



## **Nova abordagem para simulação de resíduos utilizando MCMC aplicado na geração de cenários de Energia Natural Afluente**

**Hugo Ribeiro Baldioti**

Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)  
Rua Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22430-060  
baldioti@ele.puc-rio.br

**Reinaldo Castro Souza**

Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)  
Rua Marquês de São Vicente, 225 - Gávea, Rio de Janeiro - RJ, 22430-060  
reinaldo@ele.puc-rio.br

### **RESUMO**

O estudo da simulação de cenários sintéticos de Energia Natural Afluente (ENA) vem se desenvolvendo com novas propostas metodológicas. Tais desenvolvimentos muitas vezes pressupõem Gaussianidade nos resíduos, de forma que seja possível ajustar uma distribuição paramétrica nos mesmos. Percebeu-se que na maioria dos casos reais, no contexto do Setor Elétrico Brasileiro, os resíduos não podem ser tratados desta forma, uma vez que apresentam comportamentos extremos que são de interesse para o planejamento da operação do Sistema Interligado Nacional, de forma que a assimetria intrínseca a este comportamento seja passível de reprodução. Dessa forma, este trabalho propõe uma abordagem não paramétrica para simulação e amostragem dos resíduos das séries de ENA utilizando a técnica Markov Chain Monte Carlo (MCMC) e o Kernel Density Estimation, para estimação da envoltória da função de densidade. Os resultados apresentados mostram que a metodologia proposta é uma boa alternativa em relação ao modelo vigente.

**PALAVRAS CHAVE.** Markov Chain Monte Carlo, Kernel Density Estimation, Simulação de Resíduos.

**Tópicos:** EN – PO na Área de Energia; SIM – Simulação; MP – Modelos Probabilísticos.

### **ABSTRACT**

The study of the Natural Inflow Energy (NIE) synthetic scenarios simulation has been developing with new methodological proposals. Such developments often assume Gaussianity in the residues, so that it is possible to adjust a parametric distribution in them. It was noticed that in most real cases, in the context of the Brazilian Electric Sector, the noise can not be treated like that, since they present extreme behaviors that are interesting for the National Interconnected System operational planning, so that the asymmetry intrinsic to this behavior is amenable to reproduction. Thus, this work proposes a nonparametric approach to simulate and sample the NIE series noise using the Markov Chain Monte Carlo technique (MCMC) and the Kernel Density Estimation (KDE), for estimation of the density function envelopment. The presented results show that the proposed methodology is a good alternative in relation to the current model.

**KEYWORDS.** Markov Chain Monte Carlo. Kernel Density Estimation. Noise Simulation.

**Paper topics:** EN – OR in Energy; SIM – Simulation; MP – Probabilistic Models.



## 1. Introdução

O estudo de geração de cenários sintéticos está atrelado à modelagem de séries temporais. Por muitas vezes as metodologias clássicas não conseguem incorporar a não gaussianidade no que diz respeito à envoltória e à amostragem dos ruídos. A construção de cenários pressupõe que as séries possam ser geradas a partir de resíduos que sigam uma distribuição específica, porém, nos casos reais esta aproximação pode não apresentar uma solução satisfatória.

Tomando como caso de estudo a geração de cenários sintéticos de Energia Natural Afluente (ENA), dado no contexto de planejamento da operação do Setor Elétrico Brasileiro (SEB), observou-se que os resíduos gerados a partir do ajuste do modelo podem não seguir um padrão Gaussiano em sua distribuição, sendo assim, ajustar uma função que tem como base tal pretexto não é coerente. Dessa forma, com o intuito de preservar da melhor maneira possível a característica estocástica dos resíduos gerados a partir do ajuste da modelagem PAR(p), é proposta a geração de uma envoltória nos ruídos de forma que se possa gerar amostras aleatórias fiéis a distribuição dos mesmos. Tal alternativa tem como objetivo reduzir o erro agregado quando se ajusta uma função de densidade aos dados de interesse, assim sendo, a metodologia desenvolvida utiliza uma abordagem não paramétrica.

Diferentemente da estatística paramétrica, a abordagem não-paramétrica não utiliza famílias de funções de distribuições parametrizadas ou sequer faz suposições acerca das distribuições dos dados analisados. A diferença entre os modelos paramétricos e os modelos não-paramétrico é que o primeiro tem um número fixo de parâmetros, enquanto que o último o número de parâmetros cresce com a quantidade dos dados. De outra forma, uma vez que a estrutura do modelo não é especificada a priori, mas sim determinada a partir dos dados, isto não implica em dizer que tais modelos carecem completamente de parâmetros, mas que o número e a natureza dos parâmetros são flexíveis e não fixado antecipadamente.

Como métodos não paramétricos apresentam menos hipóteses sobre os dados sua aplicabilidade é muito mais ampla do que os métodos paramétricos correspondentes. Em particular, eles podem ser aplicados em situações que não se tem muita informação sobre os dados em questão. Além disso, devido à dependência de menos hipóteses, tais métodos são mais robustos. A aplicabilidade mais ampla e maior robustez dos testes não-paramétricos tem um custo: nos casos em que um teste paramétrico seria necessário, os testes não paramétricos têm menos poder. Em outras palavras, uma amostra maior pode ser necessária para tirar conclusões com o mesmo grau de confiança.

A partir do exposto, o presente trabalho tem como objetivo utilizar a técnica de simulação Markov Chain Monte Carlo (MCMC) para gerar amostras de resíduos respeitando a distribuição exata dos mesmos. Tal amostragem é realizada após a aproximação da envoltória dos dados, estimando a densidade a partir do método Kernel Density Estimation (KDE). Tal desenvolvimento se faz necessário em função da necessidade de se reproduzir as características extremas observadas nos resíduos históricos, de modo que a assimetria possa ser reproduzida satisfatoriamente.

O trabalho está dividido como se segue: a seção 2 apresentará brevemente os conceitos básicos envolvendo as metodologias propostas (KDE e MCMC); a seção 3 apresenta o estudo de caso e por fim, na seção 4, são expostas as conclusões e os trabalhos futuros.

