

## Table of contents

### 1 Introduction

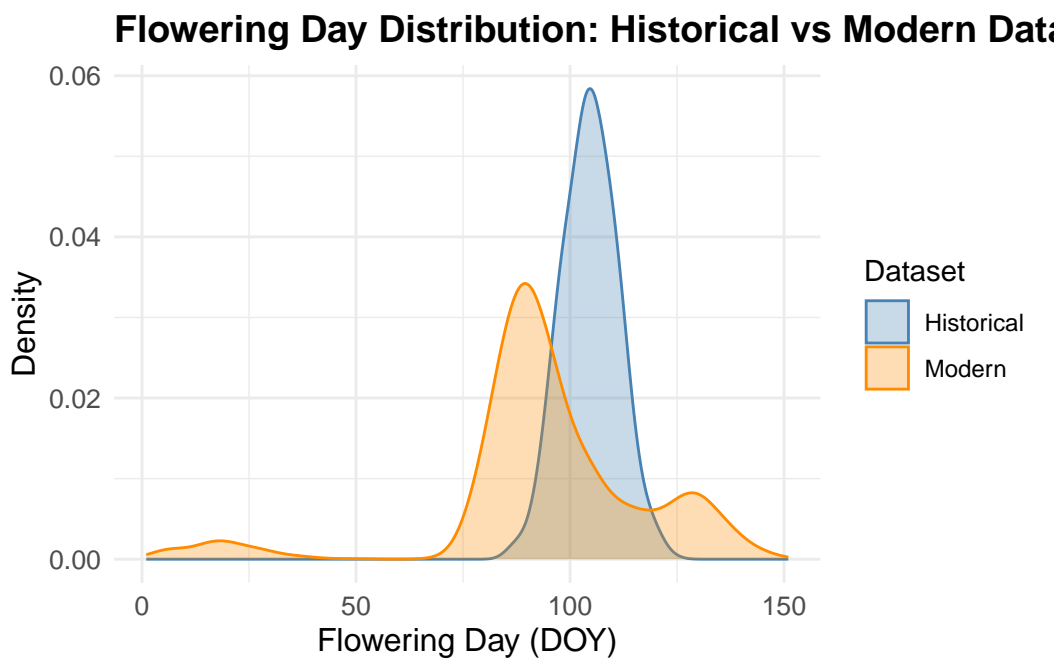
### 2 Data

#### 2.1 Overview

#### 2.2 Measurement

#### 2.3 Outcome Variables

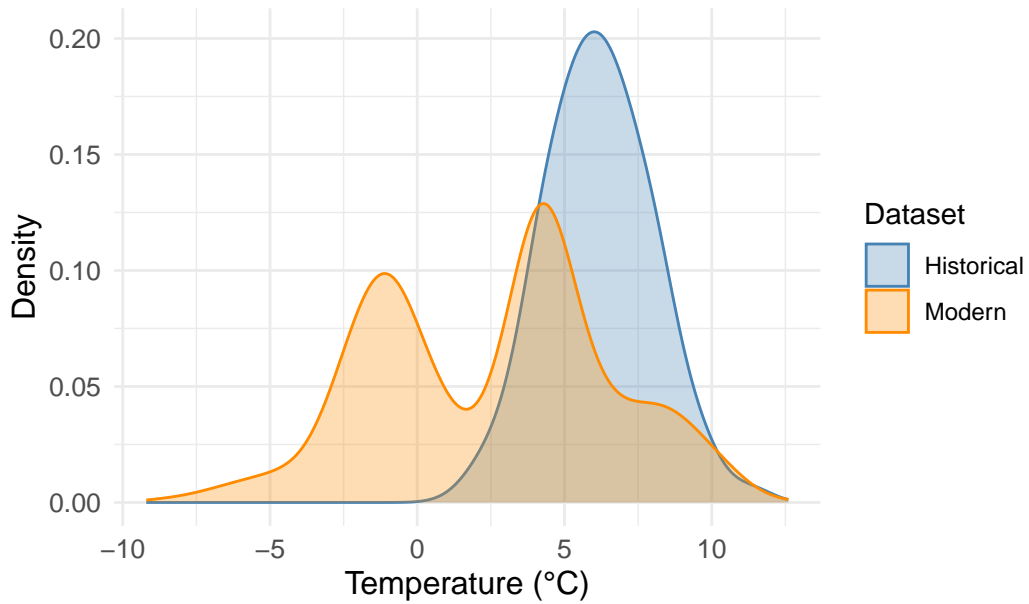
