

# MEJORAMIENTO DE LA CAPACIDAD REAL DE PROCESOS ( $C_{pk}$ ), PARA PROCESOS CON RESTRICCIONES, A TRAVES DEL METODO DE PERTURBACION SIMULTANEA DE APROXIMACIONES ESTOCASTICAS

## *IMPROVEMENT OF REAL PROCESS CAPACITY (CPK), FOR PROCESS WITH CONSTRAINTS, THROUGH THE METHOD OF SIMULTANEOUS PERTURBATION OF STOCHASTIC APPROXIMATION*

### ABSTRACT:

The Actual Process Capacity Index ( $C_{pk}$ ) is an indicator used to measure the percentage of production that meets the specifications established for the good or service generated by the process. Stochastic Approaches Algorithms (AE), correspond to iterative methods that generate a succession of random vectors that converge, with a certain type of probability: to the root of a real function (algorithms of the type: Robbins and Monro); or converge to the optimal value of an actual function (algorithms of the type: Kiefer and Wolfowitz). This article presents an AAE of the type Kiefer, Wolfowitz, with projection on the feasible space of the independent variables of the process, to test the conditions under which the process reaches its maximum  $C_{pk}$ . The convergence probability one of this algorithm to a Karush-Kuhn-Tucker point is tested). In addition, the Monte carlo simulation results are presented graphically in the application of this algorithm, where there is an increasing trend in the succession of the values  $C_{pk}$  generated in this simulation process

Simultaneous Perturbation in Stochastic Algorithms (SPSA)  
Stochastic Approach Algorithms  
Actual Process Capacity Index ( $C_{pk}$ )  
Optimization with Restrictions  
Quality

### RESUMEN:

El Índice de Capacidad Real del Proceso ( $C_{pk}$ ) es un indicador utilizado para medir el porcentaje de la producción que cumple con las especificaciones establecidas para el bien o servicio generado por el proceso. Los Algoritmos de Aproximaciones Estocásticas (AAE), corresponden a métodos iterativos que generan una sucesión de vectores aleatorios que convergen, con cierto tipo de probabilidad: a la raíz de una función real (algoritmos del tipo: Robbins y Monro); o converger al valor óptimo de una función real (algoritmos del tipo: Kiefer y Wolfowitz). En este artículo se desarrolla un AAE del tipo Kiefer, Wolfowitz, Simultaneaon proyección sobre el espacio factible de las variables independientes del proceso, para determinar las condiciones en que el proceso alcance su máximo  $C_{pk}$ . Se prueba la convergencia con probabilidad uno de este algoritmo a un punto Karush-Kuhn-Tucker). Además, se presenta de manera gráfica los resultados de simulación Monte Carlo en la aplicación de este algoritmo, donde se observa una tendencia creciente en la sucesión de los valores del  $C_{pk}$  generados en este proceso de simulación.

Perturbación Simultanea en Algoritmos de Aproximaciones Estocásticas (SPSA, por sus siglas en ingles)  
Algoritmos de Aproximaciones Estocásticas  
Índice de Capacidad Real del Proceso ( $C_{pk}$ )  
Optimización con Restricciones  
Calidad

## 1.- INTRODUCCION

### 1. INTRODUCCIÓN

Las condiciones óptimas de operación de los procesos en los que la variable de respuesta o característica de calidad depende de un conjunto de factores controlables, pueden ser consideradas como los niveles o valores de estos factores en los que el porcentaje de la producción, que está dentro de las especificaciones, es mayor o igual que un valor mínimo requerido. Esto les permite garantizar el cumplimiento de las metas de calidad (partes defectuosas por millón) establecidas y a la vez asegurar su permanencia en el mercado.

Uno de los índices, con el que puede medirse la capacidad o aptitud de los procesos para cumplir con las especificaciones, es el llamado Índice Real de Capacidad del Proceso, denotado por  $C_{pk}$ . Dicho índice puede ser usado bajo el cumplimiento de los supuestos de que las mediciones de la variable de respuesta o característica de calidad representan valores de variables aleatorias independientes con distribución normal y que el proceso se encuentra bajo control estadístico (Montgomery, 2004); (Arcidiacono y Nuzzi, 2017) presentan una revisión de los índices capacidad del proceso, además de la relación de estos con el desempeño de los procesos y los procesos sigma.

El índice de capacidad real del proceso  $C_{pk}$  representa una solución apropiada al problema de contar con una herramienta útil en la medición de la capacidad de los procesos para cumplir con las especificaciones establecidas para las características de calidad. Queda pendiente la determinación del método que resulte apropiado para determinar las condiciones de operación de los factores controlables, en las que el  $C_{pk}$  logre su máximo valor o exceda a un valor mínimo requerido, lo que genera el planteamiento del problema que se define en el siguiente apartado.

Suponga que el símbolo  $C_{pk}$  representa al índice de la capacidad real de un proceso productivo, en el que cada uno de sus  $d$  factores controlables se encuentran operando en los niveles o valores (condiciones de operación) denotados por el vector  $X = (X_1, \dots, X_d) \in \mathbb{R}^d$  ( $\mathbb{R}^d$  representa al espacio cartesiano de dimensión  $d$ ). Entonces puede asegurarse que existe una función

$$C_{pk} = f(X; LES, LEI) \quad (1.1)$$

con la que es posible, dados los valores de los parámetros Límite de Especificación Superior ( $LES$ ) y Límite de Especificación Inferior ( $LEI$ ), conocer el valor del  $C_{pk}$  al evaluarla en las condiciones de operación  $X$ . Para lograr lo anterior, es necesario medir los valores de las funciones  $\mu(X)$  y  $\sigma(X)$  donde  $\mu: \mathbb{R}^d \rightarrow \mathbb{R}$  y  $\sigma: \mathbb{R}^d \rightarrow \mathbb{R}$  son funciones reales.

En la mayoría de los casos prácticos no es posible determinar los valores de los parámetros: valor esperado  $\mu(X)$  y desviación estándar  $\sigma(X) \neq 0$  de la variable de respuesta en las condiciones de operación  $X$ , lo que trae como consecuencia que el valor del  $C_{pk}$  tampoco se logre especificar. Sin embargo, este valor puede ser estimado a través del estadístico  $\hat{C}_{pk}$ , el cual representa el valor de una variable aleatoria con distribución y varianza desconocida, pero condicionado a que su valor esperado sea igual a la función acotada  $f(X; LES, LEI)$ , es decir

$$C_{pk} = E(\hat{C}_{pk} | X; LES, LEI) \quad (1.2)$$

lo que indica que el  $C_{pk}$  puede ser expresado como el valor esperado de la variable aleatoria  $\hat{C}_{pk}$ , variable que siempre puede ser medida, dado el vector  $X$  y los valores de los parámetros  $LES$  y  $LEI$ .

