EWMA Control Charts for Monitoring Pollution

Irving Gómez irving.gomez@cimat.mx

Centro de Investigación en Matemáticas (CIMAT)

Summary

In this work, we develop an EWMA control chart with a reflecting boundary for the Weibull and lognormal distribution in presence of censored data by interval. We also develop a new efficient methodology to estimate the control limits. These control charts can be used to get more flexible rules when monitoring pollution.

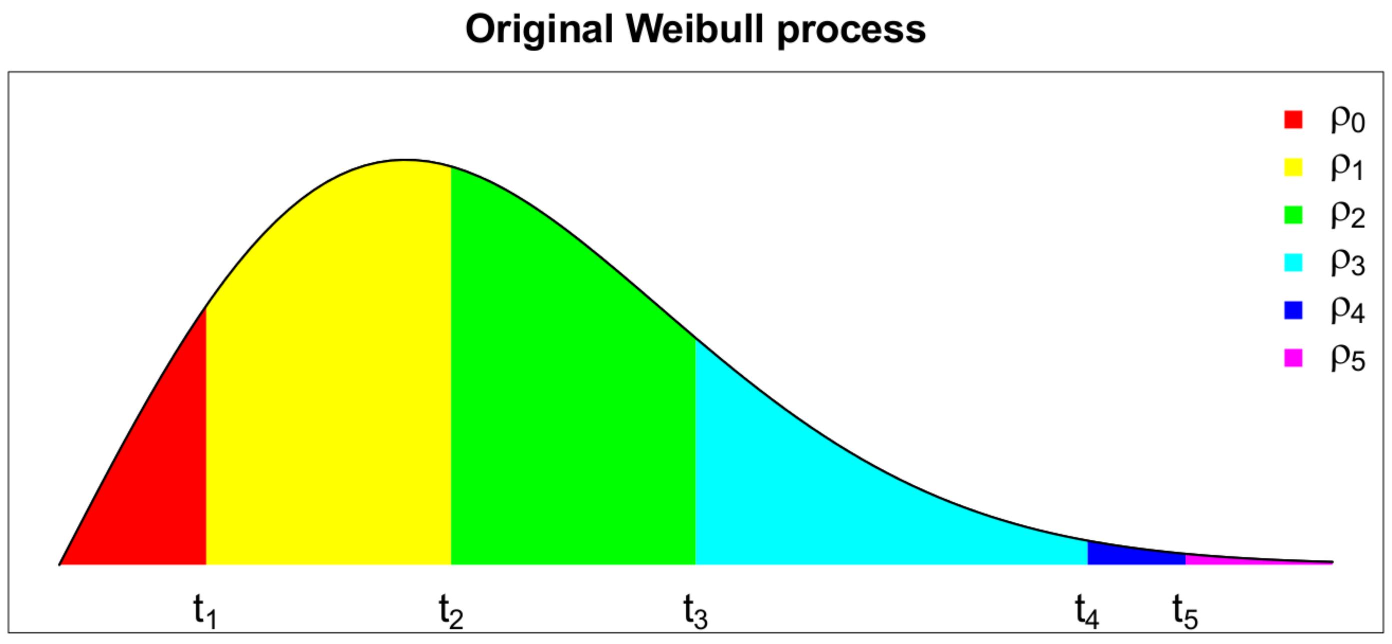
Introduction

The monitoring of pollution in big cities has become a serious problem. When certain pollutants exceed a limit, an atmospheric contingency is declared. However, when the pollutants have with a sustained high level, without exceeding the limit, usually the data is ignored.

With the use of control charts, we believe that more effective, flexible rules can be derived, when the pollutants are both at low levels or at a sustained high level. Even if part of the data is censored by intervals.

Censored data

The data that does not exceed the pollutant levels to declare an atmospheric contingency is modeled by a Weibull or lognormal distribution, with possible data censored by interval.



