from a decade-long eddy covariance measurement. *Agric. For. Meteorol.* 343, 109764. <https://doi.org/10.1016/j.agrformet.2023.109764>.

Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J., 2016. LSTM: a search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* 28 (10), 2222–2232. <https://doi.org/10.1109/TNNLS.2016.2582924>.

- Gu, Y., Zhao, Y., Guo, Z., Meng, L., Zhang, K., Wang, J., Lee, C.K.F., Xie, J., Wang, Y., Yan, Z., Zhang, H., Wu, J., 2023. The underappreciated importance of solar radiation in constraining spring phenology of temperate ecosystems in the Northern and Eastern United States. *Remote Sens. Environ.* 294, 113617. <https://doi.org/10.1016/j.rse.2023.113617>.
- Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., Yang, G.-Z., 2019. XAI—Explainable artificial intelligence. *Sci. Robot.* 4 (37), eaay7120. <https://doi.org/10.1126/scirobotics.aay7120>.
- Guo, W., Xu, Y., Zhang, H., Li, L., Wu, T., 2020. Experimental study on frost index of the marsh apple flowering period in cold and humid region of Tianshui city. *Anh. Agric. Sci.* 48 (19), 231–234. <https://doi.org/10.3969/j.issn.0517-6611.2020.19.060>.
- Guo, W., Xu, Y., Zhang, H., Wu, T., 2022. Investigation and analysis of frost damage of marsh apple during flowering period in qinzhou district. *J. Gansu Sci.* 34 (02), 66–71. <https://doi.org/10.16468/j.cnki.issn1004-0366.2022.02.012>.
- Haykin, S., 2001. *Redes neurais: principios e prática*. second ed. Bookman Editora, Canada.
- Hearst, M.A., Dumais, S.T., Osuna, E., Platt, J., Scholkopf, B., 1998. Support vector machines. *IEEE Intell. Syst.* 13 (4), 18–28. <https://doi.org/10.1109/5254.708428>.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Hoffmann, H., Rath, T., 2013. Future bloom and blossom frost risk for *Malus domestica* considering climate model and impact model uncertainties. *PLoS One* 8 (10), e75033. <https://doi.org/10.1371/journal.pone.0075033>.
- Hosseini, S., Rezaei, M., Bag-Mohammadi, M., Zendeheboudi, A., Ahmadi, G., 2019. Prediction of frost layer over flat plates under natural and forced convection conditions using intelligent and least-square fitting approaches. *Appl. Therm. Eng.* 148, 33–42. <https://doi.org/10.1016/j.applthermaleng.2018.11.022>.
- Ilhijaz, O., Du, Q., Kurban, A., Kasimu, A., Azadi, H., 2022. An explanatory spatio-temporal pattern analysis of the fruit industry in Xinjiang, China, between 1988 and 2017. *Trans. GIS* 26 (5), 2383–2409. <https://doi.org/10.1111/tgis.12967>.
- Inouye, D.W., 2000. The ecological and evolutionary significance of frost in the context of climate change. *Ecol. Lett.* 3 (5), 457–463. <https://doi.org/10.1046/j.1461-0248.2000.00165.x>.
- Jiang, S., Sweet, L.-b., Blougouras, G., Brenning, A., Li, W., Reichstein, M., Denzler, J., Shanguan, W., Yu, G., Huang, F., Zscheischler, J., 2024. How interpretable machine learning can benefit process understanding in the geosciences. *Earth's Future* 12 (7), e2024EF004540. <https://doi.org/10.1029/2024EF004540>.
- Kalma, J.D., Laughlin, G.P., Caprio, J.M., Hamer, P.J., 1992. *The Bioclimatology of Frost: Its Occurrence, Impact and Protection*. Library of Congress, Berlin Heidelberg.
- King, R.D., Orhobor, O.I., Taylor, C.C., 2021. Cross-validation is safe to use. *Nat. Mach. Intell.* 3 (4), 276. <https://doi.org/10.1038/s42256-021-00332-z>.
- Kotikot, S.M., Flores, A., Griffin, R.E., Sedah, A., Nyaga, J., Mugo, R., Limaye, A., Irwin, D.E., 2018. Mapping threats to agriculture in East Africa: performance of MODIS derived LST for frost identification in Kenya's tea plantations. *Int. J. Appl. Earth Obs. Geoinf.* 72, 131–139. <https://doi.org/10.1016/j.jag.2018.05.009>.
