

# MODELAGEM DE INSUMO DE GERAÇÃO SOLAR USANDO DISTRIBUIÇÕES DE PROBABILIDADE

JEANE S.SOUZA<sup>1</sup>, MARIANA R.B.ALENCAR<sup>2</sup>, BENEMAR A.SOUZA<sup>3</sup>, WASHINGTON A. NEVES<sup>3</sup>

<sup>1</sup>Doutoranda no Programa de Pós-Graduação em Engenharia Elétrica, PPgEE/UFCG

<sup>2</sup>Mestranda no Programa de Pós-Graduação em Engenharia Elétrica, PPgEE/UFCG

<sup>3</sup>Departamento de Engenharia Elétrica, UFCG

Campina Grande, Paraíba, Brasil

E-mails: jeane.souza@ee.ufcg.edu.br, mariana.alencar@ee.ufcg.edu.br, benemar@dee.ufcg.edu.br,

waneves@dee.ufcg.edu.br

**Abstract**— It is proposed in this paper a probabilistic modeling study of solar generation input based on probability distributions adjusted to ten-year historical data. Normal, gamma, lognormal, *t* location scale, extreme values (EV), generalized extreme values (GEV) and weibull distribution were tested. Among them the extreme generalized value distribution function proved to be very flexible and it can model ten years of decennial average global solar irradiance data in Campina Grande-PB. The main objective is to show that this modeling can be developed in any study that requires irradiance data of the studied region, both for periods (dry and rainy) and for hourly data. The Kolmogorov-Smirnov adhesion test was applied to 360 series of decennial global average solar radiation data, to verify the fit of the data to the tested distribution functions.

**Keywords**— Modeling, Probability Distributions, Solar irradiance, Kolmogorov-Smirnov, Decennial Historical Series.

**Resumo** — Neste artigo propõe-se um estudo de modelagem probabilística de insumo de geração solar com base em distribuições de probabilidade ajustadas às séries históricas decenciais. Foram utilizadas funções de distribuição de probabilidade normal, gamma, lognormal, *t* location scale, valores extremos (EV), valores extremos generalizados (GEV) e weibull. Dentre elas, a distribuição generalizada de valores extremos mostrou-se bastante flexível e foi capaz de modelar dez anos de dados de irradiância global média decennial em Campina Grande-PB. O principal objetivo é mostrar que essa modelagem pode ser desenvolvida para direcionar qualquer estudo que necessite de dados de irradiância da região estudada, tanto por períodos (seco e chuvoso) como por dados de hora em hora. Aplicou-se o teste de aderência de Kolmogorov-Smirnov nas 360 séries de dados de radiação solar global média decennial, para verificar o ajuste dos dados às distribuições testadas.

**Palavras-chave**— Modelagem, Distribuições de Probabilidade, Irradiância Solar, Kolmogorov-Smirnov, Séries Históricas Decenciais.

## 1 Introdução

A radiação solar é definida como energia emitida pelo sol, e hoje é uma das alternativas energéticas mais promissoras para fornecer energia ao desenvolvimento humano (Pinho, 2014). No entanto, o fornecimento deste tipo de energia está estritamente vinculado a fatores meteorológicos como radiação solar e temperatura. Em razão da natureza intermitente destas fontes, o primeiro passo no processo de implantação de unidades de geração distribuída (GD) é a avaliação das características de radiação solar no local de instalação.

A observação de variáveis ao longo do tempo como forma de compreensão dos fenômenos meteorológicos determina padrões de ocorrência e propicia uma adequada previsibilidade do comportamento climático de uma região. Tendo em vista estas características, a literatura explicita trabalhos (Silva, Rodrigues e Silva, 2016), (Atwa, 2010) que apontam a utilização de técnicas probabilísticas como alternativa para representação de incertezas na modelagem de fontes renováveis de energia.

A escolha do modelo mais adequado para recursos renováveis não determinísticos é condicionada a diversas aplicações e.g. localização ótima de fontes

renováveis, avaliação da confiabilidade de módulos fotovoltaicos mediante incertezas, avaliação da distorção harmônica, velocidade dos ventos, atividades destinadas a agropecuárias entre outras.