We transform the data to exponential or normal distribution and use the Conditional Expected Value (CEV) for the censored data. We use the resultant data to develop EWMA charts.

EWMA charts

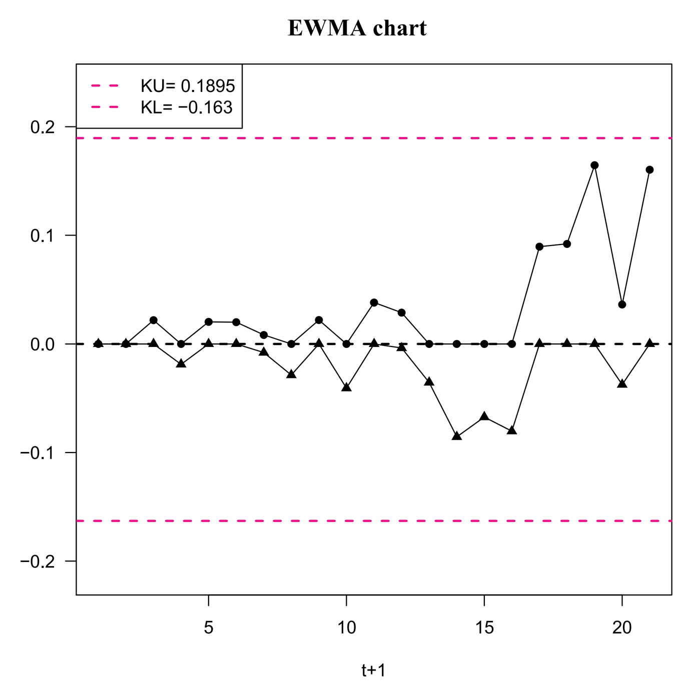
The EWMA charts are robust to non-independent data, which is the norm in this case. They have a long memory of the whole process.

One consequence of these properties is that they can detect changes even if no data has exceeded the limit, keeping under control the probability of a false positive.

Proposed EWMA chart

We propose an EWMA chart with a reflecting boundary, which minimizes the inertia effect.

$$Q_t^U = \max\{(1-\lambda)Q_{t-1}^U + \lambda\bar{w}_t,\, \omega_0\}$$
$$Q_t^L = \min\{(1-\lambda)Q_{t-1}^L + \lambda\bar{w}_t,\, \omega_0\}$$



Control Limits

The control limits K_L and K_U are selected to give a false alarm with a certain probability. When the distribution of the control statistic is unknown, the control limits must be estimated.

The traditional procedure to estimate control limits is through simulation and the bisection method, which has proven to be very ineffective.

