Journal of Apiculture
2026.41(2) : 195-210
DOI: 10.17519/apiculture.2026.06.41.2.195

Original research article

Multi-site Cross-validation of a Flowering Onset Prediction Model for *Robinia pseudoacacia*

Jun-ho Oh and Kye-Han Lee*

Department of Forest Resources, Chonnam National University, Gwangju 61186, Republic of Korea

Abstract Changes in flowering times caused by climate change deeply affect ecosystems and the beekeeping industry. *Robinia pseudoacacia* (black locust) is one of the most important honey plants in South Korea. Reliable models are needed to predict flowering onset across different sites. This study compared three predicting models with different levels of complexity — Chill Days (CD), Normal Heat Hours (NHH), and PhenoFlex — using long-term weather and flowering data (1973–2024) from seven major cities in South Korea. The results showed that making a model more complex does not always guarantee better prediction accuracy. The CD model, which uses only three parameters, was the most stable for predicting flowering in individual sites, with an average RMSE of 2.78 days. In contrast, when data from all sites were combined, the PhenoFlex model, which uses 12 parameters, showed a mean bias of 0.08 days, indicating little systematic over- or underprediction. We also observed regional differences in model performance. The CD model tended to perform better in southern sites, whereas the NHH and PhenoFlex models tended to perform better in central and northern sites. In conclusion, the CD model showed stable performance for local-scale flowering prediction, whereas the PhenoFlex model showed low systematic bias when data from all locations were combined. These results suggest that model selection should consider the spatial scale and purpose of prediction. This study suggests that model selection should consider regional differences in prediction performance and the spatial scale of application.

Keywords Climate change, Flowering prediction, Multi-site cross-validation, Phenological model

INTRODUCTION

Changes in plant phenology driven by climate change, particularly shifts in flowering timing, are recognized as the most sensitive indicators for capturing ecological impacts of global warming (Cleland *et al*., 2007; Piao *et al*., 2019). Large-scale observational studies have documented a progressive advancement of spring flowering dates (Menzel *et al*., 2006), and the finding that artificial warming experiments may underestimate such phenological sensitivity highlights the critical importance of analyses grounded in long-term observational records (Wolkovich *et al*., 2012). Moreover, interspecific phenological mismatch can exert cascading effects on plant-pollinator interactions and ecosystem functioning (Thackeray *et al*., 2016), making the quantitative analysis of regional phenological variability an indispensable undertaking for climate signal detection and ecological sensitivity assessment.

Black locust (*Robinia pseudoacacia* L.) has attracted considerable attention as a focal species in climate change adaptation research, owing to its close association with the domestic apiculture industry in South Korea. As the nation's primary nectar-source tree species, black locust accounts for more than 70% of total honey production (Jung and Cho, 2015; Kim *et al*., 2021a). In 2018 and 2020, anomalous weather events coinciding with the flowering period caused honey yields to plum-

http://journal.bee.or.kr/
Received 13 April 2026; Revised 25 May 2026; Accepted 25 May 2026
*Corresponding author. E-mail: khl@jnu.ac.kr

met to 19.1% and 12.9% of the normal-year average, respectively (Kim *et al*., 2021a). More recently, rising spring temperatures have advanced flowering dates and shortened the spring flowering duration of Korean trees, with potential implications for plant communities and beekeeping management (Kim and Jung, 2025). The growing risk of phenological mismatch between flowering and pollinator activity under climate change (Stemkovski *et al*., 2020) further underscores the need to establish a reliable, regionally resolved flowering prediction framework.

Phenological prediction for temperate woody species is fundamentally rooted in process-based models that quantify chilling accumulation during endodormancy and forcing (heat) accumulation during ecodormancy. The Chill Days (CD) model employs daily maximum and minimum temperatures together with threshold temperatures to sequentially determine the fulfillment of chilling and forcing requirements. Although its modest data requirements confer broad applicability, the model is limited by the fact that threshold temperatures are not fixed but vary with species and regional climate conditions (Cesaraccio *et al*., 2004). The Normal Heat Hours (NHH) model captures, on an hourly basis, the non-linear developmental response in which the rate of progress declines above an optimum temperature, parameterized by four cardinal temperatures (Cola *et al*., 2014; Alilla *et al*., 2022); however, determining the onset date for effective temperature accumulation remains a persistent challenge. To address this issue, the present study proposes a methodology that integrates Partial Least Squares (PLS) regression (Luedeling and Gassner, 2012; Luedeling *et al*., 2013) as a preprocessing technique coupled with the NHH model. The PhenoFlex model provides a unified framework that combines chilling accumulation based on the Dynamic Model with forcing accumulation based on Growing Degree Hours, while flexibly mediating the transition between the two processes (Luedeling *et al*., 2021). Owing to its structural complexity—requiring simultaneous estimation of numerous parameters—convergence to local optima and the derivation of biologically implausible solutions have been reported; an Enhanced Global Optimization approach, featuring a restructured parameter space and an improved scatter search algorithm, was proposed to overcome these limitations (Mojahid *et al*., 2025).

The Korean Peninsula exhibits substantial spatial variability in climatic conditions associated with latitudinal gradients, elevation differences, complex terrain, and interactions between continental and maritime air masses (Park *et al*., 2019; Kim *et al*., 2025b). In addition, urban heat island (UHI) and urbanization-related thermal effects can further modify plant phenological patterns by altering local thermal conditions, including in Seoul, where urban-outskirt temperature differences were associated with shifts in Mongolian oak leaf unfolding and cherry flowering dates (Jochner and Menzel, 2015; Zipper *et al*., 2016; Kim *et al*., 2024). Domestic cherry blossom prediction studies have confirmed that prediction errors intensify with increasing spatial scale when microclimatic variation is not accounted for (Chung *et al*., 2009). Consequently, establishing a nationwide flowering prediction system necessitates multi-regional comparison and validation that simultaneously accounts for regional climate gradients and urban effects.

The objective of this study is to develop flowering prediction models for *R. pseudoacacia* using long-term meteorological and phenological observation records (1973–2024) from seven major metropolitan cities across South Korea (Seoul, Incheon, Daejeon, Daegu, Gwangju, Ulsan, and Busan), and to compare the performance and regional applicability of the CD, NHH, and PhenoFlex models using prediction accuracy, error distribution, and city-specific model stability.

MATERIALS AND METHODS

1. Site description and data collection

Bud-burst and flowering date records from observation stations operated by the Korea Meteorological Administration (KMA) were compiled for parameter estimation and validation of the black locust (*Robinia pseudoacacia* L.) flowering prediction models. Although the overall analytical scope spans from 1973 to 2024, the effective period applied in the analysis was defined independently for each station in accordance with the commencement date of observations and data availability (Table 1, Fig. 1). Daily minimum and maximum

Table 1. Observation periods and statistics of bud-burst and flowering dates (Julian day) for *Robinia pseudoacacia* at study site

Region (Site)	Observation period	*N* (years)	Bud-burst date (Julian day)	Flowering date (Julian day)
			MEAN ± STD	MEAN ± STD
Seoul	1975–2024	50	108.51 ± 9.50	131.20 ± 6.67
Incheon	1973–2024	52	114.35 ± 9.16	136.27 ± 6.70
Daejeon	1990–2024	35	102.97 ± 5.86	125.53 ± 5.57
Daegu	1982–2024	43	96.10 ± 7.76	122.00 ± 5.99
Ulsan	1988–2024	37	97.24 ± 7.13	126.08 ± 5.82
Gwangju	1979–2024	46	95.96 ± 6.41	123.27 ± 6.46
Busan	1984–2024	41	95.36 ± 6.75	124.81 ± 6.31