Suponga, además, que el estadístico  $\hat{C}_{pk}$  puede ser expresado mediante la ecuación

$$\hat{C}_{pk} = f(X; LES, LEI) + \varepsilon \quad (1.3)$$

en la que  $f$  representa a una función real cuya expresión algebraica es desconocida y  $\varepsilon$  corresponde el valor de una variable aleatoria.

Este trabajo plantea resolver el siguiente problema

$$\underset{X \in G}{\text{maximizar}} f(X; LES, LEI) \quad (1.4)$$

la existencia de la solución a este problema se garantiza con el cumplimiento de los dos supuestos siguientes: 1) la función  $f$  es continua en  $G \subset \mathfrak{R}^d$ ; 2) el conjunto  $G = \{(X_1, X_2, \dots, X_d) | u_i \leq X_i \leq v_i; u_i, v_i \in \mathfrak{R}, i = 1, 2, \dots, d\}$  representa a una región rectangular cerrada y acotada, y esta corresponde a la región factible del proceso productivo; es decir, el conjunto de posibles condiciones de operación o de valores que puede asumir cada uno de los factores controlables, o en su caso las restricciones del proceso).

Para determinar la solución al problema presentado en la ecuación (1.4) es necesario optimizar simultáneamente los valores de  $\mu(X)$  y  $\sigma(X)$ , lo que se lleva a cabo dependiendo de lo que resulta mejor para el valor de la característica de calidad o variable de respuesta que puede resultar en cualquiera de los siguientes tres casos: *primero*, lo nominal es lo mejor; *segundo*, lo mayor es lo mejor; *tercero*, lo menor es lo mejor. En estos tres casos se requiere, respectivamente, que el valor de  $\mu(X)$  sea igual al nominal, lo más grande posible o lo más pequeño posible, mientras que en cualquier caso se pretende que el valor de  $\sigma(X)$  sea el mínimo ( $\forall X \in G$ ). Para lograr lo anterior se requiere contar con un método apropiado para optimizar simultáneamente  $\mu(X)$  y  $\sigma(X)$ .

Aquí se propone una método útil para mejorar el valor del índice de capacidad del proceso  $C_{pk}$ , además de garantizar que la condición óptima de operación del proceso generada pertenezca a  $G$ , para lo cual se utiliza un algoritmo de aproximaciones estocásticas, el cual tiene la característica de incrementar el valor del índice de capacidad  $C_{pk}$  de los procesos productivos, en los que el valor de la variable de respuesta es una función (desconocida) de un conjunto de factores controlables en  $G$ . Para verificar que el algoritmo propuesto es útil para mejorar el  $C_{pk}$ , se demuestra que la sucesión generada con este algoritmo converge con probabilidad 1 a un punto Karush-Kuhn-Tucker, además, se analiza el comportamiento del  $C_{pk}$  en las “primeras iteraciones” del método.

## 1.2. Método Propuesto

El método propuesto para el mejoramiento del valor del índice  $C_{pk}$  de los procesos productivos, es desarrollado basándose en los métodos de algoritmos de aproximaciones estocásticas, presentados por (Robbins, y Monro, 1951), para determinar las raíces de funciones cuya relación funcional es desconocida, y modificado por (Kiefer, y Wolfowitz, 1952) para determinar el valor óptimo de este tipo de funciones. Una muestra de los avances de desarrollo en este tipo de métodos de optimización, así como de sus aplicaciones en la resolución de problemas de optimización de procesos en los que únicamente se disponen valores de la variable de respuesta perturbada por el error experimental están disponibles en las siguientes bibliografías: Alaeddini y Klein, 2019; Benveniste, et.al, 1990; Borkar, et.al 2018; Chaiba, et. al, 2018; Chin,(1997) Granichin y Amelina, 2015; Ishizuka, y Kajiwara 2014; Kushner, y Clark, 1978; Kushner y Yin, 2003; Li, & Mohammad 2013; Popovic (2005); Ramaswamy 2019; Ros R., et. al. (2020); Sadegh, 1997; Spall, 1992; 1997 y 2000; Ozguven, y Ozbay 2008; Wang 2018; Zeren D. et. al 2018.

Suponga que tanto el promedio ( $\mu$ ) y la varianza ( $\sigma^2$ ) de la variable de respuesta, son funciones desconocidas, independientes entre sí, de un conjunto de  $d$  factores controlables, los cuales pertenecen a  $G$ . Entonces como consecuencia de la aplicación del método aquí presentado, se genera la sucesión de vectores aleatorios  $\{X_k\}$  en  $\mathfrak{R}^d$  que converge con probabilidad 1 a un vector  $X^* \in \mathfrak{R}^d$  ( $X^*$  es un punto Karush-Kuhn-Tucker). A continuación se definen algunos parámetros, considerados en (Andradóttir, 1996) que utiliza el método de perturbación simultánea como estimador del vector gradiente del  $C_{pk}$  (Spall, 1992) y con el método de proyección sobre la región factible de (Sadegh, 1997), y que son necesarios para la implantación de este método.

Sean:  $\{\Delta_k\}$  ( $\Delta_k \in \mathfrak{R}^d$ ) una sucesión de vectores idéntica e independientemente distribuidos  $\Delta_k = (\Delta_{k1}, \Delta_{k2}, \dots, \Delta_{kd})$ , donde cada  $\Delta_{ki} \in \mathfrak{R}$  con  $i = 1, 2, \dots, d$  son variables aleatorias independientes con distribución Bernoulli simétricas con respecto al cero, para cada  $k$ ;  $\{a_k\}$  y  $\{c_k\}$  sucesiones de números reales que convergen a cero, de las que Sadegh y Spall (1998), recomiendan que deben construirse de acuerdo al siguiente criterio

$$a_k = \frac{a}{k^\alpha}, \quad c_k = \frac{1}{k^\gamma} \quad (1.5)$$

con  $a > 0$ ,  $0 < \alpha \leq 1$ ,  $\alpha - 2\gamma > 0$ ,  $3\gamma - \alpha/2 \geq 0$ , con el cual se garantiza el cumplimiento de los supuestos para estas dos sucesiones.

