

CAPÍTULO 1

INTRODUÇÃO

1.1. CONSIDERAÇÕES INICIAIS

UM sistema elétrico de potência tem por única função suprir os consumidores com energia elétrica da forma mais econômica e confiável possível. Esses objetivos são conflitantes, pois um melhor desempenho (maior confiabilidade) requer investimentos mais acentuados, o que implica em tarifas mais elevadas. Assim, em estudos de planejamento da expansão de sistemas elétricos, para que se possa decidir quando e onde investir é necessário o emprego de metodologias capazes de quantificar os custos e a confiabilidade das diversas alternativas e escolher dentre elas as mais adequadas [M86].

Nas últimas décadas, os índices de confiabilidade da capacidade de geração (GCR) têm sido parâmetros muito úteis para auxiliar engenheiros de planejamento em muitas decisões. Diversos artigos e publicações bibliográficas têm surgido neste assunto; veja, por exemplo, [BFB01]. Quando a expansão de um sistema de potência está sendo planejada é necessário avaliar se uma dada configuração de usinas geradoras atende adequadamente aos requisitos da demanda global do mercado consumidor. Este estudo é denominado avaliação da confiabilidade de geração. A metodologia básica comum a todos os métodos consiste em supor que não há problemas de transporte de energia, i.e. as linhas de transmissão são consideradas perfeitamente confiáveis e sem restrições de capacidade.

Dentre os métodos analíticos, o mais popular é conhecido como a abordagem da LOLE (*Loss of Load Expectation*), que é uma técnica amplamente utilizada devido a sua flexibilidade e simplicidade de aplicação [E78, A90, BA96, BAS94]. Esta técnica combina o modelo de geração do sistema, representada pela *tabela de probabilidades das capacidades indisponíveis* (COPT), com as características ou modelo da carga do sistema para produzir índices de confiabilidade como a LOLP (*Loss of Load Probability*), LOLE, EPNS (*Expected Power Not Supplied*) e EENS (*Expected Energy Not Supplied*). A eficiência computacional destes métodos foi consideravelmente melhorada por meio da utilização de técnicas de convolução discreta [ALAB81, LPC88].

A análise de frequência e duração (F&D) é certamente uma abordagem analítica mais completa que a LOLE, porque também fornece a frequência de ocorrência de uma condição de capacidade insuficiente e a duração pela qual se espera que perdure. Estas são indicadas pelos índices LOLF (*Loss of Load Frequency*) e LOLD (*Loss of Load Duration*), respectivamente [BA96, SB77, CS86]. Basicamente, o método de F&D combina, através de técnicas de Markov, estados de capacidades e de carga para obter estados de reserva. Uma vez que todos os estados com margem negativa ou estados de falha sejam identificados, todos os índices previamente mencionados são avaliados.

A fim de tornar este método viável para grandes sistemas de geração encontrados na prática, as técnicas de convolução discreta foram estendidas para tratar com os conceitos de frequência e duração [LMC91, LCM92]. Evitando uma abordagem combinatória em sistemas práticos de geração, as técnicas de convolução discreta claramente aumentaram a eficiência computacional dos métodos analíticos. Estes métodos, no entanto, permanecem restritos para certos tipos de condições do sistema [PB92].

Os métodos baseados em Simulação Monte Carlo (SMC) [R91, BL94] possibilitam a avaliação da confiabilidade para as mais diversas condições do sistema, e podem ser divididos nas categorias não-cronológicas e cronológicas. A SMC não-cronológica tem sido uma ferramenta extremamente útil na avaliação da confiabilidade de geração e/ou transmissão de grandes sistemas encontrados na prática [PB92, MPL92, BL94, LRMB04, L05, SR08].

Métodos baseados em SMC cronológica são mais poderosos para avaliar sistemas elétricos de potência mais complexos, porém o esforço computacional é muito mais substancial em comparação com os métodos não-cronológicos. Para superar o problema do custo computacional das simulações cronológicas, ferramentas de SMC pseudo-cronológicas foram propostas [LMMB00].

Mais recentemente, técnicas inteligentes de busca baseadas em população foram também utilizadas para resolver problemas de confiabilidade em sistemas de potência [WS08, MCRLS09]. Estes são métodos interessantes, porém, apresentam algumas dificuldades quando aplicados em sistemas de grande porte. Além disso, o controle da convergência do processo de otimização é, em geral, muito complicado.

Sem dúvida, ferramentas baseadas em técnicas de SMC são extremamente robustas para resolver problemas de confiabilidade em sistemas de potência, principalmente em sistemas de grande porte. Porém, eles podem encontrar algumas dificuldades no que diz respeito a eventos raros: e.g. para avaliar valores muito pequenos do índice LOLP, i.e. 10^{-5} .

Embora se possa argumentar que uma configuração muito confiável para um sistema de geração não é um problema, e que a simulação deve ser parada após se verificar tal condição, em estudos de planejamento da expansão estes valores devem ser calculados considerando que diversas configurações de reforço estão sendo comparadas. Além do mais, em problemas que envolvem critérios específicos de projeto, o planejador poderia ter que medir a probabilidade de certo evento, que poderia ser muito raro, mas, a priori, esta condição é desconhecida.

A fim de melhorar o desempenho das ferramentas de SMC, diversas técnicas de redução de variância estão disponíveis [RK07, PAK08]. Algumas delas já foram testadas na avaliação da confiabilidade de sistemas de potência; veja, por exemplo, [LMMB00, AEPP90, BJ96, BJ97, ZX02]. Tendo em mente sistemas de potência reais, as eficiências de alguns métodos podem ser consideradas relevantes e outras marginais.

Em relação à raridade dos eventos de falha envolvidos, pouca discussão tem sido conduzida em confiabilidade de sistemas de potência. Um dos métodos que tem apresentado resultados interessantes é baseado em amostragem por importância (IS); uma técnica que tem evoluído consideravelmente nos últimos anos, com a utilização de algoritmos de otimização para selecionar os parâmetros da função densidade de probabilidade auxiliar de uma maneira ótima. Estes novos algoritmos estão baseados nos conceitos de Entropia Cruzada e da divergência de Kullback-Leibler, que são idéias fundamentais da moderna teoria da informação [RK04, RK07, KHN07].

Este trabalho apresenta uma nova abordagem de SMC baseada no Método da Entropia Cruzada (MEC) para avaliar índices de confiabilidade em sistemas de geração. A idéia básica é utilizar uma função massa de probabilidade auxiliar, cujos parâmetros são obtidos de um processo de otimização estocástico. Diversos aspectos como o tamanho do sistema, a raridade dos eventos de falha, o número de unidades diferentes e o perfil da curva de carga são considerados a fim de verificar o desempenho da metodologia proposta.

O IEEE-RTS [IEEE79] é utilizado para verificar a metodologia proposta. Diversos sistemas testes são criados a partir de modificações no IEEE-RTS original. Estas modificações têm como principal objetivo comparar o desempenho do método proposto com uma SMC não-cronológica e uma técnica analítica muito eficiente baseada em convolução discreta [LMC91]. Uma configuração do sistema Sul-Sudeste Brasileiro (SSB) de geração, planejada para os anos 90, é também utilizada para demonstrar a aplicação do método proposto, denominado *CE-ISMC*, em sistemas de potência encontrados na prática.

1.2. DESENVOLVIMENTO HISTÓRICO

Não se pode determinar com precisão quando foi publicado o primeiro trabalho sobre a avaliação da confiabilidade de geração utilizando métodos probabilísticos, porém, aproximadamente em 1933 o interesse nesta área tornou-se evidente. Os principais trabalhos publicados entre 1977 e 1982 estão bem documentados em [ABL84]. De maneira semelhante, os principais trabalhos publicados no período compreendido entre 1996 e 1999 estão referenciados em [BFB01].

Desde 1933, bons artigos foram surgindo a cada ano. Calabrese, Lyman, Loane e Watchorn [C47, L47, LW47] publicaram os trabalhos pioneiros que estruturam os métodos usados até o presente. Estes artigos formam a base da avaliação da confiabilidade da capacidade de geração por técnicas probabilísticas, e servem como ponto de partida de outros bons trabalhos.

Em 1958, Halperin e Adler [HA58] introduziram pela primeira vez um método de frequência e duração para a avaliação da confiabilidade de geração. Todavia a metodologia apresentada foi um tanto complicada e os índices não foram utilizados de fato até 1968-71. Nesse período novas publicações estenderam a metodologia de frequência e duração desenvolvendo técnicas recursivas para a construção do modelo de geração e para sua combinação com o modelo de carga, as quais facilitaram a implementação computacional.

Em 1959, um segundo grupo de artigos foi publicado por Baldwin, Gaver, Hoffman, entre outros [BGH59, BBGH59, BGHR60]. Estas publicações modificaram e estenderam os métodos propostos por publicações anteriores e introduziram uma metodologia mais sofisticada ao problema usando técnicas de simulação baseadas em amostragem estatística,

surgindo desta maneira, as primeiras aplicações da Simulação Monte Carlo em sistemas elétricos de potência

Em 1982, Billinton, Wee e Hamoud [BWH82] publicaram um resumo didático sobre os métodos básicos para a avaliação da confiabilidade de geração. Foi apresentado também um método recursivo para a obtenção da frequência e probabilidade acumulada dos estados de capacidade. Este modelo de geração não foi combinado com nenhum modelo de carga.

Um sistema teste para comparar as diversas metodologias e utilizado até hoje é o IEEE-RTS (i.e., IEEE Reliability Test System), publicado em 1979 pelo IEEE Subcommittee on Application of Probability Methods [IEEE79]. Em 1986 Allan, Billinton e Abdel-Gawad propuseram modificações que poderiam ser aplicadas a este sistema para o estudo da confiabilidade da geração [ABA86].

Também em 1986, Melo [M86] em sua Tese de Mestrado apresentou uma nova técnica de frequência e duração que permitiu modelar de forma precisa os diversos estados operativos das unidades geradoras, bem como os diversos estados de carga, a fim de se obter índices de confiabilidade mais realísticos para sistemas de grande porte. Este método foi baseado em técnicas de convolução discreta propostas por Allan, Leite da Silva, Abu-Nasser e Burchett [ALAB81] em 1981. Em 1991, Leite da Silva, Melo e Cunha [LMC91] apresentaram um método analítico extremamente eficiente, onde não somente as probabilidades dos estados de reserva podiam ser avaliados recursivamente por convolução discreta, mas também suas frequências incrementais, o que possibilitou a aplicação de técnicas de truncamento e arredondamento permitindo uma avaliação eficiente dos índices de frequência e duração para sistemas de grande porte.

Os métodos analíticos tradicionais precisavam muitas vezes de aproximações para contornar alguns problemas, principalmente no que diz respeito ao grande número de estados possíveis do sistema. Novos métodos baseados em técnicas de simulação, semelhantes a aqueles propostos por Baldwin et alii em 1959 [BGH59, BBGH59, BGHR60] começaram a aparecer. Estes métodos apresentavam vantagens competitivas com relação aos métodos analíticos em sistemas de grande porte (elevado número de estados) e com restrições operativas complexas.

Em 1982, Cunha, Ganes, Oliveira e Pereira [CGOP82] publicaram um dos primeiros trabalhos sobre a utilização da simulação Monte Carlo na avaliação da confiabilidade de geração. Trabalhos semelhantes começaram a ser publicados nos anos seguintes, sendo um dos mais importantes, o de Patton, Blackstone e Balu [PBB88] em 1988. Os métodos de simulação Monte Carlo eram baseados em técnicas de amostragem estatística, onde os estados eram amostrados proporcionalmente a sua probabilidade de ocorrência e possibilitaram a avaliação da confiabilidade para as mais diversas condições do sistema.

Apareceram duas categorias bem definidas: a simulação Monte Carlo não-cronológica e a cronológica. A SMC não-cronológica tem sido uma ferramenta extremamente útil na avaliação da confiabilidade de geração e/ou transmissão em sistemas reais. Exemplos de trabalhos relacionados são os de Pereira e Balu em 1992 e 1994 [PB92, BL94]. Os métodos baseados em SMC cronológica são mais poderosos para avaliar sistemas de potência mais complexos, porém o esforço computacional é muito maior em comparação com os métodos não-cronológicos, pois durante o processo de simulação toda a cronologia é levada em consideração. Por este motivo, sua aplicação é complexa, ainda hoje, na avaliação da confiabilidade composta de sistemas elétricos.

Para superar o problema do custo computacional das simulações cronológicas, ferramentas de SMC pseudo-cronológicas foram propostas por Leite da Silva, Manso, Mello e Billinton em 2000 [LMMB00]. Este método manteve a flexibilidade e a precisão da SMC cronológica reduzindo o esforço computacional a níveis similares da SMC não-cronológica.

A principal desvantagem dos métodos baseados em simulação Monte Carlo era a necessidade de um tempo de computação maior do que os métodos analíticos convencionais. Estes tempos de computação eram maiores quanto mais raros fossem os eventos envolvidos. O motivo era a necessidade de se ter um número elevado de amostras para se estimar um evento raro com relativa precisão. A raridade do evento passou a ser um problema. Vários trabalhos surgiram com propostas sobre como tratar a raridade do evento, estes tinham como principal objetivo reduzir o esforço computacional. Apareceram assim as chamadas Técnicas de Redução de Variância (VRT).

Em 1996, Billinton e Jonnavithula [BJ96] utilizaram VRT baseadas em variáveis de controle (CV) e variáveis antitéticas (AV) juntamente com a simulação Monte Carlo cronológica para avaliar a confiabilidade de sistemas elétricos. Em 1999, Lieber, Nemirovskii e Rubinstein [LNR99] utilizaram uma VRT baseada em amostragem por importância para acelerar a convergência da simulação Monte Carlo. A idéia central do IS era amostrar os estados do sistema com uma distribuição diferente da original, de tal maneira que o evento raro de interesse fosse encontrado com maior frequência, porém o problema era justamente encontrar esta nova distribuição pois ela era inicialmente desconhecida. Uma escolha errada poderia ter o efeito contrario e aumentar ainda mais o tempo de computação. Estimar esta distribuição era um processo complicado, e isto levou os pesquisadores a procurarem um método eficiente para resolver este problema.

Em 2002, Zhaohong e Xifan [ZX02] apresentaram uma nova técnica de redução de variância denominada Método da Fissão e Roleta que foi aplicada na avaliação da confiabilidade composta de sistemas elétricos de potência. Porém, a metodologia apresentada foi um tanto quanto complicada e seu desempenho não foi comparado com outras técnicas de redução de variância presentes na literatura. Além disso, este trabalho não apresentou aplicações do método proposto em sistemas de grande porte.

Também em 2002, Homem-de-Mello e Rubinstein [HR02] publicaram um trabalho de grande importância que apresentou um método para estimar a probabilidade de ocorrência de eventos raros. Neste trabalho, modelos de simulação baseados em amostragem por importância e a entropia cruzada de Kullback-Leibler foram utilizados. Inicialmente, a distribuição ótima de amostragem era estimada utilizando o CE, e a seguir, a probabilidade do evento raro era estimada utilizando a já conhecida técnica da amostragem por importância. Assim surgiu um novo e poderoso método para a simulação de eventos raros e a otimização contínua ou combinatória. Esta nova metodologia foi denominada Método da Entropia Cruzada. Em 2004, Rubinstein e Kroese [RK04] publicaram um livro que resumia os fundamentos do MEC e suas diversas aplicações.

Nos anos seguintes, vários trabalhos foram publicados sobre as aplicações do MEC nas mais diversas áreas. Exemplos recentes são os trabalhos de Hui, Bean, Kraetzl e Kroese em 2005 [HBKK05] sobre a aplicação do MEC na avaliação da confiabilidade de redes, os trabalhos de Kroese, Hui e Nariyai em 2007 [KHN07] sobre otimização da confiabilidade de

redes através da aplicação do MEC, e os trabalhos de Ernst, Glavic, Stan, Mannor e Wehenkel em 2007 [EGSMW07] sobre a aplicação do MEC na solução de problemas de otimização combinatória em sistemas elétricos de potência.

Mais recentemente, outra metodologia para avaliar a confiabilidade de geração que se afasta um pouco da linha das metodologias anteriores foi proposta por Wang e Singh em 2008 [WS08]. A proposta foi utilizar técnicas metaheurísticas de busca inteligente, baseadas em população, para avaliar a confiabilidade da geração. Nesta publicação, técnicas metaheurísticas como Algoritmos Genéticos (GA), Otimização por Enxame de Partículas (PSO), Otimização por Colônia de Formigas (ACO), entre outros, foram utilizados primeiramente para selecionar os estados dominantes de falha, ou seja, aqueles que têm maior contribuição para os índices de confiabilidade, e depois os índices são calculados analiticamente considerando os estados encontrados na etapa anterior.

De maneira semelhante a [WS08], Miranda, Carvalho, Rosa, Leite da Silva e Singh [MCRLS09] publicaram em 2009 um trabalho onde eram aplicadas variações do EPSO em confiabilidade de sistemas de potência. Ambos os métodos são interessantes, mas também possuem algumas limitações para controlar a convergência do processo de otimização e para tratar com sistemas de grande porte, pois muitas informações pertinentes a cada estado dominante devem ser salvos e levados a etapa seguinte, o que representa um custo de memória elevado, podendo inclusive inviabilizar esta abordagem em sistemas de grande porte.

O objetivo básico deste trabalho é a avaliação de índices de confiabilidade da capacidade de geração, através de uma nova abordagem que combina uma simulação Monte Carlo não-cronológica e uma técnica de amostragem por importância baseada no Método da Entropia Cruzada. Será desenvolvido um método que possa aproveitar as características positivas da simulação Monte Carlo, como a flexibilidade e a capacidade de trabalhar com sistemas de grande porte, mas sem as limitações provocadas pela raridade do evento. Espera-se que a metodologia proposta possa ser competitiva com outros métodos disponíveis na literatura. O desempenho do método proposto será verificado em diversos sistemas testes sob as mais diversas condições operativas.

1.3. ESTRUTURA DA DISSERTAÇÃO

Este trabalho apresenta uma nova metodologia, baseada em técnicas de simulação Monte Carlo e no método da Entropia Cruzada, para avaliar a confiabilidade da capacidade de geração. O principal objetivo é avaliar os índices de confiabilidade por meio da já conhecida técnica da amostragem por importância. Os parâmetros da nova função massa de probabilidade (pmf), que será utilizada no processo de amostragem, são encontrados por meio de um processo de otimização estocástico que minimiza a divergência de Kullback-Leibler entre a pmf ótima (inicialmente desconhecida) e a pmf de amostragem a cada iteração.

Uma vez posicionado o problema da avaliação da confiabilidade da capacidade de geração, feito um desenvolvimento histórico da literatura existente e definido o objetivo básico deste trabalho, cabe agora descrever em poucas linhas o conteúdo geral de cada capítulo. A presente dissertação é dividida em cinco capítulos brevemente descritos a seguir.

Este capítulo apresentou as razões pelas quais o problema de avaliar a confiabilidade da capacidade de geração vem sendo estudado na atualidade, bem como uma classificação e uma discussão geral sobre os trabalhos mais importantes relativos a este assunto, publicados na literatura especializada.

O Capítulo 2 apresenta a formulação matemática do problema, assim como um resumo dos principais métodos tradicionalmente utilizados na avaliação da confiabilidade de geração com suas respectivas vantagens e desvantagens. O principal objetivo é estabelecer fundamentos conceituais para permitir o desenvolvimento do método proposto, baseado em Entropia Cruzada, que será apresentado no capítulo subsequente.

O Capítulo 3 destina-se à apresentação do Método da Entropia Cruzada. As principais técnicas de redução de variância são brevemente discutidas com uma maior ênfase na técnica da amostragem por importância. É apresentada a formulação matemática do MEC aplicado na simulação de eventos raros. O capítulo apresenta também como este método pode ser reformulado matematicamente para ser aplicado ao problema da avaliação da confiabilidade da capacidade de geração. Com o objetivo de facilitar o entendimento do

método e também manter a formulação matemática a mais genérica possível, todo o desenvolvimento é ilustrado com base em um sistema de N_C estações geradoras.

No Capítulo 4, realizam-se aplicações numéricas com o IEEE Reliability Test System (IEEE-RTS) [IEEE79] e também com diversas modificações deste sistema. O desempenho do método proposto baseado em Entropia Cruzada é comparado com uma simulação Monte Carlo não-cronológica e também com um método analítico extremamente eficiente baseado em convolução discreta. Uma configuração do sistema de geração brasileiro da região Sul-Sudeste é também utilizada para demonstrar a aptidão da metodologia proposta em aplicações reais. Diversos testes de sensibilidade também são apresentados, principalmente no que diz respeito à raridade dos eventos de falha envolvidos e ao tamanho do sistema.

Finalmente, no Capítulo 5 serão apresentadas as principais conclusões decorrentes deste trabalho, além de sugestões para possíveis trabalhos futuros.

CAPÍTULO 2

AVALIAÇÃO DA CONFIABILIDADE DA CAPACIDADE DE GERAÇÃO

2.1. INTRODUÇÃO

NESTE capítulo serão descritos os principais métodos utilizados para a avaliação da confiabilidade da capacidade de geração. Serão brevemente discutidas as hipóteses em que eles se baseiam e suas respectivas vantagens e pontos deficientes. Antes, porém, serão feitas algumas considerações básicas sobre o problema da confiabilidade da capacidade de geração.

Uma tarefa importante no planejamento e operação de um sistema elétrico de potência é a determinação da capacidade de geração necessária para o atendimento do mercado consumidor [M86]. Este problema pode ser dividido em duas partes: determinação da capacidade estática e da capacidade operativa. A capacidade estática está relacionada com a avaliação em longo prazo das necessidades do sistema em termos globais. A capacidade operativa está relacionada com a avaliação em curto prazo da capacidade real necessária para atender um dado nível de carga.

Na avaliação da confiabilidade de capacidade estática de geração se está interessado no cálculo de índices que meçam a adequação do sistema no atendimento da carga em termos globais. Para tanto, considera-se que toda a geração e toda a carga do sistema estão concentradas em um único barramento. Isto significa supor que não existem restrições de transporte de energia e que as linhas de transmissão são 100% confiáveis. É possível admitir um nível aproximado de perdas e incluí-las na carga total do sistema.

A metodologia básica para a obtenção dos índices de adequação consiste na construção de modelos matemáticos que representem os comportamentos da geração existente e da carga do sistema. Estes modelos são a seguir combinados para se obter um modelo de risco ou de

reserva a qual expressa, em termos de probabilidades, a reserva de capacidade de geração do sistema.

2.2. FORMULAÇÃO DO PROBLEMA

Reserva, ou margem, é definida como a diferença entre a capacidade disponível e a carga. Deste modo, se os modelos de geração e de carga possuem vários estados discretos de capacidade e níveis de demanda, respectivamente, podem ser obtidas informações sobre disponibilidade, frequência e duração média para várias condições de reserva [M86]. Um estado de reserva é a combinação de um estado de carga e um estado de capacidade.

Um estado de reserva negativa representa uma condição em que a carga do sistema excede a capacidade disponível e descreve, portanto, um estado de falha do sistema. Diferentes estados individuais podem ser acumulados em um único estado equivalente. Acumulando todos os estados de reserva negativa, temos um único estado que representa as condições operativas onde o sistema se encontra em uma situação de falha. De maneira semelhante, todos os estados de reserva positiva podem ser acumulados em um único estado equivalente que representa situações de sucesso do sistema. Define-se assim, uma fronteira entre o que é considerado como *sucesso* e o que é considerado como *falha*. Estados de fronteira são todos aqueles estados que tem uma ligação direta com algum outro estado que não pertence ao seu grupo acumulado, ou seja, são aqueles estados que podem cruzar a fronteira em uma única transição.

Conforme exposto anteriormente, o principal objetivo é determinar índices que avaliem o risco do parque gerador existente não atender às necessidades do mercado consumidor. Estes índices de confiabilidade são obtidos a partir das probabilidades e frequências associadas ao estado acumulado de falha, ou seja, considerando todos os estados com reserva negativa. Os principais índices são a LOLP, EPNS e LOLF. Outros índices de confiabilidade como a LOLE, EENS e LOLD podem ser encontrados a partir dos anteriores.

O índice LOLP (*Loss of Load Probability*) é a probabilidade de existir qualquer corte de carga independentemente da magnitude do corte. Matematicamente é igual ao somatório das probabilidades individuais dos estados de reserva negativa, ou seja, onde a capacidade

de geração disponível é inferior que a carga total do sistema. A LOLE (*Loss of Load Expectation*) é o tempo esperado durante o qual o sistema reside em estados de margem negativa, considerando um período de análise T . Matematicamente é igual ao produto do índice LOLP pelo valor do intervalo T , por exemplo, um ano.

O índice EPNS (*Expected Power Not Supplied*) representa a potência esperada não suprida. O índice é calculado considerando estados de reserva negativa. Matematicamente, é igual ao somatório dos produtos das probabilidades individuais dos estados de margem negativa, pelas suas respectivas magnitudes de corte de carga. A EENS (*Expected Energy Not Supplied*), ou energia esperada não suprida, é igual ao produto do índice EPNS pelo período de análise T .

O índice LOLF (*Loss of Load Frequency*), ou frequência de perda de carga, é o número de transições ou número de vezes que o sistema saiu do estado de sucesso e foi para o estado de falha (ou vice-versa) em um período de análise T . Se o sistema for *ergódico*, um espaço de estados dividido em dois grupos, por exemplo, A e B , será balanceado em frequência, pois o número de transições de A para B é igual ao número de transições de B para A , considerando o mesmo intervalo T . No cálculo da LOLF somente os estados de fronteira com margem negativa (ou positiva) são relevantes. A LOD (*Loss of Load Duration*) é a duração média de perda de carga. É igual à razão entre os índices LOLP e LOLF.

2.3. MÉTODOS ANALÍTICOS

2.3.1. Considerações Iniciais

Como exposto anteriormente, um sistema de potência está operando satisfatoriamente enquanto houver geração disponível suficiente para atender a demanda imposta pela carga. Portanto, somente unidades geradoras são incluídas na avaliação e o resto do sistema é assumido perfeitamente confiável, com capacidade de transporte de energia ilimitada e sem perdas. É possível também admitir um nível aproximado de perdas e incluí-las na carga total do sistema.

Ao longo dos anos, diversos métodos analíticos foram surgindo, cada um deles com suas respectivas vantagens e desvantagens. As principais diferenças entre eles se encontram no

nível de detalhamento com que as unidades geradoras e as cargas são modeladas, assim como nas diversas hipóteses assumidas. Neste trabalho, o método analítico utilizado nas comparações e testes de sensibilidade é aquele descrito em [LMC91].

A referência [LMC91] apresenta um método de Frequência e Duração (F&D) em sua forma geral, i.e. as unidades geradoras são representadas por modelos de estados múltiplos e a carga é representada por níveis discretos horários. Portanto, ambos os modelos são de estados múltiplos e não necessariamente balanceados em frequência. Também foi demonstrado que não somente as probabilidades dos estados de reserva podem ser avaliadas recursivamente por técnicas de convolução discreta, mas também suas respectivas frequências incrementais. Isto possibilitou a implementação de técnicas de truncamento e arredondamento permitindo uma eficiência computacional excelente para os índices de F&D em sistemas reais.

Basicamente, o método de F&D combina, por meio de técnicas de Markov, estados de capacidade e de carga para obter estados de reserva. Uma vez que todos os estados de reserva são determinados, os principais índices de confiabilidade, LOLP, LOLE, EPNS, EENS, LOLF e LOLD são avaliados. A avaliação de índices de confiabilidade da capacidade de geração envolve três etapas básicas: (i) desenvolver um modelo adequado de capacidade de geração partindo dos parâmetros de cada unidade geradora; (ii) desenvolver um modelo adequado de carga a partir de dados históricos disponíveis em um período de análise; (iii) finalmente combinar os modelos de capacidade e carga para se obter o modelo probabilístico da reserva.

