Sakura Blossom Prediction Model for Japan*

Forecasting Sakura Blossom Using Bayesian Spline Regression

Shanjie Jiao

November 20, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

Table of contents

*Code and data are available at: https://github.com/Jie-jiao05/Sakura-Blossom-Prediction-Model.

1 Introduction

Sakura is not merely a ornamental plants but also hold profound cultural significance. In Japanese literature, poetry, and art, sakura blossoms carry deep emotional and symbolic meaning, with the aesthetic concept of "mono no aware" being particularly notable. Due to their short blooming period, sakura blossoms are often seen as a metaphor for the impermanence and fleeting beauty of life, evoking deep reflection and appreciation for the essence of existence.

Beyond their cultural significance, sakura blossoms also have a significant positive impact on Japan's economy. "Ohanami" (sakura blossom viewing) is a traditional celebration of spring that attracts a large number of domestic and international visitors every year during the blooming season from April to May. According to research by Katsuhiro Miyamoto, a professor at Kansai University, the 2024 cherry blossom season is projected to contribute up to ¥1.14 trillion (approximately $7.7 billion) to Japan's economy (Kaneko (2024)). This event not only supports the post-pandemic recovery of the tourism sector but also positively impacts related industries such as catering and retail.

Given the importance of sakura blooming times for tourism planning and economic activities, accurately forecasting these dates is essential. This study aims to utilize linear regression and Bayesian spline methods to systematically analyze the effects of temperature and geographical location on sakura blooming times. By developing a predictive model, the study seeks to provide scientific insights for sakura enthusiasts worldwide, as well as for tourism and related industries, facilitating more precise planning of viewing activities and resource allocation.

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2….

2 Data

2.1 Overview

We use the statistical programming language R (R Core Team 2023)…. Our data (**shelter?**)…. Following Alexander (2023), we consider…

Overview text

2.2 Measurement

Some paragraphs about how we go from a phenomena in the world to an entry in the dataset.

2.3 Outcome variables

Add graphs, tables and text. Use sub-sub-headings for each outcome variable or update the subheading to be singular.

Some of our data is of penguins (Figure 1), from (**palmerpenguins?**).

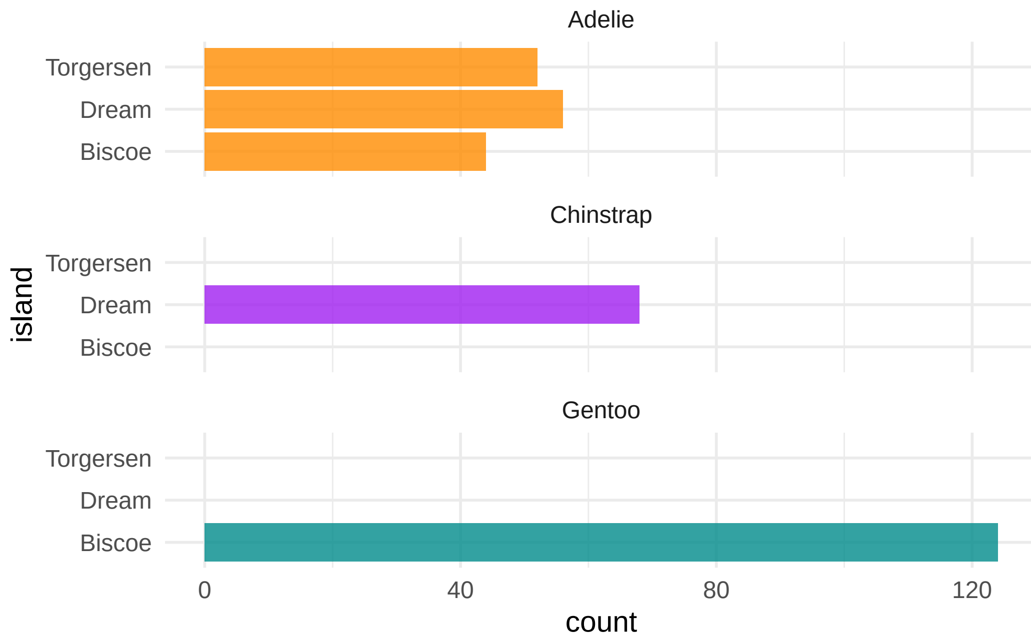


Figure 1: Bills of penguins

Talk more about it.

