# DETERMINACIÓN DEL PERIODO DE CALENTAMIENTO EN SIMULACIONES DE ESTADO ESTABLE

**PEDRO PABLO BALLESTEROS**
Ingeniero Industrial, Esp.
Profesor Asistente
Universidad Tecnológica de Pereira
ppbs@utp.edu.co

**RUBEN DARIO MEDINA**
Estadístico Informático
Profesor Asistente
Universidad Católica Popular del Risaralda
medinar@ucpr.edu.co

**RESUMEN**
Este artículo muestra la forma cómo se puede determinar el tiempo que una simulación requiere para alcanzar un estado estable y cómo este tiempo está relacionado con los tiempos y cantidad de actividades que ocurren en el sistema. Este tiempo se denomina periodo de calentamiento. Existen modelos que alcanzan el estado estable en pocas horas, mientras que otros necesitan más tiempo. Conocer el tiempo de calentamiento en varias replicaciones facilita el inicio de recolección de información estadística de las variables de respuesta y elimina el sesgo ocasionado por observaciones consideradas en el estado transitorio del modelo.

**PALABRAS CLAVES:** Sistema**,** período de calentamiento, estado estable, replicación, variables de respuesta.

**ABSTRACT**
*This article shows how you can determine the time required by a simulation to reach a stable state and how this time is related with the activities times or quantity that happen in the system. This time is denominated the "warm-up period". Models that reach the stable state in few hours exist while others need more time. To know the time of warming in several replications helps the beginning of gathering statistical information of the answer variables and it eliminates the bias caused by observations considered in the transitory state of the model.*

**KEY WORDS**: *System, warm-up period, steady state, replication, answers variables.*

## 1. INTRODUCCIÓN

El tema que es motivo de este artículo, titulado "Periodo de Calentamiento para simulaciones de estado estable", nos ha permitido conocer alguna de las técnicas empleadas para determinar el tiempo que una simulación toma para alcanzar el estado estable. Se debe tener en cuenta que este tiempo está en función de los tiempos de las actividades y la cantidad de éstas que pueden realizarse en el sistema.

Por otra parte, con la finalidad de formular conclusiones útiles y correctas de los resultados de la simulación es conveniente y necesario que los datos sean significativos y confiables para hacer un adecuado análisis.

Para poder lograr este objetivo la mayor parte de los autores que han tratado este tema, recomiendan tener en cuenta los siguientes aspectos:

- Definir si el sistema es terminal o no terminal.
- Establecer el intervalo de confianza.
- Determinar el número de corrida de la simulación.
- Determinar el periodo de calentamiento (warp-up period)

Se espera que este tema permita la aprehensión de conocimientos en forma práctica y sencilla, y comprender que a medida que se analizan situaciones diferentes se requiere considerar más información para tomar decisiones adecuadas.

El desarrollo de este tema permite a través de la revisión bibliográfica especializada la aplicación de las técnicas para la determinación del periodo de calentamiento, donde se incorporan algunas distribuciones probabilidad, que facilita entender la importancia de que la simulación alcance la condición de estado estable, y en consecuencia, el desempeño del modelo sea independiente de las condiciones iniciales de arranque.

Una vez logrado lo anterior, se puede pensar cuándo reiniciar las estadísticas y la recolección de los datos de salida, obviando el posible sesgo dado por las condiciones iniciales.

## 2. DETERMINACIÓN DEL PERIODO DE CALENTAMIENTO

Las medidas de desempeño para las simulaciones terminantes dependen explícitamente del estado del sistema en el tiempo cero, por lo que debe tenerse

*Fecha de Recibo: 15 Agosto de 2003*
*Fecha de Aceptación: 21 Noviembre de 2003*

especial atención en la escogencia apropiada de las condiciones.

Antes de iniciar el análisis de los datos de salida, los responsables del proceso de simulación deben esforzarse para asegurar que éstas representen una estimación adecuada acerca de los valores verdaderos del sistema.

