

Study of the concentration of Particulate Matter in Air

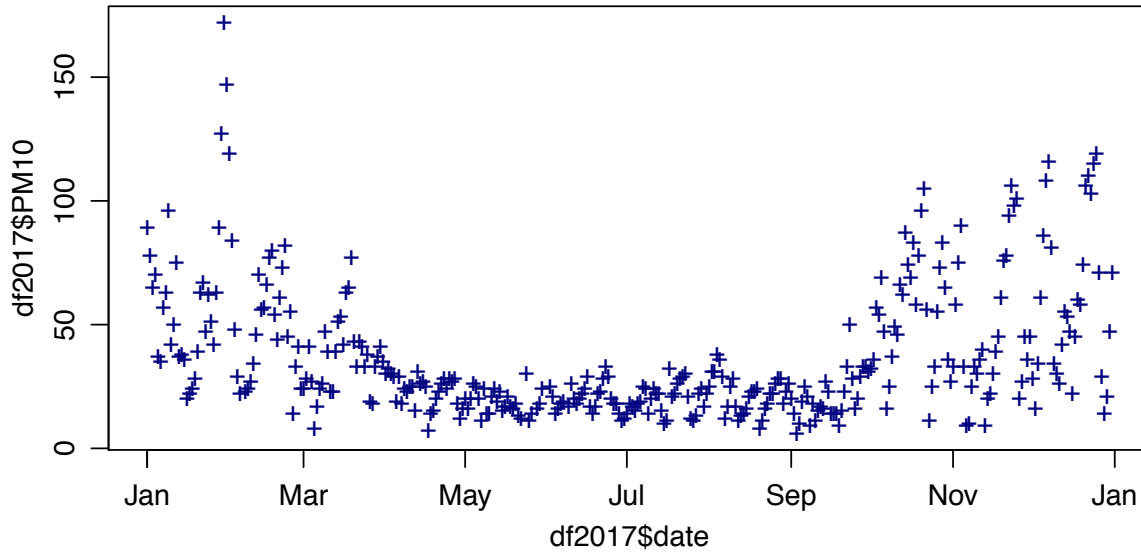
The air quality and the concentration of pollutants¹ is measured on a daily base by several monitoring stations located around the territory. Using the data collected by one monitoring station

<http://www.arpa.veneto.it/temi-ambientali/aria/stazioni-di-monitoraggio/500000079>,

it is required to study the concentration of Particulate Matter in air, called PM10, over the last fifteen years (2004-2018).

Using data which are available as EXCEL files, import them in R and perform a Gaussian Process Regression [1]-[2] on the data.

Using a Radial Basis Function (RBF) kernel (or covariance function), fit the data for different choices of the measurement error and the hyper-parameter that controls the size of the RBF. Plot both the mean function, $\mu_*(t)$, and point-wise variance, $\sigma_*(t)$, for these different choices and discuss the resulting curves.



Bibliography

- [1] Jeff Bayers, *Physics of Data Padua Lectures 8*, DFA Moodle web site
- [2] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, 2006. <http://www.gaussianprocess.org/gpml/>.
The book is available for download in electronic format in <http://www.gaussianprocess.org/gpml/chapters/RW.pdf>

¹According to European and Italian laws, the following pollutants are measured: Particulate Matters (PM₁₀ and PM_{2.5}), Carbon Monoxide (CO), Sulfure Dioxide (SO₂), Ozone (O₃) and Nitrogen Oxides (NO_x)