##### 2.3.1 Florescence



## 2.4 Predictor Variables

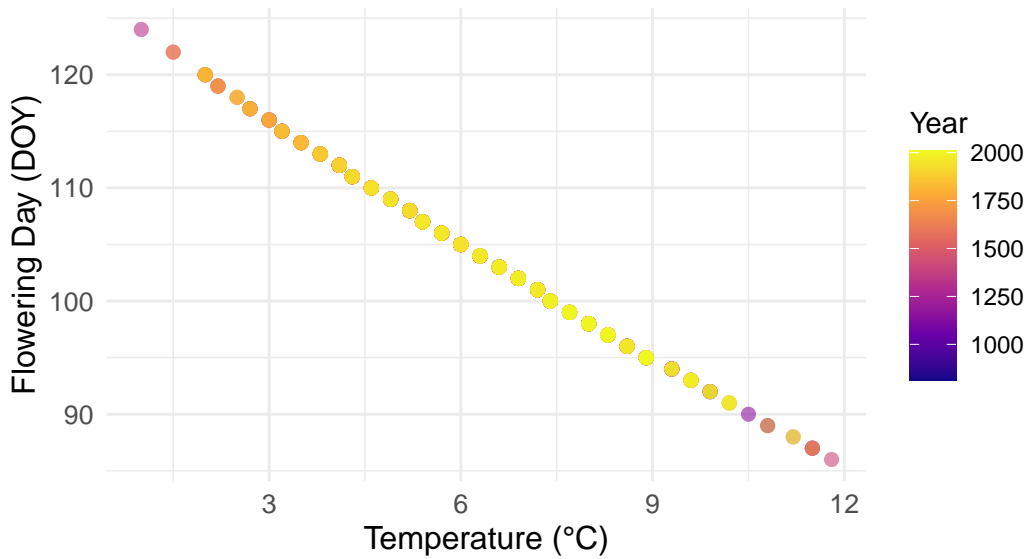
### 2.4.1 Temperature

**Temperature Distribution: Historical vs Modern Data**



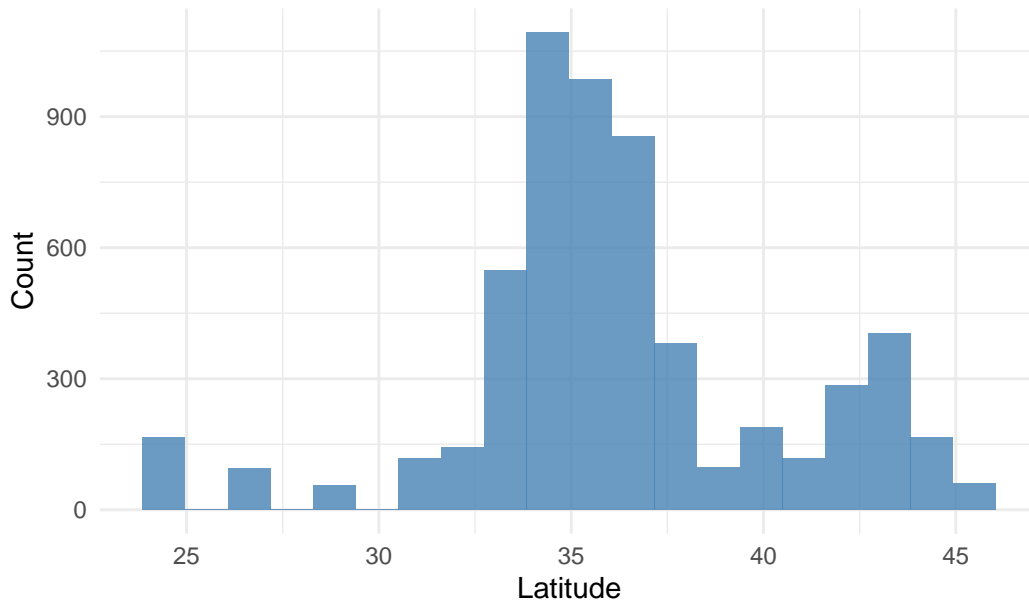
**Relationship Between Temperature and Flowering Day**