O estado da arte evidencia que a modelagem de natureza estocástica de fontes renováveis de energia é um tema atual e relevante. Ribeiro (2007) apresenta uma comparação de distribuições de probabilidade ajustadas às séries históricas decenciais e mensais, e estima as precipitações prováveis para o município de Barbacena-MG. Araújo (2010) avalia a distribuição de probabilidade que melhor se adequa a séries diárias de temperatura máxima da cidade de Iguatu-CE. Em Cordeiro (2016) propõe-se uma nova metodologia com um ano de dados de irradiância com base na distribuição generalizada de valores extremos. Poucos trabalhos relacionados a essa área explicitam avaliações realizadas com séries de dados relativas a dias ou a máxima de poucos anos; o que retrata um intervalo de tempo questionável se comparado ao tempo de operacionalidade destinado a estas aplicações (Nunes, 2006), (Oliveira, 2016), (Cordeiro, 2016).

Este trabalho apresenta um estudo de avaliação probabilística de recursos renováveis de energia para verificação do ajuste das séries de dados de radiação solar média decennial de 10 anos de dados de uma estação meteorológica do município de Campina

Grande-PB. Nesse contexto, são investigadas as distribuições de probabilidade que melhor descrevem as séries de dados históricos de radiação e temperatura utilizadas.

O conteúdo apresentado neste trabalho está organizado da seguinte forma: na seção 2 a modelagem de recursos renováveis é apresentada, na seção 3 é descrito o teste de aderência Kolmogorov-Smirnov, na seção 4 são reportados os resultados obtidos e na seção 5 são feitas as conclusões.

## 2 Modelagem de Recursos Renováveis

Dez anos de estudo selecionados são divididos em períodos chuvoso (entre janeiro e julho) e seco (entre agosto e dezembro) a fim de representar o comportamento aleatório dos diferentes recursos renováveis durante cada período. Esta seleção é realizada em compromisso com as estações explicitadas no município de Campina Grande-PB. Para o estudo, os dados horários de irradiância são testados por diferentes funções de densidade de probabilidade (FDP).

### 2.1 Processamento histórico de dados

A maior parte do território brasileiro está localizada relativamente próxima à Linha do Equador. A Paraíba possui um clima quente, com temperaturas elevadas durante todo o ano e variado conforme o relevo local. O clima é considerado como tropical, com temperaturas médias em torno de 24 °C (Geografia da Paraíba, 2018).

Para o desenvolvimento do modelo, foram utilizados dados diários de radiação solar em kJ/m<sup>2</sup> da cidade de Campina Grande-PB, obtidos pelo Instituto Nacional de Meteorologia (INMET) dos anos de 2008 a 2017, compreendendo assim, uma série de 10 anos, cuja localização é 07°13'S latitude Sul e 35°53'W longitude Oeste.

A irradiância é um processo aleatório indexado pelo tempo. Sendo assim, a rigor existe uma distribuição de probabilidade para cada instante. Na prática, isto é realizado para intervalos de uma hora, sendo que no período noturno a radiação solar é zero (Atwa, 2010), (Cordeiro, 2016). Ou seja, para descrever o fenômeno estocástico da radiação solar, faz-se necessário selecionar, o período histórico de dez anos, o qual é dividido em duas temporadas que representam a situação meteorológica da Paraíba, sendo cada estação representada pelos dados colhidos nos seus respectivos dias. Além disso, cada dia deve ser dividido em 24 segmentos (cada um representando uma hora).

### 2.2 Modelagem da Intensidade da Irradiância

As vantagens de usar o efeito fotovoltaico para gerar eletricidade são em grande parte devido à ausência de poluentes durante a operação, silenciosa,

requisitos de manutenção baixas e longa vida útil. Além disso, a energia solar é abundante, limpa e inesgotável (Villalva, 2013).