- Kotikot, S.M., Flores, A., Griffin, R.E., Nyaga, J., Case, J.L., Mugo, R., Sedah, A., Adams, E., Limaye, A., Irwin, D.E., 2020. Statistical characterization of frost zones: Case of tea freeze damage in the Kenyan highlands. *Int. J. Appl. Earth Obs. Geoinf.* 84, 101971. <https://doi.org/10.1016/j.jag.2019.101971>.
- Kumar, I., Tripathi, B.K., Singh, A., 2023. Attention-based LSTM network-assisted time series forecasting models for petroleum production. *Eng. Appl. Artif. Intell.* 123, 106440. <https://doi.org/10.1016/j.jag.2018.05.009>.
- Lagap, U., Ghaffarian, S., Gelinas-Gagne, S., Jilma, J., Liu, Z.Y., Luo, Z.Y., 2025. Towards reliable deep learning for post-disaster damage Assessment: an XAI-based evaluation. *Int. J. Disaster Risk Reduction* 130, 105839. <https://doi.org/10.1016/j.ijdrr.2025.105839>.
- Li, G., 2018. Investigation and enlightenment of frost during apple flowering period in Baoji, Shaanxi in 2018. *N.A. Hortic* 12, 5–7.
- Lindkvist, L., Gustavsson, T., Bogren, J., 2000. A frost assessment method for mountainous areas. *Agric. For. Meteorol.* 102 (1), 51–67. [https://doi.org/10.1016/S0168-1923\(99\)00087-8](https://doi.org/10.1016/S0168-1923(99)00087-8).
- Litschmann, T., Ströda, T., 2019. Surface temperature of plant tissues. Which method of air temperature measurement fits best? *Contrib. Geophys. Geodesy* 49 (1), 11–23. <https://doi.org/10.2478/congeo-2019-0002>.
- Liu, W., Li, D., Zhang, Q., 2020. Analysis of a severe freezing damage process in spring 2018. *Technol. Innov.* 20, 21–24. <https://doi.org/10.15913/j.cnki.kjycx.2020.20.007>.
- Liu, F., Feng, S., Jia, Z., Xu, Z., Ma, R., Fang, J., Lin, F., 2021a. Damage of late frost to fruit trees in Wuzhong area and preventive measures. *Mod. Hortic.* 44 (20). <https://doi.org/10.14051/j.cnki.xdy.2021.20.075> 182–183 + 186.
- Liu, Y., Liu, B., Ma, J., Cheng, C., Wang, K., Mao, L., He, Y., Qiu, M., Yang, X., 2021b. Evaluation of potential climatic production of apple during the possible growing period at Zhaotong, Yunnan across Cool Highland of Southwest China. *Chin. J. Agrometeorol.* 42 (02), 87–101. <https://doi.org/10.3969/j.issn.1000-6362.2021.02.001>.
- Liu, N., Hao, Z., Zhao, P., 2024. Explainable deep learning insights into the history and future of net primary productivity in China. *Ecol. Indic.* 166, 112394. <https://doi.org/10.1016/j.ecolind.2024.112394>.
- Lundberg, S.M., Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions. *arXiv:1705.07874*. <https://doi.org/10.1038/s42256-019-0138-9>.
- Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I., 2020. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* 2 (1), 56–67.
- Ma, H., Xu, T., Lv, D., 2018. Investigation of bud freezing damage to 'Cold Fortune' apples in north-central Liaoning after consecutive low temperatures. *North. Fruit* 03, 50–51. <https://doi.org/10.16376/j.cnki.bfsg.2018.03.023>.
- Monteith, J., Unsworth, M., 2013. *Principles of Environmental Physics: Plants, Animals, and the Atmosphere*. fourth ed. Academic press, Oxford.
- Morin, X., Chuine, I., 2014. Will tree species experience increased frost damage due to climate change because of changes in leaf phenology? *Can. J. For. Res.* 44 (12), 1555–1565. <https://doi.org/10.1139/cjfr-2014-0282>.
- Naing, W.Y.N., Huike, Z.Z., 2015. Forecasting of monthly temperature variations using random forests. *ARPN J. Eng. Appl. Sci.* 10 (21), 10109–10112.
- Natekin, A., Knoll, A., 2013. Gradient boosting machines, a tutorial. *Front. Neurobot.* 7, 21.
- Niu, Z., Zhong, G., Yu, H., 2021. A review on the attention mechanism of deep learning. *Neurocomputing* 452, 48–62. <https://doi.org/10.1016/j.neucom.2021.03.091>.