2.3.2. Modelo da Capacidade de Geração

Tendo em mente a hipótese prévia, todas as unidades de geração podem ser combinadas para produzir uma unidade equivalente G , que é então combinada com a carga total do sistema L . No método da LOLE, a unidade equivalente G pode ser descrita pela *tabela de probabilidades das capacidades indisponíveis* (COPT). No método da F&D, a frequência incremental é um parâmetro fundamental e deve ser incluído nesta tabela. Portanto, neste caso, a unidade de geração equivalente G é descrita por uma *tabela de probabilidades e frequências das capacidades indisponíveis* (COPFT) [LMC91, LCM92].

A unidade de geração equivalente G pode ser expressa pela soma de N_G variáveis aleatórias independentes G_k representando cada unidade geradora. Este somatório pode ser efetuado pelo seguinte processo recursivo, com $k = 1, \dots, N_G - 1$:

$$G'_{k+1} = G'_k + G_{k+1}, \quad (2.1)$$

onde $G'_k = \sum G_i$ (com $i = 1, k$), e o processo para quando $G'_{N_G} = G$. Para ilustrar o processo de combinação, (2.1) será simplificado para $G = G_1 + G_2$. Embora a notação seja simplificada, o desenvolvimento matemático permanece genérico, pois G , G_1 , e G_2 podem representar G'_{k+1} , G'_k , e G_{k+1} , respectivamente.

O modelo das capacidades de geração no método de F&D pode ser determinado como segue: dadas as capacidades dos estados c , probabilidades p (i.e. disponibilidades) e as frequências incrementais q de $G_1 = \{c_1; p_1; q_1\}$ e $G_2 = \{c_2; p_2; q_2\}$, determinar os mesmos parâmetros para a unidade equivalente $G = \{c_G; p_G; q_G\}$. Os parâmetros p e q são sequências de impulsos associados com as sequências de capacidades dos estados c . Ambas sequências de impulsos são igualmente espaçadas por meio de uma capacidade de arredondamento (*rounding increment*) Δ predefinida. Foi completamente demonstrado em [LMC91] que os parâmetros p e q , que caracterizam a unidade equivalente G , podem ser avaliados pelas seguintes equações de convolução (*):

$$p = p_1 * p_2, \quad (2.2)$$

$$q = [p_1 * q_2] + [q_1 * p_2]. \quad (2.3)$$

As equações (2.2) e (2.3) são então aplicadas para avaliar o processo recursivo (2.1), e no final deste procedimento, os parâmetros da unidade G são expressos pela estrutura $G = \{c_G; p_G; q_G\}$, onde, por exemplo, $p_G(g_i)$ é o i -ésimo termo de uma sequência ou do vetor p_G , com dimensão N_G , e representa a probabilidade associada com o estado de geração g_i . O processo de convolução discreta pode ser efetuado utilizando a transformada rápida de Fourier [ALAB81] para acelerar o processamento computacional.

2.3.3. Modelo da Carga do Sistema

O comportamento da carga total do sistema L pode ser expresso como uma sequência de níveis discretos de carga definidos em um período de análise, assim como ilustrado na Fig. 2.1. O *step* de discretização pode ser qualquer unidade de tempo desejada ou disponível. Também, os níveis de carga podem ser igualmente espaçados ou não. Partindo desta sequência de carga, é possível construir um modelo de carga descrito pelos mesmos parâmetros utilizados no modelo de geração, i.e. $L = \{c_L; p_L; q_L\}$, onde, por exemplo, $p_L(l_j)$ é um termo de um vetor (ou sequência) p_L , com dimensão N_L , e representa a probabilidade associada com o estado de carga l_j .

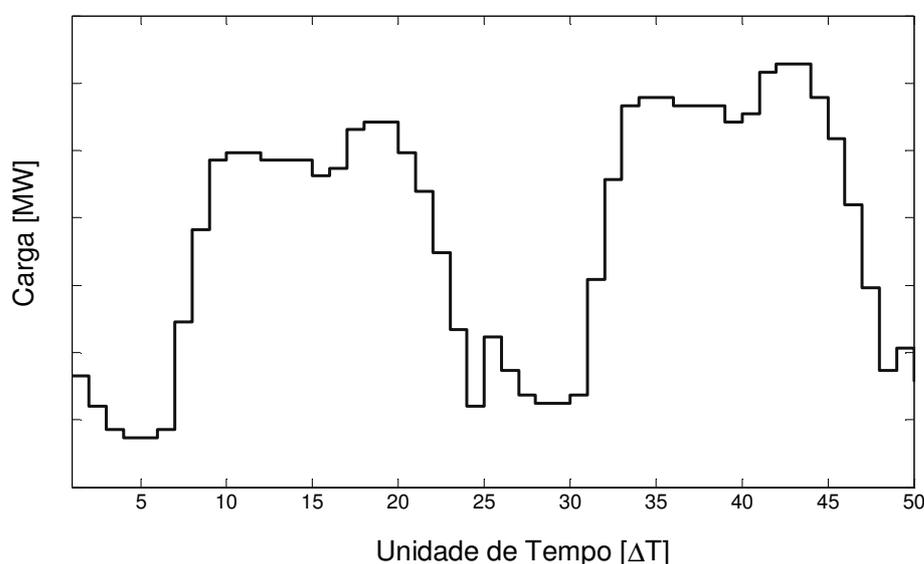


Fig. 2.1: Modelo de Carga - Sequência de Níveis Discretos de Carga

Note que, para manter a definição da frequência incremental coerente para ambos os modelos de geração e de carga, a sequência de capacidades de L (i.e. c_L) deve ser ordenada tal que o maior estado de capacidade seja associado com o nível mínimo de carga e, conseqüentemente, o menor estado de capacidade seja aquele associado com a máxima carga do sistema [LMC91].

2.3.4. Modelo da Reserva e Índices de Confiabilidade

O modelo da capacidade da geração, representado pela variável aleatória G , pode ser combinado com o modelo da carga descrito pela variável aleatória L para produzir o

modelo da reserva (estática) de potência R , i.e.

$$R = G - L, \quad (2.4)$$

com $R = \{c_R; p_R; q_R\}$, semelhante a G e L .

Os índices de confiabilidade da capacidade de geração, LOLP, EPNS, e LOLF, podem agora ser avaliados a partir dos estados individuais de reserva de geração e carga. Primeiro, as equações (2.2) e (2.3) são utilizadas para se obterem as probabilidades e as frequências incrementais que descrevem os estados da variável R . Observe que somente aqueles parâmetros do estado k de R , tal que $c_R(r_k) = c_G(g_i) - c_L(l_j) < 0$, devem ser avaliados. Considerando que existam N_R destes estados, então, para $k = 1, N_R$:

$$\text{LOLP} = \sum_k p_R(r_k), \quad (2.5)$$

$$\text{LOLF} = \sum_k q_R(r_k), \quad (2.6)$$

$$\text{EPNS} = \sum_k |c_R(r_k)| p_R(r_k). \quad (2.7)$$

Os outros três índices de confiabilidade podem ser obtidos a partir dos já previamente definidos, i.e.: $\text{LOLE} = \text{LOLP} \times T$; $\text{EENS} = \text{EPNS} \times T$; e $\text{LOLD} = \text{LOLP}/\text{LOLF}$, onde T representa o período de análise, geralmente de 8736 ou 8760 horas (um ano).

2.4. TÉCNICAS DE SIMULAÇÃO MONTE CARLO

Os métodos baseados em técnicas de Simulação Monte Carlo (SMC) são abordagens muito robustas para avaliar a confiabilidade de sistemas elétricos de potência. Modelos de SMC têm sido utilizados em uma variedade de sistemas de geração [CGOP82, PBB88]. Sem dúvida, estes métodos têm sido bastante úteis na avaliação de índices de confiabilidade, principalmente devido a sua flexibilidade e fácil implementação computacional. Outros motivos que incentivaram pesquisas nesta área são as vantagens competitivas da SMC no

tratamento de sistemas de grande porte e/ou com restrições operativas complexas. Porém, sua principal desvantagem é a necessidade (na maioria dos casos) de um tempo de processamento maior em comparação aos métodos analíticos.

O principal problema dos métodos baseados em SMC não é a quantidade de estados possíveis do sistema, mas a raridade dos eventos que se desejam capturar. Devido à utilização de um sistema de amostragem proporcional à probabilidade de ocorrência, eventos muito raros serão dificilmente amostrados, o que pode significar um problema grave, pois os esforços computacionais envolvidos nestas condições são muito expressivos. Isto quer dizer que um número muito elevado de amostras é necessário a fim de estimar as probabilidades dos eventos raros com relativa precisão, ou com intervalos de confiança estreitos.

Os métodos baseados em SMC podem ser classificados em métodos cronológicos (ou sequenciais) e métodos não-cronológicos (ou não-sequenciais). A SMC não-cronológica tem sido uma ferramenta muito útil na avaliação da confiabilidade de geração e/ou transmissão de grandes sistemas encontrados na prática [PB92, BL94, MPL92, LRMB04, SR08]. Métodos baseados em SMC cronológica são mais poderosos para avaliar a confiabilidade de sistemas de potência mais complexos, porém o esforço computacional é muito mais substancial quando comparado com os métodos não-cronológicos. Por este motivo, sua aplicação na avaliação da confiabilidade composta de sistemas de potência ainda é, em muitos casos, inviável. Para superar o problema do custo computacional das simulações cronológicas, ferramentas de SMC pseudo-cronológicas foram propostas em [LMMB00]. A seguir, cada uma destas classificações será brevemente descrita.

2.4.1. SMC NÃO-CRONOLÓGICA

Métodos como o da Enumeração de Estados e a Simulação Monte Carlo Não-Cronológica são exemplos de algoritmos baseados em espaço de estados, onde modelos markovianos são utilizados para ambas as transições entre estados dos equipamentos e da carga. Portanto, os estados são selecionados e avaliados sem considerar nenhuma conexão cronológica ou memória. Como nenhuma cronologia do sistema está sendo levada em consideração, estados não vizinhos, i.e. estados onde as taxas de transição entre eles são nulas, podem ser amostrados um depois do outro.

Em geral, algoritmos baseados em espaço de estados seguem três passos principais: (i) selecionar um estado do sistema (i.e. disponibilidade dos equipamentos e nível de carga); (ii) analisar o desempenho dos estados selecionados (i.e. verificar se a geração total disponível é capaz de satisfazer a carga associada sem violar nenhum limite operativo; se necessário, ativar medidas corretivas como, por exemplo, corte de carga); (iii) estimar índices de confiabilidade (i.e. LOLP, EPNS, etc.); se as precisões das estimativas são aceitáveis, parar o processo; caso contrário volte para o passo (i). O modelo analítico anterior é um exemplo de representação com espaço de estados, mas toda a avaliação é efetuada por meio de técnicas de convolução discreta.

Os índices GCR podem ser estimados por meio de técnicas de simulação Monte Carlo, como a média de N valores de uma função teste $H(\mathbf{Y}_i)$, considerando cada estado amostrado do sistema \mathbf{Y}_i (um vetor incluindo estados de geração e de carga), i.e.:

$$\tilde{E}[H] = \frac{1}{N} \sum_{i=1}^N H(\mathbf{Y}_i). \quad (2.8)$$

Todos os principais índices GCR podem ser representados por (2.8) dependendo da definição da função teste H . A incerteza da estimativa é dada pela variância do estimador, representado por,

$$V(\tilde{E}[H]) = V(H)/N, \quad (2.9)$$

onde $V(H)$ é a variância da função teste. Esta incerteza é geralmente representada como um coeficiente de variação denominado parâmetro β , muito utilizado como maneira de controlar a convergência da simulação. O parâmetro β é definido como,

$$\beta = \sqrt{V(\tilde{E}[H])} / \tilde{E}[H]. \quad (2.10)$$

A simulação não-cronológica pode facilmente fornecer estimativas não tendenciosas para os índices LOLP e EPNS. Neste caso, as funções testes H_{LOLP} e H_{EPNS} são dadas por [MPL92]:

$$H_{\text{LOLP}}(\mathbf{Y}_i) = \begin{cases} 0 & \text{se } \mathbf{Y}_i \in \Psi_{\text{Success}} \\ 1 & \text{se } \mathbf{Y}_i \in \Psi_{\text{Failure}} \end{cases} \quad (2.11)$$

$$H_{\text{EPNS}}(\mathbf{Y}_i) = \begin{cases} 0 & \text{se } \mathbf{Y}_i \in \Psi_{\text{Success}} \\ \Delta P_i & \text{se } \mathbf{Y}_i \in \Psi_{\text{Failure}} \end{cases} \quad (2.12)$$

onde $\Psi = \Psi_{\text{Success}} \cup \Psi_{\text{Failure}}$ é o conjunto de todos os estados possíveis \mathbf{Y}_i (i.e. o espaço de estados), dividido em dois subespaços Ψ_{Success} de estados de sucesso e Ψ_{Failure} de estados de falha; ΔP_i é a quantidade de potência cortada no estado de falha \mathbf{Y}_i .

A simulação não-cronológica pode também fornecer estimativas não tendenciosas para a LOLF. Neste caso, a função teste H_{LOLF} é dada por [MPL92]:

$$H_{\text{LOLF}}(\mathbf{Y}_i) = \begin{cases} 0 & \text{se } \mathbf{Y}_i \in \Psi_{\text{Success}} \\ \Delta \lambda_i & \text{se } \mathbf{Y}_i \in \Psi_{\text{Failure}} \end{cases} \quad (2.13)$$

onde $\Delta \lambda_i$ é a soma das taxas de transição (incluindo estados de geração e carga) entre \mathbf{Y}_i e todos os estados de sucesso, que podem ser alcançados desde \mathbf{Y}_i em uma única transição. É obvio que na avaliação de índices GCR, para cada estado selecionado $\mathbf{Y}_i \in \Psi_{\text{Failure}}$ somente as unidades indisponíveis devem ser consideradas na análise. Outro estimador não tendencioso para o índice LOLF é proposto em [LRMB04]. Finalmente, os outros três índices de confiabilidade (i.e. LOLE, EENS e LOLD) podem ser facilmente avaliados.

2.4.2. SMC SEQUENCIAL OU CRONOLÓGICA

Na subseção anterior pôde-se notar que as técnicas baseadas na representação por espaço de estados ignoram os aspectos relativos à evolução cronológica do sistema. Devido a essa restrição, a aplicação das técnicas que empregam a representação por espaço de estados limita-se a sistemas onde a dependência dos estados dos componentes com o tempo é pequena, podendo ser desprezada sem ocasionar a perda de precisão dos índices estimados.

Em uma representação cronológica, dois estados consecutivos diferem um do outro apenas pelo estado de um de seus componentes. Desta maneira, o esforço computacional requerido

é substancialmente maior que aquele apresentado pelas técnicas baseadas em representação por espaço de estados. Isso se torna mais agravante no caso da avaliação da confiabilidade composta de sistemas de potência, pois neste caso as análises de adequação dos estados envolvem a resolução de um problema de fluxo de potência ótimo, o que requer um esforço computacional elevado.

A representação cronológica permite, entretanto, uma abordagem com maior flexibilidade, pois é possível reproduzir aspectos temporais tais como diferentes padrões cronológicos de carga por área ou barra do sistema, a operação dos reservatórios, entre outros aspectos dependentes do tempo. Além disso, é possível obter as distribuições de probabilidade associadas aos índices estimados.

Essa característica torna a simulação Monte Carlo sequencial a técnica mais apropriada para lidar com sistemas nos quais a reprodução da cronologia de ocorrência dos eventos é de fundamental importância para uma correta avaliação dos índices de confiabilidade. Este é o caso dos sistemas com elevada penetração de energia eólica, ou quando se deseja avaliar a reserva operativa.

2.4.3. SMC PSEUDO-CRONOLÓGICA

A SMC cronológica é a ferramenta natural para simular os aspectos cronológicos do sistema, porém, o esforço computacional é muito elevado. A fim de superar estas dificuldades, uma técnica de SMC pseudo-cronológica foi proposta em [LMMB00]. A SMC pseudo-cronológica retém a flexibilidade e a precisão da simulação cronológica reduzindo o custo computacional.

Nesta metodologia, os estados do sistema são amostrados de maneira semelhante à SMC não-cronológica. Quando um estado de falha é encontrado, uma sequência de interrupção é obtida para este estado baseada em simulações *forward/backward* conforme apresentado em [LMMB00]. A simulação *forward* se preocupa em identificar uma possível sequência cronológica de estados de falha, partindo do estado de falha amostrado até encontrar um estado de sucesso. A simulação *backward* se preocupa em identificar uma possível sequência de estados de falha experimentados pelo sistema antes de encontrar o estado de falha amostrado, partindo deste estado até encontrar um estado de sucesso.

2.5. MÉTODOS BASEADOS EM METAHEURÍSTICAS

Mais recentemente, métodos baseados em técnicas metaheurísticas têm sido utilizados na avaliação da confiabilidade da capacidade de geração [WS08, MCRLS09]. Metaheurísticas podem ser definidas como técnicas iterativas de busca que auxiliam as heurísticas combinando diferentes conceitos para explorar (*exploring*) e aproveitar (*exploiting*) o espaço de soluções (*search space*). Uma técnica metaheurística é um método informal de solução de um problema baseado em alguma experiência prévia, ou simplesmente por bom senso, geralmente utilizado quando uma solução formal do problema não existe, ou sua implementação não é viável.

As metaheurísticas, em geral, são aplicadas para resolver problemas de otimização muito complexos, onde uma solução ótima não é necessariamente o objetivo, mas sim um conjunto de boas soluções. Exemplos de técnicas metaheurísticas que foram aplicadas na avaliação de índices GCR são: Algoritmos Genéticos (GA), Otimização por Exame de Partículas (PSO), Sistemas Imunológicos Artificiais (AIS), Otimização por Colônia de Formigas (ACO) e combinações e/ou variações dos métodos anteriormente citados.

Independentemente das particularidades da metaheurística utilizada, o algoritmo de avaliação de índices GCR segue sempre, em geral, os mesmos passos. O objetivo principal é utilizar uma metaheurística para selecionar os estados dominantes de falha do sistema, ou seja, aqueles que têm maior contribuição nos índices GCR. Diferentemente ao que ocorre na SMC, onde os estados são amostrados proporcionalmente a sua probabilidade de ocorrência, nos métodos baseados em metaheurísticas os estados são selecionados medindo os seus respectivos graus de aptidão avaliando-se uma função denominada função de *fitness*.

As metaheurísticas utilizadas são baseadas em população. O processo começa com uma população inicial aleatoriamente gerada, a seguir, o valor da função de *fitness* de cada indivíduo é avaliado baseado em uma função objetivo predefinida. Aquelos indivíduos que tem os melhores valores de *fitness* são selecionados para gerar uma nova população, seguindo um conjunto de regras definidas pela metaheurística escolhida. Este processo continua até que algum critério de parada seja atingido. Critérios de parada usualmente

utilizados são: o número máximo de gerações ou iterações, o tempo máximo de simulação, o conjunto de soluções é invariante após certo número de gerações, etc.

Uma vez que um conjunto de estados de falha sem repetição seja determinado, e este conjunto seja suficiente para ser considerado como representativo, os índices GCR podem ser calculados analiticamente considerando somente os estados selecionados. Ao contrário do que ocorre na SMC, somente os estados de falha do sistema são interessantes para o cálculo de índices e os estados de falha são mais prováveis de serem amostrados, pois apresentam maior valor de *fitness*.

Estes são métodos interessantes, porém apresentam algumas deficiências. Como estes métodos, em geral, são muito dependentes do problema em questão e da metaheurística escolhida, um esforço considerável pode ser necessário para especificar corretamente os parâmetros iniciais dos algoritmos, que serão diferentes para cada problema. O controle da convergência do processo de otimização é também complicado, pois os critérios de convergência são muito subjetivos e não estritamente definidos matematicamente. A necessidade de encontrar um vetor de resposta onde não existam estados repetidos pode ser um problema, pois um tempo considerável é gasto para verificar se um estado já forma parte da solução ou não. Finalmente, muitas informações referentes aos estados selecionados precisam ser armazenadas para posteriormente calcular os índices GCR, o que se traduz em um custo de memória muito elevado, podendo inclusive inviabilizar a aplicação destes métodos em sistemas reais.

2.6. COMENTÁRIOS FINAIS

Cada método de avaliação de índices GCR descrito neste capítulo tem suas respectivas vantagens e desvantagens. Os métodos analíticos são praticamente imbatíveis do ponto de vista computacional em sistemas pequenos, onde o número de estados possíveis não é muito elevado, porém apresentam problemas sérios quando este número aumenta mesmo com a utilização de algoritmos extremamente eficientes, como é o caso da metodologia proposta em [LMC91]. Devido a este problema, a aplicação de métodos analíticos para a avaliação da confiabilidade composta de sistemas reais é impraticável. Outra vantagem dos métodos analíticos é que, em geral, a COPT ou mesmo a COPFT do sistema é obtida como

um subproduto do algoritmo. O desempenho dos métodos analíticos não é muito sensível à raridade dos eventos de falha.

Técnicas baseadas em SMC são muito mais flexíveis que os métodos analíticos podendo simular restrições operativas complexas. Também, são muito robustos e têm grande aplicação em sistemas práticos assim como na avaliação da confiabilidade composta. O grande número de estados possíveis de um sistema não representa um problema significativo. São, em geral, de fácil programação e fornecem estimativas não tendenciosas para os índices GCR. O principal problema dos métodos baseados em SMC é o tempo de computação necessário, principalmente em simulações cronológicas. Os tempos de simulação para uma mesma precisão pré-especificada crescem exponencialmente com a raridade dos eventos envolvidos, ou seja, neste caso, a raridade do evento torna-se o principal problema.

Métodos baseados em técnicas metaheurísticas são interessantes como uma alternativa para resolver o problema, porém, têm difícil aplicação em sistemas reais devido ao grande custo de memória e tempo necessário para verificar a não repetição dos estados no vetor de respostas. No entanto, estes métodos poderiam apresentar resultados interessantes se aplicados em problemas de planejamento da expansão de sistemas elétricos de potência.

Considerando as vantagens e deficiências dos métodos anteriores, pode-se concluir que os principais problemas encontrados são: o número de estados possíveis do sistema, a raridade dos eventos de falha envolvidos, a definição de um critério de convergência adequado, o custo de memória, o tempo de simulação, entre outros. Especificamente no caso dos métodos baseados em SMC, o principal problema é a raridade dos eventos envolvidos. A fim de melhorar o desempenho destes métodos, diversas técnicas de redução de variância foram propostas, das quais, algumas delas serão tratadas no capítulo seguinte. Um novo método de avaliação de índices GCR será também apresentado que retém os pontos fortes da SMC, mas é imune a problemas decorrentes da raridade dos eventos de falha.

CAPÍTULO 3

APLICAÇÃO DO MÉTODO DA ENTROPIA CRUZADA NA CONFIABILIDADE DE GERAÇÃO

3.1. INTRODUÇÃO

CONFORME foi discutido no capítulo anterior, as técnicas baseadas em SMC são ferramentas extremamente robustas para avaliar a confiabilidade de sistemas de potência e podem ser aplicadas para as mais diversas condições operativas. Os métodos baseados em SMC apresentam como principais vantagens a flexibilidade e a facilidade de trabalhar com sistemas de grande porte, onde o número de estados possíveis é muito elevado. Porém, em sistemas muito confiáveis, o esforço computacional necessário para estimar os índices com relativa precisão pode ser expressivo. O principal motivo é a raridade dos eventos de falha, pois em sistemas muito confiáveis estes eventos têm baixíssima probabilidade de ocorrência.

A raridade dos eventos envolvidos passa a ser um problema devido à própria metodologia de amostragem na qual se baseiam os métodos de SMC. A fim de melhorar a eficiência computacional destas ferramentas, diversas técnicas de redução de variância (VRT) estão disponíveis e algumas delas serão brevemente comentadas no início deste capítulo. Dentre as principais VRT, maior atenção será conferida à denominada amostragem por importância (IS), uma técnica que tem evoluído consideravelmente nos últimos anos com a utilização de algoritmos de otimização para selecionar os elementos do vetor de parâmetros de referência.

Este capítulo tem como principal objetivo apresentar os conceitos fundamentais do Método da Entropia Cruzada (MEC) para simulação de eventos raros; um método que auxilia a IS na escolha do vetor de parâmetros de referência, e também apresentar como este método

pode ser reformulado matematicamente, partindo da sua formulação genérica, para ser aplicado no problema específico de avaliar índices de confiabilidade da capacidade de geração. Será apresentado como o MEC pode ser utilizado em conjunto com a já conhecida técnica de IS e uma SMC não-cronológica para estimar os principais índices GCR.

3.2. TÉCNICAS DE REDUÇÃO DE VARIÂNCIA

Uma técnica de redução de variância é um procedimento utilizado em conjunto com uma SMC para aumentar a precisão das estimativas que podem ser obtidas para um número definido de amostras. Cada variável aleatória obtida como resposta de uma simulação (variável aleatória de saída), tem associada uma respectiva variância que limita a precisão dos resultados obtidos. Por meio da utilização de VRT, o desempenho computacional da SMC pode ser consideravelmente melhorado, pois seu principal objetivo é fazer com que a simulação se torne estatisticamente eficiente, e assim, os intervalos de confiança das estimativas sejam os mais estreitos possíveis. Algumas das VRT mais utilizadas são brevemente comentadas a seguir.

3.2.1. Variáveis Antitéticas

A idéia central por trás das Variáveis Antitéticas baseia-se na intuição de que valores extremos amostrados podem ser menos impactantes nas estimativas, se outros valores extremos são utilizados para contrabalancear os efeitos provocados por eles [F09]. No caso mais simples, considere um vetor de comprimento $N/2$ de amostras X que seguem uma distribuição normal, e também o vetor $-X$ a fim de completar as N amostras. Note que, se as amostras de X seguem uma distribuição normal com média μ e variância σ^2 , as amostras de $-X$ seguiram uma distribuição normal com os mesmos parâmetros μ e σ^2 .

Mesmo que $-X$ não seja independente de X , é possível obter melhores estimativas para média e variância da distribuição por esta abordagem considerando um N predefinido. Essencialmente, é possível gerar um conjunto de amostras, e complementar este conjunto com outro que segue a mesma distribuição de probabilidade. Por exemplo, uma amostra X de x que segue $U(0,1)$, pode ser complementada por $(1-X)$, que também segue $U(0,1)$. Algumas aplicações desta VRT em sistemas de potência podem ser encontradas em [BJ96, BJ97].

3.2.2. Variáveis de Controle

A técnica conhecida como Variáveis de Controle procura utilizar a correlação entre duas variáveis aleatórias para se obter uma redução na variância [BJ97]. Se y é uma variável aleatória que tem uma correlação positiva ou negativa com x e tem um valor esperado conhecido $v = E[y]$, então a variável aleatória $x_c = x - a(y - v)$ é um estimador não tendencioso de μ para qualquer número real a , onde a é conhecido como coeficiente de controle e μ é o valor desconhecido da média da variável aleatória x .

O caso especial onde $a = 1$ tem sido utilizado em muitas aplicações desta técnica na avaliação da confiabilidade de sistemas de potência. A utilização de +1 ou -1 para a variável a implica que o sucesso desta técnica dependa de y , que é a variável de controle. A utilização de outros valores para o coeficiente a pode fazer a variância reduzir mais rapidamente.

Os limites para o coeficiente de controle e seu valor ótimo dependem da correlação entre a saída x e a entrada y . Um método para estimar o coeficiente de controle ótimo pode ser encontrado em [BJ97]. Uma formulação matemática formal do método pode ser encontrada em [F09]. Exemplos de aplicações desta técnica em sistemas de potência podem ser encontrados em [PAK08, BJ96, BJ97].

3.2.3. Números Aleatórios em Comum

A técnica de Números Aleatórios em Comum, também é conhecida como amostragem correlacionada, é uma VRT utilizada muito conhecida e útil quando duas ou mais configurações diferentes para um mesmo sistema estão sendo comparadas como, por exemplo, na teoria das filas. Esta VRT requer uma sincronização dos vetores de números aleatórios para garantir que os mesmos números aleatórios sejam utilizados para simular todas as configurações. Além disso, também é necessário que cada número aleatório seja utilizado sempre para o mesmo objetivo específico independentemente da configuração simulada.