**2.1 Selección de la condición inicial**

Las medidas del funcionamiento para una simulación terminante depende del estado del sistema en el tiempo 0; así, debe tenerse mucho cuidado en elegir condiciones iniciales apropiadas.

Permítanos ilustrar este problema potencial por medio de un ejemplo.

Suponga que quisiéramos estimar el tiempo promedio esperado de atraso de todos los clientes que llegan y completan su retraso entre las 12 del mediodía y 1 de la tarde. (El período más ocupado) en un banco. Puesto que el banco estará congestionado probablemente al mediodía, comenzar la simulación entonces sin los clientes presentes (las condiciones iniciales usuales para una simulación de colas) causará distorsión en nuestra estimación del promedio esperado de atraso previsto por lo bajo. Ahora se discute un acercamiento heurístico a este problema, que parece ser usado muy ampliamente.

Para este acercamiento, se asume que el banco se abre a las 9 de la mañana sin los clientes presentes. Entonces, comienza la simulación a las 9 de la mañana sin clientes presentes y se realiza durante 4 horas.

Al estimar el tiempo promedio esperado de atraso, se utilizan retrasos solamente de los clientes que llegan y terminan su retraso entre el mediodía y la 1 de la tarde. La evolución de la simulación entre las 9 de la mañana y el medio día (el período del calentamiento, denominado $l$) determina las condiciones apropiadas para la simulación al mediodía.

Una desventaja de este acercamiento es que 3 horas de tiempo simulado no son utilizadas directamente en la estimación. Consecuentemente, uno pudo comprometerse y comenzar la simulación en una cierta otra hora, digamos 11 de la mañana, sin los clientes presentes. Sin embargo, no hay garantía que las condiciones en la simulación al mediodía sean representativas de las condiciones reales en el banco a dicha hora.

El acercamiento alternativo es recoger datos sobre el número de los clientes presentes en el banco al mediodía durante varios días.

Considere $\hat{p}_i$ la proporción de los días a la que $i$ clientes ($i$ = 0, 1....) están presentes al medio día. Entonces se simula el banco del mediodía hasta la 1 p.m. con el número de los clientes presentes al mediodía que es elegido aleatoriamente de la distribución $\hat{p}_i$. (Todos los clientes que se están atendiendo al mediodía pudieron ser considerados para comenzar justo sus servicios. Comenzar todos los servicios inmediatos al mediodía da lugar a una aproximación a la situación real en el banco, puesto que los clientes que están en el proceso de ser atendidos al mediodía habrían terminado parcialmente sus servicios. Sin embargo, el efecto de esta aproximación debe ser insignificante para una simulación de longitud una hora.)

Si se desea más de una simulación desde el mediodía hasta la 1 p.m., entonces una muestra diversa de $\hat{p}_i$ se dibuja para cada corrida.

**2.2 El análisis estadístico para los parámetros de estado estables**

Considere que $Y_1, Y_2\ldots.$ sea una salida de un proceso estocástico de una corrida simple de una simulación no terminante. Suponga que $P(Y_i \leq y) = F_i(y) \rightarrow F(y) = P(Y_i \leq y)$ con $i \rightarrow \alpha$, donde $Y$ es la variable de estado estable de interés con una distribución $F$. (Hemos suprimido en nuestra notación la dependencia de $F_i$ de las condiciones iniciales.)

Entonces $\phi$ es un parámetro de estado estable si es una característica de $Y$ como $E(Y)$, $P(Y \leq y)$o una división de Y. Una dificultad para estimar $\phi$ es que la función de distribución de $Y_i$ (para $i$ = 1, 2,…) es diferente de $F$, puesto que no será generalmente posible elegir $l$ para ser representativa del comportamiento de estado estable. Estas causas del estimador de $\phi$ basado en las observaciones $Y_1, Y_2, \ldots.., Y_m$ no son representativas. Por ejemplo, la media de la muestra será un estimador en polarización negativa de $\nu = E(Y)$ para todos los valores finitos de $m$. El problema que acabamos de describir se llama el *problema de la iniciación transiente* o el *problema de lanzamiento inicial* en la literatura de la simulación.