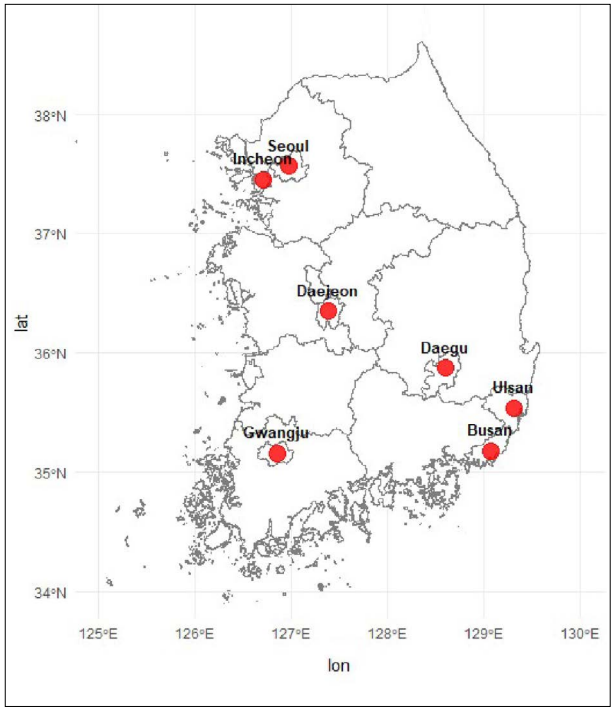

Fig. 1. Location map of Korea Meteorological Administration (KMA) ASOS stations representing the study regions for temperature-based modelling of bud-burst and flowering dates of *Robinia pseudoacacia*. Monitoring sites for bud-burst and blooming dates of *Robinia pseudoacacia*.

air temperature data corresponding to the bud-burst and flowering observation sites were obtained from the Automated Surface Observing System (ASOS) and the KMA seasonal phenological observation records. Both bud-burst and flowering dates were converted to Julian day of year (DOY) for use as model input variables.

2. Phenological models

1) Chill days model

The Chill Days (CD) model is a process-based model grounded in the dormancy physiology of temperate deciduous trees, distinguishing between two successive phases: endodormancy and ecodormancy (Cesaraccio *et al*., 2004; Jung *et al*., 2005; Kim *et al*., 2021b) (Fig. 2). The model computes chill days (C_d) and anti-chill days (C_a) through five conditional equations based on the relationship between daily minimum and maximum air temperatures and a threshold temperature (T_c) (Cesaraccio *et al*., 2004, 2005) (Table 2). When the cumulative chill day sum ($\sum C_d$) reaches the chilling requirement (C_R), endodormancy is released. Subsequently, ecodormancy is released and bud-burst is initiated at the point when the cumulative anti-chill day sum ($\sum C_a$) equals $|C_d|$. The date at which $\sum C_a$ reaches the heat requirement (H_R) following bud-burst is defined as the predicted flowering date (Cesaraccio *et al*., 2004, 2005; Jung *et al*., 2005).

The threshold temperature (T_c) and chilling requirement (C_R) were constrained to the ranges $5°C \leq T_c \leq 10°C$, $-300 \leq C_R \leq -70$, respectively, based on values reported in the literature (Cesaraccio *et al*., 2004; Jung *et al*., 2005; Kim *et al*., 2013; Chun *et al*., 2017). The optimal parameter combination was identified via an iterative procedure that minimized the root mean square error (RMSE) between predicted and observed bud-burst dates (Chun *et al*., 2017). The heat requirement (H_R) was estimated on an annual basis using the selected T_c and C_R following the approach of Jung *et al*. (2005), and the mean value across years was adopted as the fi-

nal estimate.

For model implementation, the dormancy onset date was fixed at 1 October of the preceding year (Jung *et al*., 2005). To prevent overfitting, the full observational dataset was randomly partitioned into training (approximately 60−75%) and validation (approximately 25−40%) subsets. Only splits in which the Wilcoxon rank-sum test indicated no statistically significant difference ($p > 0.05$) in the flowering date distributions between the two subsets were accepted as the final partition (Kim *et al*., 2021b).

2) Normal Heat Hours (NHH) model

The Normal Heat Hours (NHH) model was developed to overcome the limitation of the linear temperature-development rate relationship assumed by conventional growing degree day (GDD) models, by explicitly accounting for the inhibitory effect of supraoptimal temperatures on developmental rate (Wang and Engel, 1998; Mariani *et al*., 2013). The model employs a trapezoidal response function defined by four cardinal temperatures, which converts hourly air temperature into NHH values ranging from 0 to 1. NHH equals 0 below the lower cardinal temperature (LC) and above the upper cardinal temperature (UC), NHH equals 1 between the lower optimum temperature (LOC) and the upper optimum temperature (UOC). NHH increases or decreases linearly within the LC-LOC and UOC-UC intervals, respectively (Cola *et al*., 2017; Alilla *et al*., 2022).

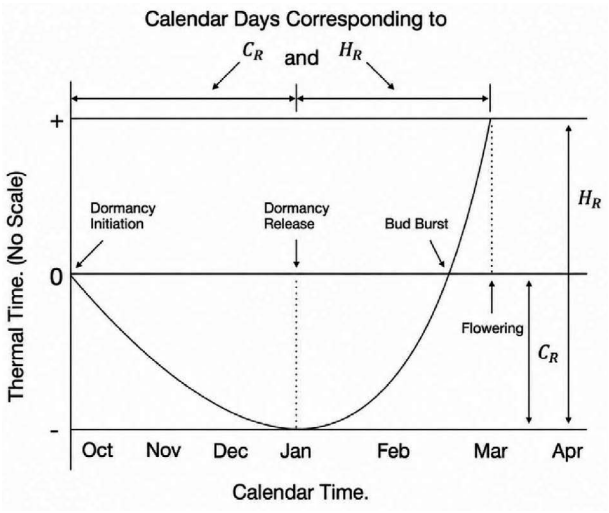

Fig. 2. Illustration of thermal time accumulation for dormancy release and flowering. The curve represents the accumulation of chill units until the chilling requirement (C_R) is met, followed by the accumulation of heat units until the heating requirement (H_R) is reached (Modified from Jung *et al*., 2005).

Since the meteorological dataset used in this study comprises only daily maximum and minimum air temperatures, hourly temperature series were generated by applying the method of Parton and Logan (1981), which employs a sine curve during daytime and an exponential decay curve during nighttime. The potential sunshine duration was calculated from the latitude of each observation site and the solar declination angle.

Model calibration followed the framework of Alilla *et al*. (2022), with partial modifications to accommodate the use of a single annual flowering date per year in the present study. Whereas the original study calibrated the model through regression analysis using multiple phenological stages (BBCH 51-69), the present study defined the flowering threshold as the median cumulative NHH accumulated up to the observed flowering date across training years, and the date on which this threshold was reached was designated as the predicted

Table 2. Calculation scheme of the chill day (C_d) and anti-chill day (C_a) model (Cesaraccio *et al*., 2004, 2005)

Number	Temperature cases	Chill days	Anti-chill days
1	$0 \le T_c \le T_n \le T_x$	$C_d = 0$	$C_a = T_M - T_c$
2	$0 \le T_n \le T_c \le T_x$	$C_d = -\left[(T_x - T_n)] - \frac{(T_x - T_c)^2}{2(T_x - T_n)}\right]$	$C_a = \frac{(T_x - T_c)^2}{2(T_x - T_n)}$
3	$0 \le T_n \le T_x \le T_c$	$C_d = -(T_M - T_n)$	$C_a = 0$
4	$T_n < 0 \le T_x \le T_c$	$C_d = -\left[\frac{T_x^2}{2(T_x - T_n)}\right]$	$C_a = 0$
5	$T_n < 0 < T_c < T_x$	$C_d = -\frac{T_x^2}{2(T_x - T_n)} - \frac{(T_x - T_c)^2}{2(T_x - T_n)}$	$C_a = \frac{(T_x - T_c)^2}{2(T_x - T_n)}$

T_x: Daily maximum temperature, T_n: Daily minimum temperature, T_c: Threshold temperature, T_M: Mean daily temperature

flowering date.

