

My title*

My subtitle if needed

Lexun Yu

November 21, 2024

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

1 Introduction

Climate change is a global challenges today. Patterns such as rising temperatures, shifting weather systems, and increased frequency of severe weather events. In 2021, floods swept through streets in Japanese cities, displacing millions, while extreme heat fueled wildfires in Siberia (Greenpeace East Asia 2021). Climate change impacts human health, ecosystems, food security, water supplies, and economic stability. Understanding the factors driving temperature changes is necessary for designing effective mitigation strategies. This requires examining the various contributors to temperature variations.

Some scholars have examined the changing climate. Xu et al. (2009) analyze the effects of rising temperatures in the Himalayas, highlighting increased frequency and duration of extreme events and shifts in ecosystems. These changes pose challenges to water supply, agriculture, and human populations. Visser et al. (2021) investigates the relationship between precipitation and temperature using data from the Australian Bureau of Meteorology. Visser’s regression model indicates that average precipitation intensities increase with temperature, suggesting more intense rainfall in a warmer climate. The role of sea level pressure is also significant. Wills et al. (2022) note that observed trends in sea level pressure have intensified warming in the Indo-Pacific Warm Pool and caused slight cooling in the eastern equatorial Pacific. However, as Zhang, Zhang, and Chen (2017) argue, much of the research has focused on temperature and precipitation. Zhang, Zhang, and Chen (2017) expands on this by incorporating additional predictors—relative humidity and wind speed—and concludes, using data from the Ministry of Agriculture of China, that these variables are important in understanding climate dynamics.

*Code and data are available at: <https://github.com/yulexun/ClimateChangeYVR>.

With the data obtained from Canadian Centre for Climate Services (2022) and Meteorological Service of Canada (2023), this research paper aims to identify the factors influencing temperature at Vancouver International Airport. Located on the west coast of Richmond, the airport sits on Sea Island, surrounded by water. As a transportation hub for passengers and freight, it is important to assess the location's safety in a warming climate.

Estimand paragraph

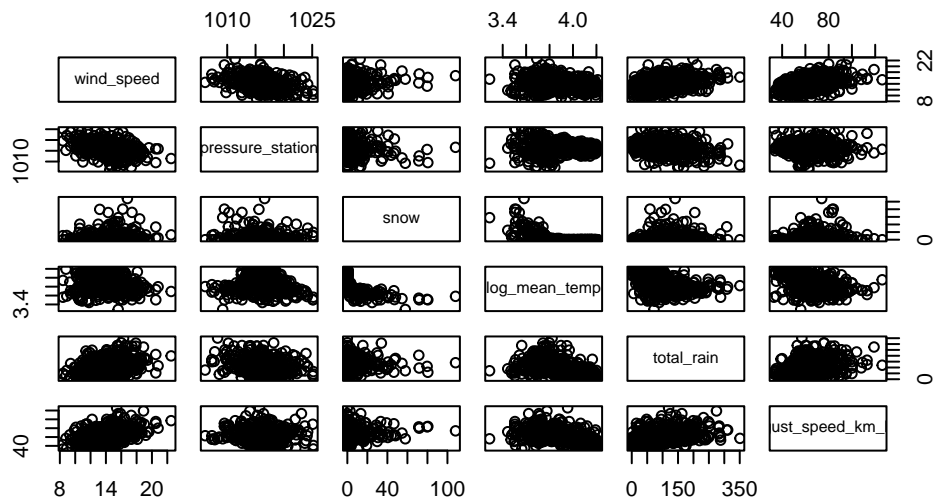
Results paragraph

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

2 Data

```
selected_data_1 <- analysis_data[, c("wind_speed", "pressure_station", "snow", "log_mean_temp",
pairs(as.data.frame(selected_data_1), main = "Pairwise Scatterplots (Excluding Max Temp, and
```

Pairwise Scatterplots (Excluding Max Temp, and Min Temp)