**2.3 El problema de la iniciación transiente**

Suponga que se desea estimar la media del estado estable que también se define generalmente por

$$\nu = \lim_{i \to \infty} E(Y)$$

Así, las medias transientes convergen a la media del estado estable. La consecuencia más seria del problema de la iniciación transiente es probablemente la $E[\overline{Y}(m,)] \neq \nu$ para cualquier $m$. La técnica sugerida más a menudo para ocuparse de este problema se llama *calentamiento del modelo* o la *cancelación inicial de datos*. La idea es suprimir algunos números de observaciones del principio de una corrida y usar solamente las observaciones restantes para estimar $\nu$. Por ejemplo, dado las observaciones $Y_1, Y_2\ldots..Y_m$ se sugiere a menudo utilizar:

$$\overline{Y}(m,l) = \frac{\sum_{i=l+1}^{m} Y_i}{m-l} \quad \left\{ \begin{array}{l} (1 \leq l \leq m-l) \text{ más bien que } \overline{Y}(m,l) \\ \text{como un estimador de } v. \end{array} \right.$$

En general, uno esperaría que $\overline{Y}(m,l)$ fuera menos sesgado que $\overline{Y}(m)$, puesto que las observaciones cerca del principio de la simulación pueden no ser muy representativas del estado estable debido a la opción de condiciones iniciales.

Por ejemplo, esto es verdad para el proceso $D_1, D_2, \ldots$ en el caso de una cola tipo *M/M/1* con $s = 0$, puesto que $E(Di)$ aumenta monótonamente a $d$, cuando $i \to \infty$.

La pregunta que debe responderse es ¿cómo elegir el período del calentamiento $l$?

Téngase en cuenta que se debe esperar hasta el periodo de calentamiento para comenzar a recolectar las estadísticas. De esta manera se elimina el sesgo debido a las observaciones tomadas durante el estado transitorio del modelo.

Se puede escoger $l$ (y $m$) tal que $E[\overline{Y}(m,l)] \approx v$. Si $l$ y $m$ son elegidos demasiado pequeños, entonces $E[\overline{Y}(m,l)]$ puede ser significativamente diferente de $v$. Por otra parte, si $l$ es escogido tan grande como sea necesario, entonces $\overline{Y}(m,l)$ tiene probablemente una variación innecesariamente grande. Existe un buen número de métodos sugeridos en la literatura para elegir $l$. Sin embargo, Gafarian. Ancker, y Morisaku (1978) encontraron que ninguno de los métodos disponibles en aquella época se realizaron bien en la práctica.

Kelton y Law (1983) desarrollaron un algoritmo para elegir $l$ (y $m$) que trabajó bien {es decir, $E[\overline{Y}(m,l)] \approx v$} para una variedad amplia de modelos estocásticos; sin embargo, una limitación teórica del procedimiento es que hace básicamente la suposición que $E(Y_i)$ es una función monótona de $i$.

La técnica más simple y más general para determinar $l$ es procedimiento gráfico debido a Welch (1981, 1983). Su meta específica es determinar un tiempo indizado $l$ tal que $E(Yi) \approx v$ para $i > l$, donde $l$ es el período del calentamiento. En general, es muy difícil determinar $l$ en una sola réplica debido a la variabilidad inherente del proceso $Y_1, Y_2, \ldots$ (Ver figura 4.).

Consecuentemente, el procedimiento de Welch se basa en la realización de $n$ réplicas independientes de la simulación y el empleo de las siguientes cuatro etapas:

1. Hacer $n$ réplicas de la simulación ($n \geq 5$), cada uno de la longitud $m$ (donde $m$ es grande). Deje $Y_{ji}$ como la observación *i-ésima* de la réplica *j-ésima* ($j = 1, 2 \ldots n$; $i = 1, 2, \ldots m$), según lo demostrado en Figura 1.