A intensidade de irradiância foi modelada baseando-se nas referências Atwa (2010) e Cargnelutti (2004). A técnica usada para modelagem de dados neste artigo pode ser implementada para diferentes fontes primárias. Os dados de irradiância são divididos em dois períodos, chuvoso (de janeiro a julho) e seco (agosto a dezembro) visando representar o comportamento aleatório do insumo. Desta forma, para cada período de cada ano é analisado a média do dia e gerada uma FDP, assim como também a melhor distribuição referente a cada hora do dia em diferentes períodos.

De acordo com Atwa (2010), não há como utilizar uma função densidade de probabilidade para descrever o comportamento aleatório da irradiância solar anual. Para usar FDP para modelar uma variável aleatória, a variável deve ser totalmente aleatória. No entanto, este não é o caso quando se lida com a irradiância solar devido à noite o valor de irradiância ser zero (devido se tornar um valor constante perde-se a aleatoriedade). Com base nisso os seguintes procedimentos foram realizados com os dados do INMET para modelar a irradiância:

**Passo 1:** Agrupar dados em dois períodos: chuvoso (entre janeiro e julho) e seco (entre agosto e dezembro);

**Passo 2:** Reagrupar os dados de irradiância em suas respectivas horas de ocorrência, para cada período (cada hora terá uma amostra);

**Passo 3:** Obter uma distribuição de probabilidade que represente de forma satisfatória os dados coletados.

As funções densidade de probabilidade consideradas candidatas mais prováveis para representar a irradiância foram: normal, *t location scale* lognormal, weibull, gamma, valores extremos e valores extremos generalizados. Os parâmetros de cada distribuição testada foram estimados pelo método de máxima verossimilhança e em seguida cada distribuição foi submetida ao teste de aderência de Kolmogorov-Smirnov.

## 3 Teste de Aderência

A simples visualização dos dados amostrais de uma variável em um histograma de frequência é insuficiente para concluir, entre as diversas funções de distribuição de probabilidade conhecidas, a que melhor se ajusta aos dados em estudo (Cargnelutti, 2004). Deste modo, faz-se necessário o uso de testes de aderência para verificar se a distribuição de probabilidade dos dados de uma variável em análise pode ser estudada por uma função de distribuição de probabilidade conhecida.

O teste de Kolmogorov-Smirnov (KS) é um teste não paramétrico, e tem como embasamento a máxima distância entre as probabilidades empíricas e as teóricas.

cas obtidas sob a função de distribuição de probabilidade em teste (Campos, 1983; Assis et al., 1996; Morettin e Bussab, 2004). Assim, menores valores da estatística fornecem maiores valores de valor-p e, consequentemente, maior evidência de não rejeição da hipótese nula ( $H_0$ ), ou seja, maior aderência dos dados à distribuição em teste. Portanto, quando o valor-p for maior que 0.05% ele aceita  $H_0$  e quando o valor-p for menor ele rejeita ( $H_1$ ). Este teste não é aplicável a variáveis aleatórias discretas (Martins et al., 2011).

A estatística de Kolmogorov–Smirnov para uma dada função distribuição acumulada  $F_o(x)$  é:

$$D = \max |F_o(x) - S_n(x)| \quad (1)$$

$F_o(x)$ : uma distribuição teórica acumulada;

$S_n(x)$ : uma distribuição observada em uma amostra de  $n$  observações (distribuição empírica).

#### 4 Avaliação da Modelagem Proposta

Uma síntese dos resultados obtidos e das análises realizadas é relatada a seguir. São apresentadas simulações de diferentes períodos:

- Com os dados diários de irradiância, estimou-se a média decenal (média de 10 dias) dos 36 decêndios de cada ano (1º decêndio de janeiro até 3º decêndio de dezembro) em seguida agrupando os dados em duas estações, uma no período seco e a outra no período chuvoso formando 360 séries temporais;
- Modelar a irradiância solar gerando uma FDP para cada hora, que representa uma dada estação;
- Selecionar a FDP associada à hora de maior intensidade de irradiância solar.

Antes de iniciar a busca por uma distribuição de probabilidade que represente de forma satisfatória os dados, é importante certificar se há uma grande semelhança entre os dados de anos adjacentes.

