

GERAÇÃO DE SÉRIES SINTÉTICAS DE IRRADIAÇÃO DIÁRIA PARA AS CONDIÇÕES CLIMÁTICAS BRASILEIRAS

Cleber Onofre Inácio – cleberoi@petrobras.com.br
Hugo Tavares Vieira Gouveia – hugo.gouveia@petrobras.com.br
Luiz Fernando Almeida Fontenele – luizfontenele@petrobras.com.br
Paulo Henrique Fernandes Ferreira – phff@petrobras.com.br
Petróleo Brasileiro S.A., Centro de Pesquisas e Desenvolvimento (CENPES)

Resumo. *Etapa essencial para o projeto e simulação de desempenho de sistemas de geração solar é a obtenção de dados climáticos para as localidades de interesse. No caso de sistemas fotovoltaicos, por exemplo, são normalmente utilizados dados em escala temporal horária para o cálculo da energia gerada. Em função da escassez de dados, frequentemente é realizada a geração sintética das séries de irradiação solar na escala temporal requerida mediante uso de diversos modelos empíricos em cascata. A literatura geralmente apresenta modelos calibrados com dados de medição de Europa e Estados Unidos. Assim, o presente trabalho teve como objetivo apresentar os resultados de um exercício de adaptação de um desses modelos às condições climáticas brasileiras com foco na etapa inicial de geração em cascata de dados sintéticos de irradiação solar, quando séries de médias mensais são convertidas em séries na escala diária. Com o uso de dados de medição nacionais foram obtidas seis variações de calibração considerando um modelo nacional e cinco regionais. O desempenho dos modelos propostos foi avaliado a partir de comparação com outras opções da literatura através de métricas relacionadas à distribuição de frequência dos valores diários e características sequenciais das séries temporais. Os modelos calibrados com dados regionais apresentaram desempenho superior aos demais, o que encoraja sua aplicação direta e eventuais trabalhos de aprimoramento com uso de técnicas de agrupamento de dados.*

Palavras-chave: *Matrizes de Markov, Radiação Solar, Séries Sintéticas*

1. INTRODUÇÃO

Os resultados do leilão de Energia de Reserva de 2014 – LER 2014, o primeiro em nível federal em que houve a contratação de empreendimentos de geração fotovoltaica, e do primeiro leilão de Energia de Reserva de 2015 – 1º LER 2015 sugerem que a energia solar no Brasil pode ter iniciado uma trajetória semelhante à observada para a geração eólica. Com uma capacidade contratada da ordem de 2 GW_p para entrada em operação em 2017, o resultado desses leilões supera com folga os montantes previstos no Plano Decenal de Energia da Empresa de Pesquisa Energética (EPE) (2014), que indicou a entrada de 500 MW_p anuais de geração fotovoltaica centralizada no período entre 2017 e 2023. No caso da geração distribuída, o mesmo plano considerou uma evolução mais tímida, da ordem de 664 MW em capacidade acumulada até 2023.

Uma etapa básica para o projeto e simulação de desempenho de sistemas de geração solar é a obtenção de dados climáticos para o local de interesse, em particular para os níveis típicos de radiação solar. Se, por um lado, dados de medição representativos das características climáticas de longo prazo não estão disponíveis para todos os locais, por outro, os melhores levantamentos do recurso solar com abrangência completa do território brasileiro, como o apresentado por Pereira *et al.* (2006), apresentam valores de radiação solar apenas em escala temporal mensal. A simulação de sistemas fotovoltaicos, por exemplo, em razão de seu comportamento não-linear, requer informação em escala temporal mais detalhada, sendo a resolução horária a mais empregada quando o objetivo é calcular a energia gerada pela planta.

O expediente comumente utilizado para superar o problema da escassez de dados é o uso de métodos “em cascata” para a geração de séries sintéticas de irradiação solar na escala temporal requerida. Um roteiro bastante comum consiste no uso de uma sequência de médias mensais que caracterize o recurso solar no longo prazo e, então, a aplicação de: o método de Aguiar *et al.* (1988) para geração de série sintética em escala diária, o método de Aguiar *et al.* (1992) para conversão em uma série sintética com escala horária, o método de Erbs *et al.* (1982) para estimativa da fração difusa, e Perez *et al.* (1990) para a transposição para o plano inclinado dos coletores.

A maioria dos trabalhos presentes na literatura, incluindo os mencionados no parágrafo anterior, emprega uma abordagem empírica para calibração de seus modelos, a partir de dados de medição proveniente de locais geralmente localizados na Europa e Estados Unidos. Assim, antes de seu uso, faz-se mister a realização de testes para avaliar se tais métodos produzem dados com características semelhantes às que seriam verificadas através de medição no local de interesse.