2. Considerar $\overline{Y}_i = \sum_{j=1}^{n} Y_{ij} / n$ para $i = 1, 2 \ldots m$ (Figura 1). El promedio del proceso $\overline{Y}_1, \overline{Y}_2, \ldots$ tiene media $E(\overline{Y}_i) = E(Y_i)$ y varianza $Var(\overline{Y}_i) = Var(Y_i)/n$.

Réplica

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ① | $Y_{11}$ | $Y_{12}$ | $Y_{13}$ | $Y_{14}$ | $Y_{1,m-2}$ | $Y_{1,m-1}$ | $Y_{1,m}$ |
| ② | $Y_{21}$ | $Y_{22}$ | $Y_{23}$ | $Y_{24}$ | $Y_{2,m-2}$ | $Y_{2,m-1}$ | $Y_{2,m}$ |
| ③ | | | | | | | |
| . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . |
| N | $Y_{n1}$ | $Y_{n2}$ | $Y_{n3}$ | $Y_{n4}$ | $Y_{n,m-2}$ | $Y_{n,m-1}$ | $Y_{n,m}$ |
| Promedio de procesos | $\overline{Y}_1$ | $\overline{Y}_2$ | $\overline{Y}_3$ | $\overline{Y}_4$ | $\overline{Y}_{m-2}$ | $\overline{Y}_{m-1}$ | $\overline{Y}_m$ |
| | ↓ | ↓ | ↓ | | | ↓ | |
| Promedio móvil w=1 | $\overline{Y}_{1(1)}$ | $\overline{Y}_{2(1)}$ | $\overline{Y}_{3(1)}$ | | $\overline{Y}_{m-1(1)}$ | | |

Figura 1. Proceso promedio y promedio móvil con $w = 1$ basado en $n$ replicas de longitud $m$.

Así, el proceso promedio tiene la misma curva de la media transiente del proceso original, pero su diagrama tiene solamente ($1/n$)avo de la varianza.

3. Para suavizar las oscilaciones de alta frecuencia en $\overline{Y}_1, \overline{Y}_2, \ldots$, (excluyendo las oscilaciones de baja frecuencia o la longitud de la tendencia de la corrida de interés), se define más lejos el promedio móvil $\overline{Y}_i(w), \ldots$ (donde $w$ es la ventana y es un número entero positivo tal que $w \leq m/4$ como sigue:

$$\overline{Y}_i(w) = \begin{cases} \dfrac{\sum_{s=-w}^{w} \overline{Y}_{i+s}}{2w+1} & \text{si } i = w+1, \ldots\ldots m-w \\[2ex] \dfrac{\sum_{s=-(i-1)}^{i-1} \overline{Y}_{i+s}}{2i-1} & \text{si } i = 1 \ldots\ldots w \end{cases}$$

Así, si $i$ no está también cerca de ser principio de las réplicas, entonces $\overline{Y}_i(w)$ es justo el promedio simple de $2w + 1$ como observaciones del promedio del proceso centrado en la observación $i$ (véase Figura 1). Se llama un promedio móvil puesto que 1 se mueve a través del tiempo.

4. Trazar $\overline{Y}_i(w)$ para $i = 1, 2 \ldots, m-w$ y elija $l$ para que el valor de $i$ vaya más allá de lo que $\overline{Y}_1(w)$, $\overline{Y}_2(w)$

parece converger. Vea Welch (1983, p. 292) para una ayuda en la determinación de convergencia.

El ejemplo siguiente ilustra el cálculo del promedio móvil.

### 2.3 Aplicación del procedimiento de Welch.

La práctica consiste en estudiar los tiempos de ciclo para un sistema compuesto por una fresadora, un torno, un área de inspección, una cola de entrada, una cola de inspección, como puede observarse en la Figura 3.

