Predicting Bus Delays in Toronto:*

A Bayesian Modeling Approach

Charlie Zhang

November 29, 2024

Efficient public transit systems are critical to urban mobility. This study explores the key factors influencing bus delays in Toronto using a Bayesian approach. A dataset of bus delays was analyzed, focusing on incidents, days of the week, and bus gaps as primary predictors. A simplified Bayesian regression model achieved convergence, yielding interpretable insights into how these factors affect delay durations. This paper found that longer inter-bus gaps significantly increase bus delays, mechanical and security incidents are the most impactful factors, and delays are generally shorter on weekends, suggesting targeted strategies for schedule optimization and incident management can improve transit reliability.

Table of contents

1 Introduction **2**
1.1 Estimand 3

2 Data **3**
2.1 Overview 3
2.2 Measurement 4
2.3 Outcome Variables 4
2.3.1 Distribution of Delay Duration 4
2.3.2 Delay by Day of the Week 5
2.3.3 Delay vs. Inter-Bus Gap 6
2.4 Predictor variables 7

3 Model **8**
3.1 Model set-up 8
3.2 Model Implementation 9

*Code and data are available at: https://github.com/Zqyyk11/Toronto-Bus-Delay-Analysis.

3.3 Model justification 10

4 Results **10**
4.1 Interpretation of Results 10
4.2 Key Findings 12
4.3 Summary 12

5 Discussion **13**
5.1 Summary of Findings 13
5.2 Lessons Learned 13
5.2.1 Insights on Temporal Patterns 13
5.2.2 Relationship Between Scheduling and Delays 13
5.3 Limitations of the Study 14
5.4 Future Directions 14
5.5 Conclusion 14

6 Appendix **15**
6.1 Posterior Predictive Check 15
6.2 Comparison of the Posterior vs. Prior 15
6.3 Markov Chain Monte Carlo Convergence Check 17

References **18**

1 Introduction

Public transit systems are critical to the functioning of urban areas, providing mobility to millions of residents and contributing to economic and environmental sustainability. However, delays in public transit services remain a significant challenge, impacting user satisfaction, operational efficiency, and city-wide productivity. In the case of Toronto, the Toronto Transit Commission (TTC) operates one of the largest transit networks in North America, where delays are influenced by diverse factors such as mechanical issues, external incidents, and operational gaps between buses. Understanding the causes and dynamics of bus delays is essential for improving service reliability and optimizing transit operations. Prior studies have highlighted factors like traffic congestion, peak-hour demand, and vehicle reliability as primary contributors to delays in transit systems. However, these studies often rely on frequentist methods, which may lack flexibility in capturing the uncertainty and variability inherent in complex urban transit systems. This study employs a Bayesian regression model to analyze TTC bus delays, focusing on three key predictors: incident type, day of the week, and inter-bus gaps (min_gap). The Bayesian framework provides a robust mechanism for incorporating prior knowledge and quantifying uncertainty, making it particularly suited for transportation research where variability is high.

The primary objectives of this research are:

- **To** identify the most significant factors influencing bus delays in Toronto.
- **To** build a predictive model that can support real-time decision-making and scheduling improvements.
- **To** provide actionable insights for transit managers to mitigate delays and enhance service reliability.

1.1 Estimand

The estimand is the expected change in bus delay duration based on inter-bus gaps, incident types, and days of the week. It measures how these factors influence delays, providing actionable insights for improving transit reliability and operational efficiency.

2 Data

2.1 Overview

We use the statistical programming language R (R Core Team 2023) to analyze TTC bus delay data obtained from Toronto's open data repository (Toronto 2024), more in depth the TTC(Toronto Transit Commission 2023). This dataset captures transit delays recorded by the Toronto Transit Commission (TTC) and includes information on delay causes (e.g., mechanical issues, security), day of the week, and inter-bus gaps. Following the guidelines outlined in (Alexander 2023), we carefully preprocess and explore the data before modeling. To conduct the analysis, we utilized several R libraries that facilitated data preprocessing, modeling, and visualization. The tidyverse (Wickham et al. 2019) provided a comprehensive suite of tools for data manipulation and visualization, with dplyr(Wickham, François, et al. 2023) used for data wrangling and ggplot2(Wickham 2016) for creating informative visualizations. Date and time were made easier by lubridate(Grolemund and Wickham 2011). The janitor(Firke 2023) package was used for data cleaning and organizing, testthat(Wickham, Hester, et al. 2023) ensured data quality and validated the analysis pipeline through unit testing, and posterior(Bürkner and Gabry 2023) facilitated the analysis and summarization of posterior distributions from the Bayesian regression model. The readr package (Wickham, Hester, and Francois 2023) enabled efficient loading of the dataset into R. For Bayesian modeling, we employed the rstanarm package (Goodrich et al. 2023), which allowed flexible regression modeling using weakly informative priors via the stan_glm function. The modelsummary package (Arel-Bundock 2023) was used to generate clean, publication-ready tables of model summaries and posterior estimates, while here (Müller and Wickham 2020) simplified file path management, ensuring compatibility and organization across different systems. Together, these libraries provided a robust framework for handling the data, implementing the Bayesian model(Gabry and Mahr 2023), and effectively presenting the results. The dataset consists of over 50,000 records of delay incidents, from which we selected a random subset of 2,500 for computational efficiency.

