**Control Chart with Time Series Methods**

Given the time series data which is auto correlated, we would be utilizing Time Series Forecasting techniques in addition to Control Chart to detect anomaly in individual sensors. [1]

Control chart (also known as Shewhart charts) is a statistical process control tool which can be used to assess if an industrial process is in a state of control. For our application, we are using the ggQC and ggplot2 package to derive the Individual-X /Moving Range pair chart (also known as XmR chart).

XmR consist of a pair of charts that enable us to monitor a process for shifts in the process that changes the mean or variance of the measured statistic. The Individual-X chart is used to analyses central location while the Moving Range chart shows the difference between consecutive readings. Together, the XmR chart is used to study system variability. [5]

* Individual-X chart displays the individual measured values (mean ± 3sigma); and
* Moving Range chart displays the difference from one point to the next

The control limits (or natural process limits) are defined by ± 3 sigma of the measured statistic. These limits indicate the levels by which the process will fall within if there are no significant changes to the process. When any given point exceeds the 3 sigma control limits, it signals that some assignable cause may have resulted a change in the process. Likewise, continuous run of points on one side of the mean line should also be interpreted as sign of change in the process. These signals support the occurrence of anomalies for a component where ratification action could be taken prevent further deterioration to the component system, within and adjacent.

Individual-X chart: red line (3 sigma limits) vs blue line (3 standard deviation limits)

* Sigma is sequence sensitive and deemphasizes systematic variation (less affected by systematic variation) which allowing us to measure the inherent random variation more clearly.[3]
* Standard Deviation (S.D) measures total variation (systematic & random variation).

**Formulas** [4]:

mR (Moving Range) = Absolute difference between points

SigmaX (Sequential Standard DeviationX) = MeanmR/1.128

Upper Control Limit for XmR (UCLX) = MeanX + 3\*Sigma

Lower Control Limit for XmR (LCLX) = MeanX - 3\*Sigma

Upper Control Limit for mR (UCLmR) = 3.267\*MeanmR

Lower Control Limit for mR (LCLmR) = 0

Time series analysis is able to breakdown the time series data in its seasonal, trend, and remainder components which allows us to analysis the components individually. [2] In this case, we wish to detect anomalous behavior within the remainder component. For our application, we utilize the Anomalize package. While Anomalize is configurable to analyse via two time series decomposition methods and anomaly assessment, our application is set to use Seasonal Decomposition of Time Series by Loess (STL) and Inter Quartile Range (IQR) for anomaly assessment.

[1] Use of Shewhart individuals control chart; <https://en.wikipedia.org/wiki/Control_chart>

[2] Anomalize package guide:

<https://cran.r-project.org/web/packages/anomalize/vignettes/anomalize_quick_start_guide.html>

[3] Formula for Sigma (Sequential\_Standard\_Deviation):

<https://r-bar.net/xmr-control-chart-tutorial-examples/>

[4] Formula for mR control limits: <https://qualityamerica.com/LSS-Knowledge-Center/statisticalprocesscontrol/moving_range_chart_calculations.php>

[5] Interpreting XmR chart: <https://qualityamerica.com/LSS-Knowledge-Center/statisticalprocesscontrol/interpreting_an_individual-x___mr_chart.php>

[6] Sigma violation rules: <http://rcontrolcharts.com/stat_qc_violations_vignette.html>

**PURPOSE OF CHART**

* XmR chart: Plot individual sensor readings against time with use control limits to detect systematic and random variation
* Sigma Violation chart: Highlight points where sigma threshold are breached
* Anomaly chart: detect anomaly after trend and seasonality factors are removed
* Compare between different Periods and/or Sensors

**How to Interpret XmR and Sigma Violation Charts**

* If the sigma line is close to standard deviation line, it signals the lack of systematic variation in the process.
* If the sigma line deviates from the standard deviation line, it signals significant systematic variation in the process
* Look at mR chart first, if mR chart is out of control, then control limits on Individual-X chart are meaningless.
* After reviewing the mR chart, look for point that exceed the 3 sigma limits
* In addition to looking for points exceeding the 3 sigma limit, we can apply additional rules via Sigma Violation Chart to provide an earlier warning that process are out of control [6]

1. Violation same side: 8 or more consecutive points on the same side
2. Violation same side: 4 or more consecutive points on the same side exceeding 1 sigma
3. Violation same side: 2 or more consecutive points on the same side exceeding 2 sigma
4. Violation same side: any points exceeding 3 sigma

For points that are deemed out of control, consider the context, possible reasons for failure, and if it warrants further actions.

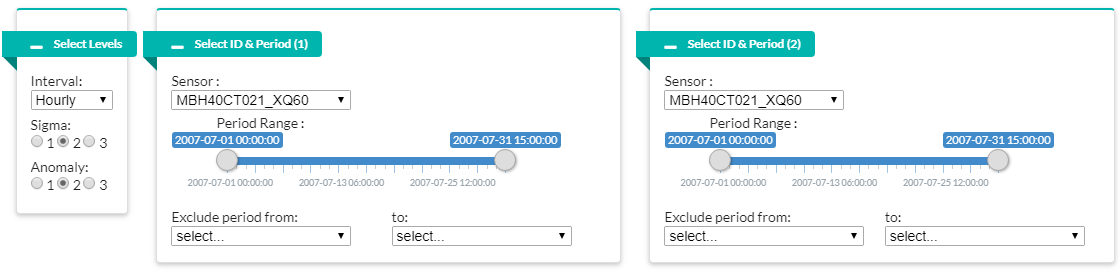
**How to Interpret Anomaly Chart**

* Points that are highlighted in red are deemed to be anomalous after considering trend and seasonality patterns.

**HOW TO USE THE PLOT**

**Allowable selections:**

1. Time Interval (default: Hourly)
2. Violation Sigma Level (default: 2)
3. Anomaly Level (default: 2)
4. Sensor x 2
5. Period Range x 2
6. Exclusion Period x 2



**EDA:**

1. Time interval of data can be specified as every 5 minutes, hourly or daily.
2. There are 3 level each which can be specified for Sigma Violation Run Test and Anomaly Detection Level.
3. Sensor of concern is specified to view their readings individually.
4. Period of concern can be adjusted (via the slider) by specifying the start and end of the period.
5. An exclusion period can be specified to exclude its data from the chart and computation.
6. Users can choose another set of Sensor and Period for ease of comparison.