

UMA REVISÃO SOBRE MÉTODOS PARA A DETECÇÃO DE PERÍODOS DE AQUECIMENTO EM MODELOS DE SIMULAÇÃO A EVENTOS DISCRETOS

Hélida Mara Gomes Norato¹, Anderson Ribeiro Duarte²

Resumo: *A seleção de um período de aquecimento adequado para uma simulação de eventos discretos ainda é uma escolha de difícil determinação. Alguns métodos já foram desenvolvidos e serão brevemente revistos aqui. Ainda não existem critérios definitivos para a escolha de um método específico dentre os existentes. Buscamos discutir alguns dos métodos existentes e sua adaptação à estrutura de um modelo experimental em estudo. Analisaremos o desempenho dos métodos expostos através de um procedimento avaliativo para comparação. Os métodos serão aplicados para as saídas geradas por um modelo simples de trabalho com diferentes níveis de utilização. Algumas métricas serão utilizadas para averiguação da eficiência de cada um dos métodos em comparação.*

Palavras-chave: Simulação; Estado Transiente; Estado Permanente; Aquecimento.

Introdução

Em estudos que utilizam técnicas de simulação, exige-se uma alta demanda de tempo computacional, bem como elevados esforços no desenvolvimento e programação dos modelos. Por outro lado, em geral, é dedicado apenas um pequeno esforço para a avaliação de dados de saída de simulação de forma a promover uma análise adequada dos resultados. Uma rodada de simulação de sistemas não terminais (aqueles que funcionam intermitentemente) é composta por um tempo de aquecimento necessário à estabilização do sistema, seguida de diversas replicações, cada uma delas, fornecendo dados para a realização de estatísticas referentes ao comportamento do modelo simulado. Em procedimentos mais rústicos, é comum se deparar com uma simulação simplista, em que o tempo de aquecimento e o número de replicações são escolhidos de forma arbitrária, e em seguida, as estimativas resultantes do procedimento de simulação são observadas como as características do modelo em estudo. Considerando que os modelos de simulação a eventos discretos são conduzidos ao longo do tempo e através de amostras aleatórias de distribuições de probabilidade, podemos observar que as estimativas são apenas realizações particulares de variáveis aleatórias que podem ter muita variabilidade.

Para duas replicações distintas em uma simulação, é possível observar diferentes características entre elas, portanto, ao menos uma delas, não deve estar representando a realidade a qual o modelo tenta reproduzir. Assim poderia haver uma probabilidade significativa de fazer inferências errôneas sobre o sistema em estudo. Um impedimento para a obtenção de estimativas

¹Departamento de Engenharia de Produção, Escola de Engenharia, UFMG, hnorato@gmail.com

²Departamento de Matemática, ICEB, UFOP, anderson@iceb.ufop.br

precisas dos parâmetros de um sistema modelado através de simulação é o tempo computacional necessário para que o sistema simulado se estabilize. Isto é necessário para tornar possível coletar a quantidade necessária de dados da simulação para estudos inferenciais. Esta dificuldade ocorre muitas vezes na simulação em problemas de médio e grande porte. Para promover melhorias nos procedimentos de coleta de dados de simulações visando estudos inferenciais, é de extrema importância a adoção de métodos para seleção do período de aquecimento (*warm-up*). Por período de aquecimento denominamos a detecção do período transiente inicial em que os dados da simulação apresentam elevada variabilidade quando comparados aos dados do período permanente. Denominamos esse período permanente por estado estacionário. Neste período os dados são capazes de produzir estimativas mais estáveis para os parâmetros de interesse.

O problema de detecção do período de aquecimento adequado tem sido discutido na literatura de simulação há muito tempo. Uma escolha adequada é importante para evitar a presença de vício nas estimativas de uma simulação quando em seu estado estacionário. Muitas abordagens sobre este assunto sugerem a execução dos experimentos de simulação para um dado período de aquecimento e a exclusão destes dados iniciais. Entretanto uma questão óbvia surge: “Como identificar o período adequado de aquecimento?” Este fato reforça a importância do tratamento adequado do período transiente, para que este período não influencie sobremaneira as estimativas dos parâmetros de interesse no procedimento de simulação, quando este atingir o seu estado estacionário. Estes métodos possuem características que podem ser adequadas a diversos tipos de modelos de simulação, considerando as especificidades associadas a cada modelo.

Existem diversas pesquisas realizadas para selecionar o período adequado de aquecimento. O principal foco é proporcionar condições apropriadas para a utilização de modelos de simulação. O principal objetivo é que estes modelos forneçam dados que efetivamente sejam capazes de refletir as verdadeiras características associadas ao sistema que se busca reproduzir através do modelo de simulação. Pode-se encontrar na literatura, associada à área de simulação, uma variedade de métodos para a detecção do período adequado de aquecimento. Este fato reforça a importância do tratamento adequado do período transiente, para que este período não influencie sobremaneira as estimativas dos parâmetros de interesse no procedimento de simulação quando este atingir o seu estado estacionário. Estes métodos possuem características que podem ser adequadas a diversos tipos de modelos de simulação, considerando as especificidades associadas a cada modelo.

Em geral, duas abordagens são utilizadas: a exclusão de dados e inicialização inteligente [11]. Na primeira abordagem, o modelo é executado por um período de aquecimento até atingir um estado estacionário e, antes do qual, esses os dados são eliminados. A maior dificuldade presente na utilização desta abordagem está em escolher a duração do período de aquecimento. Se está escolha for subestimada, haverá algum vício na estimação para as variáveis de interesse no procedimento de simulação. Por outro lado, se a escolha for superestimada, alguns dados relevantes para o procedimento de estimação serão excluídos fazendo com que os tempos de se simulação se tornem elevados de forma desnecessária. Na segunda abordagem, o modelo é colocado em uma condição “realista” no início da execução. Aqui, o desafio é determinar o que constitui um estado realista para se iniciar o procedimento de simulação. Em princípio, essas abordagens levam ao mesmo objetivo: colocar o modelo em uma condição realista a partir do momento em que os dados de saída serão coletados. Este trabalho se concentra na primeira abordagem, ou seja, com o objetivo de determinar um período adequado de aquecimento para os modelos de simulação.