Este método é muito simples e produz reduções na variância, porém, estes ganhos na maioria dos casos não são significativos. Outras técnicas de redução de variância citadas

nesta dissertação provaram ser muito mais poderosas. Um exemplo de aplicação desta VRT pode ser encontrado em [C90].

3.2.4. Amostragem Estratificada

Estratificação é o processo de agrupar membros de uma população em subgrupos relativamente homogêneos, ou seja, que apresentem aproximadamente as mesmas características. Os estratos devem ser mutuamente exclusivos e coletivamente exaustivos, i.e. cada amostra deve pertencer unicamente a um estrato e os estratos em conjunto devem representar toda a população. A seguir uma amostragem aleatória ou sistemática é aplicada em cada estrato.

Esta VRT, em geral, melhora a representatividade da amostra e pode produzir uma média ponderada que tenha menor variância que a média aritmética de uma amostragem aleatória da população. As principais vantagens deste método são a possibilidade de concentrar os esforços nos estratos importantes e a possibilidade de utilizar diferentes técnicas de amostragem para cada estrato, porém, o método requer um bom conhecimento do comportamento do sistema, não é útil se todo o sistema é relativamente homogêneo e a definição dos estratos muitas vezes é complicada. Uma aplicação desta técnica em sistemas de potência pode ser encontrada em [PAK08].

3.2.5. Amostragem por Importância

A técnica de amostragem por importância é uma VRT que pode ser utilizada em conjunto com a SMC. A idéia central por trás desta técnica é que certos valores da variável aleatória de entrada têm maior impacto, em comparação a outros, no parâmetro que se deseja estimar via SMC. Se estes valores ditos importantes são enfatizados amostrando-os com maior frequência, então a variância do estimador será reduzida. Portanto, a metodologia básica da IS é escolher uma distribuição que estimule o aparecimento destes valores significativos.

A aplicação direta desta nova distribuição na SMC resultaria em um estimador tendencioso, porém os resultados da simulação são ponderados para compensar este efeito e desta maneira garantir que o estimador final, denominado estimador de IS, seja não-

tendencioso. A ponderação é definida pela utilização da *razão de verossimilhança*, que será apresentada mais adiante.

O problema fundamental encontrado durante a implementação da IS, conforme será visto na seção seguinte, é a escolha correta desta nova distribuição de amostragem. Escolher, estimar, ou determinar esta nova distribuição é considerada como a *arte* da IS. Uma escolha correta trará como recompensa uma enorme economia em esforço computacional, porém, uma escolha equivocada será penalizada com tempos de simulação maiores aos que seriam necessários se uma SMC convencional fosse diretamente utilizada. Várias maneiras de resolver este problema foram propostas em [LNR99, ZX02, PAK08, QZW08]. Dos métodos apresentados nestas referências, a maioria deles recai em resolver algum tipo de problema de otimização, tal que a variância do estimador de IS seja minimizada.

Outra maneira de estimar a nova distribuição de amostragem foi proposta por Homem-de-Mello e Rubinstein em 2002 [HR02]. Neste trabalho, um procedimento rápido de dois estágios bem definidos foi utilizado para estimar probabilidades de eventos raros. Este procedimento era baseado em IS e a divergência de Kullback-Leibler. Os conceitos apresentados neste trabalho, mais algumas outras contribuições, foram posteriormente organizados por Rubinstein dando origem ao denominado Método da Entropia Cruzada. A formulação matemática formal deste poderoso método será sintetizada na seguinte seção.

3.3. O MÉTODO DA ENTROPIA CRUZADA

O Método da Entropia Cruzada, atribuído a Reuven Y. Rubinstein, é uma abordagem genérica da simulação Monte Carlo para problemas de otimização combinatória e amostragem por importância. O método teve suas origens na área de simulação de eventos raros, onde probabilidades muito pequenas precisam ser estimadas com precisão, por exemplo, na análise da confiabilidade de redes, na teoria das filas, ou na análise de desempenho de sistemas de telecomunicação.

Conforme exposto na seção anterior, a técnica de amostragem por importância é baseada na idéia de fazer com que a ocorrência de *eventos significativos* seja mais frequente, ou em outras palavras, acelerar o processo de convergência da simulação. Estes eventos

significativos são aqueles que têm maior contribuição nos parâmetros sendo estimados pela SMC. Tecnicamente, a IS procura selecionar uma distribuição de probabilidade diferente da original tal que a variância do seu estimador seja minimizada. A eficiência desta técnica depende da obtenção desta nova distribuição ou, pelo menos, uma próxima a ela. O principal problema é que esta nova distribuição é inicialmente desconhecida e geralmente difícil de ser encontrada.

Em geral, é útil escolher a distribuição de IS dentro da mesma família de funções densidade de probabilidade (pdf) da distribuição original. Neste caso, ambas as funções têm a mesma distribuição de probabilidade e pertencem à mesma família de pdf, sendo o conjunto de parâmetros que as descrevem a única diferença entre elas. Assim, o problema se resume em encontrar o melhor conjunto de parâmetros para esta nova pdf, denominado *vetor de parâmetros de referência ótimo*.

O Método da Entropia Cruzada fornece um processo iterativo para estimar os parâmetros de referência ótimos, sendo esta sua principal vantagem. Isto é possível minimizando a *distância* entre a pdf ótima de amostragem e a pdf de IS [BKMR05]. A seguir, os fundamentos básicos do MEC aplicado na simulação de eventos raros serão apresentados.

3.3.1. Formulação Matemática Básica

Seja $\mathbf{X} = [x_1, x_2, \dots, x_n]$ um vetor aleatório que assume valores em certo espaço \mathcal{X} . Seja $\{f(\cdot; \xi)\}$ uma família de funções de densidade de probabilidade no espaço \mathcal{X} , onde ξ é um vetor real de parâmetros. Para qualquer função mensurável H temos,

$$E_{\xi}[H(\mathbf{X})] = \int_{\mathcal{X}} H(x) f(x; \xi) dx. \quad (3.1)$$

Seja $S(\mathbf{X})$ uma função real no espaço \mathcal{X} denominada *função desempenho* e γ um número real. Considere que o objetivo seja estimar a probabilidade ℓ tal que $S(\mathbf{X}) \geq \gamma$, sob $f(\cdot; \mathbf{u})$. Esta probabilidade pode ser expressa como,

$$\ell = P_{\mathbf{u}}\{S(\mathbf{X}) \geq \gamma\} = E_{\mathbf{u}}[I_{\{S(\mathbf{X}) \geq \gamma\}}]. \quad (3.2)$$

Neste caso, $H(\mathbf{X}) = I_{\{\text{expressão lógica}\}}$ é uma *função teste*, cujo valor é igual a um se {expressão lógica} tem valor booleano *verdadeiro* e zero caso contrário. Se esta probabilidade ℓ for muito pequena, por exemplo, menor que 10^{-5} , o evento $\{S(\mathbf{X}) \geq \gamma\}$ pode ser considerado como um evento raro.

Uma maneira de estimar diretamente o valor de ℓ é utilizar a simulação Monte Carlo convencional, i.e. retirar uma amostra aleatória $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ considerando $f(\cdot; \mathbf{u})$ e, a seguir, utilizar o estimador não tendencioso (3.3) de ℓ ,

$$\hat{\ell} = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \gamma\}}. \quad (3.3)$$

Entretanto, isto implica problemas sérios quando $\{S(\mathbf{X}) \geq \gamma\}$ é um evento raro. Se este for o caso, um grande esforço computacional será necessário a fim de se estimar ℓ com relativa precisão, i.e. com um erro relativo pequeno ou com um intervalo de confiança estreito. Isto é devido ao grande número de amostras necessárias.

Uma alternativa para contornar este problema é a técnica da amostragem por importância, i.e. retirar uma amostra aleatória $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ de uma função densidade de probabilidade g , diferente de f , em \mathcal{X} , e avaliando ℓ utilizando o *estimador da razão de verossimilhança* (3.4),

$$\hat{\ell} = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \gamma\}} \frac{f(\mathbf{X}_i; \mathbf{u})}{g(\mathbf{X}_i)}. \quad (3.4)$$

É conhecido que a melhor maneira de estimar ℓ é utilizar uma *mudança de medida* com uma densidade dada por,

$$g_{opt}(x) = \frac{I_{\{S(x) \geq \gamma\}} f(x; \mathbf{u})}{\ell}. \quad (3.5)$$

De fato, se utilizando esta *mudança de medida* em (3.4),

$$\frac{I_{\{S(\mathbf{X}_i) \geq \gamma\}} f(\mathbf{X}_i; \mathbf{u})}{g_{opt}(\mathbf{X}_i)} = \ell, \quad (3.6)$$

para todo i . Em outras palavras, o estimador (3.4) teria variância zero, então seria necessário gerar $N = 1$ amostra. A dificuldade evidente é que esta nova função densidade de probabilidade g_{opt} depende do valor desconhecido de ℓ , que é justamente o valor que se deseja estimar.

Obter esta *mudança de medida ótima*, ou pelo menos uma quase ótima, partindo da pdf de amostragem original f , até encontrar a nova pdf g_{opt} , é o principal desafio. A eficiência da amostragem por importância é dependente da escolha correta desta nova pdf. Uma escolha inadequada de g poderia ter o efeito contrário e causar com que a variância do estimador (3.4) sob g seja maior do que seria sob a pdf original f [QZW07]. A idéia é agora escolher g dentro de $\{f(\cdot; \mathbf{u})\}$, i.e. g pertence à mesma família de funções densidade de probabilidade de f . Em outras palavras, é preciso encontrar um novo vetor de parâmetros de referência \mathbf{v} tal que a *distância* entre as funções de densidade de probabilidade g_{opt} e $f(\cdot; \mathbf{v})$ seja mínima.

Na teoria de probabilidades e na teoria da informação, a *divergência de Kullback-Leibler* (também conhecida como divergência da informação, ganho de informação ou entropia relativa) é uma medida não comutativa do afastamento ou da divergência entre duas funções densidade f_1 e f_2 . A divergência de Kullback-Leibler é definida como,

$$D_{KL}(f_1 \parallel f_2) = E_{f_1} \left[\ln \frac{f_1(\mathbf{X})}{f_2(\mathbf{X})} \right] = \int f_1(x) \ln f_1(x) dx - \int f_1(x) \ln f_2(x) dx, \quad (3.7)$$

onde o primeiro termo é denominado de *entropia* de f_1 e o segundo termo é denominado de *entropia cruzada* de f_1 e f_2 . Uma definição semelhante pode ser obtida no caso de variáveis aleatórias discretas, trocando as funções densidade por funções de massa (pmf) e as integrais por somatórios [HR02]. Mesmo sendo usualmente entendida como uma

métrica de distância, a divergência de Kullback-Leibler não é uma verdadeira medida de distância, pois ela não é simétrica (por isso utiliza-se o termo *divergência* em vez de *distância*).

Minimizar a divergência de Kullback-Leibler entre g_{opt} e $f(\bullet; \mathbf{v})$ em (3.5) é equivalente a escolher \mathbf{v} tal que a entropia cruzada de g_{opt} e $f(\bullet; \mathbf{v})$ seja mínima [BKMR05]. Isto é equivalente a encontrar o máximo da entropia cruzada com sinal invertido, resultando no problema de maximização (3.8),

$$\max_{\mathbf{v}} \int g_{opt}(x) \ln f(x; \mathbf{v}) dx. \quad (3.8)$$

Substituindo $g_{opt}(x)$ de (3.5) em (3.8), o seguinte problema de otimização é obtido,

$$\max_{\mathbf{v}} \int \frac{I_{\{S(x) \geq \gamma\}} f(x; \mathbf{u})}{\ell} \ln f(x; \mathbf{v}) dx, \quad (3.9)$$

que é equivalente ao problema,

$$\max_{\mathbf{v}} D(\mathbf{v}) = \max_{\mathbf{v}} E_{\mathbf{u}}[I_{\{S(\mathbf{X}) \geq \gamma\}} \ln f(\mathbf{X}; \mathbf{v})], \quad (3.10)$$

onde $D(\mathbf{v})$ está implicitamente definido em (3.10). Utilizando amostragem por importância novamente, com uma *mudança de medida* $f(\bullet; \boldsymbol{\omega})$, é possível escrever (3.10) como,

$$\max_{\mathbf{v}} D(\mathbf{v}) = \max_{\mathbf{v}} E_{\boldsymbol{\omega}}[I_{\{S(\mathbf{X}) \geq \gamma\}} \mathbf{W}(\mathbf{X}; \mathbf{u}, \boldsymbol{\omega}) \ln f(\mathbf{X}; \mathbf{v})], \quad (3.11)$$

para qualquer vetor de parâmetros $\boldsymbol{\omega}$, onde,

$$\mathbf{W}(x; \mathbf{u}, \boldsymbol{\omega}) = \frac{f(x; \mathbf{u})}{f(x; \boldsymbol{\omega})} \quad (3.12)$$

é a *razão de verossimilhança* em x , entre $f(\bullet; \mathbf{u})$ e $f(\bullet; \boldsymbol{\omega})$. A solução ótima de (3.11) pode ser escrita como,

$$\mathbf{v}^* = \underset{\mathbf{v}}{\operatorname{argmax}} E_{\boldsymbol{\omega}} [I_{\{S(\mathbf{X}) \geq \gamma\}} \mathbf{W}(\mathbf{X}; \mathbf{u}, \boldsymbol{\omega}) \ln f(\mathbf{X}; \mathbf{v})]. \quad (3.13)$$

É possível estimar \mathbf{v}^* solucionando o seguinte problema de otimização estocástico, também conhecido como *correspondente estocástico* de (3.11),

$$\hat{\mathbf{v}}^* = \max_{\mathbf{v}} \hat{D}(\mathbf{v}) = \max_{\mathbf{v}} \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \gamma\}} \mathbf{W}(\mathbf{X}_i; \mathbf{u}, \boldsymbol{\omega}) \ln f(\mathbf{X}_i; \mathbf{v}), \quad (3.14)$$

onde $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ é uma amostra aleatória considerando $f(\bullet; \boldsymbol{\omega})$. Em aplicações típicas a função $\hat{D}(\mathbf{v})$ em (3.14) é convexa e diferenciável em relação a \mathbf{v} e, portanto, a solução pode ser diretamente obtida resolvendo (em relação a \mathbf{v}) o seguinte sistema de equações,

$$\frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \gamma\}} \mathbf{W}(\mathbf{X}_i; \mathbf{u}, \boldsymbol{\omega}) \nabla \ln f(\mathbf{X}_i; \mathbf{v}) = 0, \quad (3.15)$$

onde o gradiente é em relação a \mathbf{v} .

A principal vantagem do MEC é que a solução de (3.15) pode ser frequentemente calculada analiticamente. Em particular, isto acontece se as distribuições das variáveis aleatórias envolvidas pertencem à família natural de exponenciais (NEF) ou no caso de distribuições com domínio finito. Demonstrações matemáticas formais podem ser encontradas em [HR02].

Minimizar a variância do estimador (3.4) é equivalente a minimizar o valor esperado da razão de verossimilhança condicionado à ocorrência do evento raro [QZW08]. Note que minimizar a entropia cruzada é equivalente a minimizar o valor esperado do *logaritmo natural* da razão de verossimilhança, condicionado à ocorrência do evento raro. Isto é em certa maneira próximo mas não idêntico a minimizar a variância do estimador. Ou seja, em termos da variância do estimador, o MEC não busca a solução ótima. Porém, buscar

minimizar a variância do estimador leva a tentar resolver um problema de otimização estocástico complicado e que, na maioria dos casos, não tem convexidade podendo levar a ótimos locais. Contrariamente, o MEC recai em problemas que na maioria dos casos são convexos facilitando o encontro da solução, e em muitas aplicações a solução pode ser encontrada analiticamente [HR02].

3.3.2. Abordagem de Níveis Múltiplos para Eventos Raros

É importante notar que a solução via (3.14) é útil somente quando $P_{\mathbf{u}}\{S(\mathbf{X}) \geq \gamma\}$ não representa um valor muito pequeno, i.e. $\ell \geq 10^{-5}$. Entretanto, para probabilidades de eventos raros (quando, por exemplo, $\ell < 10^{-5}$), a abordagem descrita por (3.14) é difícil de ser efetuada. De fato, devido à raridade dos eventos, a maior parte dos valores de $I_{\{S(\mathbf{X}_i) \geq \gamma\}}$, $i=1,2,\dots,N$, serão iguais a zero, para um N moderado [BKMR05].

Um algoritmo de níveis múltiplos pode ser utilizado para superar esta dificuldade. A idéia é construir uma sequência de parâmetros de referência $\{\mathbf{v}_k, k \geq 0\}$ e uma sequência de níveis $\{\gamma_k, k \geq 1\}$, e iterativamente achar os valores de \mathbf{v}_k e γ_k . Para isto, é inicializada uma variável ρ denominada *parâmetro de níveis múltiplos* com um valor, por exemplo, de $\rho = 10^{-2}$ sendo definido $\mathbf{v}_0 = \mathbf{u}$. A seguir, se faz $\gamma_1 < \gamma$ tal que, sob a função de densidade de probabilidade original $f(\bullet; \mathbf{u})$, a probabilidade de $\ell_1 = E_{\mathbf{u}}[I_{\{S(\mathbf{X}_i) \geq \gamma_1\}}]$ seja pelo menos ρ . O vetor \mathbf{v}_1 será igual ao vetor de parâmetros de referência ótimo para estimar ℓ_1 , e estes dois últimos passos são repetidos iterativamente com o objetivo de se estimar o par $\{\ell, \mathbf{v}^*\}$.

Em outras palavras, cada iteração do algoritmo consiste em duas fases principais. Na primeira fase, γ_k é atualizado, e na segunda \mathbf{v}_k é atualizado. Especificamente, começando com $\mathbf{v}_0 = \mathbf{u}$, γ_k é encontrado como segue.

Para um dado \mathbf{v}_{k-1} , seja γ_k igual ao $(1 - \rho)$ quantil de $S(\mathbf{X})$ sob \mathbf{v}_{k-1} . Ou seja, γ_k satisfaz,

$$P_{\mathbf{v}_{k-1}}\{S(\mathbf{X}) \geq \gamma_k\} \geq \rho, \quad (3.16a)$$

$$P_{\mathbf{v}_{k-1}}\{S(\mathbf{X}) \leq \gamma_k\} \geq (1 - \rho), \quad (3.16b)$$

onde $\mathbf{X} \sim f(\bullet; \mathbf{v}_{k-1})$. Uma estimativa simples $\hat{\gamma}_k$ de γ_k pode ser obtida retirando-se uma amostra $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ aleatoriamente de $f(\bullet; \mathbf{v}_{k-1})$, calculando-se os desempenhos $S(\mathbf{X}_i)$ para todo i , ordenando estes valores do menor ao maior, tal que $S_{[1]} \leq S_{[2]} \leq \dots \leq S_{[N]}$ e finalmente avaliando o $(1 - \rho)$ quantil amostral como,

$$\hat{\gamma}_k = S_{[(1-\rho)N]}. \quad (3.17)$$

Portanto, na primeira iteração e começando com $\mathbf{v}_0 = \mathbf{u}$, para se obter uma boa estimativa de $\hat{\mathbf{v}}_1$, a raridade do evento objetivo é artificialmente diminuída (temporariamente) utilizando-se um nível $\hat{\gamma}_1$ que é escolhido menor que γ . O valor de $\hat{\mathbf{v}}_1$ obtido desta maneira fará o evento $\{S(\mathbf{X}) \geq \gamma\}$ ser menos raro na próxima iteração, de tal maneira que na próxima iteração, um valor $\hat{\gamma}_2$ mais próximo de γ poderá ser utilizado. O algoritmo termina quando, para alguma iteração k , um nível $\hat{\gamma}_k$ é atingido que seja pelo menos γ e, conseqüentemente, o valor original de γ pode ser utilizado sem a preocupação de se obter poucas *amostras de elite*.

3.3.3. Algoritmo do MEC para Simulação de Eventos Raros

Até aqui, um algoritmo bem genérico do método da entropia cruzada para simulação de eventos raros pode ser resumido como segue.

Passo 1. Definir $\hat{\mathbf{v}}_0 := \mathbf{u}$. Fazer $k := 1$ (contador de iterações).

Passo 2. Gerar uma amostra aleatória $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ de acordo com a pdf $f(\bullet; \hat{\mathbf{v}}_{k-1})$. Calcular os desempenhos $S(\mathbf{X}_i)$ para todo i , e ordenar os valores do menor ao maior tal

que, $S_{[1]} \leq S_{[2]} \leq \dots \leq S_{[N]}$. Fazer $\hat{\gamma}_k$ igual ao $(1-\rho)$ quantil amostral de desempenhos, ou seja, $\hat{\gamma}_k := S_{[(1-\rho)N]}$, caso este valor seja menor que γ . Caso contrário, $\hat{\gamma}_k := \gamma$.

Passo 3. Utilizar a mesma amostra para calcular o vetor de parâmetros da iteração k . Para cada j , com $j=1,2,\dots,J$, utilizar a expressão (3.18), onde J é o número de variáveis aleatórias envolvidas.

$$\hat{\mathbf{v}}_{k,j} = \frac{\sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \hat{\gamma}_k\}} \mathbf{W}(\mathbf{X}_i; \mathbf{u}, \hat{\mathbf{v}}_{k-1}) \mathbf{X}_{ij}}{\sum_{i=1}^N I_{\{S(\mathbf{X}_i) \geq \hat{\gamma}_k\}} \mathbf{W}(\mathbf{X}_i; \mathbf{u}, \hat{\mathbf{v}}_{k-1})} \quad (3.18)$$

Passo 4. Se $\hat{\gamma}_k = \gamma$, então ir para o *Passo 5*; caso contrário, fazer $k := k + 1$ e voltar para o *Passo 2*.

Passo 5. Seja T a última iteração. Gerar uma amostra $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{N_1}$ de acordo com $f(\cdot; \hat{\mathbf{v}}_T)$ e estimar ℓ utilizando o estimador de amostragem por importância,

$$\hat{\ell} = \frac{1}{N_1} \sum_{i=1}^{N_1} I_{\{S(\mathbf{X}_i) \geq \gamma\}} \mathbf{W}(\mathbf{X}_i; \mathbf{u}, \hat{\mathbf{v}}_T). \quad (3.19)$$

Note que nos *Passos 2-4* o vetor de parâmetros de referência ótimo a ser utilizado na amostragem por importância é estimado. No passo final, este vetor de parâmetros de referência é utilizado para estimar a probabilidade do evento objetivo. É preciso fornecer o valor de ρ (usualmente entre 0,01 e 0,1) e os parâmetros N e N_1 a priori.

Também, é possível demonstrar que se as funções de densidade de probabilidade das variáveis aleatórias envolvidas pertencem à família natural de exponenciais, a solução do sistema de equações apresentado em (3.15) sempre se torna semelhante a (3.18) [BKMR05, RK04].

3.3.4. Parâmetro de Suavização α

Em problemas envolvendo variáveis aleatórias discretas é frequentemente vantajoso a utilização de um *parâmetro de suavização* α , que tem por objetivo acrescentar um efeito inercial no processo de atualização dado por (3.18). O desempenho em certos casos é melhor, principalmente quando se deseja estimar a probabilidade de eventos extremamente raros. Utilizando-se o parâmetro α se evita o aparecimento de zeros e uns no vetor de parâmetros de referência, o que pode trazer complicações, pois sempre que isso acontece o sistema, ou o problema em questão, é essencialmente modificado. Este processo de atualização com efeito inercial pode ser escrito como,

$$\hat{\mathbf{v}}_{k,j} = \hat{\mathbf{v}}_{k,j}\alpha + \hat{\mathbf{v}}_{k-1,j}(1-\alpha), \quad (3.20)$$

onde $\hat{\mathbf{v}}_{k,j}$ é o valor resultante da expressão (3.18) na iteração k , $\hat{\mathbf{v}}_{k-1,j}$ é o valor resultante da expressão (3.20) na iteração $k-1$ e α é o denominado parâmetro de suavização que usualmente assume valores reais entre 0,5 e 0,95. Claramente, para $\alpha = 1$, se tem a regra de atualização (3.18) original.

3.4. APLICAÇÃO DO MEC NA AVALIAÇÃO DE ÍNDICES DE CONFIABILIDADE

Nesta seção será apresentada uma nova metodologia para avaliar os índices de confiabilidade da capacidade de geração utilizando o MEC. Os principais conceitos do MEC aplicado na simulação de eventos raros foram apresentados na seção 3.3. A idéia da metodologia proposta é utilizar o algoritmo apresentado na seção anterior e adaptá-lo para o problema específico de estimar os índices de confiabilidade em um sistema de geração de barra única. As estações geradoras seguem distribuições binomiais e são compostas por diversas unidades independentes com características idênticas. O modelo da carga pode ser representado por um nível constante ao longo do período de análise ou por níveis discretos diferentes que variam a cada ΔT , por exemplo, de uma hora.

Dependendo do grau de confiabilidade do sistema, uma tentativa de se estimar os índices diretamente por SMC não-cronológica pode ser inviável. Em sistemas muito confiáveis

seria necessário um número muito elevado de amostras para se estimar os eventos raros envolvidos, o que se traduz em um grande esforço computacional. Utilizando amostragem por importância, é possível selecionar uma função densidade de probabilidade diferente da original, e amostrar os estados do sistema considerando esta nova distribuição. A idéia é encontrar uma nova pdf tal que os estados de falha sejam encontrados com maior frequência e, desta maneira, acelerar o processo de convergência dos principais índices de confiabilidade.

O método proposto apresenta duas etapas bem diferenciadas. Em um primeiro momento, o algoritmo do MEC para simulação de eventos raros será utilizado para se estimar um vetor de parâmetros de referência ótimo ou, pelo menos, quase ótimo. Nesta etapa, as disponibilidades das máquinas individuais serão *distorcidas* de tal maneira a aumentar a frequência com que os estados de falha são encontrados. Uma vez determinada esta *distorção ótima* para o sistema, i.e. uma *regra de amostragem ótima*, os índices LOLP, LOLE, EPNS, EENS, LOLF e LOLD são estimados utilizando-se a já conhecida técnica da amostragem por importância. A convergência dos índices será controlada pelo valor do parâmetro β especificado.

O método proposto denominado *CE-ISMC (Cross-Entropy-Importance Sampling Monte Carlo)*, será aplicado em diversos sistemas testes padrões, e seu desempenho será discutido e comparado com outros métodos tradicionalmente utilizados na avaliação da confiabilidade de geração. Os principais resultados destas comparações, dos testes de sensibilidade e das diversas simulações serão apresentados no capítulo seguinte.

3.4.1. Reformulação do Problema

Considere um sistema de geração composto por N_C estações geradoras conforme ilustrado na Fig. 3.1. Também, considere que a j -ésima estação (GS_j) está composta por n_j unidades idênticas e independentes, cada uma com capacidade C_j . O comportamento de cada unidade individual segue uma distribuição de Bernoulli com indisponibilidades u_j .

Sob estas considerações, cada estação geradora é uma variável aleatória binomial com uma função massa de probabilidade $f(\bullet; n_j, u_j)$. Note que, na definição da função massa,

indisponibilidades são utilizadas em vez de disponibilidades. Considere também que a carga do sistema é igual a L com um nível constante ao longo do período de análise.

