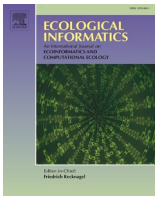
Ecological Informatics 71 (2022) 101835

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



Estimating the first flowering and full blossom dates of Yoshino cherry (*Cerasus* × *yedoensis 'Somei-yoshino'*) in Japan using machine learning algorithms

Yoshifumi Masago *, Maychee Lian

Center for Climate Change Adaptation, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan

ARTICLE INFO

Keywords:
Phenology
Yoshino cherry
Machine learning
Gradient boosting decision tree
Random forest
Artificial neural network

ABSTRACT

Climate change alters the phenology of various plants. For example, increasing temperatures shift the first flowering and full blossom days of Yoshino cherry trees and affect cultural events related to cherry blossoms. We developed models to estimate the first flowering and full blossom dates of Yoshino cherry in Japan based on temperature and phenological data observed at 82 stations in Japan for 68 years (1953–2020). Three machine learning algorithms, namely, the random forest (RF), artificial neural network (ANN), and gradient boosting decision tree (GBDT) algorithms, were utilized, and the hyperparameters were optimized using Optuna. The GBDT models produced the best estimation accuracy, with an overall root mean square error (RMSE) = 1.53 and 1.48 days for the first flowering date and full blossom date, respectively. Furthermore, our analysis using Shapley Additive Explanations (SHAP) revealed that in the RF and GBDT models, the low temperature in winter and high temperature in spring would advance the estimated first flowering and full blossom dates.

1. Introduction

Plant life cycle events such as flowering, leaf expansion, leaf coloring, and defoliation are closely related to the environment and climate. Temperature is one of the main factors affecting plant phenology (Badeck et al., 2004; Cleland et al., 2006; Primack et al., 2009; Tomczyk and Szyga-Pluta, 2019; Vitasse et al., 2011; Wang and Tang, 2019). For example, increasing temperatures due to climate change alter the phenology of various plants (Gaira and Dhar, 2020; Piao et al., 2019; Richardson et al., 2013). Changes in the phenology of a plant alter its ecology and interrelationships with other plants and animals. These changes also affect traditional cultural events associated with a plant, which may affect local communities and economies (Chung et al., 2011; Nagai et al., 2019; Primack et al., 2009).

Cherry trees, especially Yoshino cherry tree (*Cerasus* × *yedoensis* 'Somei-yoshino'), are iconic spring plants in Japan. The blooming of cherry blossoms heralds the arrival of spring, and cherry blossom viewing, "hanami", is a traditional spring event. However, the first flowering and full blossom periods of cherry trees occur earlier as temperatures increase due to climate change. Process-based models based on the biological characteristics of cherry trees have been employed to analyze the relationships between temperature and the first flowering and full blossom days in Japan (Allen et al., 2014; Aono and Moriya, 2003; Aono and Murakami, 2017; Asakura et al., 2010; Ibáñez et al., 2010; Maruoka and Itoh, 2009; Miller-Rushing et al., 2007) and other countries (Chung et al., 2011, 2016). Although these models well predict the first flowering dates, they have a problem of low prediction accuracy for the warm southern regions of Japan (Aono and Moriya, 2003; Asakura et al., 2010; Maruoka and Itoh, 2009).

Predicting plant phenology using machine learning has received increasing attention. Machine learning is widely utilized in phenological analyses, such as those related to olive tree flowering (Oses et al., 2020a, 2020b), maize yields (Kang et al., 2020), and rice yields (Guo et al., 2021). Machine learning shows higher predictive capabilities than statistical models (Kang et al., 2020). Nagai et al. (2020) applied a self-organizing map to predict the first flowering day of Yoshino cherry, marking the only application of machine learning in studying the phenology of cherry trees.

In the present study, three commonly employed machine learning algorithms for supervised regression models, namely, the random forest (RF), artificial neural network (ANN), and gradient boosting decision tree (GBDT) algorithms, were selected to develop models for estimating the first flowering and full blossom dates of Yoshino cherry trees. The first flowering and full blossom dates and daily average temperature

* Corresponding author.
E-mail address: masago.yoshifumi@nies.go.jp (Y. Masago).

https://doi.org/10.1016/j.ecoinf.2022.101835
Received 9 January 2022; Received in revised form 22 September 2022; Accepted 22 September 2022
Available online 26 September 2022

observed at 82 weather stations in Japan were employed as the response variables and explanatory variables, respectively. After optimizing the structure and hyperparameters of the machine learning models, we compared the estimation accuracies of the models using each algorithm. In addition, we investigated how the machine learning models estimated the output values by evaluating the effect of each input parameter (daily average temperature) on the output values.

2. Materials and methods

2.1. Climate and phenological observation data

The Japan Meteorological Agency (JMA) has been observing the first flowering and full blossom dates of cherry trees at meteorological offices in Japan since 1953. The first flowering and full blossom dates of three types of cherry trees (*Cerasus* × *yedoensis 'Somei-yoshino'* (Yoshino cherry), *Cerasus campanulata*, and *Cerasus sargentii*) were recorded at 102 offices. We utilized the data from the 82 observation sites of Yoshino cherry (Japan Meteorological Agency, 2021). Most sites showed that the first flowering and full blossom dates gradually occurred earlier (Figs. 1 and 2). On the contrary, we did not observe a consistent trend in the period between first flowering and full blossom dates. The period became longer over time at 44 of the 81 sites and shorter at 37 sites.

The years in which the observations were recorded differed among the observation sites. We used all available data up to 2020, resulting in 4952 and 4921 data points for the first flowering and full blossom dates, respectively. The data were provided in month-day format on the JMA website (Japan Meteorological Agency, 2021) and were converted to day-of-year format for analysis.

Temperature was selected as the explanatory variable because all process-based models used temperature as a parameter (Allen et al., 2014; Aono and Moriya, 2003; Aono and Murakami, 2017; Asakura et al., 2010; Chung et al., 2011, 2016; Ibáñez et al., 2010; Maruoka and Itoh, 2009; Miller-Rushing et al., 2007). These models utilized intermediate indices calculated from daily or hourly average temperatures, such as chill units and temperature conversion days (Allen et al., 2014; Aono and Moriya, 2003; Aono and Murakami, 2017; Asakura et al., 2010; Maruoka and Itoh, 2009), to estimate first flowering and full blossom dates. We hypothesized that we could avoid these intermediate indices by utilizing machine learning algorithms. Thus, we selected daily average temperatures at the meteorological offices for one year (365 consecutive days) as explanatory variables. The homogeneity-corrected daily average temperature data, considering the effects of relocation of the offices, changes in instruments, and changes in observation and statistical methods, were provided by JMA and employed as the observed values.