Os métodos para detecção do período de aquecimento em modelos de simulação podem ser classificados em cinco grupos: Gráficos, Heurísticos, Estatísticos, Testes de vício de inicialização e Híbridos. A apresentação e a descrição de alguns destes métodos, de acordo com sua classificação, são importantes instrumentos para a seleção adequada da metodologia para detecção do período de aquecimento em modelos de simulação.

Métodos Gráficos

São métodos que dependem de uma análise visual de séries temporais obtidas das saídas de simulação. Em geral, se baseiam em particionamento de dados em grupos para cálculo de médias móveis. Dentre estes métodos, podemos destacar as propostas de [14] e [19] que se baseiam em avaliações em séries temporais simples. Existem ainda métodos baseados em gráficos de médias cumulativas que podem ser obtidos em [1, 7, 11]. Tratam-se de métodos que tem por vantagem a sua simplicidade e a “quase” ausência de fortes suposições acerca dos dados de saída da simulação. Por outro lado, existe uma forte dependência da avaliação do analista (decisão humana), fato este que pode levar a diferentes decisões de acordo com a experiência do analista.

Heurísticas

Trata-se na verdade de estabelecer um conjunto de regras para a determinação da duração do período de aquecimento. Entre os métodos heurísticos, é possível citar a regra de Conway [2], a regra de Conway modificada [5], a regra do estimador de autocorrelação [3], a regra do cruzamento de médias [4], as regras de erros marginais (MSER e MSER-5) [20]. O uso de regras específicas tem por vantagem o fato de remover a subjetividade dos métodos gráficos. Por outro lado, isto pode significar que alguns padrões importantes podem não ser identificados, especialmente se uma representação gráfica dos dados não for utilizada em associação à regra em uso.

Métodos Estatísticos

São métodos que se baseiam em princípios estatísticos para o procedimento de estimação do período de aquecimento. É possível citar o método proposto em [9] que executa um procedimento de regressão dos dados avaliados na sequência inversa de tempo, buscando até que ponto a inclinação estimada pelo procedimento de regressão ainda é nula. O teste de aleatorização proposto por [21] avalia a diferença entre duas médias para os dados divididos em dois subgrupos ordenados ao longo do tempo de execução da simulação. Tratam-se de métodos mais complexos que as abordagens anteriores, dada exigência de cálculos extensivos, bem como a validação de algumas suposições iniciais.

Testes de vício de inicialização

Não se tratam de testes para determinação do período adequado de aquecimento, mas sim testes para verificação da presença ou não de vício que seja associado à inicialização na série de dados em estudo. Como estratégia para a detecção do período adequado de aquecimento devem ser utilizados em conjunto com algum dos métodos descritos anteriormente. Dentre possíveis formatos, pode-se mencionar o teste do máximo [15], o teste de Schruben modificado [15], o teste ótimo [16], o teste do posto [18]. Uma família de testes desta natureza considerando médias por subgrupos de dados é apresentado em [6]. Uma grande vantagem nestes métodos é a possibilidade do analista avaliar se o vício de inicialização foi completamente removido. São métodos cuja complexidade dos cálculos envolvidos é relativamente baixa, entretanto, computacionalmente instáveis e, em geral, se faz necessária uma estimativa da variância dos dados.

Métodos Híbridos

Trata-se de um conjunto de métodos que empregam os testes para vício de inicialização associados à testes gráficos ou heurísticos. Dois dos principais métodos são o método sequencial de [13] e o método para obtenção de ponto de truncamento através de uma escala invariante, proposto em [8].

Objetivo

Este trabalho tem por objetivo discutir algumas técnicas já existentes para a detecção de períodos de aquecimento; estabelecer métricas de comparação entre técnicas através de um modelo experimental de simulação para comparar técnicas.

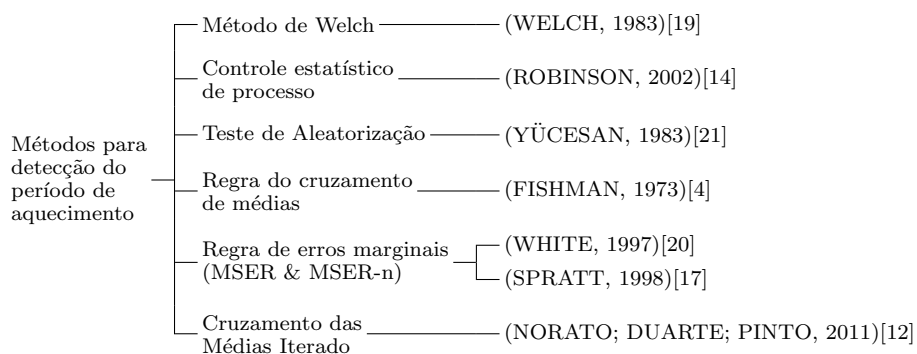
Diversos testes já foram realizados sobre os métodos anteriormente descritos obtendo-se estimativas para o período de aquecimento, com diferentes graus de sucesso. Em muitos casos, isto se deve a uma dependência da natureza dos dados analisados. Não se pode afirmar sobre a existência de um método que possa ser recomendado acima de quaisquer outros. De fato, a literatura sugere a não existência de métodos de uso comum. Apesar da grande variedade de abordagens que têm sido propostas ao longo dos últimos anos, existe uma necessidade tanto de testar estas abordagens mais profundamente como de elaborar novos métodos que poderiam ser adotados amplamente por usuários da simulação. Com o interesse de estabelecer uma estratégia comparativa e desenvolver novas metodologias, este estudo, está focado em avaliar métodos de três grupos: Gráficos, Estatísticos e Heurísticos.