El método propuesto opera de acuerdo al siguiente procedimiento:

*Paso m0.* Considere a las condiciones iniciales de operación del proceso como el punto inicial representado por  $X_0$ ;

*Paso m1.* Dadas las condiciones de operación  $X_k \in G \subseteq \mathfrak{R}^d$  ( $k = 0, 1, 2, \dots$ ) obtenga una estimación del valor promedio  $\hat{\mu}_k$  y la desviación estándar  $\hat{\sigma}_k$  de la variable de respuesta, en la iteración  $k$ , mediante

$$\hat{\mu}_k = \frac{\sum_{t=1}^n Y_t}{n} \quad (1.6)$$

$$\hat{\sigma}_k = \sqrt{\frac{\sum_{t=1}^n (Y_t - \hat{\mu}_k)^2}{n-1}} \quad (1.7)$$

$$\hat{\sigma}_k = \frac{Rango}{d_2}, \quad (si, n \leq 8) \quad (1.7b)$$

la variable  $Y_t$  representa el valor observado en la  $t$ -ésima ( $t = 1, 2, \dots, n$ ;  $n \geq 3$ ) medición de la variable de respuesta que se presenta en las condiciones de operación correspondientes a la iteración  $k$  ( $X_k$ ),  $n$  representa el total de observaciones o replicas de la variables de respuesta en cada iteración;

*Paso m2.* Obtener una estimación del índice de capacidad del proceso  $C_{pk}$  a través de

$$\hat{C}_{pk} = \min \left\{ \frac{LES - \hat{\mu}_k}{3\hat{\sigma}_k}, \frac{\hat{\mu}_k - LEI}{3\hat{\sigma}_k} \right\} \quad (1.8)$$

$LES$  y  $LEI$  representan, respectivamente, al valor del límite de especificación superior e inferior de la variable de respuesta;

*Paso m3.* Establezca las condiciones de operación

$$X_k^{(+)} = (X_1 + \Delta_{k1}c_k, X_2 + \Delta_{k2}c_k, \dots, X_d + \Delta_{kd}c_k) \quad (1.9)$$

$$X_k^{(-)} = (X_1 - \Delta_{k1}c_k, X_2 - \Delta_{k2}c_k, \dots, X_d - \Delta_{kd}c_k)$$

(1.10)

donde cada componente  $X_i$  del vector  $X_k$ , para  $X_k^+$  o  $X_k^-$  se define como

$$X_i = \begin{cases} v_i - \Delta_{ki}c_k & \text{si } X_i > v_i - \Delta_{ki}c_k \\ u_i + \Delta_{ki}c_k & \text{si } X_i < u_i + \Delta_{ki}c_k \\ X_i & \text{si } u_i + \Delta_{ki}c_k \leq X_i \leq v_i - \Delta_{ki}c_k \end{cases} \quad (1.10)$$

( $i = 1, 2, \dots, d$ ).

Observe que con lo anterior se garantiza que la condiciones de operación representadas por  $X_k$ ,  $X_k^{(+)}$  y  $X_k^{(-)}$  estén en  $G_k = \{X_k \mid u_i - \Delta_{ki}c_k \leq X_i \leq v_i + \Delta_{ki}c_k\} \subset G$ , además, de que se cumple que  $\lim_{k \rightarrow \infty} G_k = G$ .

*Paso m4.* Obtenga  $n$  mediciones o réplicas de la variable de respuesta para cada una de las condiciones de operación  $X_k^{(+)}$  y  $X_k^{(-)}$ . Después use los resultados anteriores para estimar los valores de los índices de capacidad respectivos  $\hat{C}_{pk}^{(+)}$  y  $\hat{C}_{pk}^{(-)}$  usando la ecuación (1.8). Con estos valores, a su vez estime el vector gradiente de la función  $C_{pk}$  evaluada en  $X_k$ , lo que se realiza a través del cociente

$$\frac{\partial \hat{C}_{pk}(X_k)}{\partial X} = \begin{bmatrix} \frac{(\hat{C}_{pk}^{(+)} - \hat{C}_{pk}^{(-)})}{2c_k \Delta_{k1}} \\ \frac{(\hat{C}_{pk}^{(+)} - \hat{C}_{pk}^{(-)})}{2c_k \Delta_{k2}} \\ \vdots \\ \frac{(\hat{C}_{pk}^{(+)} - \hat{C}_{pk}^{(-)})}{2c_k \Delta_{kd}} \end{bmatrix} \quad (1.11)$$

*Paso m5.* Calcular

$$X_{k+1} = X_k - a_k \cdot \frac{\partial \hat{C}_{pk}(X_k)}{\partial X}; \quad (1.12)$$

*Paso m6.* Considerar  $X_{k+1} = X_k$  y regresar al paso m1.

El método propuesto, descrito anteriormente, aplica en procesos en los que se considera que el valor nominal de la variable de respuesta es lo mejor. Para los dos restantes casos: lo más grande o lo más pequeño es lo mejor, el método propuesto sigue operando de manera similar, pero para cada uno de estos dos casos la ecuación (1.8) es substituida, respectivamente, por las ecuaciones

$$\hat{C}_{pk} = \frac{\mu_k - LEI}{3\sigma_k} \quad (1.13)$$

$$\hat{C}_{pk} = \frac{LES - \mu_k}{3\sigma_k} \quad (1.14)$$

## 2. MATERIALES Y METODOS

La utilidad del método propuesto para el mejoramiento del  $C_{pk}$  para procesos productivos con restricciones en su espacio factible, se lleva a cabo en dos partes: en la primera, se presenta de la convergencia con probabilidad 1 del método propuesto y de acuerdo con Alaeddini y Klein, (2019), quienes establecen que el algoritmo que converge con probabilidad 1, es adecuado para su aplicación en diferentes tipos de procesos; en la segunda, se evalúa, mediante simulación, esta utilidad a través de probar que el valor esperado del  $C_{pk}$  en la iteración k+1, es mayor o igual al valor correspondiente en la iteración k.

### 2.1 Convergencia con Probabilidad 1 del Método Propuesto

la demostración de que el método propuesto converge con probabilidad 1 hacia un punto Karush-Kuhn-Tucker del modelo (3.1), para lo cual se prueba que la sucesión  $\{\beta_k''\} = \{h(\pi(X_k)) - h(X_k)\}$  definida en Sadegh (1997) es acotada con probabilidad 1 y converge a cero con la misma probabilidad, dada la región factible  $G$  definida en el método propuesto y agregando el que la función

$f$  satisface la condición de Lipschitz con constante  $U$ . El acotamiento y la convergencia de esta sucesión permiten concluir con la demostración de la convergencia con probabilidad 1 de la sucesión que se genera con el método propuesto.

La demostración de la convergencia con probabilidad 1 de la sucesión generada con el método de Sadegh (1997), es equivalente a probar la convergencia, con probabilidad 1, de la sucesión generada con el método propuesto, dado que el segundo método es una adaptación del primero. Por lo que en este trabajo se hará referencia a cualquiera de los dos como uno solo, y por lo tanto, los resultados obtenidos son válidos para ambos métodos.

Para llevar a cabo la prueba de la convergencia con probabilidad 1 de la sucesión generada por el método de Sadegh (1997) en la región factible  $G$ , sucesión que es definida como en el método propuesto, primeramente se transcriben algunos de los supuestos establecidos por este autor para que la sucesión de vectores generada por su método converge con probabilidad 1 hacia un punto Karush-Kuhn-Tucker, supuestos que posteriormente son considerados en el teorema 2.1.