Um exemplo de situação em que os resultados obtidos em simulação podem se distanciar significativamente da realidade é no cálculo do desempenho de uma usina fotovoltaica que empregue sistema de rastreamento solar. Os ganhos obtidos nesse caso provêm essencialmente do melhor aproveitamento da componente direta da radiação solar. O uso de

um método para sintetizar informação em escala diária pode gerar séries com distribuição diferente do comportamento real, impactando diretamente o desempenho dos métodos a jusante até uma situação em que ocorreria uma sub ou superestimação dos níveis da componente difusa. Nesse caso, as estimativas do ganho obtido com o uso de sistemas de rastreamento poderiam levar o empreendedor a tomar uma decisão equivocada para o tipo de montagem dos módulos fotovoltaicos durante a etapa de projeto.

O trabalho aqui apresentado teve como foco a etapa inicial de geração de dados sintéticos de irradiação solar, quando séries de médias mensais são convertidas em séries na escala diária. Um exercício de adaptação do método de Aguiar *et al.* (1988) às condições climáticas brasileiras foi realizado e o desempenho obtido foi comparado com a calibração original e uma variação apresentada por Meteonorm (2014).

2. SÉRIES DIÁRIAS DE IRRADIAÇÃO ATRAVÉS DE MCMC

O método proposto por Aguiar *et al.* (1988) figura como um dos mais populares para a geração sintética de sequências diárias de irradiação solar, tendo sido o escolhido para ser implementado em uma ferramenta comercial de simulação de desempenho de sistemas fotovoltaicos largamente utilizada pela indústria, o PVsyst. A técnica emprega a simulação de Monte Carlo via cadeias de Markov (MCMC) de 1ª ordem para gerar estocasticamente índices diários de claridade (K_T) a partir de médias mensais do mesmo índice (\bar{K}_T). O índice de claridade é definido como a razão entre a irradiação global (ou irradiância) da série medida ou sintética e a irradiação global extraterrestre (passível de cálculo preciso), que é aquela que seria obtida na superfície caso não houvesse atmosfera, sendo, portanto, uma medida da transmitância atmosférica.

O método MCMC permite que, de uma forma simples e rápida, sejam reproduzidas as distribuições de frequência e características sequenciais das séries obtidas através de medição. Para seu uso, Aguiar *et al.* (1988) discretizaram os índices diários de claridade em dez estados e produziram dez matrizes de transição de Markov (MTM) P_n , indicadas pela Eq. (1):

$$P_n = \begin{bmatrix} P_{1-1} & P_{1-2} & \cdots & P_{1-10} \\ P_{2-1} & P_{2-2} & \cdots & P_{2-10} \\ \vdots & \vdots & \ddots & \vdots \\ P_{10-1} & P_{10-2} & \cdots & P_{10-10} \end{bmatrix}, \quad (1)$$

em que P_{i-j} é a probabilidade de transição do estado i ao j . Cada matriz engloba um faixa de valores de \bar{K}_T , por exemplo, $0,35 < \bar{K}_T \leq 0,40$. As linhas dessas matrizes apresentam apenas elementos entre zero e um, e o somatório de cada linha é igual a um. Os parâmetros P_{i-j} das dez matrizes foram estimados a partir de um total de 300 meses de medição de irradiação global de nove estações na Europa, África e Ásia, com predominância de climas temperados. Os elementos das matrizes foram obtidos a partir do estimador de máxima verossimilhança que consiste basicamente no cálculo da frequência de cada transição verificada nos dados históricos.

Tomando como entrada as matrizes P_n e 12 valores de \bar{K}_T , é possível iniciar o procedimento de geração das séries diárias. Inicialmente, as matrizes P_{acum-n} são calculadas pela acumulação das colunas de P_n . A seguir, para cada mês, a matriz P_{acum-n} correspondente é selecionada, e os estados são sorteados sequencialmente utilizando as linhas da matriz para determinar o próximo estado mediante uso de um gerador de números aleatórios entre zero e um com distribuição uniforme. Na Fig. 1 essa etapa é ilustrada. Por fim, os estados são convertidos em valores de K_T por interpolação linear e convertidos em irradiação diária, sendo o trecho de série aceito caso atenda ao critério de tolerância para o valor mensal estabelecido pelo usuário.

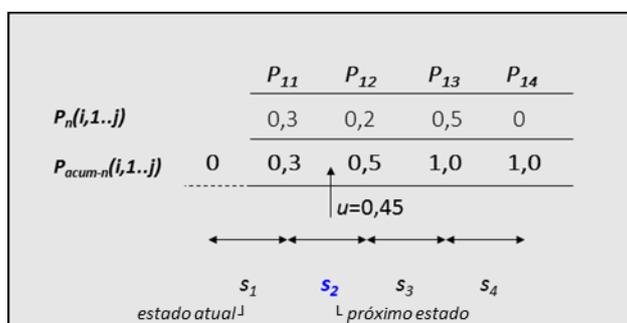


Figura 1 – Sorteio de estados com base em matrizes de transição de Markov.

A ferramenta comercial Meteonorm possibilita a obtenção de séries de radiação solar em nível mundial. Para tal, emprega uma variação do método de Aguiar *et al.* (1988) em que a variável modelada K_T foi substituída pelo índice de

céu claro, K_{CS} , de forma que fatores locais como altitude e turbidez atmosférica fossem considerados (Meteonorm, 2014). Dados de 150 estações meteorológicas de diversas regiões do globo foram utilizados na calibração de nove matrizes de tamanho 10×10 . Mapas globais de turbidez atmosférica típica e um modelo de céu claro são utilizados em conjunto com essas matrizes para a geração das séries sintéticas de irradiação diária com o emprego de algoritmo idêntico ao apresentado nos parágrafos anteriores.

