REPORT ABOUT THE ASSIGNMENT

MATHIEU GUIGUE

1. Introduction

2. Part 1: Data averaging

2.1. **Time series.** Each file corresponds to a day of data, one data point every 60 seconds. Each day has its own time stamp (number of seconds since midnight) and a hardcoded-modification of the reader function is required to get things working.

2.2. **Estimating the optimal average period.** Using an Allan representation of the data, I can estimate what the typical period where the measurements fluctuations are Gaussian is. It seems that the value at which non Gaussian fluctuations appear is defined for all three variables to be about 60 samples. For safety, I will use 50 samples for averaging.

To-Go-Further 1. *Implement an algorithmic way to extract the right number of measurements to average (something like a minimum finder).*

The averaged data (represented on Figure **??**) are then saved in a CDF file.

3. Part 2: Clustering of data

Date: March 20, 2018.

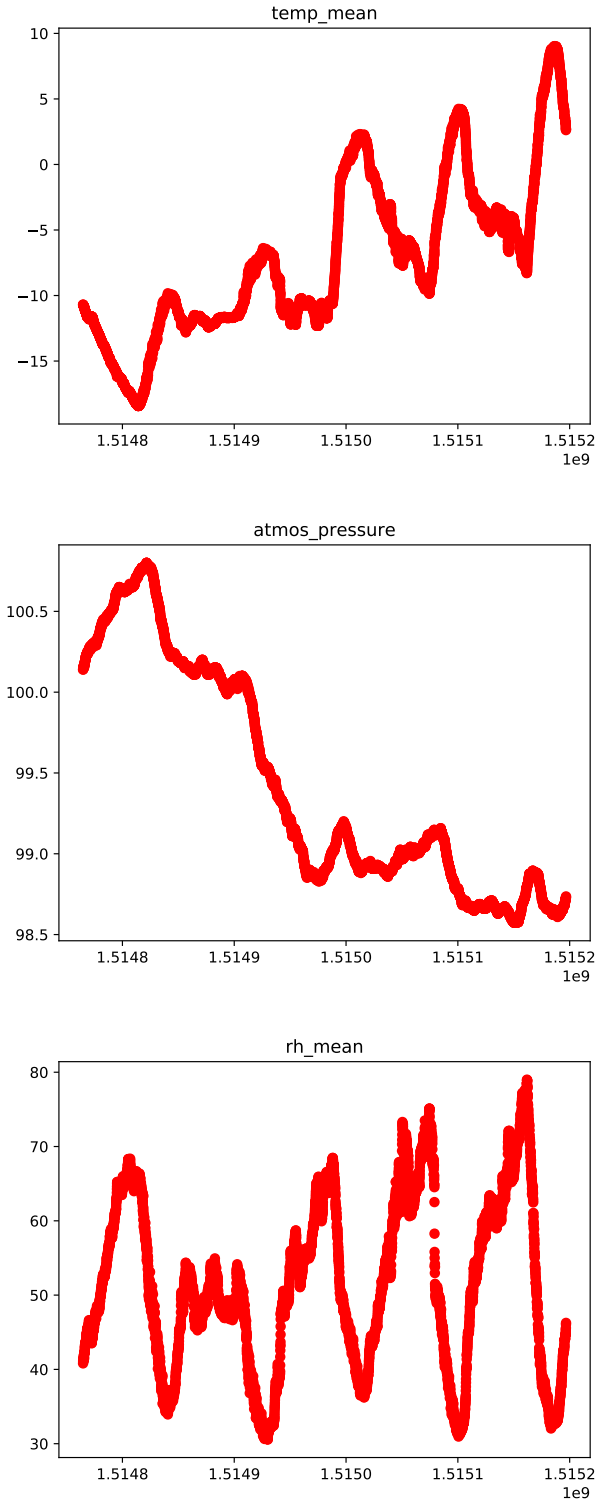

FIGURE 1. Time series of the mean temperature, mean relative humidity and atmospheric pressure as a function of time.