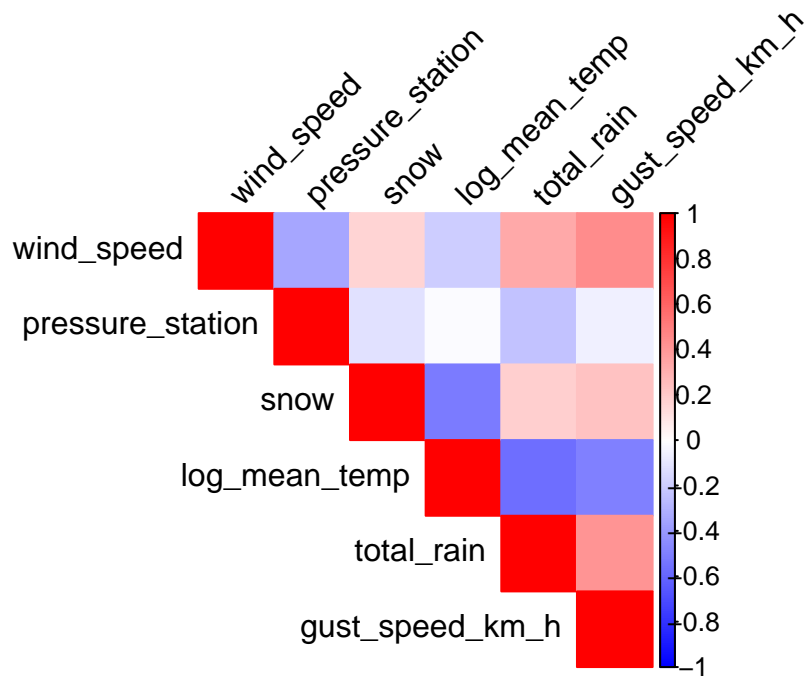


```
cor_matrix <- cor(selected_data_1)
print(cor_matrix)
```

	wind_speed	pressure_station	snow	log_mean_temp
wind_speed	1.0000000	-0.34760462	0.1648222	-0.19023484
pressure_station	-0.3476046	1.00000000	-0.1173164	-0.01396792
snow	0.1648222	-0.11731635	1.0000000	-0.51554945
log_mean_temp	-0.1902348	-0.01396792	-0.5155495	1.00000000
total_rain	0.3333823	-0.23418715	0.1843201	-0.56140560
gust_speed_kmh	0.4414225	-0.05910226	0.2300011	-0.49986221

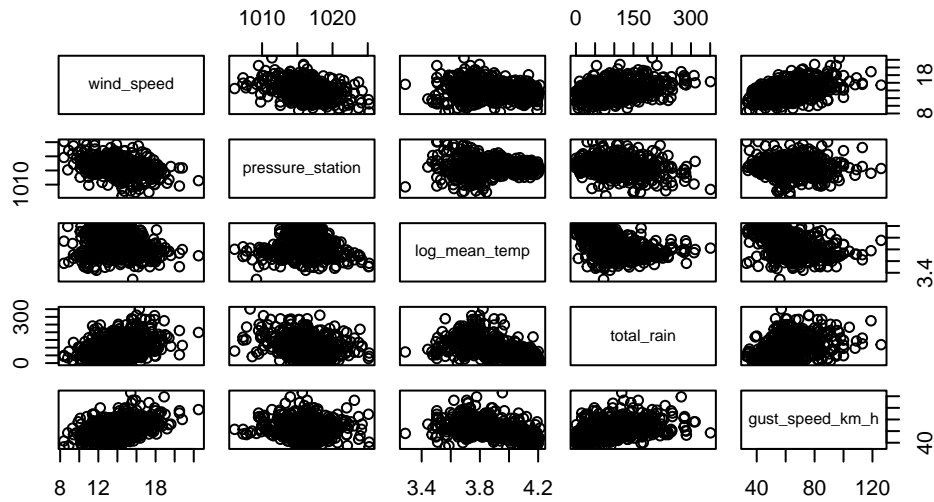
	total_rain	gust_speed_kmh
wind_speed	0.3333823	0.44142245
pressure_station	-0.2341872	-0.05910226
snow	0.1843201	0.23000113
log_mean_temp	-0.5614056	-0.49986221
total_rain	1.0000000	0.41892153
gust_speed_kmh	0.4189215	1.00000000

```
corrplot(cor_matrix, method = "color", type = "upper",
         tl.col = "black", tl.srt = 45,
         col = colorRampPalette(c("blue", "white", "red"))(200))
```



```
selected_data_3 <- analysis_data[, c("wind_speed", "pressure_station", "log_mean_temp", "total_rain", "gust_speed_kmh")]
pairs(as.data.frame(selected_data_3), main = "Pairwise Scatterplots of log transformed temper"
```

