

MNXB11 Final Project Report

Akash Arjun Aleksei Jemeljanov Ossian Malmberg
Sofia Slaniceanu Vishvam Kaushikkumar Patel

October 2025

1 Introduction

The goal of this project was to analyze Swedish weather data provided by the Swedish Meteorological and Hydrological Institute (SMHI) and extract meaningful trends from it. We used a combination of bash, C++, and ROOT for data processing, visualization, and statistical analysis. The work was carried out collaboratively through GitHub, following a release-manager workflow where Ossian Malmberg acted as the release manager responsible for maintaining the main branch and merging contributions. The datasets we worked with included measurements from Falun, Falsterbo, Uppsala, and Lund, which allowed us to explore regional temperature and precipitation patterns across Sweden. Before analysis, all datasets were cleaned to remove faulty or irrelevant entries and formatted consistently for use in ROOT.

Our project included four main tasks:

- Warmest and coldest day of Uppsala
- Temperature difference between Falun and Falsterbo
- Temperature of a given day in Falsterbo
- Rain analysis comparing Uppsala and Lund on the years 1961 and 2024

2 Code

2.1 Data Cleaning

Each dataset was first processed using separate cleaning scripts:

- `cleaning_data_Uppsala.cxx`
- `cleaning_data_Falun.cxx`
- `cleaning_data_Falsterbo.cxx`

- `rain_analysis/data_clean/Rain_data_clean.cxx`

These scripts read the raw SMHI `.csv` files, skipped header and metadata lines, handled missing or faulty data entries, and ensured consistent column formatting for date and temperature/rainfall. The cleaned output files were then stored in a dedicated directory to preserve reproducibility, following the recommendations in the MNXB11 project guidelines.

2.2 Task 1: Warmest and Coldest Day of Uppsala

This task was implemented in `warmest_coldest.cxx`. The program reads daily temperature data for Uppsala, identifies maximum and minimum temperatures each year. The program `warmest_plot.cxx` is used to plot the distribution of the warmest and coldest days in data set by using ROOT. A Gaussian fit was applied to estimate the most probable occurrence dates of the warmest and coldest days.

2.3 Task 2: Temperature Difference between Falun and Falsterbo

The comparison between Falun and Falsterbo was implemented through two main files: `FalunVSFalsterbo.cxx` and the macro `FalunVSFalsterboPlot.C`. After cleaning, both datasets were combined to calculate the mean temperature and their difference for the corresponding years. This was done by adding up the temperature readings for each day and dividing them by the number of days counted in the year. The file created by `FalunVSFalsterbo.cxx`, `FalunVSFalsterbo.csv`, contains information about the year, the temperature in Falun, the temperature in Falsterbo, and also the difference between the two each year. ROOT's `TH1F` and `TGraph` classes were used to visualize the temperature in these two cities and their difference over time as 3 distinct lines.

2.4 Task 3: Temperature of a Given Day (Falsterbo)

In `temperature_given_day.cxx` and `temperature_given_day.C`, we analyzed the distribution of temperature for a specific day across all available years. This is visualized through a ROOT histogram which shows the spread of temperatures for that chosen date, alongside mean and standard deviation calculations. This gives insight into the expected temperature range for that day in Falsterbo and how stable or variable the weather is year-to-year.

The data for the temperature on a given day was obtained from the provided Falsterbo datafile through the cleaning executed via the `cleaning_data_Falsterbo` executable. From this cleaned data, `temperature_given_day.cxx` performed largely two functions. Namely, it would take two temperature readings from the datafile for a given day and month, and calculate the average temperature of that day. This average would then be saved in an output file, so that later

analysis may be performed. This was done for a given day and month over all the years in the datafile. The day and month is chosen by the user, through a terminal input which is requested upon running the executable.

