Statistical Control of Water Quality in the Aracaju, Sergipe, Brazil

1Suzana Leitão Russo, 2Mércia Valéria Vieira dos Santos, 3Lázaro Souto Araújo, Ana Eleonora Paixão, 4José Augusto Andrade Filho, 5Maria Emilia Camargo

1Post-Graduate Program in Intellectual Property Science, Federal University of Sergipe, Brazil
2Department of Statistic Federal University of Sergipe, Brazil

A B S T R A C T

In this work, we use statistical techniques for quality control, using the Control Attributes with the analysis of control chart C. The period of analysis was from 2009 to 2013 in order to perform microbiological analyzes of water quality. The parameter collected by DESO company was Total Coliforms, which is evaluated through the variables Accepted Samples and Rejected Samples. Before the application of control charts C, we performed exploratory analysis of variables and data normality. For normality, we used the Shapiro-Wilk Normality test and the Komogorov-Smirnov test. From the quantitative point of view, in relation to the water supply, we found that the analysis done by DESO in the Total Coliforms parameter, did not find a risk factor for human consumption and results indicated that the samples were free of coliforms. Based on the obtained data from DESO’s samples, the maximum allowable value (MAV) of total coliforms present in the water was 5%, in other words, the company is in accordance with the highest standard required by the regulations. Before the application of control charts C, we performed an exploratory analysis of the variables and we tested the normality of the data. We used STATISTICA 11 software to perform the tests, by means of these results, we could choose the best fit and the most significant and representative values for our work.

INTRODUCTION

The consumption of poor quality water can cause health problemsto the population using such water. For this reason, the Ministry of Health issued the decree n. ° 518/04, which deals with rules and quality standards of drinking water, in order to ensure the supply of good quality water to preserve the health of the population. This ordinance defines allowable maximum values (VMP), based on bacteriological, physical-chemical and organoleptic characteristics from drinking water. Thus, all water intended for human consumption must meet the standards set by that ordinance (Ordinance 518/2004MS).

DESO, Basic Sanitation Company of Sergipe, provided the data by means of their drinking water analysis of the total coliform parameter, to contribute to the quantitative improvement of the treatment process of drinking water geared to consumption by the population.

Theorical Conceptions:

Exploratory Data Analysis:

According to Medri (2001), the exploratory data analysis aims at organize, present, and synthetize data, using charts, tables and descriptive measures as tools. It is used in the initial stages of analysis, focusing on obtaining information that indicates possible models to be used in the final stages. The quantitative variables use measures of location and dispersion, to characterize how data is distributed, they are: $\bar{X}$ (sample average), MED (median); MIN (minimum); MAX (Maximum); $\sigma$ deviation (sample standard deviation) and EP (standard error) (MEDRI and Costa 2001).

Statistical Quality Control:

The graph shown in Figure 1 is the Sequential Chart, a simple tool used to assess the state of statistical control of a process. A special cause is remarkable, and is generally unique, however large enough to produce...
significant disturbances in the production process (SAMOHYL, 2009; MCCRACKEN and CHAKRABORTI, 2013).

![Control Chart](image)

**Fig. 1:** Common Causes and Special Causes of Variation
Source: eliasnogueira.com / blog.

**Control Chart:**
Control charts distinguish random from non-random variation. The basis of control graphs is the sample distribution, which tends to have a probability curve associated with a Gaussian distribution. The control chart has two limits, calculated from the sample data, separating the random variation of non-random variation. The highest value corresponds to the upper control limit (UCL) and the lowest value is the lower control limit (LCL). A statistical sample located between these two limits provides the randomness of the distribution, while a value outside one of the two limits suggests the non-randomness. In control charts is common to use the LSC three times sigma above the midline ($\mu + 3\sigma$) and LIC three times sigma below that ($\mu - 3\sigma$) (REBELATO, 2006; BIROLINE, 2014).

**Control Charts for Attributes:**
There are situations where you have a product attribute. This attribute can be "defective" or "not defective" or it may be the number of defects per unit of product. In this case, control charts for attributes should be applied. The chart C for control by attributes is the most used, according to Montgomery (2008).

**Methodology:**
The methodology for the review of concepts and reasoning of the proposed theme was the literature research. The Excel program and the Statistica11 software are used to generate graphs to generate and analyze results.

The variable was classified as quantitative, in order to know the number of samples analyzed and rejected by means of the parameter analyzed and collected by the DESO water-supply company. The collected data was tested to determine the normality, then the results of the collected samples were analyzed and marked as within or outside of the control limits. The data was collected and surveyed at the city of Aracaju in the period from 2009 to 2013. Corresponding to 60 samples, two variables and a parameter obtained by the company Basic Sanitation Company of Sergipe – DESO. The company analyzed the samples collected and the results were sent to VIGIAGUA, in order to determine if samples are within the limits defined by the Decree No. 518/04 (BRASIL, 2014).