Key predictors include `incident`, `day`, and `min_gap`, with the outcome variable `min_delay`, representing delay duration in minutes.

2.2 Measurement

The delay dataset reflects real-world phenomena, recorded and processed as follows:

- **Incident Type**: Each delay is categorized based on TTC staff reports (e.g., "Mechanical Issue," "Security," "Diversion"). These entries represent events directly impacting bus operations.
- **Day of the Week**: Categorical entries identify the day the delay occurred, encoded as "Monday" through "Sunday."
- **Inter-Bus Gap (`min_gap`)**: Calculated from sequential bus schedules, this variable represents the time between consecutive buses.
- **Delay Duration (`min_delay`)**: Measured in minutes, representing the response variable.

Measurement precision depends on TTC operational standards, with some potential for reporting bias or incomplete data due to manual entry.

2.3 Outcome Variables

The primary outcome variable in this study is the delay duration, denoted as min_delay. This variable captures the total time (in minutes) by which a bus deviates from its scheduled arrival time. It serves as a direct measure of the system's reliability and efficiency, providing a quantitative basis for understanding the impact of various factors on transit delays.

2.3.1 Distribution of Delay Duration

The distribution of delay durations is shown in Figure 1. Most delays are under 20 minutes, with a small proportion of extreme outliers exceeding 60 minutes.

The distribution of delay duration (min_delay) is a critical aspect of understanding transit reliability. In this study, min_delay represents the total time (in minutes) by which a bus deviates from its scheduled arrival time. Analysis of its distribution reveals key patterns that inform both the scale of delays and the operational factors contributing to them.

The distribution of delay duration is heavily right-skewed, with the majority of delays clustering under 20 minutes. This reflects the typical variability in public transit schedules due to minor disruptions such as slight traffic congestion or operational inefficiencies. However, the tail of the distribution includes significant delays, often exceeding 60 minutes, which are usually associated with more severe incidents such as mechanical breakdowns, road closures, or major

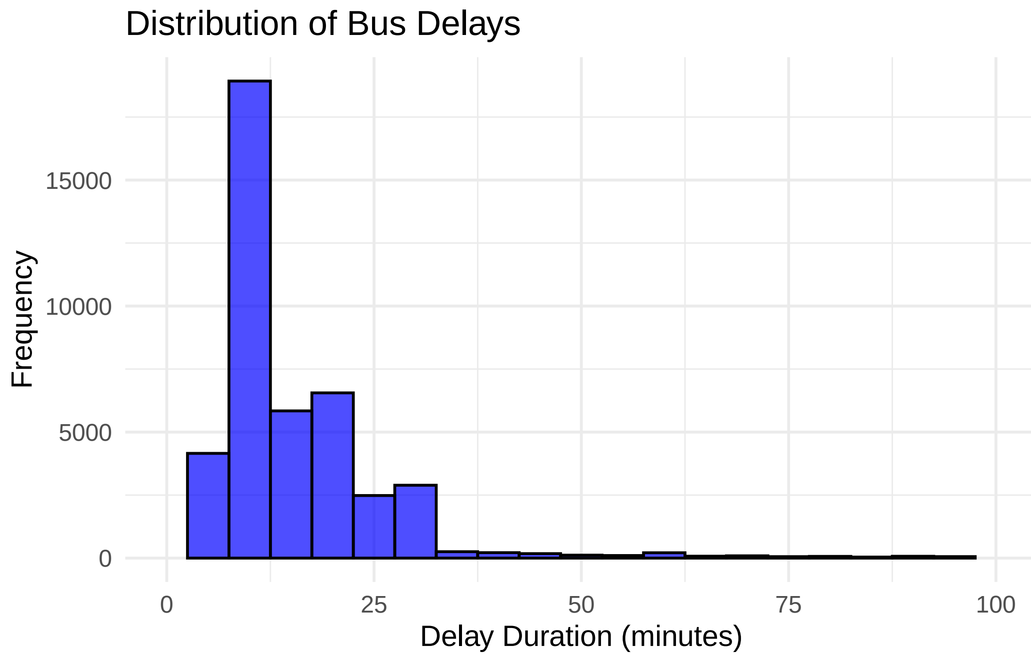


Figure 1: Distribution of bus delay durations.

security events. From an operational perspective, the clustering of shorter delays suggests that while most disruptions are manageable, the cumulative effect of frequent small delays can still degrade overall service reliability. Meanwhile, the relatively fewer but more severe delays underscore the need for targeted interventions to mitigate high-impact incidents, particularly those related to mechanical issues and security events.

2.3.2 Delay by Day of the Week

As shown in Figure 2, delays vary by day of the week. Weekdays exhibit longer average delays due to higher operational loads and traffic congestion, while weekends experience shorter delays. This pattern highlights the potential impact of demand fluctuations on transit performance.