An immediate simplification is noted with regards to how the average is calculated. By only taking into consideration two readings, at 6 AM and 6 PM, one introduces a degree of uncertainty in the average, especially if, for instance, the datafile was to be from a specific geographical region, like a desert, where morning and evening temperatures could drop relatively sharply and not necessarily reflect the daily average. This simplification was chosen as the first few yearly entries in the Falsterbo datafile contain only the aforementioned two readings, however it was later observed that more recent years contain more entries.

2.5 Task 4: Rain Analysis (Uppsala vs. Lund, 1961 and 2024)

The rain analysis was performed using the scripts in the `rain_analysis/` folder. The scripts were used for data cleaning, analysis, and generating plots, all automated via Bash scripts for reproducibility. The process began with `Rain_data_clean.cxx`, which parsed raw SMHI datasets containing daily precipitation and temperature readings and produced clean CSV files. These were then processed by `analysis.cxx`, which filtered the data to the years 1961 and 2024, computing monthly precipitation totals and average temperatures for each city and saving the results in structured CSV files.

ROOT was then used to generate graphs to show long-term rainfall trends and inter-annual temperature variation. This analysis provided a broader understanding of changes in precipitation and temperature over time between central and southern Sweden.

3 Results and Analysis

3.1 Warmest and Coldest Day of Uppsala

The graph (Figure 1) shows the distribution of the warmest and coldest days of the year in Uppsala. The red bars represent the warmest days, which are concentrated around mid-year (around day 200), corresponding to summer. The blue bars indicate the coldest days, occurring mainly at the beginning and end of the year (around days 20 and 350), corresponding to winter. The Gaussian curves show the typical timing of temperature extremes, with peaks clearly separated between summer and winter periods. This pattern reflects the strong seasonal climate variation typical of northern Europe.

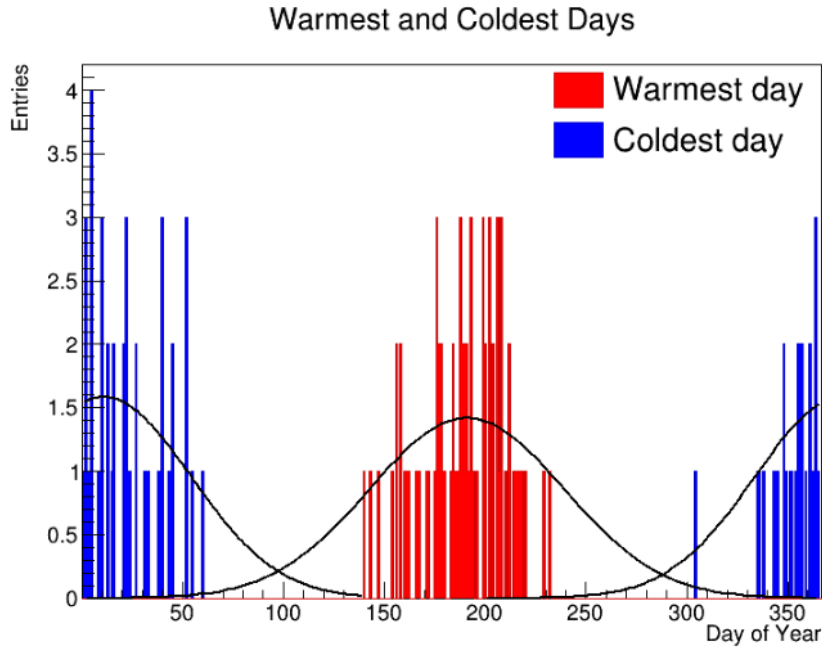


Figure 1: Warmest and coldest days of the year in Uppsala. The Gaussian fits indicate the typical timing of temperature extremes.

3.2 Temperature Difference between Falun and Falsterbo

The resulting plot (Figure 2) clearly demonstrates how the inland vs. coastal locations in Sweden differ—Falun generally exhibits more extreme temperature variation than the coastal Falsterbo. The difference in temperature does not appear to have changed much over time, although it should be noted that this only shows the last 20 years.