Para tanto, foram selecionados dois anos de dados para apresentar essa aleatoriedade. No entanto, foram usados a irradiância média do dia no período chuvoso de 2008 e 2014 e a irradiância às 12 horas, para representar essa comparação de resultados obtidos para cada ano. O resultado está exposto nas Figuras 1 e 2.

Vale ressaltar, que modelar a irradiância solar usando uma função densidade de probabilidade adequada fornece alguns parâmetros chave. É por isso que é desejável usar essa modelagem nos problemas de planejamento a longo prazo. A função de distribuição weibull é uma das FDP mais comuns e usadas.

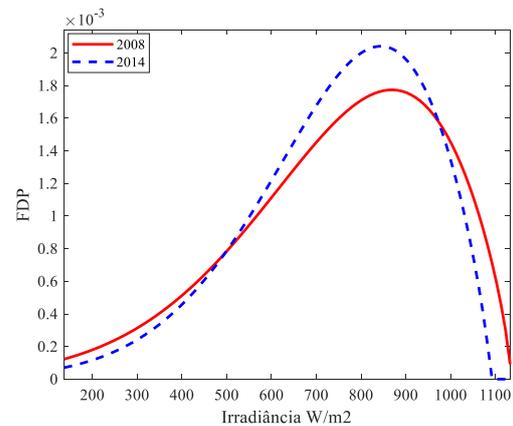


Figura 1: Comparativo entre os dados de irradiância solar média às 12 horas no período seco de 2008 e 2014

Em sequência, foram testadas as distribuições de probabilidade candidatas. Os parâmetros de cada distribuição avaliada foram estimados pelo método de máxima verossimilhança e em seguida cada distribuição foi submetida ao teste de Kolmogorov Smirnov com nível de significância de 5%.

As distribuições lognormal e gamma apresentaram diversos erros, elas foram, portanto, descartadas. Na Figura 2 é apresentado o comportamento da irradiância média decenal no período chuvoso referente aos anos de 2008 a 2017.

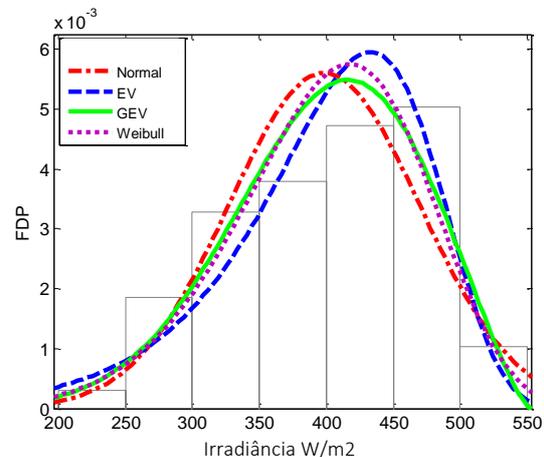


Figura 2: Histograma da irradiância média decenal no período chuvoso e diferentes distribuições de probabilidade.

Na Figura 3 é apresentado o histograma da irradiância média decenal no período seco de 2008 a 2017.

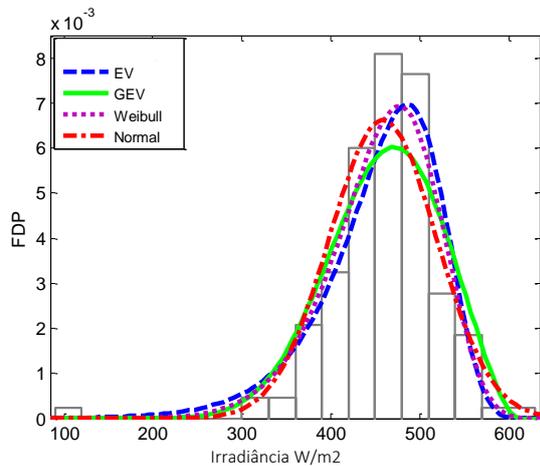


Figura 3: Histograma da irradiância média decidual no período seco e diferentes distribuições de probabilidade.

