International Cartographic Association
Association Cartographique Internationale

AI-Based Plant Phenology Mapping and Analysis Using Satellite Imagery and Meteorological Data

Geunah Kim[a], Yangwon Lee [b,*]

[a] *Major of Spatial Information Engineering, Division of Earth Environmental System Science, Pukyong National University, 45 Yongso-ro, Busan 48513, Republic of Korea, topho@pukyong.ac.kr*
[b] *Major of Geomatics Engineering, Division of Earth Environmental System Science, Pukyong National University, 45 Yongso-ro, Busan 48513, Republic of Korea, modconfi@pknu.ac.kr*

* Corresponding author

Keywords: Remote Sensing, Climate Change, Phenology, Artificial Intelligence, Phenology Map

Abstract:

Recently, global warming has led to a rise in average temperatures, and problems related to climate change have become increasingly serious. Global warming is also affecting temperature-sensitive organisms and environments, and changes in ecosystems caused by this are being observed. Among the various indicators used to assess the impacts of climate change, seasonal change is one of the most common and easily recognized.

Plant phenology is strongly influenced by energy cycles and meteorological conditions, while also affecting the geographic and seasonal variation of carbon and water cycles. To more effectively analyze and communicate changes in plant phenological events, visualizing them in the form of maps is considered to be an effective approach.

Therefore, this study integrates meteorological data and satellite-based vegetation indices to predict and estimate specific physiological events such as flowering and peak autumn foliage. AutoML-based machine learning models and the TabNet deep learning model were compared and evaluated. Model performance was analyzed not only in terms of prediction and estimation accuracy, but also with a focus on the reproducibility of spatial distribution.

Furthermore, considering that few previous studies have applied deep learning approaches to phenological mapping, this study expands the applicability of map-based prediction techniques and presents their practical potential as spatial decision-support tools for responding to climate change and managing ecosystems.

The phenology data used in this study were derived from flowering and peak foliage observation records collected at ASOS stations operated by the Korea Meteorological Administration (KMA). The analysis revealed that while the temporal trends in phenological data were not strongly pronounced, flowering dates tended to occur earlier, and peak foliage dates tended to occur later. Given the limited temporal autocorrelation and the strong linearity between phenological events and meteorological or vegetation states, regression models were deemed more appropriate than time-series models. Consequently, the study utilized Extremely Randomized Trees (ERT), AutoML, and TabNet models.

ERT is a tree-based ensemble method that excels in computational efficiency and preventing overfitting. AutoML is a framework that automates data preprocessing, model training, and hyperparameter optimization to select the optimal model. TabNet combines the strengths of artificial neural networks and tree-based models, enabling the selection of important features and achieving high predictive performance and interpretability through sparse constraints.

For model training, various temperature-related variables were generated to identify the most suitable variables for each phenological event through correlation analysis and feature importance evaluation. Vegetation indices were assessed using the Normalized Difference Vegetation Index (NDVI), a widely used indicator of vegetation vigor. Considering the study period, NDVI data were obtained from MODIS products, available since 2003. Missing values in the vegetation index data, primarily caused by cloud cover, were reconstructed using the DCT-PLS (Discrete Cosine Transform-Partial Least Squares) method.

The final explanatory variables, including selected meteorological and vegetation indices, year, and XY coordinates, were used to train and optimize the machine learning models. The trained models were then applied to spatially continuous input data to map plant phenological events. Examples of the generated plant phenology maps are presented in Fig. 1–Fig. 6.

https://doi.org/10.5194/ica-abs-10-145-2025

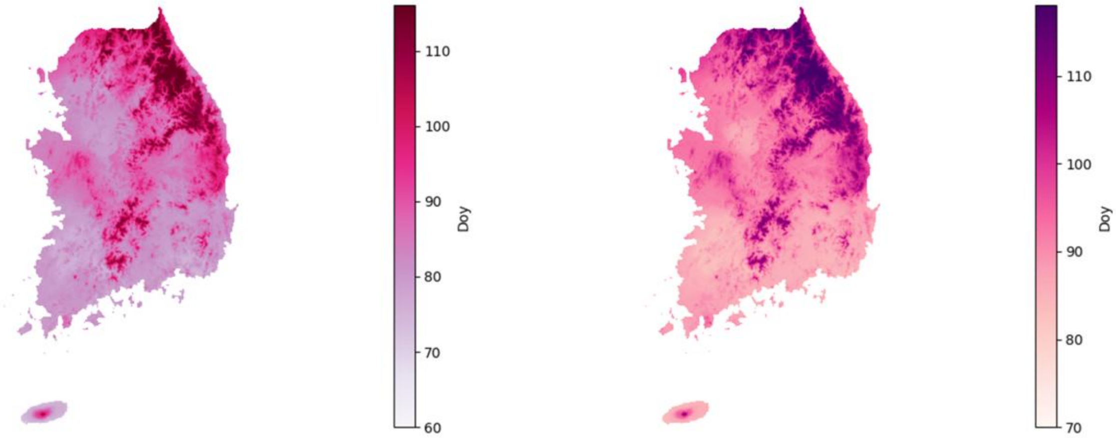

Figure 1. Examples of AI Model-based Phenology Maps for Cherry Blossoms and Azale(as.