Color represents temporal progression



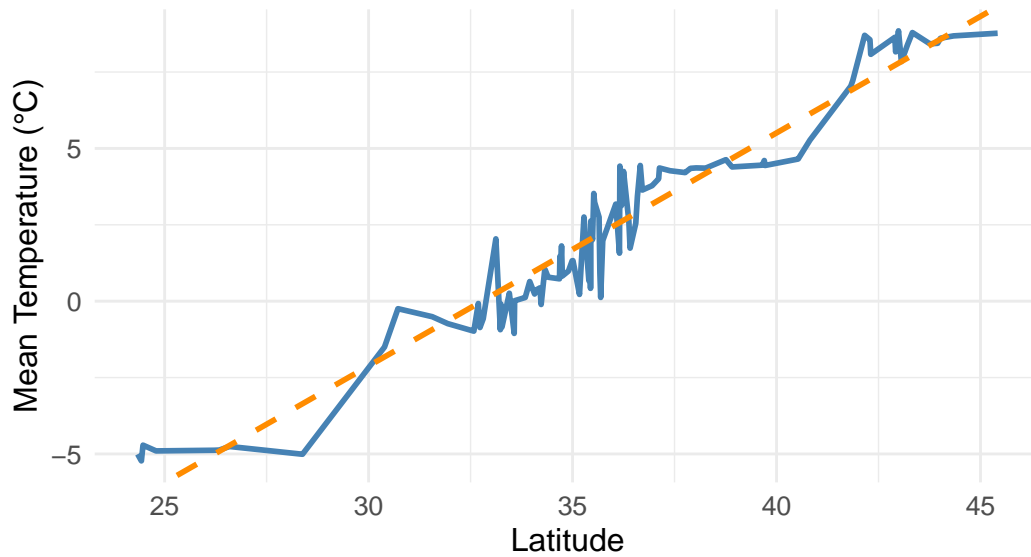
## 2.4.2 Latitude

Latitude Distribution



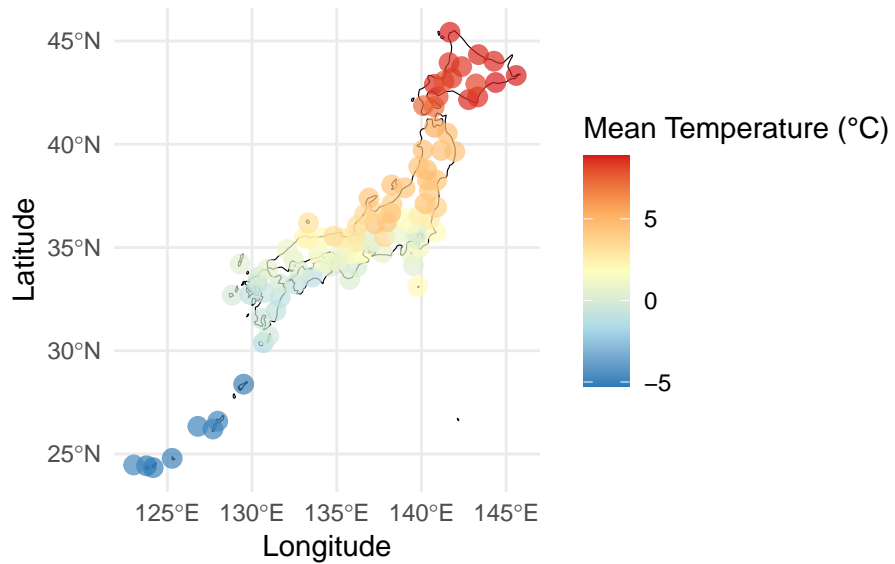
### Mean Temperature by Latitude

Relationship between Latitude and Mean Temperature with Linear Ti



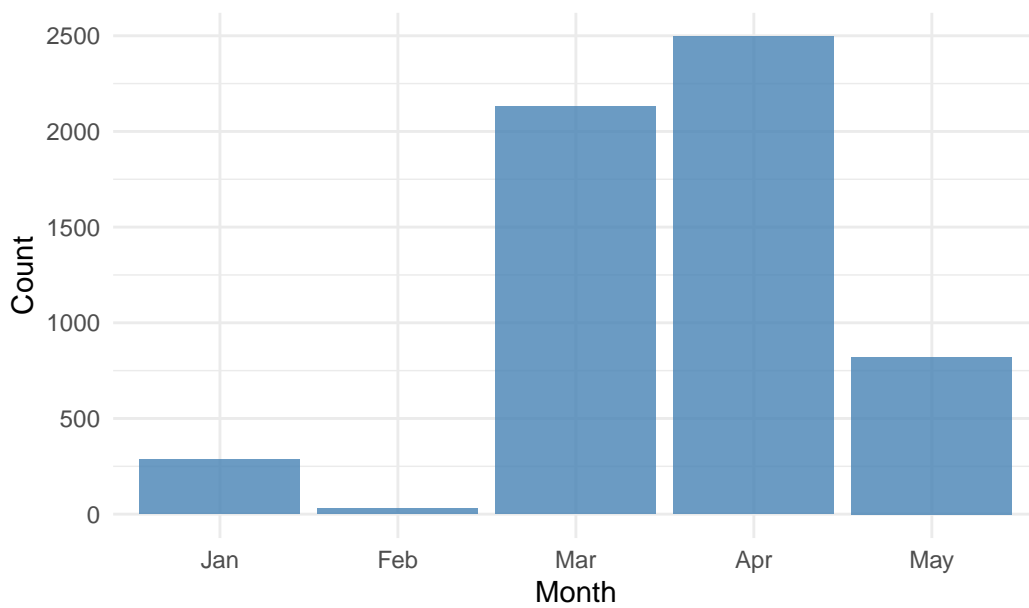
## Average Monthly Temperatures Across Japan