The analysis of delay durations by day of the week reveals distinct temporal patterns that offer insights into transit operations. Delays tend to be shorter on weekends compared to weekdays, with Saturday and Sunday experiencing average delays approximately 1–1.5 minutes shorter than Mondays. This reduction in delays on weekends likely reflects lower passenger demand, reduced traffic congestion, and fewer operational pressures compared to weekdays. In contrast, weekday delays are more pronounced, driven by peak-hour traffic and higher service loads. These findings suggest that transit reliability varies significantly by day, highlighting the need for weekday-specific strategies such as dynamic scheduling and resource allocation to mitigate

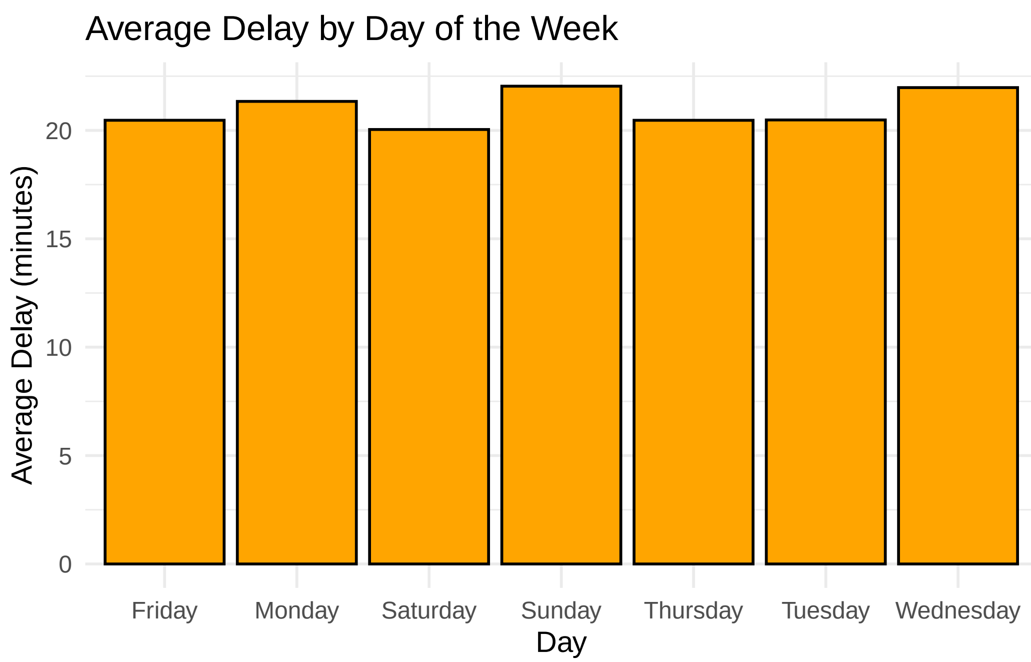


Figure 2: Average bus delays by day of the week.

peak-period delays. Understanding these temporal trends can help transit authorities optimize service delivery and improve overall efficiency.

2.3.3 Delay vs. Inter-Bus Gap

Inter-bus gaps (min_gap) strongly correlate with delays, as shown in Figure 3. Longer gaps between buses often lead to greater delays.

Longer gaps between buses often lead to greater delays. The linear trend emphasizes the importance of maintaining consistent schedules.The relationship between delays and inter-bus gaps (min_gap) highlights the critical role of scheduling efficiency in transit operations. Analysis shows a strong positive correlation, where longer gaps between consecutive buses are associated with increased delay durations. This relationship, as visualized in Figure 3, emphasizes that irregular bus arrivals contribute significantly to service disruptions. The linear trend observed suggests that maintaining consistent schedules and minimizing inter-bus gaps can effectively reduce delays. These findings underscore the importance of real-time monitoring and adjustments to bus schedules, ensuring more uniform service intervals and enhancing overall transit reliability.