- Omazić, B., Anić, M., Telišman Prtenjak, M., Kvakić, M., Blašković, L., 2024. Analysis of different existing measurement-based methods and a new approach for frost probability detection. *Agric. For. Meteorol.* 347, 109898. <https://doi.org/10.1016/j.agrformet.2024.109898>.
- Parker, L., Pathak, T., Ostojica, S., 2021. Climate change reduces frost exposure for high-value California orchard crops. *Sci. Total Environ.* 762, 143971. <https://doi.org/10.1016/j.scitotenv.2020.143971>.
- Piao, S., Liu, Q., Chen, A., Janssens, I.A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., Zhu, X., 2019. Plant phenology and global climate change: current progresses and challenges. *Glob. Chang. Biol.* 25 (6), 1922–1940. <https://doi.org/10.1111/gcb.14619>.
- Qiu, M., Liu, B., Liu, Y., Wang, K., Pang, J., Zhang, X., He, J., 2020. Simulation of first flowering date for apple and risk assessment of late frost in main producing areas of northern China. *Trans. Chin. Soc. Agric. Eng.* 36 (21), 154–163. <https://doi.org/10.11975/j.issn.1002-6819.2020.21.019>.
- Qu, Z., Zhou, G., 2016. Climate suitability for potential Fuji apple cultivation in China. *Acta Meteor. Sin.* 74 (3), 479–490. <https://doi.org/10.11676/qxb.2016.027>.
- Rodrigo, J., 2000. Spring frosts in deciduous fruit trees—morphological damage and flower hardiness. *Sci. Hortic.* 85 (3), 155–173. [https://doi.org/10.1016/S0304-4238\(99\)00150-8](https://doi.org/10.1016/S0304-4238(99)00150-8).
- Ru, X., Jiang, Y., Luo, Q., Wang, R., Feng, X., Wang, J., Wang, Z., Li, M., Qu, Z., Su, B., Feng, H., Zhang, D., Liu, D., Yu, Q., He, J., 2023a. Evaluating late spring frost risks of apple in the Loess Plateau of China under future climate change with phenological modeling approach. *Sci. Hortic.* 308, 111604. <https://doi.org/10.1016/j.scienta.2022.111604>.
- Ru, X., Zhou, J., Gong, K., He, Z., Dai, Z., Li, M., Feng, X., Yu, Q., Feng, H., He, J., 2023b. Climate warming may accelerate apple phenology but lead to divergent dynamics in late-spring frost and poor pollination risks in main apple production regions of China. *Eur. J. Agron.* 150, 126945. <https://doi.org/10.1016/j.eja.2023.126945>.
- Shang, Q., 2019. Investigation and Remedial Measures of Freezing Injury in Apple Flowering Period in Luochuan, Shaanxi Province in 2018. Northwest A&F University.
- Shapley, L., 1997. *A Value for n-Person Games*. Contributions to the Theory of Games II (1953) 307–317. In: Harold William, K. (Ed.), *Classics in Game Theory*. Princeton University Press, Princeton, pp. 69–79.
- Shi, X., Huang, Q., Chang, J., Wang, Y., Lei, J., Zhao, J., 2015. Optimal parameters of the SVM for temperature prediction. *Proc. Int. Assoc. Hydrol. Sci.* 368, 162–167. <https://doi.org/10.5194/piahs-368-162-2015>.
- Simões, D.d.S., Fontana, D.C., Vicari, M.B., 2015. Use of LST images from MODIS/AQUA sensor as an indication of frost occurrence in RS. *Rev. Bras. Eng. Agric. Ambient.* 19 (10), 920–925. <https://doi.org/10.1590/1807-1929/agriambi.v19n10p920-925>.
- Štrumbelj, E., Kononenko, I., 2014. Explaining prediction models and individual predictions with feature contributions. *Knowl. Inf. Syst.* 41, 647–665. <https://doi.org/10.1007/s10115-013-0679-x>.
- Tang, W., Wang, W., Chen, D., Cui, N., Yang, H., Hu, X., 2021. Evaluation of the regional-scale optimal K rate based on sustainable apple yield and high-efficiency K use in loess plateau and Bohai Bay of China: a meta-analysis. *Agron* 11 (7), 1368. <https://doi.org/10.3390/agronomy11071368>.