A FDP associada à hora de maior intensidade de irradiância foi a FDP de 12 horas da tarde. Deste modo, as FDP da hora 12 do período seco e chuvoso foram apresentadas nas Figuras 4 e 5, nas quais as amostras de dados foram divididas em decêndios em todos os valores de irradiância referentes a 12 horas durante os 10 anos de dados.

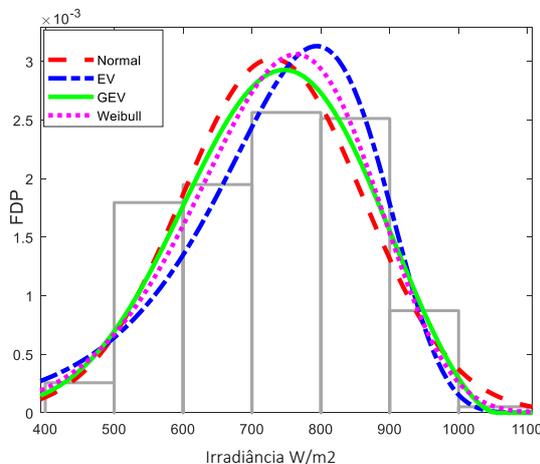


Figura 4: Histograma da irradiância média às 12 no período chuvoso e diferentes distribuições de probabilidade.

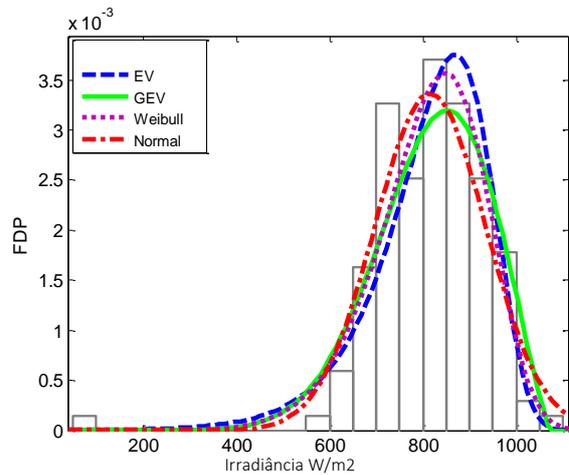


Figura 5: Histograma da irradiância média às 12 no período seco e diferentes distribuições de probabilidade.

Aplicou-se o teste de aderência Kolmogorov-Smirnov para todas as distribuições de probabilidade citadas anteriormente. As distribuições normal, lognormal, gamma e *t location scale* apresentaram diversos erros em diferentes horários, inclusive erros às 12 horas. Elas foram, portanto, descartadas.

Por fim, a distribuição weibull, apresentou erros apenas entre 16 horas, 17 horas e 18 horas para ambas as estações.

Na Tabela 1 é apresentado os resultados do teste de aderência de Kolmogorov-Smirnov para ambos os períodos (chuvoso e seco) e para os horários de maior incidência. Uma vez realizado o teste, as distribuições gamma e lognormal foram rejeitadas devido seu valor-p ser menor que 5%. Em seguida, observando bastante consistência nos resultados obtidos utilizando a distribuição GEV, optou-se por utilizá-la neste trabalho de acordo com as Tabelas 2 e 3, que contém os parâmetros da distribuição e as suas características (entropia, média e variância).

Tabela 1: Valor-p do teste de aderência para os dados de irradiância solar de diferentes distribuições.

	Valor – p dos testes de aderência						
	T location scale	Normal	EV	GEV	Gamma	Lognormal	Weibull
<b>Período Chuvoso</b>	0,1168	0,1208	0,2398	0,2873	0,0496	0,0349	0,1891
<b>Período Seco</b>	0,04338	0,4192	0,2339	0,3135	0,0974	0,0216	0,4785
<b>Chuvoso 12 horas</b>	0,1796	0,1963	0,4262	0,4115	0,1600	0,1345	0,4073
<b>Seco 12 horas</b>	0,4829	0,6064	0,1789	0,4588	0,0240	0,0008	0,3592

Tabela 2: Valores dos parâmetros encontrados e Valor – p do teste de aderência para os dados no período chuvoso.