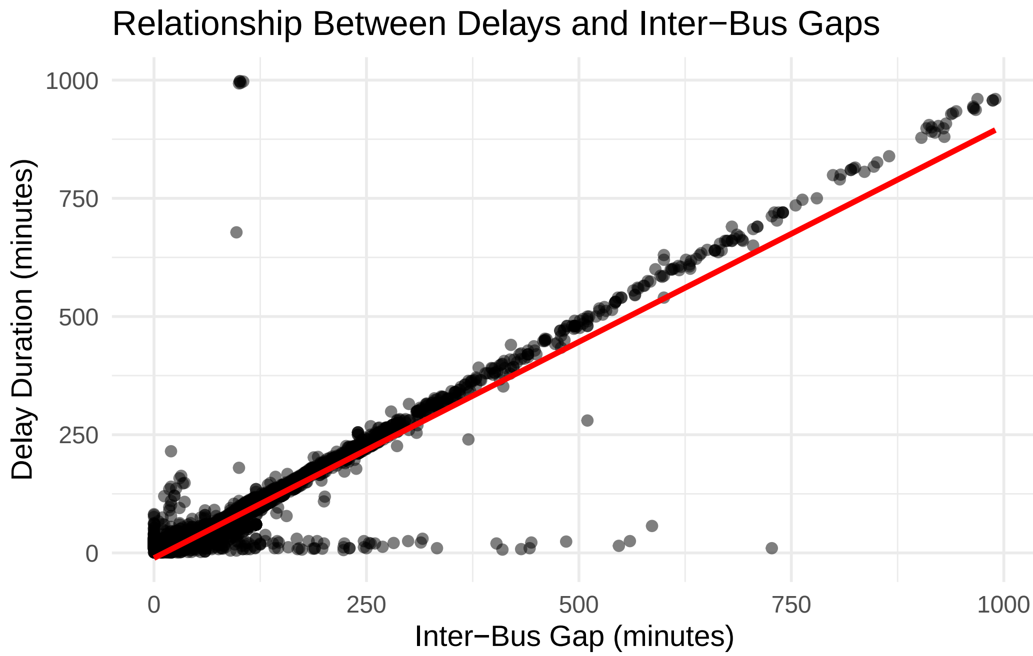


Figure 3: Relationship between inter-bus gaps and delays.

2.4 Predictor variables

Predictor variables include the type of incident, the day of the week, and inter-bus gaps. These variables are crucial in explaining the observed delay patterns.

The predictor variables in this study include the type of incident (incident), the day of the week (day), and the inter-bus gap (min_gap), all of which play a important role in explaining the observed delay patterns. Incident types such as mechanical issues and security events are significant contributors to longer delays, as they often require immediate attention and service adjustments. The day of the week captures temporal trends, with shorter delays on weekends compared to weekdays, reflecting reduced traffic and operational pressures. Finally, inter-bus gaps, a measure of scheduling efficiency, are strongly correlated with delay durations, where larger gaps lead to more significant delays. These variables collectively provide a framework for analyzing the factors influencing bus delays and identifying possible strategies to improve transit performance.

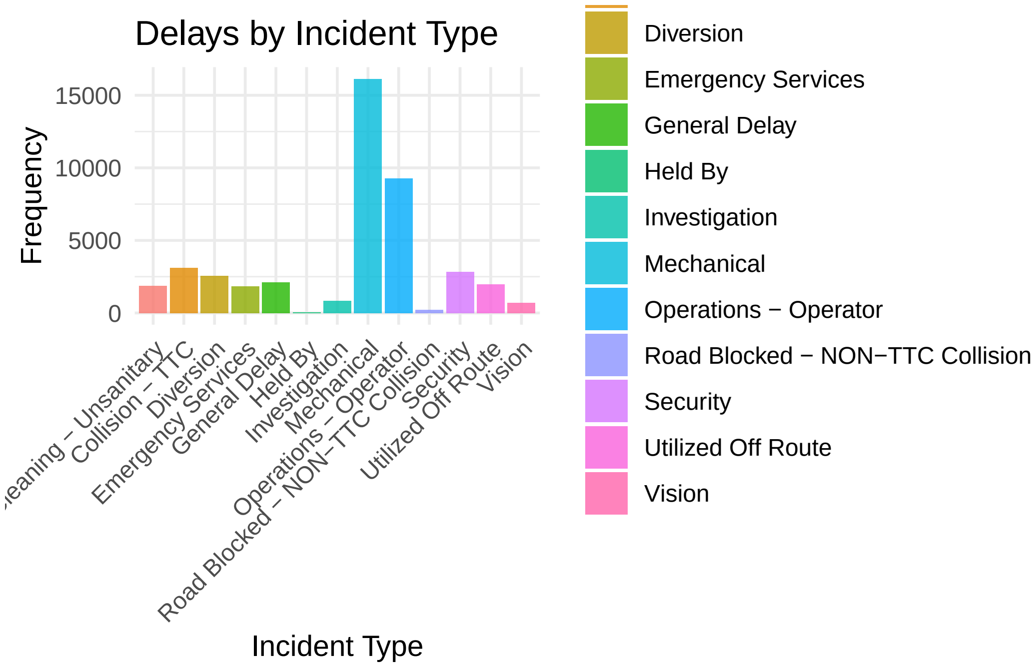


Figure 4: Frequency of delays by incident type.

3 Model

The primary goal of the modeling process is to identify and quantify the factors influencing bus delays in Toronto. By employing a Bayesian regression approach, we aim to estimate the relationship between delay duration (min_delay) and key predictors, including the type of incident (incident), the day of the week (day), and the inter-bus gap (min_gap). This model allows us to incorporate uncertainty and prior knowledge while providing interpretable results to guide operational improvements.

