

Consider: remove passive voice. "has been shown to" --> "has varied"

Global: Remove highlight

Consider more concrete phrasing. "To be linked"--> in what way? Does temperature, wind, and precipitation cause the trees to bloom? Potentially find a synonym that is more active. Additionally, the statement in the second sentence could possibly be moved to the first sentence and restate "climate" as "climate warmth, wind speed, and levels of precipitation".

Introduction

Cherry blossom blooming has been shown to vary with species and linked to have supported significant association of flowering in cherry trees with climate temperature, wind, and precipitation. Miller-Rushing, et al. (2007)¹ found evidence of temperature impacting flowering, especially during February and March. Previous works have applied different approaches to predicting peak bloom for varied cherry tree species in various locations. For instance, Chung, et al. (2011)² applied a two-step phenology model, per research by Cesaraccio, et al. (2004)³ and Jung, et al. (2005)⁴, to predict peak bloom dates for cherry trees in Washington, D.C. and the Mid-Atlantic region. This report focuses on predicting peak bloom dates for the next decade of cherry blossoms at four sites in Washington, D.C. (USA), Kyoto (Japan), Liestal-Weideli (Switzerland), and Vancouver, BC (Canada), with consideration of multiple climate factors.

Consider a sentence that links this statement to the rest of the paragraph. Is this two-step model referenced because your model intends to use a two-step model?

Consider more active phrase. "This report focuses on predicting peak bloom dates using a multivariate analysis using XYZ factors...", etc., rather than "consideration of multiple...factors".

Global: I don't believe your parenthetical citations need a footnote anchor, based on my reading of the APA style guide. Possible I am wrong, though

Data for modeling the day of peak bloom were obtained through a variety of sources. Historical data for peak bloom dates were obtained through a GitHub repository for the Cherry Blossom Prediction Competition hosted by George Mason University. Historical data for day of peak bloom dates back to 812 AD for Kyoto, 1894 for Liestal, and 1921 for Washington, DC. No historical data for day of peak bloom were available for Vancouver, BC. For each site, peak bloom differed for each site. For Kyoto, peak bloom is defined as the date when 70% of the cherry blossoms are in full bloom. In D.C. and Vancouver, BC, the date when 70% of the cherry trees and in Liestal, 25% of the trees. In statistical modeling, peak bloom dates were weighted, with 70% bloom set as the baseline. Research and past data have shown that the time from first bloom to peak bloom for cherry blossoms typically lasts about one week. It was assumed that cherry blossoms progressed from first bloom to peak bloom linearly, so weights were assigned to historical dates based the percentage of linear progression to (or from) 70% bloom.

Who assumed? The researchers in the prior sentence?

Citation?

I really like the demarcation on different definition of peak bloom since I missed that myself!

For the three locations with available peak bloom date data as well as Vancouver, BC (Canada), publicly available climate data on factors including wind speed, sea level pressure, precipitation, snow depth, temperature, sun light duration, plant hardiness, and growing degree days are explored for this project. The data for each location are based on a nearby weather station, and those stations' observations on all the mentioned factors are gathered from multiple datasets. The weather stations are at the following locations: Reagan National Airport for Washington, D.C., Basal Binningen for Liestal-Weideli, Vancouver International Airport for Vancouver, and in Kyoto for Kyoto. Wind speed and sea level pressure observations come from the Global Surface Summary of the Day (GSOD) dataset. Precipitation, snow depth, and temperature observations primarily come from the Global Historical Climatology Network (GHCN) dataset. Sunlight duration data used is from the National Oceanic Atmospheric Administration's (NOAA) solar calculator. Plant hardiness zones for the locations are found through the

Typo?

Remove

Consider referencing the units these factors are measured in by using parenthetical

Consider a footnote referencing the hyperlink and date of access

Be more specific. "Not length"--how many values is that: 3? 4? 5? 20?

Plantmaps website. Growing degree days data is obtained through the chillR package in R. Synop reports from Ogimet obtained through the climate package in R is used as available when temperature data is missing in the GSOD dataset for nearby weather stations. A summary of the sites where peak bloom predictions were modeled are shown in Table 1.

Remove

"there were"?