Hora	GEV				
	k	$\sigma$	$\mu$	Média	Valor-p
6 h	0,001	10,38	19,49	20,28	0,85
7 h	-0,05	30,50	110,32	126,47	0,84
8 h	-0,27	57,80	269,63	290,46	0,82
9 h	-0,50	103,8	427,5	451,04	0,38
10 h	-0,48	118,3	580,3	608,36	0,76
11 h	-0,42	133,1	663,96	699,46	0,51
12 h	-0,37	136,1	688,72	729,29	0,42
13 h	-0,40	134,9	657,85	695,29	0,07
14 h	-0,34	121,9	560,69	599,10	0,30
15 h	-0,31	105,2	423,01	458,03	0,49
16 h	-0,18	77,67	256,82	289,46	0,37
17 h	0,05	45,24	80,74	109,57	0,07
18 h	1,22	2,63	1,85	9,40	0,06

Tabela 3: Valores dos parâmetros encontrados e Valor – p do teste de aderência para os dados no período seco.

Hora	GEV				
	k	$\sigma$	$\mu$	Média	Valor-p
6 h	0,003	20,48	29,49	41,38	0,88
7 h	-0,16	45,34	156,11	175,9	0,72
8 h	-0,35	106,26	315,14	347,8	0,55
9 h	-0,31	106,78	502,67	538,4	0,25
10 h	-0,38	120,66	657,58	692,7	0,34
11 h	-0,49	135,25	757,32	788,5	0,22
12 h	-0,45	129,98	783,41	816,4	0,34
13 h	-0,64	111,84	755,93	773,6	0,51
14 h	-0,50	100,89	642,56	665,0	0,59
15 h	-0,48	90,57	501,51	522,9	0,35
16 h	-0,47	59,55	301,40	315,7	0,54
17 h	-0,30	38,92	96,34	109,4	0,28
18 h	0,11	4,26	3,38	24,26	0,23

Verifica-se, portanto, que a distribuição mais adequada aos dados de irradiância é a GEV tanto para os dados divididos em períodos como para os dados de hora em hora, pois é a única distribuição que consegue representar todos os horários de cada período e apresenta o maior valor-p e o menor valor de estabilidade dentre as testadas.

A função GEV possui três parâmetros: forma, escala e posição. O parâmetro de forma representa onde a maioria dos dados está concentrada. O parâmetro de escala descreve quão espalhado a distribuição é; na medida em que o valor do parâmetro cresce, a função se torna mais espalhada. O parâmetro de posição descreve a posição da distribuição em uma dada direção do eixo horizontal (Millington et al, 2011). As Equações 2 e 3 denotam respectivamente, a distribuição GEV.

$$F_y(y) = \exp\{-[1-k(y-\beta/\alpha)]^{1/k}\} \quad (2)$$

$$x(TR) = \beta + \frac{\alpha}{k} \{1 - [-\ln(1-1/TR)]\}^k \quad (3)$$

$F_y(y)$ : probabilidade acumulada de ocorrência de um evento  $y$ ;

$x(TR)$ : quantil para determinado tempo de retorno TR;

$k, \alpha, \beta$ : parâmetros de forma, escala e posição respectivamente;

TR: tempo de retorno.

Naturalmente, o fato de a distribuição GEV conter três parâmetros confere a ela mais um grau de liberdade quando comparada às distribuições beta, Weibull e normal. Tal característica enseja explicar a flexibilidade da GEV, bem como sua maior aderência aos dados de radiação. No entanto, outra distribuição que também contém três parâmetros (*t location scale*) foi testada e não apresentou resultados semelhantes comparadas a GEV.

## 5 Conclusões

Um modelo para irradiância solar foi desenvolvido utilizando uma distribuição de probabilidade generalizada de valores extremos. Este modelo apresentou os melhores resultados de representar os dados de radiação solar em um intervalo de dez anos de dados, quando comparada com outras distribuições de probabilidade comumente apresentadas na literatura.

Os procedimentos para testes de distribuição de probabilidade apresentados constituem ainda fonte de contribuição uma vez que são passíveis de replicação em aplicações semelhantes. Em relação aos tempos computacionais não implica devido ser um processo continuado.