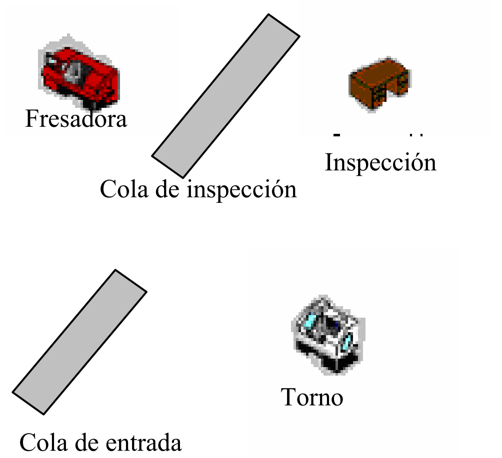


Figura 3. Distribución de máquinas en una pequeña empresa.

En este ejemplo los tiempos de operación no son constantes y siguen las distribuciones normales teóricas, que a continuación se muestran (Tabla 1):

| Entidad | Estación | Operación |
|---|---|---|
| Parte | Fresadora | N(3, 0.1) |
| Parte | Torno | N(4, 0.1) |
| Parte | Inspección | N(1, 0.1) |

Tabla 1. Tiempos de operación y distribuciones teóricas

El tiempo de ciclo está definido como el tiempo durante el cual la entidad estuvo en el sistema.

La simulación se hizo durante 10 horas con cinco réplicas. Los resultados relacionados con el comportamiento del tiempo de ciclo se muestran a continuación: (ver Tabla 2)

Tabla 2. Comportamiento del tiempo de ciclo para 35 periodos

| Periodo | Promedio del perio-do | Promedio móvil de Tiempo de ciclo durante el i -ésimo periodo $\bar{y}_i(w)$ | | Periodo | Promedio del perio-do | Promedio móvil de Tiempo de ciclo durante el i -ésimo periodo $\bar{y}_i(w)$ | |
|---|---|---|---|---|---|---|---|
| | $\bar{y}_i = \bar{x}(5)$ | w=5 | w=10 | | $\bar{y}_i = \bar{x}(5)$ | w=5 | w=10 |
| 1 | 4,30 | 4,30 | 4,30 | 18 | 4,58 | 4,40 | 4,39 |
| 2 | 4,24 | 4,29 | 4,29 | 19 | 4,39 | 4,40 | 4,39 |
| 3 | 4,33 | 4,34 | 4,34 | 20 | 4,38 | 4,40 | 4,39 |
| 4 | 4,44 | 4,35 | 4,35 | 21 | 4,33 | 4,40 | 4,39 |
| 5 | 4,39 | 4,36 | 4,36 | 22 | 4,41 | 4,41 | 4,39 |
| 6 | 4,39 | 4,37 | 4,37 | 23 | 4,40 | 4,40 | 4,39 |
| 7 | 4,39 | 4,37 | 4,37 | 24 | 4,39 | 4,38 | 4,40 |
| 8 | 4,41 | 4,38 | 4,37 | 25 | 4,42 | 4,38 | 4,40 |
| 9 | 4,39 | 4,39 | 4,37 | 26 | 4,35 | 4,38 | |
| 10 | 4,38 | 4,38 | 4,38 | 27 | 4,41 | 4,39 | |
| 11 | 4,36 | 4,38 | 4,38 | 28 | 4,34 | 4,38 | |
| 12 | 4,39 | 4,38 | 4,38 | 29 | 4,38 | 4,40 | |
| 13 | 4,37 | 4,40 | 4,39 | 30 | 4,36 | 4,39 | |
| 14 | 4,37 | 4,40 | 4,40 | 31 | 4,41 | | |
| 15 | 4,35 | 4,40 | 4,39 | 32 | 4,38 | | |
| 16 | 4,38 | 4,39 | 4,39 | 33 | 4,38 | | |
| 17 | 4,41 | 4,40 | 4,39 | 34 | 4,59 | | |
| | | | | 35 | 4,31 | | |

A continuación se pueden observar en la figura 4 el promedio del tiempo de ciclo y en las figuras 5 y 6 el promedio móvil del tiempos de ciclo para w=5 y w=10.