The search parameter used by the company are:

<table>
<thead>
<tr>
<th>Parameter Limits and Ordinance No. 518/04 of the Ministry of Health.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter Limits of the ordinance</td>
</tr>
<tr>
<td>Limitsof the ordinance</td>
</tr>
<tr>
<td>Parameter Description of Analyses Performed</td>
</tr>
</tbody>
</table>

The data collected by the company must follow the MS n. 518/04 norm. The total coliforms variable has a rigorous water treatment, which was collected daily for analysis and to be distributed to the population. The technical evaluation of the bacteriological water quality to comply with the drinking water standards.

**RESULTS AND DISCUSSION**

**Exploratory Analysis of Variables:**
We performed an exploratory data analysis. The highest monthly average was 3779.5 in the year of 2011 and the lowest one was 2615 samples in 2013. There was a difference of 1164.5 samples collected in a two years period. All variation coefficients are above 25%, what indicates a set of heterogeneous data.
**Pareto Chart:**
The Pareto Chart (Figure 2) shows the order of occurrence frequencies, from largest to smallest. Which allows the problem prioritization. This chart shows the cumulative curve percentages and its greatest utility is to allow easy viewing and identification of the rejected samples.

**Table 1:** Statistical Summary of the Total Coliforms Parameter.

<table>
<thead>
<tr>
<th>Year</th>
<th>Valid N</th>
<th>X</th>
<th>MED</th>
<th>MIN</th>
<th>MAX</th>
<th>σ</th>
<th>CV</th>
<th>EP</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>2</td>
<td>3276</td>
<td>3276</td>
<td>83</td>
<td>6469</td>
<td>4515.58</td>
<td>137.84</td>
<td>3193</td>
</tr>
<tr>
<td>2010</td>
<td>2</td>
<td>3249.5</td>
<td>3249.5</td>
<td>86</td>
<td>6413</td>
<td>4473.86</td>
<td>137.68</td>
<td>3163.5</td>
</tr>
<tr>
<td>2011</td>
<td>2</td>
<td>3779.5</td>
<td>3779.5</td>
<td>75</td>
<td>7484</td>
<td>5238.95</td>
<td>138.62</td>
<td>3704.5</td>
</tr>
<tr>
<td>2012</td>
<td>2</td>
<td>3083.5</td>
<td>3083.5</td>
<td>45</td>
<td>6122</td>
<td>4297.09</td>
<td>139.36</td>
<td>3038.5</td>
</tr>
<tr>
<td>2013</td>
<td>2</td>
<td>2615</td>
<td>2615</td>
<td>39</td>
<td>5191</td>
<td>3643.01</td>
<td>139.31</td>
<td>2576</td>
</tr>
</tbody>
</table>

Source: elaborated by the author.

**Normality Check:**
Regarding to the Shapiro-Wilk (W) normality test, the calculated W was 0.96692, which is greater than the calculated p-value 0.10300. Therefore, we can state that Total Coliforms parameter comes from a normal sample. The p-value calculated by means of the hypothesis test was greater than 0.05, i.e. there is no evidence to reject the null hypothesis (H₀), so this data follows a normal distribution.

**Application of Attribute Control Graphic:**
The ratio identifies the level of water quality that was observed variability, where two points are out of control. It was also noted that there is evidence of a lack of randomness however, this presents a supposedly
random behavior of the points found outside the control limits or is in an unstable process, means that a special cause variation is present.

It was found that due to high incidence of rain in August 2009 and January 2011 period, the riverbed was transported organic and inorganic materials, resulting in an increase of total coliforms due to failure of the treatment plant - ETA.

![Control Chart](chart1)

**Fig. 4**: Control Chart C Source: elaborated by the author.

We found that 6.66% of the samples taken did meet the required specifications, and 33.33% were out of control. Even outside the standards define in the Decree 518/04 MS, such quantitative poses no risk to the population, given that the events were caused due to the rainy season and these are only to be considered harmful when applicant.

![Graph of C Control Revised](chart2)

**Fig. 5**: Graph of C Control Revised Source: elaborated by the author.

Thus, therevisedchartshowsthatthe water qualitymeetsestablished standards, whatensures that the water deliveredby the supply companyduring the years analyzedis adequate, and aseasonal variationdoes not affectthe health of the population.

**Conclusions**:

From the exploratory analysis of the variables, it was observed that the samples achieved the highest monthly average in the year 2011 and the lowest monthly average in 2013, which makes us understand that in two years there have been differences of total coliforms found in the treated water. The collected data are normal at a significance level of 5 %, and the two points outside the upper control limit were identified in the graph in August 2009 and January 2011, due to a failure in the treatment station and then we revised the control chart, which shows the quality of the water provided by the company. From the quantitative point of view, we found that there was a risk factor for human consumption. The results indicated that the samples were free of coliform therefore with no indication of fecal contamination. According to the Decree 518/2004 of the Ministry of Health, this variability presents no health risk to the population of Aracaju since it meets the standards specified.

**REFERENCES**


Rebelato, M.G., 2006. Study on the application of control charts for processes of saturation of paper, XIII SIMPEP, Bauru, SP.