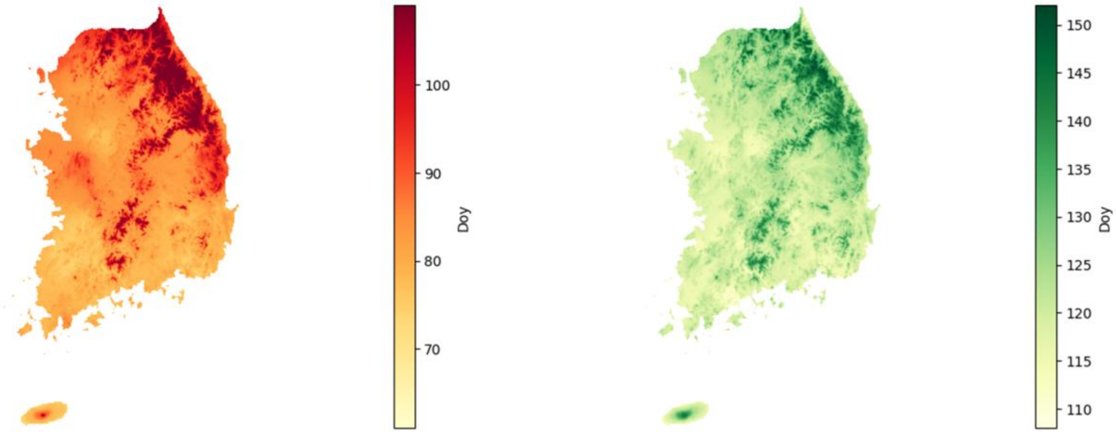

Figure 2. Examples of AI Model-based Phenology Maps for forsythias and Acasiafalses

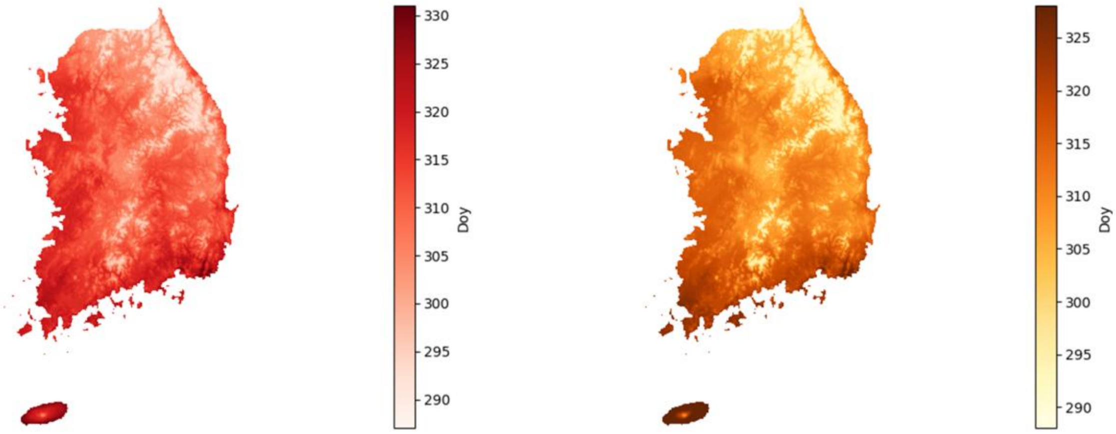

Figure 3. Examples of AI Model-based Phenology Maps for Maple And Ginkgo

Traditional plant phenology maps often aggregate information at specific points or regions or rely on interpolation methods such as kriging to create contour-based representations. In contrast, the proposed method extends point-based observational data into spatially continuous surfaces, providing significant advantages. This approach allows for the creation of plant phenology maps that reflect detailed spatial characteristics and facilitates the analysis of local variations.

Abstracts of the International Cartographic Association, 10, 145, 2025
32nd International Cartographic Conference (ICC 2025), 17–22 August 2025, Vancouver, Canada.
https://doi.org/10.5194/ica-abs-10-145-2025

Such maps can be used to investigate regional differences in plant growth, understand the spatial distribution of climate change impacts, and simultaneously observe large-scale and localized changes. Additionally, this method enables tracking temporal changes in plant phenology, further enhancing its utility for long-term monitoring and analysis.

The plant phenology data used in the study were obtained from the flowering and peak foliage dates measured at ASOS stations by the Korea Meteorological Administration. The analysis of plant phenology revealed that, while time-series trends were not very distinct, flowering was occurring earlier, and peak foliage was occurring later. When selecting models for plant phenology, it was considered more appropriate to use regression models rather than time-series models because the time-series trends were not strong, but there was a strong linear relationship between weather conditions and vegetation status. Therefore, Extremely Randomized Trees (ERT), AutoML, and TabNet models were selected. ERT is a tree-based ensemble technique known for its computational efficiency and strength in preventing overfitting. AutoML is a framework that automates the processes of data preprocessing, model training, and hyperparameter optimization to select the best model. TabNet is a model that combines the advantages of neural networks and tree-based models, selecting important features from the given data and providing high predictive performance and interpretability through sparsity constraints.

The meteorological data used for model training consisted of temperature data. To select the most suitable temperature variables for each plant phenology event, various temperature variables were created, and correlation analysis and feature importance analysis were performed to select the best variables for each event. For the vegetation index, NDVI, which is most commonly used to assess vegetation vitality, was utilized. Given the research period, MODIS vegetation index data, available since 2003, were used. To handle missing values in the vegetation index data, the DCT-PLS technique was applied. Finally, using the selected meteorological and vegetation index variables, along with year and XY coordinates, the AI models were trained and optimized. The trained models were then applied to spatial continuous input data to map and visualize plant phenology events.

In addition, the ERT model, which was identified in this study as having the highest spatial reproducibility, shows potential for application across various climates and ecosystems. To achieve this, it is necessary to secure plant phenology observation data for each region and to select temperature and satellite-based vegetation index variables that are most suitable for the region’s climate characteristics, ecosystem types, and dominant species. Through this process, it is expected that prediction and estimation modeling of regional plant phenological events will be sufficiently feasible.

Acknowledgements

This work was carried out with the support of the "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ0162342025)" by the Rural Development Administration, Republic of Korea.

https://doi.org/10.5194/ica-abs-10-145-2025