Missing data was handled in a variety of ways depending on the length and number of missing values and the nature of the distribution of the known data. Where there were only a relatively limited number of missing values and the runs of missing values were not lengthy, linear interpolation was used to fill in missing values. Linear interpolation of missing values was used particularly for imputing missing Tmin, Tmax and Tavg data as well as data on other covariates including SLP and precipitation. When there longer stretches of missing data such as long runs of missing daily data or years of missing data, missing values were sometimes imputed if the run was not too long and when the trend of the data approximated a stationary series with white noise. Trended series or series with a clear random walk or moving average were either imputed over flat areas of the trend only or were predicted through fitting the series to an ARIMA model and forecasting the missing values from the time series. When the series for the data was relatively flat with no trend and if the distribution of the data was approximately normally distributed, we took the empirical mean and standard deviation of the data and imputed missing values by randomly sampling from a normal distribution with the same mean and standard deviation as the data. When the trend of the time series was relatively flat and runs were short but the distribution of the data not quite normal, we performed Monte Carlo simulation by taking 10,000 random samples from the known historic data and then we randomly sampled data for missing values from the sampled distribution of the data. When gaps in missing data were serious and lengthy and a time series of the data was highly stochastic or showed signs of unpredictable random walks, moving averages, or nonconstant variance, missing values were left in the data. If there were only a couple of years of missing data for the covariate, a model was sometime run with the variable if the variable was deemed important to prediction. Often, however, variables with missing data had to be dropped from the analysis entirely due to poor data quality.

Specify

Unclear how all these data imputation steps overlapped with each other or how many observations they were used on. It seems intricate and I compliment the effort (I especially like the idea of using a dataset's distribution to randomly draw values of a simulated normal distribution), but this seems to be a possibly too complex of a solution. Especially since it is so hard to disentangle what steps were taken and when. Did you run Monte Carlo simulations at the same time as time series imputation? Or did you run Monte Carlo analysis after some of the data was imputed? How did that impact the imputation process?

Location	Latitude (°)	Longitude (°)	Altitude (m)	Country
Kyoto	35.0120	135.6761	44	Japan
Liestal-Weideli	47.4814	7.730519	350	Switzerland
Washington, D.C.	38.8853	-77.0386	0	USA
Vancouver, BC	49.2237	-123.1636	24	Canada

Table 1. Locations and Species of Cherry Trees of Interest.

Methodology

make singular

One main predictor considered was temperature. To begin, we attempted to explore temperature as a predictor based on the chill days model developed by Cesaraccario, et al.³ and modifications to the model by Chung, et al.² This two-step phenology model proposes that buds enter a dormancy phase typically on a set date in the fall, when temperature goes below a threshold temperature (T_c), and accumulate chill days (D_c) until a chill requirement (R_c) is satisfied. Once this chill requirement is satisfied, the buds begin to develop with accumulation of anti-chill days (D_h) and may reach peak bloom once a heat requirement (R_h) is met. The daily chill days or anti-chill days values were calculated based on daily maximum, minimum, and average temperatures as described by Cesaraccario, et al.³. **Define what a Growing Degree Day is** tiation dates and cal peak bloom date data available) were identified as those which minimized root mean square error between observed and predicted peak bloom dates. As an alternate, temperature was also explored based on the Growing Degree Day (GDD) model through the GDD function in the chillR R package. **function** uses average hourly temperatures, and GDD values were obtained through the equation $GDD = \text{Thourly}/24 - T_{\text{base}}$, where “Thourly” refers to hourly temperatures for the day and “Tbase” refers to the base temperature above which growing degree days accrue. The chillR package generated hourly temperatures from daily temperature data available based on an idealized temperature curve using a sine function for warming occurring between sunrise and sunset and a logarithmic decay function for cooling overnight⁵⁻⁶

Footnote anchors don't have corresponding footnotes

I believe the beginning of this sentence was accidentally removed

Covariates

Remove extra comma and empty parenthesis

Remove empty parenthesis

Remove "[equation]"

Several covariates were considered for statistical modeling of the research on *Prunus* phenology has strongly suggested that temperature timing of blooming. A number of temperature-related and temperature for model fit including cumulative average temperature (\bar{T}), cumulative daily minimum temperature [equation], cumulative daily maximum temperature [equation], absolute difference of cumulative daily max and min temperatures (TDIFF) ($T_{\text{max}} - T_{\text{min}}$), plant hardiness zone, growing degree days (GDD), heating days, chill days, and an optimized estimate of peak bloom day based on the sum of heating days and chill days. We chose to use cumulative daily measures of temperature rather than average or median measures of temperature because research on plant phenology suggests that bloom times are likely due to an accumulation of temperature rather than daily average temperature.[citation] Several studies have even shown that many plant species exhibit minimum/maximum temperature thresholds after which there is no further daily accumulation towards the chilling/heating requirement.[citation] As a result, measures of temperature that reflect some cumulative accumulation of temperature are more likely to have predictive power of bloom time than average or median temperatures.

