



**LUND**  
UNIVERSITY

LUND'S UNIVERSITY  
THE FACULTY OF SCIENCE, DEPARTMENT OF MATHEMATICS

# MNXB01 Project Report, Group F

*Paulina Benthem Ciano, Dustin Lindner Daii,  
Edmund Lehsten, Anton Palets*

COURSE: MNXB01

November 4, 2021

## Contents

<b>1</b>	<b>Intro</b>	<b>2</b>
<b>2</b>	<b>Temperature of a given day</b>	<b>2</b>
<b>3</b>	<b>The temperature for every day of the year</b>	<b>2</b>
<b>4</b>	<b>Covariance between cities</b>	<b>3</b>

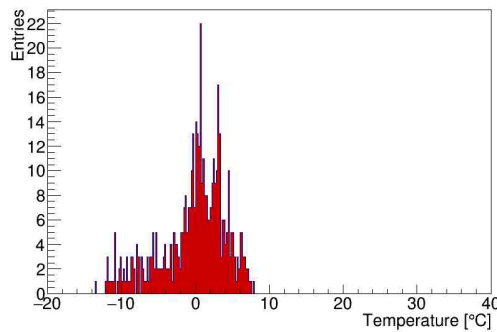
# 1 Intro

In this project we were given data from SMHI to write a program which extracts climate data from Sweden. After cleaning the data we make use of ROOT to create plots and perform some statistical analysis. There are three primary results which we study. The first is the temperature of a given day, we then expand upon this to see the temperature for every day of the year rather than one day. Finally, we investigate the covariance between the weather in Lund and Falsterbo as well as the covariance in Lund and Umeå and see what conclusions can be drawn.

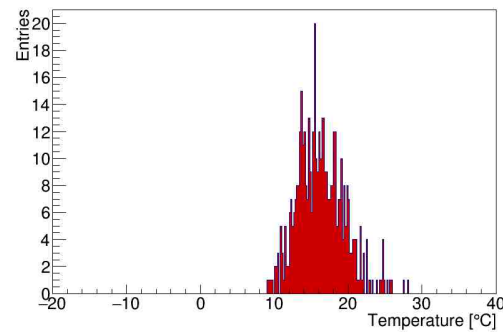
## 2 Temperature of a given day

Following the first example in the project instructions, the goal of this exercise is to calculate the histogram of the temperature of a given day, the mean temperature of that given day and the standard deviation of the temperature. The function `tempOnDay` can take as input either the date by month and day, or the day of the year. The histogram gets filled with the temperature of the given day over all years with available data.

In Figure 1, the resulting histogram for input date (8/23) is presented. The probability of observing a particular temperature on a given day can be calculated by the number of entries in the histogram for the particular temperature divided by the total number of entries.



(a) Falsterbo



(b) Göteborg

## 3 The temperature for every day of the year

To see the average trend of temperature throughout the year, we can average out the temperature for each date and add it to a histogram. The way this is implemented is by using a while loop that goes line by line of cleaned up .csv file. Within the while loop we check whether the current line has the same date as the previous, and if so we add the temperature to the running total, if not we find the mean of the previous running total and add it to the correct bin of the histogram. Due to this implementation, the last group of entries will not be added to the histogram, as the while loop would have to run again, but this is easy to fix because the necessary information is still stored in the variables. We address this by adding the last group of entries to the histogram outside the loop. Since all the mean temperatures are added to the bins, we need to then normalize the bin

values by the number of weighted entries. We create a normalization array and in it count how many times a day was added to the histogram. See the plots in the Figures section. One can see that the climate is milder in cities that are not on or close to the coast. For example, it might be that Kiruna has the biggest spread of temperatures because the presence of mountains generally makes the weather harsher.

City	Max	Min
Lund	14.3896	1.35988
Falsterbo	15.2263	1.98313
Kiruna	13.4096	-14.0109
Göteborg	14.6356	1.43963
Gävle	13.8721	-2.83569
Jönköping	12.6589	-0.163235

Table 1: Maximum and minimum mean temperatures in different cities

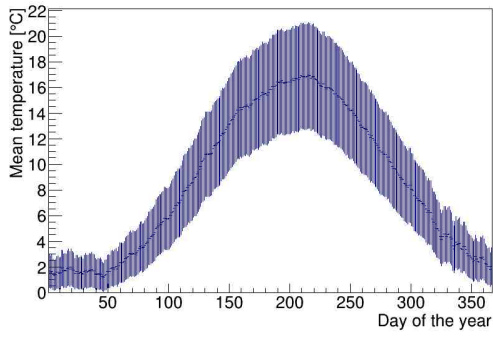
## 4 Covariance between cities

For our third result we wished to study the covariance of the temperature between different cities with possibly a time lag. We made use of three cities, Lund, Falsterbo and Umeå. The covariance  $Cov(X, Y)$  of two real-valued random variables  $X, Y$  is formally defined as  $Cov(X, Y) = E((X - E(X))(Y - E(Y)))$ , where  $E(X)$  is the expectation of  $X$ . Intuitively, or as an interpretation, the covariance tells us how  $X$  and  $Y$  relate to each other, that is, it is the degree by which the random variables change with respect to each other.

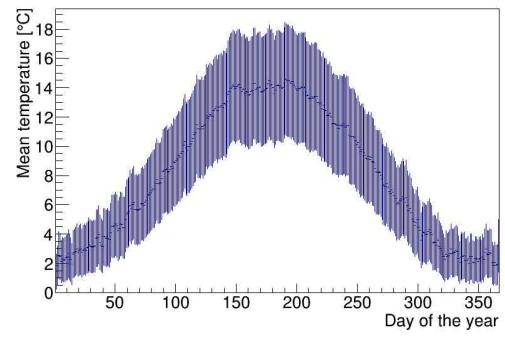
The way we did this was by implementing an extra function into our `tempTrender` object. This function takes as inputs a data path for another measuring station and then plots the covariance between the two.

