# Exploring the Relationship Between Turbidity and Environmental Factors: A Linear Regression Analysis of Beach Observation Data in Toronto (2015-2023)*

Bolin Shen

March 25, 2024

This study investigates the "Beach Observation" dataset sourced from the Toronto Open portal conducted by the City Clerk's Office from 2015 to 2023, comprising records of beach observations. The main objective is to explore the relationship between turbidity and key environmental factors, namely rain amount, wind speed, water temperature, and air temperature. Employing a linear regression model, the investigation reveals significant associations between turbidity levels and variations in the rain amount, wind speed, air temperature, and water temperature. Notably, rainfall amount emerges as the environmental factor exhibiting the strongest correlation with turbidity levels in water bodies. The findings suggest it is necessary for policy makers, particularly in keeping close eye to these observations for avoiding natural disasters happening. These findings underscore the importance of close monitoring of such observations by policymakers to mitigate the risk of natural disasters. While this study focuses on beaches in Toronto, future research incorporating additional variables and encompassing beaches in diverse geographical regions could yield a more nuanced understanding of the factors influencing turbidity.

## Table of contents

**1 Introduction** **2**

*Code and data supporting this analysis is available at: https://github.com/Brian031205/Exploring-Beach-Observation

**2 Data** 3
2.1 Variables . . . 3
2.2 Measurement . . . 4

**3 Model** 4
3.1 Model set up . . . 4
3.2 Model justification . . . 4

**4 Result** 4

**5 Discussion** 5
References . . . 7

# 1 Introduction

Beach observation stands as a paramount concern for city planners, policymakers, and residents alike, reflecting a commitment to safeguarding public health and safety. Through regular monitoring of beach conditions, municipalities aim to swiftly identify and address potential hazards or risks, including issuing advisories or closures if water quality or other conditions pose a threat to public well-being. The accumulation of observations over time yields a valuable dataset for analyzing trends and patterns in beach conditions, informing decision-making processes related to beach management, infrastructure improvements, and environmental protection initiatives. Compliance with regulatory standards and guidelines established by government agencies is essential, particularly for beaches designated for recreational use, ensuring they meet specific criteria to uphold their status and accessibility to the public.

This study seeks to examine the effect of environmental factors on water turbidity, a critical indicator of water quality. By investigating how changes in environmental factors influence turbidity, insights can be gained into the dynamics of beach conditions. Leveraging a dataset sourced from the Toronto Open Data portal, encompassing 853 records of beach observations, this study employs a multi-linear regression model to predict turbidity changes for Toronto beaches in 2024. Through this analysis, the aim is to quantify the impact of environmental factors on beach turbidity.

The analysis unveils a significant positive relationship between environmental factors and water turbidity, with rainfall amount and wind speed exhibiting stronger positive correlations with turbidity than air and water temperature. These findings underscore the imperative for regular inspections to uphold beach safety, emphasizing the need to prioritize inspections during extreme rainy and windy conditions. Furthermore, the study provides insights into urban development and environmental protection, suggesting that policymakers may need to reassess beach management strategies. It emphasizes the importance of ongoing monitoring and evaluation by urban authorities to sustain and enhance beach observation efforts citywide.

These insights carry substantial implications for urban planning and environmental protection, highlighting the necessity of proactive measures to shield beaches from natural disasters.

In conclusion, this paper significantly advances our understanding of the relationship between turbidity and various environmental factors, emphasizing the importance of proactive strategies to maintain Toronto's beaches. The subsequent sections of this paper delve into the data collection and cleaning process, the model employed for analysis, the results of the regression model, and a discussion of the limitations and potential avenues for future research.

# 2 Data

The data I used in this paper are obtained from the Open Data Toronto Portal, accessed through the library `opendatatoronto` (Gelfand 2022). The dataset, which covers polls conducted by the city from 2015, was first cleaned before analyzing by the open-source statistical programming language R (R Core Team 2022), using functionalities from `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2022), `readr` (Wickham, Hester, and Bryan 2022), `tibble` (Müller and Wickham 2022), `modelsummary` (Arel-Bundock 2022), `here` (Kirill Müller 2020), `janitor` (Firke 2021) and `knitr` (Xie 2014).

## 2.1 Variables

In the field of environmental research, understanding the intricate interplay between various factors is important. In this study, I explore the dynamics shaping water quality in Toronto's water bodies. At the investigation, the pivotal variable is turbidity. Turbidity, representing the clarity of water, stands as the dependent variable under scrutiny, serving as a lens through which we analyze the impact of several independent variables. In this study, I identify five key variables, each contributing unique insights into the complex ecosystem of Toronto's aquatic environments. While turbidity takes center stage as the focal point of our analysis, our exploration extends beyond its singular influence. Rainfall amount, wind speed, air temperature, and water temperature emerge as crucial independent variables, each wielding its own potential to shape turbidity levels over the eight-year period spanning from 2015 to 2023. This study aims to discover the intricate relationships between these variables, shedding light on the intricate dynamics governing water quality in Toronto's aquatic ecosystems. Through examination and statistical analyses, I strive to uncover the interactions between turbidity and its environmental counterparts, providing valuable insights for environmental management and policy making.