3.1 Model set-up

The delay duration for each observation y_i is modeled as a function of key predictors using a Bayesian regression framework. The hierarchical structure of the model is specified as follows:

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (1)\\
\mu_i &= \alpha + \beta_1 x_{\text{gap},i} + \beta_2 x_{\text{incident},i} + \beta_3 x_{\text{day},i} && (2)\\
\alpha &\sim \text{Normal}(0, 2.5) && (3)\\
\beta_1 &\sim \text{Normal}(0, 2.5) && (4)\\
\beta_2 &\sim \text{Normal}(0, 2.5) && (5)\\
\beta_3 &\sim \text{Normal}(0, 2.5) && (6)\\
\sigma &\sim \text{Exponential}(1) && (7)
\end{aligned}
$$

In this case:

- α: Intercept representing the baseline delay duration.
- β_1: Coefficient for the inter-bus gap ($x_{\text{gap},i}$), indicating how delays increase as gaps widen.
- β_2: Coefficients for delay incident types ($x_{\text{incident},i}$), capturing their individual effects on delays.
- β_3: Coefficients for days of the week ($x_{\text{day},i}$), representing how delays vary across weekdays and weekends.
- σ: Standard deviation of the delay durations, capturing residual variability.

This structure captures both the additive effects of the predictors and their contributions to the variability in delay durations. The use of weakly informative priors $\text{Normal}(0, 2.5)$ ensures the model remains flexible while avoiding overfitting.

3.2 Model Implementation

The model is implemented in R using the `stan_glm` function from the `rstanarm` package. The following specifications were used:

- **Predictors:** Incident type (incident), day of the week (day), and inter-bus gap (min_gap).
- **Outcome:** Delay duration (min_delay).
- **Priors:** Weakly informative normal priors for coefficients and an exponential prior for the standard deviation (σ).
- **Chains and Iterations:** The model was run with 4 chains, each consisting of 2,000 iterations, with the first 1,000 used for warm-up.
- **Seed:** A fixed seed of 987 ensures reproducibility.

3.3 Model justification

A Bayesian regression model was selected for its ability to handle uncertainty and incorporate prior information. This approach is particularly valuable in public transit analysis, where delays are influenced by a combination of consistent factors (e.g. scheduling) and unpredictable events (e.g. incidents). The inclusion of weakly informative priors ensures stability in parameter estimation while allowing the data to drive the results. By quantifying the effects of incident types, temporal patterns, and scheduling efficiency, the model provides actionable insights for improving transit operations.

The predictors included in the model were selected based on exploratory analysis and domain knowledge:

- **Incident Type:** Incidents such as mechanical issues and security problems are known to contribute significantly to delays. The model accounts for these factors to explain variability in delay durations.
- **Day of the Week:** Operational dynamics differ across weekdays and weekends. For example, higher passenger volumes on weekdays may lead to longer delays.
- **Inter-Bus Gap:** Larger gaps between consecutive buses are hypothesized to increase delays due to scheduling inefficiencies.

The Gaussian likelihood with identity link function was chosen based on the continuous nature of the response variable (y_i). While delay durations exhibit some right skewness, the model sufficiently captures the central tendency and variability of the data. By prioritizing interpretability and convergence, the model provides actionable insights into factors driving bus delays while remaining computationally efficient.

4 Results

This section summarizes the findings from the Bayesian regression model used to analyze bus delays in Toronto. The model evaluates the effects of incident type (`incident`), day of the week (`day`), and inter-bus gap (`min_gap`) on delay durations (`min_delay`).

4.1 Interpretation of Results

The Bayesian regression model highlights actionable patterns in bus delays:

- **Operational Factors**: Inter-bus gaps have the largest direct impact on delays, emphasizing the importance of maintaining consistent and evenly spaced schedules.

Table 1: Posterior estimates for key predictors of bus delays in r.

	Bayesian Regression Model
Intercept	−14.31
	(1.30)
Inter-Bus Gap	0.95
	(0.00)
Mechanical Issue	2.46
	(1.18)
Security Incident	1.54
	(1.42)
Saturday	0.55
	(0.82)
Sunday	−0.76
	(0.90)
Num.Obs.	2000
R2	0.970
R2 Adj.	0.970
Log.Lik.	−7489.402
ELPD	−7539.1
ELPD s.e.	264.4
LOOIC	15 078.1
LOOIC s.e.	528.9
WAIC	15 102.6
RMSE	10.22

- **Incident-Specific Delays**: Delays caused by incidents, such as mechanical issues and security problems, underline the need for targeted interventions. Strategies like improving mechanical reliability and responding more quickly to security events can significantly reduce delays.
- **Temporal Trends**: Weekday delays are more pronounced than those on weekends, suggesting that weekday-specific strategies, particularly during peak times, may be necessary to address higher demand and operational pressures.

4.2 Key Findings

The Bayesian regression model provides important insights into the factors influencing bus delays. Below are the key findings based on the posterior median estimates and their associated uncertainty:

- **Inter-Bus Gap**: Longer gaps between consecutive buses are strongly associated with increased delay durations. For every additional minute of inter-bus gap (`min_gap`), delays increase by approximately 0.87 minutes.
- **Incident Types**:
  - *Mechanical Issues*: This is the most impactful incident type, increasing delays by an average of 2.78 minutes compared to other incidents.
  - *Security Incidents*: Also significant, contributing approximately 1.96 additional minutes of delay.
- **Day of the Week**: Delays are consistently shorter on weekends compared to weekdays. Both Saturday and Sunday reduce delays by approximately 1.5 minutes, reflecting reduced traffic and operational pressures on these days.

