

---

## Process Improvement Through Algorithms And Statistical Tools

Luis Alberto Muñoz Bedoya<sup>1</sup>, Naslesly Liliana Cardenas Parada<sup>2</sup>

<sup>1</sup> MSc(c) Controles Industriales, [luismb@unipamplona.edu.co](mailto:luismb@unipamplona.edu.co), Grupo de investigación de Ingeniería Biomédica GIBUP, Universidad de Pamplona, Km 1 Vía Bucaramanga Ciudad Universitaria, Pamplona, Colombia

<sup>1</sup> MSc(c) Ingeniería Industrial, [Naslesly.cardenas@unipamplona.edu.co](mailto:Naslesly.cardenas@unipamplona.edu.co), Grupo de investigación INGPRO-GES, Universidad de Pamplona, Km 1 Vía Bucaramanga Ciudad Universitaria, Pamplona, Colombia

**Article History:** Received:11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

---

**Abstract:** Today, large, medium and small companies face an accelerated development and multiple challenges that test their ability to overcome and permanence. For this reason, Companies which seek to remain in a consumer market must keep in constant change and continuous improvement of their processes, procedures and other internal and external factors that influence the development of products, services and designs. It is remarkable to see how many of these companies have decided to design and implement tools that allow them to maintain control of their processes for better results, which ultimately could translate into the optimization of the processes. This paper shows how algorithms such as ant colony, k-means and other statistical tools influence and improve the expected results, the research was supported by four phases of development, starting with a diagnosis of the processes to be analyzed and ending with the validation of the tools, whose results show an improvement in the processes.

---

**Keywords:** Product, Algorithm, Statistics, Improvement, Processes.

---

### Introduction

Organizations need to continuously improve their processes and procedures in order to enhance their ability to cope with risk factors, i.e. factors that can affect them negatively or positive factors that represent an opportunity or advantage, which also have a positive impact or influence [1]. In this sense, organizations need to adapt and implement qualitative and quantitative tools that improve their processes, allowing them to better control them. For this purpose, the study focuses on showing how algorithms such as ACO, corresponding to the ant colony algorithm, influence and improve a process, as do tools such as six sigma, control charts or k-means.

As mentioned above, the ant colony algorithm or as it is known (ACO) is used in this case, performing an application of the same, focused on solving the problem of optimal routing, considering the actual direction of the streets. By implementing the ACO algorithm, it is possible to simulate the place of origin and destination of a user, based on multiple iterations [2].

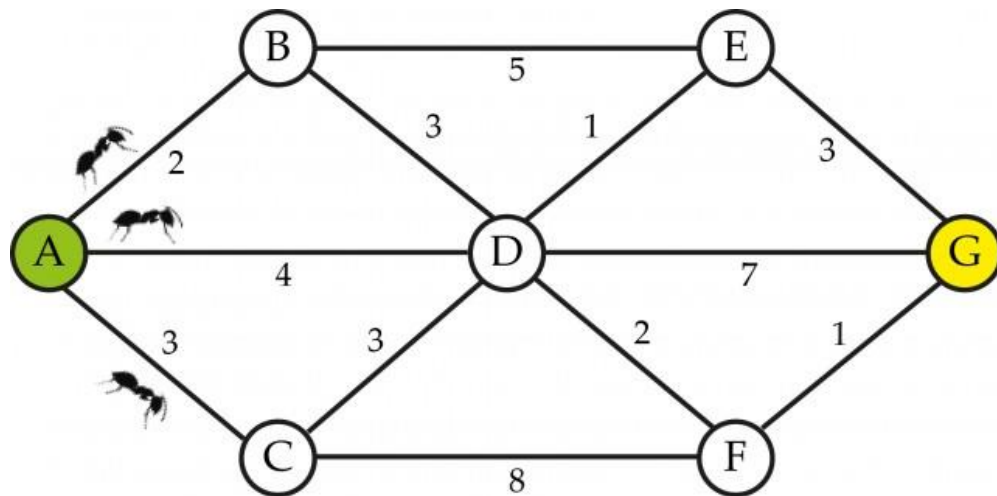
As a complement, the application of statistical tools to a production process allowed to determine the variation of the process, thus allowing to measure the variables that were outside the company's specification limits, which in turn determined the status of the process. The application of these tools such as six sigma, control charts, sipoc diagrams, made it easier for the company to take preventive and corrective actions on the process. On the other hand, these data were supported by means of the grouping method known as k-means.

However, there are multiple algorithms and ways to solve optimization problems. Among them are the Taboo Search [3], Simulated Annealing [4], GRASP [5], Genetic Algorithms [6], Particle Swarm Optimization (PSO) [7] and Ant Colony-Based Optimization (ACO) [8] [9], among others [2].