Com o intuito comparativo, testes podem ser realizados sobre os métodos de mais ampla utilização, obtendo-se estimativas para o período de aquecimento, com diferentes graus de sucesso. Em muitos casos, este grau de sucesso, se deve a uma dependência da natureza dos dados analisados. Não se pode afirmar sobre a existência de um método que possa ser recomendado acima de quaisquer outros. De fato, a literatura sugere a não existência de métodos de uso comum. Apesar da grande variedade de abordagens que têm sido propostas ao longo dos últimos anos, existe uma necessidade tanto de testar estas abordagens mais profundamente, como de elaborar novos métodos que poderiam ser adotados amplamente por usuários da simulação.

Neste trabalho, serão considerados alguns métodos de grande utilização para a adequada detecção do período de aquecimento [10] e também o recentemente proposto método do cruzamento de médias iterado [12]. Será apresentada uma análise comparativa entre os métodos utilizando um modelo experimental e algumas métricas para avaliação do desempenho dos métodos para detecção de período de aquecimento.

Metodologia

Para uma melhor compreensão da avaliação comparativa dentre as técnicas para detecção do período de aquecimento, uma descrição sucinta dos seis métodos aqui discutidos será apresentada. Serão avaliados os seguintes métodos:



Método de Welch

O Método de Welch [19] é uma técnica bastante simples para a determinação do adequado período de aquecimento. Podemos descrever o método através dos seguintes passos:

1. Execute n replicações do procedimento de simulação, cada uma destas replicações terá tamanho m . Seja Y_{ij} a i -ésima observação da variável de interesse na j -ésima replicação, neste caso $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$;
2. Calcule as médias de cada uma das observações entre as replicações, estas médias serão denotadas por \bar{Y}_i , em que: $\bar{Y}_i = \frac{\sum_{j=1}^n Y_{ij}}{n}$ para $i = 1, 2, \dots, m$;
3. Defina a média móvel $\bar{Y}_i(w)$ para suavizar as oscilações de alta frequência em $\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_m$, neste caso, w será a medida de janela que é um inteiro positivo, a proposta utiliza $w \leq \frac{m}{4}$.
A média móvel $\bar{Y}_i(w)$ será então: $\bar{Y}_i(w) = \frac{\sum_{s=-w}^w \bar{Y}_{i+s}}{2w+1}$ para $i = w+1, w+2, \dots, m-w$;
4. Plote os valores $\bar{Y}_i(w)$ para $i = w+1, w+2, \dots, m-w$ e escolha L (comprimento do período de aquecimento) como o valor i no qual $\bar{Y}_i(w)$ parece convergir. Maiores detalhes sobre a convergência podem ser obtidos em [19].

Controle Estatístico de Processo SPC

O método baseado no Controle Estatístico de Processo [14] consiste na avaliação da saída de simulação através de Cartas de Controle para a média, podendo ser descrito em quatro etapas:

1. Defina Y_{ij} e as médias \bar{Y}_i , como no método anteriormente descrito;
2. Agora as médias serão agrupadas em b grupos de tamanho k , e representaremos as médias por grupo como $\bar{Y}_x(k)$ para $x = 1, 2, \dots, b$. O tamanho destes grupos é selecionado de forma a garantir que as médias por grupo sejam aceitas em teste de aderência para normalidade e em teste quanto a ausência de correlação. Recomenda-se, inclusive, um mínimo de 20 grupos. Após a divisão em grupos, denotamos a série temporal resultante por:

$$\bar{Y}(k) = \{\bar{Y}_1(k), \bar{Y}_2(k), \dots, \bar{Y}_b(k)\}$$

3. Será gerada então uma carta de controle para esta série temporal. Considera-se as estimativas de média populacional $\hat{\mu}$ e desvio padrão $\hat{\sigma}$ obtidas da última metade da série $\bar{Y}(k)$. De posse destas estimativas os seguintes limites de controle LC são obtidos: $LC = \hat{\mu} \pm z \frac{\hat{\sigma}}{\sqrt{\frac{b}{2}}}$ para $z = 1, 2, 3$;
4. A carta de controle é construída com os três limites de controle. São seguidas as seguintes regras para a verificação de saída de controle:
 - (a) Existência de pontos fora do limite de controle de 3σ ;
 - (b) Dois pontos fora do limite de controle de 2σ de três consecutivos;
 - (c) Quatro pontos fora do limite de controle de 1σ de cinco consecutivos;
 - (d) Existência de oito pontos consecutivos acima da média ou então oito pontos consecutivos abaixo da média.

Entende-se que o processo atingiu o equilíbrio quando está em controle.