ots of log transformed temperature (Excluding Snow, Max Te



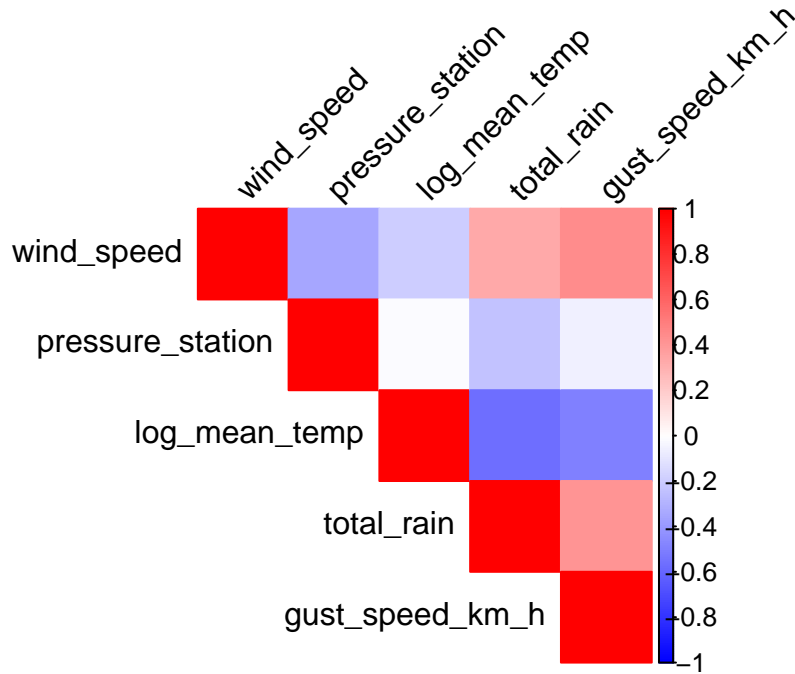
```
cor_matrix <- cor(selected_data_3)
print(cor_matrix)
```

```

          wind_speed pressure_station log_mean_temp total_rain
wind_speed      1.0000000      -0.34760462    -0.19023484   0.3333823
pressure_station -0.3476046       1.00000000    -0.01396792  -0.2341872
log_mean_temp   -0.1902348      -0.01396792     1.00000000  -0.5614056
total_rain       0.3333823      -0.23418715    -0.56140560   1.0000000
gust_speed_km_h  0.4414225      -0.05910226    -0.49986221   0.4189215

          gust_speed_km_h
wind_speed      0.44142245
pressure_station -0.05910226
log_mean_temp   -0.49986221
total_rain       0.41892153
gust_speed_km_h  1.00000000
```

```
corrplot(cor_matrix, method = "color", type = "upper",
         tl.col = "black", tl.srt = 45,
         col = colorRampPalette(c("blue", "white", "red"))(200))
```



2.1 Measurement

The measurement of Canadian weather data involves a network of weather stations and data collection methods managed by Environment and Climate Change Canada (ECCC). These stations continuously measure meteorological parameters such as temperature, precipitation, wind speed, and pressure (Meteorological Service of Canada 2023).

According to the glossary published by Meteorological Service of Canada (2023), Each day, measurement of temperature, rain, snow, precipitation, and gust speed are recorded. The wind and gust speed is measured in km/h with anemometer dials at a standard height of 10 meters above the ground. Rain and precipitation are measured in millimeter using the standard Canadian rain guage, a cylindrical container 40 cm high and 11.3 cm in diameter. Snow is measured in centimeters at several points that appear representative of the immediate area and then averaged. These raw data are combined to one entry and added to the historical climate database with a generated climate id and the station's location and id. Each row also have month and year of the data measured.