El Método Propuesto es planteado para resolver el problema

$$\begin{aligned} & \text{maximizar } C_{pk}(X) \\ & X \in G \end{aligned} \quad (2.1)$$

donde  $G = \{X = (X_1, \dots, X_d) \in \mathbb{R}^d \mid -\infty < u_i \leq X_i \leq v_i < \infty; i = 1, \dots, d\}$ .

La determinación del punto Karush-Kuhn-Tucker del modelo (2.1), utilizando el método propuesto, se lleva a cabo a través de proyectar los vectores  $X_k$ , vectores que son generados de acuerdo a la fórmula recursiva presentada en la ecuación (1.12), sobre la región  $G_k = \{X = (X_1, \dots, X_d) \in \mathbb{R}^d \mid u_i + \Delta_{ki} c_k \leq X_i \leq v_i - \Delta_{ki} c_k; i = 1, \dots, d\} \subseteq G$ , con lo que se garantiza que en cualquier iteración siempre es posible estimar el vector gradiente de  $C_{pk}(X)$  con el método de perturbación simultánea.

Considere los siguientes supuestos que aparecen en la demostración de la convergencia con probabilidad 1 de la sucesión generada por el algoritmo de aproximaciones estocásticas con el método de perturbación simultánea (Spall, 1992).

2.1.1. Las sucesiones de números reales  $\{a_k\}$  y  $\{c_k\}$  convergen a cero,

$$\sum_{k=1}^{\infty} a_k = \infty \quad \text{y} \quad \sum_{k=1}^{\infty} a_k^2 / c_k^2 < \infty.$$

2.1.2. Para  $\alpha_0, \alpha_1, \alpha_2 > 0, \forall k \quad E(\varepsilon_k^{(\pm)^2}) \leq \alpha_0,$

$$E(f(X_k \pm \Delta_k c_k))^2 < \alpha_1 \quad \text{y} \quad E(\Delta_{ki}^2) \leq \alpha_2 \quad (i = 1, 2, \dots, d).$$

2.1.3. El vector  $X^*$  es una solución asintóticamente estable de la ecuación diferencial  $\frac{dX(t)}{dt} = -h(X)$

2.1.4. Sea  $D(X^*) = \{x_0 : \lim_{t \rightarrow \infty} x(t \mid x_0) = X^*\}$ , donde  $x(t/x_0)$  denota la solución de la ecuación diferencial del

supuesto 3.1.3, basada en la solución inicial  $x_0$ . Existe un conjunto compacto  $S \subseteq D(X^*)$ , tal que

$$P[(X_k \in S) \text{ i.o.}] = 1.$$

$$2.1.5. \quad P\left[\sup_k \|X_k\| < \infty\right] = 1$$

2.1.6. Considérense todas las  $k \geq K$  para algún  $K < \infty$ . Supóngase que para cada  $k$  la sucesión  $\{\Delta_{ki}\}$  ( $i = 1, 2, \dots, d$ ) es idénticamente independientemente y simétricamente distribuidas alrededor del cero con  $P[|\Delta_{ki}| < \alpha_0] = 1$ ,  $E|\Delta_{ki}^{-1}| \leq \alpha_1$ . Además, que para casi todos los  $X_k$  estimados ( $k \geq K$ ), ocurre que para casi todo  $X$  en una vecindad abierta de  $X_k$ , la cual no está en función de  $k$  o  $w$ ,  $\frac{\partial^3 f}{\partial X^3}(X)$  existe, es continua y el valor absoluto de cada una de sus componentes es menor que  $\alpha_2$ .

2.1.7. Para algunos valores reales positivos  $\alpha_0, \alpha_1, \alpha_2$  y  $\forall k$  ocurre que  $E(\varepsilon_k^{(\pm)^2}) \leq \alpha_0$ ,  $E(f(X_k + c_k \Delta_k))^2 < \alpha_1$  y  $E(\Delta_{ki})^2 \leq \alpha_2$ , para  $i = 1, 2, \dots, d$

2.1.8. La función  $h(X) = \frac{\partial f}{\partial X}$  es continua;

2.1.9. El conjunto  $G$  es la cerradura de su interior y es acotado. Además las  $q_r(X)$  son continuamente diferenciables. Para cada  $X \in \partial G$  ( $\partial G$  representa la frontera de  $G$ ), los gradientes de las restricciones activas son linealmente independientes;

En la proposición 1 de Sadegh (1997) establece que dados los supuestos, 2.1.1, 2.1.2, 2.1.4, 2.1.5, 2.1.6, 2.1.7., 2.1.8. y 2.1.9. la sucesión de vectores  $\{X_k\}$  generada por el método que este autor propone converge con probabilidad 1 hacia un punto karush-Kuhn-Tucker, si se logra demostrar que para la sucesión  $\{\beta_k\}$ , descrita en el párrafo siguiente, ocurre que  $P\left[\sup_k \beta_k < \infty\right] = 1$  y que  $P\left[\lim_{k \rightarrow \infty} \beta_k = 0\right] = 1$ .

Considere a  $Y_k(\pi(X_k))$  como la estimación del vector gradiente del  $C_{pk}(X_k)$ , en las condiciones  $X_k$  proyectadas en  $G_k$ , en la que cada una de sus componentes está dada por  $Y_k(\pi(X_k))_i = \left[ \frac{\hat{C}_{pk}^+(\pi(X_k)) - \hat{C}_{pk}^-(\pi(X_k))}{2c_k \Delta_{ki}} \right] (\pi(X_k))_i$  (representa la proyección del vector  $X_k$  en  $G_k$ ). Sea  $\{\beta_k\} = \{Y_k(\pi(X_k)) - h(X_k)\}$  ( $h(X_k) = \frac{\partial C_{pk}}{\partial X}(X_k)$ ), la sucesión de elementos que corresponden al error de estimación del vector gradiente en cada iteración.