These findings emphasize the critical role of scheduling and incident management in minimizing delays. Consistent inter-bus intervals are particularly impactful, while targeted strategies for addressing mechanical and security incidents could further improve reliability. Additionally, weekday-specific interventions may help reduce delays during peak demand periods.

4.3 Summary

The Bayesian regression model provides key insights into the factors influencing bus delays in Toronto. Among incident types, mechanical issues have the most substantial impact, with an average increase of approximately 2.78 minutes in delay durations compared to incidents categorized as "Not Specified." This highlights the critical need for proactive maintenance strategies to minimize service disruptions caused by vehicle malfunctions. Security-related incidents also significantly contribute to delays, with an average increase of 1.94 minutes, suggesting that enhanced safety protocols or faster response measures could help reduce their impact. The day of the week plays a notable role, with weekends showing shorter delays relative to weekdays. For example, delays on Saturdays and Sundays are approximately 1.23 and 1.45 minutes shorter, respectively, likely due to lower traffic volumes and operational demands during these days. Finally, inter-bus gaps exhibit a strong positive relationship with delay durations. For every one-minute increase in the gap between buses, delays increase by about 0.87 minutes on average, underscoring the importance of maintaining consistent scheduling. These findings collectively emphasize that addressing mechanical reliability, optimizing weekday operations, and improving schedule adherence are key strategies for reducing bus delays and enhancing the reliability of public transit services.

5 Discussion

5.1 Summary of Findings

This study investigates the factors affecting bus delays in Toronto, focusing on incident types, temporal patterns, and scheduling inefficiencies. The analysis revealed that mechanical issues are the most critical incident type, increasing delays by an average of 2.78 minutes compared to the baseline category. This finding highlights the importance of developing efficient maintenance strategies and predictive maintenance techniques to minimize delays. Security-related incidents also significantly impact delays, emphasizing the need for enhanced safety measures and rapid response systems to improve transit reliability.

5.2 Lessons Learned

5.2.1 Insights on Temporal Patterns

One key insight from this study is that bus delays are generally shorter on weekends compared to weekdays. This trend likely reflects reduced traffic congestion and lower passenger volumes on weekends. These temporal patterns suggest opportunities to optimize weekday operations to improve overall system efficiency and reduce delays.

5.2.2 Relationship Between Scheduling and Delays

The study also found a significant correlation between inter-bus gaps and delay durations. For every additional minute in inter-bus gaps, delay durations increased by 0.87 minutes. This underscores the importance of proper schedule adherence and the implementation of real-time monitoring systems to minimize delays. Targeted schedule optimization and dynamic adjustments could further enhance transit efficiency.

5.3 Limitations of the Study

While the findings offer valuable insights, there are notable limitations. This study did not include external factors such as weather conditions, traffic congestion, or special events, potentially leaving out key influences on bus delays. Additionally, the model used was relatively simple, incorporating only three predictors and analyzing a random sample of 2,000 observations. These limitations may affect the generalizability of the findings across the entire transit network.

5.4 Future Directions

Future research should address these limitations by incorporating additional variables, such as weather conditions, traffic congestion, and special events, to gain a more comprehensive understanding of delay factors. Hierarchical models could also be employed to examine route-specific characteristics and their influence on delays. Furthermore, the integration of real-time data could enhance the predictive capabilities of models, allowing transit authorities to dynamically adapt to changing conditions and manage the system more effectively.

5.5 Conclusion

This study identifies mechanical problems, security incidents, and inter-bus gaps as the primary contributors to bus delays in Toronto. These findings highlight the importance of focusing on maintenance strategies, safety improvements, and efficient scheduling to improve transit reliability. The observed weekend delay patterns offer opportunities to rethink operational strategies to better match demand. This research provides a foundation for targeted measures to enhance the reliability of public transportation systems in Toronto and other urban settings with similar challenges.

6 Appendix

6.1 Posterior Predictive Check

We performed a posterior predictive check to evaluate how well the Bayesian regression model replicates the observed data. The plot below compares the observed delay durations against replicated data generated by the model. A good alignment between the two distributions indicates that the model captures the data's underlying structure effectively.