- Tang, J., Wang, P., Li, X., Yang, J., Wu, D., Ma, Y., Li, S., Jin, Z., Huo, Z., 2023. Disaster event-based spring frost damage identification indicator for tea plants and its applications over the region north of the Yangtze River, China. *Ecol. Indic.* 146, 109912. <https://doi.org/10.1016/j.ecolind.2023.109912>.
- Tian, H.R., Wang, P.X., Tansey, K., Wang, J., Quan, W.T., Liu, J.M., 2024. Attention mechanism-based deep learning approach for wheat yield estimation and uncertainty analysis from remotely sensed variables. *Agric. For. Meteorol.* 356, 110183. <https://doi.org/10.1016/j.agrformet.2024.110183>.
- Vanoni, M., Bugmann, H., Nötzli, M., Bigler, C., 2016. Drought and frost contribute to abrupt growth decreases before tree mortality in nine temperate tree species. *For. Ecol. Manag.* 382, 51–63. <https://doi.org/10.1016/j.foreco.2016.10.001>.
- Vapnik, V., 1995. *The Nature of Statistical Learning Theory*. Springer, N. Y., p. 10 978–971.
- Vaswani, A., 2017. Attention is All You Need. *arXiv preprint arXiv:1706.03762*. <https://doi.org/10.48550/arXiv.1706.03762>.
- Vitasse, Y., Schneider, L., Rixen, C., Christen, D., Rebetez, M., 2018. Increase in the risk of exposure of forest and fruit trees to spring frosts at higher elevations in Switzerland over the last four decades. *Agric. For. Meteorol.* 248, 60–69. <https://doi.org/10.1016/j.agrformet.2017.09.005>.
- Wang, M., Liu, B., Liu, Y., Yang, X., Han, S., Qiu, M., 2020. Assessment on the freezing injury risk during apple flowering in Likan and Xunyi. *Chin. J. Agrometeorol.* 41 (6), 381. <https://doi.org/10.3969/j.issn.1000-6362.2020.06.005>.
- Wang, H., Liu, Q., Gui, D., Feng, X., Qu, J., Zhao, J., Wang, G., Wei, G., 2023. Simulation of potential vegetation distribution in arid areas at regional scale. *Catena* 232, 107428. <https://doi.org/10.1016/j.catena.2023.107428>.
- Wen, K., Ding, Y., 2008. *Encyclopedia of Meteorological Disasters in China: Comprehensive Volume*. first ed. China Meteorological Press, Beijing.
- Xu, R., Zhu, Y., Wang, J., Duan, X., Ma, G., Li, H., 2019. Investigation and analysis of late frost during 'Fuji' apple flowering period of Ningxia in 2018. *J. Ningxia Agric. For. Sci. Technol.* 60 (08), 17–20. <https://doi.org/10.3969/j.issn.1002-204x.2019.08.006>.
- Yang, B., Li, Y., Sun, Y., He, D., Dai, S., 2020a. Analysis on the frost damage in florescence of Jixian apple trees in 2018 and climate characteristics. *J. Shanxi Norm. Univ. (Nat. Sci. Ed.)* 34 (03), 79–82. <https://doi.org/10.16207/j.cnki.1009-4490.2020.03.015>.

- Yang, A., Sun, Z., Li, G., Yang, L., He, F., 2020b. Discussion on meteorological service and defense measures for decision making of freezing injury in apple flowering period. *J. Agric. Catastrophol.* 10 (09), 46–48. <https://doi.org/10.19383/j.cnki.nyzhyj.2020.09.021>.
- Yang, A., Li, G., Yang, L., 2022. Causes and prevention measures of two apple freezing injuries in Yan'an region in 2020. *J. Shaanxi Meteorol.* 04, 75–79.
- Yang, Y., Zhang, X., Liang, X., Ma, M., Wei, J., 2023. Studies on the sensitivity of apple flowers and fruits to low temperatures from flowering to young fruiting stage. *Acta Botanica Boreali-occident. Sin.* 43 (10), 1694–1703. <https://doi.org/10.7606/j.issn.1000-4025.2023.10.1694>.
- Yin, H., 2019. An investigation of the selling price of apple in baishui country in 2018. *J. Fruit Resour.* 03, 47–48. <https://doi.org/10.16010/j.cnki.14-1127/s.2019.03.018>.
- Yoshino, M., 1984. Thermal belt and cold air drainage on the mountain slope and cold air lake in the basin at quiet, clear night. *Geojournal* 8, 235–250. <https://doi.org/10.1007/BF00446473>.