This paragraph seems to overlap with the prior paragraph. GDD are already mentioned and their acronym defined. No need to include here.

Remove

Out of all the temperature-related variables considered, cumulative average temperature, cumulative absolute temperature difference, GDD, heating days, chill days, and optimized peak bloom date were all significant predictors in modeling, though not all were used. Heating days and optimized peak bloom day, for instance, both displayed high collinearity with GDD. But since

Based on a specific model? No reference to this model is included in the report

Provide the evidence for this in a parenthetical such as the test value for the test you used to measure collinearity

GDD was a much stronger predictor for PBD, heating days and optimized peak bloom were not used in modeling. Cumulative minimum and cumulative maximum temperatures showed predictive power in some models, but not all. Cumulative min and max temperatures were abandoned in favor of much stronger covariates such as cumulative average temperature and cumulative absolute temperature difference. GDD, and TAVG were significant predictors in all models used in this study. GDD was found to correlate highly with the peak bloom dates (Liestal: $r=0.996$; Kyoto: $r=0.950$; Washington: $r=0.950$) and proved to be the strongest predictor in our models. Chill days was significant and appropriate for predicting peak bloom day in Washington, Kyoto, and Liestal, but not in Vancouver. TDIFF was a significant predictor for Washington and Liestal, but not the other cities. Plant hardiness zone was found to be a significant predictor but there were problems incorporating a categorical variable into time series modeling so the variable was ultimately not included in any of the models.

Global: there are a number of times when statistical significance, or lack thereof, is used to justify a decision in the modeling process. Provide evidence of statistical significance/non-significance by including test values

Consider including what those problems were

suggested other factors may contribute to the timing of blooming such as hours of peak, and the urban heat island effect. [citation] Several sunlight-related variables model fit including average number of sunlight hours between January 1st and April 1st of hours of sunlight between May 1st and September 30th from the prior summer, number of possible sunshine, and total hours of sunlight from Jan 1st to March 31st. The number of sunlight hours from May-September proved to be a significant predictor for Liestal and Vancouver, but for Kyoto, total number of sunlight hours (Jan-Mar) had more predictive power. Total hours of sunlight was only fit as a covariate for Kyoto so it is possible that this predictor has more predictive power for other sites as well. Average hours of sunlight from Jan-Mar was also significant for Washington, DC.

on the effect

Remove

Sea level air pressure (0.1*mbar) was a significant predictor of bloom time in every model where there was data available for this metric. Unfortunately, sea level air pressure was not available for a long enough series for Liestal or Vancouver so the covariate could not be incorporated into models for these sites. We are not currently aware of any research on effect atmospheric air pressure may have on bloom times for cherry trees, but the fact that these predictors were significant in modeling is perhaps indicative that some association exists, and it is perhaps an area that merits more research.

What is mbar? Avoid technical jargon unless proper definitions are provided

A study has suggested that wind speed may be a significant predictor for bloom time. [citation] However, due to the paucity of data on wind speed, we were only able to include wind speed as a covariate in the model for Washington, DC.

Variable selection was performed using backward stepwise regression on linear models. Covariates present in the models with a balance of the lowest AIC and highest adjusted [R-squared] were selected for modeling. If variables displayed high variance inflation factor (VIF), these covariates were removed from the model and backward stepwise regression was reimplemented with the offending variable left out. The process was iterated until a model of best fit had no covariates displaying collinearity.

Consider reducing this section. There is a lot of discussion on covariates that ultimately not used. Unless they are covariates that were previously shown in another study to show statistical significance, but your model proves otherwise, or you excluded variables that would intuitively seem to be significant based on the science, I don't believe it is beneficial to discuss a lot of covariates that ultimately weren't used. It gives the impression a lot of covariates were thrown against the wall to see what stuck

Large statement. What is considered "reasonable" and by whom?