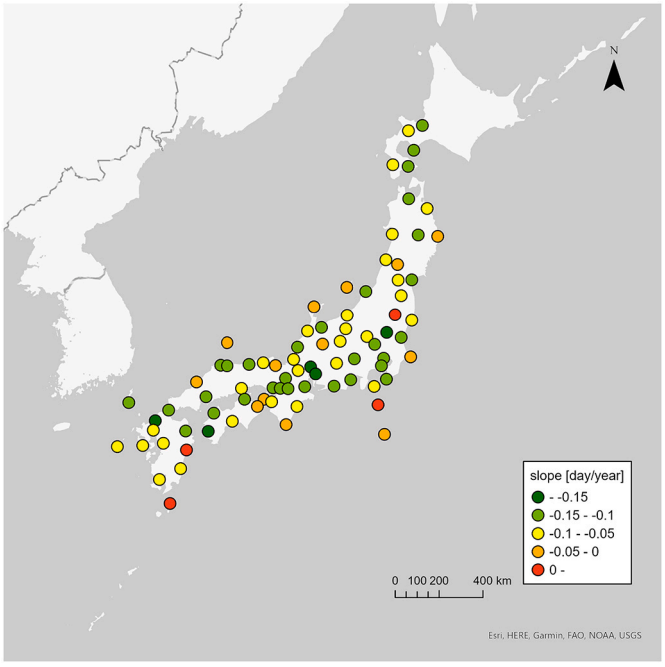


Fig. 1. The observation sites of the first flowering dates (n = 82) with the slope of the temporal trends [day/year]. The slope was calculated using all available data and the time interval varies depending on the site.

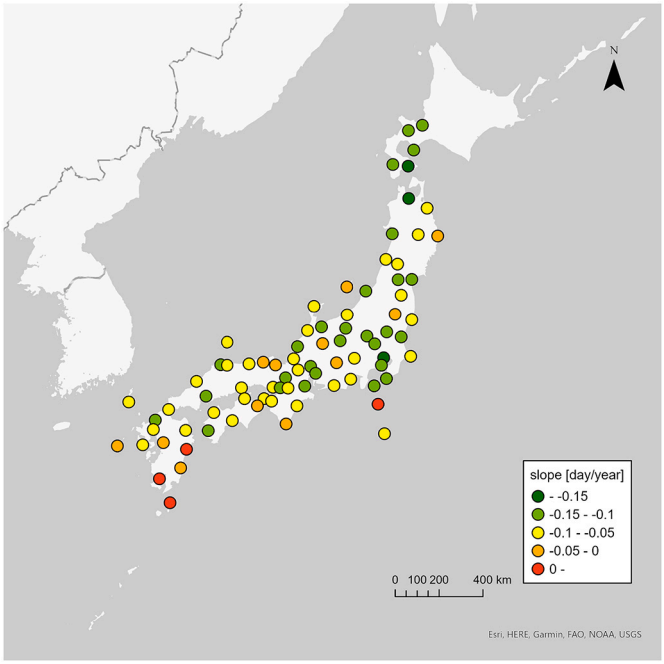


Fig. 2. The observation sites of the full blossom dates ($n = 82$) with the slope of the temporal trends [day/year]. The slope was calculated using all available data and the time interval varies depending on the site.

2.2. Data preprocessing

The first flowering and full blossom dates ranged from March 11th to May 26th and from March 20th to May 31st, respectively. The daily average temperature for 365 consecutive days from June 1st of the previous year to May 31st of the following year was used as the explanatory variable to ensure that the period of daily average temperature that was used as the explanatory variable included these observation periods and contained as many preblooming data points as possible. One year was selected because all process-based models considered environmental factors within one year prior to the events. Leap days were excluded from the analysis. Some daily average temperature data sets were incomplete and contained missing observations. In these cases, the data points were excluded from the study. As a result, the data sets for the first flowering and full blossom dates contained 4844 and 4814 data points, respectively. Before applying the data sets to the models, all data (daily average temperature and first flowering and full blossom dates) were rescaled using z-score normalization, which is a standard scaling method compared with others (e.g., min-max scaling).

2.3. Construction of machine learning models

Three machine learning algorithms, RF, ANN, and GBDT, were selected to develop the estimation models. All algorithms included the daily average temperature for 365 consecutive days as the explanatory variables. The models were coded in Python (version 3.8.12) using Scikit-learn 1.0.1 (Pedregosa et al., 2011), TensorFlow 2.3.0 (Abadi

et al., 2016), and LightGBM 3.2.1 (Ke et al., 2017) libraries. In building a machine learning model, the structure and hyperparameters of the model must be determined, such as the number of decision trees and the depth of trees for RF models and the number of intermediate layers and nodes for ANN models. We optimized the structure and hyperparameters of the RF and ANN models using Optuna 2.10.0 (Akiba et al., 2019) with the parameters and ranges shown in Table 1 as candidates. For the LightGBM models, we used Optuna's LightGBM Tuner (optuna.integration.lightgbm) for hyperparameter fitting. The performance of the algorithms was evaluated by comparing the loss function (mean squared error (MSE)) of the optimized models. In addition, the mean absolute error (MAE) and the coefficient of determination (R^2) were calculated for reference.

Next, the first flowering and full blossom dates for each station and year were calculated using the optimized models. Finally, the root mean square error (RMSE) was calculated for all observation sites to evaluate the overall performance of the optimized models, and the RMSE for each observation site was calculated to evaluate the performance of the models for each site.

2.4. Understanding the machine learning models using SHAP values

The constructed models included the daily average temperature for 365 consecutive days as the explanatory variables. However, these variables are not equally important, and a limited number of input values may significantly affect the output values. The recent development of explainable artificial intelligence (AI) technology has facilitated the development of methods to evaluate the relative importance of input values. We applied Shapley Additive Explanations (SHAP) (Lundberg and Lee, 2017), which evaluates the magnitude of the effect of each input value on the output value. Using this method, we identified periods when the daily average temperature significantly affected the first flowering or full blossom dates.

First, the SHAP values were calculated for each of the 365 explanatory variables for a given response value. A positive SHAP value for a given input value indicates that the variable increased the response value. In contrast, a negative SHAP value indicates that the variable decreased the output value. The absolute SHAP value shows the magnitude of the effect. Next, we calculated the average of the absolute SHAP value (AA_SHAP) for each input value using eq. 1 to evaluate the magnitude of the effect of an explanatory variable:

Table 1
Parameters of the machine learning algorithms optimized using Optuna.

Algorithm	Parameter name	Description	Candidates
Random forest	n_estimators	Number of trees	int, 1 to 100
	max_depth	Maximum depth of the tree	int, 1 to 20
ANN	n_layers	Number of middle layers	int, 0 to 5
	units (initial layer)	Number of nodes in the input layer	int, 10, 20, … 100
	units (middle layers)	Number of nodes in the middle layers	int, 10, 20, … 100
	activation	Activation function in the input and middle layers	ReLu or linear
	optimizer	Optimizer	Adam or RMSprop
GBDT	num_leaves	Maximum tree leaves for base learners	
	lambda_l1	L1 regularization	
	lambda_l2	L2 regularization	
	feature_fraction	A subset of features on each iteration (tree)	
	bagging_fraction	Randomly select a part of data without resampling	
	bagging_freq	Frequency for bagging	
	min_child_samples	Minimal number of data in one leaf	

$$AA\_SHAP_j = \frac{1}{n_i}\sum_i \left|SHAP_{i,j}\right| \quad (1)$$

where AA_SHAP_j is the average of the absolute SHAP values for input variable j (i.e., the daily average temperature on day j), n_i is the number of response values (4844 for the first flowering date models and 4814 for the full blossom date models), and $SHAP_{i,j}$ is the SHAP value for the jth input parameter of the ith response value.

By comparing the AA_SHAP values of the 365 input variables, we determined the daily average temperature for the days that strongly influenced the response values (first flowering and full blossom dates). In addition, we calculated the 21-day moving average of the AA_SHAP to identify the periods with high impacts on the response values.