Warmer colors indicate higher temperatures

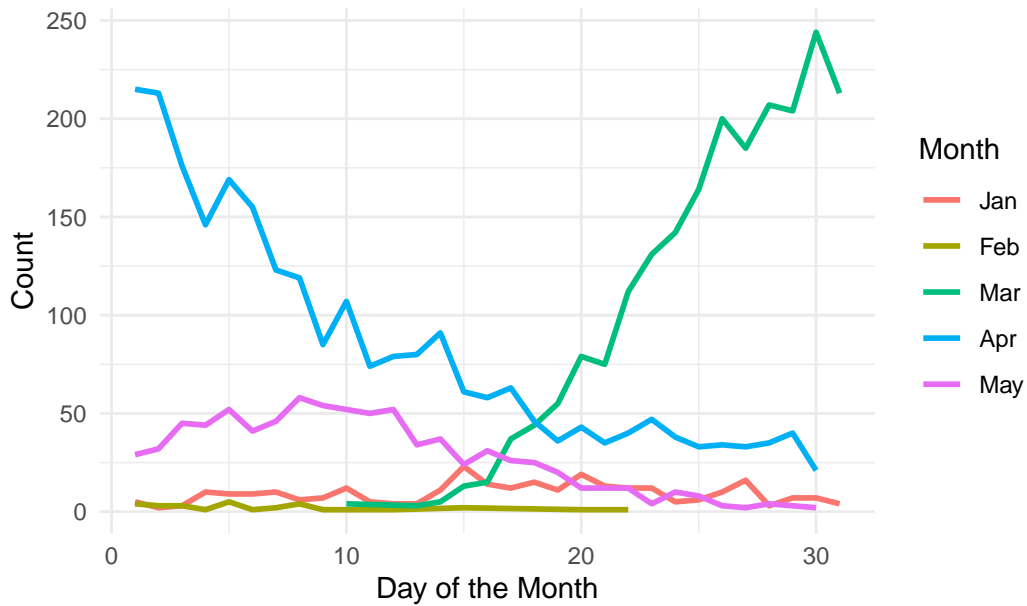


### 2.4.3 Time of Blossom

#### Monthly Distribution of Flower Date

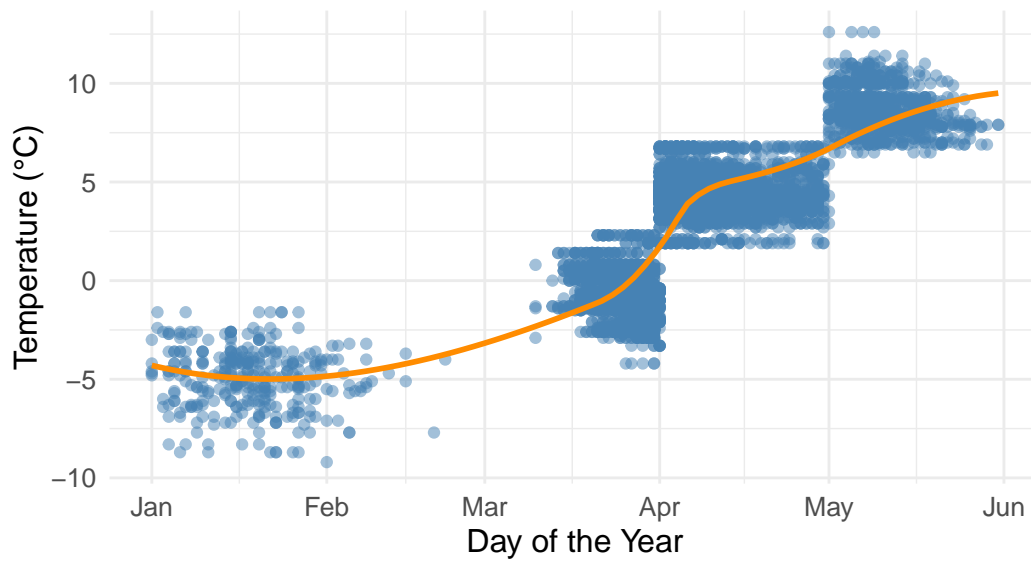


Flower Date Distribution by Month and Day



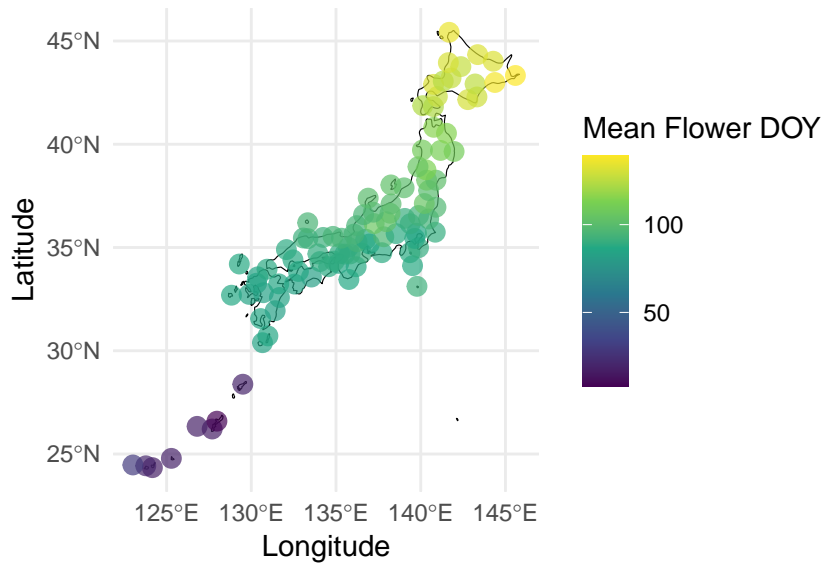
Temperature Distribution by Date

Scatter plot with temperature trend over the year



## Average Flowering Dates Across Japan

Darkers points indicate later flowering dates

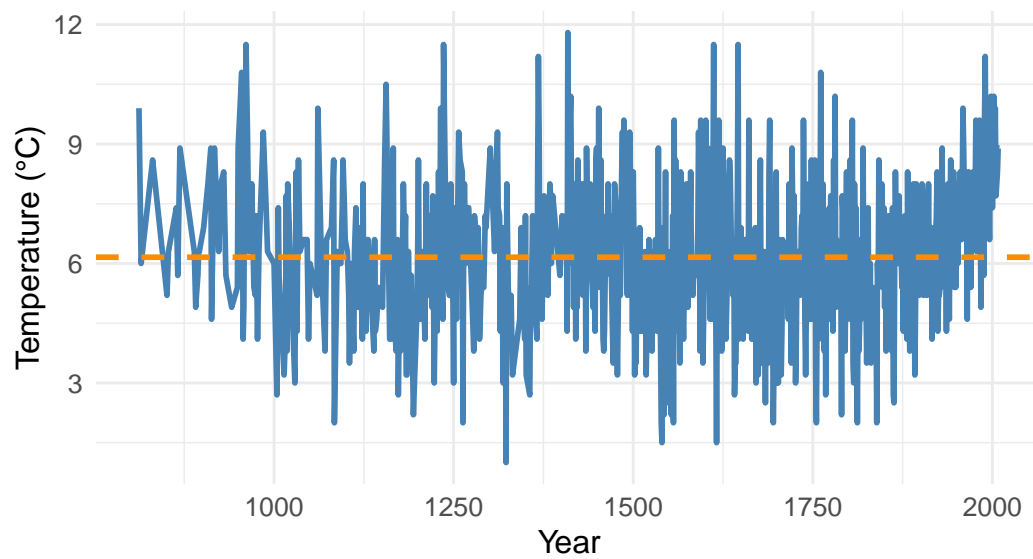


## 2.5 Excluded Variables

### 2.5.1 Year

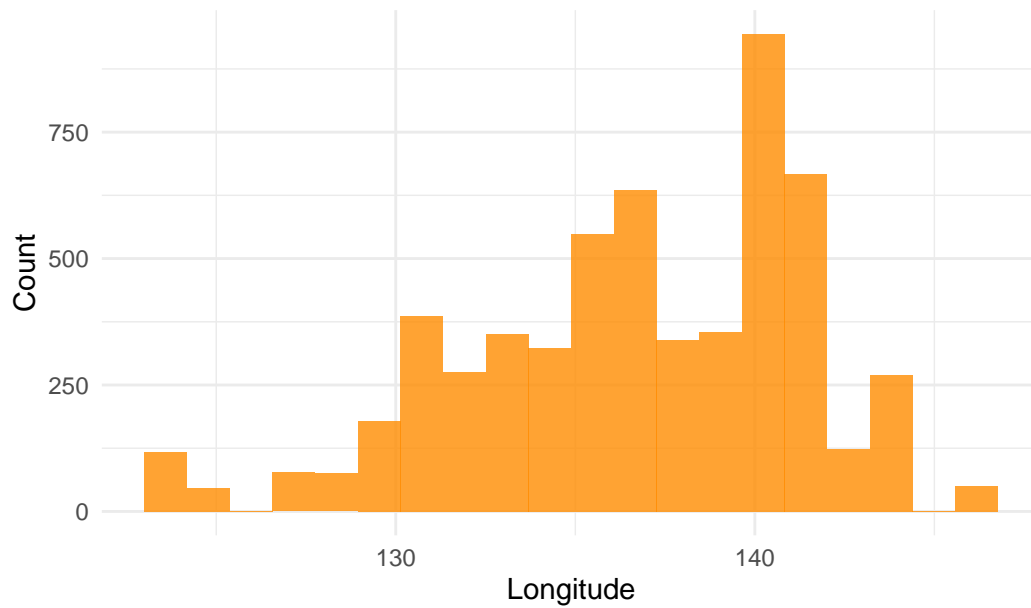
#### Temperature Trend Over Years

Yearly average temperature with reference line