3. CALIBRAÇÃO DAS MTM A PARTIR DE DADOS DE MEDIÇÃO NACIONAIS

Para a calibração das MTM adaptadas às condições climáticas brasileiras, foram utilizados os dados das estações meteorológicas de observação de superfície automáticas do Instituto Nacional de Meteorologia (INMET) (2014). As variáveis meteorológicas fornecidas pelo INMET são registradas em base horária, assim, realizou-se a integração diária da irradiação global para o cálculo do K_T e \bar{K}_T . Na Tab. 1 são apresentadas, para cada região do Brasil, as quantidades de estações meteorológicas, números de dias utilizados para cálculo do valor de K_T e também o número de meses para o cálculo de \bar{K}_T .

Tabela 1 – Informações referentes às estações de medição utilizadas para a elaboração das MTM.

Região	Nº de estações	Nº de dias	Nº de meses
Norte	68	39.207	1.296
Nordeste	128	81.247	2.684
Sudeste	112	72.265	2.394
Centro-Oeste	89	55.411	1.832
Sul	82	51.002	1.688
Brasil	479	299.132	9.894

Os valores apresentados na Tab. 1 são referentes aos anos de 2012 e 2013, após a realização de uma etapa de pré-tratamento dos dados do INMET. Nessa etapa, inicialmente foram excluídos os dias que continham lacunas (ou dados inconsistentes) em pelo menos uma hora do dia. Posteriormente, excluíram-se os meses que violaram os seguintes critérios de exclusão: mais do que três dias sem o valor de K_T ; e meses contendo dias com $K_T < 0,01$ ou $K_T > 0,90$.

Após a etapa de pré-tratamento, foram elaborados histogramas com as distribuições dos valores de K_T para o território brasileiro para diversos intervalos (ou classes) de \bar{K}_T . Com base na comparação desses histogramas com aqueles obtidos através da aplicação das MTM (e respectivas distribuições limite) propostas por Aguiar *et al.* (1988), observou-se que as distribuições dos valores sintéticos de K_T não são perfeitamente aderentes àquelas obtidas através da análise dos dados de medição. Como exemplo, são apresentados os histogramas para três diferentes classes de \bar{K}_T na Fig. 2, na qual se observam diversas diferenças entre as distribuições de frequência dos dados observados no Brasil (e suas diferentes regiões) e das séries sintéticas obtidas pelas MTM de Aguiar *et al.* (1988).

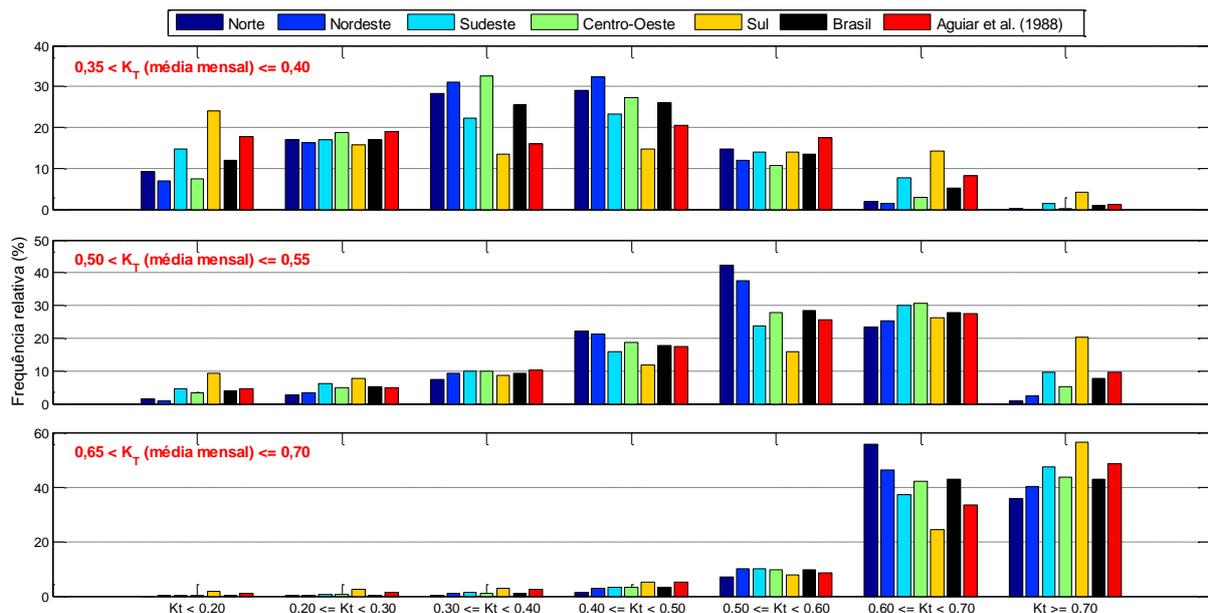


Figura 2 – Histogramas dos valores de K_T para 3 diferentes classes de \bar{K}_T .