Next, for each of the 365 input values, the correlation coefficient between the SHAP values and the daily average temperature was calculated. If the correlation coefficient was positive, the SHAP value increased (i.e., the first flowering or full blossom date occurred later) when the daily average temperature increased. Conversely, an explanatory variable with a negative correlation coefficient indicates that the higher the daily average temperature is, the earlier the first flowering or full blossom date. Therefore, by examining the correlation coefficients for the 365 days, we evaluated the daily average temperature for specific periods that positively or negatively affected the first flowering or full blossom date.

3. Results

3.1. Estimation accuracy of the three algorithms

We optimized the three machine learning algorithms (RF, ANN, and GBDT) using Optuna. Table 2 shows the optimal model parameters and their performance (MSE, MAE, and R^2). For both the first flowering date and the full blossom date, the GBDT models showed the best accuracy (MSE = 0.0381 and 0.0424, MAE = 0.148 and 0.158, R^2 = 0.964 and

Table 2
The best-fit parameters and the performance of the optimized models.

Algorithm	Best parameters	MSE	MAE	R^2
First flowering date				
Random forest	n_estimators = 71, max_depth = 16	0.0530	0.175	0.942
ANN	n_layers = 0, units (initial layer) = 90, units (middle layers) = 70, activation = ReLu, optimizer = Adam	0.0429	0.156	0.958
GBDT	num_leaves = 31, lambda_l1 = 0.000192 lambda_l2 = 0.0215 feature_fraction = 1.0 bagging_fraction = 0.506 bagging_freq = 3 min_child_samples = 20	0.0381	0.148	0.964
Full blossom date				
Random forest	n_estimators = 92, max_depth = 15	0.0586	0.175	0.940
ANN	num_layer = 0, units (initial layer) = 90, units (middle layers) = 60, activation = ReLu, optimizer = Adam	0.0485	0.170	0.951
GBDT	num_leaves = 31, lambda_l1 = 0 lambda_l2 = 0 feature_fraction = 0.5 bagging_fraction = 0.573 bagging_freq = 7 min_child_samples = 20	0.0424	0.158	0.955

0.965, respectively), followed by the ANN and RF models, while the differences were not significant. The optimized ANN models did not have any intermediate layers, and the model with the input layer followed by the output layer was determined to be the best.

3.2. SHAP values

Figs. 3 and 4 show the absolute SHAP values for the optimized models of the first flowering and full blossom dates using the three algorithms. For the first flowering date models, all models showed the largest peaks in the 21-day moving average of the SHAP values from February to April. The maximum SHAP value exceeded 0.1 in the RF and GBDT models (RF: 0.134 (February 19th), GBDT: 0.136 (March 2nd)), but such a high peak was not observed in the ANN model (the maximum value was 0.0510 on March 6th). Apart from these high peaks, the RF model showed small peaks in June–July and November, and the GBDT model also showed small peaks in June–July and November–January. In contrast, the ANN model did not show these peaks. According to the correlation coefficients between daily average temperatures and the SHAP values, 300 (82%) of the 365 data points showed negative correlations in the RF model. For almost all days, the higher the daily average temperature was, the earlier the first flowering date. Positive correlation coefficients were mainly observed from October to December. The ANN model showed negative correlation coefficients for 200 of 365 days (55%), which clustered from February to April. The GBDT model produced negative correlation coefficients for 188 of 365 days (52%). Positive correlation coefficients were observed in 102 of 123 days (83%) from October to January, and negative correlation coefficients were observed in 115 of 120 days (96%) from February to May.

These characteristics were similar in the models for the full blossom dates. All models revealed high peaks for the absolute SHAP values from February to April. In the RF and GBDT models, the maximum SHAP values were approximately 0.1 (RF: 0.108 (March 2nd); GBDT: 0.0926 (March 2nd)), while the value for the ANN model was smaller (0.0599 on March 30th). The RF and GBDT models also showed small peaks in

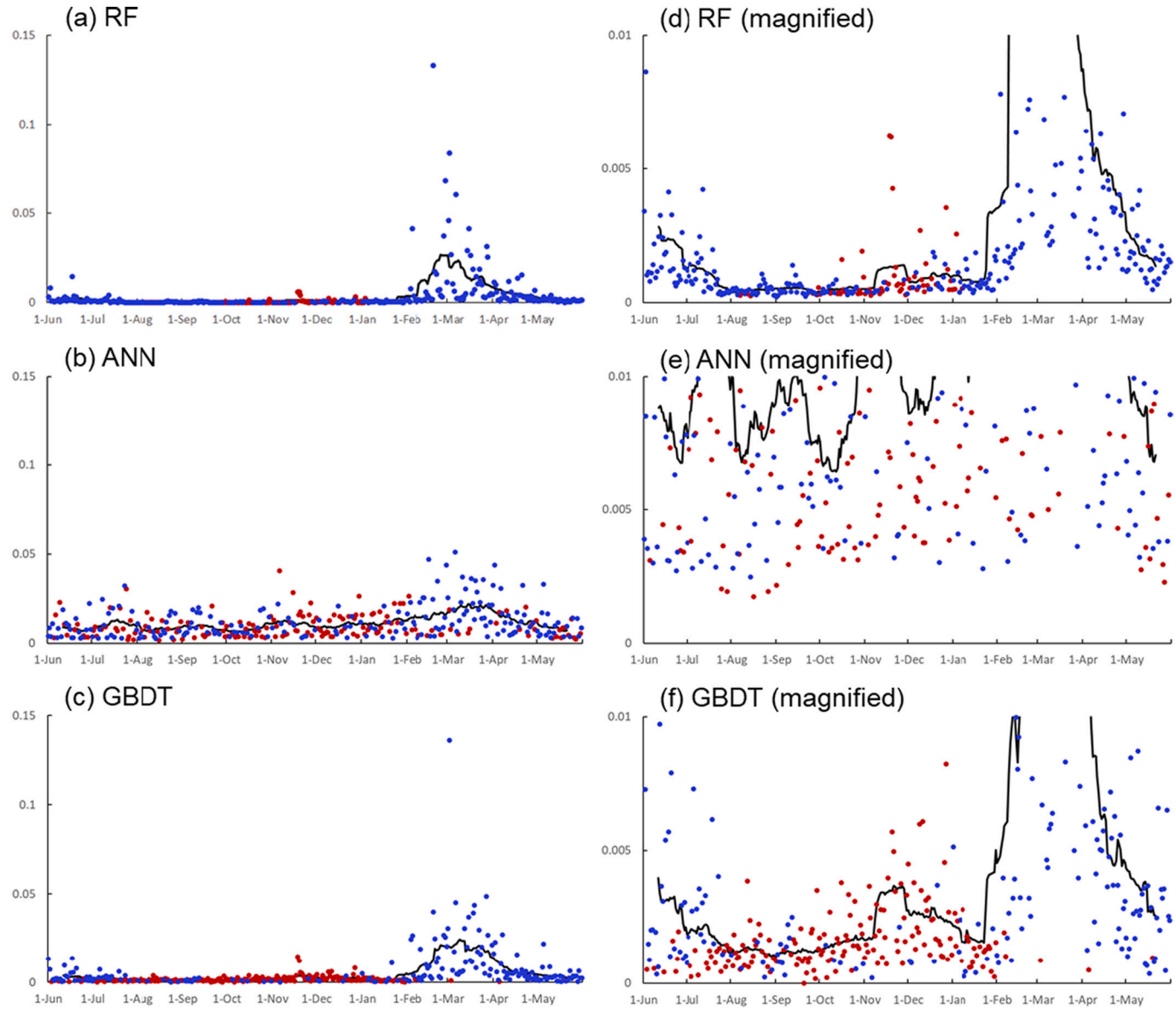


Fig. 3. The absolute SHAP values for each day derived from the models of the first flowering dates using the three algorithms (RF, ANN, and GBDT). The red plots indicate that the correlation coefficient between the daily average temperature and the absolute SHAP values was positive, and the blue dots indicate negative correlation coefficients. The black line is the 21-day moving average of the absolute SHAP values. Figures (d) – (f) are magnifications of Figures (a) – (c) in the y-axis direction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