These last two meta-heuristics fall into the category of bio-inspired or artificial life and collective intelligence algorithms, since the potential of these models to solve problems is given by the cooperation between individuals in a

direct or indirect way. Specifically, in the ACO meta-heuristic, cooperation is performed between ants in an indirect way [2].

For the sake of continuity, it is important to understand how each of the tools works. According to Gambardella et al. in [9], This kind of algorithms are inspired by the real behavior of ants, which always find the shortest path between the place where they live and the food sources, due to the indirect exchange of information through pheromone deposition.



*Image 1: Behavior of the ants*

Source: [10]

The ant colony optimization method, also known as ACO, was introduced by Marco Dorigo in early 1990, whose purpose was based on the solution of complex optimization problems [9] [11]. The method belongs to the class of heuristic methods, which correspond to algorithms used to obtain better results to complex problems, considering time as the main variable.

On the other hand, there are tools such as six-sigma. In 1988, Motorola won the Malcolm Baldrige national quality award for the six-sigma tool (SS) [12], due to the results that were evidenced, which increased the level of quality of several products and allowed significant savings of money. By the year 2000, many companies were preparing the initiative with respect to SS and by the year 2003, more than 100,000 million dollars in combined savings had been accounted for because the SS tool became the worldwide standard of quality business practice [13] [14].

Six-Sigma is a tool that reduces variation, improves product quality, reduces waste, significantly improves process performance, aligns employee efforts to organizational objectives, emphasizes measurement and results, promotes the use of statistical analysis, develops focused products, improves market share, and improves customer retention [15].

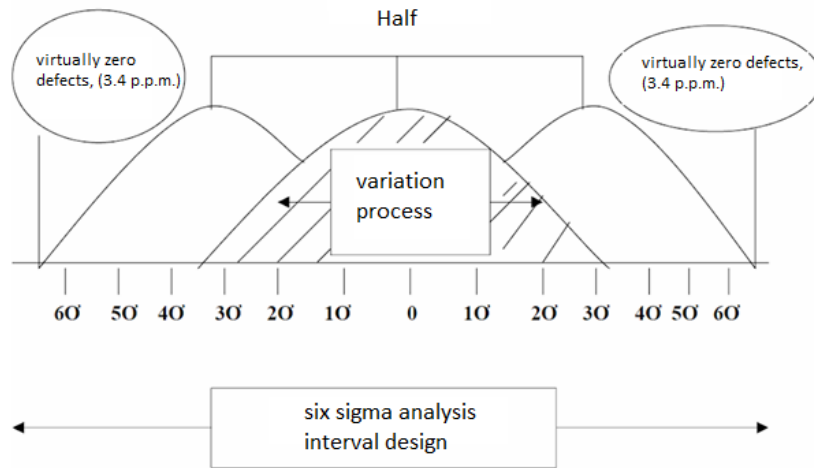


Illustration 1. Six-Sigma

Source: [16]

On the other hand, the grouping method or K-means, as this algorithm is known, was proposed by MacQueen in 1967. This algorithm allows the division and grouping of data in several sets, taking into account the characteristics of the study [17]. In this case, the algorithm was used to look at the dispersion with respect to the weight variable of the product being evaluated, allowing to determine if the variable was within the required specifications.

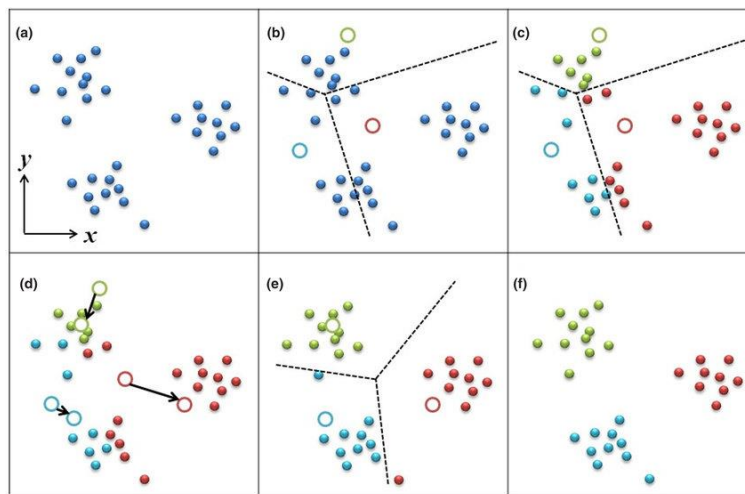


Image 2: K-means Method

Source: [18]

The purpose of K-means is to separate the exact data in order to group them into different groups, which are known as k-groups. Each group is determined by the average of the points that compose it. For each one of them the Centroid will be found and it is the one that marks or determines if the data belongs or not to that group. The number of groups that are determined depends on the person who is conducting the study [19].