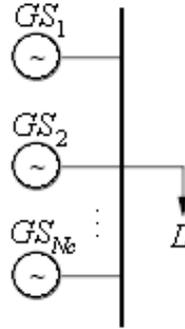


Fig. 3.1: Sistema de Geração - Barra Única

O problema analítico de avaliar o índice LOLP pode ser descrito por,

$$\text{LOLP} = E_{\mathbf{u}}[H_{\text{LOLP}}(\mathbf{X})] = \sum_{\mathbf{X}_k \in \Omega} H_{\text{LOLP}}(\mathbf{X}_k) f(\mathbf{X}_k; \mathbf{n}, \mathbf{u}), \quad (3.21)$$

onde:

\mathbf{X}_k é um possível resultado do vetor aleatório $\mathbf{X} = [x_1, x_2, \dots, x_j, \dots, x_{N_c}]$ que pertence ao espaço de estados Ω , cujo elemento genérico x_j representa o número de unidades disponíveis na estação geradora GS_j , com $x_j \in \mathbb{Z}$ e $0 \leq x_j \leq n_j$;

LOLP (*Loss of Load Probability*) é o somatório das probabilidades de todos os estados de Ω onde a capacidade de geração é insuficiente para atender L ;

$H_{\text{LOLP}}(\mathbf{X}_k) = I_{\{S(\mathbf{X}_k) < L\}}$, conhecida como função teste para o índice LOLP. $H_{\text{LOLP}}(\mathbf{X}_k) = 1$, se $\{S(\mathbf{X}_k) < L\}$ é verdadeiro e $H_{\text{LOLP}}(\mathbf{X}_k) = 0$, caso contrário. Esta função teste é similar à definida em (2.11), mas foi adaptada conforme a notação utilizada em [RK04] e [RK07];

$f(\mathbf{X}_k; \mathbf{n}, \mathbf{u})$ é a probabilidade ou proporção do tempo que o sistema reside em \mathbf{X}_k considerando a função massa de probabilidade associada a \mathbf{X} , $f(\bullet; \mathbf{n}, \mathbf{u})$;

$\mathbf{u} = [u_1, u_2, \dots, u_j, \dots, u_{N_c}]$, é um vetor que contém os parâmetros originais (indisponibilidades) do sistema com $u_j \in \mathbb{R}$ e $0 < u_j < 1$;

$\mathbf{n} = [n_1, n_2, \dots, n_j, \dots, n_{N_c}]$ é um vetor constante com o número de unidades existentes em cada estação geradora GS_j , com $n_j \in \mathbb{Z}$ e $n_j > 0$.

A expressão $S(\mathbf{X}_k)$ representa uma função desempenho no espaço de estados Ω . Neste caso $S(\mathbf{X}_k)$ é basicamente o somatório de todas as capacidades de geração disponíveis associadas com o estado \mathbf{X}_k dado por,

$$S(\mathbf{X}_k) = \mathbf{X}_k \mathbf{C}^T = [x_1, x_2, \dots, x_j, \dots, x_{N_c}] [C_1, C_2, \dots, C_j, \dots, C_{N_c}]^T. \quad (3.22)$$

Como todas as variáveis aleatórias envolvidas seguem uma distribuição binomial, o valor da função massa de probabilidade definida em \mathbf{X}_k sob $f(\bullet; \mathbf{n}, \mathbf{u})$ é da forma,

$$f(\mathbf{X}_k; \mathbf{n}, \mathbf{u}) = \prod_{j=1}^{N_c} \frac{n_j!}{x_j!(n_j - x_j)!} (1 - u_j)^{x_j} (u_j)^{n_j - x_j}. \quad (3.23)$$

Considere que o objetivo seja estimar o valor do índice LOLP por simulação, tal que,

$$\text{LOLP} = P_{\mathbf{u}}\{S(\mathbf{X}) < L\} = E_{\mathbf{u}}[I_{\{S(\mathbf{X}) < L\}}], \quad (3.24)$$

i.e. a probabilidade da capacidade de geração disponível ser insuficiente para atender L .

A fim de estimar o índice LOLP em (3.24), uma simulação Monte Carlo não-cronológica pode ser utilizada, i.e. retirar uma amostra $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ onde $\mathbf{X} \sim f(\bullet; \mathbf{n}, \mathbf{u})$, e em seguida utilizar o estimador não tendencioso para o índice LOLP,

$$\hat{\text{LOLP}} = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) < L\}}. \quad (3.25)$$

No entanto, se o nível de carga L e/ou as indisponibilidades u_j são pequenas, o índice LOLP terá um valor muito baixo, pois o evento $\{S(\mathbf{X}) < L\}$ será extremamente raro. Em consequência, o estimador (3.25) requer um grande esforço computacional, i.e. N precisa ser muito grande para poder estimar o índice LOLP com um baixo nível de incerteza β . A fim de acelerar o processo de convergência, técnicas de IS devem ser utilizadas, e outro estimador não tendencioso para o índice LOLP é dado por,

$$\hat{\text{LOLP}} = \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) < L\}} \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \mathbf{v}), \quad (3.26)$$

onde $\mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \mathbf{v})$ é a denominada *razão de verossimilhança*, e representa uma correção que deve ser realizada no processo de amostragem devido à utilização da função massa de probabilidade (pmf) $f(\bullet; \mathbf{n}, \mathbf{v})$, que é diferente da pmf $f(\bullet; \mathbf{n}, \mathbf{u})$ original. No caso das variáveis aleatórias binomiais, o valor da razão de verossimilhança definido para um estado \mathbf{X}_k e dado por,

$$\mathbf{W}(\mathbf{X}_k; \mathbf{n}, \delta, \mathbf{v}) = \frac{f(\mathbf{X}_k; \mathbf{n}, \delta)}{f(\mathbf{X}_k; \mathbf{n}, \mathbf{v})} = \frac{\prod_{j=1}^J \frac{n_j!}{x_j!(n_j - x_j)!} p_j^{x_j} (1 - p_j)^{n_j - x_j}}{\prod_{j=1}^J \frac{n_j!}{x_j!(n_j - x_j)!} p'_j{}^{x_j} (1 - p'_j)^{n_j - x_j}}. \quad (3.27)$$

No caso específico do problema apresentado nesta seção, a expressão (3.27) finalmente resulta em,

$$\mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \mathbf{v}) = \frac{f(\mathbf{X}_i; \mathbf{n}, \mathbf{u})}{f(\mathbf{X}_i; \mathbf{n}, \mathbf{v})} = \frac{\prod_{j=1}^{N_c} (1 - u_j)^{x_j} (u_j)^{n_j - x_j}}{\prod_{j=1}^{N_c} (1 - v_j)^{x_j} (v_j)^{n_j - x_j}}. \quad (3.28)$$

Uma distorção das probabilidades do espaço de estados Ω é determinada pelo vetor de parâmetros $\mathbf{v} = [v_1, v_2, \dots, v_j, \dots, v_{N_c}]$. O problema agora consiste em escolher o melhor

conjunto de parâmetros de referência \mathbf{v} tal que o esforço computacional da simulação seja reduzido o máximo possível.

3.4.2. Algoritmo da Abordagem Baseada no MEC

Utilizando os conceitos do MEC apresentados na seção anterior, a idéia é estimar o vetor de parâmetros de referência ótimo \mathbf{v}_{opt} tal que a divergência de Kullback-Leibler seja mínima entre $f(\bullet; \mathbf{n}, \mathbf{v}_{opt})$ e $f(\bullet; \mathbf{n}, \mathbf{v})$. Isto é atingido pelo subsequente algoritmo de dez passos, onde o *Passo 1* é o processo de inicialização, *Passos 2-6*, representam a abordagem de otimização do MEC, e os *Passos 7-10* são a SMC não-cronológica otimizada com IS. Além disso, todo o desenvolvimento será baseado no índice LOLP e os outros índices GCR serão discutidos na subseção seguinte.

Passo 1. As características das estações geradoras, o nível de carga L e todos os parâmetros iniciais são especificados. Devem ser especificadas as capacidades das unidades (C_j), taxas de falha e reparo (λ_j e μ_j), número de unidades por estação geradora (n_j) e o nível de carga (L). As disponibilidades e indisponibilidades são calculadas para o regime permanente a partir das taxas de falha e reparo em (3.29a) e (3.29b), respectivamente.

$$1 - u_j = \frac{\mu_j}{\lambda_j + \mu_j} \quad (3.29a)$$

$$u_j = \frac{\lambda_j}{\lambda_j + \mu_j} \quad (3.29b)$$

O parâmetro de suavização α , em geral, é igual a um. Valores diferentes de um (tipicamente entre 0,80 e 0,99) são utilizados somente nos casos onde a distorção ótima do espaço de estados é crítica. Uma distorção exagerada pode levar à aparição de zeros ou uns no vetor de parâmetros \mathbf{v} , modificando desta maneira o problema original. O parâmetro de níveis múltiplos ρ (com valores típicos entre 0,01 e 0,1) e o número de amostras N (por exemplo, 20000 amostras) devem ser especificados para os *Passos 2-6* do algoritmo. O

coeficiente de variação β_{LOLP} (e.g. entre 1% e 5%) deve ser especificado para os *Passos* 7-10. Para estes mesmos passos é possível definir um número máximo de iterações N_{MAX} .

Passo 2. Definir $\hat{\mathbf{v}}_0 := \mathbf{u}$, ou seja, $\hat{\mathbf{v}}_0$ será igual ao vetor de indisponibilidades originais do sistema de geração; logo fazer $k := 1$ (contador de iterações do processo de otimização do MEC).

Passo 3. Gerar N amostras aleatórias $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ de estados de geração, de acordo com a função massa de probabilidade $f(\bullet; \mathbf{n}, \hat{\mathbf{v}}_{k-1})$. Avaliar a função desempenho $S(\mathbf{X}_i)$ para todo \mathbf{X}_i de acordo com (3.22) e ordenar os resultados em ordem decrescente tal que $\mathbf{S} = [S_{[1]}, S_{[2]}, \dots, S_{[N]}]$ e $S_{[1]} \geq S_{[2]} \geq \dots \geq S_{[N]}$.

Passo 4. O próximo passo é encontrar o valor de \hat{L}_k , ou seja, o nível de carga que será considerado na iteração k . O valor de \hat{L}_k será uma aproximação do (ρ) quantil amostral tal que $\hat{L}_k := S_{[(1-\rho)N]}$. Em outras palavras, calcular o valor de $(1-\rho)N$, e utilizar esse valor como um ponteiro r . \hat{L}_k será igual ao valor da r -ésima posição do vetor \mathbf{S} caso $S_{[(1-\rho)N]} > L$. Se $S_{[(1-\rho)N]} \leq L$, fazer $\hat{L}_k := L$. Avaliar a função teste $H(\mathbf{X}_i) = I_{\{S(\mathbf{X}_i) < \hat{L}_k\}}$ para todo \mathbf{X}_i . Avalie também a razão de verossimilhança $\mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_{k-1})$ para todo \mathbf{X}_i de acordo com (3.28).

Passo 5. O vetor de parâmetros de referência $\hat{\mathbf{v}}_k$ será agora encontrado resolvendo o problema de otimização estocástico (3.30),

$$\hat{\mathbf{v}}_k = \max_{\hat{\mathbf{v}}_k} \frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) < \hat{L}_k\}} \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_{k-1}) \ln f(\mathbf{X}_i; \mathbf{n}, \hat{\mathbf{v}}_k). \quad (3.30)$$

Considerando que a distribuição binomial pertence à família natural de exponenciais, a solução de (3.30) pode ser encontrada analiticamente resolvendo o sistema de equações (3.31),

$$\frac{1}{N} \sum_{i=1}^N I_{\{S(\mathbf{X}_i) < \hat{L}_k\}} \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_{k-1}) \nabla \ln f(\mathbf{X}_i; \mathbf{n}, \hat{\mathbf{v}}_k) = 0. \quad (3.31)$$

No caso da distribuição binomial, a solução de (3.31) finalmente resulta em,

$$\hat{\mathbf{v}}_{k,j} = \frac{1}{n_j} \left[\frac{\sum_{i=1}^N I_{\{S(\mathbf{X}_i) < \hat{L}_k\}} \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_{k-1}) X_{ij}}{\sum_{i=1}^N I_{\{S(\mathbf{X}_i) < \hat{L}_k\}} \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_{k-1})} \right], \quad (3.32)$$

com $j = 1, 2, \dots, N_C$. Caso $\alpha \neq 1$, corrigir $\hat{\mathbf{v}}_{k,j}$ de maneira semelhante a (3.20). Note que no processo de estimação do vetor $\hat{\mathbf{v}}_k$, definido em (3.32), somente *amostras de elite* são significativas, i.e. \mathbf{X}_i tal que $I_{\{S(\mathbf{X}_i) < \hat{L}_k\}} = 1$.

Passo 6. Caso $\hat{L}_k = L$, então é o fim do processo de otimização baseado no MEC (em $k = K$) e continuar no *Passo 7*; caso contrário incrementar o contador de iterações $k := k + 1$ e recomeçar o processo iterativo desde o *Passo 3*.

Passo 7. O vetor de referência ótimo \mathbf{v}_{opt} foi estimado utilizando o MEC nos *Passos 1-6*. Seja K a última iteração da primeira etapa do algoritmo. Considere que \mathbf{v}_{opt} seja aproximadamente igual a $\hat{\mathbf{v}}_K$. Agora, uma SMC não-cronológica e a VRT de IS serão utilizadas para estimar o índice de confiabilidade LOLP. Fazer o novo contador de iterações $N_1 := 0$.

Passo 8. Gerar uma amostra \mathbf{X}_{N_1} de acordo com a pmf $f(\bullet; \mathbf{n}, \hat{\mathbf{v}}_K)$, fazer $N_1 := N_1 + 1$.

Passo 9. Avaliar $H_{LOLP}(\mathbf{X}_{N_1}) = I_{\{S(\mathbf{X}_{N_1}) < L\}}$, $\mathbf{W}(\mathbf{X}_{N_1}; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K)$ e também o estimador não tendencioso para o índice LOLP na iteração i como segue:

$$\hat{LOLP} = \frac{1}{N_1} \sum_{i=1}^{N_1} I_{\{S(\mathbf{X}_i) < L\}} \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K). \quad (3.33)$$

Passo 10. Calcule o coeficiente de variação β_{N_1} . Se $\beta_{N_1} \leq \beta_{LOLP}$ ou $N_1 \geq N_{MAX}$, pare o algoritmo, caso contrário, volte ao *Passo 8*. A fim de tornar o processo da SMC mais eficiente, a convergência pode ser verificada em blocos de, por exemplo, 1000 amostras.

3.4.3. Distorção do Espaço de Estados e Índices de Confiabilidade

Até aqui, o método de avaliação da confiabilidade da capacidade de geração via o MEC tem mencionado somente o índice LOLP. Além disso, a carga foi fixada somente em um nível. Note que a distorção do espaço de estados é estimada para o pico de carga L . A distorção poderia ser baseada em um nível L' diferente de L , porém, no caso da carga ser constante no período de análise, o melhor desempenho em termos de tempo de simulação é sempre obtido ajustando o algoritmo para L .

Suponha que a distorção do espaço de estados é ajustada para um nível de carga L' . Basicamente aumentando L' , partindo do nível ótimo L , a distorção do espaço de estados será cada vez menor, pois $\hat{\mathbf{v}}_K$ vai se aproximando cada vez mais de \mathbf{u} . De fato, no limite onde L' é igual à capacidade instalada do sistema mais um acréscimo infinitesimal (neste caso todas as amostras serão *elite*), quando $N \rightarrow \infty$, temos $\hat{\mathbf{v}}_K \approx \mathbf{u}$ e $\mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K) \approx 1$ para todo \mathbf{X}_i , o que implica trabalhar com a simulação Monte Carlo não-cronológica convencional.

Além disso, diferentes estados de falha não contribuem na mesma proporção para os índices de confiabilidade. Diminuindo L' , partindo do nível ótimo L , fará com que estados \mathbf{X}_i extremamente raros sejam amostrados, porém, uma vez compensados com $\mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K)$ será verificado que estes estados em quase nada contribuem para os índices de confiabilidade, pois $\mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K)$ será próximo de zero. Em conclusão, em um sistema de geração com carga constante L , o melhor desempenho sempre é obtido fazendo $L'=L$, ou seja, ajustando a distorção do espaço de estados para o pico de carga.

A Fig. 3.2 apresenta os tempos de simulação necessários para avaliar o índice LOLP quando L' assume valores acima e abaixo da carga pico L . Estes tempos foram obtidos simulando o sistema de geração apresentado em [IEEE79], composto por 32 unidades

geradoras, com uma capacidade instalada de 3405MW e uma carga constante em 1995MW. Para as simulações, um $\beta_{LOLP}=1\%$ foi especificado. O valor da LOLP nestas condições é aproximadamente $8,86 \times 10^{-5}$. Note que o melhor desempenho é obtido quando $L'=L=1995MW$. O tempo necessário nestas condições foi aproximadamente 2,6 segundos, dos quais 0,7 segundos foram necessários para estimar o vetor de parâmetros de referência ótimo. A SMC convencional atingiu o mesmo critério de convergência em aproximadamente 2,1 horas.

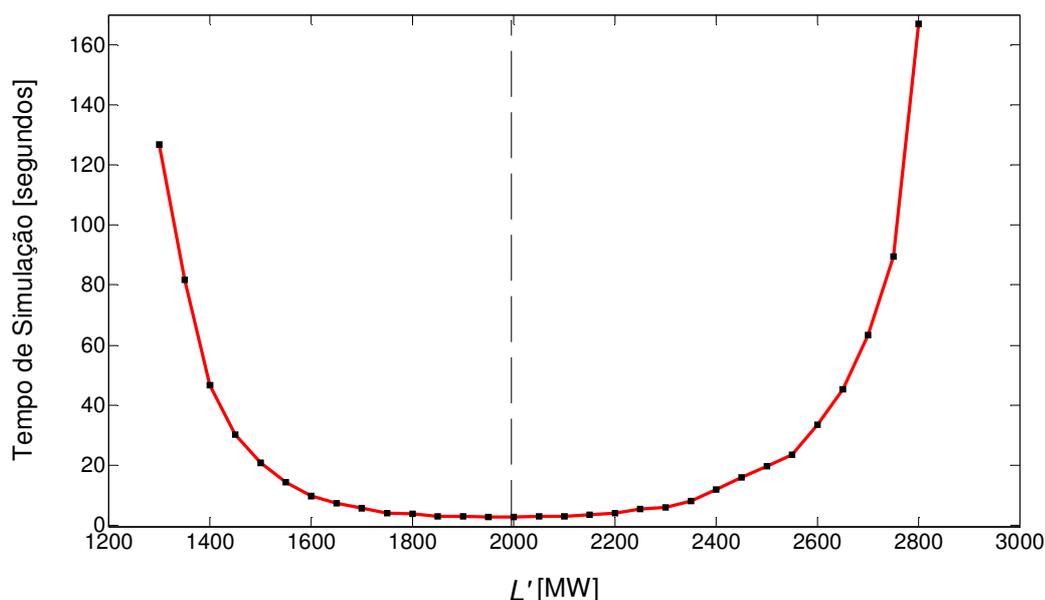


Fig. 3.2: Distorção Ótima do Espaço de Estados em $L'=L$

A Fig. 3.3a ilustra as probabilidades dos estados de geração sujeitas ao vetor \mathbf{u} . Note que a probabilidade de amostrar um estado \mathbf{X}_i tal que $S(\mathbf{X}_i) < L$ é baixa, o que explica a demorada convergência na SMC. A Fig. 3.3b apresenta a distorção do espaço de estados obtida quando a amostragem é baseada no vetor de parâmetros de referência ótimo $\hat{\mathbf{v}}_K$.

Pode-se concluir que o processo de amostragem baseado no vetor $\hat{\mathbf{v}}_K$ é muito mais eficiente e o número de amostras necessárias para atingir a convergência é drasticamente reduzido.

Os efeitos provocados por esta mudança de medida são facilmente compensados pela utilização da razão de verossimilhança, garantindo desta forma uma estimativa não tendenciosa para o índice LOLP.

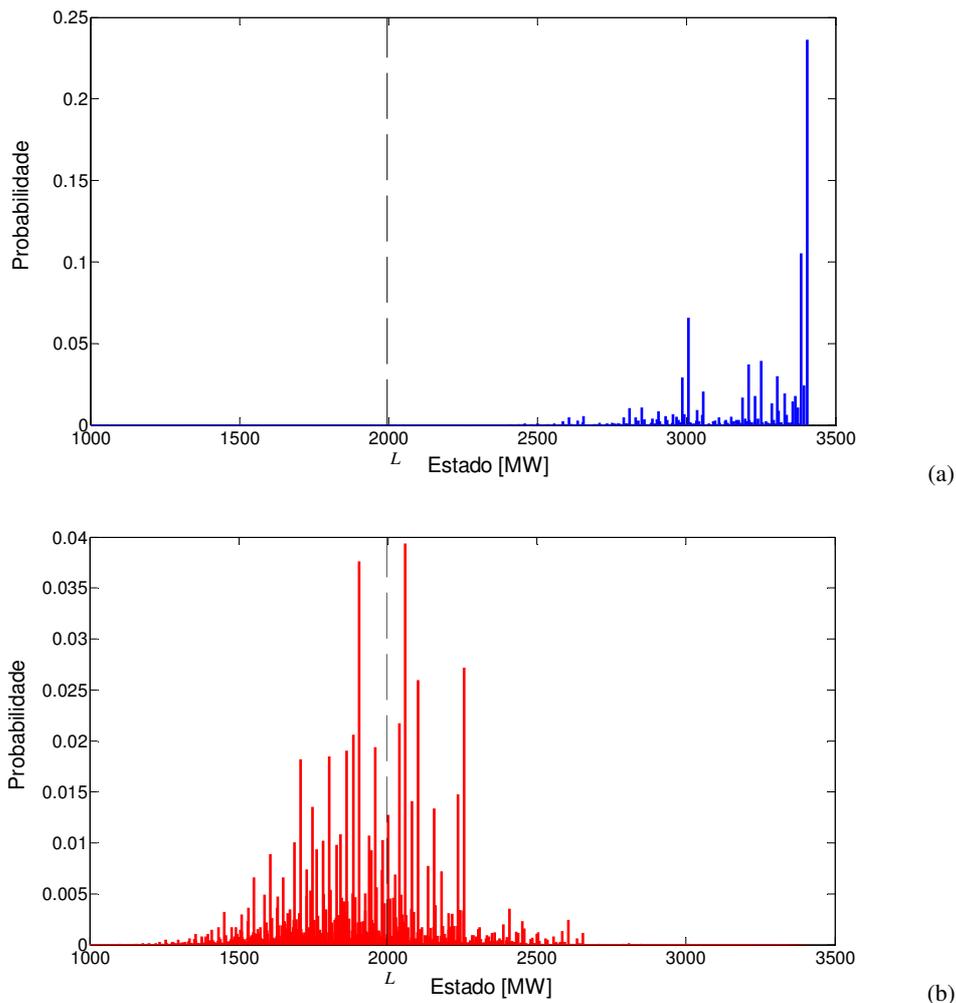


Fig. 3.3: Probabilidades dos Estados de Geração

Note também que a *distorção* do espaço de estados foi baseada no índice LOLP. Isto levanta uma dúvida sobre as distorções ótimas para os demais índices de confiabilidade, pois a distorção ótima para o índice LOLP, i.e \hat{v}_K^{LOLP} , não é necessariamente ótima para estimar os índices EPNS e LOLF.

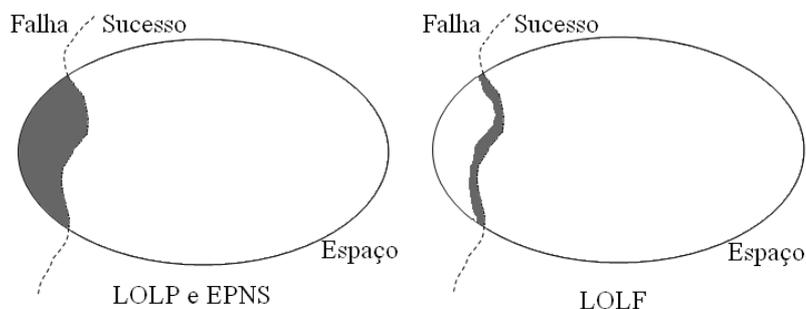


Fig. 3.4: Estados de Falha com Contribuições para os Índices GCR

A Fig. 3.4 ilustra os diferentes estados do sistema que têm contribuição nos índices de confiabilidade. No caso do índice LOLF, qualquer lado da fronteira poderia ter sido indicado. Os estados de falha que contribuem para o índice LOLP são os mesmos que contribuem para o índice EPNS, porém não na mesma proporção. Enquanto o índice LOLP é obtido pelo somatório das probabilidades de todos os estados de falha, o índice EPNS é o somatório dos produtos das probabilidades dos estados pelo tamanho do corte nesses estados.

No caso do índice LOLP, os estados de falha que mais contribuem são os que se encontram de certa forma próximos à fronteira, pois estes são os mais prováveis. Em geral, afastando-se da fronteira, os eventos tendem a ser cada vez mais raros e suas contribuições vão ficando desprezíveis. No caso do índice EPNS, existe um compromisso entre a raridade do evento e o tamanho do corte. Próximos à fronteira, os estados apresentam probabilidades mais significativas, porém, o tamanho do corte é baixo. Longe da fronteira, os estados apresentam cortes expressivos, porém, têm baixíssima probabilidade de ocorrência. Contudo, as probabilidades acabam influenciando mais que as magnitudes dos cortes e conclui-se então que estados dominantes para estimar o índice EPNS também estão localizados próximos à fronteira.

No caso do índice LOLF, somente os estados de falha que têm ligação direta com algum estado de sucesso apresentam contribuições, ou seja, necessariamente um estado que contribui para o índice LOLF deve estar localizado próximo da fronteira.

Da discussão acima se conclui que, em geral, as distorções ótimas para cada índice serão diferentes, porém, considerando que para os três índices os estados dominantes estão localizados próximos à fronteira, é de se esperar que $\hat{\mathbf{v}}_K^{\text{LOLP}} \cong \hat{\mathbf{v}}_K^{\text{EPNS}} \cong \hat{\mathbf{v}}_K^{\text{LOLF}}$, então a mesma distorção encontrada para o índice LOLP pode ser utilizada para estimar os índices EPNS e LOLF. Note que, se as distorções ótimas do espaço de estados fossem muito diferentes para cada índice, seria necessário executar um algoritmo específico para cada um deles, o que poderia tornar esta abordagem inviável. Considerando que $\hat{\mathbf{v}}_K^{\text{LOLP}} \cong \hat{\mathbf{v}}_K^{\text{EPNS}} \cong \hat{\mathbf{v}}_K^{\text{LOLF}}$, podemos modificar o algoritmo apresentado na subseção 3.4.2 para estimar todos os índices.

Passo 1. Repetir os *Passos 1-6* do algoritmo apresentado na subsecção 3.4.2 para estimar $\hat{\mathbf{v}}_K^{\text{LOLP}}$ e considerar que $\hat{\mathbf{v}}_K^{\text{LOLP}} \cong \hat{\mathbf{v}}_K^{\text{EPNS}} \cong \hat{\mathbf{v}}_K^{\text{LOLF}}$.

Passo 2. Gerar uma amostra \mathbf{X}_{N_1} de acordo com a função massa de probabilidade $f(\bullet; \mathbf{n}, \hat{\mathbf{v}}_K^{\text{LOLP}})$ e fazer $N_1 := N_1 + 1$.

Passo 3. Avaliar $H_{\text{LOLP}}(\mathbf{X}_{N_1})$, $H_{\text{EPNS}}(\mathbf{X}_{N_1})$, $H_{\text{LOLF}}(\mathbf{X}_{N_1})$ conforme (2.11), (2.12) e (2.13) respectivamente. Também avaliar $\mathbf{W}(\mathbf{X}_{N_1}; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K^{\text{LOLP}})$ e estimar os índices GCR como segue.

$$\hat{\text{LOLP}} = \frac{1}{N_1} \sum_{i=1}^{N_1} H_{\text{LOLP}}(\mathbf{X}_i) \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K^{\text{LOLP}}) \quad (3.34)$$

$$\hat{\text{EPNS}} = \frac{1}{N_1} \sum_{i=1}^{N_1} H_{\text{EPNS}}(\mathbf{X}_i) \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K^{\text{LOLP}}) \quad (3.35)$$

$$\hat{\text{LOLF}} = \frac{1}{N_1} \sum_{i=1}^{N_1} H_{\text{LOLF}}(\mathbf{X}_i) \mathbf{W}(\mathbf{X}_i; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K^{\text{LOLP}}) \quad (3.36)$$

As expressões $H_{\text{LOLP}}(\mathbf{X}_i)$, $H_{\text{EPNS}}(\mathbf{X}_i)$ e $H_{\text{LOLF}}(\mathbf{X}_i)$ representam as funções testes da LOLP, EPNS e LOLF respectivamente. Especificando um intervalo de tempo T , por exemplo, de 8736 horas, temos,

$$\hat{\text{LOLE}} = \hat{\text{LOLP}} \cdot T \quad (3.37)$$

$$\hat{\text{EENS}} = \hat{\text{EPNS}} \cdot T \quad (3.38)$$

$$\hat{\text{LOLD}} = \frac{\hat{\text{LOLP}}}{\hat{\text{LOLF}}} \quad (3.39)$$

Passo 4. Calcule os coeficientes de variação $\beta_{N_1}^{\text{LOLP}}$, $\beta_{N_1}^{\text{EPNS}}$ e $\beta_{N_1}^{\text{LOLF}}$. Caso os três coeficientes sejam menores ou iguais as respectivas tolerâncias especificadas, β_{LOLP} , β_{EPNS} e β_{LOLF} ou se $N_I \geq N_{\text{MAX}}$, pare o algoritmo, caso contrário, volte ao *Passo 2*. A fim de tornar o processo do CE-ISMC mais eficiente, a convergência é verificada em blocos de, por exemplo, 1000 amostras.