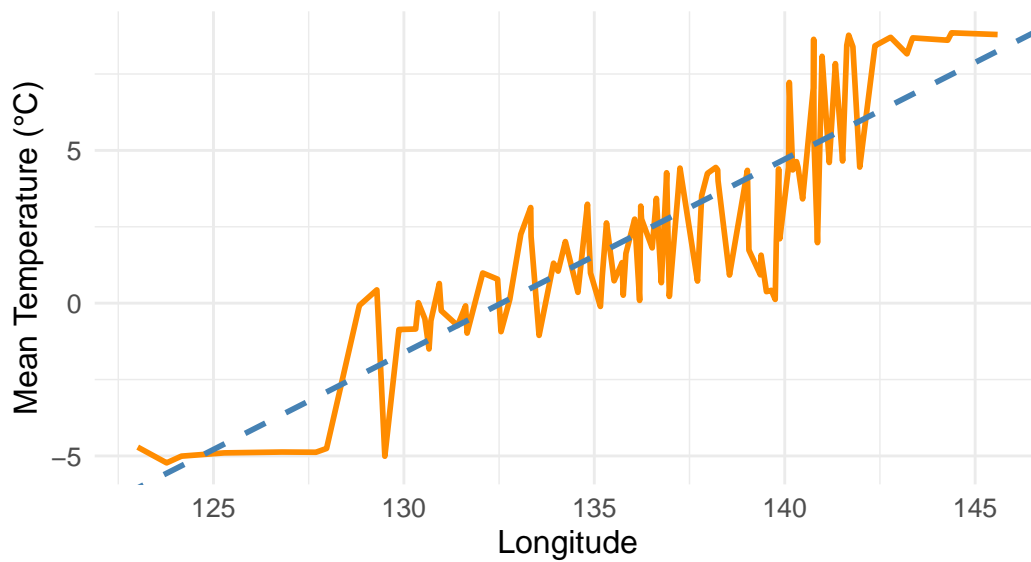
## 2.5.2 Longitude

Longitude Distribution



**Mean Temperature by Longitude**

Relationship between Longitude and Mean Temperature with Linear



### 3 Model

#### 3.1 Alternative Model

$$\text{Alternative Model: mean\_temp\_month} = \beta_1 \cdot \text{latitude} + \beta_2 \cdot \text{longitude} + \beta_0 \quad (1)$$

where:

- latitude  $\in [20, 50]$ , representing geographical latitude in degrees ( $^{\circ}\text{N}$ ).
- longitude  $\in [120, 150]$ , representing geographical longitude in degrees ( $^{\circ}\text{E}$ ).