## MATERIALS AND METHODS

The research was developed through four main stages, the first stage was supported by qualitative tools that allowed to know the current status of the processes to which the algorithms and statistical tools were going to be applied. During this stage, a survey and evaluation of the processes was also carried out to determine if it was really necessary to apply tools. The second and third stages were supported by the application of the statistical tools and algorithms to the different evaluation processes. The fourth and last stage was the validation of these algorithms, which made it possible to verify that the tools and algorithms really allowed for an improvement in the processes.

Free tools such as NodeJS as JavaScript execution environment, ReactJS as modern web development library, Open Street Maps and the Open Stack, and student Matlab were used for its development [2] [20]. Images (3), (4) show a brief description of the processes:

### Mat-Lab (Matrix Lab)



*Image 3. Matlab for students (trial version)*

Source: [21]

### Mat-Lab (Matrix Lab) Laboratory of Matrices

It is a scientific, engineering and mathematical software that has laid the foundation for today's technological and multidisciplinary development. Its development was based on commercial libraries: BLAS (Basic Linear Algebra), LAPACK (Linear Algebra Package), JIT (Just In Time Accelerator), which make algorithms and operations based on matrices and vectors much faster and optimal in execution [2].



*Imagen 4 Logo Open Street Maps*

### Open Street Maps (Freeware)

OpenStreetMap (also known as **OSM**) is a collaborative project to create free, editable maps. Instead of the map itself, the data generated by the project is considered its main output.

In this software, maps are created using geographic information captured from mobile GPS devices, orthophotographs and other free sources.

### Ant Colony Optimization Algorithm

The ACO algorithm includes the evaporation mechanism, to decrease the value of all pheromones and avoid accumulation on any component. There are also mechanisms to perform centralized actions, such as local optimization procedures [2].

$$p_{xy} = \frac{(\tau_{xy})(\eta_{xy})}{\sum(\tau_{xy})(\eta_{xy})} \quad (1)$$

*Equation 1. The probability of selecting a path*

$$\tau_{xy} \leftarrow (1 - \rho)\tau_{xy} + \sum \tau_{xy} \quad (2)$$

$$\Delta\tau_{xy} = \begin{cases} \frac{Q}{L} & \text{If the ant traveled the path } xy \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

*Equation 3. Pheromone contribution of each ant*

Where,  $p_{xy}$  is the probability of selecting a path from x to y (1)

$\tau$  is the pheromone per ant,

$\eta$  is the visibility which in another word is the inverse of the length of a path from x to y,

$\tau_{xy}$  is the amount of pheromone per ant and  $\rho$  is the frequency of operation (2),

$\Delta\tau_{xy}$  the pheromone contribution of each ant (3),

$Q$  is the learning parameter, usually equal to 1 (3), and

$L$  is the path cost per ant learning parameter, usually equal to 1 (3).

### K-means algorithm

The basic description corresponding to the method states: Objects are represented by d-dimensional real vectors  $(x_1, x_2, \dots, x_n)$  and the k-means algorithm constructs k groups where the sum of distances of objects, within each group  $S = \{S_1, S_2, \dots, S_k\}$  to their centroid is minimized.

$$\min_S E(\mu_i) = \min_S \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (4)$$

*Equation 4. K-means Algorithm*

**Source:** [22]

where S is the data set whose elements are the objects x<sub>j</sub> represented by vectors, where each of its elements represents a characteristic or attribute. There will be k groups or clusters with its corresponding centroid μ<sub>i</sub>.

## RESULTS AND ANALYSIS

In a practical way, the following steps were stipulated for the elaboration of the algorithm in M language:

1. Label each marker and obtain its geographic coordinate (latitude and longitude).
2. Determine the possible paths between markers.
3. Define paths and list them.
4. Calculate distances between listed paths with the haversine formula.
5. Indicate the start and end of each path.
6. Drop the ants on the map.
7. Iterate until all the thrown ants arrive from the origin to the destination.
8. Measure the costs of the path traveled by each ant.
9. Make the collective contribution of pheromones according to the number of times that each path was traveled or not by each ant.
10. Iterate cyclically until the lowest possible cost of all the paths is obtained, making the pheromone contribution in each cycle