Considera-se tema de trabalhos futuros a formalização matemática que justifique a utilização da distribuição generalizada de valores extremos para modelagem de dados resultantes da média decenal da radiação solar.

## Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e a Coordenação de Pós-Graduação em Engenharia Elétrica da UFCG (COPELE), e aos revisores pelas recomendações sugeridas.

## Referências Bibliográficas

Araújo, E. M., Silva, I. N., Oliveira, J. B., et al. (2010). Aplicação de seis distribuições de probabilidade a séries de temperatura máxima em Igua-

- tu – CE. Revista Ciência Agrônômica, Vol. 41, n.1, pp. 36-45.
- Assis, F.N., Arruda, H.V., Pereira, A.R (1996). Aplicações de estatística à climatologia: teoria e prática. Pelotas: UFPEL, pp.161.
- Atwa, Y. M., El-Saadany, E. F., Salama e Seethapath, R (2010). Optimal Renewable Resources Mix for Distribution System Energy Loss Minimization. IEEE Transactions on Power Systems, Vol. 25, n. 1, pp. 360- 370.
- Cargnelutti, A, Matzenauer, R., Trindade, J.K (2004). Ajustes de funções de distribuição de probabilidade à radiação solar global no Estado do Rio Grande do Sul. Pesquisa Agropecuária Brasileira. v.39, n.12, p.1157-1166.
- Campos, H. (1983). Estatística experimental não-paramétrica. 4.ed. Piracicaba: Departamento de Matemática e Estatística - ESALQ, pp. 349.
- Cordeiro, I.A (2016). Modelo Generalizado de Valores Extremos Aplicado para Irradiação Solar e Análise de Confiabilidade de Sistemas Fotovoltaicos Residenciais. Dissertação, Universidade Federal de Campina Grande.
- Geografia da Paraíba. Portal Iparaíba (2018) Disponível em: <<http://www.paraibatotal.com.br/a-paraiba/geografia>>. Acesso em: 11/03/2018.
- INMET. Instituto Nacional de Meteorologia. (2018). Disponível em: <<http://www.inmet.gov.br>>. Acesso em: 20/03/2018.
- Martins, C. A; Uliana, E. M; Reis, E. F (2011). Estimativa da vazão e da precipitação máxima utilizando modelos probabilísticos na bacia hidrográfica do rio Benevente. Enciclopédia biosfera, Vol.7, pp. 1130-1142.
- Millington, N; Das, S; Simonovic, S, P (2011). The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models. Department of Civil and Environmental Engineering, The University of Western Ontario – Canada, report n 77.
- Morettin, P.A. and Bussab, W.O (2004). Estatística básica. 5.ed. São Paulo: Saraiva, pp. 526.
- Nunes, F. A. H (2006). Estimativa da Evapotranspiração de Referência em Campina Grande-PB. Dissertação em Meteorologia. UFCG.
- Oliveira, S. A (2016). Geração Distribuída Fotovoltaica: Avaliação dos Impactos Técnicos e um Breve Estudo sobre Viabilidade Econômica. Dissertação de Mestrado. UFCG.
- Pinho, J. T. and Galdino, M. A (2014). Manual de Engenharia para Sistemas Fotovoltaicos, CEPEL–CRESESB, Rio de Janeiro.
- Ribeiro, B, T., Avanzi, J, C., Melo, C, R., et al.(2007). Comparação de distribuições de probabilidade e estimativa da precipitação provável para região de Barbacena, MG. Revista Ciência e Agrotecnologia, Vol.31, n.5, pp.1297-1302.
- Silva, E. N, M., Rodrigues, A. B. e Silva, M. G (2016). Estimação Probabilística do Nível de Distorsão Harmônica Total de Tensão em Redes de Distribuição Secundárias. XXI Congresso Brasileiro de Automática-CBA, UFES, Vitória – ES, pp. 54-59.
- Villalva, M. G. Energia solar fotovoltaica: Conceito e Aplicações. São Paulo: Érica, 2013.