3.4.4. Modelos de Carga Variável

Qualquer modelo de carga é uma aproximação da carga real do sistema. Sua precisão depende da quantidade e qualidade dos dados disponíveis. Hipóteses markovianas sempre podem ser verificadas não somente para saídas forçadas de equipamentos, mas também para as cargas. Se elas são aceitáveis, as estimativas baseadas nestes modelos serão satisfatórias. O algoritmo de SMC baseado no MEC proposto na subseção anterior apresentou a avaliação dos índices GCR considerando somente um nível de carga L , i.e. a carga pico. O algoritmo proposto será estendido para tratar um modelo de carga de múltiplos estados.

Nesta subseção, L será considerada como uma sequência cronológica de T níveis discretos de carga, cada um com duração ΔT , por exemplo, de uma hora, semelhante ao ilustrado na Fig. 3.5.

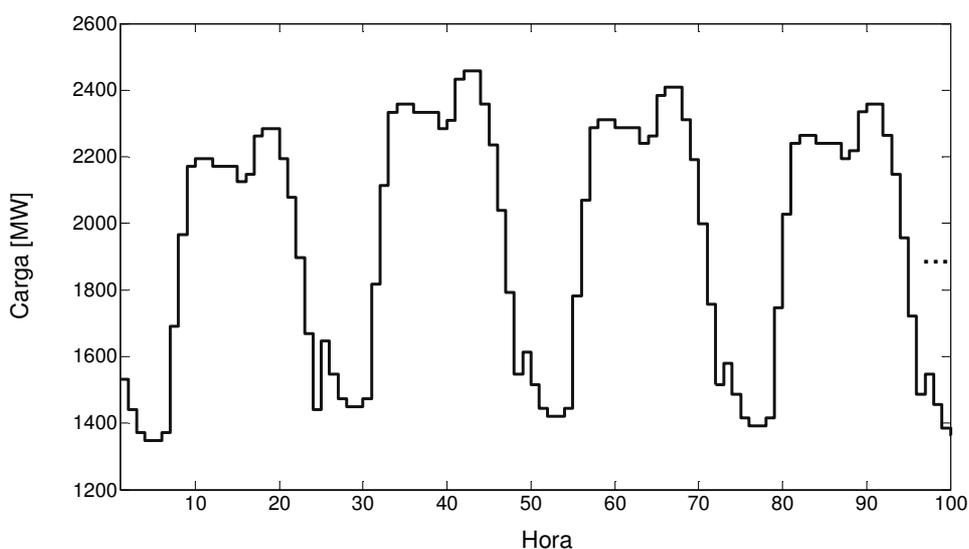


Fig. 3.5: Modelo de Carga de Níveis Discretos

A curva de carga L contém, por exemplo, 8736 pontos que representam os níveis de carga para 8736 horas. Diferentemente do modelo de carga onde L era representado por um único estado (constante no período de análise), este modelo de L será representado por diferentes estados possíveis. Isto quer dizer que agora os estados de carga precisam ser amostrados também. Uma maneira de amostrar os estados de carga é o *mapeamento* de L por uma distribuição uniforme discreta. Primeiramente, um número inteiro a uniformemente distribuído no intervalo $[1,8736]$ é sorteado, a seguir, verifica-se o nível de carga associado. Este procedimento é ilustrado na Fig. 3.6.

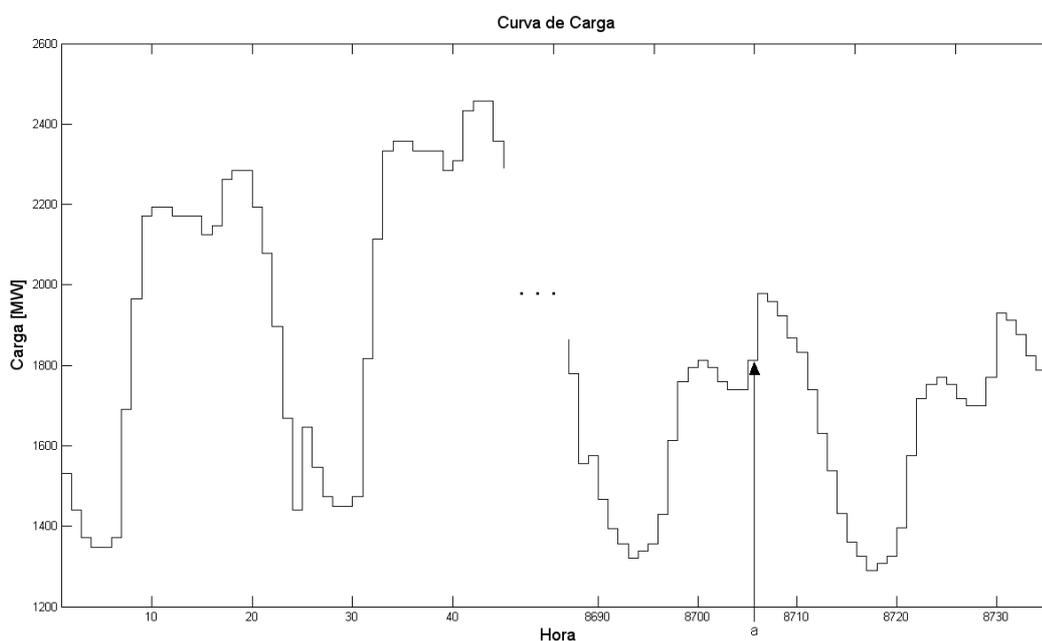


Fig. 3.6: Mapeamento da Curva de Carga por uma Distribuição Uniforme Discreta

Outra possibilidade é utilizar o modelo markoviano sequencial ilustrado na Fig. 3.7. Ele descreve um conjunto de T níveis de carga, representando, por exemplo, 8736 horas, sequencialmente conectadas na mesma ordem cronológica com que aparecem na sequência histórica.

O modelo utiliza uma taxa constante de transição de carga $\lambda_L = 1/\Delta T$, onde ΔT representa a unidade de tempo utilizada para discretizar o período T , e.g. $\lambda_L = 1$ transição por hora. Como todas as taxas de transição são as mesmas, a carga irá permanecer, em média, ΔT horas (e.g. 1 hora) em cada nível, e T horas (e.g. 8736 horas), em média para o período total de análise. Este conceito é um caso particular do modelo de Markov para níveis múltiplos de carga não agregado proposto em [LMMB00].

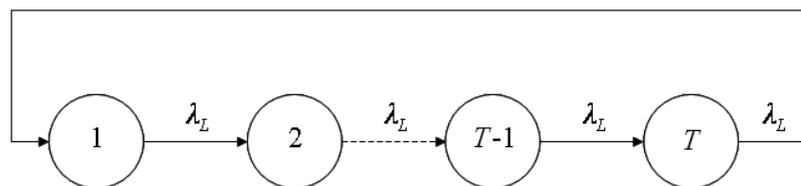


Fig. 3.7: Modelo Markoviano Sequencial de Carga

Na implementação computacional da metodologia proposta, como o número de simulações é elevado, será considerado que, em média, cada estado de carga tem duração de ΔT , por exemplo, de uma hora, não havendo necessidade de sortear os tempos individuais de duração de cada estado.

O principal problema em considerar L com múltiplos estados é que, em geral, L não obedece a nenhuma distribuição de probabilidade padrão conhecida. Caso L seguisse uma distribuição binomial, os parâmetros de L poderiam ser distorcidos de maneira semelhante aos de G . Neste cenário hipotético, o modelo da geração seria combinado com o modelo da carga, e o modelo resultante, o da reserva R , teria seus parâmetros distorcidos para encontrar $P_u\{S'(\mathbf{X}) < 0\}$, onde S' é uma função real que retorna o nível de reserva para os estados \mathbf{X} (agora estados de geração e de carga). Em conclusão, considerando que L , em geral, não segue uma distribuição binomial, as probabilidades do modelo da carga L não serão distorcidas.

Note também que o algoritmo apresentado na subseção anterior determina a distorção do espaço de estados de geração baseado em um único nível de carga L' . O problema agora é determinar qual é o nível de carga L' que terá o melhor desempenho considerando que agora o modelo L é representado por múltiplos estados.

É de conhecimento geral que os estados de carga onde há maior contribuição para os índices de confiabilidade são aqueles onde os níveis de carga são elevados. Dificilmente um estado de carga com nível baixo terá uma contribuição significativa. Considere que a carga L tenha a distribuição de probabilidades apresentada na Fig. 3.8. Observe que a maior parcela da massa de probabilidade está localizada em estados com níveis de carga menores a 2500MW (aproximadamente 97,4%), ou seja, L passa a maior parte do tempo

em estados com cargas de nível médio ou baixo que não contribuem significativamente para os índices.

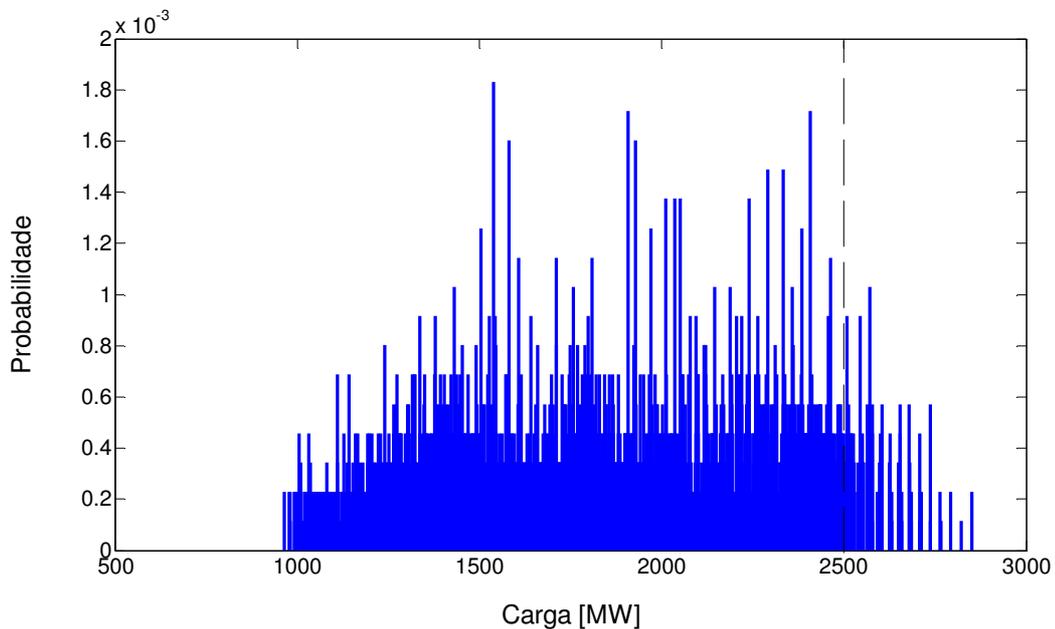


Fig. 3.8: Distribuição de Probabilidade de L

Quando uma distorção de G é obtida para L' os estados de geração amostrados serão próximos a L' , conforme foi ilustrado na Fig. 3.3b. Portanto, uma escolha de L' na faixa de cargas elevadas deve fornecer um bom desempenho. De fato, foi verificado por diversas simulações que um L' igual ou muito próximo ao pico da carga apresenta o melhor desempenho. O nível L' ótimo é desconhecido inicialmente, mas em nenhuma das simulações foi menor a $0,96L_{MAX}$, onde L_{MAX} é a carga pico do sistema. De um modo geral, fazendo $L'=L_{MAX}$ sempre se obtém um excelente desempenho e isto será adotado para todas as simulações futuras. Note que qualquer L' tal que $L'>L_{MAX}$ sempre apresentará um pior desempenho, pois na medida em que L' se aproxima da capacidade instalada do sistema $\hat{\mathbf{v}}_K^{LOLP}$ se aproxima de \mathbf{u} .

3.4.5. O Parâmetro Φ

Como as probabilidades dos estados de carga não são distorcidas, a amostragem de L pode ser ineficiente dependendo do fator de carga do sistema, pois nosso interesse limita-se aos estados com cargas elevadas. No caso apresentado na Fig. 3.8, aproximadamente 97,4% das amostras de L não serão relevantes. Em outras palavras, os estados de falha continuam

sendo raros, mas agora por causa da distribuição de probabilidades de L e não mais de G , pois seus parâmetros foram distorcidos.

Uma maneira de contornar este problema é utilizando um parâmetro adicional denominado Φ definido como,

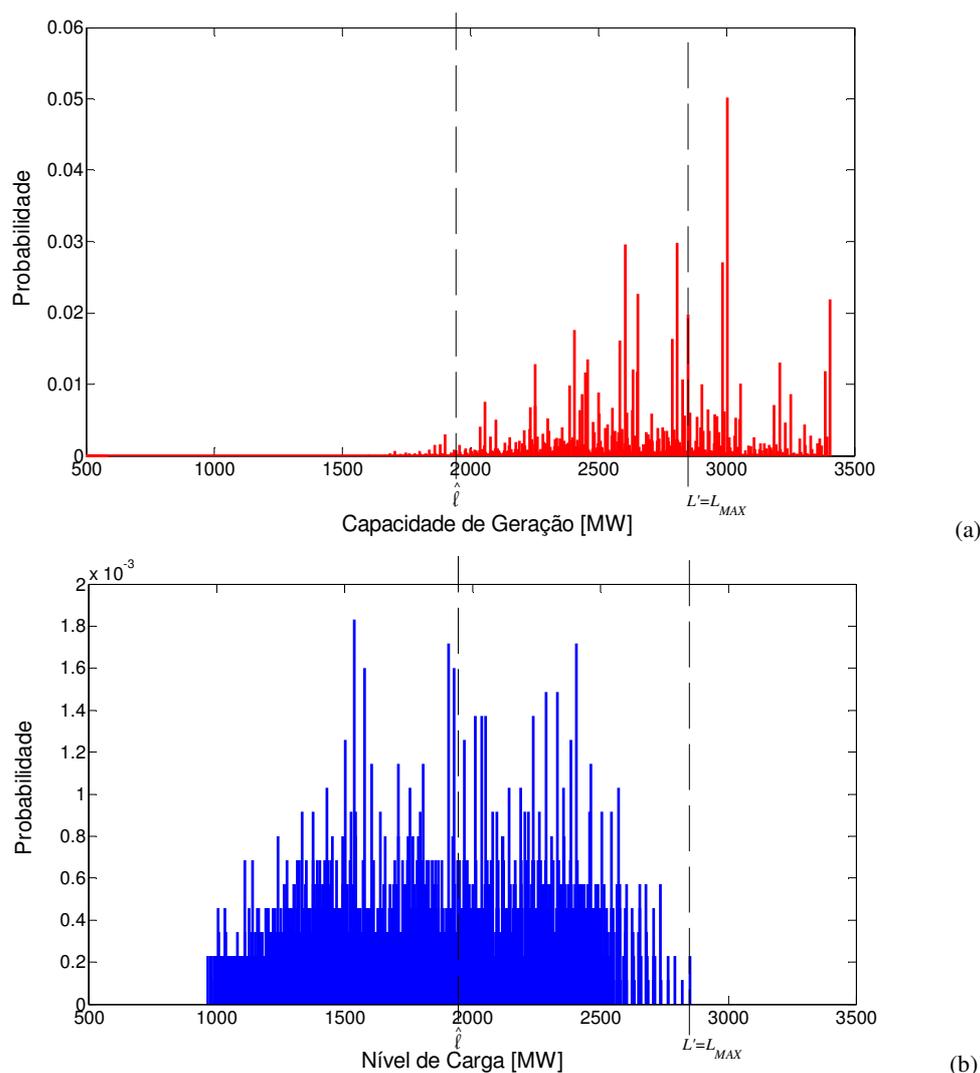
$$\Phi = P_{\hat{\mathbf{v}}_K} \{S(\mathbf{X}_i) < \ell\}. \quad (3.40)$$

Note que, na definição deste parâmetro, as probabilidades são em relação ao vetor de parâmetros de referência ótimo $\hat{\mathbf{v}}_K$, ou seja, o parâmetro Φ é definido no *mundo distorcido* de probabilidades dos estados de geração.

Se a probabilidade Φ for especificada, uma estimativa $\hat{\ell}$ de ℓ pode ser encontrada por N amostras (e.g. 10000) dos estados de geração, i.e. $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ de acordo com $f(\bullet; \mathbf{n}, \hat{\mathbf{v}}_K^{\text{LOLP}})$, que é a distribuição obtida pela abordagem proposta (*Passos 2-6*). A função de desempenho $S(\mathbf{X}_i)$ é avaliada para cada \mathbf{X}_i , e os resultados são ordenados do menor ao maior valor: $S_{[1]} \leq S_{[2]} \leq \dots \leq S_{[N]}$. O valor de $\hat{\ell}$ será tal que $\hat{\ell} = S_{[\Phi N]}$.

Todos aqueles estados de carga amostrados com níveis de carga menores que $\hat{\ell}$ serão automaticamente considerados como estados de sucesso e não serão verificados durante o processo da SMC baseado em IS (i.e. *Passos 7-10* do algoritmo apresentado na subseção 3.4.2). Este é um procedimento simples, mas muito eficiente do ponto de vista computacional.

A Fig. 3.9a apresenta a distorção obtida para o mesmo sistema original apresentado na Fig. 3.3a, porém, para um $L' = L_{MAX} = 2850\text{MW}$. Considerando um $\Phi = 0,05$, temos um $\hat{\ell} = 1945\text{MW}$ e todos os estados do sistema com níveis de carga menores que $\hat{\ell}$ serão diretamente considerados como estados de sucesso. Na Fig. 3.9b pode-se observar os estados de carga considerados como sucesso para o mesmo modelo de L apresentado na Fig. 3.8.


Fig. 3.9: Aplicação do Parâmetro Φ

O valor da probabilidade Φ depende da raridade dos eventos envolvidos. Valores típicos são de 0,01 a 0,05 para uma LOLP entre 10^{-3} a 10^{-5} ; de 0,05 a 0,15 para uma LOLP entre 10^{-5} a 10^{-8} ; e de 0,15 a 0,2 para $LOLP < 10^{-8}$. O excelente desempenho computacional do parâmetro Φ é devido ao fato de um estado de geração ser sorteado somente se um estado de carga maior ou igual a $\hat{\ell}$ for encontrado primeiro, pois a probabilidade de amostrar um estado \mathbf{X}_i , tal que $S(\mathbf{X}_i) < \hat{\ell}$, considerando a pmf $f(\cdot; \mathbf{n}, \hat{\mathbf{v}}_K^{LOLP})$, é aproximadamente Φ .

É sempre importante ter em mente que utilizar o parâmetro Φ introduz uma pequena imprecisão no processo de estimação, pois alguns estados amostrados são diretamente considerados como sucesso. Esta imprecisão, no entanto, pode ser insignificante se os valores de Φ são pequenos. Além disso, as contribuições já desprezadas no mundo

distorcido serão ainda mais insignificantes uma vez compensadas com $\mathbf{W}(\mathbf{X}_{N_t}; \mathbf{n}, \mathbf{u}, \hat{\mathbf{v}}_K^{\text{LOLP}})$. Quando não existe conhecimento da ordem de grandeza do índice LOLP, um $\Phi = 0,01$ pode ser utilizado. Quando a carga for constante ao longo do período de análise este parâmetro deve obviamente ser ignorado.

3.5. COMENTÁRIOS FINAIS

Neste capítulo foi apresentada a estrutura básica do Método da Entropia Cruzada aplicado na simulação de eventos raros. O principal objetivo foi demonstrar como este poderoso método pode ser aplicado ao problema específico da avaliação de índices de confiabilidade em sistemas de geração.

Foi apresentado o algoritmo da metodologia proposta baseada no MEC para a avaliação do índice LOLP. Posteriormente este algoritmo foi estendido para estimar também os demais índices GCR, i.e. LOLE, EPNS, EENS, LOLF e LOLD, considerando que o vetor ótimo de indisponibilidades encontrado para o índice LOLP, $\hat{\mathbf{v}}_K^{\text{LOLP}}$, mesmo não sendo ótimo para os demais índices, pode ser aplicado satisfatoriamente para estimar também os índices EPNS e LOLF. Foi demonstrado também que a carga pico do sistema, L_{MAX} , sempre pode ser utilizada como referência para estimar o vetor de parâmetros ótimo, independentemente se o modelo L for representado por um único estado ou por múltiplos estados.

Com o objetivo de melhorar a eficiência do processo de amostragem dos estados de L , o parâmetro Φ foi introduzido, minimizando problemas encontrados em sistemas muito confiáveis e/ou com fatores de carga baixos. Este procedimento resultou simples, prático e extremamente eficiente do ponto de vista computacional. Erros introduzidos no processo de estimação serão desprezíveis para um Φ moderado, conforme será demonstrado pelas diversas simulações apresentadas no capítulo seguinte.

No Capítulo 4 serão apresentados os principais resultados obtidos, onde o desempenho computacional da metodologia proposta é comparado com um método analítico extremamente eficiente e uma SMC não-cronológica convencional.

CAPÍTULO 4

RESULTADOS

4.1. INTRODUÇÃO

N O presente capítulo a metodologia proposta baseada no MEC será implementada e seu desempenho computacional será comparado com outras metodologias tradicionalmente utilizadas na avaliação de índices GCR. Um método analítico extremamente eficiente baseado em convolução discreta e uma simulação Monte Carlo não-cronológica serão utilizados para esta finalidade. Ambos os métodos foram brevemente descritos no Capítulo 2 desta dissertação.

A fim de avaliar a precisão e a eficiência do método proposto, diferentes sistemas de geração foram testados. Primeiramente, o sistema *IEEE Reliability Test System* (IEEE-RTS) [IEEE79] será utilizado com o principal objetivo de se obter valores padrões para os índices de confiabilidade. Este sistema, porém, não é o mais indicado para verificar completamente a capacidade da metodologia proposta, pois a raridade dos eventos de falha envolvidos não é significativa e o número de estados possíveis não é elevado.

Em seguida, algumas modificações do IEEE-RTS são utilizadas para realizar diversas comparações e testes de sensibilidade. Estas modificações são determinadas pelo denominado *fator de escala*, que será apresentado mais detalhadamente nas seções seguintes. O principal objetivo é verificar o desempenho da metodologia proposta quando aplicada a diversas condições operativas do sistema. A eficiência computacional do parâmetro Φ também será discutida. Por fim, uma configuração do sistema Sul-Sudeste Brasileiro de geração será utilizada para verificar a aplicação da metodologia proposta em aplicações práticas.

Os principais resultados das simulações e dos testes de sensibilidade serão apresentados na forma de gráficos e tabelas. Em todos os casos, a implementação computacional foi desenvolvida em uma plataforma MatLab utilizando um processador Intel[®] Core 2 Duo 2,66 GHz.

4.2. IEEE RELIABILITY TEST SYSTEM

Este sistema de geração consiste em 32 unidades geradoras com capacidade total instalada de 3405MW. Os principais parâmetros que descrevem o sistema, i.e. indisponibilidades e capacidades das unidades, podem ser encontrados em [IEEE79]. A carga é representada como uma curva horária com 8736 níveis discretos e cronologicamente conectados, cujo valor pico é de 2850MW. Este sistema é primeiramente utilizado com o único objetivo de se ter valores padrões de referência para os índices de confiabilidade e tempos de CPU. No entanto, o IEEE-RTS não é o melhor sistema para mostrar a capacidade total do método proposto assim como será discutido mais adiante.

4.2.1. IEEE-RTS – Modelo de Carga Constante

Primeiramente, será considerado o sistema IEEE-RTS com um modelo de carga L de um único nível no período de análise ($T=8736$ horas). A Tabela 4.1 apresenta os resultados dos índices GCR considerando um nível constante de carga, i.e. 2850MW, para os métodos: analítico, SMC convencional e CE-ISMC.

TABELA 4.1 – IEEE-RTS: ÍNDICES DE CONFIABILIDADE – MODELO DE CARGA CONSTANTE

IEEE-RTS	LOLP	LOLE [horas/ano]	EPNS [MW]	EENS [MWh/ano]	LOLF [oc./ano]	LOLD [horas/oc.]	Tempo de CPU [s]
Método Analítico	$8,45778 \times 10^{-2}$	$7,38872 \times 10^2$	$1,46936 \times 10^1$	$1,28364 \times 10^5$	$1,95123 \times 10^1$	$3,78669 \times 10^1$	0,03
SMC Conv.	$8,42043 \times 10^{-2}$ (0,73572%)	$7,35609 \times 10^2$	$1,47239 \times 10^1$ (0,98733%)	$1,28628 \times 10^5$	$1,92890 \times 10^1$ (0,92441%)	$3,81362 \times 10^1$	6,90
CE-ISMC	$8,45349 \times 10^{-2}$ (0,58412%)	$7,38497 \times 10^2$	$1,46499 \times 10^1$ (0,41708%)	$1,27981 \times 10^5$	$1,94612 \times 10^1$ (0,99820%)	$3,79472 \times 10^1$	2,90

Para o método analítico, um incremento $\Delta=1$ MW é utilizado de maneira semelhante ao descrito em [LMC91]. O parâmetro β é especificado como menor ou igual a 1% para todos os índices GCR, e seus valores correspondentes são apresentados entre parêntesis. Em geral, os índices EPNS e LOLF requerem maior número de amostras para atingir o critério de convergência. Para o CE-ISMC, os parâmetros $\alpha = 1$, $\rho = 0,1$, e $N = 2 \times 10^4$ foram utilizados.

Todos os métodos atingem um valor próximo de $8,4 \times 10^{-2}$ para o índice LOLP. Os métodos analítico, SMC convencional e CE-ISMC gastam 0,03, 6,9 e 2,9 segundos, respectivamente. Como os eventos envolvidos não são raros (i.e. 10^{-2}), não há ganho computacional significativo entre a SMC convencional e o CE-ISMC: um ganho aproximado de 2,4. Verifica-se que em sistemas pequenos o método analítico apresentado em [LMC91] resulta praticamente imbatível.