For climate research, including climate change studies, Environment and Climate Change Canada (2021a) has developed the Adjusted and Homogenized Canadian Climate Data (AHCCD) dataset. This dataset undergoes rigorous quality control and homogenization processes to address non-climatic factors that can affect long-term data consistency, such as station relocations or changes in instrumentation. The AHCCD ensures that observed trends reflect actual climate changes rather than artificial shifts in the data. In the AHCCD

Table 1: Column Headers of Raw Climate Data

Longitude (x)	Latitude (y)	Station Name
Climate ID	Date/Time	Year
Month	Mean Max Temp (°C)	Mean Max Temp Flag
Mean Min Temp (°C)	Mean Min Temp Flag	Mean Temp (°C)
Mean Temp Flag	Extr Max Temp (°C)	Extr Max Temp Flag
Extr Min Temp (°C)	Extr Min Temp Flag	Total Rain (mm)
Total Rain Flag	Total Snow (cm)	Total Snow Flag
Total Precip (mm)	Total Precip Flag	Snow Grnd Last Day (cm)
Snow Grnd Last Day Flag	Dir of Max Gust (10's deg)	Dir of Max Gust Flag
Spd of Max Gust (km/h)	Spd of Max Gust Flag	Longitude (x)

dataset, the precipitation, rain, pressure, snow and wind speed are adjusted with models to account for missing data and other non-climate factors. The detailed adjustments and corrections are documented in Section C. For example, precipitation measurements, which are often underestimated, are adjusted to ensure accuracy, especially in regions like the Arctic (Environment and Climate Change Canada 2021b). In the AHCCD dataset, parameters measured are recorded with the units, date, station ids, location and unique identifiers. The AHCCD data maintains a one-to-one correspondence with the historical weather dataset by a matching station id system, ensuring that each entry in the AHCCD aligns directly with a specific observation in the historical dataset.

2.2 Raw Data

In this project, we focus on weather data from YVR Airport, extracting only the datasets containing measurements taken at this specific location from the database. In both datasets, each row corresponds to a single averaged observation for a specific month and year. Each entry includes climate information such as temperature and wind speed, with their respective units recorded alongside the values. Additionally, a unique station ID and geographic coordinates (x, y) are included at the beginning of each entry for reference. The column headers of the raw historical weather dataset is displayed in Table 1. The column headers of the AHCCD dataset is displayed in Table 2.

The variables in the two datasets contains the following:

- Geographical Information: Longitude (x) and Latitude (y), with corresponding identifiers for location (Station Name in Table 1, station_id and province in Table 2).
- Temperature Metrics: Mean, maximum, and minimum temperatures (Mean Temp, Mean Max Temp, Mean Min Temp, Extr Max Temp, Extr Min Temp) and associated flags for data validity in Table 1. Similar metrics (temp_mean, temp_max, temp_min) in Table 2, with additional units included.

Table 2: Column Headers of Raw AHCCD Data

x	y
temp_mean_units__temp_moyenne_unites	temp_max_units__temp_max_unites
total_precip__precip_totale	temp_min__temp_min
rain__pluie	total_precip_units__precip_totale_unites
pressure_sea_level_units__pression_niveau_mer_unite	snow_units__neige_unites
temp_max__temp_max	lat__lat
identifier__identifiant	pressure_station__pression_station
lon__long	wind_speed__vitesse_vent
period_value__valeur_periode	period_group__groupe_periode
wind_speed_units__vitesse_vent_unites	temp_mean__temp_moyenne
pressure_station_units__pression_station_unites	province__province
station_id__id_station	temp_min_units__temp_min_unites
pressure_sea_level__pression_niveau_mer	snow__neige
date	rain_units__pluie_unites

- Precipitation and Snowfall: Total precipitation (Total Precip) and total snow (Total Snow), with flags for data quality in Table 1. Equivalent precipitation and snow variables (total_precip, snow) in Table 2, with units explicitly defined.
- Wind and Gust Metrics: Direction and speed of maximum gusts (Dir of Max Gust, Spd of Max Gust) in Table 1, with units and flags. Wind speed (wind_speed) and related metrics in Table 2, with units included.
- Pressure Information: Sea level and station pressure variables in Table 2 (pressure_sea_level, pressure_station) with units.
- Temporal Information: Date and time variables (Date/Time in Table 1, date, period_value in Table 2) to track observations across time periods.
- Flags and Identifiers: Flags for data validity in both tables, such as precipitation flags, temperature flags, and identifiers like Climate ID or identifier.

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 (?@fig-bills), from (palmerpenguins?).

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 [?@sec-model-details](#).