## 2. Metodologia

### 2.1 Kernel Density Estimation

O processo de estimação da densidade de kernel é uma forma não paramétrica de estimação da densidade de probabilidade de uma variável aleatória.

Definindo-se formalmente o KDE. Seja  $(x_1, x_2, \dots, x_n)$  uma amostra i.i.d. gerada por alguma distribuição com uma densidade desconhecida  $f$ . O objetivo é estimar o formato da função  $f$ . Essa técnica é semelhante à construção de um histograma, porém com uma característica suavizada. Formalmente:



$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right)$$

Onde:

$K(\cdot)$  é o kernel (uma função simétrica mas não necessariamente positiva que integra um);  
 $h$  é o parâmetro de suavização chamado bandwidth.

Diversas funções podem ser utilizadas, alguns exemplos de funções kernel: uniforme; triangular; Epanechnikov, gaussiana (normal) etc.

Em outras palavras, a metodologia KDE traça uma função predeterminada (kernel) em cada observação e ao final do processo soma as componentes de forma que a área sobre a curva integre um. A figura 1 apresenta um exemplo dessa aplicação utilizando um kernel gaussiano, comparando tem-se a esquerda o histograma e o KDE a direita, construído utilizando o mesmo conjunto de dados.

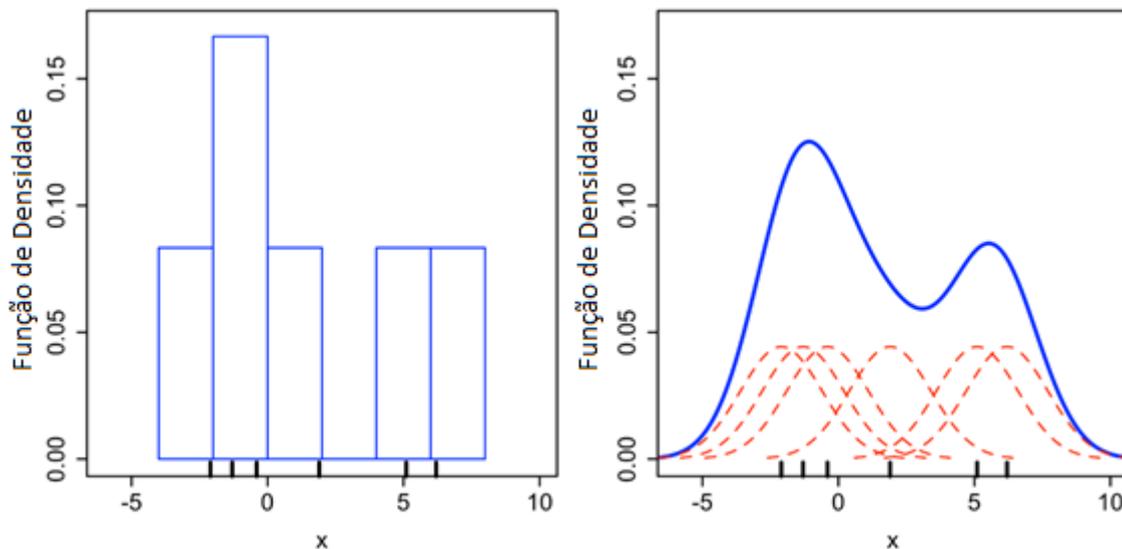


Figura 1 – Histograma e KDE para um mesmo conjunto de dados [Drleft, 2010].

A escolha correta do parâmetro  $h$  influencia expressivamente na “qualidade” do processo. Existem técnicas de otimização que retornam o melhor ajuste do parâmetro. O critério mais comumente utilizado para seleção deste parâmetro é conhecida como MISE (mean integrated squared error). Outro critério prático para estimação do bandwidth considera que, caso seja utilizada funções gaussianas no kernel e a função a ser estimada também é gaussiana, pode-se mostrar que:

$$h = \left(\frac{4\hat{\sigma}^5}{3n}\right) \approx 1.06\hat{\sigma}n^{-\frac{1}{5}}$$

Onde:

$\hat{\sigma}$  é o desvio padrão da amostras

A figura 2 apresenta, para uma mesma população, a aplicação do KDE utilizando três kernels diferentes (Normal, triângulo e Epanechnikov) com três diferentes valores de  $h$ . Observe-se que a escolha do kernel e da largura de banda influenciam a qualidade da envoltória desenvolvida.

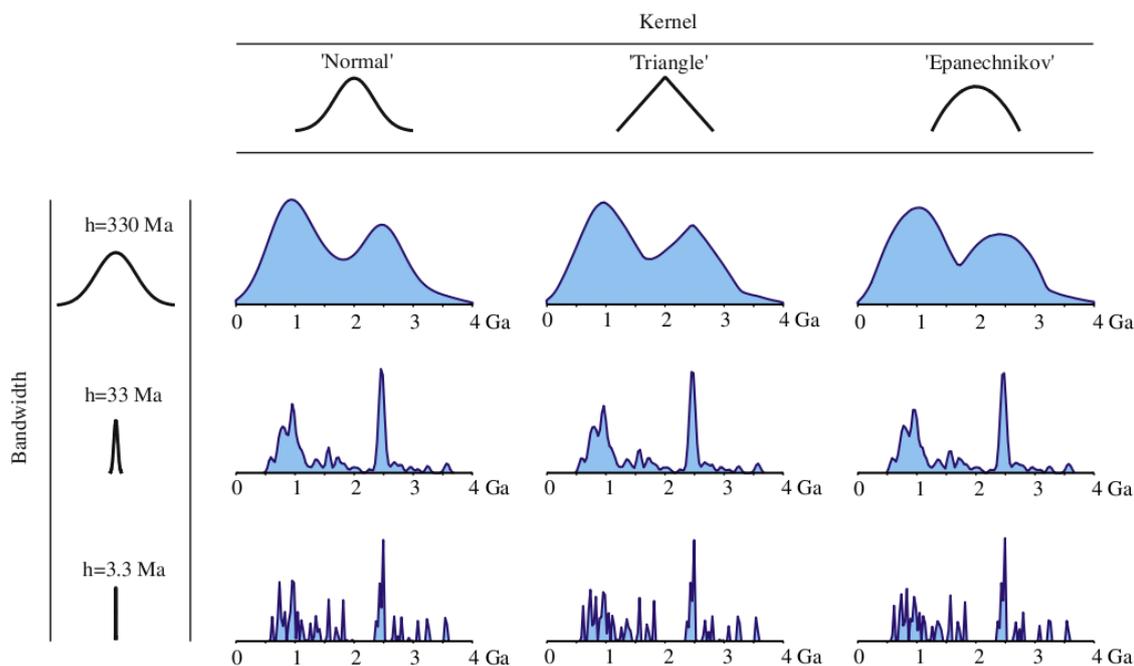


Figura 2 – Comparação entre três kernels e três valores de  $h$  para uma dada população [Vermeesch, 2012].