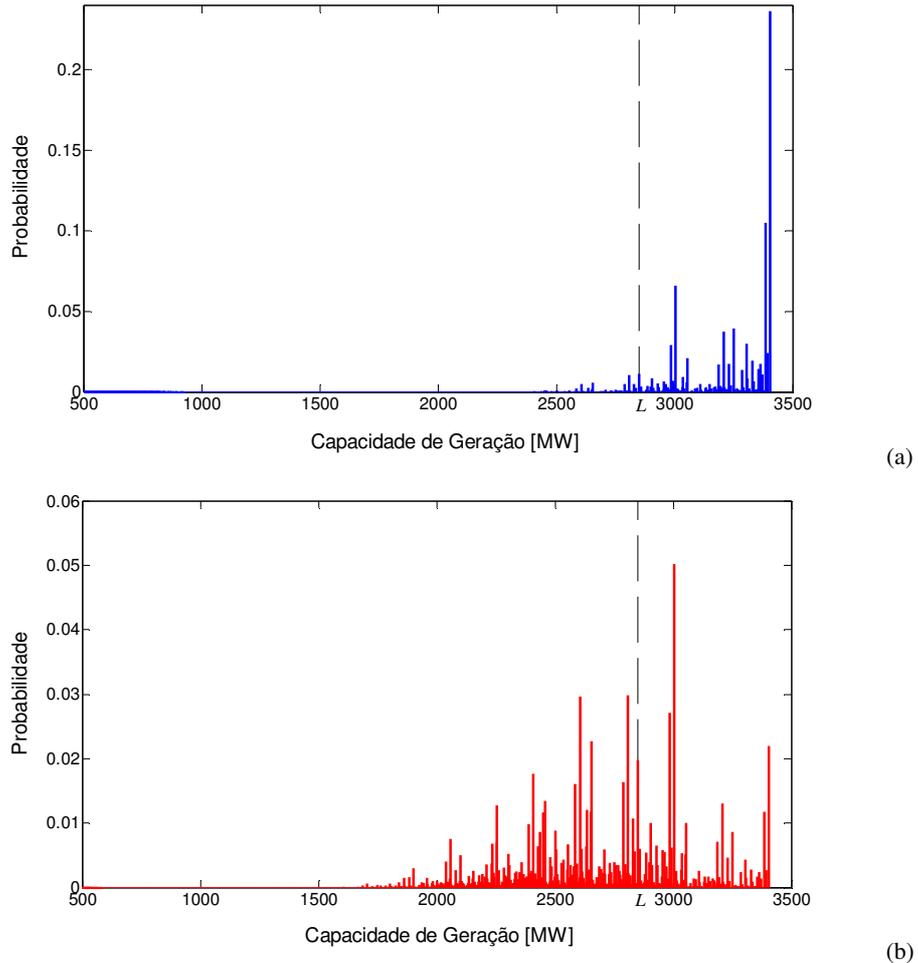


Fig. 4.1: Probabilidades dos Estados de Geração

A Fig. 4.1a ilustra as probabilidades dos estados de geração sob a função massa de probabilidade original $f(\bullet; \mathbf{n}, \mathbf{u})$. A Fig. 4.1b apresenta o mesmo espaço de estados de geração, mas sob a função massa de probabilidade ótima $f(\bullet; \mathbf{n}, \hat{\mathbf{v}}_K^{\text{LOLP}})$ encontrada pelo processo de otimização (*Passos 1-6* do algoritmo apresentado na subsecção 3.4.2).

Conclui-se a partir das Fig. 4.1a e Fig. 4.1b, que utilizando $f(\bullet; \mathbf{n}, \hat{\mathbf{v}}_K^{\text{LOLP}})$ como a função

massa de probabilidade amostral, o desempenho computacional deve ser sempre melhor, pois os estados dominantes de falha estão sendo amostrados com maior frequência. Uma quantidade de esforço computacional até 10 segundos é um tempo de espera aceitável para qualquer usuário. Contrariamente, tempos de espera de horas podem realmente causar alguma impaciência aos usuários e isto será ilustrado a seguir.

4.2.2. IEEE-RTS – Modelo de Carga de Níveis Múltiplos

A Tabela 4.2 apresenta os resultados para os índices GCR considerando o modelo markoviano de níveis múltiplos para L apresentado na Fig. 3.7, aplicado aos dados do IEEE-RTS [IEEE79]. O fator de carga é aproximadamente 0,6144 com uma carga pico de 2850MW e um período de 8736 horas. O parâmetro β é também especificado em 1% para todos os índices GCR, e seus valores correspondentes são apresentados entre parêntesis.

Na abordagem analítica, $\Delta=1\text{MW}$ é utilizado e não é especificada nenhuma probabilidade de truncamento. Para o CE-ISMC, os parâmetros $\alpha = 1$, $\rho = 0,1$, $N = 2,5 \times 10^4$ e $\Phi = 0,04$ foram utilizados. Note que neste caso, a abordagem proposta baseada no MEC tem um desempenho muito melhor do que a SMC convencional, pois a convergência foi atingida em um tempo de simulação 40,34 vezes menor (ganho).

TABELA 4.2 – IEEE-RTS: ÍNDICES DE CONFIABILIDADE – MODELO DE CARGA VARIÁVEL

IEEE-RTS	LOLP	LOLE [horas/ano]	EPNS [MW]	EENS [MWh/ano]	LOLF [oc./ano]	LOLD [horas/oc.]	Tempo de CPU [s]
Método Analítico	$1,07258 \times 10^{-3}$	$9,37008 \times 10^0$	$1,34649 \times 10^{-1}$	$1,17629 \times 10^3$	$2,01619 \times 10^0$	$4,64743 \times 10^0$	0,89
SMC Conv.	$1,06849 \times 10^{-3}$ (0,48375%)	$9,33433 \times 10^0$	$1,32769 \times 10^{-1}$ (0,65554%)	$1,15987 \times 10^3$	$1,99962 \times 10^0$ (0,99983%)	$4,66805 \times 10^0$	1440
CE-ISMC ($\Phi=0,04$)	$1,07047 \times 10^{-3}$ (0,41364%)	$9,35163 \times 10^0$	$1,34107 \times 10^{-1}$ (0,38734%)	$1,17156 \times 10^3$	$2,00882 \times 10^0$ (0,99968%)	$4,65528 \times 10^0$	35,7

As Fig. 4.2, Fig. 4.3 e Fig. 4.4 ilustram o processo de convergência de ambos os métodos, SMC convencional e CE-ISMC, para os índices LOLP, EPNS e LOLF, respectivamente. Neste caso, o CE-ISMC utiliza um conjunto de parâmetros definidos como segue: $\alpha = 1$, $\rho = 0,1$, $N = 2,5 \times 10^4$, $\Phi = 0,02$. Em cada gráfico, a ilustração superior representa os valores para o índice e a inferior os respectivos valores de β .

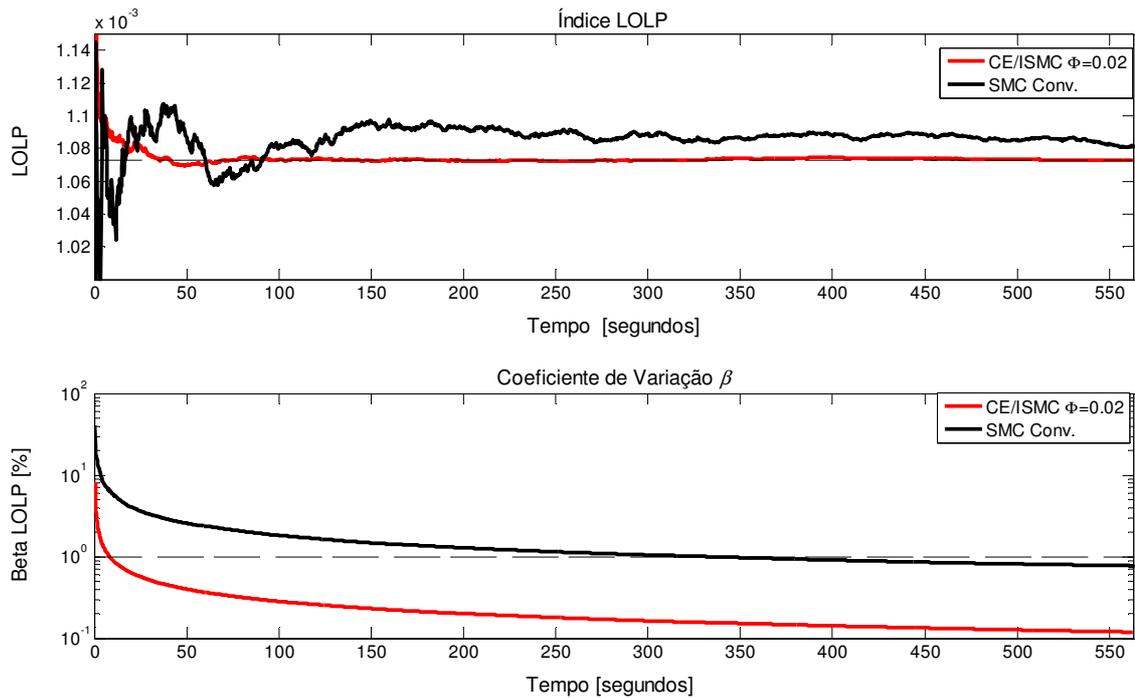


Fig. 4.2: Processo de Convergência do Índice LOLP

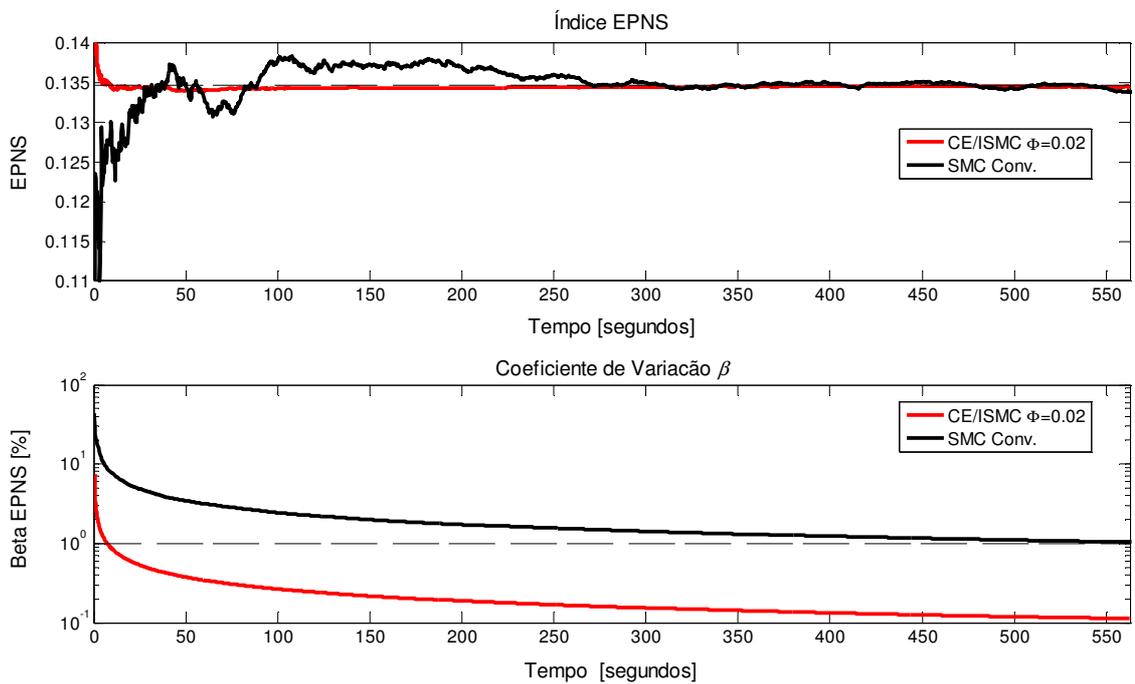


Fig. 4.3: Processo de Convergência do Índice EPNS

A abordagem de SMC baseada no MEC atingiu um $\beta_{LOLP} = 1\%$ em aproximadamente 8,5 segundos para o índice LOLP, que representa aproximadamente $6,22 \times 10^5$ amostras,

enquanto a SMC convencional atingiu o mesmo β em 332 segundos, ou aproximadamente $9,21 \times 10^6$ amostras.

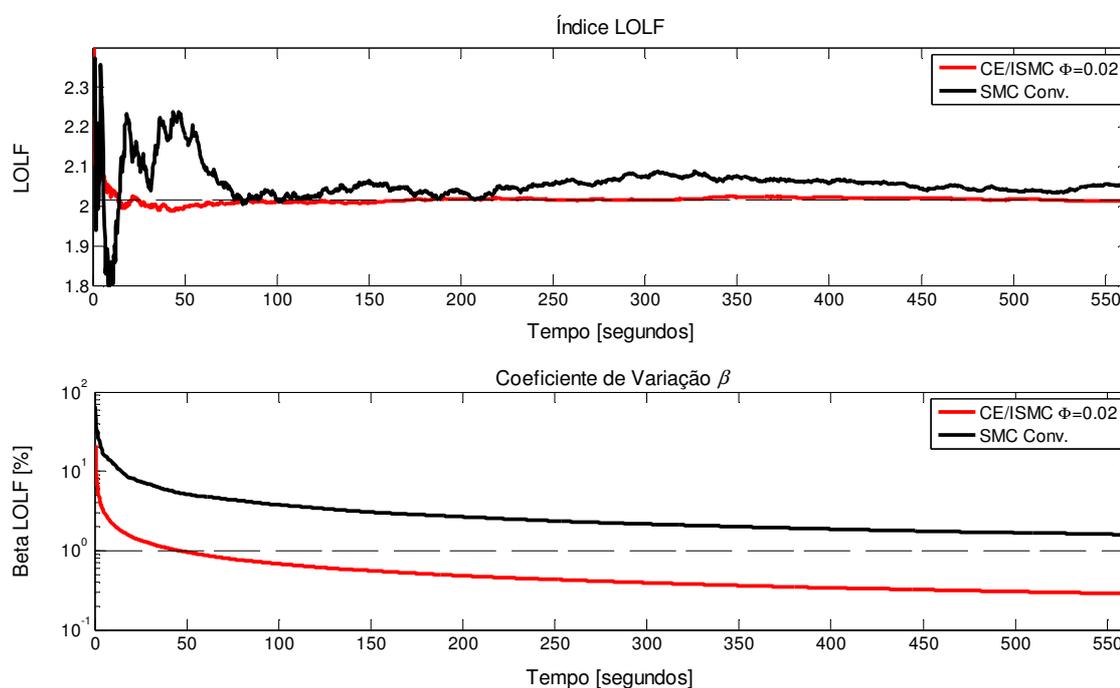


Fig. 4.4: Processo de Convergência do Índice LOLF

Em ambos os métodos o índice LOLF requer o maior número de amostras para atingir a convergência. O CE-ISMC atingiu o critério $\beta_{LOLF}=1\%$ em aproximadamente 46 segundos, que é um tempo um pouco maior ao apresentado na Tabela 4.2, devido à utilização de um Φ menor. A SMC convencional atinge o mesmo critério em 1440 segundos conforme a Tabela 4.2, porém este valor não aparece na Fig. 4.4 devido à necessidade de manter uma escala conveniente.

Note que para todos os índices GCR a variância do estimador obtida com o CE-ISMC é muito menor, o que leva a uma rápida convergência dos índices. De fato, observe que quando os coeficientes de variação atingem um nível de 1% na SMC convencional, estes mesmos coeficientes são próximos a 0,1% no CE-ISMC para o mesmo tempo de simulação. Se um critério de convergência de $\beta_{LOLF}=5\%$ fosse definido, a convergência seria atingida em aproximadamente 2,3 segundos considerando o CE-ISMC.

É importante destacar que as taxas de amostragem, i.e. amostras/segundo, são diferentes para ambos os métodos principalmente porque $\Phi \neq 0$ e também pela implementação

computacional específica de cada um deles. Por isto, a melhor maneira de comparar o processo de convergência para cada método foi baseada em uma escala de tempo, e não em uma escala definida pelo número de amostras como é usualmente encontrado na literatura. No entanto, o número de amostras necessárias para atingir a convergência é evidentemente menor no CE-ISMC.

4.3. MODIFICAÇÕES NO IEEE-RTS ORIGINAL

A fim de comparar completamente o desempenho do CE-ISMC com ambos os métodos analítico e SMC convencional, diversas modificações do IEEE-RTS são criadas. Estas modificações estão representadas por diferentes *fatores de escala*. O fator de escala é um número inteiro que indica a quantidade de sistemas idênticos IEEE-RTS que serão adicionados nos modelos de L e G .

Por exemplo, se um fator de escala de cinco está sendo utilizado (denominado como IEEE-RTS×5), isto quer dizer que o número de unidades e níveis de carga são multiplicados por este fator. Isto tem um impacto direto no grau de confiabilidade e no número de estados possíveis do sistema.

A Tabela 4.3* apresenta os resultados para o IEEE-RTS modificado por 1 (i.e. o sistema original), 5, 10, 15, e 20 fatores de escala. Aumentar este fator implica no aumento do número de estados do sistema, enquanto os valores do índice LOLP decrescem. Isto ocorre porque a capacidade total do sistema se torna imensa em comparação a uma única unidade geradora, o que significa que mesmo as saídas forçadas das máquinas de maior capacidade não têm um impacto significativo no desempenho operativo normal do sistema.

A Tabela 4.3 apresenta os valores para o índice LOLP calculados com o método analítico e com o CE-ISMC (com $\beta=1\%$ para os três índices). Também apresenta os tempos de processamento obtidos com: o método analítico ($\Delta=1\text{MW}$); SMC convencional com $\beta=5\%$; CE-ISMC com $\beta=5\%$; e CE-ISMC com $\beta=1\%$. Para os métodos de CE-ISMC, diferentes valores do parâmetro Φ são apresentados entre parêntesis. Os parâmetros $\alpha = 1$,

* Estes valores, expressos em meses (me.) e anos (a), foram obtidos por extrapolação das tendências em gráficos semi-log correspondentes.

$\rho = 0,1$ e $N = 2,5 \times 10^4$ foram utilizados em todos os casos.

TABELA 4.3 – IEEE-RTS \times FATOR DE ESCALA – MODELO DE CARGA VARIÁVEL

Fator de Escala	LOLP Método Analítico	LOLP CE-ISM $C_{\beta=1}$	Tempo de CPU			
			Método Analítico	SMC Conv. $\beta=5$	CE-ISM $C_{\beta=5}$ (Φ)	CE-ISM $C_{\beta=1}$ (Φ)
1	$1,07258 \times 10^{-3}$	$1,06406 \times 10^{-3}$	0,89 s	61,0 s	2,39 s (0,01)	34,6 s (0,05)
5	$1,24395 \times 10^{-6}$	$1,21271 \times 10^{-6}$	4,70 s	6,75 h	4,21 s (0,05)	36,6 s (0,13)
10	$1,06714 \times 10^{-8}$	$1,06797 \times 10^{-8}$	19,4 s	3 me.*	5,77 s (0,05)	34,7 s (0,15)
15	$1,58195 \times 10^{-10}$	$1,57368 \times 10^{-10}$	43,8 s	26 a*	5,38 s (0,07)	30,6 s (0,15)
20	$2,69301 \times 10^{-12}$	$2,69051 \times 10^{-12}$	85,0 s	707 a*	4,96 s (0,1)	28,2 s (0,2)

Pode ser observado na Tabela 4.3 que os resultados obtidos com o método analítico e com o CE-ISM C (com $\beta=1\%$) são basicamente os mesmos. A Fig. 4.5 também ilustra estas comparações em termos de um diagrama de barras semi-log. Considerando o aumento do fator de escala desde 1 até 20, o índice LOLP vai de 10^{-3} a 10^{-12} . De fato, desde um fator de escala igual a 5, os eventos de falha no sistema de geração podem ser considerados como raros (i.e. menores que 10^{-5}), e os ganhos computacionais obtidos com a SMC baseada no MEC tornam-se evidentes.

Note que os desempenhos de ambos os métodos, analítico e SMC convencional, decrescem quando o tamanho do sistema aumenta. O método analítico não é consideravelmente sensível à raridade dos eventos envolvidos, porém, ele é ao aumento do número possível de estados, mesmo que um algoritmo numérico extremamente eficiente que evita a *maldição combinatória* (*combinatorial curse*) seja utilizado. O contrário ocorre com a abordagem da SMC convencional, pois ela não é muito sensível ao número de estados do sistema, mas é dramaticamente sensível à raridade dos eventos envolvidos.

O desempenho do CE-MCS é indiferente a ambos os problemas. Ele não tem problemas com o número de estados do sistema, pois é basicamente uma SMC. A raridade dos eventos envolvidos não apresenta problemas devido à distorção ótima aplicada as indisponibilidades das unidades. Problemas envolvendo sistemas com fatores de carga baixos podem ser minimizados pelo uso adequado do parâmetro Φ .

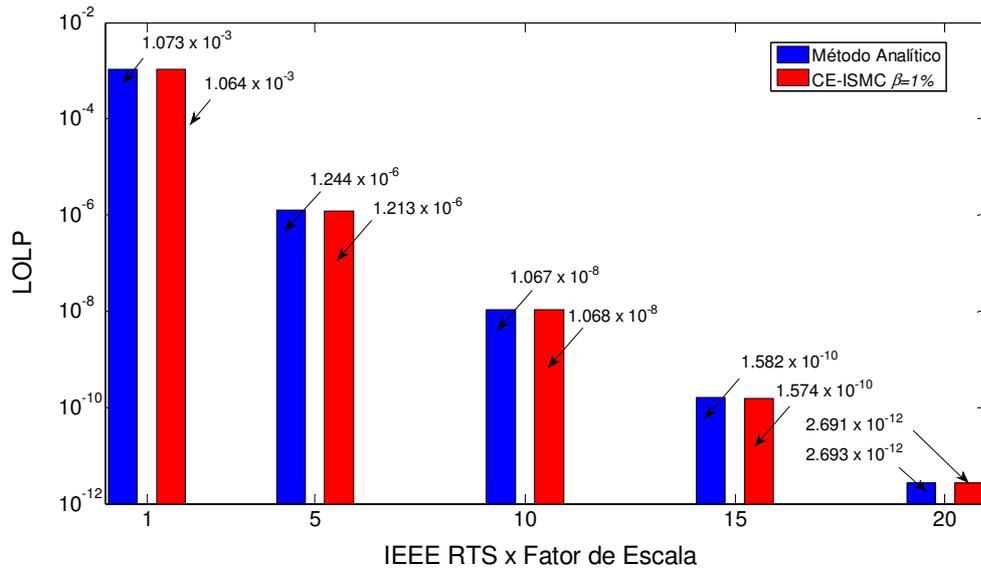


Fig. 4.5: Valores da LOLP para Diferentes Modificações no IEEE-RTS

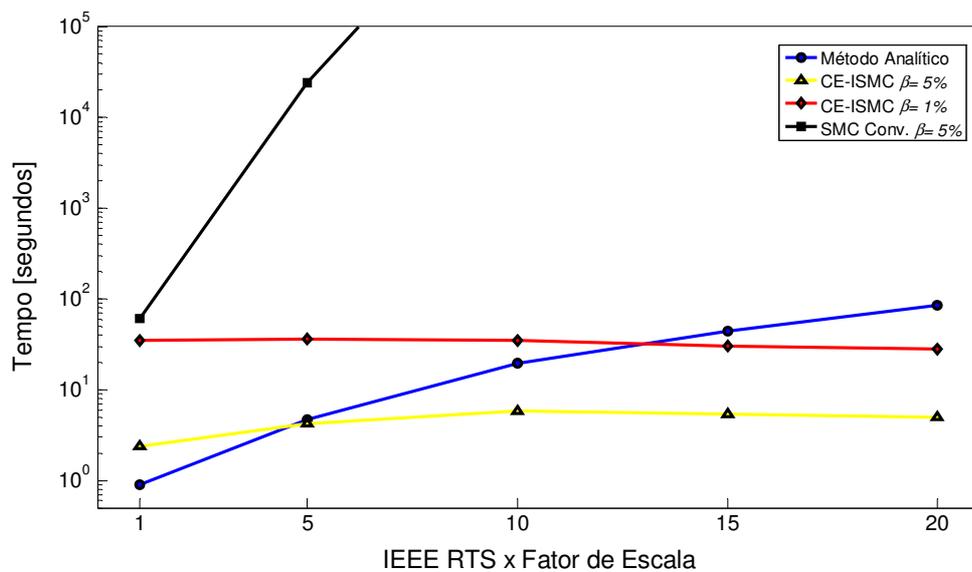


Fig. 4.6: Comparação de Desempenhos – IEEE-RTS com Diferentes Fatores de Escala

A Fig. 4.6 ilustra os desempenhos das metodologias quando aplicadas ao IEEE-RTS com diferentes fatores de escala. Pode ser visto que, enquanto os desempenhos de ambos os métodos (analítico e SMC convencional) ficam piores, o desempenho do CE-ISMC é praticamente constante, o qual neste caso depende principalmente dos valores especificados de β e Φ .

É impressionante observar que enquanto a SMC convencional pode gastar horas (IEEE-RTS×5) ou mesmo anos (IEEE-RTS×20) para concluir a avaliação da confiabilidade, o

CE-ISMC levará alguns segundos para *qualquer sistema*. Além disso, o CE-ISMC torna-se ainda mais veloz do que um método analítico extremamente eficiente, por exemplo, considerando $\beta = 5\%$, o ponto de *corte* é o IEEE-RTS $\times 5$ (i.e. 4,70s > 4,21s).

Observe a Fig. 4.7 onde é ilustrado o processo de convergência para o índice LOLP considerando o IEEE-RTS $\times 5$. É possível observar claramente nesta ilustração a dificuldade que a SMC convencional encontra quando os eventos de falha são raros. Sem dúvida, o desempenho do método baseado em Entropia Cruzada é excepcional.

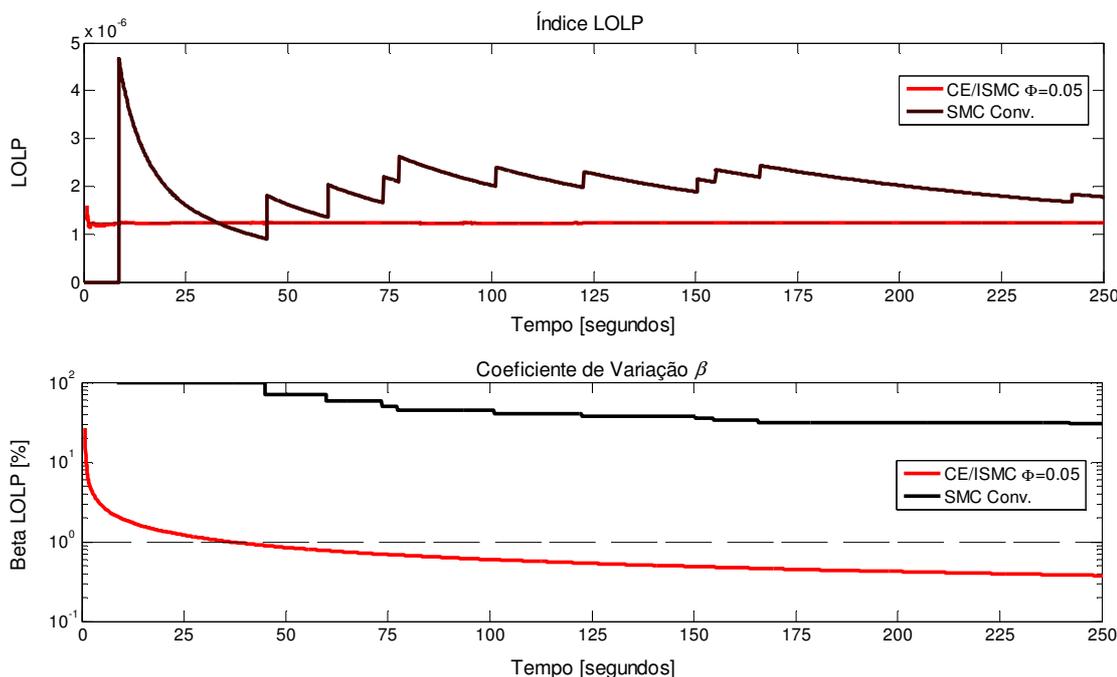


Fig. 4.7: Processo de Convergência do Índice LOLP – IEEE-RTS $\times 5$

4.4. SISTEMA SUL-SUDESTE BRASILEIRO

A metodologia proposta baseada no MEC foi aplicada em duas configurações planejadas, normal e fortemente reforçada, do sistema Sul-Sudeste Brasileiro (SSB) para os anos 90: veja alguns detalhes em [LMC91]. Existem 67 usinas geradoras: 53 usinas hidráulicas e 14 usinas térmicas. Existem 290 unidades com capacidades que variam de 15MW até 700MW (unidades de ITAIPU).