## 2.2 Measurement

The data utilized in this study were sourced from the Open Data Toronto Portal, accessed through the library `opendatatoronto` (Gelfand 2022). These data encompass beach observation records spanning from 2015 to 2023, meticulously collected and maintained in accordance with the city's standards by-law, aimed at ensuring public health and safety in Toronto. The beach observations were conducted by Bylaw Enforcement Officers, who utilized rigorous methodologies for the measurement and collection of data, adhering to established protocols to maintain consistency and validity. Measurement were taken at regular intervals. Consistent recordings of the data facilitate the explorations of the potential influence of the environmental factors on water turbidity. The dataset resulting from these measurements offers valuable insights into the dynamics of turbidity under the influence of air and water temperature, wind speed and rainfall amount. It is beneficial for researchers and analysts looking to investigate interactions between turbidity and environmental factors, and provide valuable insights for environmental management and policy making.

# 3 Model

The objective of the model is to discern the relationship between winter speed, rainfall amount, air and water temperature with water turbidity and to quantify the relationship's strength and direction. The study utilizes a single linear regression model to analyse the dataset obtained from the Toronto Open Data portal.

## 3.1 Model set up

## 3.2 Model justification

# 4 Result

These four figures provide a clear visual relationship between the water turbidity and rainfall amount, wind speed, air and water temperature. Displayed with rainfall amount, wind speed, air and water temperature scores on the x-axis and turbidity score on the y-axis, the data points represent individual beaches. The line running through the cluster of points represents the linear regression line, which captures the average effect of environmental factors' score on turbidity. Notably, the slope of the line is upward, suggesting a positive correlation, indicating that the scores of rainfall amount, wind speed, air and water temperature get higher, the turbidity tend to increase. This trend is consistent across the dataset, indicating a general increase in scores as turbidty increases.

The scatterplots reveal insights into the relationships between turbidity and various environmental factors. Notably, it appears that the correlation between turbidity and rainfall amount, as well as wind speed, exhibits a stronger association compared to turbidity with air and water temperature. There are several potential explanations for this observed pattern. Firstly, rainfall and wind speed may directly influence the transport and sedimentation of particles in water bodies, thereby impacting turbidity levels more directly. Additionally, the effects of rainfall and wind speed on turbidity might be more immediate and pronounced due to their dynamic and fluctuating nature, leading to more apparent correlations in the scatterplots. Conversely, air and water temperature may exert more indirect or delayed effects on turbidity, such as through influencing biological processes or chemical reactions in the water, resulting in weaker correlations in the plotted data. Furthermore, the observed differences in correlation strength could also be attributed to the spatial and temporal variability of these environmental factors, highlighting the complex interplay of multiple variables in shaping turbidity levels in aquatic ecosystems.

# 5 Discussion

The visual analysis of the dataset provides valuable insights into the relationship between water turbidity and various environmental factors, shedding light on the dynamics of water quality in Toronto's aquatic ecosystems. The upward slope of the linear regression line across all four figures indicates a positive correlation between turbidity and the scores of rainfall amount, wind speed, air temperature, and water temperature. This consistent trend underscores the influence of environmental factors on turbidity levels and offers several implications for understanding and managing water quality in urban coastal areas.

The stronger correlations observed between turbidity and rainfall amount, as well as wind speed, compared to air and water temperature, suggest that precipitation and wind dynamics play pivotal roles in shaping turbidity levels. Rainfall and wind speed may directly impact the transport and suspension of particles in water bodies, leading to increased turbidity levels. These findings emphasize the need for robust stormwater management strategies to mitigate the adverse effects of urban runoff and erosion on water quality in coastal areas.

Furthermore, the immediate and pronounced effects of rainfall and wind speed on turbidity highlight the importance of timely monitoring and response mechanisms to mitigate potential pollution events and safeguard public health. By incorporating real-time data on weather patterns and turbidity levels, authorities can implement proactive measures to minimize the risk of contamination and enhance recreational water safety.

Conversely, the weaker correlations observed between turbidity and air and water temperature suggest more indirect or delayed influences on turbidity dynamics. Biological processes and chemical reactions influenced by temperature variations may contribute to fluctuations in turbidity levels over longer time scales. Understanding these complex interactions is essential for developing comprehensive management strategies that address both short-term fluctuations and long-term trends in water quality.

While the findings of this study provide valuable insights into the relationship between turbidity and environmental factors, several limitations warrant consideration for future research and improvement. The analysis focused solely on linear relationships between variables, overlooking potential nonlinear dynamics and interactions. Additionally, the study's temporal scope from 2015 to 2023 may not capture longer-term trends or emerging patterns in turbidity dynamics. Future research efforts should explore more sophisticated analytical techniques and expand the temporal and spatial scope of data collection to enhance our understanding of the multifaceted drivers of water turbidity in urban coastal environments. Moreover, collaboration between researchers, policymakers, and local communities is crucial for implementing evidence-based strategies to preserve and enhance water quality for current and future generations.

## References


Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Kirill Müller, Jennifer Bryan. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Müller, Kirill, and Hadley Wickham. 2022. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

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, and Kirill Müller. 2022. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2022. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.