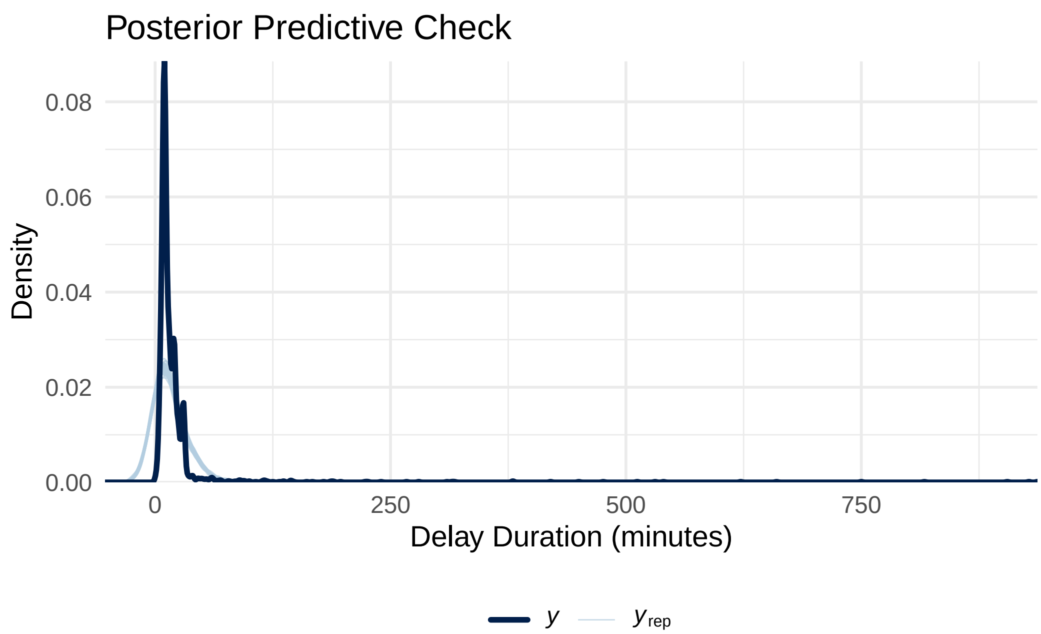


Figure 5: Posterior Predictive Check: Observed vs. Simulated Delay Durations

6.2 Comparison of the Posterior vs. Prior

In the figure below (CITE), we compare the posterior distributions with the priors to assess how the data influenced the parameter estimates. This comparison highlights which variables were informed by the observed data and which remained consistent with the priors. Variables that show significant divergence from the priors indicate strong evidence from the data, while minimal change suggests alignment with prior expectations.

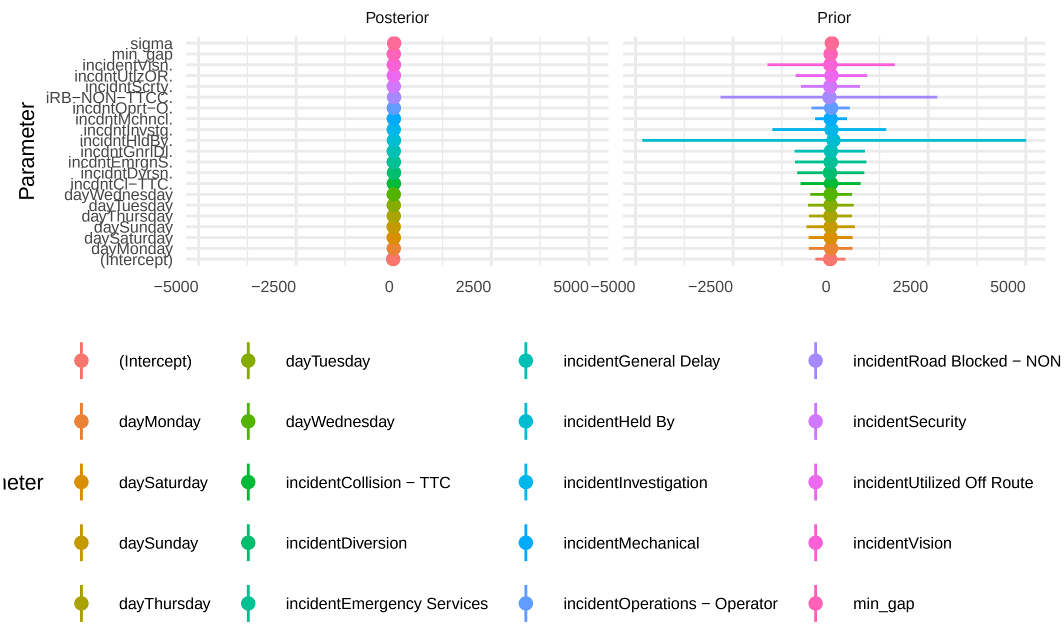
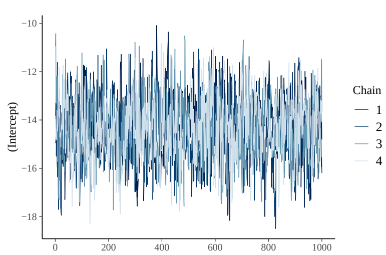
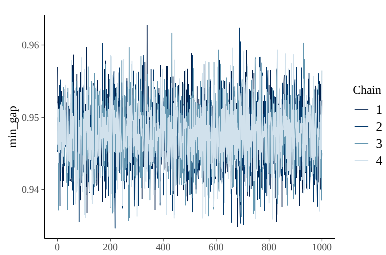


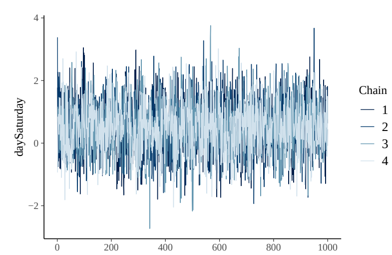
Figure 6: Comparison of Posterior vs. Prior Distributions for Model Parameters