Calibration was conducted in two stages. In the first stage, Partial Least Squares (PLS) regression was employed to objectively determine the onset date for NHH accumulation (Luedeling and Gassner, 2012; Luedeling *et al*., 2013). An 11-day moving average of daily mean temperature was used as the predictor variable to suppress short-term weather noise while preserving the seasonal dormancy signal, and the flowering date (Julian day of year) served as the response variable. The analysis spanned 1 July of the preceding year to 30 June of the target year, covering a full phenological cycle from endodormancy induction to bloom, which allowed the PLS procedure to objectively detect the chilling-to-forcing transition (Luedeling and Gassner, 2012; Luedeling *et al*., 2013). Based on the resulting Variable Importance in Projection (VIP) scores and standardized regression coefficients, a continuous period in which VIP≥0.8 and rising temperatures exerted a promoting effect on flowering was defined as the forcing phase, and its start date was fixed as the onset date for NHH accumulation.

In the second stage, a grid search was performed over the four cardinal temperature parameters (LC, LOC, UOC, and UC) using the fixed accumulation onset date established in the first stage. The search ranges, following Alilla *et al*. (2022), were set as LC (0−11°C), LOC (14−19°C), UOC (20−25°C), and UC (26−31°C), with a grid resolution of 1°C and the constraint LC＜LOC＜UOC＜UC enforced throughout (Table 3). For each parameter combination, the median cumulative NHH across training years was defined as the flowering threshold, and the combination yielding the minimum RMSE relative to observed flowering dates was selected as the optimal parameter set. This two-stage procedure provides an independent statistical basis for determining the accumulation onset date, thereby mitigating the risk of overfitting during the grid search process. The training-validation data splitting procedure followed the same Wilcoxon rank-sum test-based protocol applied in the CD model (Kim *et al*., 2021b).

Table 3. Parameter ranges tested for model calibration (Modified from Alilla *et al*., 2022)

Parameters	Description	Range
DOY	Starting date	Fixed by PLS
LC	Lower cardinal	0−11°C
LOC	Lower optimal cardinal	14−19°C
UOC	Upper optimal cardinal	20−25°C
UC	Upper cardinal	26−31°C

3) PhenoFlex model

The PhenoFlex model (Luedeling *et al*., 2021) is an integrated framework combining the Dynamic Model (DM), which describes chilling accumulation, and the Growing Degree Hours (GDH) model, which describes heat (forcing) accumulation. The model is defined by a total of 12 parameters: six parameters associated with the DM (A_0, A_1, E_0, E_1, T_f and *slope*), three parameters associated with the GDH model (T_b, T_u and T_c), and three parameters mediating the transition between the two sub-models (y_c, s_1 and z_c). The mathematical algorithms and equations of the PhenoFlex model applied in this study follow those described in Luedeling *et al*. (2021).

Since the PhenoFlex model requires hourly temperature as input, hourly temperature series were generated from daily maximum and minimum temperatures following the method of Linvill (1990), which simulates daytime temperature using a sine curve and nocturnal cooling using a logarithmic decay function (Luedeling *et al*., 2021). Approximately 80% of the available years for each city were allocated to calibration, with the remaining 20% reserved for validation. Parameter bounds were adopted from Mojahid *et al*. (2025) (Table 4). Initial parameter values for the DM parameters were taken from relevant prior studies (Anderson *et al*., 1985; Fishman *et al*., 1987a, 1987b; Erez *et al*., 1989), while those for the transition parameters were set according to the values recommended by Luedeling *et al*. (2021).

To overcome the convergence instability of conventional global optimization approaches arising from inter-parameter scale disparities and susceptibility to local optima, this study adopted the Enhanced Global Optimization method proposed by Mojahid *et al*. (2025). The method can be summarized by three key features.

First, the activation energy (E) and amplitude (A) parameters of the original DM were replaced with intermediate parameters possessing clear physical interpretations (θ^*, θ_c, π_c and τ), thereby constraining the search space to a physiologically meaningful domain. For each derived parameter combination, Q_{10} values were

computed and restricted to the range 1.5–3.5 to verify biological plausibility.

Second, the MEIGO toolbox (Egea *et al*., 2014) was employed as the optimization algorithm. MEIGO operates through a structure in which the Enhanced Scatter Search metaheuristic drives global exploration, supplemented by local solvers such as Dynamic Hill Climbing.

Third, the optimization was run independently 10 times (each with 30,000 iterations), after which a weighted ensemble prediction approach was applied to prevent overfitting associated with reliance on a single optimal solution. The final predicted values were derived by weighting the predicted outputs of each run by the Ratio of Performance to Inter-Quartile distance (RPIQ) computed on the validation dataset (Mojahid *et al*., 2025).

4) Model evaluation

Model performance was assessed using three common evaluation metrics: root mean square error (RMSE), coefficient of determination (R^2), and bias. For the PhenoFlex model, the Ratio of Performance to Interquartile Range (RPIQ) was additionally calculated for use in the weighted ensemble prediction, with the interquartile range (IQR) computed based on the full dataset (Mojahid *et al*., 2025).

$$\text{Bias} = \frac{\sum_{i=1}^{n}(P_i - O_i)}{n}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}}$$

$$\text{RPIQ} = \frac{Q_3 - Q_1}{SE}$$

N: Number of observation years during the analysis period
O_i: Observed flowering date in the i-th year
P_i: Predicted flowering date in the i-th year
Q_1: First quartile of observed values (25th percentile)
Q_3: Third quartile of observed values (75th percentile)

Table 4. Parameter bounds and initial values used in the PhenoFlex Enhanced Global Optimization (EGO) based on Mojahid *et al*. (2025)

Parameter	Lower bound	Upper bound	Initial values
y_c	10	90	40
z_c	30	500	190
s (1/K)	0.1	1.0	0.5
T_f (°C)	0	10	4
slope	0.5	50	1.6
T_b (°C)	0	10	4
T_c (°C)	0	40	36
T_u (°C)	0	30	25
θ*	276	281	279
θ_c	284	287	286
π_c	24	50	28
τ	16	48	35

RESULTS

1. Chill days model

The optimal parameters and validation performance of the CD model are presented in Table 5. The threshold temperature (T_c) ranged from 5.0°C (Gwangju) to 8.4°C (Seoul), while the chilling requirement (C_R) ranged from −108 (Seoul) to −199 (Ulsan), with cities located farther south tending to exhibit larger absolute values of C_R. Validation RMSE of the CD model ranged from 1.95 days (Busan) to 4.28 days (Incheon), with a mean RMSE of 2.78 days across the seven cities. The coeffi-

Table 5. Estimated parameters and goodness-of-fit measures for the Chill Days (CD) model across 7 cities