Testes de Aleatorização

O método baseado em testes de Aleatorização [21] é formulado como um conjunto de testes de hipóteses acerca da média do processo. São executados testes de Aleatorização visando testar a validade da hipótese nula de que a média se mantém sem alteração durante toda a simulação. Uma grande vantagem sobre este método seria a possibilidade de relaxar a hipótese de normalidade. O método pode ser descrito através dos seguintes passos:

1. Execute a simulação considerando um tamanho pré-estabelecido de m horas. Defina Y_i como a i -ésima observação da variável de interesse na simulação;
2. Divida a série Y_1, Y_2, \dots, Y_m em b grupos de tamanho k e obtenha médias $\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_b$;
3. Particione o conjunto de médias em dois conjuntos C_1 e C_2 . Para a primeira iteração, o primeiro conjunto conterá apenas \bar{Y}_1 e o segundo conjunto as outras $b-1$ médias. Obtenha as médias \bar{Y}_{C_1} e \bar{Y}_{C_2} dentro dos dois conjuntos;
4. Para cada iteração, as médias \bar{Y}_{C_1} e \bar{Y}_{C_2} são comparadas. Se existir diferença estatisticamente significativa entre elas então a hipótese nula é rejeitada;
5. Se a hipótese nula for rejeitada os dois conjuntos C_1 e C_2 são modificados; é acrescentada a média \bar{Y}_2 ao conjunto C_1 enquanto o conjunto C_2 passa a ser composto por $b-2$ elementos. O teste é executado novamente agora para os novos conjuntos C_1 e C_2 ;
6. Enquanto a hipótese nula estiver sendo rejeitada, o conjunto C_1 é aumentado como no passo anterior. Quando a hipótese nula for aceita pela primeira vez, entende-se que o grupo C_2 é composto pelas observações do estado estacionário.

Cruzamento das Médias

A regra do cruzamento das médias [4] é um procedimento heurístico com objetivo de obter a quantidade de vezes em que o valor para a variável de interesse saiu de um valor menor que média e foi para um valor maior que a média em apenas um passo, e também quando a variável de interesse saiu de um valor maior que média e foi para um valor menor que a média em apenas um passo. Se esta contagem atinge um valor pré-estabelecido l , então se assume que o processo se tornou estacionário. Entretanto, este valor pré-estabelecido é uma escolha do analista. De uma forma mais prática podemos descrever o procedimento através dos seguintes passos:

1. Execute a simulação considerando um tamanho pré-estabelecido de m horas. Defina Y_i como a i -ésima observação da variável de interesse na simulação;
2. Obtenha a série temporal Y_1, Y_2, \dots, Y_m e a sua média $\bar{Y} = \frac{\sum_{i=1}^m Y_i}{m}$;
3. Para $i = 1, 2, \dots, m$ defina ω_i tal que $\omega_i = 1$ se $Y_{i-1} < \bar{Y} < Y_i$ ou $Y_i < \bar{Y} < Y_{i-1}$ e ainda $\omega_i = 0$ caso contrário;
4. O número de vezes que a série cruza a média é dado por $\Omega_m = \sum_{i=1}^m \omega_i$ para i de 1 até m ;
5. Calcule $\Omega_1, \Omega_2, \dots, \Omega_k$, até algum k , tal que Ω_k atinja o valor pré-estabelecido l ;

Regra de Erros Marginais

A Regra de erros marginais [20] é um procedimento denominado (MSER), e possui uma versão modificada (MSER-n) [17]. O método MSER se baseia em obter o período de aquecimento adequado d^* que melhor equilibra o balanço entre a ausência de vício (alta precisão) e redução do tamanho da amostra (baixa precisão) considerando a série $Y_{d^*+1}, Y_{d^*+2}, \dots, Y_m$ obtida de uma execução de tamanho m com as observações Y_1, Y_2, \dots, Y_m . Os dois casos são métodos que escolhem um ponto de truncamento d^* que minimiza a largura do intervalo de confiança marginal sobre a média da amostra truncada (Note que dada a série truncada de observações, existe uma correlação seqüencial, portanto o intervalo de confiança marginal não é uma estimativa adequada para a média truncada. Este intervalo está sendo usado aqui como uma medida

de homogeneidade para a série truncada que será analisada). Dada uma série de saída finita Y_1, Y_2, \dots, Y_m , o ponto de truncamento ideal para a seqüência é dado por:

$$d^* = \arg \min \left[\frac{1}{(m-d)^2} \sum_{i=d+1}^m (Y_i - \bar{Y}_{m,d})^2 \right] \text{ em que } \bar{Y}_{m,d} = \frac{\sum_{i=d+1}^m Y_i}{m-d}$$

O método MSER- n considera novamente a equação anterior, entretanto com as observações divididas em b grupos de tamanho k , são consideradas as b médias dos grupos.

Cruzamento das Médias Iterado

O método do Cruzamento das Médias Iterado [12] avalia os cruzamentos ocorridos em um passo em relação a média, mas não da forma proposta no método do cruzamento de médias. Este procedimento é feito de uma forma iterada. Para tanto, considere n replicações do procedimento de simulação, cada uma delas de tamanho m . Defina as médias das observações entre replicações por \bar{Y}_i com $i = 1, 2, \dots, m$. Denote por \bar{Y}_j a média descartando as primeiras $j - 1$ médias entre replicações, neste caso $\bar{Y}_j = \frac{\sum_{i=j}^m \bar{Y}_i}{m - (j - 1)}$. É fácil ver que para a série $\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_m$, algumas das observações iniciais são pertencentes ao período transiente. Em geral, todas essas observações iniciais são menores ou então, são todas maiores que \bar{Y}_1 . O objetivo é descartar estas observações de uma forma iterada como veremos a seguir através dos seguintes passos:

1. Execute n replicações da simulação considerando um tamanho pré-estabelecido m para as replicações e obtenha os valores \bar{Y}_i para $i = 1, 2, \dots, m$ e o valor \bar{Y}_1 ;
2. Verifique se \bar{Y}_1 é inferior ou superior à \bar{Y}_1 . Se inferior, determine o menor k tal que \bar{Y}_{k+1} seja superior a \bar{Y}_1 e descarte as k primeiras observações, o procedimento é análogo para \bar{Y}_1 superior a \bar{Y}_1 ;
3. Considere agora a média das médias \bar{Y}_{k+1} e também a série truncada \bar{Y}_i para os valores $i = k + 1, k + 2, \dots, m$, repita o procedimento até que o número de valores em seqüência, que sejam inferiores ou então superiores a média das médias, seja limitado por 2, ou seja, considerando a média das médias dos últimos valores, é possível cruzar a média em pelo menos dois passos. Neste momento o método atingiu sua estimativa para o comprimento do período de aquecimento. Este método tem uma formulação bastante simples, sem a dependência de escolha de parâmetros pelo analista.