A new methodology to estimate control limits

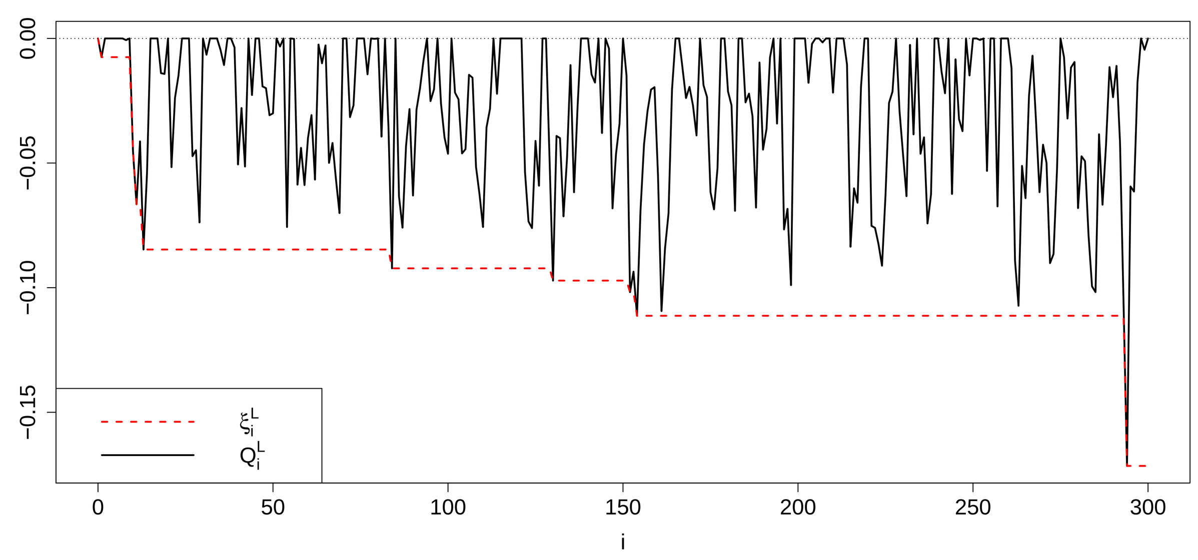
Consider the statistics

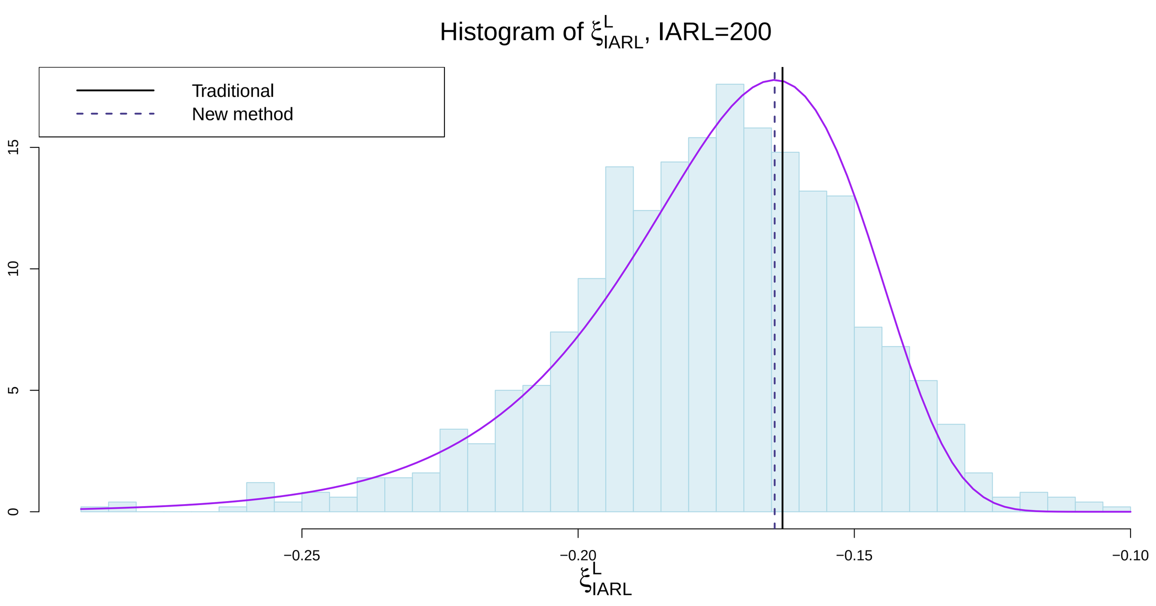
$$\xi_T^L = \min\{Q_t^L \,|\, t \le T\}$$
$$\xi_T^U = \max\{Q_t^U \,|\, t \le T\}$$