We show some figures of the covariance between Lund and Falsterbo, Further we show the covarriance between Lund at different time lags.

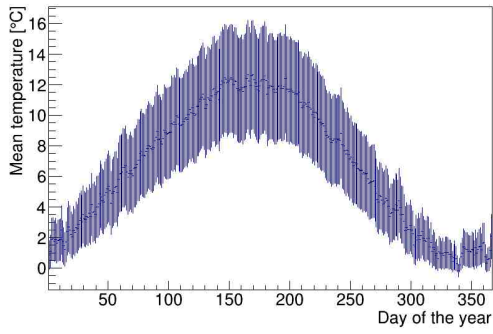
Due to the close proximity of Lund and Falsterbo we expected the covariance for Lund and Falsterbo to be fairly good. Further we should see a better correlation in the graphs with a time lag with 1 day than 1 week, while the 6 months time lag should have very week negative correlation which we see.



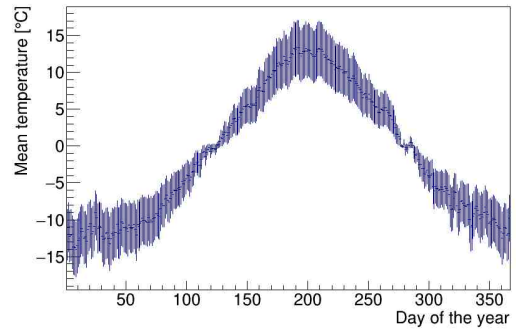
(a) Falsterbo



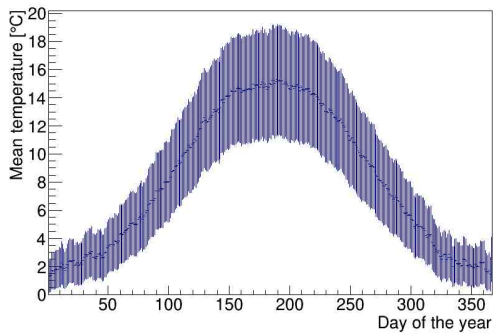
(b) Göteborg



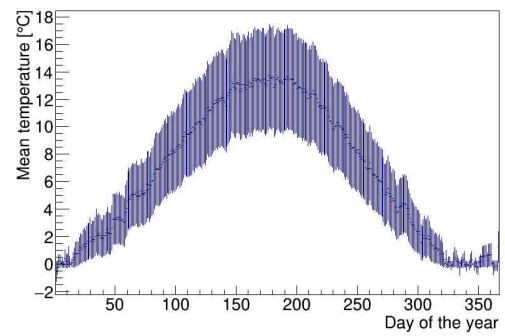
(c) Jönköping



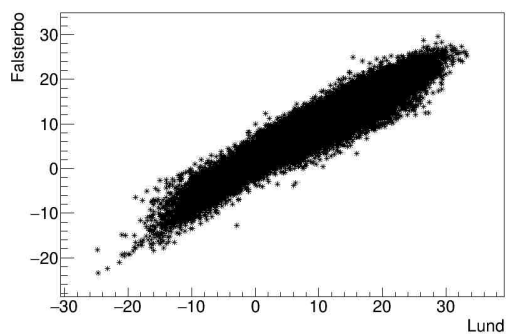
(d) Kiruna



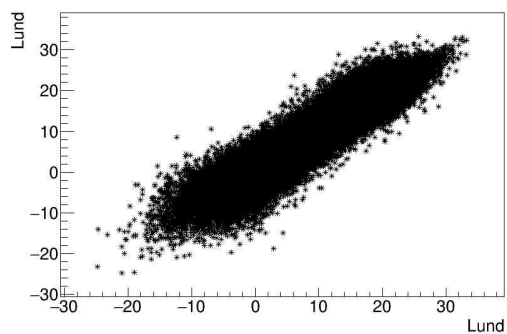
(e) Lund



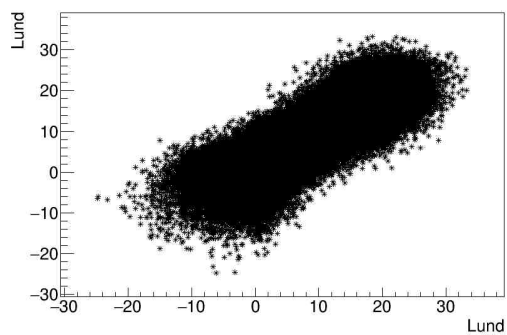
(f) Stockholm



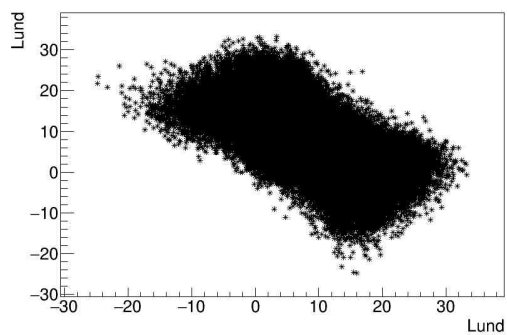
(a) Covariance between Lund and Falsterbo



(b) Lund with time lag 1 Day



(c) Lund with time lag 1 Week



(d) Lund with time lag 6 Months