Modelo Experimental

A estratégia de comparação será baseada na utilização do modelo experimental apresentado em [12] para avaliar os métodos. No referido modelo experimental, os métodos são testados em modelos simples de trabalho. São cinco procedimentos em série, cada procedimento tem diferentes números de máquinas (recursos) e existem três tipos distintos de peças a serem produzidas. As chegadas são distribuídas conforme Poisson(λ) e os tempos de serviço para os procedimentos são exponenciais com médias $\mu_{i,j}$ em que i representa a peça e j o procedimento. O modelo foi construído e executado através do software Arena. As peças que chegam são divididas nos três tipos, com probabilidades: 0,5; 0,3 e 0,2 respectivamente. As peças são processadas em tempos de serviço diferentes. Após o tratamento, as peças saem do sistema. Os métodos foram aplicados ao modelo com diferentes níveis de utilização como segue:

- Tipo I: Modelo tem um alto nível de utilização. A utilização média de todos os recursos, esta próxima de 90% com utilizações individuais entre 80% e 95%;

- Tipo II: Modelo tem um nível moderado de utilização. A utilização média de recursos, esta próxima de 70% com utilizações individuais entre 65% e 80%;
- Tipo III: Modelo tem um baixo nível de utilização. A utilização média de recursos, esta próxima de 50% com utilizações individuais entre 45% e 65%;

Quanto às condições de execução da simulação, o sistema foi iniciado vazio e ocioso. Foi considerado um tamanho inicial para cada simulação de 5000 horas. O parâmetro de interesse θ avaliado foi o tempo médio de permanência das entidades no sistema. Trata-se de um parâmetro de desempenho escolhido de forma arbitrária com intuito apenas de testar os métodos. Para outros parâmetros, a análise pode revelar resultados diferentes.

Métricas comparativas entre os métodos

Os seis métodos descritos serão avaliados através de execuções do modelo experimental descrito anteriormente. A principal métrica de comparação utilizada será o erro quadrático médio para o estimador do parâmetro de interesse θ dado por $EQM(\hat{\theta}) = \mathbb{E}(\theta - \hat{\theta})^2 = \left[\mathbb{E}(\theta - \hat{\theta}) \right]^2 + Var(\hat{\theta})$. Obviamente, o valor θ é desconhecido, para tanto utilizamos uma estimativa deste parâmetro, considerando um período de aquecimento que descarta a metade inicial da série, ou seja, uma escolha completamente conservadora para que a análise comparativa fosse justa e adequada. Já os valores $\hat{\theta}$ são obtidos considerando a média das realizações descartando o período de aquecimento sugerido por cada um dos métodos.

Vale ressaltar, que em situações nas quais o erro quadrático médio apresenta valores muito próximos de zero, a comparação através desta métrica poderia não ser suficientemente conclusiva, para tanto utilizamos a segunda métrica, a variação percentual do erro quadrático médio da estimativa $\hat{\theta}$, fornecida por cada método em relação ao erro quadrático médio considerando a estimativa $\hat{\theta}_{ini}$ que leva em conta todas as observações da simulação.

Avaliação da variação relativa entre o Erro quadrático médio inicial (EQM_{ini}) (sem deleção de dados), e o Erro quadrático médio posterior (EQM_{pos}) (considerando a deleção de dados por período de aquecimento).

$$V_r(EQM) = \frac{(EQM_{pos} - EQM_{ini})}{EQM_{ini}}$$

Resultados Numéricos

Foram executadas 30 replicações, cada uma com 5000 horas de simulação para cada nível de utilização do Modelo Experimental. A avaliação foi conduzida considerando os diferentes tipos de utilização para o Modelos Experimental e também, as diferentes técnicas para detecção do comprimento do período de aquecimento. Denotaremos por L , o comprimento detectado por cada uma das técnicas em estudo.

Modelo Experimental com alto nível de utilização

Considerando o Modelo Experimental Tipo I, com alto nível de utilização do sistema a ser simulado e considerando todas as observações da simulação obtivemos uma estimativa inicial para o parâmetro em avaliação dada por $\hat{\theta}_{ini} = 134,7813$. Quando considerando o descarte da primeira metade da série obtivemos $\theta = 135,0477$. Nestas condições o erro quadrático médio para $\hat{\theta}_{ini}$ foi $EQM(\hat{\theta}_{ini}) = 4,155896$.

A Tabela 1 apresenta os resultados para estas medidas e também a redução do erro quadrático médio em relação à $EQM(\hat{\theta}_{ini})$ pela variação relativa.

A Figura 1 apresenta os diferentes períodos de aquecimento obtidos através de cada um dos métodos em avaliação.

Tabela 1: Resultados avaliativos para o Modelo Experimental Tipo I.