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.

$$y_i | \mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$

$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$

$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$

$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$

$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$

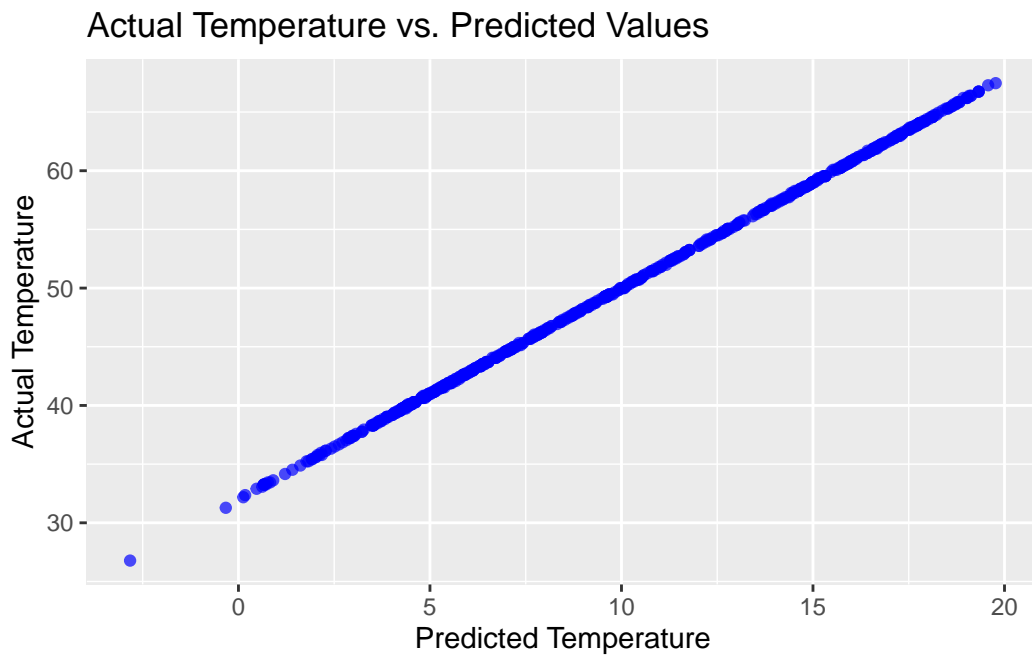
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

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

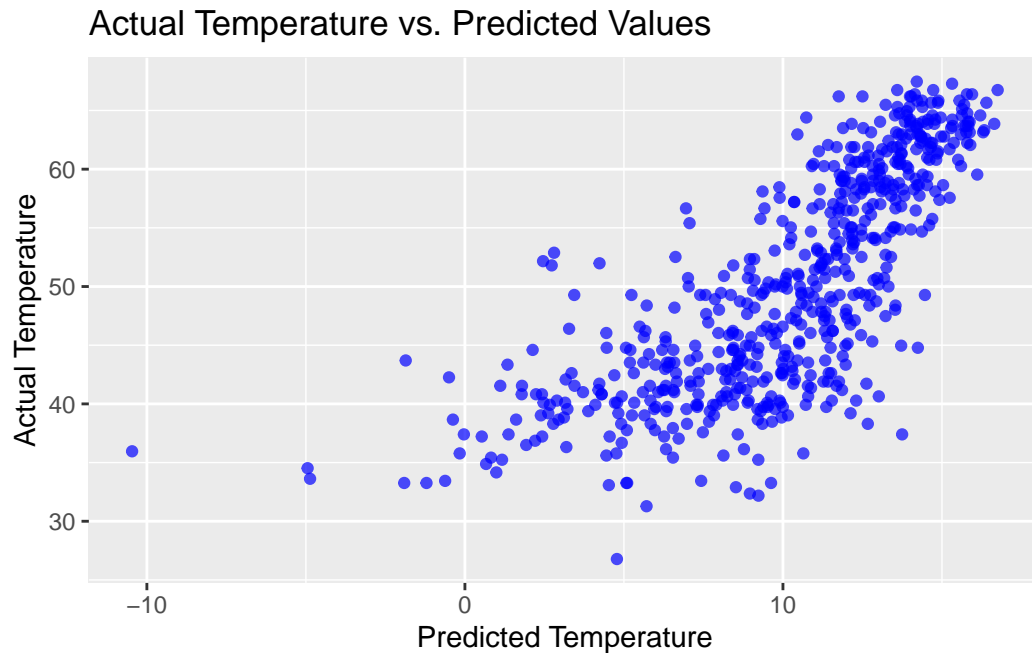
3.2 MLR Model


```
fitted_values <- fitted(m1)
ggplot(data = analysis_data, aes(x = fitted_values, y = mean_temp_F)) +
  geom_point(color = "blue", alpha = 0.7) +
  geom_abline(intercept = 0, slope = 1, color = "red", linetype = "dashed", size = 1) +
  labs(x = "Predicted Temperature", y = "Actual Temperature", title = "Actual Temperature vs
```

Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
 i Please use `linewidth` instead.



```
fitted_values <- fitted(m2)
ggplot(data = analysis_data, aes(x = fitted_values, y = mean_temp_F)) +
  geom_point(color = "blue", alpha = 0.7) +
  geom_abline(intercept = 0, slope = 1, color = "red", linetype = "dashed", size = 1) +
  labs(x = "Predicted Temperature", y = "Actual Temperature", title = "Actual Temperature vs
```



3.2.1 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 .

4 Results

Our results are summarized in `?@tbl-modelresults`.

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.

A Appendix

B License

Contains information licensed under the [Open Government Licence – Canada](#)

C Methodology of ECCC

The Adjusted and Homogenized Canadian Climate Data (AHCCD) is a collection of climate datasets developed by Environment and Climate Change Canada (2021a). These datasets provide long-term, quality-controlled data that have been adjusted to correct for non-climatic influences.

C.1 Population, Frame, and Sample

The population of interest in the AHCCD is the entirety of Canada’s climate data, representing diverse geographical regions and climate conditions. The frame of the dataset are the climatological stations maintained by the ECCC that span across the countries in important locations such as airports, and banks of lakes or rivers. These stations record data on climate elements such as temperature, precipitation, surface pressure, and wind speed over extended periods. The sample is the selected stations across Canada, with adjustments applied to address inconsistencies. The datasets cover periods extending back to 1895 for precipitation, while other variables like wind speed and surface pressure start from 1953 or later. The recorded sample consists of monthly, seasonal, and annual data about surface air temperature, precipitation, pressure, and wind speed, according to Environment and Climate Change Canada (2021a).

C.2 Sample Corrections and Adjustments

The original data for AHCCD are extracted from the National Climate Data Archive of Environment Canada. These data include daily observations, such as maximum and minimum temperatures, precipitation, surface pressure, and wind speed. Observations are quality-controlled and adjusted to correct for biases due to changes in instruments, observation procedures, and other factors.

Precipitation data adjustments account for wind undercatch, evaporation, and gauge-specific losses. According to Environment and Climate Change Canada (2021b), corrections to account for wind undercatch, evaporation, and gauge specific wetting losses were implemented, especially in snowy conditions where snowfall is not fully captured by standard gauges. Corrections are made with the study by Devine and Mekis.

Surface air temperature adjustments apply Quantile-Matching techniques to remove inhomogeneities. According to Environment and Climate Change Canada (2021c), With Vincent and Wang’s third generation homogenized temperature, Quantile-Matching ensures that the temperature data remain consistent across different periods, even when observation practices change.

Surface pressure and wind speed data undergo adjustments based on metadata and statistical tests for systematic shifts. According to Environment and Climate Change Canada (2021e), wind speed is first adjusted with a logarithmic wind profile, then tested for homogeneity using a technique based on regression models. It involves the identification of variation due to changes in anemometer and location change. The pressure data is corrected due to systematic shifts of non-updated station elevation and relocation, as stated by Environment and Climate Change Canada (2021d).

C.3 Sampling Approach and Trade-offs

According to the published methodology and the webpage by Dunbar (2020), they employ a systematic sampling approach by selecting specific climatological stations with long-term, consistent data records. In some cases, observations from neighboring or overlapping stations are merged to extend time series. The AHCCD dataset may also contain missing values, which can vary depending on the variable, station, and time. Additionally, the AHCCD dataset is site-specific, meaning it provides data specific to individual observation stations.