### Visualization on maps

For the visualization of the map, the openStreetMaps provided by MapBox was used since it provides a query API for web browsers. Using the ReactJS library, the map visualization interface was programmed by creating a custom hook that optimizes the loading of the maps, the geographic coordinates and additionally the location of the munchers. The following is a part of the hook code used to optimize the map utilization:

```
import { useRef, useState, useEffect, useCallback } from 'react';
import mapboxgl from "mapbox-gl";
import { v4 } from 'uuid';
import { Subject } from "rxjs";
```

```

mapboxgl.accessToken =
'pk.eyJ1IjoiYmFzZGVzbnHkiLCJhIjoiY2tqdDlkc2hoMXQyYjJ5azM1OWp0MG8wYSJ9.v80WhyW_NMFj3wyiQVa
Ryw';

export const useMapBox = ( puntoInicial ) => {

    // referencia al Div del mapa
    const mapaDiv = useRef();
    const setRef = useCallback(( node ) => {
        mapaDiv.current = node;
    }, []);

    // Referencia a marcadores
    const marcadores = useRef({});

    // Observables de RXJS
    const movimientoMarcador = useRef( new Subject() );
    const nuevoMarcador = useRef( new Subject() );

    // Mapa y coords
    const mapa = useRef();
    const [ coords , setCoords ] = useState( puntoInicial );

    // Función para agregar marcadores
    const agregarMarcador = useCallback((ev, id) => {
        const { lng, lat } = ev.lngLat || ev;

        const marker = new mapboxgl.Marker();
        marker.id = id ?? v4(); // TODO: verificar si el marcador tiene id
        marker
            .setLngLat(
                [lng, lat]
            )
            .addTo( mapa.current )
            .setDraggable( true );

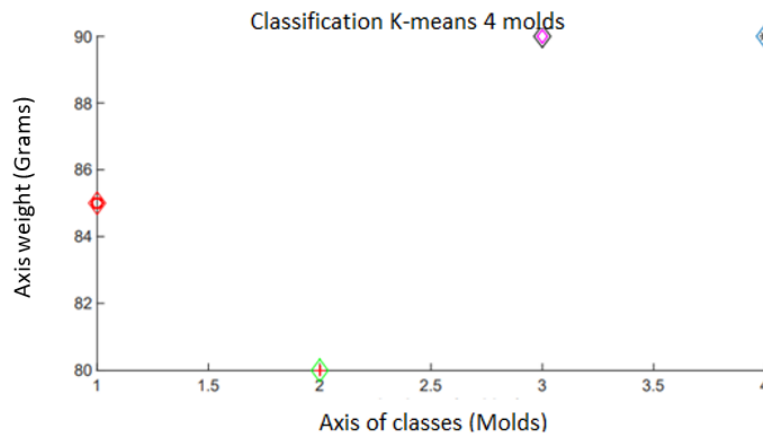
        marcadores.current[ marker.id ] = marker;

        // Si el marcador no tiene ID no emitir para evitar ciclo infinito
        if(!id){
            nuevoMarcador.current.next({
                id: marker.id,
                lng,
                lat
            })
        }
    }
}

```

The function previously created in Matlab for the ACO algorithm was migrated to Javascript language in order to be used as a function in the web browser. Some native Matlab functions had to be implemented to make the algorithm execution identical. Finally, the validation tests performed contain the start node, the end node, the blocking of a street to search for an alternate route, the visualization of the route on the map and finally the comparison with Google Maps. Each test was carried out with 8 ants with an operating frequency of  $\rho = 0.01$ , learning parameter  $Q = 1$ , and finally proving to be of great use in route planning.

Finally, the result of the application of the K-means algorithm to the selected process is shown. In this case, the algorithm was worked with the variable weight, in the Matlab, one of the obtained results was the following



*Illustration 2. K-means Classification*

Source: [19]

The centroids in this case were 4 because the 4 production molds were taken as reference, 80 values were taken, corresponding to 80 study units. Now 4 points were found, each point representing a centroid, and next to each point the other points closest to the corresponding weights are located. Although the algorithm works with more variables and is very useful for grouping data, in this study the purpose was to check if the algorithm could identify variability in the process, which was correct.

## CONCLUSIONS

- Statistical tools and other tools allow to know the of a process, which is very useful when applying corrections to those failures or causes weaknesses. The use of these tools extends to all organizations that want to know, optimize or improve processes.
- Although many companies do not use this type of tools, it was possible to verify the great utility that they could have, being such that their use could improve the results with respect to a process or a procedure, this being the basis for optimization and making possible to obtain better results.
- The ant colony algorithm for the purpose of optimizing transport logistics processes of any nature is a very useful tool if it is combined with modern tools and even mobile applications.