Statistical Analysis

Basic linear models were first fit to the data for Washington, DC, Kyoto, Japan, and Liestal, Switzerland. Models of best fit were selected for each site from backward stepwise regression. Covariates in the model of best fit were subsequently forecasted to 2032 using multiple linear regression with AR, MA, ARMA, ARIMA, or SARIMA adjusted errors. Covariates were predicted using all reasonably associated variables for which historic climate data was available for the site. Models for covariate forecasting were checked for assumptions of linear modeling such as independent, normal, and homoscedastic errors, and linearity. Forecasted covariates were then refit to the original linear models of best fit, yielding forecasted fitted values for the day of peak bloom.

Use the full name with parenthetical references to acronyms in their first appearance of the paper rather than in later paragraphs

To improve prediction accuracy, monthly autocorrelation was included in a monthly autoregressive integrated moving average (SARIMA) model for Washington, Kyoto, and Liestal. Linear models of best fit were first selected for each site. The best fitting model was then fit as a multiple linear regression model with SARIMA errors on the same set of covariates. Peak bloom dates were then forecasted from these model using historical data alone and the SARIMA structure of these models.

This may be my statistical ignorance, but why are SARIMA s.e's used? Why not regular and usual standard errors? Consider providing justification of this approach for the reader

BC were much more difficult as there are presently no historical data for this site. We took two different approaches to forecast bloom times for each site. The first approach employed a linear regression model. To forecast peak bloom dates for each site, a linear regression model of best fit was fitted to peak bloom dates and a design matrix of covariates from Washington, Kyoto, and Liestal. The model of best fit was selected through backward stepwise regression and the model with the best balance between lowest AIC and highest adjusted [R-squared] was selected for modeling. Covariates for Vancouver were subsequently forecasted from historic climatological data from Vancouver. The forecasted values for covariates from Vancouver were then fit to the multiple linear regression model and fitted values yielded forecasted bloom dates for Vancouver.

Use superscript "2" rather than "-squared"

As a second method, a vector autoregressive (VAR) model was fit to historic data from Washington, Kyoto, and Liestal. PBD for Washington, Kyoto, and Liestal were treated as the endogenous variables in the model and so the model was a K=3 dimensional VAR. Forecasted peak bloom dates for each of the three cities was extrapolated from the model. We then took the average of the peak bloom days over forecasted values for the three cities as our prediction for the peak bloom dates in Vancouver to 2032.

Statistical analysis was performed in R, version 4.1.2 (11-01-2021), "Bird Hippie."

What is bird hippie?

Results

Global: No need to add "data" when referencing data; redundant phrasing

For three locations with historical peak bloom dates **data** available, October 1st was found as the optimal dormancy initiation date and Table 2 summarizes the optimal parameters found. These parameters were subsequently used to calculate heating days, and chill days from historic temperature data.

Location	Tc (°C)	Rc (chill days)	Rh (anti-chill days)
Kyoto	6	-125	235
Liestal-Weideli	5	-68	210.5
Washington, D.C.	5	-13.5	264.6

Table 2. Estimated Parameters for the Cherry Blossoms per the Phenology Model

Coefficients and standard errors for multiple linear regression models fit to Kyoto, Washington, and Liestal are shown in Tables 3, 4, and 5 respectively. Forecasting results from extrapolation of these models are shown in Table 6. As can be seen from Table 3, GDD was highly significant to predicting peak bloom times in Kyoto. Average temperature also had a p-value less than 0.05. The model shows some signs of underfitting. Unfortunately, the data available for Kyoto was limited.

The model for Washington, DC is a much more developed model. GDD was still the most significant predictor by far for predicting peak bloom day. But several other predictors were also significant to the model at the 95% confidence level. Average temperature was significant at the 0.01 level. Sea level pressure, wind speed, cumulative precipitation, and absolute temperature difference were all significant at the 0.05 level. These covariates match more closely with our expectations from research that has been done on the topic previously. The ample availability and quality of the data for this site allowed us to build a fuller, better fitting model that explains more of the variance we see in peak bloom dates.