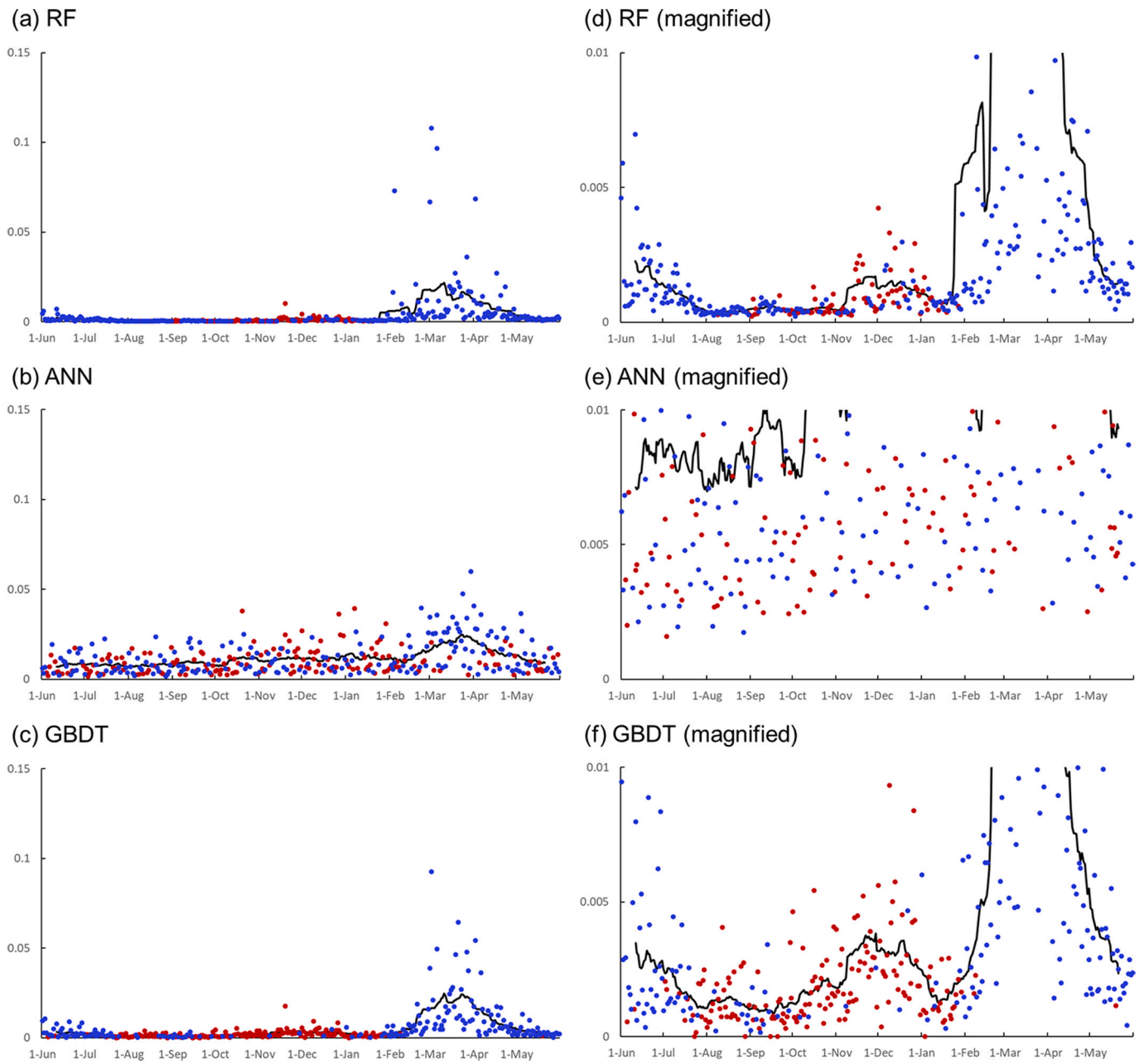


Fig. 4. The absolute SHAP values for each day derived from the models of the full blossom dates using the three algorithms (RF, ANN, and GBDT). The red plots indicate that the correlation coefficient between the daily average temperature and the absolute SHAP values was positive, and the blue dots indicate negative correlation coefficients. The black line is the 21-day moving average of the absolute SHAP values. Figures (d) – (f) are magnifications of Figures (a) – (c) in the y-axis direction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

June–July and November–December, while the ANN model did not show these peaks. The correlation coefficients between the daily average temperature and the SHAP values from the RF model were negative for 281 (77%) of the 365 data points. The ANN model produced negative correlation coefficients for 199 (55%) of 365 days and tended to be negative from February to April. The GBDT model showed negative correlation coefficients for 194 days (53%), with positive correlation coefficients observed on 108 (88%) of 123 days from October to January and negative correlation coefficients on 117 (98%) of 120 days from February to May.

3.3. Estimation accuracy for each observation site

Fig. 5 shows the overall RMSE and RMSE per location obtained using the three algorithms. The overall RMSE was lowest for the GBDT model (1.53 for the first flowering date and 1.48 for the full blossom date), followed by the RF (1.60 and 1.50) and ANN (2.20 and 2.10) models. For all algorithms, the overall RMSE was smaller for the full blossom date model than for the first flowering date model.

Differences in the variability of the RMSE were observed between locations, depending on the model. The RF models contained the largest number of sites with very accurate estimates (RMSE <1), followed by the GBDT and ANN (Table 3) models. On the other hand, the GBDT models contained the fewest locations with a low estimation accuracy (RMSE >2), followed by the RF and ANN models. The standard deviation of the RMSE for each site was the smallest for the GBDT models (0.46 for the first flowering date and 0.43 for the full blossom date), followed by the RF (0.58 and 0.58) and ANN (0.63 and 0.60) models. A