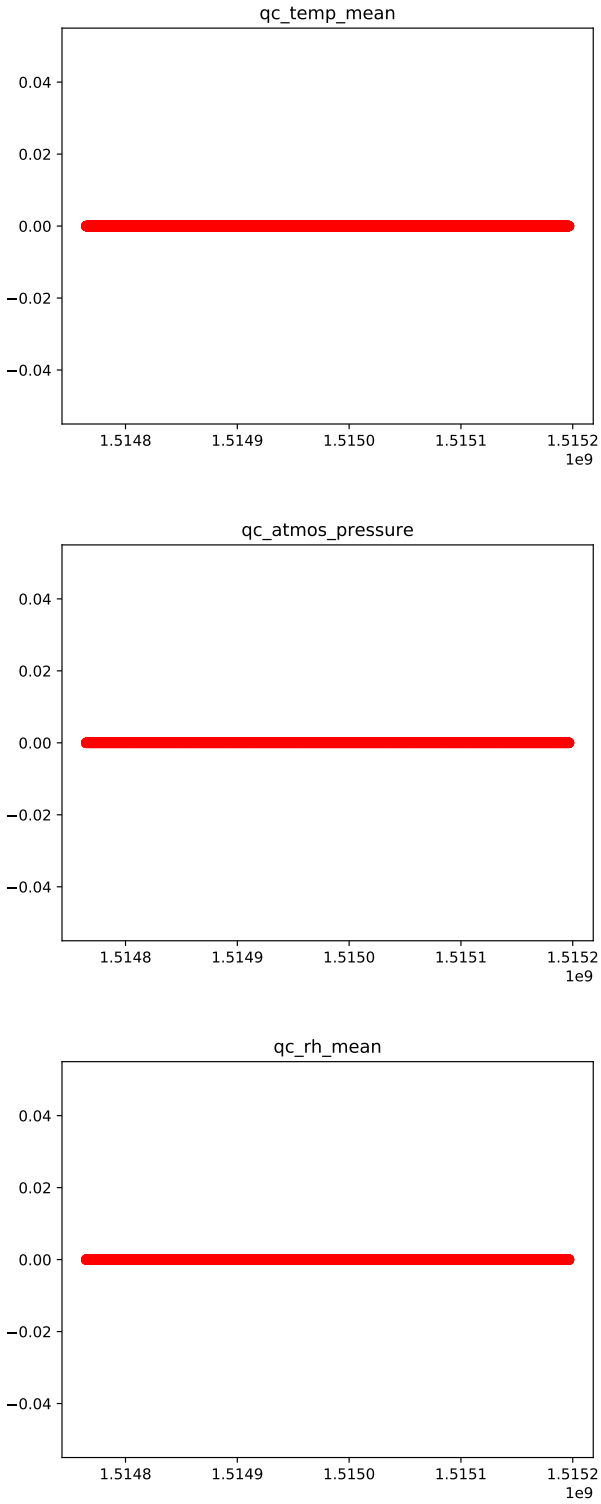

FIGURE 2. Time series of the quality checks for mean temperature, mean relative humidity and atmospheric pressure as a function of time.