## 2.2 Markov Chain Monte Carlo

A ideia da utilização da técnica MCMC para amostragem dos resíduos é realizar uma simulação de Monte Carlo utilizando cadeias de Markov. Esse método permite simulações de uma distribuição ao embuti-la como distribuição limite da cadeia de Markov. Essa técnica se mostra bastante útil quando a densidade é muito complicada para se amostrar e/ou é definida com um número muito alto de dimensões.

O objetivo central é gerar amostras de uma distribuição ou aproximar  $E[f(X)]$ ,  $X \sim \pi$ . Este método produz uma cadeia de Markov onde a distribuição limite ( $\pi$ ) pode ser parcialmente conhecida e a sequência de amostras é obtida através de uma cadeia de Markov.

Para a aplicação da técnica a cadeia de Markov deve satisfazer as seguintes propriedades: irredutível; aperiódica e homogênea (recorrente). As duas formas mais utilizadas para construir uma cadeia de Markov com essas propriedades são: o algoritmo Metropolis-Hastings; e o amostrador de Gibbs. Como foi utilizado o primeiro algoritmo, será exposto abaixo a conceituação do mesmo.

### 2.2.1 Algoritmo Metropolis-Hastings

Inicialmente supõe-se que o estado atual da cadeia de Markov seja  $x_n$  e quer se gerar  $x_{n+1}$  da densidade  $P(x)$ . A geração desse novo ponto é um processo de dois estágios.

O primeiro é referente a geração de um ponto candidato,  $x^*$ . O valor deste ponto é gerado por uma distribuição proposta, denotada  $Q(x^* | x_n)$ , que depende do estado corrente da cadeia de Markov,  $x_n$ . Existem algumas condições a serem satisfeitas sobre a distribuição proposta a se utilizar. Uma forma típica para gerar esses pontos é utilizar uma distribuição normal centrada no estado corrente,  $x_n$ , e com um desvio padrão definido pelo usuário.

O segundo estágio é a realização de um teste de hipótese, ou seja, aceitar ou rejeitar o ponto escolhido. Primeiramente calcula-se a probabilidade de aceitação  $A(x_n \rightarrow x^*)$ , que é dada por:



$$A(x_n \rightarrow x^*) = \min \left( 1, \frac{P(x^*) Q(x_n | x^*)}{P(x_n) Q(x^* | x_n)} \right)$$

Nota-se no termo referente à distribuição proposta que, se a mesma for simétrica, a razão  $Q(x_n | x^*) / Q(x^* | x_n)$  será 1. Para o estudo de caso utilizou-se uma normal, portanto essa razão é considerada como sendo unitária. Este caso particular do algoritmo Metropolis-Hastings é chamado de Algoritmo Metropolis.

Dando continuidade ao desenvolvimento do algoritmo. Tendo proposto um ponto candidato e calculando a probabilidade de aceitação, decide-se aceitar ( $x_{n+1} = x^*$ ), ou rejeitar ( $x_{n+1} = x_n$ ) o ponto candidato. Para tomar essa decisão gera-se um valor aleatório (de uma distribuição uniforme) entre 0 e 1, denotado por  $u$ . Assim tem-se:

$$x_{n+1} = \begin{cases} x^* & \text{se } u \leq A(x_n \rightarrow x^*) \\ x_n & \text{se } u > A(x_n \rightarrow x^*) \end{cases}$$

Resumindo, o algoritmo pode ser escrito como:

- 1) Inicialize a cadeia;
- 2) Gere um ponto candidato da distribuição proposta,  $x^*$ ;
- 3) Gere um ponto a partir de uma uniforme (0, 1),  $u$ ;
- 4) Se  $u \leq A(x_n \rightarrow x^*)$  então  $x_{n+1} = x^*$ , caso contrário  $x_{n+1} = x_n$ .
- 5) Incremente  $t = t+1$  e volte ao passo 2.

Isto posto, a seguir é apresentado um estudo de caso envolvendo a geração de amostras de resíduos utilizando a técnica de simulação MCMC.

### 3. Estudo de Caso

Para o estudo de caso proposto neste trabalho são utilizadas séries de Energia Natural Afluente (ENA). Tais séries são utilizadas no contexto do planejamento da operação de médio/longo prazo do Setor Elétrico Brasileiro. A partir da simulação de cenários é possível calcular o preço da energia através da PDDE (Programação Dinâmica Dual Estocástica).

Neste contexto as séries de ENA são ajustadas através do modelo autorregressivo periódico de ordem  $p$  (PAR( $p$ )). Uma vez ajustado o modelo, tem-se uma matriz de resíduos que é utilizada para geração dos cenários. O SEB ajusta uma distribuição lognormal de três parâmetros aos resíduos de forma a garantir a não negatividade das séries geradas. Alguns estudos têm sido realizados no sentido de propor outras distribuições e métodos de amostragem para utilização no processo de simulação. O presente trabalho se posiciona nesse contexto, a medida que foram identificados diversos períodos que não podem ser caracterizados por distribuições paramétricas. Na figura 3 é apresentado um exemplo da utilização do KDE, com  $h=0.5$ , para geração da envoltória em comparação com a distribuição lognormal ajustada a um conjunto de dados.