Sadegh (1997), propone descomponer la estimación del vector gradiente  $\{\beta_k\}$  como  $\{\beta_k\} = \{\beta_k^I\} + \{\zeta_k\} + \{\beta_k^{II}\}$ , con:

$$\{\beta_k^I\} = \{E([Y_k(\pi(X_k))]/X_k) - h(\pi(X_k))\}; \quad (2.2)$$

$$\{\zeta_k\} = \{Y_k(\pi(X_k)) - E(Y_k(\pi(X_k)))/X_k\}; \quad (2.3)$$

$$\{\beta_k^{II}\} = \{h(\pi(X_k)) - h(X_k)\}. \quad (2.4)$$

por lo que demostrar la proposición 1 de Sadegh (1997) es equivalente a mostrar que las sucesiones definidas en las ecuaciones (2.2), (2.3) y (2.4) son acotadas y convergen hacia cero con probabilidad 1, lo que es presentado en el siguiente teorema

**Teorema 2.1** Sea  $f: \mathfrak{R}^d \rightarrow \mathfrak{R}$ ,  $h: \mathfrak{R}^d \rightarrow \mathfrak{R}^d$  ( $h$  representa el vector gradiente de  $f$ ). Si  $h$  es continua,  $P\left[\sup_k \|X_k\| < \infty\right] = 1$  ( $X_k \in \mathfrak{R}^d$ ) y  $f$  satisface la condición de Lipschitz con constante  $U$ , entonces  $P\left[\sup_k \|\beta_k^{II}\| < \infty\right] = 1$  y  $P\left[\lim_{k \rightarrow \infty} \beta_k^{II} = 0\right] = 1$ .

Demostración: la demostración de este teorema es consecuencia de los siguientes de los siguientes puntos:

(a) Sea  $\Omega$  el espacio muestral fundamental. Sea  $\omega \in \Omega$  fijo, tal que

$$\lambda(\omega) = \max \left[ \sup_k \|X_k(\omega)\|, \sup_k \|\pi(X_k(\omega))\| \right] < \infty, \text{ cuya existencia se garantiza por el hecho de que}$$

$$P \left[ \sup_k \|\pi(X_k)\| < \infty \right] = 1 \quad (\pi(X_k) \in G_k). \text{ Ahora si se considera la bola cerrada } E_\omega = \{X \in \mathfrak{R}^d : \|X\| \leq \lambda(\omega)\}$$

, se obtiene que  $\{X_k(\omega)\} \in E_\omega$  y  $\{\pi(X_k(\omega))\} \in E_\omega$ .

Por otra parte, puesto que  $h$  es continua y  $E_\omega$  es un conjunto compacto (cerrado y acotado), existe

$$M(\omega) \in \mathfrak{R} \text{ tal que } \|h(X_k(\omega))\| \leq M(\omega) \text{ y } \|h(\pi(X_k(\omega)))\| \leq M(\omega).$$

Finalmente, puesto que  $P[\|\lambda(\omega)\| < \infty] = 1$  se tiene que

$$P \left[ \sup_k \|h(X_k)\| < \infty \right] = 1 \text{ y } P \left[ \sup_k \|h(\pi(X_k))\| < \infty \right] = 1.$$

b) Ahora observe que si

$$P \left[ \lim_{k \rightarrow \infty} \|X_k - \pi(X_k)\| = 0 \right] = 1, \text{ entonces } P \left[ \lim_{k \rightarrow \infty} \|h(X_k) - h(\pi(X_k))\| = 0 \right] = 1 \quad (2.1.b)$$

Esto se debe a que  $h$  es uniformemente continua sobre conjuntos compactos.

Sean  $\omega$ ,  $\lambda(\omega)$  y  $E_\omega$  como en el inciso (a). Entonces por la continuidad uniforme de  $h$  sobre  $E_\omega$ , para cada número real  $\varepsilon > 0$  existe otro número real  $\delta > 0$ , tal que si

$$\|X - Y\| < \delta, \text{ entonces } \|h(X) - h(Y)\| < \varepsilon \quad \forall (X, Y \in E_\omega) \quad (2.2.b)$$

Por otra parte, de (2.1.b) se tiene que si

$$\|X_k(\omega) - \pi(X_k(\omega))\| < \delta \text{ para todo } k \geq N_\delta \text{ y para algún número natural } N_\delta, \text{ entonces (2.2.b)}$$

implica que  $\|h(X_k(\omega)) - h(\pi(X_k(\omega)))\| < \varepsilon$ , para todo  $k \geq N_\delta$ .

c) La fórmula recursiva para la sucesión  $\{X_k\}$  está dada por  $X_k - \pi(X_{k-1}) = a_k Y(\pi(X_{k-1}))$ . Cada componente de la estimación  $Y(\pi(X_k))$  del vector gradiente de  $f$  satisface la siguiente relación

$$\|Y(\pi(X_{k-1}))\|_i = \frac{\|f(\pi(X_{k-1}) + \Delta_{ki} c_k) - f(\pi(X_{k-1}) - \Delta_{ki} c_k)\|}{\|(\pi(X_{k-1}) + \Delta_{ki} c_k) - (\pi(X_{k-1}) - \Delta_{ki} c_k)\|} \leq U$$

para  $i = 1, 2, \dots, n$  de lo cual se concluye que  $\|Y(\pi(X_{k-1}))\|$  está acotada. Esto implica que

$$P \left[ \lim_{k \rightarrow \infty} \|X_k - \pi(X_{k-1})\| = 0 \right] = 1.$$

d)  $\|X_k - \pi(X_k)\| < \|X_k - \pi(X_{k-1})\|$ . Esta desigualdad se deriva del hecho de que  $\pi(X_k) \in G_k$  es el vector más cercano al vector  $X_k$ . De lo anterior y del inciso (c) se concluye que,  $P \left[ \lim_{k \rightarrow \infty} \|X_k - \pi(X_k)\| = 0 \right] = 1$ .

La demostración del teorema se sigue del hecho: combinando los resultados del inciso (b) y del inciso (d) se tiene que



$$P\left[\lim_{k \rightarrow \infty} \|h(\pi(X_k)) - h(X_k)\| = 0\right] = P\left[\lim_{k \rightarrow \infty} \beta_k = 0\right] = 1, \text{ lo que a su vez implica que}$$

$$P\left[\sup_k \|h(\pi(X_k)) - h(X_k)\| < \infty\right] = P\left[\sup_k \|\beta_k\| < \infty\right] = 1.$$

## 2.2 Evaluación del Método Propuesto a través de Simulación

La afirmación de que el método propuesto logra mejorar el valor del índice de capacidad del proceso  $C_{pk}$ , se verifica mediante simulación, en la cual el valor de este índice, en las condiciones de operación  $X \in G = \{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 4, 0 \leq x_2 \leq 4\}$ , es estimado a través de

$$\hat{C}_{pk} = f(X) = i[v(X), w(X)] \quad (2.5)$$

con  $v, w: \mathbb{R}^d \rightarrow \mathbb{R}$  e  $i: \mathbb{R}^2 \rightarrow \mathbb{R}$ .  $\forall w(X) \neq 0$

Las funciones  $v$  y  $w$  que son usadas para construir el modelo de simulación (función de prueba) son las funciones de dos variables usadas para la comparación numérica de técnicas experimentales (Sánchez, 1991), mismas que son listadas a continuación:

1) Función Valle Parabólico de Rosenbrock,  $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ ; 2) Función Cúbica de Rosenbrock,  $f(x_1, x_2) = 100(x_2 - x_1^3)^2 + (1 - x_1)^2$  y 3) Función Suma de Componentes Lineales,  $f(x_1, x_2) = x_1 + x_2$ . Las dos primeras funciones son diseñadas para dificultar la determinación de sus valores mínimos con la implantación de técnicas experimentales; mientras que en la tercera ocurre lo contrario, es decir, los valores deseados se obtienen con mayor facilidad, por lo que ésta es usada como una medida relativa de comparación de efectividad de los algoritmos, además, el uso de estas funciones permite estimar la utilidad de las técnicas experimentales en la optimización de los procesos productivos (Sánchez, 1991).