Método	L estimado	$\hat{\theta}$	$EQM(\hat{\theta})$	$V_r(EQM)$
Welch	1164905	135,18100	0,01844658	-0,9955613
SPC	399450	135,00940	0,01103734	-0,9973442
Aleatorização	365390	135,01630	0,01132434	-0,9972751
Cruzamento	154428	135,11140	0,06317613	-0,9847984
MSR-n	5013	134,86980	0,83913980	-0,7980845
Cruzamento iterado	163101	135,11260	0,06359186	-0,9846984

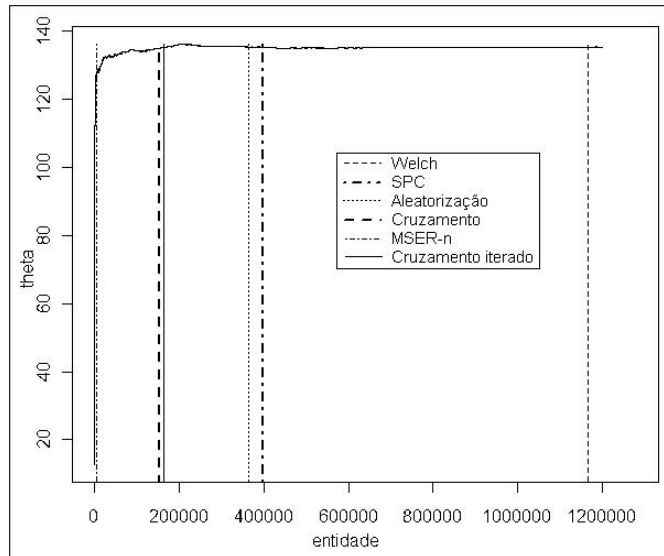


Figura 1: Períodos de aquecimento para o Modelo Experimental Tipo I.

Observa-se que os melhores resultados foram obtidos através do Método SPC, entretanto, este método propõe o descarte de uma quantidade de dados muito grande se comparada aos demais, isto decorre da necessidade de se obter grupos que atendam os requisitos de normalidade e ausência de correlação. No caso específico desta simulação, para que estes requisitos fossem satisfeitos, foi necessário construir grupos muito extensos. Neste caso a exclusão de uma pequena quantidade de grupos pode levar a exclusão de um volume muito grande de dados.

É fácil ver que, se a quantidade de dados descartada for muito grande as estimativas produzidas serão muito próximas do valor θ , que considera apenas a segunda metade da série, por outro lado, se for possível produzir estimativas satisfatórias para θ com um volume significativamente menor de dados descartados, o tamanho total da simulação poderia ser bastante inferior. Desta forma, não podemos considerar a solução proposta através do método SPC como uma solução adequada. Os métodos de Welch e de Aleatorização também apresentam resultados bons, mas sofrem de um problema semelhante ao método SPC com um grande descarte de dados.

O gráfico apresentado na Figura 1 ilustra estes fatos. A regra dos erros marginais MSR-n apresenta os piores resultados. A regra do cruzamento das médias e o cruzamento das médias iterado apresentam resultados bastante semelhantes. Dentre as métricas utilizadas para a avaliação seria difícil escolher uma destas. Se considerarmos o fato da regra do cruzamento das médias exigir a calibração do parâmetro associado a quantos cruzamentos devem ser obtidos até que se entenda atingir o período de aquecimento, enquanto a regra do cruzamento das médias iterado é livre deste parâmetro, isto nos leva a optar pela segunda opção.

Modelo Experimental com moderado nível de utilização

Considerando o Modelo Experimental Tipo II, com moderado nível de utilização e todas as observações da simulação obtivemos uma estimativa inicial $\hat{\theta}_{ini} = 37,07618$. Quando considerando o descarte da primeira metade da série obtivemos $\theta = 37,07334$. O erro quadrático médio para $\hat{\theta}_{ini}$ foi $EQM(\hat{\theta}_{ini}) = 1,1143684 \times 10^{-2}$.

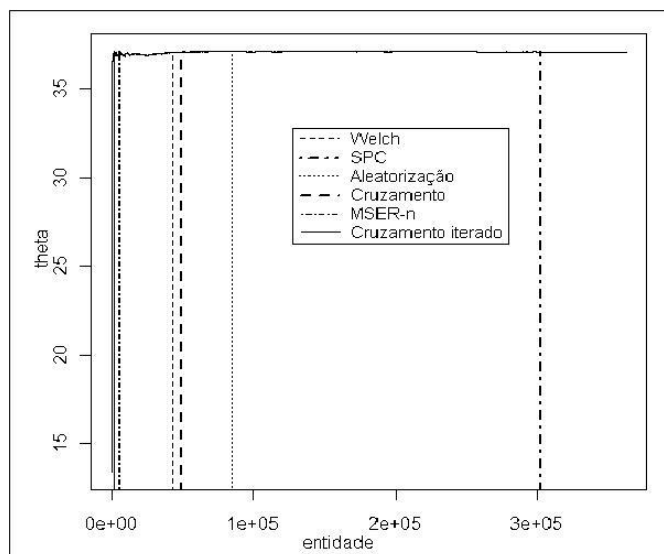


Figura 2: Períodos de aquecimento para o Modelo Experimental Tipo II.

A Figura 2 apresenta os diferentes períodos de aquecimento obtidos através de cada um dos métodos em avaliação. A Tabela 2 apresenta os resultados para estas medidas e também a redução percentual do erro quadrático médio em relação à $EQM(\hat{\theta}_{ini})$.

Tabela 2: Resultados avaliativos para o Modelo Experimental Tipo II.

Método	L estimado	$\hat{\theta}$	$EQM(\hat{\theta})$	$V_r(EQM)$
Welch	42506	37,08409	$3,667991 \times 10^{-4}$	-0,9679283
SPC	301988	37,07547	$3,829701 \times 10^{-5}$	-0,9966514
Aleatorização	84645	37,08309	$3,025090 \times 10^{-4}$	-0,9999272
Cruzamento	49273	37,08427	$3,685085 \times 10^{-4}$	-0,9677788
MSER-n	5000	37,07906	$9,238176 \times 10^{-4}$	-0,9192244
Cruzamento iterado	1900	37,07860	$9,901725 \times 10^{-4}$	-0,9134225

Novamente o método SPC apresenta o melhor resultado obtido através de um descarte excessivo de dados, isto pode ser ilustrado através do gráfico da Figura 2. Podemos observar que os melhores resultados foram obtidos através do Método de Welch, Método da Aleatorização e da Regra do cruzamento das médias, entretanto novamente estes métodos propõem o descarte de uma quantidade de dados grande se comparada aos demais. O volume de dados descartados é pelo menos 20 vezes maior que através do cruzamento das médias iterado.