O modelo PAR( $p$ ) pode ser descrito matematicamente, de forma simplificada como se segue:

$$\left( \frac{Y_t - \mu_m}{\sigma_m} \right) = \varphi_1^m \left( \frac{Y_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \varphi_{p_m}^m \left( \frac{Y_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right) + a_t^m$$

Onde:

$Y_t$  é a série sazonal de período  $s = 12$ ;

$t$  é o índice de tempo,  $t = 1, 2, \dots, sN$ , função do ano  $T$  ( $T = 1, 2, \dots, N$ ) e do período

$m$  ( $m = 1, 2, \dots, s$ );



$N$  é o número de anos;

$\mu_m$  é a média sazonal do período  $m$ ;

$\varphi_m^i$  é o  $i$ -ésimo coeficiente auto regressivo do período  $m$ ;

$p_m$  é a ordem do operador auto regressivo do período  $m$ ;

$a_t^m$  é a série de ruídos independentes com média zero e variância  $\sigma_a^2(m)$ .

O modelo apresentado é aplicado ao contexto do SEB nas séries de ENA para cada um dos 4 subsistemas (Sudeste/Centro-Oeste, Sul, Nordeste e Norte). As etapas de identificação das ordens “ $p$ ” e a estimação dos parâmetros do modelo são apresentadas detalhadamente em [Penna, 2009], [Oliveira, 2010], [Oliveira, 2013], [Ferreira, 2013].

Uma vez ajustado o modelo, para uma série de 120 períodos (10 anos de planejamento) e 84 anos, tem-se uma matriz de resíduos de igual ordem. Os elementos de tal matriz podem ser calculados como se segue:

$$a_t^m = \left( \frac{Y_t - \mu_m}{\sigma_m} \right) - \varphi_1^m \left( \frac{Y_{t-1} - \mu_{m-1}}{\sigma_{m-1}} \right) + \dots + \varphi_{p_m}^m \left( \frac{Y_{t-p_m} - \mu_{m-p_m}}{\sigma_{m-p_m}} \right)$$

Como o modelo PAR( $p$ ) trata cada período como um modelo AR( $p$ ), tem-se que a matriz de resíduos pode ser encarada periodicamente. É ajustada uma envoltória aos resíduos periódicos a partir da técnica KDE. Posteriormente, a partir de uma função que descreva exatamente a envoltória dos dados gerados, amostram-se os resíduos utilizando o MCMC, a partir do algoritmo Metropolis-Hastings.

Neste trabalho foram analisados dois conjuntos de amostras com tamanhos 4.000 e 100.000 para verificar qual conjunto representa de maneira mais fiel a envoltória dos dados históricos. O desvio padrão do espaço de busca do passo referente a Cadeia de Markov foi definido em 0,5 a partir do mesmo estudo que será apresentado a seguir. Tais valores foram considerados padrões para o desenvolvimento, porém são passíveis de alteração em trabalhos futuros.

Também é verificado no conjunto de dados originais do ajuste do modelo, a necessidade do desenvolvimento de tal técnica de amostragem alternativa. Verifica-se principalmente que a representação de dados extremos é mais fidedigna em relação ao histórico do que o simples ajuste de uma função de distribuição conhecida

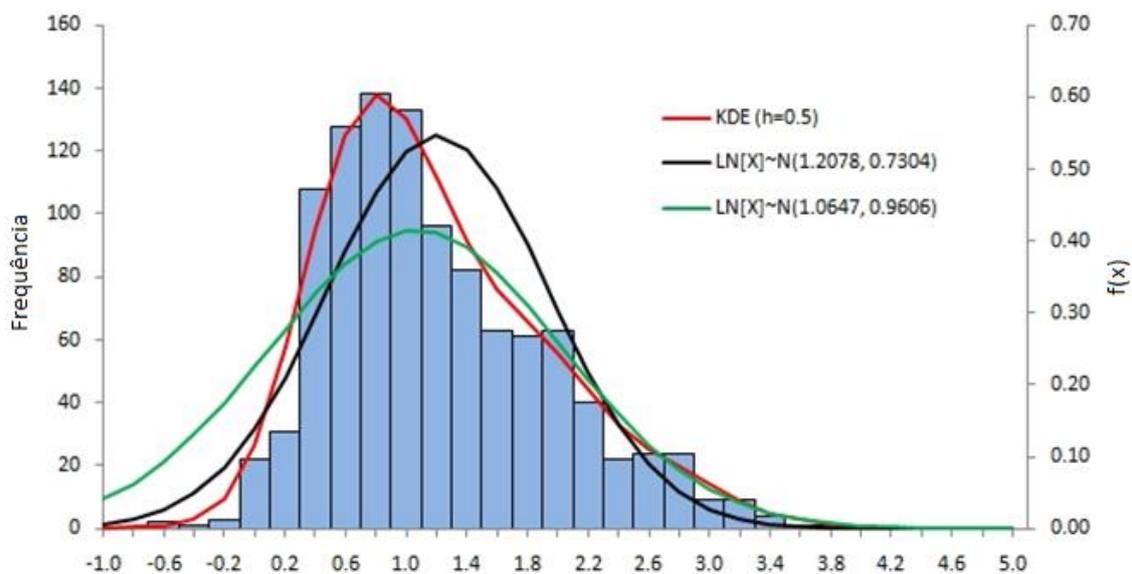


Figura 3 – Comparação entre o ajuste da envoltória utilizando KDE e Lognormal [Redfern, 2010]



### 3.1 Resultados

Primeiramente verificou-se a condição de não-gaussianidade no conjunto de resíduos para justificar o desenvolvimento do modelo. Para exemplificar tal questão, segue na figura 4 os histogramas dos resíduos para períodos aleatórios do conjunto de dados.