We have discovered that the lower control limit (K_L) and the upper control limit (K_U) satisfy

$$K_L = Q_{0.6321}^{\xi_{IARL}^L} \qquad K_U = Q_{0.3679}^{\xi_{IARL}^U}$$

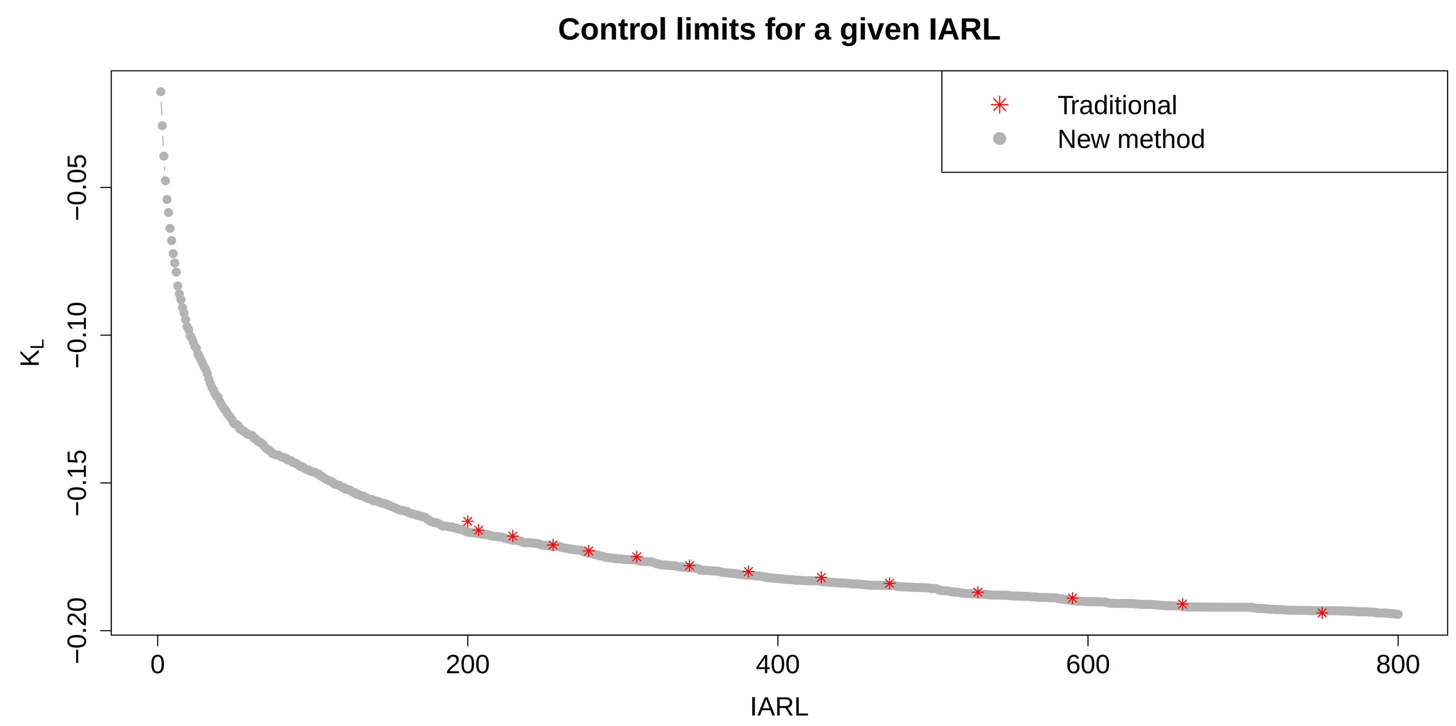
for a desired In-control Average Run Length (IARL).





This methodology has been proven to be faster than the traditional procedure and very accurate, using the asymptotic extreme value theory.

With this methodology, it is also possible to identify the complete relation between control limits and their IARL.



Conclusions

- The interval censored data is achieved for the first time.
- The developed control charts have properties which detect changes not only when a pollutant level has been exceeded, but also in other cases.
- The memory property of the charts makes them a valuable option when dealing with non-independent samples.
- The methodology we propose to estimate the control limits has proven to be very effective, and faster than the traditional procedure. We also think this methodology could be extended to other cases.

Synthesis of Statistics, Data Mining and Environmental Sciences in Pursuit of Knowledge Discovery from October 29 to November 03, 2017, Casa Matemática Oaxaca, Mexico

LaTeX TikZposter