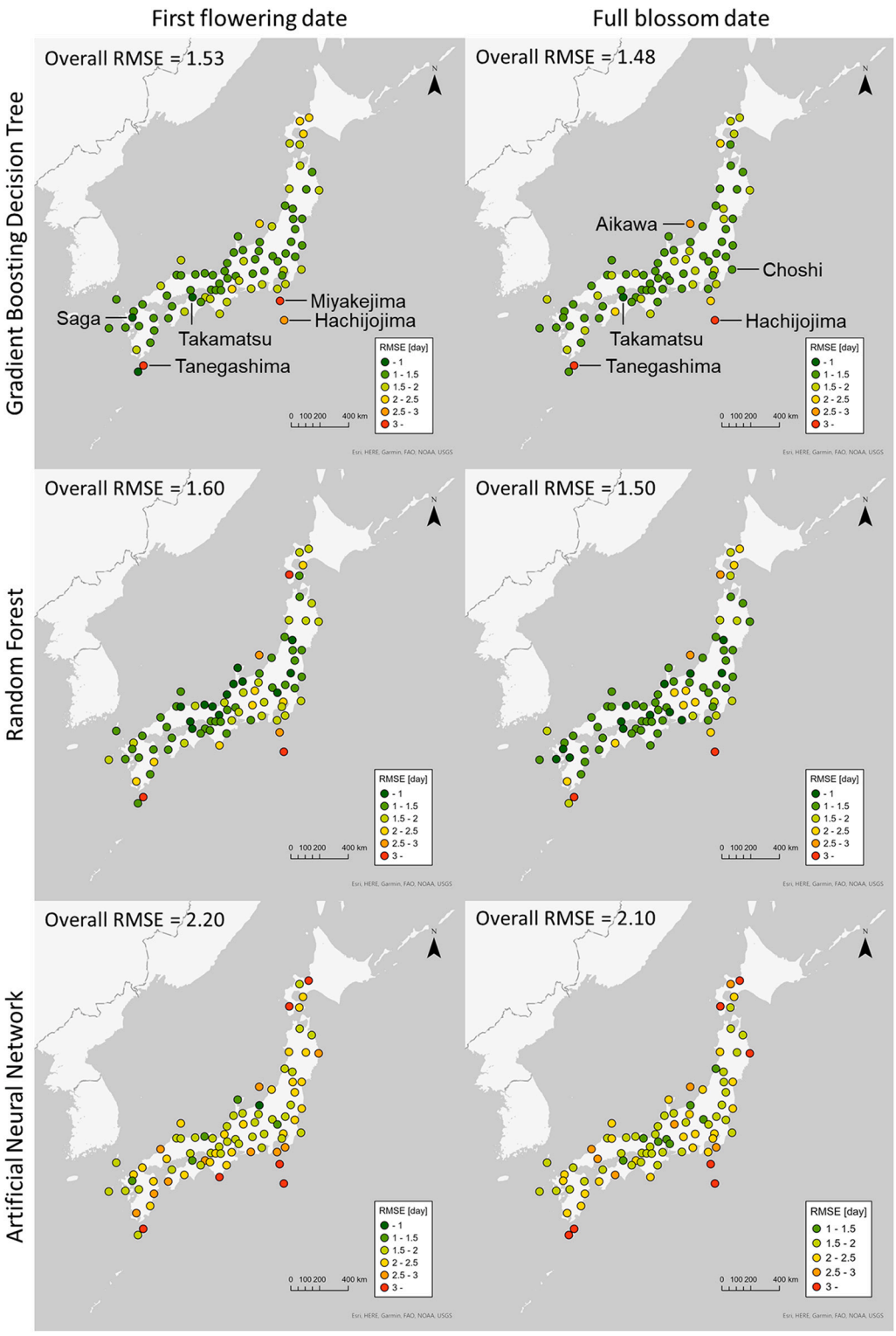


Fig. 5. The RMSE for each observation site derived from the first flowering and full blossom date models using the three machine learning algorithms. The names of the observation sites with RMSE >2.5 or the lowest RSME using the GBDT models are indicated.

comparison of the GBDT model with the RF model revealed that the RF model could estimate the dates with good accuracy at some sites but also showed low accuracy at several sites. In contrast, the GBDT model estimated the dates with stable accuracy at many sites.

The GBDT algorithm showed the lowest overall RMSE among the three algorithms utilized in this study. The RMSE values for the first flowering model did not significantly differ by region, but three sites (Miyakejima, Hachijojima, and Tanegashima), which are located on small islands, had relatively high RMSEs (RMSE >2.5). The model for the full blossom dates was also inaccurate (RMSE >2.5) at three sites on small islands (Aikawa, Hachijojima, and Tanegashima). Two sites, Hachijojima and Tanegashima, which are located on islands in the Pacific Ocean, showed high RMSEs for both models. Figs. 6 and 7 show the observed and estimated dates for the sites with RMSE >2.5 and those

Table 3
Performance of the optimized models using the three algorithms.

	Overall RMSE		Number of sites with RMSE <1		Number of sites with RMSE >2		Standard deviation of the site RMSE	
	FF	FB	FF	FB	FF	FB	FF	FB
GBDT	1.53	1.48	3	1	8	6	0.46	0.43
RF	1.60	1.50	13	13	12	13	0.58	0.58
ANN	2.20	2.10	1	0	40	37	0.63	0.60

FF: first flowering date, FB: full blossom date.

with the lowest RMSE, respectively. Even at the poor performance sites, the models generated accurate estimates in most years. However, some years showed a substantial discrepancy, especially when the observed first flowering or full blossom dates occurred extraordinarily early (Fig. 6 (a), (c), (f)) or late (Fig. 6 (b), (c), (e), (f)), which resulted in a large RMSE.

4. Discussion

Machine learning-based modeling has been increasingly applied in various fields, including phenological studies. The selection of an appropriate algorithm for the target event and the data used are necessary to build a model with good estimation accuracy. We applied