3.1 Exemplo de uma MTM calibrada para o Brasil

Devido à restrição de espaço, será apresentada no presente artigo apenas uma das matrizes calibradas com o intuito de prover ao leitor um melhor entendimento de como é o formato da MTM. Na Tab. 2 é apresentada a MTM referente à classe $0,54 < \bar{K}_T \leq 0,57$, cujos estados são formados pelos valores de K_T pertencentes aos dez intervalos linearmente espaçados entre 0,0184 e 0,8422.

Tabela 2 – MTM calibrada para o Brasil, com $0,54 < \bar{K}_T \leq 0,57$.

0,050	0,091	0,099	0,099	0,156	0,149	0,149	0,112	0,087	0,008
0,028	0,059	0,106	0,084	0,156	0,154	0,154	0,124	0,109	0,026
0,018	0,050	0,081	0,110	0,146	0,172	0,171	0,146	0,088	0,018
0,011	0,037	0,058	0,095	0,149	0,177	0,189	0,182	0,092	0,010
0,007	0,027	0,040	0,087	0,138	0,203	0,228	0,188	0,077	0,005
0,005	0,018	0,026	0,054	0,103	0,201	0,280	0,230	0,079	0,004
0,004	0,013	0,020	0,034	0,073	0,156	0,325	0,285	0,086	0,004
0,004	0,011	0,015	0,025	0,053	0,109	0,235	0,390	0,154	0,004
0,007	0,011	0,014	0,021	0,039	0,062	0,120	0,297	0,411	0,018
0,007	0,007	0,021	0,019	0,019	0,040	0,067	0,126	0,399	0,295

4. RESULTADOS

Com o objetivo de avaliar o desempenho dos modelos calibrados para as condições brasileiras, foram elaboradas diversas simulações para oito diferentes locais do Brasil. As séries utilizadas nas simulações fazem parte do banco de dados da rede SONDA, desenvolvida e coordenada pelo Instituto Nacional de Pesquisas Espaciais (INPE) (2015). Na Tab. 3 são apresentadas informações referentes aos períodos das séries utilizadas.

Tabela 3 – Séries utilizadas nas simulações dos modelos.

Estação	Símbolo	Estado	Região	Período	
Brasília	BRB	DF	CO	nov/2008	out/2009
Caicó	CAI	RN	NE	jan/2004	dez/2004
Campo Grande	CGR	MS	CO	jan/2007	dez/2007
Florianópolis	FLN	SC	S	jan/2011	dez/2011
Ourinhos	ORN	SP	SE	dez/2009	nov/2010
Palmas	PMA	TO	N	jan/2010	dez/2010
Petrolina	PTR	PE	NE	fev/2011	jan/2012
São Martinho da Serra	SMS	RS	S	jan/2010	dez/2010

4.1 Critérios para comparação entre os desempenhos dos modelos

Para a comparação entre os desempenhos dos modelos utilizados foram utilizadas, além da autocorrelação de primeira ordem, as métricas *KSI* e *OVER*, ambas propostas por Espinar *et al.* (2009). O parâmetro *KSI* (*Kolmogorov-Smirnov test Integral*) é definido pela integração das diferenças entre duas funções de probabilidade acumulada. As diferenças são calculadas a partir da Eq. (2):

$$D_n = \max |R(K_{T_i}) - S(K_{T_i})|, \quad K_{T_i} \in [K_{T_{min}} + (n-1)p, K_{T_{min}} + np], \quad (2)$$

em que $R(K_{T_i})$ é a função de densidade de probabilidade acumulada para a série de referência; $S(K_{T_i})$ é a função de densidade de probabilidade acumulada para a série sintética; K_{T_i} é o i -ésimo valor de K_T pertencente ao intervalo definido por n (inteiro positivo menor ou igual a 100), $p = 0,01(K_{T_{max}} - K_{T_{min}})$, $K_{T_{max}} = 1$ e $K_{T_{min}} = 0$.

De posse dos valores de D_n , é possível calcular o parâmetro *KSI* a partir da Eq. (3):

$$KSI(\%) = 100 \left[\frac{\int_{K_T^{min}}^{K_T^{max}} D_n dK_T}{V_c(K_T^{max} - K_T^{min})} \right], \quad (3)$$

em que V_c é o valor crítico para o qual é formulada a hipótese de que se todos os valores de D_n forem menores do que o valor crítico, então as distribuições podem ser consideradas estatisticamente idênticas. O valor crítico depende do número da população da amostra, sendo calculado para um nível de confiança de 99 % a partir da Eq. (4), conforme Massey Jr. (1951). No presente trabalho, cada amostra contém 365 valores de K_T , obtendo-se um valor crítico igual a 0,0853.

$$V_c = 1,63/\sqrt{N}. \quad (4)$$

A normalização do KSI pela área crítica (denominador da Eq. (3)) permite a comparação dos valores percentuais para diferentes conjuntos de teste. Segundo Espinar *et al.* (2009), o valor mínimo para o KSI é zero, indicando que as funções de distribuição de probabilidade acumulada dos dois conjuntos sob análise são iguais.