Region (Site)	Period	T_c	C_R	H_R	RMSE (day)	Bias (day)	R^2
Seoul	1975–2024	8.4	−108	252.09	2.63	−0.08	0.81
Incheon	1976–2024	6.0	−163	248.88	4.28	2.50	0.73
Daejeon	1990–2024	7.5	−130	245.73	2.40	0.89	0.83
Daegu	1976–2024	6.4	−159	286.70	3.67	2.10	0.77
Ulsan	1988–2024	5.2	−199	293.58	2.06	0.22	0.84
Gwangju	1980–2024	5.0	−172	268.83	2.45	0.44	0.87
Busan	1984–2024	6.5	−169	304.63	1.95	−0.40	0.88

Table 6. Estimated parameters and goodness-of-fit measures for the Normal Heat Hours (NHH) model across 7 cities

Region (Site)	DOY	LC (°C)	LOC (°C)	UOC (°C)	UC (°C)	NHH	RMSE	Bias	R^2
Seoul	79	7	19	25	29	626.8	2.43	0.85	0.82
Incheon	92	11	18	24	26	441.3	4.29	0.38	0.53
Daejeon	70	5	18	25	31	704.6	3.90	−0.09	0.58
Daegu	61	10	19	25	27	497.4	3.30	−1.30	0.66
Ulsan	77	11	19	25	30	447.6	2.98	0.00	0.60
Gwangju	87	11	19	24	31	377.7	4.96	0.86	0.52
Busan	45	7	17	23	31	910.4	2.39	−0.10	0.72

Table 7. Performance evaluation of the PhenoFlex model across seven cities

Region (Site)	RMSE	Bias	R^2
Seoul	4.23	0.05	0.609
Incheon	3.90	0.36	0.673
Daejeon	2.73	−0.02	0.771
Daegu	3.18	0.05	0.717
Ulsan	2.95	−0.04	0.746
Gwangju	3.38	−0.03	0.731
Busan	3.80	−0.03	0.644

cient of determination (R^2) ranged from 0.73 (Incheon) to 0.88 (Busan), indicating generally high explanatory power. Notably, the model achieved the best predictive performance in Busan (RMSE = 1.95 days, R^2 = 0.88), Ulsan (RMSE = 2.06 days, R^2 = 0.84), and Daejeon (RMSE = 2.40 days, R^2 = 0.83). In contrast, Incheon exhibited comparatively poor performance, characterized by the highest RMSE (4.28 days) and a pronounced positive bias (Bias = +2.50 days). Positive biases were observed in most cities except Seoul and Busan, suggesting a systematic tendency of the CD model to predict flowering dates later than observed. Seoul (Bias = −0.08 days) and Busan (Bias = −0.40 days), however, showed negligible or marginally early predictions.

2. Normal Heat Hours (NHH) model

The optimal parameters and validation results of the NHH model are presented in Table 6. The onset date of forcing accumulation (DOY) was earliest in Busan at DOY 45 (mid-February) and latest in Incheon at DOY 92 (early April), revealing an inter-city difference of approximately 47 days. Among the parameters defining the temperature response curve, the optimum temperature range (LOC−UOC) exhibited relatively consistent values across most cities, ranging from 17−19°C to 23−25°C. In contrast, the lower cardinal temperature (LC) showed substantial inter-city variability, spanning from 5°C (Daejeon) to 11°C (Incheon, Ulsan, and Gwangju). The cumulative NHH requirement varied considerably among cities, ranging from 377.7 (Gwangju) to 910.4 (Busan). The notably high NHH requirement observed in Busan is attributable to the low ambient temperatures prevailing during the early forcing period, as the accumulation onset in this city begins as early as DOY 45.

In terms of validation performance, the RMSE of the NHH model ranged from 2.39 days (Busan) to 4.96 days (Gwangju), with a seven-city mean RMSE of 3.46 days. The coefficient of determination (R^2) ranged from 0.52 (Gwangju) to 0.82 (Seoul), indicating an overall lower explanatory power compared to the CD model. Satisfactory predictive performance was achieved in Seoul (RMSE = 2.43 days, R^2 = 0.82) and Busan (RMSE = 2.39 days, R^2 = 0.72); however, markedly diminished accuracy was observed in Gwangju (RMSE = 4.96 days, R^2 = 0.52) and Incheon (RMSE = 4.29 days, R^2 = 0.53). The bias remained relatively small for most cities, ranging from −1.30 days (Daegu) to +0.86 days (Gwangju), suggesting minimal systematic error. Nevertheless, a moderately pronounced negative bias in Daegu indicated that the model tended to predict flowering dates earlier than the observed values.

3. PhenoFlex model

The estimated 12 parameters and validation performance of the PhenoFlex model are presented in Table 7 and Fig. 3. Among the chilling-related parameters, the

Fig. 3. Estimated parameters of the PhenoFlex model across seven cities.

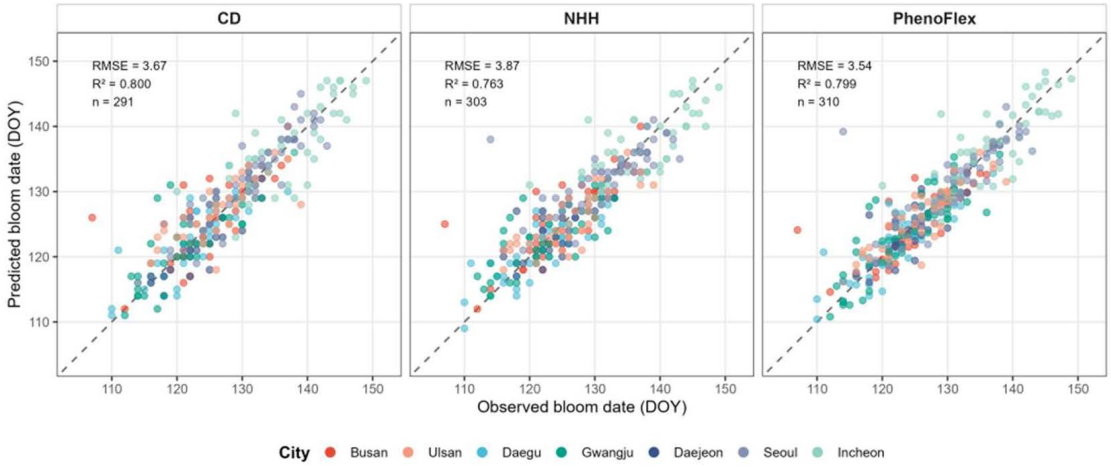

Fig. 4. Observed versus predicted flowering dates of *Robinia pseudoacacia* for the CD, NHH, and PhenoFlex models across seven cities in South Korea. The dashed line represents a perfect match (1 : 1 line) between the observed and predicted dates. Colors indicate the respective cities.

critical chilling portion (y_c) was lowest in Seoul (25.83 ± 6.04) and highest in Ulsan (51.64 ± 13.83), suggesting that the chilling requirement is fulfilled with a relatively smaller accumulation in Seoul. The critical forcing accumulation (z_c) ranged from 313.15 (Ulsan) to 444.18 (Daejeon).

The activation energy parameters, E_0 (5,590.89–6,258.64 K) and E_1 (7,368.14–8,660.04 K), remained at relatively comparable levels across cities. In contrast, the frequency factors (A_0 and A_1) exhibited substantial inter-city variability, spanning orders of magnitude of 10^8–10^{10} and 10^{11}–10^{14}, respectively. Their associated standard deviations were also large, indicating limited convergence stability in the optimization results. Regarding the forcing-related parameters, the base temperature (T_b) ranged from 1.71°C (Daejeon) to 6.32°C (Gwangju), remaining within a relatively low range across all cities. The upper critical temperature (T_c) ranged from 25.55°C (Incheon) to 36.83°C (Daejeon), and the optimum temperature (T_u) from 20.01°C (Incheon) to 29.76°C (Daejeon), such that the relationship $T_b < T_u < T_c$ held consistently across all cities. Incheon and Ulsan displayed relatively low values of T_c (25.55°C and 26.59°C, respectively) and T_u (20.01°C and 20.02°C, respectively), reflecting a narrow forcing response range. Conversely, Daejeon (T_c = 36.83°C, T_u = 29.76°C) and Seoul (T_c = 34.60°C, T_u = 27.77°C) exhibited a considerably wider forcing response range.