Table 3. referenced with no corresponding Tables 1 & 2

Table 3. Multiple Linear Regression Model: Kyoto, Japan				
	β	s.e.	t	p
(Intercept)	219.44	172.028	1.28	0.207
Avg. Temperature	-0.01	0.003	-2.51	0.015 *
Chill days	0.02	0.011	1.95	0.056
GDD	4.25	0.171	24.80	<2e-16 ***
Total hrs. sunlight	-0.01	0.004	-1.74	0.087
SLP (Jan-Mar)	-2.00	1.681	-1.19	0.238
AIC	230.8			
Adj. R	0.917			

This table looks really professional and well-done

*** p < 0.001 ** p < 0.01 * p < 0.05

Was the data scaled in any way?

The s.e. on the intercept seems really, really high. If all variables were zero, you have a confidence range of multiple years!

Table 4. Multiple Linear Regression Model: Washington, DC

	β	s.e.	t	p	
(Intercept)	442.30	2084.000	0.21	0.833	
SLP	-3.38	1.614	-2.10	0.043	*
Cum. precipitation	0.01	0.004	2.13	0.039	*
Wind speed	-0.95	0.395	-2.42	0.021	*
Avg. temperature	-0.01	0.003	-3.09	0.004	**
Abs Temperature Diff	0.12	0.054	2.22	0.032	*
Chill days	-0.04	0.020	-2.00	0.052	
GDD	4.36	0.344	12.68	0.000	***
Sunlight (Jan-Mar)	18.05	59.670	0.30	0.764	
Sunlight (May-Sep)	-20.24	107.900	-0.19	0.852	
AIC	201.3				
Adj. R	0.931				

*** p < 0.001 ** p < 0.01 * p < 0.05

The linear regression model for Lie Typo? has the data as can be seen from Table 5. Only GDD and average hours of sun is it are significant in this model. GDD is highly significant at the 0.001 level. Average sunlight hours is significant at the 0.01 level.

Tables 7, 8, and 9 show the coefficients and standard errors for the multiple linear regression models with SARIMA adjusted errors. For Kyoto, the model of best fit had 5 non-seasonal autoregressive lags, 4 non-seasonal moving average coefficients, 1 seasonal autoregressive lag and 2 seasonally adjusted moving average terms. The model showed seasonality with a period of 5 years. In Table 7, we note that the coefficient for GDD was considerably larger than its standard error and is therefore highly significant to the model. In fact, GDD dominates over all other covariates in the model. SLP is also significant in this model, but it is not as important to this model as is it for DC. We also observe that the 5th order lag is highly important for this model.

"it was for the model fitted to the Washington D.C. peak bloom data"

How is the seasonal component defined?

Table 5. Multiple Linear Regression Model: Liestal, Switzerland

	β	s.e	t	p	
(Intercept)	-798.70	273.100	-2.92	0.005	**
Sunlight (May-Sep)	54.77	18.650	2.94	0.005	**
Abs temperature diff	0.02	0.012	1.38	0.173	
GDD	4.47	0.064	70.34	<2.00E-16	***
Avg. temperature	0.00	0.001	-1.54	0.128	
Chill days	0.01	0.007	1.29	0.202	
AIC	162.6				
Adj. R	0.996				

*** p < 0.001 ** p < 0.01 * p < 0.05

Consider including a table plotting the residuals. I'd be curious to see if these really, really high adjusted R-squared values are representative of the model fit

Table 7. Coefficients for Multiple Linear Regression Model with SARIMA Adjusted Errors for Cherry Blossom Peak Bloom: Kyoto, Japan

Model: SARIMA (5,0,4) (1,0,2) [5]			
	β	s.e.	
AR1	0.324	0.077	
AR2	-0.424	0.075	
AR3	-0.331	0.075	
AR4	0.352	0.066	
AR5	-0.865	0.074	
MA1	0.164	0.144	
MA2	1.243	0.179	
MA3	-0.133	0.169	
MA4	0.427	0.160	
SAR1	0.327	0.138	
SMA1	1.191	0.158	
SMA2	1.000	0.166	
Intercept	427.477	17.613	
Drift	-0.027	Inf	
Avg. Temperature	-0.003	0.001	
Chill Days	0.044	0.003	
GDD	3.737	0.029	
Total Sunlight Hrs. (Jan-Mar)	-0.007	0.001	
SLP (Jan-Mar)	-3.975	0.186	
σ^2	0.3817		
AIC	186.58		
BIC	231.26		
Log lik	-73.29		
RMSE	0.5229		
MAPE	0.4424		

Again, I really love these professionally made tables.