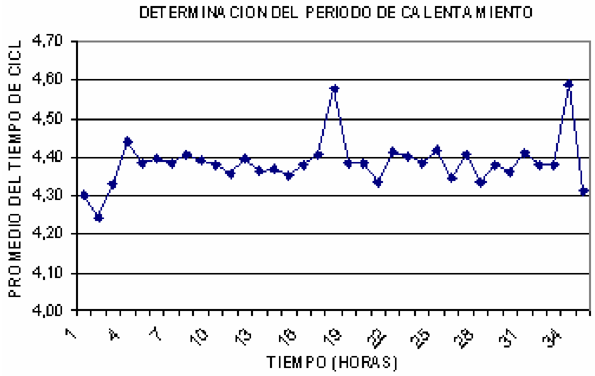


Figura 4. Comportamiento del tiempo promedio de ciclo.

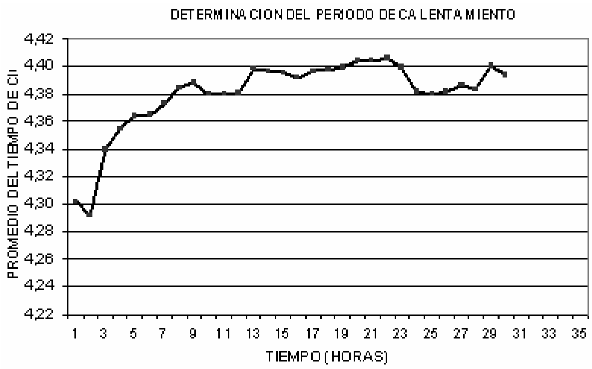


Figura 5. Comportamiento del promedio móvil del tiempo de ciclo para w = 5.

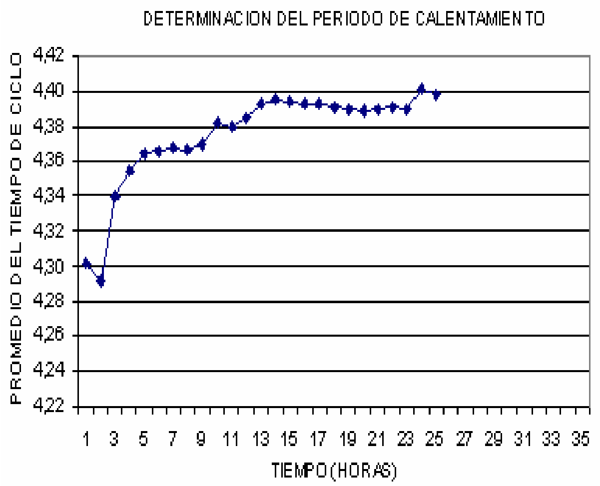


Figura 6. Comportamiento del promedio móvil del tiempo de ciclo para $w = 10$.

Se observa cómo en la figura 6 la simulación tiende a estabilizarse. Si se tomará w = 18, el comportamiento del promedio móvil se suavizará aún más, sin embargo se puede notar en la figura 6 que el período de calentamiento se sitúa aproximadamente en 14 horas; aunque se presenta una pequeña variación en los últimos dos períodos, recordemos que lo que nos interesa, es el tiempo a partir del cuál se estabiliza el sistema.