The validation results showed that the RMSE of the PhenoFlex model ranged from 2.73 days (Daejeon) to 4.23 days (Seoul), with a seven-city mean RMSE of 3.45 days. The bias remained within ±0.36 days for most cities—the smallest among the three models evaluated—demonstrating well-balanced predictions with minimal systematic error. The coefficient of determination (R^2) was reasonably high in Daejeon (0.771), Ulsan (0.746), Gwangju (0.731), and Daegu (0.717), whereas Seoul (0.609) and Busan (0.644) exhibited comparatively lower explanatory power. Notably, Daejeon recorded the best validation performance among all seven cities across all three metrics: RMSE (2.73 days), bias (−0.02 days), and R^2 (0.771). In contrast, Seoul exhibited the highest RMSE (4.23 days) and the lowest R^2 (0.609), indicating that the PhenoFlex model was unable to adequately capture the interannual variability in flowering dates in Seoul.

4. Comparison of prediction performance and spatial characteristics among models

Fig. 4 presents scatter plots of observed versus predicted bloom dates for the three models. The CD model (n = 291) showed data points clustered around the 1 : 1 line, with an overall RMSE of 3.67 days and R^2 = 0.800. A slight tendency toward overprediction was observed in the early-blooming range (around DOY 110–120)

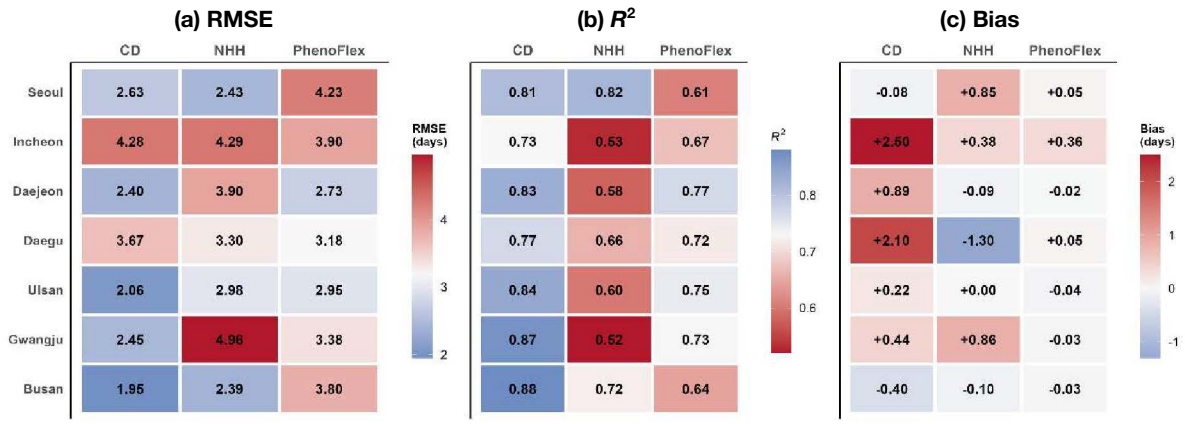

Fig. 5. Heatmap comparing the validation performance of the CD, NHH, and PhenoFlex models across seven metropolitan cities: (a) RMSE (days), (b) R^2, and (c) Bias (days). Cities on the y-axis are arranged by latitude from north (Seoul) to south (Busan). In panels (a) and (c), blue indicates better performance (lower RMSE or near-zero Bias), while red indicates poorer performance. In panel (b), the color scale is reversed so that blue corresponds to higher R^2 values.

across several cities.

The NHH model (n = 303) yielded an RMSE of 3.87 days and $R^2 = 0.763$, comparable to those of the CD model. A somewhat wider scatter around the 1 : 1 line was apparent in the DOY 120–140 range, which may reflect greater sensitivity of the NHH model to temperature variability during the forcing period.

The PhenoFlex model (n = 310) achieved the lowest overall RMSE (3.54 days) with $R^2 = 0.799$, similar to that of the CD model. Among the three models, PhenoFlex showed the least systematic deviation from the 1 : 1 line, indicating relatively smaller bias across specific bloom date ranges.

Fig. 5 compares the validation results by city and model in a heatmap format. All three models performed well in Daejeon and Ulsan, whereas Incheon exhibited the highest RMSE across all models (CD: 4.28 days, NHH: 4.29 days, PhenoFlex: 3.90 days). The optimal model varied by city: the CD model recorded the lowest RMSE in four cities—Busan (1.95 days), Ulsan (2.06 days), Daejeon (2.40 days), and Gwangju (2.45 days)—while PhenoFlex performed best in Incheon (3.90 days) and Daegu (3.18 days), and NHH in Seoul (2.43 days). The superiority of the CD model was particularly evident in southern cities (Busan, Ulsan, and Gwangju), whereas NHH or PhenoFlex showed relatively better performance in central and northern cities (Seoul and Incheon).

The mean RMSE across the seven cities ranked in the order of CD (2.78 days), PhenoFlex (3.45 days), and NHH (3.46 days), and the mean R^2 followed the same ranking: CD (0.82), PhenoFlex (0.70), and NHH (0.63), indicating that the CD model achieved the highest overall explanatory power. In contrast, PhenoFlex recorded the lowest pooled RMSE, which can be attributed to its more uniform inter-city RMSE range (2.73–4.23 days) compared to that of CD (1.95–4.28 days) and its smaller bias, allowing errors to offset one another when data were aggregated. In terms of bias, the mean absolute bias across the seven cities was 0.08 days for PhenoFlex, markedly smaller than those of CD (0.95 days) and NHH (0.51 days).

DISCUSSION

1. Model structure and prediction performance

This study compared three process-based models of differing structural complexity under identical conditions, demonstrating that increasing model complexity does not necessarily lead to improved prediction performance. The CD model, with only three parameters, achieved the best city-averaged RMSE and R^2, whereas the 12-parameter PhenoFlex model was the most stable in terms of pooled RMSE and bias. This indicates that the CD model delivers robust predictive performance against meteorological variability at individual city scales, while PhenoFlex effectively controls

systematic bias in region-wide aggregation by continuously simulating the chilling-to-forcing transition.

The finding that an increased number of parameters does not guarantee superior prediction performance is corroborated by previous studies. Yang *et al*. (2024) reported that a single-parameter GDD model exhibited predictive accuracy comparable to that of nonlinear complex models for rice maturity prediction. Mo *et al*. (2023) confirmed that simpler models yielded performance similar to that of complex models in a comparison of 17 spring phenology models. García-Gutiérrez and Meza (2023) also reported that a two-parameter model outperformed a ten-parameter model based on the BIC criterion for grapevine. The consistently good performance of the CD model in the present study over long-term datasets spanning 35−52 years is also attributable to the lower risk of overfitting associated with a limited number of parameters.