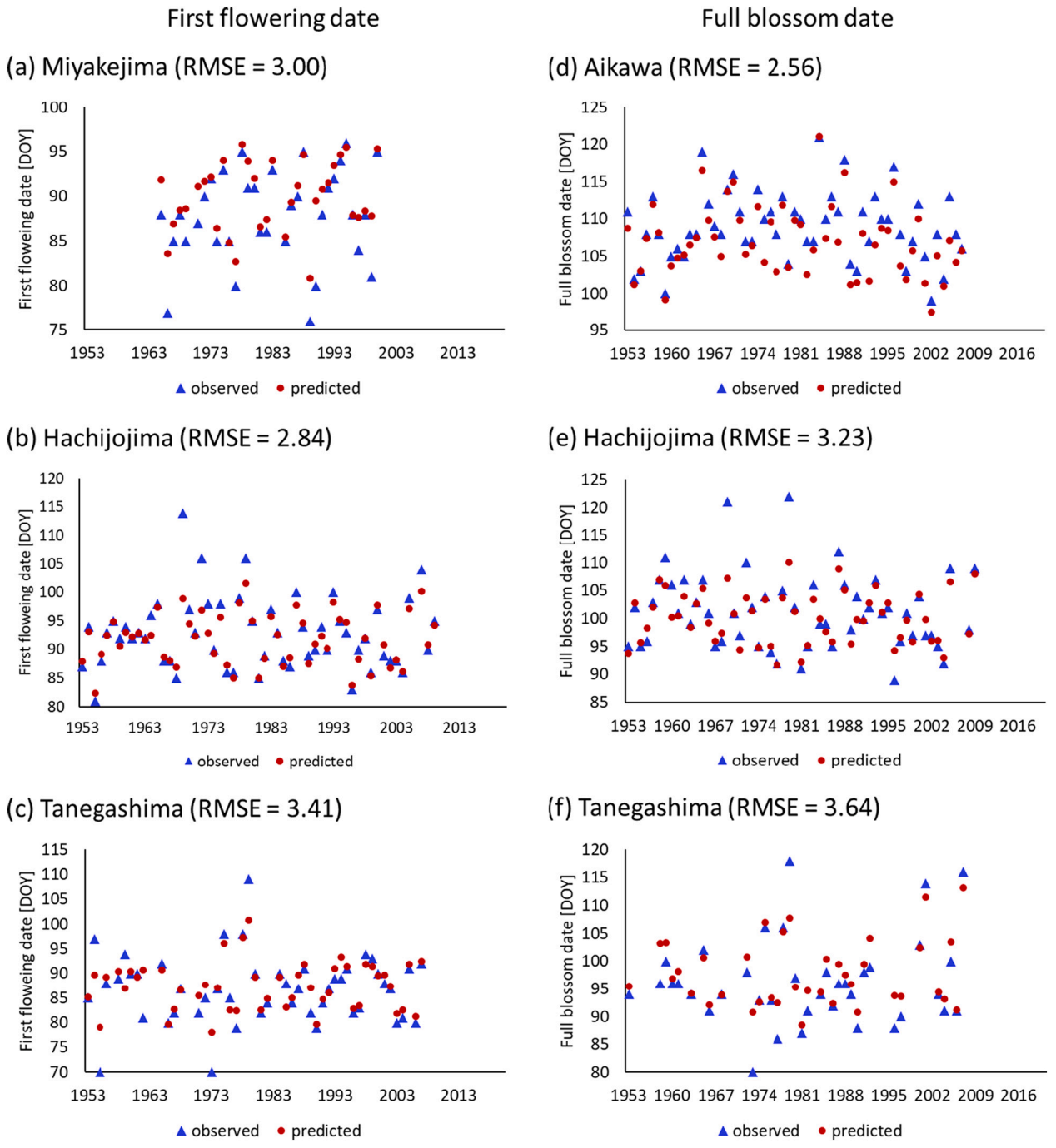


Fig. 6. The observed and estimated first flowering and full blossom dates using the GBDT models at the observation sites with RMSE >2.5. The locations of these sites are shown in Fig. 5.

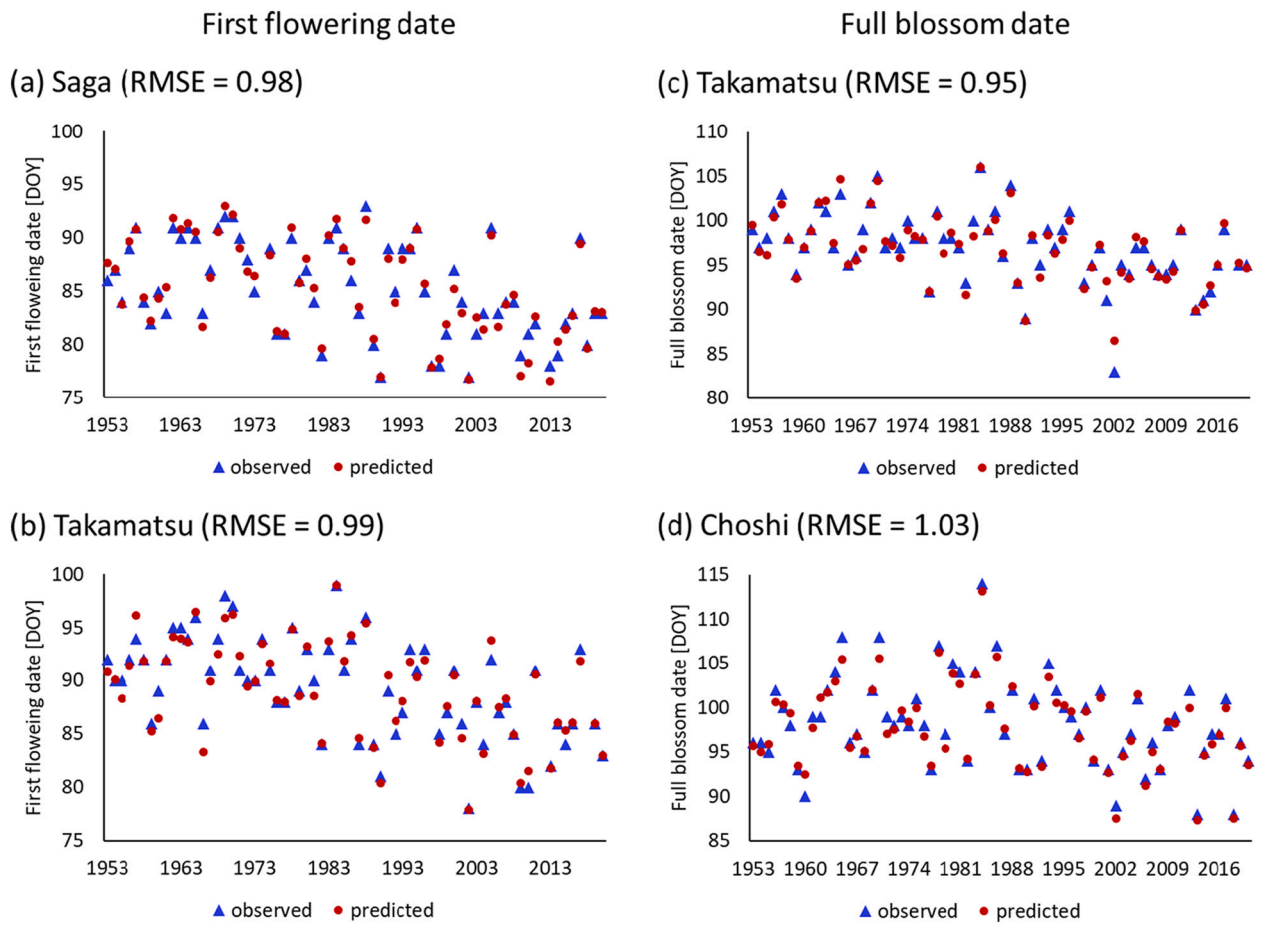


Fig. 7. The observed and estimated first flowering and full blossom dates using the GBDT models at the observation sites with the lowest RMSE. The locations of these sites are shown in Fig. 5.

three machine learning algorithms that are commonly employed in supervised regression models, RF, ANN, and GBDT, to estimate the first flowering and full blossom dates of Yoshino cherry across Japan. Since it was developed in 2017 (Ke et al., 2017), LightGBM, which is a GBDT method, has attracted attention because of its high model estimation accuracy and fast computation. This is the first study to use LightGBM for estimating the first flowering and full blossom dates of cherry trees. The GBDT (LightGBM) models were more accurate (lowest overall RMSE) and stable (lowest standard deviation of the RMSE for each site) than the RF and ANN models. The GBDT was also the best algorithm for predicting the leaf unfolding dates observed in Herbin among these three machine learning algorithms in addition to statistical methods (Dai et al., 2019).