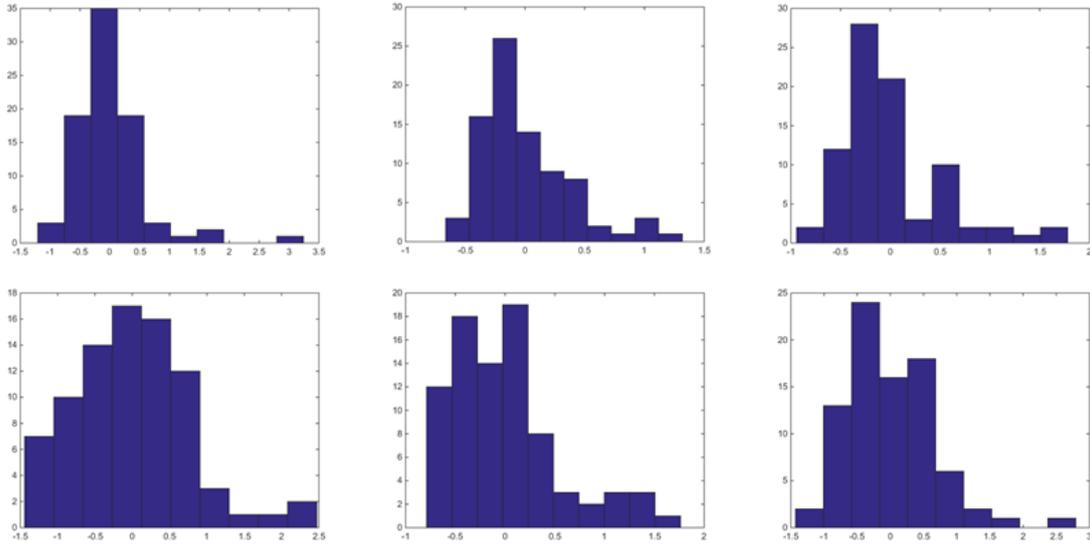


Figura 4 – Histograma de períodos aleatórios retirados da matriz de resíduos ajustados

Observa-se que os períodos acima apresentam comportamentos limite característicos que podem não ser capturados pelo ajuste de uma distribuição paramétrica. Tendo como proposta gerar amostras a partir da densidade dos ruídos mensais, foi utilizado primeiramente o KDE para estimar o kernel estocástico dessas séries, a partir de uma distribuição Normal com desvio padrão 1. Após, foi feita uma aproximação dessa densidade por uma função polinomial de ordem suficientemente grande para descrever bem a envoltória dos ruídos. De posse desse polinômio, foi possível avaliar os pontos candidatos para geração de amostras aleatórias que respeitem a distribuição exata dos ruídos.

A geração da envoltória a partir do KDE é exemplificada da figura 5, para um outro período aleatório. Observa-se que foi possível capturar o comportamento extremo da série, bem como picos de densidade.

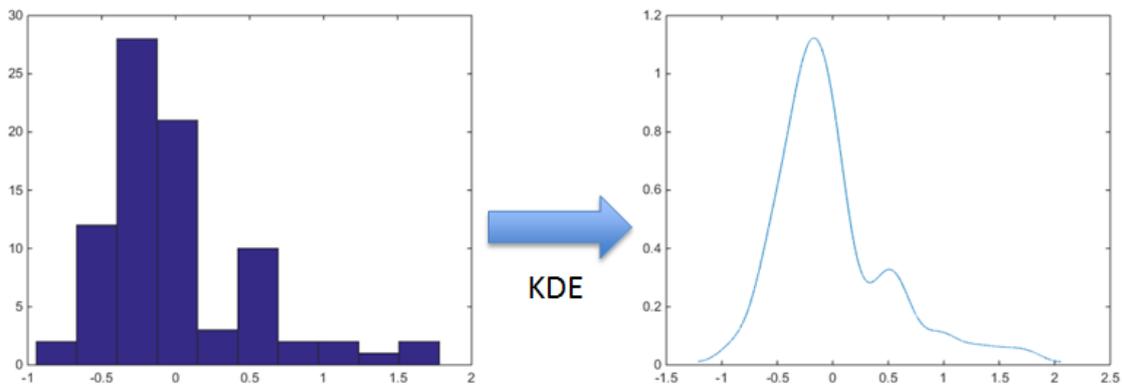


Figura 5 – Comparação do histograma gerado (esquerda) com a envoltória gerada pelo KDE (direita)



A partir do ajuste de uma função suficientemente grande, para capturar o comportamento exato da envoltória, o próximo passo da metodologia proposta é gerar amostras aleatórias desta distribuição utilizando o MCMC.

Inicialmente foram realizados testes utilizando um conjunto de 1000 amostras aleatórias para validação do código. As figuras 6 e 7 apresentam as envoltórias geradas para dois períodos distintos dos resíduos e a amostra aleatória obtida a partir da técnica MCMC (algoritmo Metropolis-Hastings).

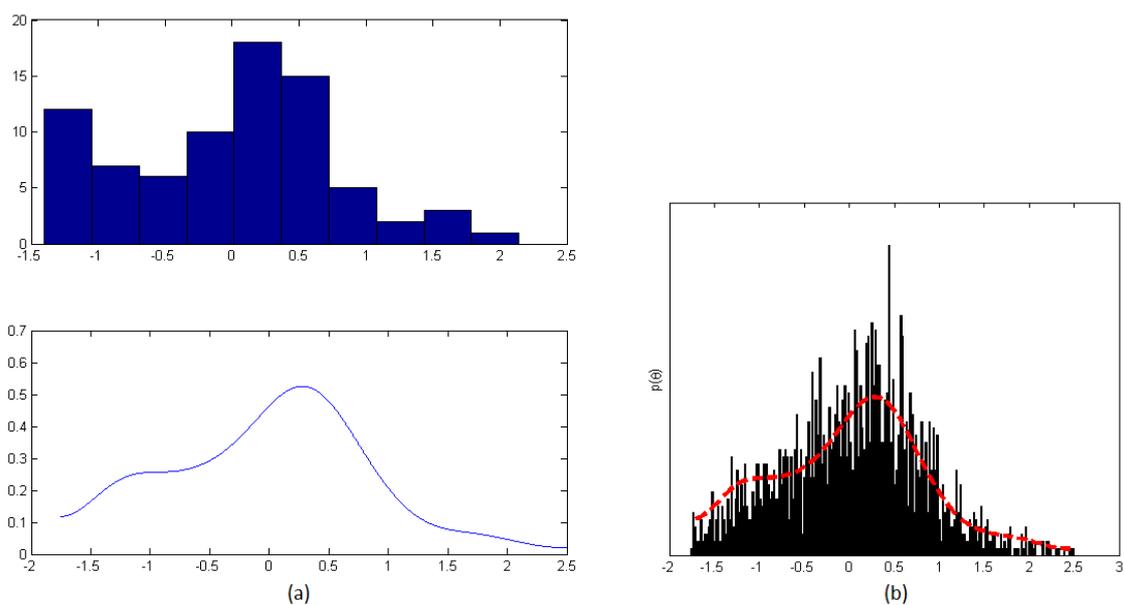


Figura 6 - (a) Histograma e envoltória calculada; (b) Amostra aleatória geradas MCMC