A confiabilidade do sistema SSB é avaliada pelos três métodos: o analítico (com $\Delta=1\text{MW}$),

SMC convencional (com $\beta = 1\%$ ou $N_{MAX} = 5,425 \times 10^8$ amostras), e CE-ISMC (com $\alpha = 1$, $\rho = 0,1$, $N = 2,5 \times 10^4$, $\beta = 1\%$, $\Phi = 0,01$ para a configuração normal e $\Phi = 0,10$ para a configuração fortemente reforçada).

Um modelo de carga horária com 8736 níveis é considerado com uma carga pico de 41,2GW. A configuração normal tem uma capacidade total instalada de 42,8GW, e a fortemente reforçada tem 45,6GW. Os correspondentes valores do índice LOLP são $3,443 \times 10^{-3}$ e $1,909 \times 10^{-5}$, respectivamente. Os principais resultados deste estudo são apresentados nas Tabelas 4.4* e 4.5*.

TABELA 4.4 – RESULTADOS PARA A CONFIGURAÇÃO NORMAL DO SSB

SSB - Normal	LOLP	EPNS [MW]	LOLF [oc./ano]	Tempo de CPU	Número de Estados Visitados
Método Analítico	$3,44369 \times 10^{-3}$	$1,92595 \times 10^0$	$2,98303 \times 10^1$	30,2s	-
SMC Conv. ($\beta=1\%$)	$3,40922 \times 10^{-3}$ 0,74741%	$1,91135 \times 10^0$ 0,99978%	$2,95463 \times 10^1$ 0,78065%	912,6s	$5,2329 \times 10^6$
CE-ISMC ($\beta=1\%$)	$3,45031 \times 10^{-3}$ 0,92227%	$1,93725 \times 10^0$ 0,90152%	$2,98685 \times 10^1$ 0,99986%	17,6s	$1,7606 \times 10^6$
CE-ISMC ($\beta=5\%$)	$3,53631 \times 10^{-3}$ 4,59173%	$1,92110 \times 10^0$ 4,57637%	$3,08155 \times 10^1$ 4,96402%	1,8s	$6,8085 \times 10^4$

TABELA 4.5 – RESULTADOS PARA A CONFIGURAÇÃO REFORÇADA DO SSB

SSB - Reforçada	LOLP	EPNS [MW]	LOLF [oc./ano]	Tempo de CPU	Número de Estados Visitados
Método Analítico	$1,90930 \times 10^{-5}$	$7,04975 \times 10^{-3}$	$1,89561 \times 10^{-1}$	31,2s	-
SMC Conv. ($\beta=1\%$)	$1,90938 \times 10^{-5}$ 0,98251%	$7,02623 \times 10^{-3}$ 1,36264%	$1,89792 \times 10^{-1}$ 0,99995%	26h	$5,4253 \times 10^8$
CE-ISMC ($\beta=1\%$)	$1,89273 \times 10^{-5}$ 0,93130%	$6,98889 \times 10^{-3}$ 0,72333%	$1,87814 \times 10^{-1}$ 1,00000%	31,6s	$5,9896 \times 10^6$
CE-ISMC ($\beta=5\%$)	$1,88883 \times 10^{-5}$ 4,61376%	$7,03400 \times 10^{-3}$ 3,59083%	$1,88757 \times 10^{-1}$ 4,98483%	3,3s	$2,3904 \times 10^5$

Os tempos de processamento para os métodos, analítico, SMC convencional e CE-ISMC,

* Para a abordagem CE-ISMC, na coluna *Número de Estados Visitados*, os valores apresentados consideram somente aqueles estados visitados durante a segunda etapa do algoritmo, i.e. a SMC baseada em IS.

para a configuração normal, são respectivamente: 30,2s, 912,6s, e 17,6s. Os tempos de processamento para a configuração reforçada obtidos para os métodos, analítico, SMC convencional e CE-ISMC, são respectivamente: 31,2s, 26 horas, e 31,6s. Considerando um $\beta = 5\%$, o CE-ISMC gasta 3,3 segundos, o que é muito mais rápido que a abordagem analítica. Torna-se evidente o excelente desempenho da SMC baseada no MEC.

4.5. COMENTÁRIOS ADICIONAIS

Nesta seção serão apresentados os resultados dos principais testes de sensibilidade. Estes resultados são organizados em ilustrações com o objetivo de visualizar melhor o comportamento dos diversos métodos quando certas características dos sistemas testes variam, principalmente, no que diz respeito à raridade dos eventos de falha envolvidos e ao número de estados possíveis do sistema. Os resultados apresentados nesta seção têm como principal objetivo reforçar hipóteses levantadas em seções e capítulos anteriores.

A Fig. 4.8 ilustra os desempenhos obtidos para o IEEE-RTS×1, com L constante no período de análise, quando o valor do índice LOLP varia de aproximadamente 10^{-1} e 10^{-4} . Isto é atingido modificando o pico de carga utilizada. Note que o desempenho do método analítico ($\Delta=1\text{MW}$) é praticamente invariante quando o índice LOLP decresce. Neste caso, como o sistema é pequeno, este método apresenta o melhor desempenho.

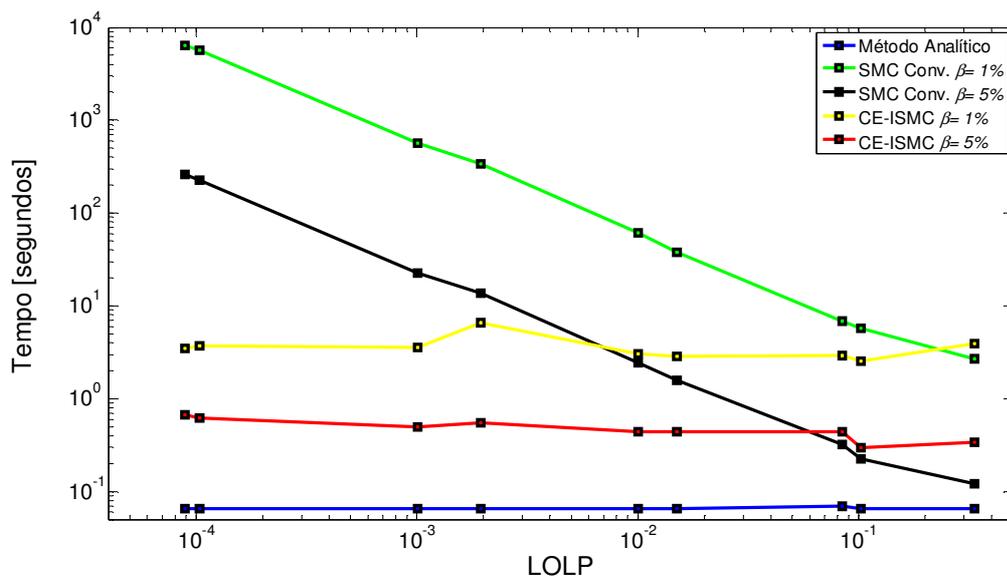


Fig. 4.8: Raridade do Evento de Falha – IEEE-RTS×1

O CE-ISMC ($\alpha = 1$, $\rho = 0,1$, $N = 2,5 \times 10^4$) também apresenta um desempenho relativamente constante. Porém a SMC convencional tem seu desempenho afetado de maneira quase linear considerando uma escala logarítmica em ambos os eixos. Quando os eventos de falha são da ordem de 10^{-1} o SMC convencional apresenta melhor desempenho que o CE-ISMC sendo esta diferença basicamente o tempo necessário para estimar $\hat{\mathbf{v}}_K^{\text{LOLP}}$.

A Fig. 4.9 ilustra os desempenhos obtidos para o IEEE-RTS $\times 20$ com L constante no pico, quando o valor do índice LOLP varia de aproximadamente 10^{-1} e 10^{-4} . Note que os desempenhos dos métodos baseados em simulação não são significativamente alterados pelo aumento da dimensão do sistema, porém o método analítico teve uma redução expressiva no seu desempenho.

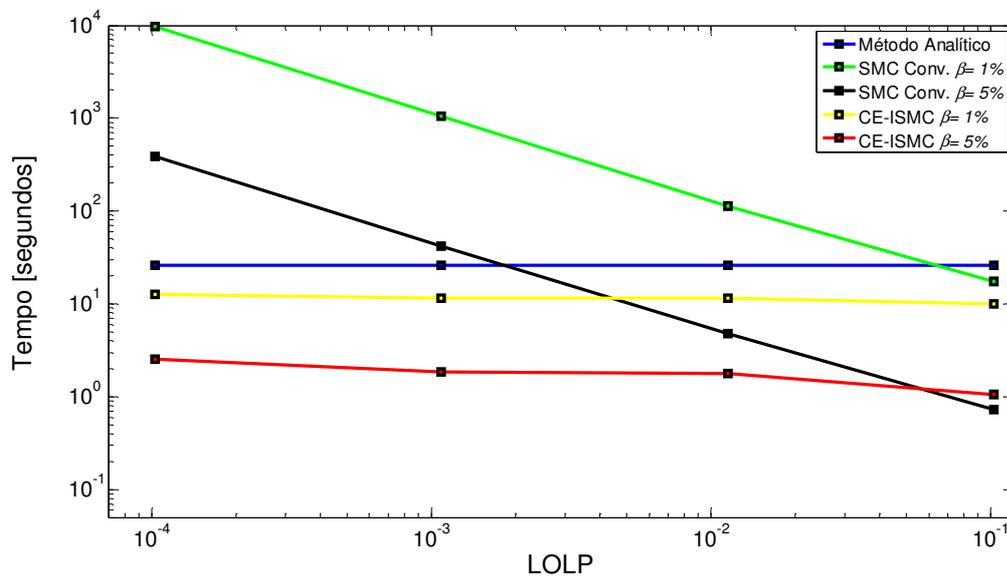


Fig. 4.9: Raridade do Evento de Falha – IEEE-RTS $\times 20$

Considerando ambas as ilustrações pode-se concluir o já exposto na seção anterior. O método analítico apresenta problemas se o número de estados possíveis do sistema é elevado, mas é indiferente ao valor da LOLP. Na SMC convencional ocorre o contrário sendo a raridade dos eventos de falha o principal problema e finalmente no CE-ISMC nenhum destes problemas representam limitações para o método.

Quando o modelo de L é representado por múltiplos estados, o parâmetro Φ e também o processo de amostragem dos estados de L passam a ser relevantes. Duas maneiras de amostragem diferentes foram implementadas computacionalmente. A primeira consiste no mapeamento da curva de carga pela distribuição uniforme discreta conforme a Fig. 3.6. A

segunda consiste em utilizar o modelo markoviano sequencial ilustrado na Fig. 3.7 e considerar que o número de amostras é grande.

A Fig. 4.10 apresenta os tempos de simulação necessários para avaliar os índices GCR ($\beta=5\%$) para o IEEE-RTS $\times 10$, considerando o algoritmo do CE-ISMC implementado com e sem o parâmetro Φ para ambas as técnicas de amostragem de L . O valor do índice LOLP é controlado variando o perfil da curva de carga mantendo constante o fator de carga do sistema. Para todas as simulações, a seguinte configuração de parâmetros foi utilizada: $\alpha = 1$, $\rho = 0,1$, $N = 2,5 \times 10^4$, $\Phi = 0,01$ (se for o caso).

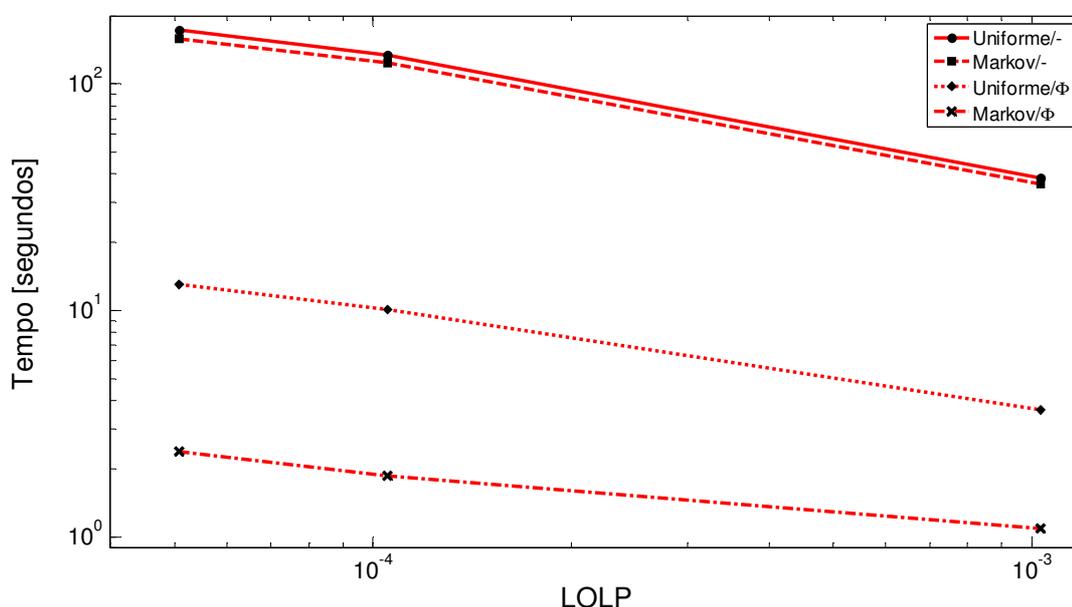


Fig. 4.10: Eficiência do Parâmetro Φ

É possível verificar que o melhor desempenho é obtido quando o parâmetro Φ é utilizado em conjunto com o modelo markoviano sequencial (denominado Markov/ Φ). O mesmo ocorreu em todas as simulações realizadas e por este motivo esta configuração foi utilizada em todos os resultados apresentados anteriormente (exceto no caso de L ser constante no período de análise).

A Tabela 4.6 apresenta os resultados obtidos para o pior caso ilustrado na Fig. 4.10. Nesta condição o valor do pico de carga é 31070MW (o IEEE-RTS $\times 10$ original tem carga pico de 28500MW e $LOLP=1,06714 \times 10^{-8}$). O tempo necessário para a avaliação analítica é aproximadamente de 21 segundos. A SMC convencional com $\beta=5\%$ requer 764 segundos. O

CE-ISMIC sem o parâmetro Φ apresenta um melhor desempenho que a SMC convencional, mas este desempenho é muito inferior se comparado com o método analítico. O principal motivo é a ineficiente amostragem dos estados de L .

Quando um parâmetro Φ é especificado mesmo com um valor baixo (para o índice LOLP considerado), o desempenho do CE-ISMIC melhora drasticamente sem comprometer a precisão das estimativas. Considerando a melhor configuração (Markov/ Φ), o CE-ISMIC atinge o critério de convergência $\beta=5\%$ em aproximadamente 2,4 segundos.

TABELA 4.6 – IEEE-RTS $\times 10$ COM CARGA PICO DE 31070MW

IEEE-RTS $\times 10$	LOLP	LOLE [horas/ano]	EPNS [MW]	EENS [MWh/ano]	LOLF [oc./ano]	LOLD [horas/oc.]	Tempo [s]
Método Analítico	$5,08688 \times 10^{-5}$	$4,44390 \times 10^{-1}$	$1,94609 \times 10^{-2}$	$1,70010 \times 10^2$	$2,10121 \times 10^{-1}$	2,11492	20,89
SMC Conv.	$5,00875 \times 10^{-5}$	$4,37564 \times 10^{-1}$	$1,98512 \times 10^{-2}$	$1,73420 \times 10^2$	$2,04341 \times 10^{-1}$	2,14135	764,43
CE-ISMIC Uniforme/-	$5,03952 \times 10^{-5}$	$4,40252 \times 10^{-1}$	$1,89205 \times 10^{-2}$	$1,65290 \times 10^2$	$2,12664 \times 10^{-1}$	2,07018	172,50
CE-ISMIC Markov/-	$5,05938 \times 10^{-5}$	$4,41988 \times 10^{-1}$	$1,95323 \times 10^{-2}$	$1,70634 \times 10^2$	$2,07032 \times 10^{-1}$	2,13488	158,67
CE-ISMIC Uniforme/Φ	$5,01788 \times 10^{-5}$	$4,38362 \times 10^{-1}$	$1,91362 \times 10^{-2}$	$1,67174 \times 10^2$	$2,14206 \times 10^{-1}$	2,04646	13,05
CE-ISMIC Markov/Φ	$5,00086 \times 10^{-5}$	$4,36875 \times 10^{-1}$	$1,95510 \times 10^{-2}$	$1,70797 \times 10^2$	$2,13394 \times 10^{-1}$	2,04727	2,37

Considerando que neste caso o valor do índice LOLP é aproximadamente $5,1 \times 10^{-5}$, o parâmetro Φ poderia ter sido especificado com um valor maior que 0,01 e desta maneira melhores desempenhos poderiam ser obtidos. Contudo existe um limite ou *saturação* onde aumentar o valor do parâmetro Φ não apresenta mais ganhos significativos. Foi verificado por diversas simulações que este limite depende principalmente do valor do índice LOLP, possibilitando assim definir valores típicos por faixas (conforme apresentado na subseção 3.4.5 deste trabalho).

Conforme exposto anteriormente, utilizar o parâmetro Φ implica um compromisso entre a qualidade do resultado e o tempo computacional. Este compromisso é ilustrado na Fig. 4.11 considerando o sistema IEEE-RTS $\times 10$ em sua configuração normal, com o seguinte conjunto de parâmetros: $\alpha = 1$, $\rho = 0,1$, $N = 2,5 \times 10^4$, $\beta = 1\%$. A Fig. 4.11a apresenta os valores das estimativas obtidas para o índice LOLP, e a Fig. 4.11b, os respectivos tempos de

simulação. Observe que aumentado o valor do parâmetro Φ melhores desempenhos são obtidos em termos de tempo de simulação, porém, após certo valor de Φ , a relação custo/benefício não justifica seguir aumentando o valor deste parâmetro.

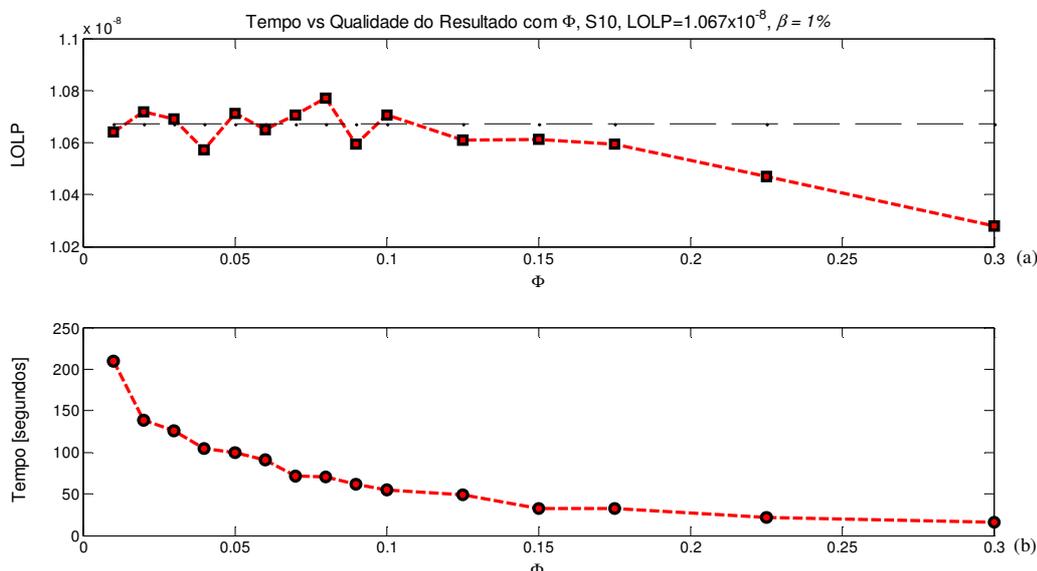


Fig. 4.11: Parâmetro Φ – Compromisso entre Qualidade e Esforço Computacional

Note que a estimativa encontrada para $\Phi = 0,3$ não é necessariamente *ruim* (dependendo do critério adotado), porém, o resultado obtido para $\Phi = 0,15$ é muito melhor não existindo diferença significativa entre os tempos de simulação.

4.6. COMENTÁRIOS FINAIS

Este capítulo apresentou a metodologia proposta, CE-ISMC, aplicada em diversos sistemas testes e também em uma configuração planejada do sistema Sul-Sudeste Brasileiro para os anos 90. Em cada caso, o desempenho do CE-ISMC foi comparado com um método analítico extremamente eficiente e com uma SMC não-cronológica convencional.

A metodologia proposta apresentou vantagens competitivas evidentes em comparação a SMC convencional, principalmente em sistemas de geração muito confiáveis. O método também obteve um excelente desempenho em sistemas com elevado número de estados, sendo inclusive mais eficiente que o método analítico em certas condições.

Análises de sensibilidade foram apresentadas e discutidas, principalmente no que diz respeito à raridade dos eventos de falha e ao tamanho do sistema. A eficiência computacional do parâmetro Φ foi comprovada assim como sua habilidade para tornar eficiente o processo de amostragem dos estados de carga.

Diversos testes e simulações foram realizados que não se encontram documentados neste trabalho. O principal objetivo foi verificar os limites da metodologia proposta e detectar possíveis problemas numéricos. Sistemas hipotéticos de geração extremamente confiáveis foram testados utilizando fatores de escala maiores e adicionando novas unidades com diferentes capacidades. Diferentes perfis de curva de carga foram também utilizados modificando convenientemente o valor do pico e do fator de carga.

Em todos os sistemas testados a metodologia proposta apresentou ótimos desempenhos. Foi verificado que com um conjunto conveniente de parâmetros (principalmente N , Φ e β), e considerando os mesmos recursos computacionais, o CE-ISMC fornece boas estimativas dos índices GCR para qualquer sistema de geração em não mais de 2 minutos.

CAPÍTULO 5

CONCLUSÕES

AS ferramentas baseadas em Simulação Monte Carlo (SMC) são extremamente robustas para resolver problemas em confiabilidade de sistemas de potência, particularmente para sistemas de grande porte, onde o grande número de estados possíveis dificulta a aplicação de métodos analíticos. Porém, eles podem encontrar alguns obstáculos no que diz respeito a eventos raros, por exemplo, para avaliar valores muito pequenos para os índices LOLP, i.e. 10^{-5} . Embora se possa argumentar que uma configuração muito confiável para um sistema de geração não é um problema, e que a simulação deve ser suspensa após verificar tal condição, em estudos de planejamento da expansão estes valores devem ser calculados considerando que diversas configurações de reforço estão sendo comparadas. Além do mais, em problemas que envolvem critérios específicos de projeto, o planejador poderia ter que medir a probabilidade de certo evento, que poderia ser muito raro, mas, a priori, esta condição é desconhecida.

Esta dissertação apresentou uma nova abordagem para avaliar os índices de confiabilidade da capacidade de geração baseada em técnicas de SMC e no Método da Entropia Cruzada (MEC). A idéia básica da metodologia proposta foi utilizar uma função massa de probabilidade auxiliar cujos parâmetros foram obtidos através de um processo de otimização estocástico. Esta nova função foi utilizada em conjunto com a já conhecida técnica da amostragem por importância (IS) a fim de reduzir a variância amostral. Foi demonstrado que reduzindo a variância do estimador é possível obter ganhos significativos no desempenho das simulações.

O desempenho da abordagem proposta baseada no MEC foi testado em diversas condições diferentes, incluindo sistemas de pequeno, médio e grande porte com diferentes níveis de confiabilidade. Os resultados obtidos foram comparados com um método analítico extremamente eficiente e também com uma simulação Monte Carlo convencional. Foi demonstrado o desempenho excepcional do método proposto baseado em Entropia Cruzada em comparação às abordagens anteriores. É mostrado que a abordagem proposta atinge a convergência especificada em poucos segundos, enquanto a SMC convencional

poderia levar horas, meses ou até anos para garantir a mesma precisão.

No Capítulo 2 foi discutida a formulação matemática do problema, assim como um resumo dos principais métodos utilizados no presente para avaliar a confiabilidade de geração com suas respectivas vantagens e desvantagens. Em geral, os métodos analíticos apresentam problemas quando o número de estados possíveis do sistema aumenta, mas não apresentam problemas no que diz respeito à raridade dos eventos de falha envolvidos. No caso da SMC ocorre o contrário, pois o esforço computacional desta técnica não é muito sensível ao tamanho dos sistemas, mas é extremamente sensível à raridade dos eventos de falha. Foi demonstrado por diversas simulações que o método proposto baseado em Entropia Cruzada é indiferente a ambos os problemas, permitindo a avaliação dos índices de confiabilidade em poucos segundos para quaisquer que sejam as características do sistema de geração.

O Capítulo 3 apresentou o Método da Entropia Cruzada para simulação de eventos raros e como este método pode ser reformulado matematicamente, partindo da sua formulação básica, para ser aplicado ao problema de avaliar a confiabilidade da capacidade de geração. Foi discutido e verificado que a distorção ótima do espaço de estados encontrada para o índice LOLP pode ser utilizada para estimar os outros índices de confiabilidade, i.e. EPNS e LOLF. Portanto, todo o desenvolvimento matemático e computacional pode ser feito considerando somente o índice LOLP, e somente um vetor de parâmetros de referência ótimo é necessário.

Considerando um modelo de carga variável no tempo, diversas simulações demonstraram que o vetor de parâmetros de referência ótimo encontrado para um nível fixo L , no pico de carga ou muito próximo deste pico, apresenta um excelente desempenho computacional. Neste capítulo também foi introduzido o parâmetro Φ e ficou demonstrada a sua eficiência em resolver problemas inerentes ao baixo fator de potência em alguns sistemas de geração.

No Capítulo 4, foram realizadas diversas aplicações numéricas com o *IEEE Reliability Test System* (IEEE-RTS) [IEEE79] e também com diversas modificações deste sistema. O desempenho do método proposto baseado em Entropia Cruzada foi comparado com uma simulação Monte Carlo não-sequencial e também com um método analítico extremamente eficiente que evita a computação combinatória. Uma configuração do sistema de geração

brasileiro da região Sul-Sudeste também foi utilizada para demonstrar a aptidão do método proposto em aplicações reais.

A partir dos diversos testes realizados, pode-se concluir que o método proposto baseado em Entropia Cruzada pode ser satisfatoriamente aplicado para avaliar os índices de confiabilidade de qualquer sistema de geração, independentemente do maior ou menor grau de confiabilidade do mesmo, e também do seu tamanho ou características de carregamento. Foi demonstrado que na medida em que o tamanho do sistema aumenta, os desempenhos dos métodos analítico e SMC se deteriora, enquanto o desempenho da metodologia proposta permanece praticamente constante para qualquer sistema, sendo que neste caso o desempenho depende principalmente dos valores especificados para os parâmetros β e Φ .

Algumas sugestões para trabalhos futuros são as seguintes:

- Avaliar a confiabilidade da capacidade de geração através da simulação Monte Carlo Cronológica, auxiliada pelo Método da Entropia Cruzada;
- Avaliar a confiabilidade composta (i.e. geração e transmissão) de sistemas elétricos de potência através do Método da Entropia Cruzada;
- Comparar os desempenhos computacionais de diversos métodos existentes na literatura com métodos baseados em Entropia Cruzada, por exemplo, com a SMC pseudo-cronológica;
- Aplicação do Método da Entropia Cruzada no planejamento da expansão de sistemas elétricos de potência.

REFERÊNCIAS BIBLIOGRÁFICAS

- [A90] G.J. Anders, “Probability Concepts in Electric Power Systems”, New York: John Wiley and Sons, 1990.
- [ABA86] R.N. Allan, R. Billinton, N.M.K. Abdel-Gawad, “The IEEE Reliability Test System - Extensions to and Evaluation of the Generating System”, *IEEE Transactions on Power Systems*, Vol. 1, No. 4, pp. 1-7, November 1986.
- [ABL84] R.N. Allan, R. Billinton, S.H. Lee, “Bibliography on the Application of Probability Methods in Power System Reliability Evaluation 1977-1982”, *IEEE Transactions on Power Apparatus and Systems*, Vol. 103, No. 2, pp. 275-282, February 1984.
- [AEPP90] G.J. Anders, J. Endrenyi, M.V.F. Pereira, L.M.V.G. Pinto, “Fast Monte Carlo Simulation Techniques for Power System Reliability Studies”, *CIGRE Symposium*, Paris, Paper 38-205, 1990.
- [ALAB81] R.N. Allan, A.M. Leite da Silva, A.A. Abu-Nasser, R.C. Burchett, “Discrete Convolution in Power System Reliability”, *IEEE Transactions on Reliability*, Vol. 30, No. 5, pp. 452-456, December 1981.
- [BA96] R. Billinton, R.N. Allan, “Reliability Evaluation of Power Systems”, 2nd Edition, New York: Plenum, 1984, 1996.
- [BAS94] R. Billinton, R.N. Allan, L. Salvaderi (Editors), “Applied Reliability Assessment in Electric Power Systems”, New York: IEEE Press, 1994.