And also planes (**?@fig-planes**). (You can change the height and width, but don't worry about doing that until you have finished every other aspect of the paper - Quarto will try to make it look nice and the defaults usually work well once you have enough text.)

Talk way more about it.

2.4 Predictor variables

Add graphs, tables and text.

Use sub-sub-headings for each outcome variable and feel free to combine a few into one if they go together naturally.

3 Model

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (1)\\
\mu_i &= \alpha + \beta_i + \gamma_i && (2)\\
\alpha &\sim \text{Normal}(0, 2.5) && (3)\\
\beta &\sim \text{Normal}(0, 2.5) && (4)\\
\gamma &\sim \text{Normal}(0, 2.5) && (5)\\
\sigma &\sim \text{Exponential}(1) && (6)
\end{aligned}
$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of (**rstanarm?**). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ.

Table 1: Explanatory models of flight time based on wing width and wing length

	First model
(Intercept)	1.12
	(1.70)
length	0.01
	(0.01)
width	−0.01
	(0.02)
Num.Obs.	19
R2	0.320
R2 Adj.	0.019
Log.Lik.	−18.128
ELPD	−21.6
ELPD s.e.	2.1
LOOIC	43.2
LOOIC s.e.	4.3
WAIC	42.7
RMSE	0.60

4 Results

Our results are summarized in Table 1.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows…

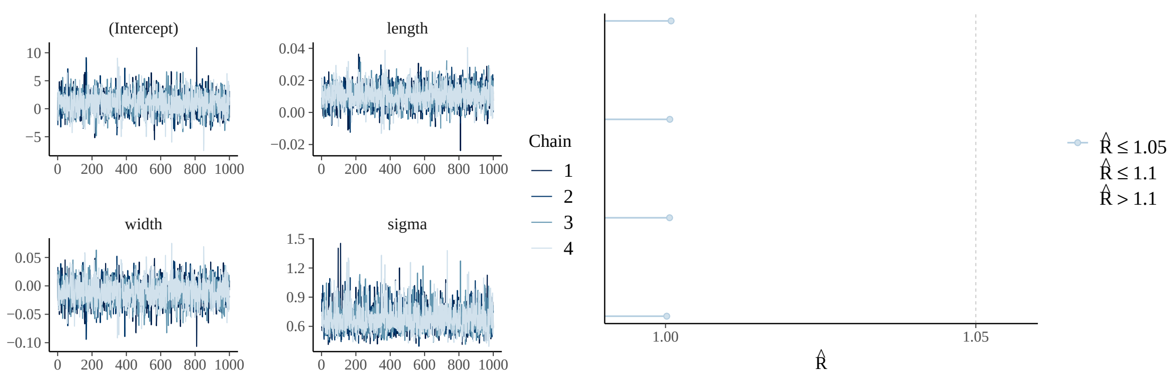
In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows…

Examining how the model fits, and is affected
by, the data

B.2 Diagnostics

Figure 2a is a trace plot. It shows… This suggests…

Figure 2b is a Rhat plot. It shows… This suggests…



(a) Trace plot (b) Rhat plot

Figure 2: Checking the convergence of the MCMC algorithm

References

Alexander, Rohan. 2023. "Telling Stories with Data." Telling Stories with Data. https://tellingstorieswithdata.com/.

Kaneko, Karin. 2024. *Economic impact of hanami expected to double this year. The Japan Times*. https://www.japantimes.co.jp/news/2024/03/15/japan/society/hanami-economic-impact/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.