## REFERENCES

- [1] J. Duarte, «USS-Universidad San Sebastian,» [En línea]. Available: <http://www.ipsuss.cl/ipsuss/columnas-de-opinion/jonathan-duarte/los-complejos-desafios-que-enfrentan-las-organizaciones/2017-02-27/135754.html>.
- [2] L. A. M. Bedoya, «ANÁLISIS Y ESTIMACIÓN DE RUTAS ALTERNAS USANDO EL ALGORITMO DE OPTIMIZACIÓN DE COLONIA DE HORMIGAS (OCH) EN LA ZONA CÉNTRICA DE LA CIUDAD DE PAMPLONA,» 2021.
- [3] F. Glover, «Heuristics for Integer Programming Using Surrogate Constraints,» *Decision Sciences*, 1977. 8: p. 156-166..



- [4] S. C. M. .. V. e. Kirkpatrick, «Optimization by simulated. annealing.,» *Science*, 1983. 220(4598): p. 671-680..
- [5] T. Resende, «A probabilistic heuristic for a computational difficult set covering problems. *Operations research letters*,» 1989. 8: p. 67–71. .
- [6] D. Goldberg, « *Genetic Algorithms in Search. Optimization and Machine Learning.*,» University of Alabama: Addison-Wesley Publishing Company. , 1998.
- [7] J. Eberhart, « *Particle swarm optimization. in on neural networks.*,» Piscataway, NJ. , 1995..
- [8] M. Stützle, «ACO Algorithms for the Traveling Salesman Problem. *Evolutionary Algorithms in Engineering and Computer Science: Recent Advances in Genetic Algorithms, Evolution Strategies, Evolutionary Programming, Genetic Programming and Industrial*,» 1999.
- [9] M. Dorigo y L. M. Gambardella, «Sistema de colonia de hormigas: Un enfoque de aprendizaje cooperativo para el problema del vendedor ambulante».
- [10] G. P. Ripollés, 2016 Octubre 18. [En línea].
- [11] M. Dorigo y C. Blum, «Teoría de optimización de colonias de hormigas,» *sciencedirect*.
- [12] C. G. S. y K. M. Gowen, «Simultaneous implementation of Six Sigma and knowledge management in hospitals».*International Journal of Production Research*.
- [13] R. y. S. P. Reosekar, «Six Sigma methodology: a structured review».
- [14] J. L.-R. M. M.-M. y. Y. B.-L. Carlos Gastelum-Acosta, «Seis Sigma en Instituciones de Educación Superior en México,» 2018.
- [15] F. O. IBARRA, «REDUCCIÓN DE ARTICULOS DEFECTUOSOS Y DESPERDICIOS EN LA EMPRESA GAMA EDITORES REYES MEDINA CIA LTDA.UTILIZANDO LA METODOLOGIA SEIS SIGMA.».
- [16] J. Diaz y G. López. [En línea]. Available: <https://www.emprendices.co/metodologia-six-sigma-calidad-industrial/>.
- [17] J. MacQueen, «Facultad de Informática. UNLP.,» 2016. [En línea].
- [18] Y.-C. Lai, «researchgate,» 2018. [En línea]. Available: [https://www.researchgate.net/figure/A-schematic-illustration-of-the-K-means-algorithm-for-two-dimensional-data-clustering\\_fig2\\_324073652](https://www.researchgate.net/figure/A-schematic-illustration-of-the-K-means-algorithm-for-two-dimensional-data-clustering_fig2_324073652).
- [19] N. L. c. Parada, «ANÁLISIS PARA LA OPTIMIZACIÓN DEL PROCESO DE PRODUCCIÓN EN LA FABRICACIÓN DE HELADOS,» 2020.
- [20] J. M. Cerón, 2019. [En línea].
- [21] recluit, «recluit-Mariana Velazquez,» 2020. [En línea]. Available: <https://recluit.com/que-es-matlab/#.YEL9nWgzZPY>.
- [22] unioviedo, «unioviedo,» [En línea]. Available: [http://www.unioviedo.es/compnum/laboratorios\\_py/kmeans/kmeans.html](http://www.unioviedo.es/compnum/laboratorios_py/kmeans/kmeans.html).
- [23] C. r. L. o. s. A. t. u. r. o. r. b. e. s. A. g. A. r. í. n, «opTimización por colonia de Hormigas: aplicaciones y Tendencias».
- [24] «webstorm,» [En línea]. Available: <https://www.jetbrains.com/webstorm/>.

[25] M. Salathé, Nature in Code, Leanpub ed., 2016.

[26] «mathworks,» [En línea]. Available: <https://la.mathworks.com/>.

[27] <https://reactjs.org/>, «reactjs,» [En línea]. Available: <https://reactjs.org/>.