What scale is this on? Were your predictions off by 0.44% or 44.24%?

The model for Washington, DC is quite similar and has 6 non-seasonal and 3 seasonal autoregressive terms, and 4 non-seasonal and 5 seasonal moving average terms. As can be seen from Table 8, the coefficients for GDD, wind speed, and SLP were therefore highly significant to predicting peak bloom day. It is unfortunate that data for wind speed and SLP were not available for all sites as it would have been interesting to see whether these covariates are similarly significant in other locations. As for the SARIMA coefficients, it is clear that the seasonally adjusted autoregressive terms, particularly SAR2, SAR4, and SAR5, are highly important to the model's predictive power. As with Kyoto, we observe a 5-year seasonal cycle with peak bloom day in DC. The model for Liestal differs from the other two sites.

Where are these located in the tables?

Table 8. Coefficients for Multiple Linear Regression
Model with SARIMA Adjusted Errors for Cherry Blossom
Peak Bloom: Washington, DC

Model: SARIMA (6, 0, 4) (3, 0, 5) [5]			
	β	s.e.	
AR1	-1.765	Inf	
AR1	-1.331	Inf	
AR1	-1.384	0.068	
AR1	-1.841	Inf	
AR1	-1.223	Inf	
AR1	-0.392	0.090	
AR1	1.060	0.108	
AR1	1.030	0.091	
AR1	0.826	0.119	
AR1	-0.136	0.079	
AR1	-1.241	0.013	
AR1	-0.923	0.007	
AR1	-0.153	0.012	
AR1	-0.925	0.007	
AR1	-1.241	0.013	
AR1	-0.999	Inf	
AR1	0.930	Inf	
AR1	-0.691	Inf	
AR1	-0.674	Inf	
Intercept	4842.508	Inf	
Drift	0.034	0.001	
SLP (Jan-PBD)	-6.635	0.043	
Cum. Precipitation	0.015	Inf	
Wind Speed	-0.983	0.004	
Avg. Temperature	-0.011	Inf	
Mean Absolute Temp Diff			
(*0.1)	0.081	0.002	
Chill Days	0.006	0.001	
GDD	4.264	0.019	
Avg. Hrs. Sunlight (Jan-Mar)	-97.944	Inf	
Avg. Hrs. Sunlight (May-Sep)	-221.542	Inf	
sigma^2	0.0017		
AIC	94.28		
BIC	152.92		
Log lik	-16.14		
RMSE	0.0257		
MAPE	0.0228		

Typo? Don't believe you intended to write the model as having only one autoregressive component

The model for Liestal has 1 non-seasonal and 2 seasonal autoregressive lags, and 0 non-seasonal and 5 seasonal moving average coefficients. The period for seasonality is only 1 year but the model coefficients and fit differ from a non-seasonal ARMA(3, 5) model. GDD dominates over all other covariates in this model and the seasonally adjusted autoregressive order 2 lag is highly significant.

Table 9. Coefficients for Multiple Linear Regression Model with SARIMA Adjusted Errors for Cherry Blossom Peak Bloom: Liestal, Switzerland

Model: SARIMA (1, 0, 0) (2, 0, 5) [1]			
	β	s.e.	
AR1	0.323	0.172	
SAR1	-0.374	0.016	
SAR2	-1.000	0.004	
SMA1	-0.198	0.230	
SMA2	1.081	0.958	
SMA3	-1.173	0.477	
SMA4	0.021	0.968	
SMA5	-0.723	0.681	
Intercept	-930.457	199.144	
Avg. Hrs. Sunlight (May-Sep)	63.711	13.546	
Mean Absolute Temp Diff (*0.1)	0.002	0.013	
GDD	4.533	0.073	
Avg. Temperature	-0.001	0.002	
Chill Days	0.011	0.005	
sigma^2	0.3848		
Log lik	-60.14		
AIC	150.27		
BIC	183.56		
RMSE	0.5506		
MAPE	0.4049		

Why so many different covariates across locations? I understand there is missing data and not all sites have the same data, but why not consider a model with fixed effects and random effects? For instance, sometimes you use average hrs sunlight and sometimes total hours. Was there something in the data that made you consider this approach?