### 2.4 Aplicación del softaware Promodel

El software Promodel tiene dos opciones para fijar el tiempo de calentamiento:

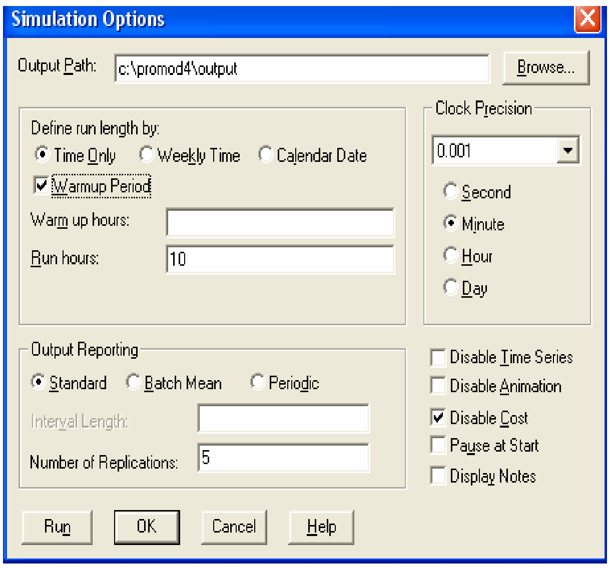


Figura 7. Caja de dialogo de opciones de simulación desde promodel.

Para llegar a esta ventana, se sigue la ruta Promodel→ Simulation→Options y estando allí, como se observa en la figura 7 se establecen los parámetros correspondientes.

La otra opción es emplear el término Warm up simultáneamente con el tiempo de calentamiento con la expresión If clock = "número", then Warm up.

## 3. CONCLUSIONES Y RECOMENDACIONES

Los resultados de la simulación presentada en este artículo, destacan la importancia de que ésta alcance un estado estable, que al mismo tiempo es de mucho interés para analizar su comportamiento, debido a que las variables de respuesta del sistema muestran lo que los especialistas en este tema han denominado "regularidad estadística"

Se ha comprobado que el tiempo que se requiere para lograr el estado estable, está en función de los tiempos de las distintas actividades y de la cantidad que de éstas se ejecutan en el sistema. Cuando se considera el periodo de calentamiento se observan dos ventajas: por un lado se elimina el sesgo (bias) generado por las observaciones consideradas durante el periodo de transición y por otro lado, este tiempo nos indica cuándo se debe comenzar la recolección de la información estadística de la simulación.

Sería de mucho interés científico que se intentara profundizar en la obtención de una solución óptima a este problema, pues los aportes de los especialistas han mostrado sólo aproximaciones a través del uso de determinadas heurísticas. En este caso, quizá uno de las mayores dificultades que presentan los modelos es que algunos pueden llegar al estado estable en cuestión de tiempos cortos, mientras que otros requieren de tiempos más largos.

Se recomienda, para determinar el periodo de calentamiento correr una simulación preliminar del sistema con tres a cinco réplicas y observar en qué momento éste alcanza la estabilidad estadística. Este método aunque es fácil y directo no necesariamente es el más confiable. Igualmente se debe tener en cuenta que la longitud de cada replica sea lo suficientemente larga como para facilitar la presencia de eventos "atípicos o raros" entre los que pueden citarse interrupciones bruscas del sistema como paradas por averías en las máquinas o tiempos ocioso por paradas en las estaciones de servicio, con frecuencia de por lo menos dos o tres veces durante el proceso de simulación.

El método exige el monitoreo de una o más variables de respuesta por determinados periodos durante el tiempo de simulación, por ejemplo el número promedio de clientes en una línea de espera.

Otro de los riesgos que se corre en la determinación del periodo de calentamiento está en su subestimación, porque afecta la recolección de estadísticas de un sistema que no ha alcanzado el estado estable, originando sesgo en los resultados.

## 4. BIBLIOGRAFÍA

[1] LAW, Averill y W. David Kelton. *Simulation Modeling & Analysis*, 3rd edition, páginas 518 – 526 y 539 - 541, Mc Graw Hill, New York, 2000.

[2] SOTO M., José A. *Fundamentos Teóricos de Simulación Discreta*. Universidad Tecnologota de Pereira. 2003.