However, caution is warranted regarding the equifinality problem in complex models such as PhenoFlex, wherein multiple parameter combinations can produce identical outputs (Luo *et al*., 2009). Luedeling *et al*. (2021) noted that the 12 parameters of PhenoFlex pose risks of overfitting and derivation of physiologically unrealistic values. Kamimori and Hosomi (2025) reported unrealistically wide effective temperature ranges estimated for grapevine, and Fernandez *et al*. (2022) demonstrated that thermal response curves varied drastically depending on the temperature range of the calibration data. Therefore, parameters derived from complex models should be interpreted not as absolute physiological values of the trees, but rather as mathematically optimized combinations for achieving optimal predictions under given meteorological conditions.

2. Climatological interpretation of city-specific prediction performance

The spatial heterogeneity in city-specific prediction performance is considered attributable to differences in temperature variability and climatic characteristics among regions. Spring temperature is the most critical factor explaining bloom timing, and a consistent tendency for earlier flowering with rising temperatures has been documented (Pearse *et al*., 2023; Kim *et al*., 2025a).

Incheon, which exhibited the highest RMSE across all three models, is a coastal city directly facing the Yellow Sea, where amplified temperature variability driven by land-sea breeze circulation is considered the primary source of prediction error (Cerlini *et al*., 2022; Guralnick *et al*., 2024). Temperature fluctuations induced by maritime influence can accumulate errors in thermal time calculations within single-station-based models, a finding supported by the CD model's bias of +2.50 days—the largest among all cities. In contrast, the consistently good performance of all three models in Daejeon (an inland basin type) and Ulsan (a southeastern coastal type) is interpreted as reflecting the stable interannual variability of winter-to-spring temperatures, which provides favorable conditions for parameter estimation.

The fact that PhenoFlex exhibited its highest RMSE among all seven cities in Seoul suggests that temporal changes in urban heat island intensity associated with rapid urbanization over the past 50 years may have introduced non-stationary temperature-bloom relationships that cannot be captured by the model's static parameter framework (Jeong *et al*., 2011; Meng *et al*., 2020; Kim *et al*., 2024).

In terms of latitudinal patterns, the superiority of the CD model was pronounced in southern cities (Busan, Ulsan, Gwangju), whereas NHH or PhenoFlex performed relatively better in central and northern cities (Seoul, Incheon). In southern regions, chilling requirements are fulfilled early and consistently, allowing the sequential structure of the CD model to operate effectively. In central and northern regions, by contrast, greater winter temperature variability and a more gradual chilling-to-forcing transition favor the nonlinear response of NHH or the flexible transition structure of PhenoFlex, which capture these dynamics more adequately.

3. PLS-based forcing start date determination for the NHH model

The PLS-based forcing start date determination applied in this study is significant as an attempt to objectively define the input period for the NHH model. PLS regression can evaluate the influence of temperature at specific periods on phenological responses through

VIP scores and regression coefficients, even in daily temperature data characterized by strong autocorrelation (Luedeling and Gassner, 2012), thereby enabling the identification of the chilling-to-forcing transition point (Luedeling *et al*., 2013). The start dates derived from PLS ranged approximately 47 days, from DOY 45 in Busan to DOY 92 in Incheon, indirectly compensating for the structural limitation of NHH, which lacks an explicit chilling module.

However, the large variation in city-specific performance of the NHH model (Seoul: 2.43 days vs. Gwangju: 4.96 days) reflects the fact that the regional suitability of PLS-determined start dates was not uniform across cities. Furthermore, when the PLS-identified start date is used as a fixed starting point, the inherently gradual chilling-to-forcing transition may not be adequately captured (Luedeling *et al*., 2013). The lower mean R^2 of the NHH model compared to those of CD and PhenoFlex is interpreted as partly reflecting this structural sensitivity.

4. Application of PhenoFlex enhanced global optimization

This study represents the first application of the Enhanced Global Optimization (EGO) framework of Mojahid *et al*. (2025) to black locust (*Robinia pseudoacacia*) on the Korean Peninsula. By substituting the original parameters with physiologically interpretable intermediate parameters (θ^*, θ_c, π_c, τ), the framework sought to simultaneously improve search efficiency and biological plausibility.

Nevertheless, high variability was still observed in the frequency factors (A_0 and A_1). Because A_0 and A_1 are tightly coupled with the activation energies (E_0 and E_1) within the Arrhenius equation, different values can yield identical chilling accumulation through compensatory combinations. In practice, city-specific estimates of A_0 and A_1 varied by several orders of magnitude, whereas E_0 and E_1 remained within a relatively stable range. This constitutes a classic case of equifinality, and field observational data alone have inherent limitations in fully disentangling such compensatory relationships (Fernandez *et al*., 2022).

To mitigate this issue, RPIQ-based weighted-average prediction was applied to the parameter combinations derived from 10 independent optimization runs (Mojahid *et al*., 2025). This ensemble approach is considered to have suppressed the risk of overfitting inherent in any single parameter set and contributed to the mean absolute bias of PhenoFlex being the smallest among the three models. However, this remains a statistical remedy for prediction stabilization and does not resolve the physiological uncertainty underlying the chilling accumulation mechanism itself.

5. Comparison with previous studies

The results of this study are broadly consistent with the trends reported in previous research, yet are distinguished by the comparison of three process-based models under identical conditions using long-term, multi-site data for black locust.

For the CD model, Kim *et al*. (2021b) reported that a single-parameter pooled model (SM) across 26 nationwide sites failed to adequately capture inter-regional climatic differences, with the correction coefficient exhibiting a strong positive correlation with mean spring temperature ($R^2=0.71$), resulting in pronounced systematic underestimation in southern regions. In the present study, the base temperature was estimated to be higher in southern cities and lower in central cities, confirming the same directional pattern. The CD RMSE of 3.02–3.22 days reported by Chun *et al*. (2017) for peach is in good agreement with the mean RMSE of 2.78 days obtained in this study. In contrast, Jeong *et al*. (2023) reported that the CD model's dormancy release timing for apple ('Fuji') was more than 30 days earlier than that of other models, yielding an RMSE of 8.7 days with poor predictive accuracy, demonstrating that even under the same model structure, performance can vary considerably depending on species-specific chilling requirements and base temperature settings.

Regarding PhenoFlex, Luedeling *et al*. (2021) reported RMSE values of 3.8–4.0 days with small bias for apple and pear in Germany, which is consistent with the finding in this study that PhenoFlex was the most stable in terms of pooled RMSE and bias. Considered alongside the excellent performance of the DM+GDH combined model (RMSE 1.6–2.0 days) reported by Jeong *et al*. (2023), model structures that simulate both chilling and forcing nonlinearly may be advantageous for reduc-

ing bias across diverse climatic conditions. However, Fernandez *et al.* (2022) and Mojahid *et al.* (2025) noted increased RMSE under warm winter conditions and the trade-off between generalizability and regional specificity, respectively, which are consistent with the high variability in frequency factors and the equifinality observed in the present study.

For the NHH model, Alilla *et al.* (2022) modeled the entire BBCH 51−69 phenological stages using an NHH-based approach in Italy, although their study differed from the present one in that it did not include a chilling module. The highest mean RMSE of NHH in this study is attributable to its structural limitation of not directly simulating the cumulative mechanism of dormancy release, which is also consistent with the finding of Jeong *et al.* (2023) that models incorporating a chilling phase exhibited greater nationwide applicability than forcing-only models.

6. Limitations and future directions

This study has the following limitations.

First, the PhenoFlex and NHH models estimate hourly temperatures from daily maximum and minimum temperatures, and phenomena such as urban heat islands and land-sea breeze circulations can distort such idealized diurnal temperature curves (Jeong *et al.*, 2011). Furthermore, temperatures recorded by the Automated Synoptic Observing System (ASOS) may not adequately represent the microclimate of forest margins or urban parks where black locust actually grows (De Frenne *et al.*, 2021; Kemppinen *et al.*, 2024). Future studies should incorporate high-resolution gridded meteorological data or in-situ microclimate measurements from forested environments.