- [BBGH59] C.J. Baldwin, J.E. Billings, D.P. Gaver, C.H. Hoffman, "Mathematical Models for Use in the Simulation of Power Generating Outages: II – Power Systems Forced Outage Distributions", *AIEE Transactions*, Vol. PAS-78, pp. 1258-1272, 1959.
- [BFB01] R. Billinton, M. Fotuhi-Firuzabad, L. Bertling, "Bibliography on the Application of Probability Methods in Power System Reliability Evaluation 1996–1999", *IEEE Transactions on Power Systems*, Vol. 16, No. 4, pp. 595-602, November 2001.
- [BGH59] C.J. Baldwin, D.P. Gaver, C.H. Hoffman, "Mathematical Models for Use in the Simulation of Power Generating Outages: I – Fundamental Considerations", *AIEE Transactions*, Vol. PAS-78, pp. 1251-1258, 1959.
- [BGHR60] C.J. Baldwin, D.P. Gaver, C.H. Hoffman, J.A. Rose, "Mathematical Models for Use in the Simulation of Power Generating Outages: III – Models for Large Interconnection", *AIEE Transactions*, Vol. PAS-78, pp. 1645-1650, 1960.
- [BJ96] R. Billinton, A. Jonnavithula, "Variance Reduction Techniques for use with Sequential Monte Carlo Simulation in Bulk Power System Reliability Evaluation", *Canadian Conference on Electrical and Computer Engineering 1996*, Vol. 1, pp. 416-419, Calgary, May 1996.
- [BJ97] R. Billinton, A. Jonnavithula, "Composite System Adequacy Assessment Using Sequential Monte Carlo Simulation with Variance Reduction Techniques", *IEE Proceedings on Generation, Transmission and Distribution.*, Vol. 144, No. 1, pp.1-6, January 1997.
- [BKMR05] P.T. de Boer, D.P. Kroese, S. Mannor, R.Y. Rubinstein, "A Tutorial on the Cross-Entropy Method", *Annals of Operations Research*, Vol. 134, No.1, pp. 19-67, January 2005.

- [BL94] R. Billinton, W. Li, "Reliability Assessment of Electrical Power Systems Using Monte Carlo Methods", New York: Plenum, 1994.
- [BWH82] R. Billinton, C.L. Wee, G. Hamoud, "Digital Computer Algorithms for the Calculation of Generating Capacity Reliability Indices", *IEEE Transactions on Power Apparatus and Systems*, Vol. 101, No. 1, pp. 203-211, January 1982.
- [C47] G. Calabrese, "Generating Reserve Capacity Determined by the Probability Method", *AIEE Transactions*, Vol. 66, pp. 1439-1450, 1947.
- [C90] G.M. Clark, "Use of common random numbers in comparing alternatives", *Proceedings of the 1990 Winter Simulation Conference*, pp. 367-371, New Orleans, U.S.A., December 1990.
- [CGOP82] S.H.F. Cunha, F.B.M. Ganes, G.C. Oliveira, M.V.F. Pereira, "Reliability Evaluation in Hydrothermal Generating Systems", *IEEE Transactions on Power Apparatus and Systems*, Vol. 101, No. 12, pp. 4665-4673, December 1982.
- [CS86] Q. Chen, C. Singh, "Equivalent Load Method for Calculating Frequency & Duration Indices in Generation Capacity Reliability Evaluation", *IEEE Transactions on Power Systems*, Vol. 1, No. 1, pp. 101-107, February 1986.
- [E78] J. Endrenyi, "Reliability Modeling in Electric Power Systems". Chichester, U.K.: John Wiley and Sons, 1978.
- [EGSMW07] D. Ernst, M. Glavic, G. Stan, S. Mannor, L. Wehenkel, "The Cross-Entropy Method for Power System Combinatorial Optimization Problems", *Power Tech, 2007 IEEE Lausanne*, pp. 1290-1295, Lausanne, Switzerland, July 2007.

- [F09] P.L. Fackler, “Notes on Simulation and Monte Carlo”, *Notas de Aula*, disponível no site <http://www4.ncsu.edu/unity/users/p/pfackler/www/ECG790C/MCNotes.pdf>, January 2009.
- [HA58] H. Halperin, H.A. Adler, “Determination of Reserve Generating Capacity”, *IAEE Transactions*, Vol. PAS-77, pp. 530-544, 1958.
- [HBKK05] K.P. Hui, N. Bean, M. Kraetzl, D.P. Kroese, “The Cross-Entropy Method for Network Reliability Estimation”, *Annals of Operations Research*, Vol. 134, pp. 101-118, 2005.
- [HR02] T. Homem-de-Mello, R. Y. Rubinstein, “Estimation of Rare Event Probabilities Using Cross-Entropy”, *Proceedings of the 2002 Winter Simulation Conference*, Vol. 1, pp. 310-319, San Diego, U.S.A., December 2002.
- [IEEE79] Reliability Test System Task Force of the IEEE Subcommittee on the Application of Probability Methods, “IEEE Reliability Test System”, *IEEE Transactions on Power Apparatus and Systems*, Vol. 98, pp. 2047-2054, November/December 1979.
- [KHN07] D.P. Kroese, K.P. Hui, S.Nariai, “Network Reliability Optimization via the Cross-Entropy Method”, *IEEE Transactions on Reliability*, Vol. 56, No. 2, pp. 275-287, June 2007.
- [L05] W. Li, “Risk Assessment of Power Systems – Models, Methods, and Applications”, New York: IEEE Press, 2005.
- [L47] W.J. Lyman, “Calculating Probability of Generating Capacity Outages”, *AIEE Transactions*, Vol. 66, pp. 1471-1477, 1947.

- [LCM92] A.M. Leite da Silva, J. Coelho, A.C.G. Melo, "Uncertainty Considerations in Frequency and Duration Analysis for Large Hydrothermal Generating Systems", *IEE Proceedings Part C*, Vol. 138, No. 3, pp. 277-285, May 1992.
- [LMC91] A.M. Leite da Silva, A.C.G. Melo, S.H.F. Cunha, "Frequency and Duration Method for Reliability Evaluation of Large-Scale Hydrothermal Generating Systems", *IEE Proceedings on Generation, Transmission and Distribution*, Vol. 138, No. 1, pp. 94-102, January 1991.
- [LMMB00] A.M. Leite da Silva, L.A.F. Manso, J.C.O. Mello, R. Billinton, "Pseudo-Chronological Simulation for Composite Reliability Analysis with Time Varying Loads", *IEEE Transactions on Power Systems*, Vol. 15, No. 1, pp. 73-80, February 2000.
- [LNR99] D. Lieber, A. Nemirovskii, R.Y. Rubinstein, "A Fast Monte Carlo Method for Evaluating Reliability Indexes", *IEEE Transactions on Reliability*. Vol. 48, No. 3, pp. 256-261, September 1999.
- [LPC88] A.M. Leite da Silva, F.A.F. Plazo Blanco, J. Coelho, "Discrete Convolution in Generating Capacity Reliability Evaluation - LOLE Calculations and Uncertainty Aspects", *IEEE Transactions on Power Systems*, Vol. 3, No. 4, pp. 1616-1624, November 1988.
- [LRMB04] A.M. Leite da Silva, L.C. Resende, L.A.F. Manso, R. Billinton, "Well-Being Analysis for Composite Generation and Transmission Systems", *IEEE Transactions on Power Systems*, Vol. 19, No. 4, pp. 1763-1770, November 2004.
- [LW47] E.S. Loane, C.W. Watchorn, "Probability Methods Applied to Generating Capacity Problems of a Combined Hydro and Steam Systems", *AIEE Transactions*, Vol. 66, pp. 1645-1657, 1947.

- [M86] A.C. Geber de Melo, “Avaliação da Confiabilidade de Geração através do Método da Frequência e Duração”, *Tese de Mestrado*, Pontifícia Universidade Católica do Rio de Janeiro, PUC-RJ, Agosto 1986.
- [MCRLS09] V. Miranda, L.M. Carvalho, M.A. Rosa, A.M. Leite da Silva, C. Singh, “Improving Power System Reliability Calculation Efficiency with EPSO Variants”, *IEEE Transactions on Power Systems – aceito para publicação*, 2009.
- [MPL92] A.C.G. Melo, M.V.F. Pereira, A.M. Leite da Silva, “Frequency and Duration Calculations in Composite Generation and Transmission Reliability Evaluation”, *IEEE Transactions on Power Systems*, Vol. 7, No. 2, pp. 469-476, May 1992.
- [PAK08] M. Perninge, M. Amelin, V. Knazkins, "Comparing Variance Reduction Techniques for Monte Carlo Simulation of Trading and Security in a Three-Area Power System", *Transmission and Distribution Conference and Exposition: Latin America 2008 IEEE/PES*, pp. 1-5, Bogota, Colombia, August 2008.
- [PB92] M.V.F. Pereira, N.J. Balu, “Composite Generation/Transmission Reliability Evaluation (invited paper)”, *Proceedings of the IEEE*, Vol. 80, No. 4, pp. 470-491, April 1992.
- [PBB88] A.D.Patton, J.H. Blackstone, N.J. Balu, “A Monte Carlo Simulation to the Reliability Modeling of Generating Systems Recognizing Operating Considerations”, *IEEE Transactions on Power Systems*, Vol. 3, No. 3, pp. 1174-1180, August 1988.
- [QZW07] Y. Qiu, H. Zhou, Y. Wu, “An Importance Sampling Method with Applications to Rare Event Probability”, *Proceedings of 2007 IEEE International Conference on Grey Systems and Intelligent Services*, pp. 1381-1385, Nanjing, China, November 2007.

- [QZW08] Y. Qiu, H. Zhou, Y.Wu, "An Importance Sampling Method Based on Martingale with Applications to Rare Event Probability", *Proceedings of the 7th World Congress on Intelligent Control and Automation*, pp. 4041-4045, Chongqing, China, June 2008.
- [R91] R.Y. Rubinstein, "Simulation and the Monte Carlo Method", New York: John Wiley and Sons, 1991.
- [RK04] R.Y. Rubinstein, D.P. Kroese, "The Cross-Entropy Method. A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning", New York: Springer, 2004.
- [RK07] R.Y. Rubinstein, D.P. Kroese, "Simulation and the Monte Carlo Methods", Second Edition, New York: John Wiley and Sons, 2007.
- [SB77] C. Singh, R. Billinton, "System Reliability Modelling and Evaluation", London, U.K.: Hutchinson, 1977.
- [SR08] M.Th. Schilling, A.M. Rei, "Reliability Assessment of the Brazilian Power System Using Enumeration and Monte Carlo", *IEEE Transactions on Power Systems*, Vol. 23, No. 3, August 2008.
- [WS08] L.Wang, C. Singh, "Population-Based Intelligent Search in Reliability Evaluation of Generation Systems With Wind Power Penetration", *IEEE Transactions on Power Systems*, Vol. 23, No. 3, pp. 1336-1345, August 2008.
- [ZX02] B. Zhaohong, W. Xifan, "Studies on Variance Reduction Technique of Monte Carlo Simulation in Composite System Reliability Evaluation", *Electric Power Systems Research*, Vol. 63, No. 1, pp. 59-64, August 2002.



UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

REINALDO ANDRÉS GONZÁLEZ FERNÁNDEZ

**AVALIAÇÃO DA CONFIABILIDADE DA CAPACIDADE DE
GERAÇÃO VIA O MÉTODO DA ENTROPIA CRUZADA**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica como parte dos requisitos para obtenção do Título de Mestre em Ciências em Engenharia Elétrica

Área de Concentração: Sistemas Elétricos de Potência

Orientador: Prof. Dr. Armando Martins Leite da Silva

Junho 2009

Itajubá – MG – Brasil

Ficha catalográfica elaborada pela Biblioteca Mauá –
Bibliotecária Margareth Ribeiro- CRB_6/1700

G643a

González Fernández, Reinaldo Andrés

Avaliação da confiabilidade da capacidade de geração via o
Método da Entropia Cruzada / Reinaldo Andrés González Fernán_
dez. -- Itajubá, (MG) : [s.n.], 2009.

84 p. : il.

Orientador: Prof. Dr. Armando Martins Leite da Silva.
Dissertação (Mestrado) – Universidade Federal de Itajubá.

1. Confiabilidade da capacidade de geração. 2. Simulação de
Monte Carlo. 3. Método da Entropia Cruzada. 4. Amostragem por
importância. I. Silva, Armando Martins Leite da, orient. II. Univer_
sidade Federal de Itajubá. III. Título.

CDU 621.311(043)



Ministério da Educação
UNIVERSIDADE FEDERAL DE ITAJUBÁ
Criada pela Lei nº 10.435, de 24 de abril de 2002

A N E X O II

FOLHA DE JULGAMENTO DA BANCA EXAMINADORA

Título da Dissertação: **“Avaliação da Confiabilidade da Capacidade de Geração Via o Método da Entropia Cruzada”**

Autor: **REINALDO ANDRÉS GONZALEZ FERNÁNDEZ**

JULGAMENTO

Examinadores	Conceito	Rubrica
	A = Aprovado - R = Reprovado - I = Insuficiente	
1º	A	
2º	A	
3º	A	AML

Observações:

- (1) O Trabalho será considerado Aprovado se todos os Examinadores atribuírem conceito A.
 - (2) O Trabalho será considerado Reprovado se forem atribuídos pelos menos 2 conceitos R.
 - (3) O Trabalho será considerado Insuficiente (I) se for atribuído pelo menos um conceito R. Neste caso o candidato deverá apresentar novo trabalho. A banca deve definir como avaliar a nova versão da Dissertação.
- Este documento terá a validade de 60 (sessenta) dias a contar da data da defesa da Dissertação.**

Resultado Final: Conceito: A, ou seja, Aprovado

Observações: _____

Itajubá, 12 de Junho de 2009.

Prof. Dr. Leônidas Chaves de Resende
1º Examinador – UFSJ

Prof. Dr. João Guilherme de Carvalho Costa
2º Examinador – UNIFEI

Prof. Dr. Armando Martins Leite da Silva
3º Examinador – UNIFEI – (Orientador)

AGRADECIMENTOS

- Ao Professor Armando Martins Leite da Silva, por toda a dedicação, ensinamentos, compreensão e pelo excelente trabalho de orientação ao longo do desenvolvimento desta dissertação.
- Aos colegas e amigos Luís Lima, Luís Nascimento, Warley e Leandro por toda a ajuda e apoio prestados desde o meu ingresso no programa de pós-graduação.
- Ao Programa de Apoio ao Plano de Reestruturação e Expansão das Universidades Federais (REUNI) pelo apoio financeiro.
- Ao meu querido pai Reinaldo Andrés J. González Cocian, por todo o amor, apoio e grande incentivo em todos meus projetos e decisões importantes.
- À minha mãe Miryan R. Fernández de González por todo o amor, carinho, educação e pelas suas constantes orações que sempre me deram forças para seguir adiante.
- Aos meus irmãos Paulo Cesar, Delia Diana, Eveline Maria e à minha noiva Marta Esperanza, por todo o amor, paciência e compreensão em todos os momentos de minha vida.

*Aos meus queridos pais Reinaldo e Miryan,
Aos meus irmãos Paulo, Delia e Eveline, e
Ao meu grande amor Marta Esperanza.*

RESUMO

Nos estudos de planejamento da expansão de sistemas elétricos de potência, técnicas são necessárias para quantificar os custos e a confiabilidade das diversas alternativas. Nas últimas décadas, os índices de confiabilidade da capacidade de geração (*generating capacity reliability* - GCR) têm sido parâmetros muito úteis para auxiliar engenheiros de planejamento em muitas decisões. Estes índices são medidas estatísticas que indicam se a capacidade de geração disponível é suficiente para atender à carga total do sistema dentro de padrões técnico-operativos predeterminados. Em problemas que envolvem critérios específicos de projeto, o planejador poderia ter que medir a probabilidade de ocorrência para certo evento que poderia ser raro, mas, a priori, esta condição é desconhecida.

Esta dissertação apresenta uma nova metodologia para avaliar os índices GCR baseada em técnicas de Simulação Monte Carlo (SMC) e no Método da Entropia Cruzada. A idéia básica é utilizar uma função massa de probabilidade auxiliar, cujos parâmetros são obtidos através de um processo de otimização estocástico. Esta nova função é utilizada em conjunto com a já conhecida técnica da amostragem por importância a fim de reduzir drasticamente a variância amostral e, conseqüentemente, o esforço computacional do processo de estimação baseado em SMC.

A fim de verificar completamente o desempenho da metodologia proposta, diversos aspectos como o tamanho do sistema, raridade dos eventos de falha, número de unidades diferentes e perfil da curva de carga são considerados. O IEEE-RTS é utilizado para testar a metodologia proposta. Também, diversas modificações são feitas neste sistema a fim de comparar o desempenho da abordagem proposta com uma SMC não-cronológica e uma técnica analítica extremamente eficiente baseada em convolução discreta. Uma configuração do sistema Sul-Sudeste Brasileiro (SSB) de geração é também utilizada para demonstrar a aptidão do método proposto em aplicações reais.

ABSTRACT

In power systems expansion planning studies, techniques are needed to quantify the costs and reliabilities for each of the several alternatives. In the last few decades, the generation capacity reliability (GCR) indices have been very useful parameters to help planning engineers in many decisions. These GCR indices are statistical measures that indicate if the available generation capacity is sufficient to meet the total system load, considering predetermined technical-operative standards. In problems involving specific design criteria, the planner may have to measure the probability of occurrence for a certain event which may be rare, but, a priori, this condition is unknown.

This dissertation presents a new methodology to evaluate the GCR indices based on Monte Carlo Simulation (MCS) techniques and the Cross-Entropy Method. The basic idea is to use an auxiliary probability mass function, whose parameters are obtained from a stochastic optimization process. This new function is used in combination with the well known Importance Sampling technique to drastically reduce the sample variance and, consequently, the computational effort of the estimation process based on MCS.

In order to fully verify the performance of the proposed methodology, several aspects such as system size, rarity of the failure events, number of different units and load shape are considered. The IEEE-RTS is used to test the proposed methodology. Also, several modifications are made in this system in order to compare the performance of the proposed approach against a non-chronological MCS and an extremely efficient analytical technique based on discrete convolution. A configuration of the Brazilian South-Southeastern generating system is also used to demonstrate the capability of the proposed method in real applications.

SUMÁRIO

Lista de Figuras.....	ix
Lista de Tabelas	x
Lista de Abreviaturas e Símbolos.....	xi
Capítulo 1: Introdução	1
1.1. Considerações Iniciais	1
1.2. Desenvolvimento Histórico	4
1.3. Estrutura da Dissertação	9
Capítulo 2: Avaliação da Confiabilidade da Capacidade de Geração.....	11
2.1. Introdução	11
2.2. Formulação do Problema	12
2.3. Métodos Analíticos	13
2.3.1. Considerações Iniciais	13
2.3.2. Modelo da Capacidade de Geração	14
2.3.3. Modelo da Carga do Sistema	16
2.3.4. Modelo da Reserva e Índices de Confiabilidade	16
2.4. Técnicas de Simulação Monte Carlo	17
2.4.1. SMC Não-Cronológica	18
2.4.2. SMC Sequencial ou Cronológica.....	20
2.4.3. SMC Pseudo-Cronológica	21
2.5. Métodos Baseados em Metaheurísticas	22
2.6. Comentários Finais	23
Capítulo 3: Aplicação do Método da Entropia Cruzada na Confiabilidade de Geração	25
3.1. Introdução	25
3.2. Técnicas de Redução de Variância	26

3.2.1.	Variáveis Antitéticas.....	26
3.2.2.	Variáveis de Controle	27
3.2.3.	Números Aleatórios em Comum	27
3.2.4.	Amostragem Estratificada.....	28
3.2.5.	Amostragem por Importância	28
3.3.	O Método da Entropia Cruzada	29
3.3.1.	Formulação Matemática Básica.....	30
3.3.2.	Abordagem de Níveis Múltiplos para Eventos Raros.....	35
3.3.3.	Algoritmo do MEC para Simulação de Eventos Raros	36
3.3.4.	Parâmetro de Suavização α	38
3.4.	Aplicação do MEC na Avaliação de Índices de Confiabilidade.....	38
3.4.1.	Reformulação do Problema	39
3.4.2.	Algoritmo da Abordagem Baseada no MEC	43
3.4.3.	Distorção do Espaço de Estados e Índices de Confiabilidade	46
3.4.4.	Modelos de Carga Variável	51
3.4.5.	O Parâmetro Φ	54
3.5.	Comentários Finais	57
Capítulo 4: Resultados.....		58
4.1.	Introdução	58
4.2.	IEEE Reliability Test System	59
4.2.1.	IEEE-RTS – Modelo de Carga Constante	59
4.2.2.	IEEE-RTS – Modelo de Carga de Níveis Múltiplos.....	61
4.3.	Modificações no IEEE-RTS Original	64
4.4.	Sistema Sul-Sudeste Brasileiro	67
4.5.	Comentários Adicionais.....	69
4.6.	Comentários Finais	73
Capítulo 5: Conclusões		75
Referências Bibliográficas.....		78

LISTA DE FIGURAS

Fig. 2.1: Modelo de Carga - Sequência de Níveis Discretos de Carga.....	16
Fig. 3.1: Sistema de Geração - Barra Única	40
Fig. 3.2: Distorção Ótima do Espaço de Estados em $L=L$	47
Fig. 3.3: Probabilidades dos Estados de Geração	48
Fig. 3.4: Estados de Falha com Contribuições para os Índices GCR	48
Fig. 3.5: Modelo de Carga de Níveis Discretos	51
Fig. 3.6: Mapeamento da Curva de Carga por uma Distribuição Uniforme Discreta	52
Fig. 3.7: Modelo Markoviano Sequencial de Carga	53
Fig. 3.8: Distribuição de Probabilidade de L	54
Fig. 3.9: Aplicação do Parâmetro Φ	56
Fig. 4.1: Probabilidades dos Estados de Geração	60
Fig. 4.2: Processo de Convergência do Índice LOLP	62
Fig. 4.3: Processo de Convergência do Índice EPNS	62
Fig. 4.4: Processo de Convergência do Índice LOLF.....	63
Fig. 4.5: Valores da LOLP para Diferentes Modificações no IEEE-RTS	66
Fig. 4.6: Comparação de Desempenhos – IEEE-RTS com Diferentes Fatores de Escala..	66
Fig. 4.7: Processo de Convergência do Índice LOLP – IEEE-RTS×5	67
Fig. 4.8: Raridade do Evento de Falha – IEEE-RTS×1	69
Fig. 4.9: Raridade do Evento de Falha – IEEE-RTS×20	70
Fig. 4.10: Eficiência do Parâmetro Φ	71
Fig. 4.11: Parâmetro Φ – Compromisso entre Qualidade e Esforço Computacional	73

LISTA DE TABELAS

Tabela 4.1 – IEEE-RTS: Índices de Confiabilidade – Modelo de Carga Constante	59
Tabela 4.2 – IEEE-RTS: Índices de Confiabilidade – Modelo de Carga Variável	61
Tabela 4.3 – IEEE-RTS × Fator de Escala – Modelo de Carga Variável.....	65
Tabela 4.4 – Resultados para a Configuração Normal do SSB	68
Tabela 4.5 – Resultados para a Configuração Reforçada do SSB	68
Tabela 4.6 – IEEE-RTS × 10 com Carga Pico de 31070MW	72

LISTA DE ABREVIATURAS E SÍMBOLOS*

$\hat{\mathbf{v}}_k$	Vetor de parâmetros de referência ótimo estimado na iteração k
$f(\bullet; \xi)$	Função massa/densidade de probabilidade definida por ξ
$f(x; \xi)$	Valor da função massa/densidade de probabilidade $f(\bullet; \xi)$ no ponto x
∇	Operador diferencial vetorial (Gradiente)
\in	Símbolo de <i>pertence</i>
*	Operador de convolução
$:=$	Operador de definição ou atribuição
$\{ f(\bullet; \xi) \}$	Família de funções massa/densidade à qual <i>pertence</i> $f(\bullet; \xi)$
a	Anos
ACO	<i>Ant Colony Optimization</i>
AIS	<i>Artificial Immune System</i>
AV	<i>Antithetic Variates</i>
CE	<i>Cross-Entropy</i>
CE-ISMC	<i>Cross-Entropy – Importance Sampling Monte Carlo</i>
COPFT	<i>Capacity Outage Probability and Frequency Table</i>
COPT	<i>Capacity Outage Probability Table</i>
CPU	<i>Central Processing Unit</i> – Tempo de Simulação
CV	<i>Control Variates</i>
$D_{KL}(\parallel)$	Divergência de Kullback-Leibler
$\tilde{E}[\]$	Valor esperado amostral
EENS	<i>Expected Energy Not Supplied</i>
EPNS	<i>Expected Power Not Supplied</i>
$E_{\xi}[\]$	Valor esperado considerando o vetor de parâmetros de referência ξ
F&D	Frequência e Duração
G	Modelo matemático que representa o conjunto de unidades geradoras

* Várias siglas são mantidas na língua inglesa por serem mais conhecidas nesta forma.

GA	<i>Genetic Algorithms</i>
GCR	<i>Generation Capacity Reliability</i>
GHz	Gigahertz
h	Horas
$H()$	Função teste
$I_{\{ \}}$	Função indicadora
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IS	<i>Importance Sampling</i>
K	Última iteração do processo de estimação de $\hat{\nu}$
L	Modelo matemático que representa a carga do sistema
L'	Nível no qual será baseada a distorção das probabilidades de G
L_{MAX}	Nível de L que representa o pico de carga do sistema
ln	Logaritmo natural
LOLD	<i>Loss of Load Duration</i>
LOLE	<i>Loss of Load Expectation</i>
LOLF	<i>Loss of Load Frequency</i>
LOLP	<i>Loss of Load Probability</i>
me	Mês
MEC	Método da Entropia Cruzada
MW	Megawatt
MWh	Megawatt-hora
N	Número total de amostras
\mathbf{n}	Vetor que contém o número de unidades em cada estação geradora
N_1	Contador de estados visitados com repetição na fase II do CE-ISMC
N_C	Número de estações geradoras do modelo G
NEF	<i>Natural Exponential Family</i>
N_{MAX}	Critério de convergência definido pelo máximo número de iterações
oc	Ocorrências
pdf	<i>Probability density function</i>
pmf	<i>Probability mass function</i>

PSO	<i>Particle Swarm Optimization</i>
$P_{\xi\{ \}}$	Probabilidade de um evento considerando o vetor de parâmetros ξ
R	Conjunto dos números reais
R	Modelo matemático da reserva estática do sistema
RTS	<i>Reliability Test System</i>
s	Segundos
$S()$	Função de desempenho
SMC	Simulação Monte Carlo
SSB	Sul-Sudeste Brasileiro
T	Intervalo de tempo que representa o período total de análise
\mathbf{u}	Vetor de parâmetros de referência original
$U()$	Distribuição uniforme de probabilidade
$V()$	Variância
VRT	<i>Variance Reduction Techniques</i>
$\mathbf{W}()$	Razão de verossimilhança
Z	Conjunto dos números inteiros
α	Parâmetro de suavização
β	Critério de convergência denominado coeficiente de variação
Δ	Incremento de arredondamento
ΔT	Subdivisões de igual comprimento do período de análise T
ξ	Vetor de parâmetros de referência genérico que define uma pmf/pdf
ρ	Parâmetro de níveis múltiplos
Φ	Parâmetro que determina os estados mais significativos de L