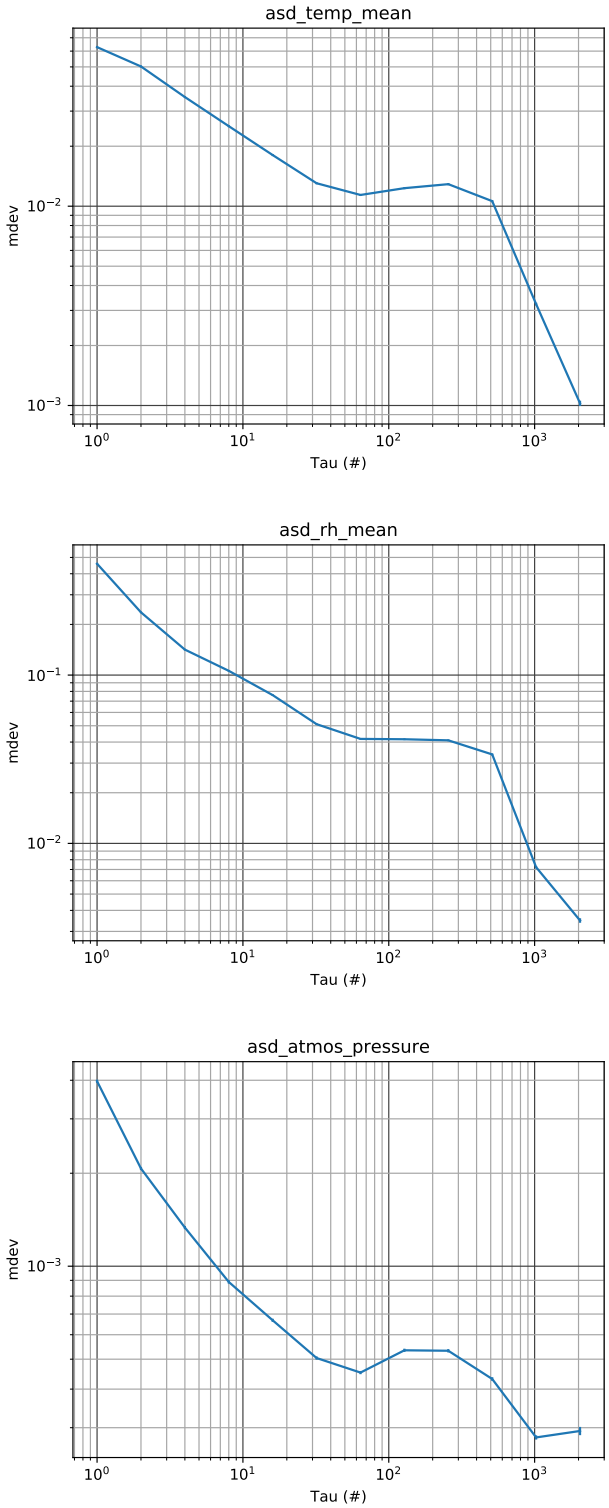

FIGURE 3. Allan Standard Deviation of the mean temperature, mean relative humidity and atmospheric pressure as a function of the period (in number of samples).

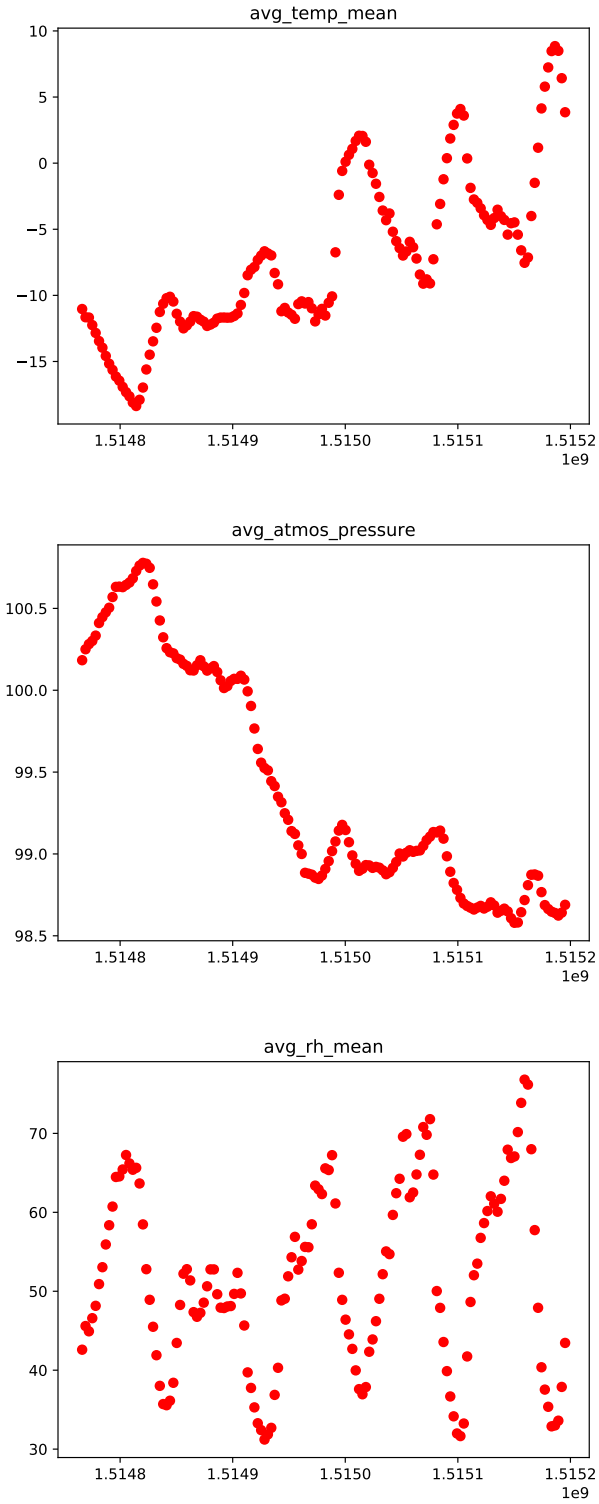

FIGURE 4. Time series of the average of mean temperature, mean relative humidity and atmospheric pressure as a function of time.