Table 10 shows a comparison of the predictive power of the linear models and the linear models with SARIMA adjusted errors and shows the root mean squared error (RMSE) of the actual peak bloom dates compared to the PBDs predicted by the model. The RMSE is much lower for the SARIMA adjusted models, indicating that these models have a much better fit to the historical data. We argue that the SARIMA adjusted models give more accurate forecasts due to the fact that these models are able to detect seasonal trends whereas simple linear models are not. Global weather patterns such as El Niño are well known to occur cyclically with quasi-predictive

No need to include "We are that"...you're the ones writing the paper! We, the reader, will trust it is you making the argument if you state a conclusion or inference

seasonality. We believe that modeling bloom dates—which are closely tied to cumulative heating and cooling temperatures—are likely to behave with quasi-regular seasonality as well.

Table 10. Comparison of Predictive Power with Linear Regression Model and Linear Regression Model with SARIMA Adjusted Errors

	RMSE Linear Model	RMSE SARIMA Model
Kyoto		0.5229
Liestal	0.7462	0.5506
Washington	1.5085	0.0257

I really like showing the statistical evidence against each other. This allows me as the reader to see the data in action rather than take your word for it. Nice job.

Forecasts for the day of peak bloom from linear regression models with SARIMA adjusted errors are shown in Table 11. Forecasts from the VAR model are shown in Table 12.

"Forecasts based on a linear regression model for the peak bloom date"

Remove space between the table and the paragraph

Table 11. Forecasts for Cherry Blossom Peak Bloom (2022-2032): Linear Regression Model

Year	Washington, DC	Kyoto, Japan	Liestal, Switzerland	Vancouver, BC
2022	98.3	97.2	114.7	91.6
2023	93.4	99.3	116.9	90.0
2024	93.4	94.2	117.3	89.2
2025	89.8	98.8	84.5	86.9
2026	89.3	100.5	114.4	88.5
2027	107.8	97.5	91.4	90.4
2028	94.2	91.0	94.3	90.2
2029	93.9	93.5	79.0	88.2
2030	96.9	98.0	116.8	88.9
2031	97.3	103.4	114.3	87.5
2032	101.8	99.8	107.4	88.5

Table 12. Forecasts for Cherry Blossom Peak Bloom (2022-2032): VAR Model

Year	Kyoto, Japan	Washington, DC	Liestal, Switzerland	Vancouver, BC
2022	94.5	90.0	91.7	92.0
2023	95.0	92.1	99.5	95.5
2024	96.0	92.7	100.7	96.5
2025	96.2	93.0	101.5	96.9
2026	96.3	93.1	101.8	97.0
2027	96.3	93.1	101.9	97.1
2028	96.4	93.1	101.9	97.1
2029	96.4	93.1	101.9	97.1
2030	96.4	93.1	101.9	97.1
2031	96.4	93.1	101.9	97.1
2032	96.4	93.1	101.9	97.1

Discussion

Missing text for the conclusion section

Conclusion

References

(1) Miller-Rushing AJ, warming on a group of Takao, Japan. American

(2) Chung U, Mack L, DC and mid-Atlantic st

(3) Cesaraccio C, Span of crop and forest speci

(4) Jung JE, Kwon EY, Phonology Model. Kor

(5) Luedeling E, Kunz A, Blanke MM. Identification of chilling and heat requirements of cherry trees—a statistical approach. International Journal of Biometeorology. 2013 Sep;57(5):679-89.

(6) Luedeling E, Fernandez E. Statistical methods for phenology analysis in temperate fruit trees[r package chillr version 0.72.7] [Internet]. The Comprehensive R Archive Network.

Thank you for giving me the opportunity to read your paper. Your team put in a lot of work and it seems like you're very comfortable with time series models. Would love to read a longer discussion on the justifications of the model you chose and the approach.

The paper jumps around a bit, and I suspect it is due to multiple people writing the paper, but try to have one reviewer read through and edit to synthesize as one voice. The paper jumps around between casualness in some sections to highly statistical jargon in the next, so it can be difficult to follow along.

Lastly, typically academic papers will provide some light discussion on the theoretical background of a statistical approach taken. Would love to have read your teams' thoughts on why an SARIMA model was chosen. Is there a theoretical reason why this data fit that approach?

al
t Mt.

ngton,
439.

burst
.

t
55.

Comprehensive R Archive Network (CRAN); 2022 [cited 2022Mar7]. Available from:
<https://cran.r-project.org/web/packages/chillR/>