The RMSE values calculated for all 82 stations using the GBDT models were 1.53 for the first flowering date and 1.48 for the full blossom date. Overall, the full blossom date model produced a lower RMSE. These values were lower than those of existing process-based models (Aono and Moriya, 2003; Aono and Murakami, 2017; Asakura et al., 2010; Maruoka and Itoh, 2009) and machine learning models (Nagai et al., 2020), which revealed the high accuracy of the GBDT model. In addition, we did not observe a geographical difference in the accuracy of the model, with the exception of some island sites. In previous studies, the estimation accuracy was lower in the southern regions of Japan (Aono and Moriya, 2003; Asakura et al., 2010; Maruoka and Itoh, 2009), which is the southern limit of Yoshino cherry. Due to its high winter temperatures, this region lacks the low-temperature environment necessary to release endodormancy. As a result, the first flowering dates considerably fluctuate, which reduces prediction accuracy (Aono and Moriya, 2003; Asakura et al., 2010; Maruoka and Itoh, 2009). Some studies have mitigated this effect by using different models or parameters to reflect regional characteristics (Aono and Moriya, 2003; Aono and Murakami, 2017).

In the GBDT models, the accuracy was relatively low at certain island sites, namely, Miyakejima, Hachijojima, Tanegashima, and Aikawa. However, this reduction was minimal compared with that reported in other studies (Aono and Moriya, 2003; Maruoka and Itoh, 2009). Among the 82 stations, the largest RMSE was observed in Tanegashima, with an RMSE of 3.41 for the first flowering date and an RMSE of 3.64 for the full blossom date. This station also showed the lowest estimation accuracy among other models (Aono and Moriya, 2003; Aono and Murakami, 2017; Maruoka and Itoh, 2009).

By comparing the estimated and measured values at these poor-performance sites (Fig. 6), we discovered that the model estimated the first flowering and full blossom dates in most years with reasonable accuracy, but a few years had substantial deviations, which deteriorated the model performance. Since all Yoshino cherry trees are clones originating from a single tree (Innan et al., 1995; Iketani et al., 2007), it is not likely that the extraordinal events occurred due to biological differences of the trees, such as local genotypes. We speculate that local climatic events contributed to the advance or delay of the first flowering or full blossom dates because all sites with high RMSE are on small islands. If this is the case, we could improve the accuracy by including other parameters, such as wind speed or precipitation that could reflect local climatic events.

The low temperatures in autumn and winter, which are necessary for endodormancy release, and subsequent high temperatures are factors that affect the flowering date of *Cerasus* or *Prunus* species, including cherry trees (Fadón et al., 2020). Therefore, chill units are often used in

process-based models to represent the effect of low temperatures in winter on releasing endodormancy (Aono and Moriya, 2003; Aono and Murakami, 2017; Asakura et al., 2010; Maruoka and Itoh, 2009). In addition, the number of days transformed to standard temperature (DTS) (Aono and Moriya, 2003; Aono and Murakami, 2017; Maruoka and Itoh, 2009) and growing degree hours (GDHs) (Asakura et al., 2010) have been utilized to represent the accumulated temperature in spring.

Although our machine learning models did not directly depict these factors in the model structure, the machine learning models generated estimates similar to the process-based models. According to the SHAP values (Figs. 3 and 4) for both the first flowering and full blossom date models, the temperature from February to April exerted the most substantial effect on the output values for all algorithms. The correlation coefficients between the SHAP values and the daily average temperature showed that the higher the daily average temperature during this period was, the earlier the first flowering and full blossom dates. In addition to the spring peak, the RF models and GBDT models also showed small peaks in November and November to January, respectively. The correlation coefficients between the SHAP values and the daily average temperature during these periods were positive, indicating that lower daily average temperatures in early winter led to earlier first flowering and full blossom dates. In the RF and GBDT models, the low temperature in winter and high temperature in spring would shift the estimated first flowering and full blossom dates earlier.

Our models used the daily average temperature for 365 consecutive days from June 1st to May 31st as the explanatory variables. Since the model used temperatures after the first flowering and full blossom dates for estimation, predicting the forthcoming first flowering and full blossom dates in a particular year is impossible using ongoing observations of the daily average temperature. However, these models can assess the impact of climate change on the first flowering and full blossom dates of cherry trees using the projected daily average temperatures from future climate scenarios, while it requires further evaluation on the range of temperature applicable to the models. Cherry blossoms are familiar to Japanese people, and cherry blossom viewing is a springtime tradition. However, due to climate change, the timing of the full blossom date of cherry trees and the timing of cultural events related to cherry blossoms are projected to be mismatched (Nagai et al., 2019; Wang et al., 2017). In addition, since the dates of the first flowering of cherry blossoms are covered in the news every year, assessing the impact of climate change on such phenological events allows citizens to more closely feel the impact of climate change. We believe that the models developed in this study will help increase citizens' awareness of climate change and encourage them to take action against it.

5. Conclusion

In this study, three machine learning algorithms were selected to estimate the first flowering and full blossom dates of Yoshino cherry. We showed that the GBDT models generate more accurate estimates than existing process-based models simply by inputting 365 days of daily average temperatures into the model. We also showed that the three algorithms differ in accuracy, with the GBDT model providing the most accurate and stable estimates.

Technological developments in explainable AI have allowed us to obtain insights into how machine learning models calculate output values, which had previously been considered a black box. Biological studies have shown that endodormancy release by accumulating low winter temperatures and subsequent high temperatures is a biological mechanism of cherry trees blooming. Although this information was not used for GBDT and RF model construction, our SHAP analysis revealed that the low winter temperatures and high spring temperatures caused the estimated first flowering and full blossom dates to occur earlier in the GBDT and RF models. A similar approach can be employed to build models for other phenological events. We propose that the method developed in this study can be used to build estimation models for various phenological events and to assess the impacts of climate change, thereby promoting activities by citizens to mitigate climate change.

CRediT authorship contribution statement

Yoshifumi Masago: Conceptualization, Methodology, Software, Writing – original draft, Project administration, Funding acquisition. **Maychee Lian:** Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Yoshifumi Masago reports financial support was provided by Environmental Restoration and Conservation Agency of Japan.

Data availability

Data will be made available on request.

Acknowledgments

This work was partially supported by the Environment Research and Technology Development Fund (JPMEERF20S11802) of the Environmental Restoration and Conservation Agency of Japan.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zhang, X., 2016. TensorFlow: a system for large-scale machine learning. In: Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, pp. 265–283.

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: A next-generation hyperparameter optimization framework, in. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631. https://doi.org/10.1145/3292500.3330701.

Allen, J.M., Terres, M.A., Katsuki, T., Iwamoto, K., Kobori, H., Higuchi, H., Primack, R.B., Wilson, A.M., Gelfand, A., Silander, J.A., 2014. Modeling daily flowering probabilities: expected impact of climate change on Japanese cherry phenology. Glob. Chang. Biol. 20, 1251–1263. https://doi.org/10.1111/gcb.12364.

Aono, Y., Moriya, C., 2003. A generalized model to estimate flowering for cherry tree (Prunus yedoensis) considering both processes of Endodormancy completion and development. J. Agric. Meteorol. 59, 165–177. https://doi.org/10.2480/agrmet.59.165.

Aono, Y., Murakami, N., 2017. A simplified method to estimate cherry blossom phenology considering temperature during endodormancy process. Clim. Biosphere 17, 25–33. https://doi.org/10.2480/cib.j-17-033.

Asakura, T., Sugiura, H., Sakamoto, D., Sugiura, T., Gemma, H., 2010. A universal model for predicting the full bloom date of Japanese flowering cherry. J. Agric. Meteorol. 66, 269–277. https://doi.org/10.2480/agrmet.66.4.7.