O cálculo do parâmetro $OVER$ é semelhante àquele utilizado para o cálculo do KSI , porém o parâmetro $OVER$ considera em sua integração apenas as diferenças D_n cujos valores são superiores ao limite V_c , conforme Eq. (5).

$$OVER(\%) = 100 \left[\frac{\int_{K_T^{min}}^{K_T^{max}} z dK_T}{V_c(K_T^{max} - K_T^{min})} \right], \quad (5)$$

em que $z = D_n - V_c$ se $D_n > V_c$, ou $z = 0$ se $D_n \leq V_c$.

4.2 Nomenclatura dos modelos

Conforme descrito anteriormente, foram utilizados dados de diversas localidades do Brasil para a calibração das MTM adaptadas às condições climáticas brasileiras. No total, foram calibradas 60 MTM, sendo dez para cada um dos seis modelos propostos neste trabalho.

Um dos modelos é denominado “BR-Nacional”, e a calibração de suas dez matrizes foi composta por todos os dados descritos na última linha da Tab. 1. Trata-se, portanto, de um modelo generalista. Os outros cinco modelos são referentes às regiões do Brasil e para a calibração de suas matrizes também foram utilizados os respectivos dados descritos na Tab. 1. O objetivo da calibração regional é verificar se as informações regionais que estão implícitas nos dados medidos se refletem nas séries sintéticas de K_T , melhorando assim os resultados. Os nomes dos modelos regionais propostos são: “BR-CO”, “BR-N”, “BR-NE”, “BR-S” e “BR-SE”, para as regiões Centro-oeste, Norte, Nordeste, Sul e Sudeste, respectivamente.

Os modelos da literatura que servirão como base de comparação foram denominados de acordo com a referência bibliográfica, logo, foram utilizados os trabalhos de Aguiar *et al.* (1988) e Meteonorm (2014).

4.3 Desempenho dos modelos

Para a comparação entre os desempenhos, foram geradas, para cada um dos oito locais descritos na Tab. 3, dez séries sintéticas anuais de K_T com cada um dos modelos. Utilizou-se como critério de convergência um erro de 0,1 % do valor de \bar{K}_T , ou seja: para cada mês, a diferença máxima entre o valor de \bar{K}_T (valor medido) e a média dos valores sintéticos de K_T não pode ultrapassar $0,001\bar{K}_T$. Para cada uma das séries foram calculados os valores KSI e $OVER$, e posteriormente calcularam-se as médias desses valores. Esse processo foi denominado rodada típica.

Cabe ressaltar que as matrizes de transição do modelo de Meteonorm (2014) não são baseadas no K_T , mas sim no índice de céu claro K_{cs} , que varia entre zero e um. Para o cálculo de K_{cs} e aplicação do modelo de Meteonorm (2014) utilizou-se o modelo proposto por Ineichen *et al.* (2002), que utiliza como um dos dados de entrada o índice de turbidez de Linke. Nesse caso, os índices de turbidez utilizados foram apresentados por Remund *et al.* (2003). A relação entre K_T e K_{cs} é dada pela Eq. (6).

$$K_T = \frac{K_{cs}H_{cs}}{H_o}, \quad (6)$$

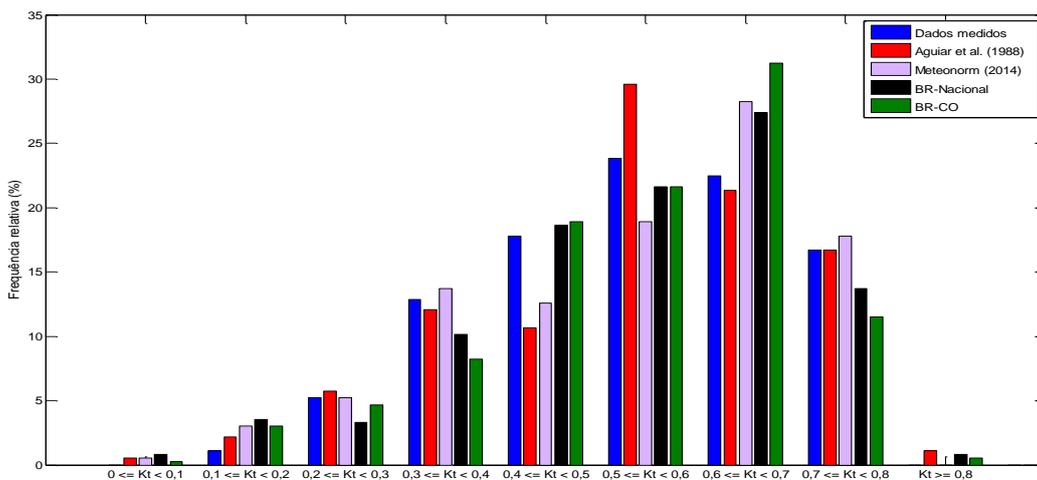
em que H_{cs} é a irradiação horizontal diária calculada segundo Ineichen *et al.* (2002), e H_o é a irradiação horizontal diária extraterrestre.