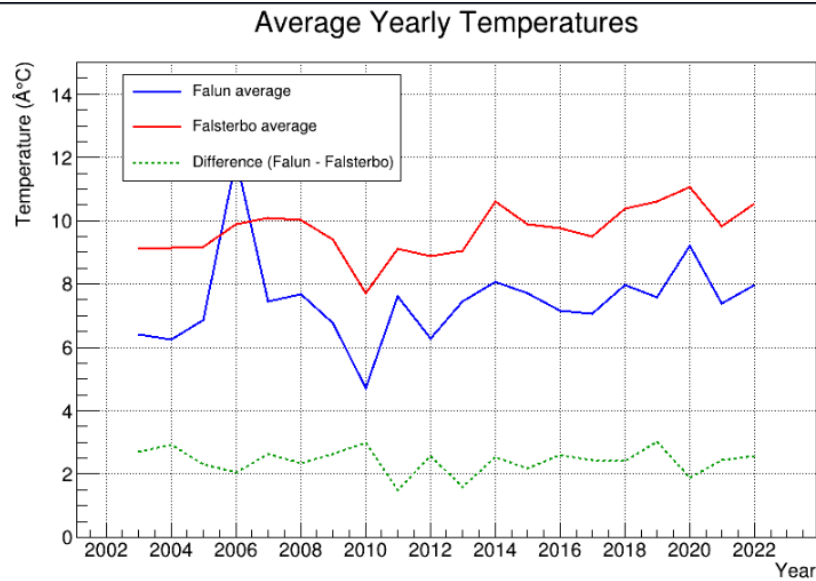


Figure 2: Temperature difference between Falun and Falsterbo. Falun exhibits larger temperature variations due to its inland climate.

3.3 Temperature on a Given Day

The temperature of a given day can be recorded over multiple years and plotted on a histogram as in figure 3. Qualitative analysis can readily be made by inspecting the histogram. It is evident, for instance, that the temperature in Falsterbo on the 5th of January tends to be above freezing, which we would expect given that the town is situated on the Southern coast of Sweden. We additionally observe that the temperatures maintain within a certain range for the given timeframe, which is again a phenomenon to be expected.

There are, however, possible improvements to be made with regards to the code and analysis. For instance, the average temperature of a given day is calculated by taking the average between two temperature readings, one at 6 AM, the other at 6 PM. For earlier years, those are the only two readings available, however the data grows more abundant for more recent years, and by expanding the average one would obtain more accurate averages and thus more accurate histograms.

In addition, it would be useful for investigation to plot the averages vs. the years. This would make it possible to see if the average temperature of a day has gone up during the years, which would be a reasonable hypothesis given the pattern of global warming. For that, however, perhaps more cleaned data is needed, as it would be useful to have access to more than 40 entries, since

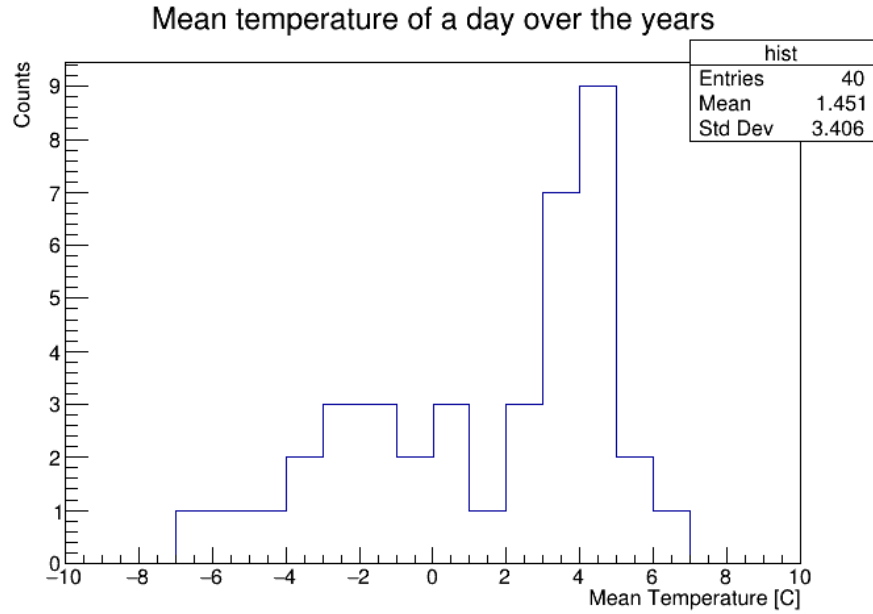


Figure 3: Histogram of the average temperature in Falsterbo for January 5th, for the years 1984-2022.