Badeck, F., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., Sitch, S., 2004. Responses of spring phenology to climate change. New Phytol. 162, 295–309. https://doi.org/10.1111/j.1469-8137.2004.01059.x.

Chung, U., Mack, L., Yun, J.I., Kim, S.-H., 2011. Predicting the timing of cherry blossoms in Washington, DC and mid-Atlantic states in response to climate change. PLoS One 6, e27439. https://doi.org/10.1371/journal.pone.0027439.

Chung, U., Kim, J.-H., Kim, K.-H., 2016. Variation and uncertainty in the predicted flowering dates of cherry blossoms using the CMIP5 climate change scenario. Asia-Pac. J. Atmos. Sci. 52, 509–518. https://doi.org/10.1007/s13143-016-0033-9.

Cleland, E.E., Chiariello, N.R., Loarie, S.R., Mooney, H.A., Field, C.B., 2006. Diverse responses of phenology to global changes in a grassland ecosystem. Proc. Natl. Acad. Sci. 103, 13740–13744. https://doi.org/10.1073/pnas.0600815103.

Dai, W., Jin, H., Zhang, Y., Liu, T., Zhou, Z., 2019. Detecting temporal changes in the temperature sensitivity of spring phenology with global warming: application of machine learning in phenological model. Agric. For. Meteorol. 279, 107702 https://doi.org/10.1016/j.agrformet.2019.107702.

Fadón, E., Herrera, S., Guerrero, B.I., Guerra, M.E., Rodrigo, J., 2020. Chilling and heat requirements of temperate stone fruit trees (Prunus sp.). Agronomy 10, 409. https://doi.org/10.3390/agronomy10030409.

Gaira, K.S., Dhar, U., 2020. Phenological change modelling for selected Himalayan medicinal herbs using herbarium records: a case study. Ecol. Inform. 60, 101177 https://doi.org/10.1016/j.ecoinf.2020.101177.

Guo, Y., Fu, Y., Hao, F., Zhang, X., Wu, W., Jin, X., Bryant, C.R., Senthilnath, J., 2021. Integrated phenology and climate in rice yields prediction using machine learning methods. Ecol. Indic. 120, 106935 https://doi.org/10.1016/j.ecolind.2020.106935.

Ibáñez, I., Primack, R.B., Miller-Rushing, A.J., Ellwood, E., Higuchi, H., Lee, S.D., Kobori, H., Silander, J.A., 2010. Forecasting phenology under global warming. Philos. Trans. R. Soc. B Biol. Sci. 365, 3247–3260. https://doi.org/10.1098/rstb.2010.0120.

Iketani, Hiroyuki, Ohta, Satoshi, Kawahara, Takayuki, Katsuki, Toshio, Mase, Nobuko, Sato, Yoshihiko, Yamamoto, Toshiya, 2007. Analyses of clonal status in 'Somei-yoshino' and confirmation of genealogical record in other cultivars of *Prunus* × *yedoensis* by microsatellite markers. Breed. Sci. 57, 1–6. https://doi.org/10.1270/jsbbs.57.1.

Innan, H., Terauchi, R., Miyashita, N.T., Tsunewaki, K., 1995. DNA fingerprinting study on the intraspecific variation and the origin of *Prunus yedoensis* (Someiyoshino). Jpn. J. Genet. 70, 185–196. https://doi.org/10.1266/jjg.70.185.

Japan Meteorological Agency, 2021. Information on Phenological Observation. URL. https://www.data.jma.go.jp/sakura/data/index.html (accessed 11.15.2021).

Kang, Y., Ozdogan, M., Zhu, X., Ye, Z., Hain, C., Anderson, M., 2020. Comparative assessment of environmental variables and machine learning algorithms for maize yield prediction in the US Midwest. Environ. Res. Lett. 15, 064005 https://doi.org/10.1088/1748-9326/ab7df9.

Ke, G., Meng, Q., Finely, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree.

Lundberg, S., Lee, S.-I., 2017. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, CA, USA, pp. 4768–4777.

Maruoka, T., Itoh, H., 2009. Impact of global warming on flowering of cherry trees (Prunus yedoensis) in Japan. J. Agric. Meteorol. 65, 283–296. https://doi.org/10.2480/agrmet.65.3.5.

Miller-Rushing, A.J., Katsuki, T., Primack, R.B., Ishii, Y., Lee, S.D., Higuchi, H., 2007. Impact of global warming on a group of related species and their hybrids: cherry tree (Rosaceae) flowering at Mt. Takao, Japan. Am. J. Bot. 94, 1470–1478. https://doi.org/10.3732/ajb.94.9.1470.

Nagai, S., Saitoh, T.M., Yoshitake, S., 2019. Cultural ecosystem services provided by flowering of cherry trees under climate change: a case study of the relationship between the periods of flowering and festivals. Int. J. Biometeorol. 63, 1051–1058. https://doi.org/10.1007/s00484-019-01719-9.

Nagai, S., Morimoto, H., Saitoh, T.M., 2020. A simpler way to predict flowering and full bloom dates of cherry blossoms by self-organizing maps. Ecol. Inform. 56, 101040 https://doi.org/10.1016/j.ecoinf.2019.101040.

Oses, N., Azpiroz, I., Marchi, S., Guidotti, D., Quartulli, M., Olaizola, I.G., 2020a. Analysis of Copernicus' ERA5 climate reanalysis data as a replacement for Weather Station temperature measurements in machine learning models for olive phenology phase prediction. Sens. Basel Switz. 20, 6381. https://doi.org/10.3390/s20216381.

Oses, N., Azpiroz, I., Quartulli, M., Olaizola, I., Marchi, S., Guidotti, D., 2020b. Machine learning for olive phenology prediction and base temperature optimisation. 2020 Glob. Intern. Things Summit Giots 00, 1–6. https://doi.org/10.1109/giots49054.2020.9119611.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

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, 1922–1940. https://doi.org/10.1111/gcb.14619.

Primack, R.B., Higuchi, H., Miller-Rushing, A.J., 2009. The impact of climate change on cherry trees and other species in Japan. Biol. Conserv. 142, 1943–1949. https://doi.org/10.1016/j.biocon.2009.03.016.

Richardson, A.D., Keenan, T.F., Migliavacca, M., Ryu, Y., Sonnentag, O., Toomey, M., 2013. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. Agric. For. Meteorol. 169, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012.

Tomczyk, A.M., Szyga-Pluta, K., 2019. Variability of thermal and precipitation conditions in the growing season in Poland in the years 1966–2015. Theor. Appl. Climatol. 135, 1517–1530. https://doi.org/10.1007/s00704-018-2450-4.

Vitasse, Y., François, C., Delpierre, N., Dufrêne, E., Kremer, A., Chuine, I., Delzon, S., 2011. Assessing the effects of climate change on the phenology of European temperate trees. Agric. For. Meteorol. 151, 969–980. https://doi.org/10.1016/j.agrformet.2011.03.003.

Wang, C., Tang, Y., 2019. Responses of plant phenology to nitrogen addition: a meta-analysis. Oikos 128, 1243–1253. https://doi.org/10.1111/oik.06099.

Wang, L., Ning, Z., Wang, H., Ge, Q., 2017. Impact of climate variability on flowering phenology and its implications for the schedule of blossom festivals. Sustainability 9, 1127. https://doi.org/10.3390/su9071127.