Predicting the Water Temperature of Toronto Beaches*

Richard Guo

November 29, 2024

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

Table of contents

1 Introduction **2**

2 Data **2**
2.1 Overview . . . 2
2.2 Measurement . . . 2
2.3 Outcome variables . . . 3
2.4 Predictor variables . . . 3

3 Model **4**
3.1 Model set-up . . . 4
3.1.1 Model justification . . . 4

4 Results **4**

5 Discussion **6**
5.1 First discussion point . . . 6
5.2 Second discussion point . . . 6
5.3 Third discussion point . . . 6
5.4 Weaknesses and next steps . . . 6

Appendix **7**

A Additional data details **7**

*Code and data are available at: https://github.com/Richard-Guo1/predicting-beach-water-temperature.

B Model details **7**

References **8**

1 Introduction

Overview paragraph

My estimand is the efficacy of using easily accessible weather information (such as from watching a news channel) to predict the temperature of the water at a beach. By analyzing this dataset,

Results paragraph

Why it matters paragraph

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2 discusses how the dataset was obtained and analyzed. Section 3

2 Data

2.1 Overview

The data used in this paper is derived from Open Data Toronto and is read into this paper through the opendatatoronto library (Gelfand 2022). The particular data set used to analyze the observations made by city staff on the conditions of all guarded Toronto beaches between the months of May and September (Toronto Department of Parks, Forestry and Recreation 2024). All the data analysis was done through R (R Core Team 2023) with the aid of the following packages: tidyverse (Wickham et al. 2019), fastDummies(Kaplan 2023), here (Müller 2020), dplyr (Wickham et al. 2023), tibble (Müller and Wickham 2023), janitor (Firke 2024), ggplot2 (Wickham 2016), and knitr (Xie 2024).

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.

.

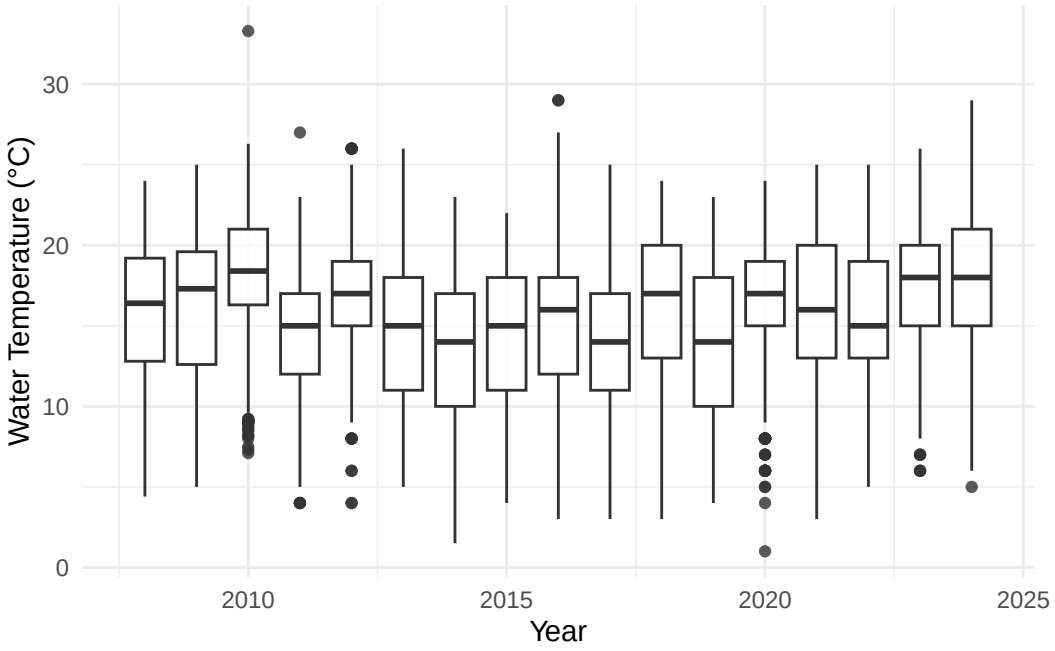

Figure 1: Water Temperature Quantiles by Year

Talk more about it.