climate patterns are most noticeable over large periods of time.

3.4 Comparison of Monthly Rainfall and Temperature for Lund

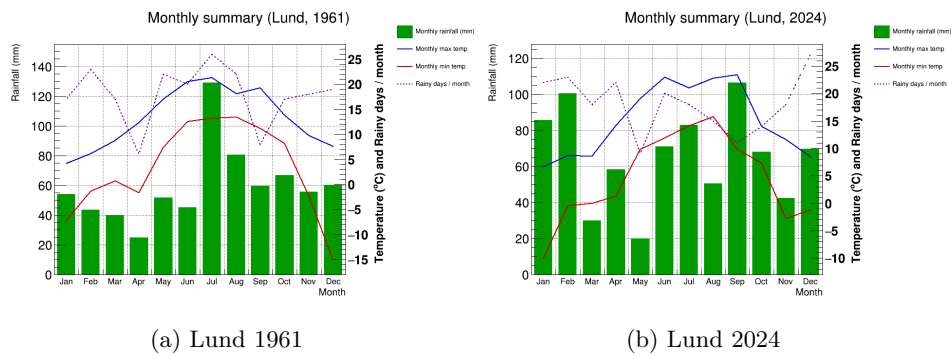
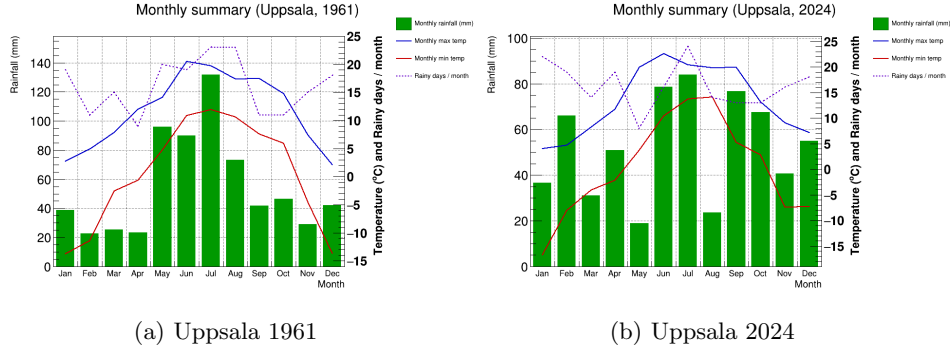


Figure 4: Monthly rainfall distribution for Lund in 1961 and 2024. The seasonal pattern is consistent, though 2024 shows slightly higher precipitation.

As seen in Figure 4, Lund exhibits a clear seasonal rainfall pattern for both 1961 and 2024, with precipitation peaking during the summer months and lower values in winter. The overall pattern remains consistent across the two years, but the 2024 data indicate a slight increase in rainfall intensity, particularly from June to August. Additionally, the temperature curves show that warmer summer months tend to coincide with higher rainfall, suggesting more frequent or intense summer precipitation events in recent years.

3.5 Comparison of Monthly Rainfall and Temperature for Uppsala



A similar trend is observed for Uppsala, where the seasonal cycle of precipitation and temperature is maintained between 1961 and 2024. However, Uppsala shows a more pronounced increase in summer rainfall and a wider temperature range across the year compared to Lund. When comparing both cities, Lund generally experiences slightly higher average rainfall, while Uppsala shows sharper temperature contrasts between summer and winter.