Na Tab. 4, são apresentados os valores do KSI e $OVER$ para uma rodada típica dos diversos modelos. De acordo com a análise do parâmetro $OVER$, observa-se claramente que o modelo de Aguiar *et al.* (1988) só conseguiu produzir séries sintéticas com parâmetro $OVER$ nulo para localidades da região Centro-Oeste, enquanto que para localidades da região Sul seu desempenho foi o pior entre todos os modelos. O modelo de Meteonorm (2014) apresentou melhores desempenhos do que o de Aguiar *et al.* (1988) para todas as localidades exceto na região Centro-Oeste. Em relação às médias do $OVER$, o modelo de Meteonorm (2014) apresentou uma redução de 66 % em relação ao modelo de Aguiar *et al.* (1988). O modelo “BR-Nacional” apresentou resultados melhores do que os dois modelos anteriormente citados para alguns locais e piores para outros. Para localidades da região Sul, por exemplo, o seu desempenho foi pior do que o desempenho de Meteonorm (2014). Em relação às médias do $OVER$, o modelo “BR-Nacional” apresentou uma redução de 59 % em relação ao modelo de Aguiar *et al.* (1988). Os modelos regionais foram aqueles que, de uma maneira geral, apresentaram os melhores desempenhos dentre os testados. Apenas três dos oito locais apresentaram parâmetro $OVER$ ligeiramente diferente de zero após a simulação de uma rodada típica com os modelos regionais. Cabe um destaque para o desempenho do modelo “BR-S”, que reduziu significativamente o parâmetro $OVER$ para a estação SMS. Em relação às médias do $OVER$, os modelos regionais apresentaram uma redução de aproximadamente 99 % em relação ao modelo de Aguiar *et al.* (1988).

Tabela 4 – Valores do KSI e $OVER$ para uma rodada típica dos diversos modelos.

Estação (Região)	Modelo							
	Aguiar <i>et al.</i> (1988)		Meteonorm (2014)		BR-Nacional		BR-Regional	
	KSI (%)	$OVER$ (%)	KSI (%)	$OVER$ (%)	KSI (%)	$OVER$ (%)	KSI (%)	$OVER$ (%)
BRB (CO)	18,79	0,00	25,40	0,09	28,80	0,00	10,18	0,00
CAI (NE)	36,77	5,94	22,25	0,06	34,17	2,76	13,90	0,66
CGR (CO)	20,33	0,00	23,86	2,13	19,30	0,00	20,52	0,03
FLN (S)	68,08	18,61	29,54	0,00	41,30	1,73	23,86	0,00
ORN (SE)	35,13	4,81	25,02	0,34	20,62	0,00	17,85	0,00
PMA (N)	34,17	5,85	14,16	0,00	31,82	1,07	15,57	0,00
PTR (NE)	33,40	2,80	37,44	2,39	28,80	0,76	9,06	0,00
SMS (S)	88,18	28,55	76,01	17,60	80,92	20,82	31,79	0,21
Média	41,85	8,32	31,71	2,83	35,72	3,39	17,84	0,11

Apenas a título ilustrativo, na Fig. 3 são apresentadas as distribuições dos valores de K_T para a estação de Brasília para apenas uma das dez simulações da rodada típica. Na cor azul estão representados os dados referentes às medições, as distribuições obtidas com os modelos de Aguiar *et al.* (1988), Meteonorm (2014), “BR-Nacional” e “BR-CO” estão representadas nas cores vermelha, roxa, preta e verde, respectivamente. Observa-se claramente na Fig. 3 que os modelos produzem distribuições diferentes, e dependendo da faixa de valores do K_T , alguns se aproximam da distribuição real e outros se distanciam. Embora as distribuições sejam diferentes, nesse caso, o valor do parâmetro $OVER$ foi nulo para os quatro modelos testados, assim, tais diferenças podem ser resultado apenas de variabilidade estatística.

Figura 3 – Distribuições dos valores de K_T para a estação de Brasília.

Na Fig. 4 são apresentadas as probabilidades acumuladas da série proveniente de medição em São Martinho da Serra no ano de 2010, as probabilidades acumuladas de séries sintéticas produzidas com os modelos propostos, além das diferenças calculadas para cada um dos modelos, conforme a Eq. (2). Os resultados apresentados na Fig. 4 também são ilustrativos e foram obtidos para uma das dez simulações da rodada típica. Observa-se, na Fig. 4, que o modelo “BR-S”

conseguiu reproduzir adequadamente a distribuição de probabilidades acumulada, sendo o único dentre os modelos testados que apresentou valor nulo para a métrica *OVER*.