And also planes (Figure 2). (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.

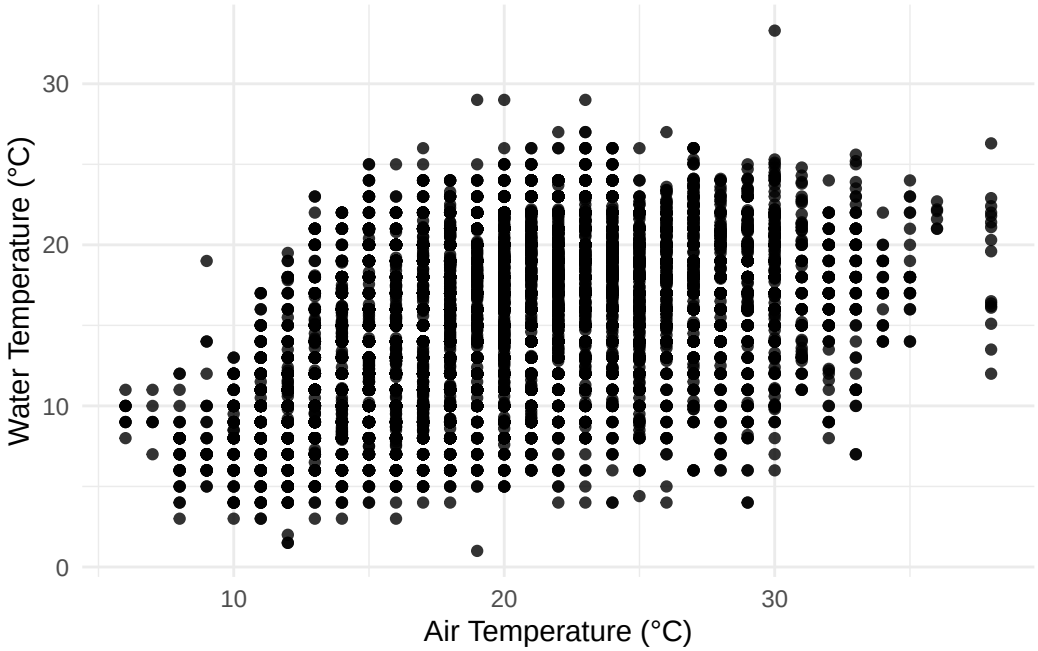

Figure 2: Relationship between Air Temperature and Water Temperature

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

3.1.1 Model justification

4 Results

Our results are summarized in Table 1.

Table 1: Explanatory models of water temperature based on air temperature, rain, and location

	(1)
(Intercept)	7.547
	(0.185)
rain	−1.737
	(0.280)
airTemp	0.359
	(0.008)
isCentreIsland	1.744
	(0.135)
isCherry	2.746
	(0.134)
isGibraltarPoint	1.325
	(0.135)
isHanlansPoint	0.986
	(0.135)
isKewBalmy	−0.382
	(0.134)
isMarieCurtis	0.403
	(0.135)
isSunnyside	0.985
	(0.133)
isWardsIsland	2.049
	(0.134)
isWoodbine	−0.313
	(0.134)
rain × airTemp	0.092
	(0.014)
Num.Obs.	15 911
R2	0.230
R2 Adj.	0.230
AIC	87 525.1
BIC	87 632.5
Log.Lik.	−43 748.536
RMSE	3.78

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

References

Firke, Sam. 2024. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*.https://sharlagelfand.github.io/opendatatoronto/, https://github.com/sharlagelfand/opendatatoronto/.

Kaplan, Jacob. 2023. *fastDummies: Fast Creation of Dummy (Binary) Columns and Rows from Categorical Variables*.

Müller, Kirill. 2020. *Here*. https://here.r-lib.org/index.html.

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://tibble.tidyverse.org/, https://github.com/tidyverse/tibble.

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

Toronto Department of Parks, Forestry and Recreation. 2024. *Toronto Beach Observations*. https://open.toronto.ca/dataset/toronto-beaches-observations/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org.

Xie, Yihui. 2024. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.