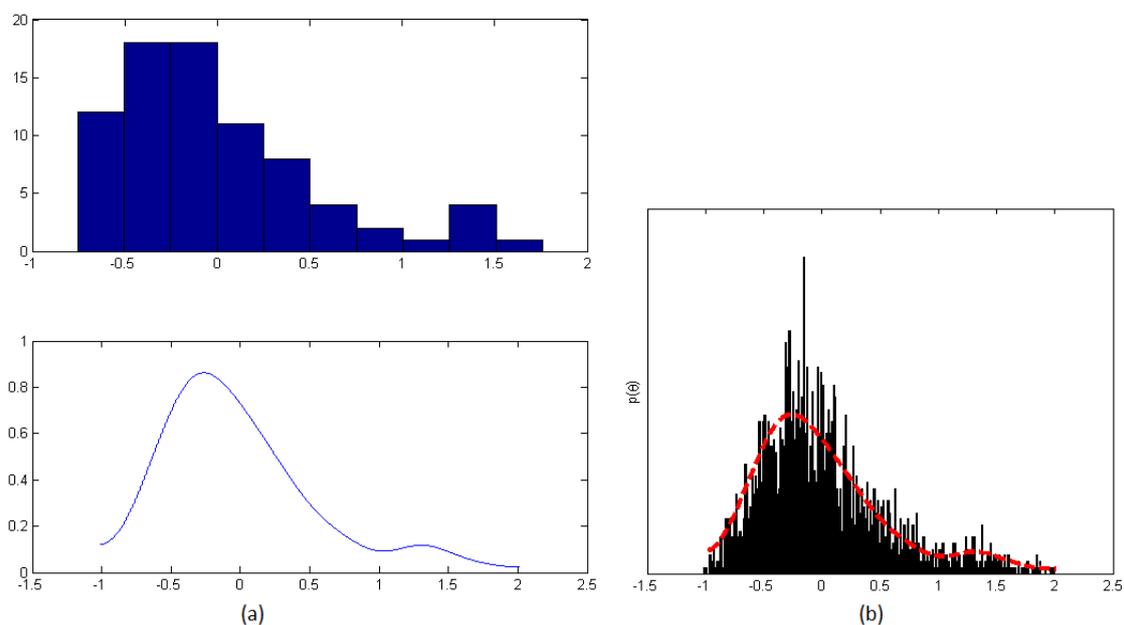


Figura 7 - (a) Histograma e envoltória calculada; (b) Amostra aleatória geradas MCMC

Percebe-se que, utilizando essa abordagem, é possível gerar amostras a partir de uma densidade contínua que respeita o comportamento estocástico do período em estudo.

Para aprimorar a qualidade das amostras geradas, no que diz respeito a reprodução do comportamento estocásticos foram realizados testes com 4.000 e 100.000 amostras bem como variando o tamanho do passo da cadeia de 1 para 0,5. Segue na figura 8 os resultados deste teste.

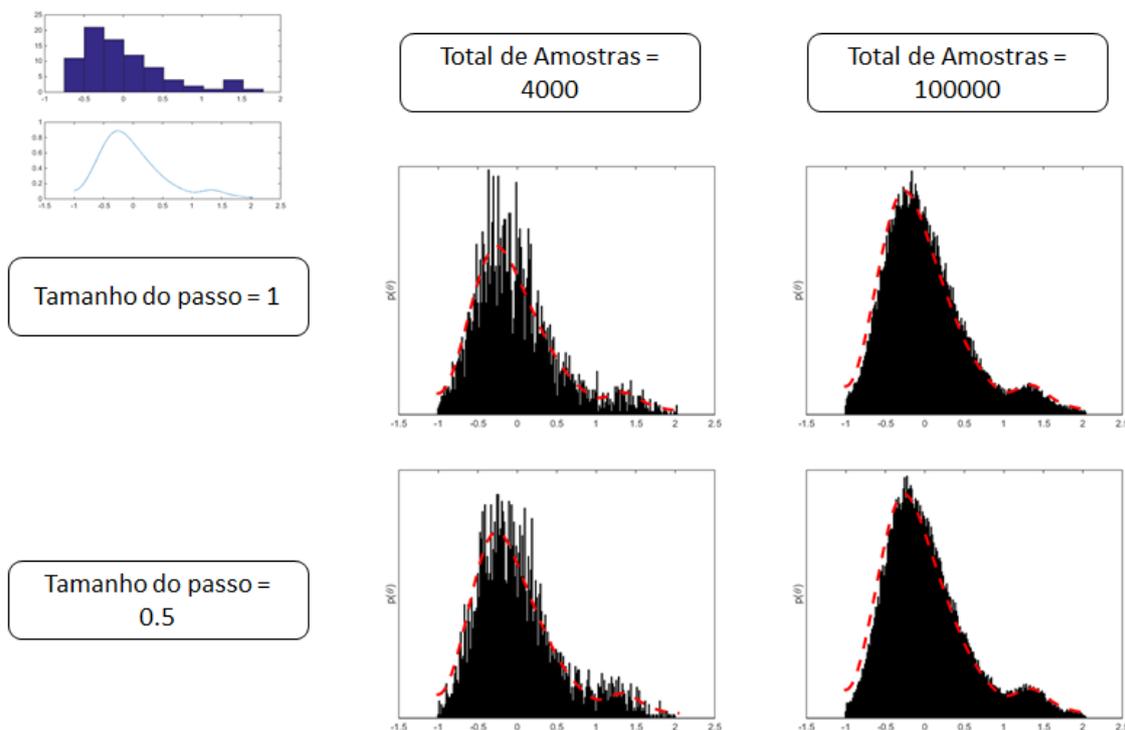


Figura 8 – Estudo sobre a variação do passo da cadeia de Markov em relação ao tamanho da amostra gerada para um período aleatório

Dando continuidade ao processo de validação das amostras geradas, verifica-se a envoltória produzida pelo conjunto amostral gerado em relação ao resíduo histórico com o intuito de avaliar a concordância entre real e simulado. Dessa forma tem-se na figura 9 a comparação da envoltória gerada para o histórico e para a amostra.