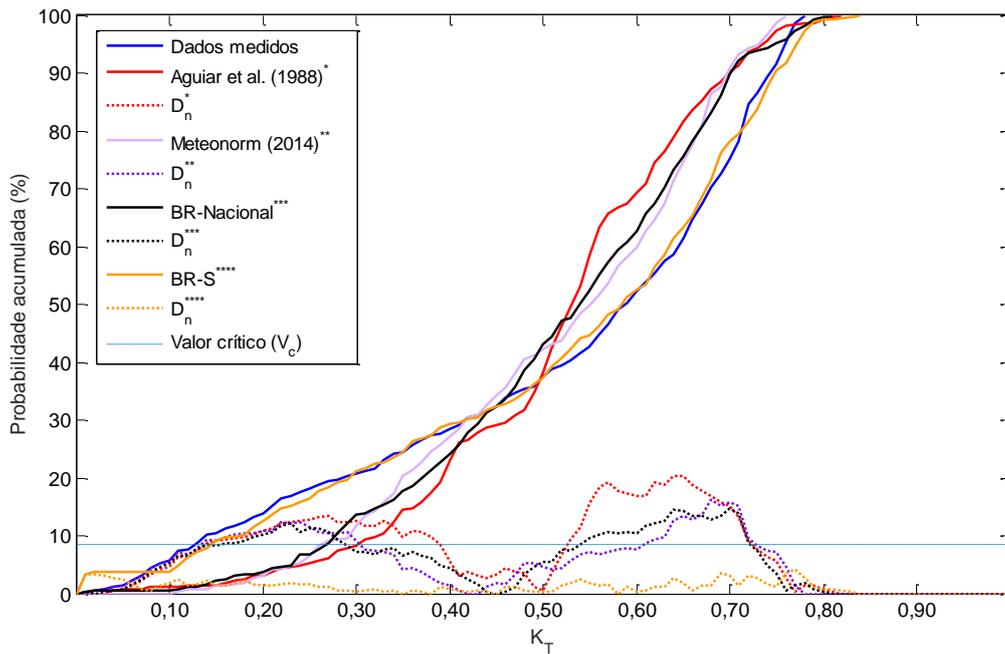


Figura 4 – Probabilidades acumuladas e diferenças para a estação de São Martinho da Serra.

Na Fig. 5 são apresentados os desvios obtidos para a métrica autocorrelação de primeira ordem (r_1) em relação aos dados de medição. É possível observar que o modelo de Aguiar *et al.* (1988) gerou séries com menor conexão com o valor de K_T do dia anterior, assim como os modelos Meteororm e BR-Nacional, porém com menores desvios. O valor médio de r_1 para os dados de medição foi de 0,44, e de 0,3 para Aguiar *et al.* (1988), 0,36 para Meteororm (2014), 0,39 para BR-Nacional e 0,4 para BR-Regional.

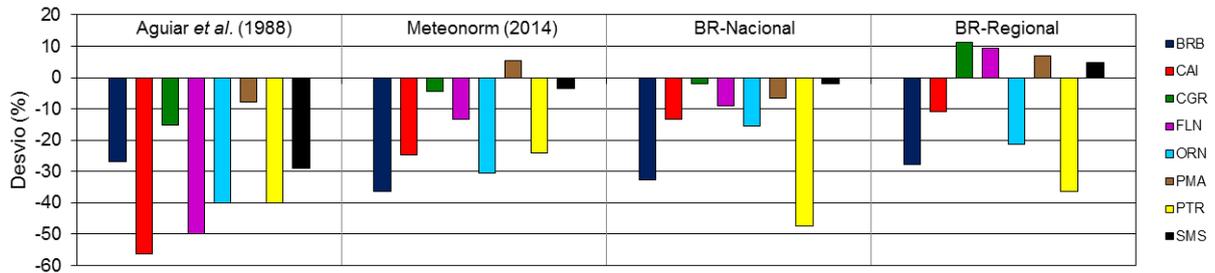


Figura 5 – Desvios da autocorrelação de primeira ordem para uma rodada típica.

5. CONCLUSÃO

O modelo de referência da literatura para a geração de séries sintéticas de irradiação solar diária foi proposto por Aguiar *et al.* (1988) e emprega a técnica de simulação de Monte Carlo via cadeias de Markov (MCMC) de 1ª ordem para gerar estocasticamente os índices diários de claridade a partir de médias mensais do mesmo índice. Para seu uso, Aguiar *et al.* (1988) discretizaram os índices diários de claridade em dez estados e produziram dez matrizes de transição de Markov (MTM). A principal contribuição do presente trabalho foi a adaptação do método proposto por Aguiar *et al.* (1988) às condições climáticas brasileiras. Para tal, utilizando-se dados das estações meteorológicas de observação de superfície automáticas do INMET, foram calibradas 60 MTM para o território nacional, sendo dez matrizes para cada um dos seguintes modelos: “BR-Nacional”, “BR-CO”, “BR-N”, “BR-NE”, “BR-S” e “BR-SE”.