Las estimaciones de los valores que asumen los parámetros  $\mu$  y  $\sigma$  en las condiciones de operación  $X$  (iteración  $k$ ) y usadas en esta simulación se obtienen, respectivamente, al generar una muestra de diez valores de la variable de respuesta

$$y_i = v(X) + \text{normal}[w(X)] \quad (2.6)$$

Que corresponden a valores de variables aleatorias con distribución normal con media  $v(X)$  y desviación estándar  $w(X)$ , donde  $\text{normal}[w(X)]$  es un número aleatorio distribuido normalmente con media igual a cero y desviación estándar igual a  $w(X)$ , generado en el software *CSS:STATISTICA*.

### 2.2.1 Verificación de la Hipótesis de que el Método Propuesto Logra Mejorar el $C_{pk}$

El proceso de simulación se lleva a cabo con la implantación del método propuesto para cada una de las dos combinaciones de las funciones de prueba consideradas en la tabla 4.1. En cada iteración que se genera con la aplicación de este método, treinta y tres estimaciones del valor del  $C_{pk}$  en igual número de puntos iniciales  $X_0 = (X_1, X_2)$ , y la media aritmética de estos treinta y tres valores del  $\hat{C}_{pk}$ , es utilizada para probar la Hipótesis Nula: el valor esperado del  $C_{pk}$  en la iteración  $X_k$  es igual al valor esperado del  $C_{pk}$  en la iteración  $X_{k+1}$ , contra la Hipótesis Alternativa: el valor esperado del  $C_{pk}$  en la iteración  $X_k$  es mayor al valor esperado del  $C_{pk}$  en la iteración  $X_{k+1}$ , misma que verificada con un nivel de significancia del 5% (esto se lleva a cabo para las primeras diez iteraciones). Cada una de las componentes de  $X_0$  requeridas para la implantación del método propuesto, se

generan en forma independiente y aleatoria con una distribución uniforme a través del generador de números aleatorios del *software* CSS:STATISTICA.

Tabla 2.1 Funciones de Prueba Usadas en el Proceso Simulación

Función de Prueba	$v(X)$	$w(X)$
VRCL	Función Valle Parabólico de Rosenbrock	Función Suma de Componentes Lineales
VRCR	Función Valle Parabólico de Rosenbrock	Función Cúbica de Rosenbrock

### 3 RESULTADOS

#### 3.1 Análisis de Resultados Para la Función de Prueba VRCL

En la tabla 3.1 se muestran los valores de la media aritmética del  $\hat{C}_{pk}$  correspondientes a cada una de las primeras diez iteraciones, donde además se observa que el valor esperado del  $C_{pk}$  correspondiente a la iteración  $X_{k+1}$  no es menor al valor esperado de este parámetro en la iteración  $X_k$ , con lo que se verifica que este método mejora el valor esperado de este índice en las primeras iteraciones.

	Iteración								
	1	2	3	4	5	6	7	8	10
$\bar{C}_{pk}$	-208	-59.4	-37.3	-29.1	-26.8	-20.2	-20.5	-21.0	-13.2
$H_1 : C_{pk}^{k+1} < C_{pk}^k$		no	no	no	no	no	no	no	no

Tabla 3.1 Valores del  $\bar{C}_{pk}$  por Iteración para la Función de Prueba VRCL.

Si = significativa al 0.05%

No = no significativa al 0.05%.

En la figura 3.1 se muestra el comportamiento del valor del  $\bar{C}_{pk}$  que se obtiene para cada una de las cien iteraciones llevadas a cabo mediante simulación, en la que se puede observar un crecimiento de este indicador hasta estabilizarse en, relativamente, pocas iteraciones.

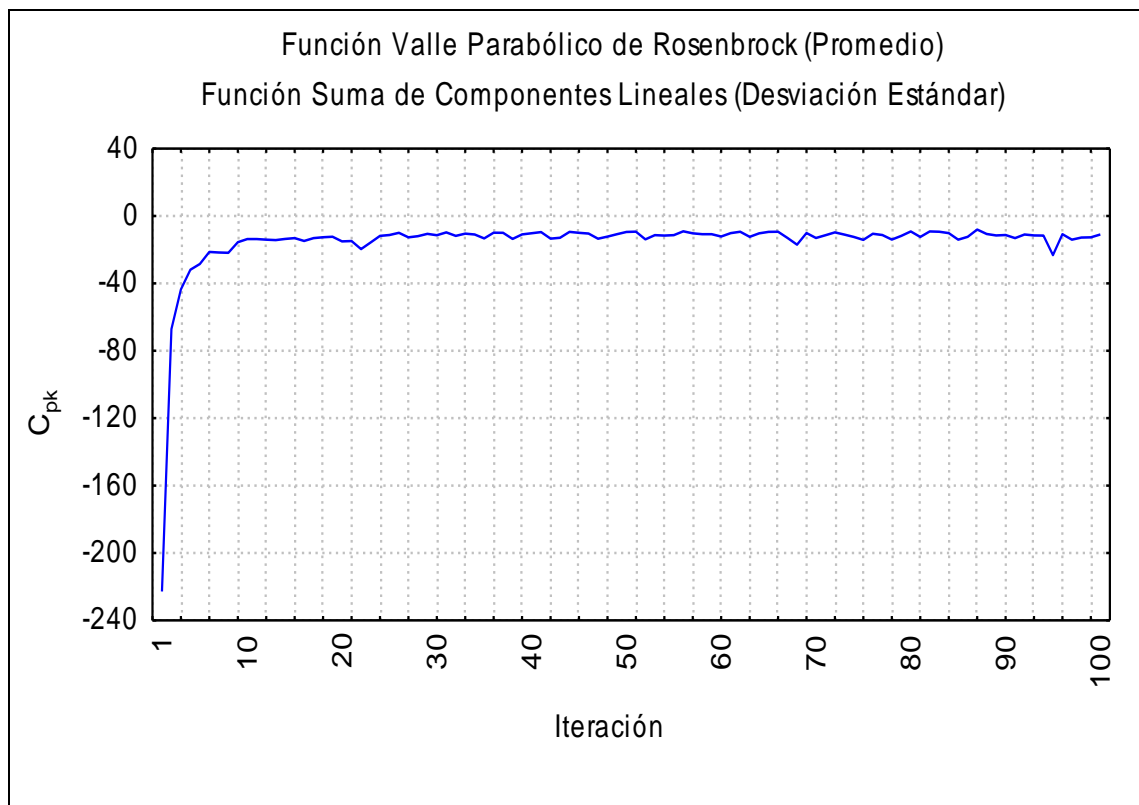


Figura 3.1 Gráfica de los  $\bar{C}_{pk}$  Generados para la Función de Prueba VRCL.