A partir da validação da modelagem, verifica-se a aplicabilidade dessa abordagem. Foi realizado um pequeno estudo utilizando o PMO de março de 2016, gerando 2000 cenários, a partir da amostra de resíduos MCMC de passo 0,5 e tamanho 100.000, de 12 meses para o subsistema sudeste. Ou seja, ajusta-se um PAR(p) para o conjunto de dados históricos, calcula-se a matriz de resíduos, ajusta-se uma envoltória (KDE), simula-se os resíduos (MCMC) e conseqüentemente gera-se os cenários sintéticos de ENA.

Os resultados gráficos para média e assimetria são exibidos na figura 10. A avaliação da assimetria foi realizada como proposto em [Baldioti, 2014].

No que diz respeito a aderência dos cenários ao histórico, o desempenho para média é bom. Para assimetria o comportamento é razoavelmente acompanhado por cada período do cenário, de modo que o resultado para análise da diferença percentual da assimetria totalizou 50%. Para questões comparativas, sob as mesmas condições, a tabela 1 apresenta a diferença percentual para outros modelos, para o mesmo período de estudo. Os resultados apontam que o PAR(p) – MCMC, proposto neste trabalho, como alternativa metodológica, apresenta o melhor resultado referente reprodução da assimetria.

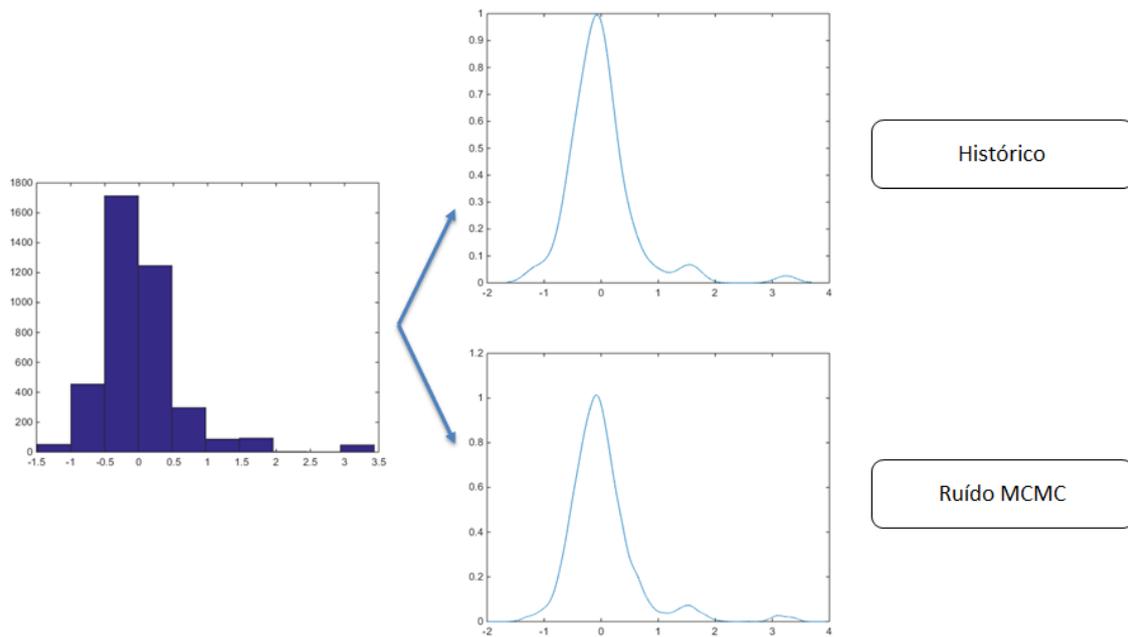


Figura 9 – Comparação das envoltórias Histórico X Amostra MCMC

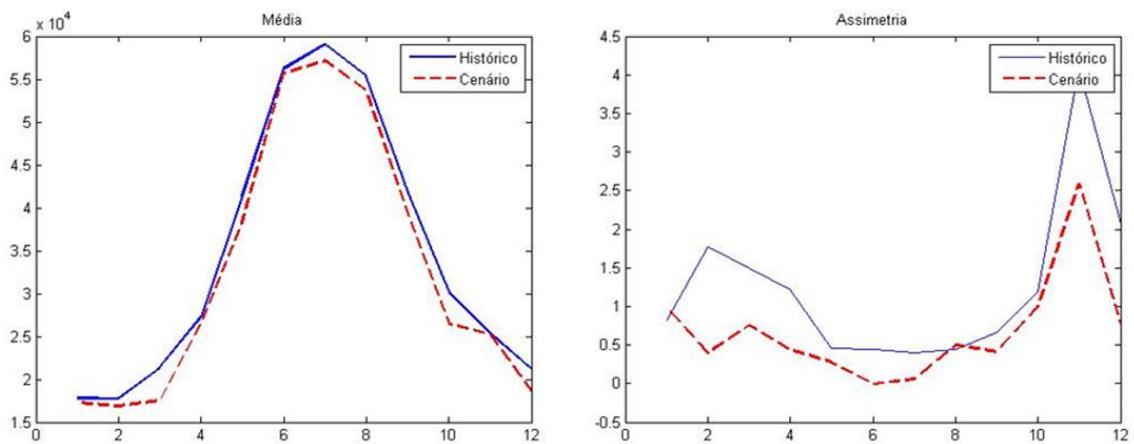


Figura 10 – Comparação entre médias e assimetrias do histórico em relação aos cenários MCMC.