O desempenho de cada um dos seis modelos propostos foi avaliado através das métricas *KSI*, *OVER* e r_1 , ambos propostos por Espinar *et al.* (2009). Os dados de irradiação utilizados para cálculo do desempenho fazem parte do banco de dados da rede SONDA, do INPE. Ao todo foram utilizados os dados de um ano de medição para cada uma das oito estações localizadas em diferentes regiões do Brasil. Com base nas médias obtidas para todas as rodadas típicas de simulações, os modelos regionais apresentaram reduções de aproximadamente 54 % para o *KSI* e de 99 % para o *OVER*, ambos em relação ao modelo de Aguiar *et al.* (1988).

O estudo aqui apresentado constitui um primeiro passo no trabalho de adaptação para condições brasileiras de toda a cadeia de modelos de radiação solar utilizada em estudos de sistemas de energia solar. Os resultados encorajam, também, o aprimoramento com calibração de modelos específicos para cada localidade não necessariamente obtidos através de agrupamento de dados com base em divisões políticas do território nacional.

Agradecimentos

Este trabalho foi realizado com recursos financeiros da Petrobras vinculados ao programa de P&D do Setor Elétrico regulado pela ANEEL, no âmbito do projeto PD-0553-0017/2011 para atendimento à chamada de projeto estratégico nº 013/2011 da ANEEL. Os autores agradecem ao INPE e INMET pela valiosa contribuição para a pesquisa nacional ao construírem a infraestrutura, mantê-la e disponibilizarem os dados de uma rede de estações meteorológicas de qualidade e com cobertura de todo o território brasileiro.

REFERÊNCIAS

- Aguiar, R., Collares-Pereira, M., 1988. Simple procedure for generating sequences of daily radiation values using a library of markov transition matrices, *Solar Energy*, vol. 40, n.3, pp. 269-279.
- Aguiar, R., Collares-Pereira, M., 1992. TAG: a time-dependent, autoregressive, gaussian model for generating synthetic hourly radiation, *Solar Energy*, vol. 49, n. 3, pp. 167-174.
- Empresa de Pesquisa Energética (EPE), 2014. Plano decenal de expansão de energia 2023. Rio de Janeiro: EPE.
- Erbs, D. G., Klein, S. A., Duffie, J. A., 1982. Estimation of diffuse radiation fraction for hourly, daily and monthly - sverage global radiation, *Solar Energy*, vol. 28, n. 4, pp. 293 - 302.
- Espinar, B., Ramírez, L., Drews, A. *et al.*, 2009. Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric station, *Solar Energy*, vol 83, n. 1, pp. 118-125.
- Ineichen, P., Perez, R., 2002. A new airmass independent formulation for the Linke turbidity coefficient, *Solar Energy*, vol, 73, n. 3, pp. 151-157.
- Instituto Nacional de Meteorologia (INMET), 2014. INMET. Disponível em <http://www.inmet.gov.br>. Acesso em dezembro de 2014.
- Instituto Nacional de Pesquisas Espaciais (INPE), 2015. Disponível em Sistema de Organização Nacional de Dados Ambientais - SONDA: <<http://sonda.ccst.inpe.br>>. Acesso em outubro de 2015.
- Massey Jr., F. J., 1951. The Kolmogorov-Smirnov test for goodness of fit, *Journal of the American Statistical Association*, vol. 46, n. 253, pp. 68-78.
- Meteonorm, 2014. Handbook part II: theory – version 7.1. Disponível em <http://meteonorm.com/images/uploads/downloads/mn71_theory.pdf>. Acesso em dezembro de 2014.
- Pereira, E. B., Martins, F. R., de Abreu, S. L., Rüther, R., 2006. Atlas brasileiro de energia solar. São José dos Campos: INPE.
- Perez, R., Ineichen, P., Seals, R., Michalsky, J., Stewart, R., 1990. Modeling daylight availability and irradiance components from direct and global irradiance, *Solar Energy*, vol. 44, n. 5, pp. 271-289.
- Remund, J., Wald, L., Lefevre, M., *et al.*, 2003. Worldwide Linke turbidity information. ISES Solar World Congress 2003, Jun 2003, Göteborg, Sweden. International Solar Energy Society (ISES), CD-ROM, 13 p, 2003.

GENERATING SYNTHETIC SERIES OF DAILY SOLAR RADIATION FOR BRAZILIAN CLIMATIC CONDITIONS

Abstract. *A basic step for design and performance simulation of solar generation systems is to obtain climate data for the site of interest. The simulation of photovoltaic systems, for instance, often makes use of data in the hourly temporal resolution when the goal is to compute the energy generated by the plant. The procedure commonly used to solve the problem of lack of data is the use of some “chain methods” in order to generate synthetic series of solar irradiation in the timescale required. Most of the studies presented in the literature use an empirical approach to calibration of models from local measurement data, generally from Europe and United States. Thus, the work presented here aimed to adapt one of these models to Brazilian climate with focus on the initial stage of generation of synthetic data of solar radiation, when average monthly series are converted into series in the daily scale. Local measurement allowed to propose six different models, one national and five regional. The performance of each of the six proposed models was evaluated against popular models available in literature using metrics to compare probability density functions and sequential characteristics of the daily series. Regional models showed superior performance and its direct application or even improvement using more advanced data clustering methods is encouraged.*

Key words: *Markov Matrices, Solar Radiation, Synthetic Series*