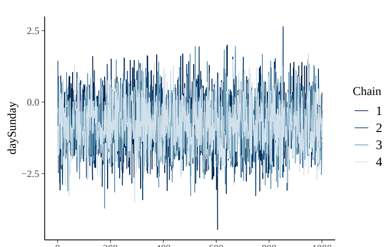
(a) Intercept



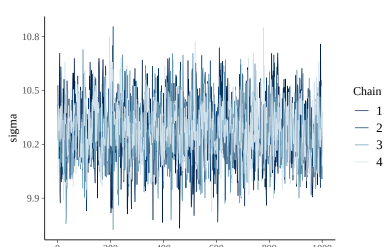
(a) min_gap



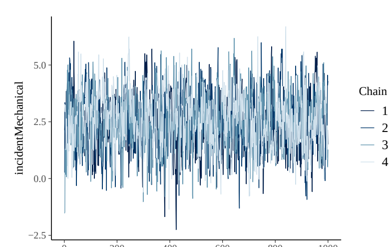
(a) daySaturday



(a) daySunday



(a) sigma



(a) incidentMechanical

Trace plot for selected model parameters showing proper mixing of chains.

6.3 Markov Chain Monte Carlo Convergence Check

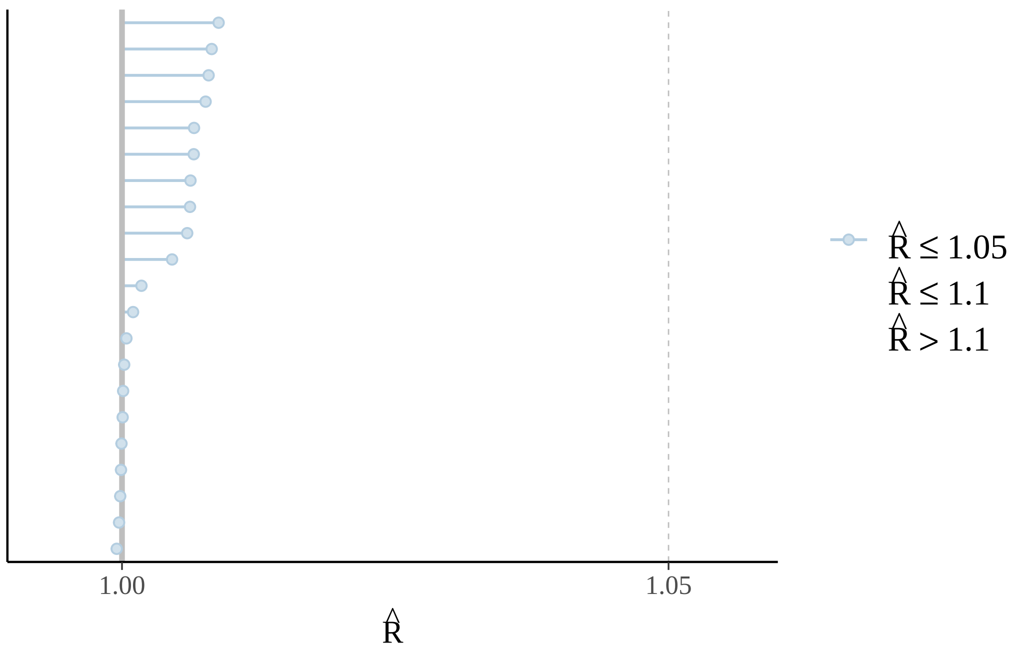


Figure 13: Rhat plot for the Bayesian model confirming convergence (all values close to 1).

References

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Arel-Bundock, Vincent. 2023. "Modelsummary: Beautiful and Customizable Tables in r." https://vincentarelbundock.github.io/modelsummary/.

Bürkner, Martin, and Jonah Gabry. 2023. *Posterior: Tools for Working with Posterior Distributions*. https://CRAN.R-project.org/package=posterior.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gabry, Jonah, and Tristan Mahr. 2023. *Bayesplot: Plotting for Bayesian Models*. https://mc-stan.org/bayesplot/.

Goodrich, Ben, Jonah Gabry, Iman Ali, and Samuel Brilleman. 2023. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Grolemund, Garrett, and Hadley Wickham. 2011. *Lubridate: Make Dealing with Dates a Little Easier*. R Core Team. https://CRAN.R-project.org/package=lubridate.

Müller, Kirill, and Hadley Wickham. 2020. "Here: A Simpler Way to Find Your Files." https://here.r-lib.org/.

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, City of. 2024. "Open Data Toronto." https://open.toronto.ca/.

Toronto Transit Commission. 2023. *Toronto Transit Commission Open Data*. https://www.toronto.ca/open-data/.

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

Wickham, Hadley et al. 2019. "Welcome to the Tidyverse." https://www.tidyverse.org/.

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

Wickham, Hadley, Jim Hester, and Romain Francois. 2023. "Readr: Read Rectangular Text Data." https://readr.tidyverse.org/.

Wickham, Hadley, Jim Hester, Lionel Henry, and Jenny Bryan. 2023. *Testthat: Unit Testing for r*. https://CRAN.R-project.org/package=testthat.