- $\beta_1$  and  $\beta_2$  are coefficients for latitude and longitude, respectively.
- $\beta_0$  is the intercept term.
- Both latitude and longitude are numerical variables.

Variable	Model Summary	
	Estimate	P-Value
latitude	0.778101	<2e-16
longitude	-0.003319	0.806
R-squared	0.672200	
Adjusted R-squared	0.672100	

#### 3.2 Model 1

$$\text{Model 1: flower\_doy} = \beta_1 \cdot \text{temp} + \beta_0 \quad (2)$$

where:

- temp  $\in [1, 12]$ , representing the temperature range in degrees Celsius ( $^{\circ}\text{C}$ ).
- $\beta_1$  is the coefficient of the variable temp, and  $\beta_0$  is the intercept term.
- The variable temp is a numerical variable.



### 3.3 Model 2

$$\text{Model 2: mean\_temp\_month} = \beta_1 \cdot \text{day} + \beta_2 \cdot \text{latitude} + \beta_3 \cdot \text{month} + \beta_0 \quad (3)$$

where:

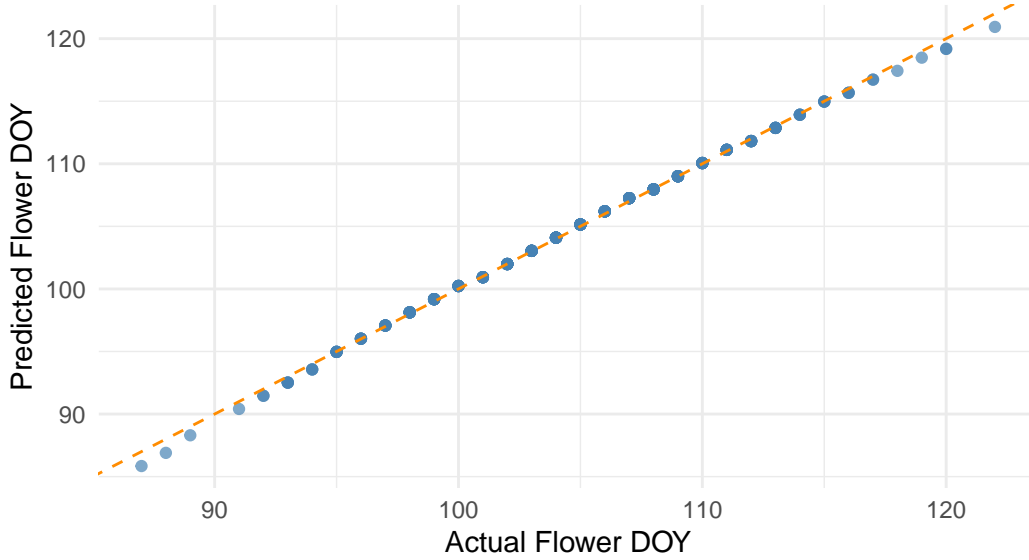
- latitude  $\in [20, 50]$ , representing geographical latitude in degrees ( $^{\circ}\text{N}$ ).
- day  $\in [1, 31]$ , representing the day of the month.
- month is a categorical variable representing months (January to May).
- $\beta_1, \beta_2, \beta_3$  are coefficients of the linear model, and  $\beta_0$  is the intercept term.
- All variables are numerical variables, except for month, which is a categorical variable.