Por outro lado mesmo com este volume bem inferior de dados descartados o cruzamento das médias iterado fornece resultados bastante adequados. A regra dos erros marginais MSER-n apresenta resultados similares aos do cruzamento das médias iterado, entretanto para obter estes resultados foi necessário um descarte de 2,5 vezes mais dados que com o cruzamento das médias iterado. Novamente frisamos que boas estimativas com um pequeno comprimento de período de aquecimento levam um custo computacional menor para a execução de todo o procedimento.

Modelo Experimental com baixo nível de utilização

Considerando o Modelo Experimental Tipo III, com baixo nível de utilização do sistema a ser simulado e considerando todas as observações da simulação obtivemos uma estimativa inicial para o parâmetro em avaliação dada por $\hat{\theta}_{ini} = 28,26932$. Quando considerando o descarte da primeira metade da série obtivemos $\theta = 28,27379$. Nestas condições o erro quadrático médio para $\hat{\theta}_{ini}$ foi $EQM(\hat{\theta}_{ini}) = 4,985163 \times 10^{-3}$.

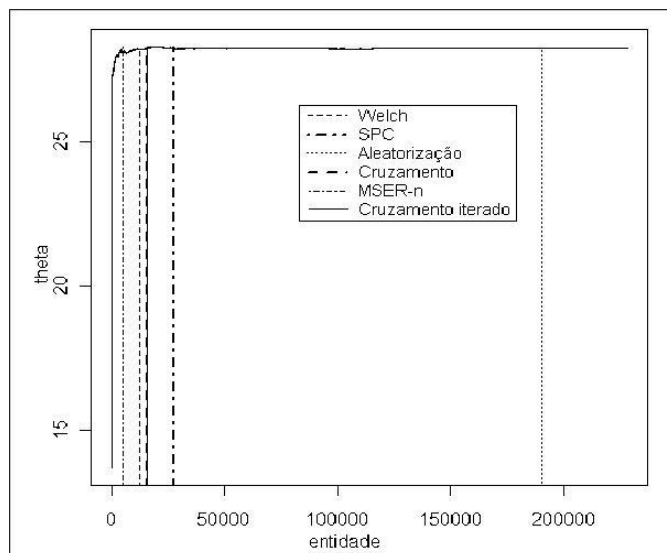


Figura 3: Períodos de aquecimento para o Modelo Experimental Tipo III.

A Figura 3 apresenta os diferentes períodos de aquecimento obtidos através de cada um dos métodos em avaliação. A Tabela 3 apresenta os resultados para estas medidas e também a redução percentual do erro quadrático médio em relação à $EQM(\hat{\theta}_{ini})$.

Tabela 3: Resultados avaliativos para o Modelo Experimental Tipo III.

Método	L estimado	$\hat{\theta}$	$EQM(\hat{\theta})$	$V_r(EQM)$
Welch	12030	28,27354	$7,422078 \times 10^{-5}$	-0,9851117
SPC	27163	28,27323	$5,230056 \times 10^{-5}$	-0,9895088
Aleatorização	190120	28,27549	$3,341882 \times 10^{-5}$	-0,9932963
Cruzamento	15975	28,27370	$6,995527 \times 10^{-5}$	-0,9859673
MSER-n	4935	28,27256	$1,887736 \times 10^{-4}$	-0,9623411
Cruzamento iterado	16022	28,27370	$6,995527 \times 10^{-5}$	-0,9859673

Os melhores resultados foram obtidos pelo método de Aleatorização, mas com um volume de dados descartado muito maior que os outros métodos. A Regra dos erros marginais MSER-n apresentou resultados satisfatórios, mas ligeiramente inferiores aos demais métodos. Excluindo estes dois métodos observamos resultados semelhantes entre os demais.

Conclusões

A estratégia de avaliação utilizada neste trabalho tem uma implementação bastante simples e pode ser utilizada para testar diversos métodos. As métricas aqui utilizadas descrevem de forma bastante clara uma classificação entre os métodos. Não buscamos concluir um método de senso comum acima de quaisquer outros, entretanto o método nota-se que o método do cruzamento das médis iterado se mostrou bastante eficiente para todos os níveis de utilização do modelo

experimental. São apresentadas conclusões preliminares que buscam atestar as qualidades e deficiências de cada um dos métodos em análise. É possível notar que a influência do analista na calibração de alguns parâmetros associados ao método de detecção do período de aquecimento tende a afetar sobremaneira a qualidade de resultados. Este é o motivo preponderante para a escolha de uma técnica livre da calibração de parâmetros pelo analista.