### 3.2 Análisis de Resultados Para la Función de Prueba VRCL

En la tabla 3.2 se muestran los valores de la media aritmética del  $\hat{C}_{pk}$  correspondientes a cada una de las primeras diez iteraciones, donde además se observa que el valor esperado del  $C_{pk}$  correspondiente a la iteración  $X_{k+1}$  no es menor al valor esperado de este parámetro en la iteración  $X_k$ , con lo que se verifica que este método mejora el valor esperado de este índice en las primeras iteraciones.

Iteración	1	2	3	4	5	6	7	8	10
$\bar{C}_{pk}$	144	363.9	279.9	230.1	548.4	333.7	209.8	268.4	236
$C_{pk}^{k+1} < C_{pk}^k$		no	no	no	no	no	no	no	no

Tabla 3.2 Valor del  $\bar{C}_{pk}$  por Iteración para la Función de Prueba VRCL.

Si = significativa al 0.05%

No = no significativa al 0.05%.

En la figura 3.2 se muestra el comportamiento del valor del  $\bar{C}_{pk}$  que se obtiene para cada una de las cien iteraciones llevadas a cabo mediante simulación, en la que se puede observar que ésta es de naturaleza oscilante con una tendencia creciente.

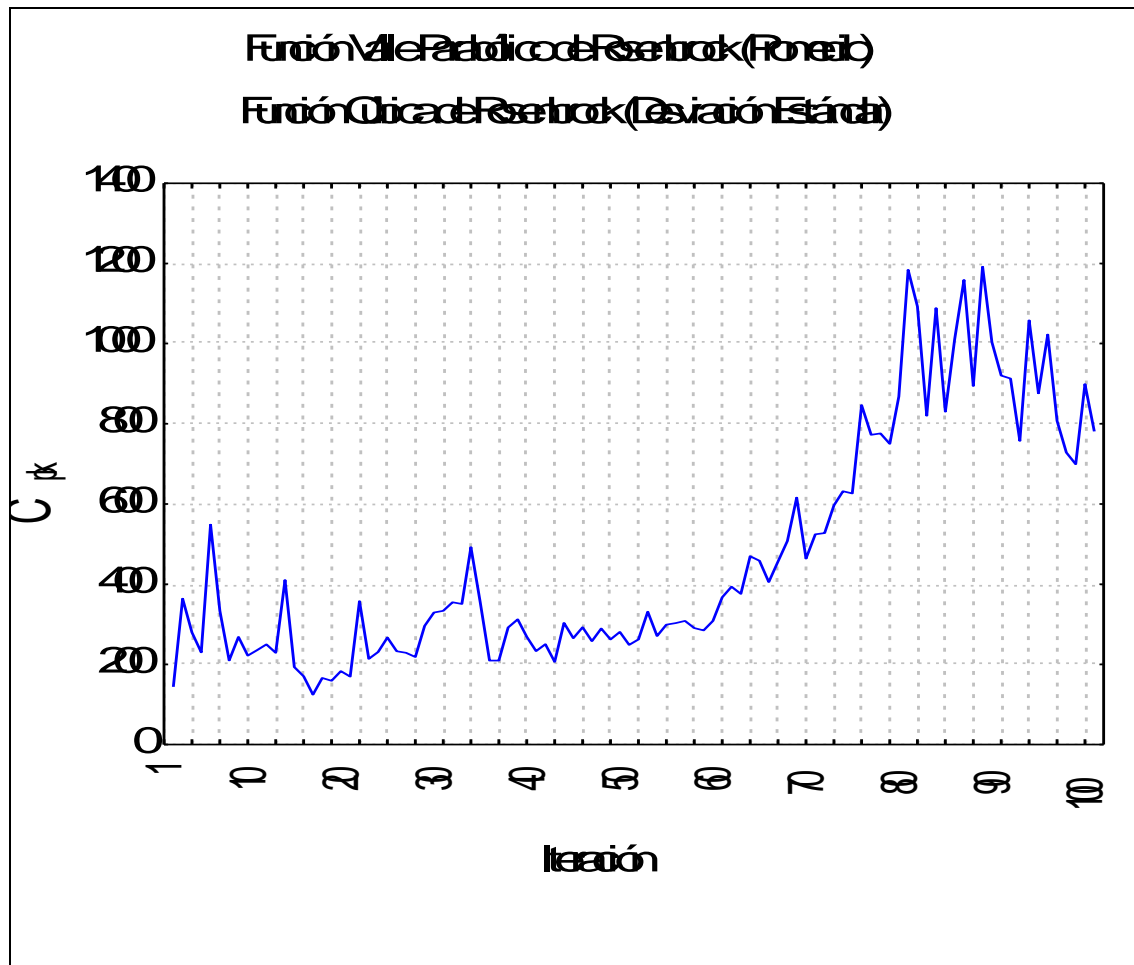


Figura 3.2 Gráfica de los  $\bar{C}_{pk}$  Generados para la Función de Prueba VRCR.

#### 4. DISCUSION

Los resultados muestran evidencia que el método propuesto para el mejoramiento del  $\bar{C}_{pk}$  para procesos con restricciones en su región factible, es adecuado para su aplicación en procesos donde se requiera mejorar el valor de este índice, por lo que se considera una herramienta estadística útil para la implementación de la filosofía six sigma. Lo anterior se deriva del hecho de se logra demostrar que este método para mejorar el  $\bar{C}_{pk}$ , converge con probabilidad uno; además de verificarse la hipótesis de que el valor esperado de este indicador en cada iteración es mayor o igual al valor esperado correspondiente de la iteración inmediata anterior.