Second, all three models use air temperature as the sole driving variable, and other environmental factors potentially involved in flowering—such as soil moisture and photoperiod—are not accounted for (Chuine and Régnière, 2017). Given the projected increase in extreme weather events under climate change, extending the modeling framework to incorporate multiple environmental variables warrants consideration.

Third, to fundamentally reduce the parameter uncertainty inherent in PhenoFlex, temperature-response experiments conducted under controlled environments such as phytotrons should be integrated into the model calibration process (Mojahid *et al.*, 2025). For the NHH model, structural improvements—such as direct coupling with a dynamic chilling module or the implementation of year-specific variable forcing start dates—could be explored to reduce the regional variability in PLS-derived start dates. Moreover, a multi-model ensemble approach that combines the predictions of the three models using city-specific performance-based weights could serve as a viable alternative for enhancing prediction stability.

CONCLUSION

This study compared the bloom date prediction performance of three process-based models (CD, NHH, and PhenoFlex) for black locust (*Robinia pseudoacacia*) using long-term observational data (1973−2024) from seven metropolitan cities across the Korean Peninsula. The results demonstrated that greater structural complexity in a model does not necessarily translate into improved predictive performance at the local scale. For practical prediction at the individual city level, the three-parameter CD model (mean RMSE of 2.78 days) was the most stable, whereas for constructing a nationally applicable pooled model across multiple regions, PhenoFlex exhibited the clearest justification by effectively suppressing systematic bias (mean absolute bias of 0.08 days).

Furthermore, spatial heterogeneity aligned with the climatic gradient of the Korean Peninsula was observed: the CD model performed best in southern regions, while NHH and PhenoFlex were superior in central and northern regions. Accordingly, rather than uniformly applying a single optimal model, a strategic approach is warranted—one that selects and combines models according to the research objective and regional climatic characteristics. Future research priorities include controlled physiological experiments in phytotrons to resolve the parameter uncertainty identified during model optimization, along with coupled analyses under Shared Socioeconomic Pathways (SSPs) climate change scenarios informed by such experimental findings.

ACKNOWLEDGEMENTS

This study was carried out with the support of the Rural Development Administration project (Development of an optimal model for honey production using prediction of major nectar plant flowering periods, RS-2023-00230940).

LITERATURE CITED

Alilla, R., F. De Natale, C. Epifani, B. Parisse and G. Cola. 2022. The flowering of black locust (*Robinia pseudoacacia* L.) in Italy: A phenology modeling approach. Agronomy 12(7): 1623.

Anderson, J., E. Richardson and C. Kesner. 1985. Validation of chill unit and flower bud phenology models for 'Montmorency' sour cherry. I International Symposium on Computer Modelling in Fruit Research and Orchard Management 184.

Cerlini, P., M. Saraceni, F. Orlandi, L. Silvestri and M. Fornaciari. 2022. Phenological response to temperature variability and orography in Central Italy. Int. J. Biometeorol. 66(1): 71-86.

Cesaraccio, C., D. Spano, R. L. Snyder and P. Duce. 2004. Chilling and forcing model to predict bud-burst of crop and forest species. Agric. For. Meteorol. 126(1-2): 1-13.

Cesaraccio, C., D. Spano, R. L. Snyder and P. Duce. 2005. Corrigendum to "Chilling and forcing model to predict bud-burst of crop and forest species" [Agric. For. Meteorol. 126(2004) 1-13]. Agric. For. Meteorol. 129(3-4): 211.

Chuine, I. and J. Régnière. 2017. Process-based models of phenology for plants and animals. Annu. Rev. Ecol. Evol. Syst. 48: 159-182.

Chun, J. A., K. Kang, D. Kim, H.-H. Han and I.-C. Son. 2017. Prediction of full blooming dates of five peach cultivars (*Prunus persica*) using temperature-based models. Sci. Hortic. 220: 250-258.

Chung, U., J.-E. Jung, H.-C. Seo and J. I. Yun. 2009. Using urban effect corrected temperature data and a tree phenology model to project geographical shift of cherry flowering date in South Korea. Clim. Change 93(3): 447-463.

Cleland, E. E., I. Chuine, A. Menzel, H. A. Mooney and M. D. Schwartz. 2007. Shifting plant phenology in response to global change. Trends Ecol. Evol. 22(7): 357-365.

Cola, G., L. Mariani, F. Salinari, S. Civardi, F. Bernizzoni, M. Gatti and S. Poni. 2014. Description and testing of a weather-based model for predicting phenology, canopy development and source-sink balance in *Vitis vinifera* L. cv. Barbera. Agric. For. Meteorol. 184: 117-136.

Cola, G., O. Failla, D. Maghradze, L. Megrelidze and L. Mariani. 2017. Grapevine phenology and climate change in Georgia. Int. J. Biometeorol. 61: 761-773.

De Frenne, P., J. Lenoir, M. Luoto, B. R. Scheffers, F. Zellweger, J. Aalto, M. B. Ashcroft, D. M. Christiansen, G. Decocq and K. De Pauw. 2021. Forest microclimates and climate change: Importance, drivers and future research agenda. Glob. Change Biol. 27(11): 2279-2297.

Egea, J. A., D. Henriques, T. Cokelaer, A. F. Villaverde, A. MacNamara, D.-P. Danciu, J. R. Banga and J. Saez-Rodriguez. 2014. MEIGO: An open-source software suite based on metaheuristics for global optimization in systems biology and bioinformatics. BMC Bioinformatics 15(1): 136.

Erez, A., S. Fishman, G. Linsley-Noakes and P. Allan. 1989. The dynamic model for rest completion in peach buds. In: II International Symposium on Computer Modelling in Fruit Research and Orchard Management, p. 276.

Fernandez, E., K. Schiffers, C. Urbach and E. Luedeling. 2022. Unusually warm winter seasons may compromise the performance of current phenology models-Predicting bloom dates in young apple trees with PhenoFlex. Agric. For. Meteorol. 322: 109020.

Fishman, S., A. Erez and G. Couvillon. 1987a. The temperature dependence of dormancy breaking in plants: Mathematical analysis of a two-step model involving a cooperative transition. J. Theor. Biol. 124(4): 473-483.

Fishman, S., A. Erez and G. Couvillon. 1987b. The temperature dependence of dormancy breaking in plants: Computer simulation of processes studied under controlled temperatures. J. Theor. Biol. 126(3): 309-321.

García-Gutiérrez, V. and F. Meza. 2023. Modeling phenology combining data assimilation techniques and bioclimatic indices in a Cabernet Sauvignon vineyard (*Vitis vinifera* L.) in Central Chile. Remote Sens. 15(14): 3537.

Guralnick, R., T. Crimmins, E. Grady and L. Campbell. 2024. Phenological response to climatic change depends on spring warming velocity. Commun. Earth Environ. 5(1): 634.

Jeong, J. H., C. H. Ho, H. W. Linderholm, S. J. Jeong, D. Chen and Y. S. Choi. 2011. Impact of urban warming on earlier spring flowering in Korea. Int. J. Climatol. 31(10): 1488.

Jeong, J. H., J. H. Han, J. G. Cho, D. Y. Lee, S. G. Lee, S. H. Jang and S. H. Ryu. 2023. Evaluation of models for predicting full bloom dates of apples in major production areas of South Korea under temperature changes. J. Bio-Env. Con. 32(4): 501-512.