Tabela 1 – Análise da Diferença Percentual da Assimetria para o Subsistema Sudeste

Modelo	DPA (%)
<b>PAR(p) - LogNormal</b>	29
<b>PAR(p) - Bootstrap</b>	14
<b>PAR(p) - MCMC</b>	50



#### 4. Conclusão e Trabalhos Futuros

Tendo como objetivo principal o desenvolvimento de uma metodologia alternativa para geração de amostras aleatórias para utilização na simulação de cenários sintéticos de ENA, pode-se dizer que o mesmo foi cumprido

Observa-se que a geração de amostras aleatórias através da metodologia proposta se comporta como esperado, gerando resultados que condizem com a envoltória gerada a partir da aproximação da função de probabilidade pela metodologia KDE.

A implementação da metodologia MCMC (Markov Chain Monte Carlo), como alternativa na geração de cenários sintéticos, com o intuito de incorporar a assimetria sem perder a qualidade de outras características estocásticas, como a média, se mostrou a partir de estudos comparativos, uma alternativa para ser utilizada no SEB.

A continuação na exploração desta metodologia se justifica pelos resultados encontrados e pela quantidade de parâmetros disponíveis para análise. Dessa forma, é importante ressaltar que tal metodologia ainda se encontra em desenvolvimento. Os próximos passos envolvem os seguintes desenvolvimentos: variação da quantidade de amostras de ruídos gerada; variação do desvio padrão (equivalente ao passo da Cadeia de Markov) utilizado para busca de pontos candidatos; adequação dos parâmetros da metodologia KDE para geração da envoltória; adequação da função utilizada para aproximar a envoltória encontrada a partir da metodologia KDE; aplicação de outros testes estatísticos nos cenários gerados, e por fim; implementação no contexto dinâmico do SEB.

É válido ressaltar que o desenvolvimento apresentado faz parte de uma das etapas do projeto estratégico de pesquisa e desenvolvimento intitulado “Modelos de Otimização do Despacho Hidrotérmico”. Os autores gostariam de agradecer ao suporte dado pelas empresas contratantes.

#### Referências

Baldioti, Hugo Ribeiro, (2014). “Abordagem multicritério para avaliação de modelos geradores de cenários aplicados ao planejamento da operação hidrotérmica a médio prazo” Dissertação de Mestrado, PUC-Rio.

Baldioti, H. R., Ribeiro, B. A. & Souza, R. C., (2014). “Multicriteria Approach for Evaluation of Scenarios Generation Models Applied to the Medium-Term Hydrothermal Operation Planning”. Washington, DC: Proceedings on The International Symposium on the Analytic Hierarchy Process.

Box, G. E P. & Tiao G. C., (1992) “Bayesian Inference in Statistical Analysis”, John Wiley and Low Inc.

Cao, R.; Cuevas, A.; Manteiga, W. G. (1994). “A comparative study of several smoothing methods in density estimation”. Computational Statistics and Data Analysis.

Corder, G. W.; Foreman, D. I. (2014). “Nonparametric Statistics: A Step-by-Step Approach”. Wiley.

Diaconis, Persi & Freedman, David, (1997) “On Markov Chains with Continuous State Space”, Technical Report, University of California.

Drleft, (2010). Kernel density estimation. [Online] Available at: [https://en.wikipedia.org/wiki/Kernel\\_density\\_estimation](https://en.wikipedia.org/wiki/Kernel_density_estimation)

Ferreira, P. G. C., (2013). A estocasticidade associada ao Setor Elétrico Brasileiro e uma nova abordagem para a geração de fluências via Modelos Periódicos Gama. Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro: Brasil.



- Finch, P. D., (1963) “A limit theorem for Markov Chains with Continuous State Space”, Journal of the Australian Mathematical Society.
- Gamerman, D., (1997) “Markov Chain Monte Carlo – Stochastic Simulation for Bayesian Inference”, ed. Chapman & Hall.
- Gibbs, W. R., Richardson, S. & Spiegelhalter, D. J., (1998) “Markov Chain Monte Carlo in Practice”, ed. Chapman & Hall.
- Gibbons, Jean Dickinson; Chakraborti, Subhabrata (2003). “Nonparametric Statistical Inference”, 4th Ed. CRC Press.
- Navarro, Dan. & Perfors, Amy, “The Metropolis-Hastings Algorithm”, University of Adelaide.
- Oliveira, Fernando Luiz Cyrino Oliveira, (2010) “Nova abordagem para geração de cenários de afluências no planejamento da operação energética de médio prazo”, tese de mestrado.
- Oliveira, F. L. C., (2013). Modelo de Séries Temporais para Construção de Árvores de Cenários Aplicadas à Otimização Estocástica. Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro: Brasil.
- Penna, Débora Dias Jardim, (2009) “Definição da Árvore de Cenários de Afluências para o Planejamento da Operação Energética de Médio Prazo”, tese de doutorado, PUC-Rio.
- Qi Li and Racine, Jeffrey S. (2007) “Nonparametric Econometrics: Theory and Practice”. Princeton University Press.
- Redfern, N., (2010). Some notes on cinematics IV. Research into film. An empirical approach to film studies: <https://nickredfern.wordpress.com/2010/11/25/some-notes-on-cinematics-iv/>.
- Santos, Reginaldo J., (2006) “Cadeias de Markov”, Departamento de Matemática – ICEx, UFMG.
- Scott, D.W. (1992) “Multivariate Density Estimation. Theory, Practice and Visualization”. New York: Wiley.
- Silverman, B.W. (1986) “Density Estimation”. London: Chapman and Hall.
- Silverman, B.W. (1998) “Density Estimation for Statistics and Data Analysis”. London: Chapman & Hall/CRC.
- Simonoff, J.S. (1996). “Smoothing Methods in Statistics”. Springer.
- Vermeesch, P., (2012). On the visualisation of detrital age distributions. Chemical Geology, Volume 312-313, pp. 190-194.
- Wand, M.P; Jones, M.C. (1995). “Kernel Smoothing”. London: Chapman & Hall/CRC.
- Zucchini, W. (2003) “Applied Smoothing Techniques”, Notes.