O método de Welch necessita da definição de um parâmetro associado ao comprimento da janela, espera-se que este parâmetro não influencie os resultados de forma significativa. Por outro lado, definir de forma específica um critério de convergência para a série das médias é um problema pontual da utilização deste método. Modificações na escolha destes critérios afetam de forma significativa a solução proposta através desta técnica. O método baseado no Controle estatístico de processo parece bastante justificável, entretanto pode se tornar difícil definir a quantidade de grupos que garantiria sucesso em testes de normalidade e correlação. Em particular, pode ser necessária a execução de um tempo maior de simulação até que se consiga adequar a amostra simulada e a divisão em grupos para os referidos testes. Os Testes de Aleatorização necessitam da escolha da quantidade b de grupos na qual a amostra será dividida. Em particular para este método, uma variação na quantidade de grupos pode alterar de forma significativa a conclusão do teste. Vale ressaltar, que se o comprimento proposto inicialmente para cada grupo for superior ao adequado período de aquecimento para os dados, automaticamente a solução apresentada será superestimada em relação a este adequado período de aquecimento. A Regra do cruzamento das médias é bastante dependente do parâmetro l a ser escolhido pelo analista. O valor l determina quantos cruzamentos da média em apenas um passo se deve observar para que se considere ter atingida a estacionariedade. A Regra dos erros marginais é um procedimento que na sua forma original se torna livre da escolha de parâmetros pelo analista. Entretanto, tende a subestimar o adequado período de aquecimento. Este fato é uma das principais razões para a postulação da técnica MSER- n dividindo em n grupos, entretanto voltamos ao problema da subjetividade da escolha do valor n . Novamente, se trata de um parâmetro que pode afetar sobremaneira as conclusões do método. Desta análise, concluí-se pela necessidade de um método simples quanto a sua utilização, robusto quanto a sua eficiência ao estimar o comprimento adequado para o período de aquecimento e livre da escolha de parâmetros pelo analista, ou pelo menos, muito pouco dependente de parâmetros que devam ser escolhidos previamente.

Foi escolhido de forma arbitrária um parâmetro θ associado ao sistema em estudo, bem como um modelo experimental específico. Uma proposta de continuidade lógica é um estudo com a estimação de outros parâmetros de interesse e a postulação de modelos experimentais mais complexos. Acredita-se que exista uma ligação entre a complexidade do modelo e a adequação da técnica para detecção do período de aquecimento. Esta ligação somente pode ser avaliada através da construção de mais modelos e da execução de mais experimentos. A considerar o modelo experimental em uso neste trabalho, atesta-se a grande capacidade de adequação da técnica do cruzamento das médias iterado em todos os níveis de utilização. Os resultados obtidos são bastante promissores e levam a um estudo mais profundo acerca desta técnica.

Referências

- [1] BANKS, J.; CARSON, J. S.; NELSON, B. L.; NICOL, D. M., *Discrete-Event System Simulation*, third ed, Upper Saddle River: Prentice-Hall, 2001.
- [2] CONWAY, R. W., Some Tactical Problems in Digital Simulation, *Management Science* **10:1**, 47-61, 1963.
- [3] FISHMAN, G. S., Estimating sample size in computing simulation experiments, *Management Science* **18:1**, 21-38, 1971.

- [4] FISHMAN, G. S., Concepts and Methods in discrete Event Digital Simulation, New York: John Wiley and Sons, 1973.
- [5] GAFARIAN, A. V.; ANCKER, C. J.; MORISAKU, T., Evaluation of commonly used rules for detecting steady-state in computer simulation, *Naval Research Logistics Quarterly* **25**, 511-529, 1978.
- [6] GOLDSMAN, D.; SCHRUBEN, L. W.; SWAIN, J. J., Tests for transient means in simulated time series, *Naval Research Logistics Quarterly* **41**, 171-187, 1994.
- [7] GORDON, G., System Simulation, Englewood Cliffs: Prentice-Hall, 1969.
- [8] JACKWAY, P. T.; DESILVA, B. M., A methodology for initialization bias reduction in computer simulation output, *Operations Research* **31:6**, 1167-1178, 1983.
- [9] KELTON, W. D.; LAW, A. M., A new approach for dealing with the startup problem in discrete event simulation, *Naval Research Logistics Quarterly* **30**, 641-658, 1983.
- [10] MAHAJAN, P. S., INGALLS, R. G. Evaluation of methods used to detect warmup period in steady state simulation. In Ingalls, R. G., Rossetti, M. D., Smith, J. S., Peters, B. A. (org.). Proceedings of the 2004 Winter Simulation Conference, Los Alamitos-CA: *IEEE Computer Society*, 663-671, 2004.
- [11] NELSON, B. L. Statistical analysis of simulation results. In Salvendy, G. (org.). Handbook of Industrial Engineering, New York: John Wiley and Sons, cap. 102, 2004.
- [12] NORATO, H. M. G., DUARTE, A. R., PINTO, L. R. Uma proposta para a detecção do adequado período de aquecimento em modelos de simulação. In Anais do XVIII Simpósio de Engenharia de Produção-SIMPEP, Bauru-SP, 1-13, 2011.
- [13] PAWLIKOWSKI, K., A survey of problems and solutions, *Computing Surveys* **22:2**, 123-170, 1990.
- [14] ROBINSON, S. A statistical process control approach for estimating the warmup period. In Yücesan, E., Chen, C. H., Snowdon, J. L., Charnes, J. M. (org.). Proceedings of the 2002 Winter Simulation Conference, San Diego-CA: *IEEE Computer Society*, 663-671, 2004.
- [15] SCHRUBEN, L. W., Detecting initialization bias in simulation output, *Operations Research* **30:3**, 569-590, 1982.
- [16] SCHRUBEN, L. W.; SINGH, H.; TIERNEY, L., Optimal tests for initialization bias in simulation output, *Operations Research* **31:6**, 1167-1178, 1983.
- [17] SPRATT, S. C., Heuristic for startup problem, Master thesis, University of Virginia, Virginia, 1998.
- [18] VASSILACOPOULOS, G., Testing for initialization bias in simulation output, *Simulation* **52:4**, 151-153, 1989.
- [19] WELCH, P. The statistical analysis of simulation results. In Lavenberg, S. (org.). The computer modeling handbook, London: Elsevier Academic Press, 268-328, 1983.
- [20] WHITE, K. P., An effective truncation heuristic for bias reduction in simulation output, *Simulation* **69:6**, 323-334, 1997.
- [21] YÜCESAN, E., Randomization testes for initialization bias in simulation output, *Naval Research Logistics Quarterly* **40**, 643-663, 1993.