Peak Bloom Prediction 2026
Executive Summary and Final Solution

March 2026

1 Executive Summary

The timing of spring phenological events, particularly the peak bloom of cherry blossoms, serves as a sensitive indicator of local climate change and interannual weather variability. Accurate forecasts of these events are essential for cultural, economic, and scientific purposes. This study presents a reproducible forecasting pipeline designed to predict the 2026 peak cherry blossom bloom, reported as both day-of-year (DOY) and calendar date, across five geographically diverse locations.

The primary biological hypothesis posits that bloom timing is mainly determined by late-winter and early-spring temperatures, with additional influences from precipitation, winter chill, and species genetics. To enhance the accuracy of climate data, temperature values for each site are adjusted according to the altitude difference between the meteorological station and the blossom location. Feature engineering incorporates biologically meaningful climate summaries, including early-spring temperature statistics (mean and maximum daily temperatures), cumulative pre-bloom precipitation, and chill-day variables that quantify cold exposure required to break endodormancy. Species and continental indicators are also included to account for systematic differences, thereby supporting reliable predictions under changing climate conditions.

Given the inherent model risk in ecological forecasting, we tested a multi-model ensemble integrating several algorithm types. Linear models (Ordinary Least Squares, Ridge, Lasso, Bayesian Ridge, and a weighted linear model) provided interpretable baselines and broad trend identification. Non-linear methods (Gradient Boosting Quantile Regression and Random Forests) and time-series models (ARIMAX) were also evaluated, but these underperformed according to MAE, RMSE, and R^2 metrics. Process-based phenological models were used to impose biological constraints. Ultimately, the weighted linear model and regularized linear models were emphasized, as they consistently outperformed non-linear and time-series approaches.

Since no single model performs best everywhere, we use block-based cross-validation for evaluation. For the final holdout, we use the most recent 10 years as an unseen test set. Base models are ranked by Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2. Only the time-weighted linear model (WLM), OLS, Ridge, Bayesian Ridge, and Lasso are used in the final stacked ensemble, which is combined by a RidgeCV meta-learner trained exclusively on out-of-sample predictions.

To support decision-making, point forecasts are provided with statistically calibrated prediction intervals. We analyze residuals from holdout predictions to estimate empirical error quantiles and generate reliable bounds for each location. The entire framework follows open-science principles. Implemented as a Quarto document, the pipeline installs dependencies, trains the ensemble, validates outputs, and renders final prediction tables, allowing independent reviewers to easily reproduce the 2026 results.

2   Final Solution

The selected final approach is a stacked ensemble that combines the top-performing linear-family base models using a RidgeCV meta-learner. This design balances interpretability, robustness, and predictive accuracy while reducing model-specific risk.

Final Model Performance (Holdout)

For the final selected model (`stacked_ensemble`), the primary holdout metrics are:

- MAE: 3.8089 days

- RMSE: 4.4273 days

Final 2026 Predictions (Stacked Ensemble)

Location	Predicted Date	Predicted DOY	90% PI Lower	90% PI Upper
Kyoto	Apr 02	92.8	86.3	99.3
Liestal	Apr 03	93.0	86.5	99.4
New York City	Apr 09	99.7	93.2	106.2
Vancouver	Mar 30	89.5	83.0	96.0
Washington, DC	Apr 02	92.2	85.7	98.7

These predictions correspond to the submission-ready stacked ensemble output in:

- `data/model_outputs/predictions/final_2026_predictions_stacked_ensemble.csv`