El método aquí propuesto, es un método iterativo desarrollado para mejorar el  $C_{pk}$  de los procesos en los que es posible obtener valores de la función objetivo en condiciones de operación que pertenecen a la región factible, es decir el conjunto de los valores de operación que admite cada uno de los factores controlables que influyen en la función objetivo, de esta última, no se conoce la relación funcional, y los valores de ella que se pueden obtener en condiciones de operación específicas, están afectadas o contaminadas por el valor de una variable aleatoria. Y para este desarrollo, se utiliza los métodos de perturbación simultanea propuesto por (Spall, 1992), y optimización con restricciones vía SPSA, los cuales se adecuan para su aplicación en este indicador, lo que es llamado el método propuesto, mismo que de acuerdo al párrafo anterior es viable su aplicación en procesos que se busca mejorar el valor del  $C_{pk}$ . En la introducción se señala que desde que (Spall, 1997) presenta el SPSA, se han venido desarrollando variaciones de este método tanto como para estimar el vector gradiente y/o la matriz Hessiana de la función objetivo, que presentan ventajas en la tasa de convergencia con respecto al método original, pero con condiciones o restricciones mas fuertes para la función objetivo como las modificaciones propuestas por Spall, 1997 y 2000; Alaeddini y Klein, 2019; Benveniste, et.al, 1990; Borkar, et.al 2018; Chin, (1997); Li, & Mohammad 2013; Ramaswamy 2019; Popovic (2005); Wang 2018; Zeren D. et. al 2018. Se recomienda evaluar el desempeño de este método aquí propuesto contra los métodos que se desarrollan al aplicar los métodos por los autores anteriores.

#### Bibliografía.

Alaeddini A. y Klein D., (2019), Parallel Simultaneous Perturbation Optimization, Asia-Pacific Journal of operation Research, volumen 36, número 3, páginas 1-16.

Arcidiacono Gabriele, y Nuzzi Stefano (2017), A Review of the Fundamentals on Process Capability, Process Performance, and Process Sigma, and an Introduction to Process Sigma Split, International Journal of Applied Engineering Research ISSN 0973-4562. Volumen 12, Numero 14, páginas 4556-4570.

Benveniste A.(2019), Metiver M, y Priouret P (1990). Adaptive Algorithms and Stochastic Approximations, Editorial Springer-Verlag, Berlin Heidelberg.

Borkar Vivek S. (2018), Dwaracherla Vikranth R., Neeraja Sahasrabudhe, (2018), Gradient Estimation with Simultaneous Perturbation and Compressive Sensing (Gradient Estimation), Journal on Machine Learning Research 18, páginas 1-20.

Chaiba, S., Ayad, A., Ziani, D. *et al.* (2018). Eddy Current Probe Parameters Identification Using a Genetic Algorithm and Simultaneous Perturbation Stochastic Approximation. J Nondestruct Eval 37, 55 . <https://doi.org/10.1007/s10921-018-0506-0>

Chin, D.C.,  
(1997), Comparative Study of Stochastic Algorithms for System Optimization Based on Gradient Approximation, IEEE Transaction on Systems, Man, and Cybernetics-partb: Cybernetics, volumen 27, número 2, páginas 244-249.

Granichin O. and Amelina N., (2015). Simultaneous Perturbation Stochastic Approximation for Tracking Under Unknown but Bounded Disturbances, in IEEE Transactions on Automatic Control, vol. 60, no. 6, páginas 1653-1658.

Ishizuka S., y Kajiwaru I. (2014), Online adaptive PID control for MIMO systems using simultaneous perturbation stochastic approximation, Journal of Advanced mechanical Design Systems, and Manufacturing, volumen 9, número 2, páginas 1-16.

Kiefer, J. Y Wolfowitz, J., (1952), Stochastic Estimation of the Maximun of a Regression Function, Annals of Mathematical Statistic, volumen 23, páginas 462-466.

Kushner, H.J. y Clark, D.J. (1978), Stochastic Approximation for Constrained and Unconstrained Systems, Springer-Verrlag, New York, N.Y.

Kushner, H.J. y Yin, G. G. (2003), Stochastic Approximation and Recursive Algorithms and Applications, segunda edición, Editorial Springer. Nueva York, USA.

Li, L., Jafarpour, B. & Mohammad-Khaninezhad, M.R. (2013). A simultaneous perturbation stochastic approximation algorithm for coupled well placement and control optimization under geologic uncertainty. Comput Geosci 17, 167–188. <https://doi.org/10.1007/s10596-012-9323-1>

Montgomery, D. C., (2004), Diseño y Análisis de Experimentos, Editorial Limusa Wiley, segunda edición, México.

Ozguven Eren Erman, y Ozbay Kaan, 2008, Simultaneous Perturbation Stochastic Approximation Algorithm for Solving Stochastic Transportation Network Analysis Problems: Performance Evaluation. Transportation Research Record Journal of Transportation Research Board, volumen 2085, página 12-20.

Popovic D., Teel A. y JankovicAn (2005) M., An Improved SPSA Algorithm for Stochastic Optimization with Bound Constraints, 16<sup>th</sup> triennial World Congress, Elsevier IFAC Publications, páginas 47-52

Ramaswamy A., (2019), DSPG: Decentralized Simultaneous Perturbations Gradient Descent Scheme, arXiv: 1903.07050v2, páginas 1-16.

Robbins, H. Y S. Monro, (1951), A Stochastic Approximation Method, Annals of Mathematical Statisitica, volumen 22, páginas 400-407.

Ros R., Montero L., Barceló J. (2020) Investigating the quality of Spiess-like and SPSA approaches for dynamic OD matrix estimation, Journal Transportmetrica A: Transport Science.

Sadegh, P., (1997), Constrained Optimization via Stochastic Approximation with a Simultaneous Perturbation gradient Approximation, *Automatica*, volumen 33, páginas 889-892.

Sadegh, P. and Spall J. (1998), Optimal Sensor Configuration for Complex Systems, *Proceedings of the American Control Conference Philadelphia, Pennsylvania*, páginas 3575- 3579

Spall, J. C., (1992), Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation, *IEEE trans. Automática*, volumen 37, número 3, páginas 332-341.

Spall, (1997), A one-measurement form of simultaneous perturbation stochastic approximation, *Automatica*, Volumen 33, Issue 1, Páginas 109-112

Spall, J. C., (2000), Adaptive Stochastic Approximation by the Simultaneous Perturbation, *IEEE Transactions on Automatic Control*, volumen 45, número 10, páginas 1839-1853.

Sánchez, L. J., (1991) *A Method of Experimental Optimization*, Ph. D. dissertation, University of New Mexico at Las Cruces.

StatSoft, (1998), *CSS:STATISTICA Release 5.1*, serie SW8088444813M51, Estados Unidos.

Wang L., Zhu J. Y J. Spall J. C., (2018) Mixed Simultaneous Perturbation Stochastic Approximation for Gradient-Free Optimization with Noisy Measurements, *Annual American Control Conference (ACC)*, Milwaukee, WI, 2018, pp. 3774-3779.

Zeren D. Yenice, Niranjana Ahikari, Yong Kai Wong, Vural Aksakalli, Alev Taskin Gumus, Babak Abbasi, (2018), SPSA-FSR Simultaneous Perturbation Stochastic Approximation for Feature Selection and Ranking, print arXiv:1804.05589, páginas 1-56.