C.4 Missing Data Handling

Non-response, such as gaps in the data due to missing records, is managed by employing statistical and physical methods to homogenize the data. For instance, the AHCCD adjusts for shifts detected through historical evidence and metadata analysis. For large amount of missing data, ECCC mark the data as NA in the dataset (Canadian Centre for Climate Services 2022).

C.5 Strengths and Weaknesses

The AHCCD by Dunbar (2020) provides long-term, high-quality climate records adjusted for non-climatic factors such as changes in instrumentation, observation procedures, and station relocations, ensuring consistency and reliability for trend analysis in climate change.

The documentation acknowledges the possibility of missing values, which naturally arise in long-term observational datasets due to factors such as station interruptions, relocation, or equipment malfunctions (Environment and Climate Change Canada 2021a). Moreover, the dataset’s coverage in Arctic regions is limited to the restricted to the mid-1940s to present, as

this limitation reflects the historical absence of earlier systematic observations in these remote regions.

D 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...

References

- Canadian Centre for Climate Services. 2022. “Adjusted and Homogenized Canadian Climate Data.” <https://climate-change.canada.ca/climate-data/#/adjusted-station-data>.
- Dunbar, Alyssa. 2020. “Adjusted and Homogenized Canadian Climate Data (AHCCD) - Data Collection Methodology.” <https://open.canada.ca/data/en/dataset/9c4ebc00-3ea4-4fe0-8bf2-66cfe1cddd1d/resource/26545adf-e689-4d83-8f2d-9aad3dfa6f57>.
- Environment and Climate Change Canada. 2021a. “Adjusted and Homogenized Canadian Climate Data.” Datasets. <https://www.canada.ca/en/environment-climate-change/services/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data.html>.
- . 2021b. “Climate Data: Adjusted Precipitation Data.” Research;program results. <https://www.canada.ca/en/environment-climate-change/services/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data/precipitation.html>.
- . 2021c. “Climate Data: Homogenized Surface Air Temperature Data.” Program descriptions;program results. <https://www.canada.ca/en/environment-climate-change/services/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data/surface-air-temperature.html>.
- . 2021d. “Climate Data: Homogenized Surface Pressure Data.” Research. <https://www.canada.ca/en/environment-climate-change/services/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data/surface-pressure.html>.
- . 2021e. “Climate Data: Homogenized Surface Wind Speed Data.” Research;datasets. <https://www.canada.ca/en/environment-climate-change/services/climate-change/science-research-data/climate-trends-variability/adjusted-homogenized-canadian-data/surface-wind-speed.html>.
- Greenpeace East Asia. 2021. “5 Ways the Climate Crisis Will Change Asia.” *Greenpeace East Asia*. <https://www.greenpeace.org/eastasia/blog/6802/5-ways-the-climate-crisis-will-change-asia/>.
- Meteorological Service of Canada. 2023. “Past Weather and Climate.” <https://www.canada.ca/en/services/environment/weather.html>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Visser, Johan B., Conrad Wasko, Ashish Sharma, and Rory Nathan. 2021. “Eliminating the ‘Hook’ in Precipitation-Temperature Scaling.” *Journal of Climate*, September, 1–42. <https://doi.org/10.1175/JCLI-D-21-0292.1>.
- Wills, Robert C. J., Yue Dong, Cristian Proistosescu, Kyle C. Armour, and David S. Battisti. 2022. “Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change.” *Geophysical Research Letters* 49 (17): e2022GL100011. <https://doi.org/10.1029/2022GL100011>.
- Xu, Jianchu, R. Edward Grumbine, Arun Shrestha, Mats Eriksson, Xuefei Yang, Yun Wang,

- and Andreas Wilkes. 2009. “The Melting Himalayas: Cascading Effects of Climate Change on Water, Biodiversity, and Livelihoods.” *Conservation Biology* 23 (3): 520–30. <https://doi.org/10.1111/j.1523-1739.2009.01237.x>.
- Zhang, Peng, Junjie Zhang, and Minpeng Chen. 2017. “Economic Impacts of Climate Change on Agriculture: The Importance of Additional Climatic Variables Other Than Temperature and Precipitation.” *Journal of Environmental Economics and Management* 83 (May): 8–31. <https://doi.org/10.1016/j.jeem.2016.12.001>.