Jochner, S. and A. Menzel. 2015. Urban phenological studies-past, present, future. Environ. Pollut. 203: 250-261.

Jung, C. and S.-k. Cho. 2015. Relationship between honeybee population and honey production in korea: A historical trend analysis. J. Apic. 30(1): 7-12.

Jung, J. E., E. Y. Kwon, U. Chung and J. I. Yun. 2005. Predicting cherry flowering date using a plant phenology model. Korean J. Agric. For. Meteorol. 7(2): 148-155.

Kamimori, M. and A. Hosomi. 2025. Predicting budbreak dates for 'Delaware'grape considering chilling and heat requirements using PhenoFlex. Ecol. Model. 501: 110985.

Kemppinen, J., J. J. Lembrechts, K. Van Meerbeek, J. Carnicer, N. I. Chardon, P. Kardol, J. Lenoir, D. Liu, I. Maclean and J. Pergl. 2024. Microclimate, an important part of ecology and biogeography. Glob. Ecol. Biogeogr. 33(6): e13834.

Kim, A.-R., J. Seol, B.-S. Lim, C.-H. Lim, G.-S. Kim and C.-S. Lee. 2024. Response of plant phenology on microclimate change depending on land use intensity in Seoul, central Korea. Forests 15(4): 718.

Kim, J. H., E. J. Lee and J. I. Yun. 2013. Prediction of spring flower bloom dates using digital forecast and bioclimatological models. Korean J. Agric. For. Meteorol. 15(1): 40-49.

Kim, J.-H., E.-J. Yun, D. G. Kang, J.-H. Han, K.-M. Shim and D.-J. Kim. 2025a. Evaluation of phenology models for predicting full bloom dates of 'Niitaka' pear using orchard image-based observations in South Korea. Atmosphere 16(9): 996.

Kim, K. M., M. Y. Lee, Y. S. Choi, E. J. Kang, H. G. Park, B. S. Park, J. K. Kim, S. M. Han, S. O. Woo and S. G. Kim. 2021a. Analysis of annual acacia honey production status and environmental factors. J. Apic. 36(1): 11-16.

Kim, M.-J. and C. Jung. 2025. "Beyond the timing of flowering: Shortening of spring flowering duration of Korean trees and potential community effects." Ecology 106(9): e70194.

Kim, S., T. K. Kim, S. Yoon, K. Jang, H. Lim, J.-H. Lim, W. Y. Lee, M. Won and H. S. Kim. 2021b. Recent changes in bloom dates of *Robinia pseudoacacia* and bloom date predictions using a process-based model in South Korea. J. Korean Soc. For. Sci. 110(3): 322-340.

Kim, S., J.-Y. Shin, G. Lee, J. Park and K. Sung. 2025b. Future changes in precipitation extremes over South Korea based on observations and CMIP6 SSP scenarios. Water 17(11): 1702.

Linvill, D. E. 1990. Calculating chilling hours and chill units from daily maximum and minimum temperature observations. HortScience 25(1): 14-16.

Luedeling, E. and A. Gassner. 2012. Partial least squares regression for analyzing walnut phenology in California. Agric. For. Meteorol. 158: 43-52.

Luedeling, E., A. Kunz and M. M. Blanke. 2013. Identification of chilling and heat requirements of cherry trees—a statistical approach. Int. J. Biometeorol. 57(5): 679-689.

Luedeling, E., K. Schiffers, T. Fohrmann and C. Urbach. 2021. PhenoFlex-an integrated model to predict spring phenology in temperate fruit trees. Agric. For. Meteorol. 307: 108491.

Luo, Y., E. Weng, X. Wu, C. Gao, X. Zhou and L. Zhang. 2009. Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models. Ecol. Appl. 19(3): 571-574.

Mariani, L., R. Alilla, G. Cola, G. Dal Monte, C. Epifani, G. Puppi, F. Osvaldo. 2013. IPHEN—A real-time network for phenological monitoring and modelling in Italy. Int. J. Biometeorol. 57: 881-893.

Meng, L., J. Mao, Y. Zhou, A. D. Richardson, X. Lee, P. E. Thornton, D. M. Ricciuto, X. Li, Y. Dai and X. Shi. 2020. Urban warming advances spring phenology but reduces the response of phenology to temperature in the conterminous United States. Proc. Natl. Acad. Sci. U.S.A. 117(8): 4228-4233.

Menzel, A., T. H. Sparks, N. Estrella, E. Koch, A. Aasa, R. Ahas, K. Alm-Kübler, P. Bissolli, O. g. Braslavská and A. Briede. 2006. European phenological response to climate change matches the warming pattern. Glob. Change Biol. 12(10): 1969-1976.

Mo, Y., J. Zhang, H. Jiang and Y. H. Fu. 2023. A comparative study of 17 phenological models to predict the start of the growing season. Front. For. Glob. Change 5: 1032066.

Mojahid, H., L. Caspersen, A. Delgado, E. Dapena, E. Luedeling, K. Schiffers and E. Fernandez. 2025. Cross-regional validation of the PhenoFlex framework for flowering date prediction in apple: A study across Germany and Spain. Agric. For. Meteorol. 373: 110746.

Park, S., H. Park, J. Im, C. Yoo, J. Rhee, B. Lee and C. Kwon. 2019. Delineation of high resolution climate regions over the Korean Peninsula using machine learning approaches. PLoS One 14(10): e0223362.

Parton, W. J. and J. A. Logan. 1981. A model for diurnal variation in soil and air temperature. Agric. Meteorol. 23: 205-216.

Pearse, W. D., M. Stemkovski, B. R. Lee, R. B. Primack and S. D. Lee. 2023. Consistent, linear phenological shifts across a century of observations in South Korea. New Phytol. 239: 824-829.

Piao, S., Q. Liu, A. Chen, I. A. Janssens, Y. Fu, J. Dai, L. Liu, X. Lian, M. Shen and X. Zhu. 2019. Plant phenology and global climate change: Current progresses and challenges. Glob. Change Biol. 25(6): 1922-1940.

Stemkovski, M., W. D. Pearse, S. R. Griffin, G. L. Pardee, J. Gibbs, T. Griswold, J. L. Neff, R. Oram, M. G. Rightmyer and C. S. Sheffield. 2020. Bee phenology is predicted by climatic variation and functional traits. Ecol. Lett. 23(11): 1589-1598.

Thackeray, S. J., P. A. Henrys, D. Hemming, J. R. Bell, M. S. Botham, S. Burthe, P. Helaouet, D. G. Johns, I. D. Jones and D. I. Leech. 2016. Phenological sensitivity to climate across taxa and trophic levels. Nature 535(7611): 241-245.

Wang, E. and T. Engel. 1998. Simulation of phenological development of wheat crops. Agric. Syst. 58: 1-24.

Wolkovich, E. M., B. I. Cook, J. M. Allen, T. Crimmins, J. L. Betancourt, S. E. Travers, S. Pau, J. Regetz, T. J. Davies and N. J. Kraft. 2012. Warming experiments underpredict plant phenological responses to climate change. Nature 485(7399): 494-497.

Yang, C., N. Lei, C. Menz, A. Ceglar, J. A. Torres-Matallana, S. Li, Y. Jiang, X. Tan, L. Tao and F. He. 2024. Regional uncertainty analysis between crop phenology model structures and optimal parameters. Agric. For. Meteorol. 355: 110137.

Zipper, S. C., J. Schatz, A. Singh, C. J. Kucharik, P. A. Townsend and S. P. Loheide. 2016. Urban heat island impacts on plant phenology: Intra-urban variability and response to land cover. Environ. Res. Lett. 11(5): 054023.