## 4 Result

### 4.1 Model Result

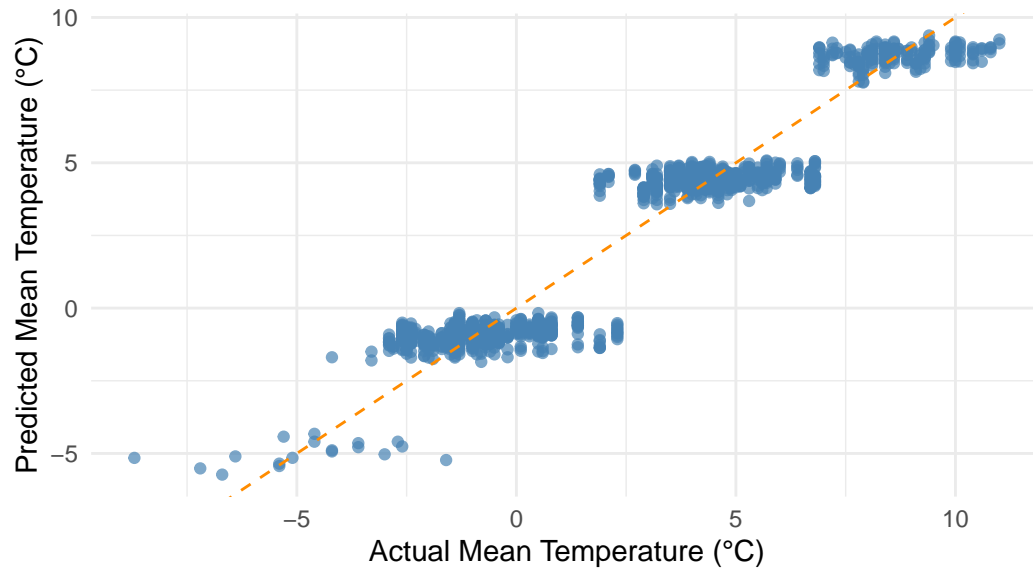
#### Model 1: Actual vs Predicted Flower DOY

Prediction accuracy for flower day using temperature data



## Model 2: Actual vs Predicted Mean Temperature

Prediction accuracy for mean temperature using latitude & time data



### 4.2 Example of Prediction

Metric	Prediction Summary	
	Actual	Predicted
Mean Temperature (°C)	3.9	4.254538
Flower DOY	119.0	111.271828

## **5 Discussion**

### **5.1 Limitation**

### **5.2 Interpretation and Future Research Directions**

## **6 Appendix**

### **6.1 Survey Methodology Overview**

### **6.2 Idealized Survey**

### **6.3 Model Details**

Residuals for Model 1:

Residuals for Model 1	
Statistic	Value
Min	-0.25208
1Q	-0.14651
Median	-0.04095
3Q	0.04606
Max	1.30709

Coefficients for Model 1:

Coefficients for Model 1				
Variable	Estimate	Std. Error	t Value	P-Value
(Intercept)	126.202188	0.034396	3669.1	<2e-16
temp	-3.509279	0.005355	-655.3	<2e-16

Model Summary for Model 1:

Model Summary for Model 1	
Metric	Value
Residual Standard Error	0.2337 on 572 degrees of freedom
Multiple R-squared	0.9987
Adjusted R-squared	0.9987
F-statistic	4.294e+05 on 1 and 572 DF
P-value	<2e-16

Residuals for Model 2:

Residuals for Model 2	
Statistic	Value
Min	-4.2185
1Q	-0.6993
Median	-0.0729
3Q	0.6641
Max	4.1114

Coefficients for Model 2:

Coefficients for Model 2				
Variable	Estimate	Std. Error	t Value	P-Value
(Intercept)	-9.000439	0.289464	-31.093	<2e-16
day	-0.049223	0.003291	-14.956	<2e-16
latitude	0.189083	0.012089	15.637	<2e-16
month2	-0.617982	0.219831	-2.811	0.00496
month3	2.869383	0.114364	25.088	<2e-16
month4	6.966108	0.164984	42.223	<2e-16
month5	10.047363	0.237453	42.313	<2e-16

Model Summary for Model 2:

Model Summary for Model 2	
Metric	Value
Residual Standard Error	1.116 on 4256 degrees of freedom
Multiple R-squared	0.9255
Adjusted R-squared	0.9254
F-statistic	8814 on 6 and 4256 DF
P-value	<2e-16