

ESTUDO COMPARATIVO DE MODELOS DE ESTIMATIVA DA RADIAÇÃO DIFUSA NO SERTÃO DA PARAÍBA, REGIÃO NORDESTE DO BRASIL

Diego Rodrigues^{1,2}, Leonardo Petribú^{1,2}, Janis Galdino^{1,2}, Olga de C. Vilela^{1,2}

¹Departamento de Energia Nuclear da Universidade Federal de Pernambuco (DEN-UFPE).
Av. Prof. Luiz Freire, s/n, Recife-PE, Brasil.

²Centro de Energias Renováveis da Universidade Federal de Pernambuco (CER-UFPE).
Av. da Arquitetura s/n, Recife-PE, Brasil.

Recibido 14/08/19, aceptado 22/10/19

RESUMO: Modelos de separação da irradiância global horizontal em irradiância difusa horizontal são importantes para o cálculo de energia produzida por sistemas de geração solar. Neste estudo, sete modelos de separação foram aplicados ao sertão da Paraíba, na Região Nordeste do Brasil. Os modelos foram divididos em modelos de um parâmetro e multiparâmetros, gerando-se para cada um a dispersão da razão difusa (k_d) em função do índice de claridade (k_t). O desempenho dos modelos foi avaliado por meio de análise gráfica, utilizando o diagrama de Taylor, e estatísticos como o *skill score*, raiz do erro quadrático médio, erro médio absoluto e razão entre os desvios padrão. O modelo multiparâmetro de Skartveit apresentou os melhores resultados.

Palavras-chave: irradiância difusa horizontal, índice de claridade, razão difusa, multiparâmetros e diagrama de Taylor.

INTRODUÇÃO

A irradiância global horizontal (I_g) e a temperatura das células são os parâmetros mais importantes para simulação da potência elétrica de saída de um módulo fotovoltaico nos projetos de geração centralizada. Todavia, para calcular a quantidade de irradiância incidente no plano inclinado, é necessário decompor a irradiância global horizontal em suas frações de irradiância direta normal (I_{bn}) e difusa horizontal (I_d). A modelagem da irradiância difusa – muitas vezes ausente de medições terrestres – é o grande desafio nesse cálculo, haja vista o seu comportamento estocástico.

Um dos trabalhos pioneiros na modelagem da radiação difusa foi o de Liu e Jordan (1960). Os autores propuseram uma relação para a razão difusa, k_d (I_d/I_g), em função do índice de claridade, k_t (I_g/I_{oh}) – aqui, I_{oh} faz menção à irradiância extraterrestre no plano horizontal. Essa modelagem da razão difusa em função de um parâmetro (índice de claridade), resultou no desenvolvimento de diferentes modelos propostos por outros autores que utilizam k_t como único parâmetro. Dentre os mais citados, Orgill e Hollands (1976), Collares-Pereira e Rabl (1979) e Erbs et al. (1982). Os modelos de um único parâmetro são também os mais utilizados comercialmente em *softwares* usuais para dimensionamento e simulação de usinas de geração fotovoltaica centralizada.

A dispersão de k_d em função de k_t gera uma nuvem de pontos que apresenta uma alta variabilidade para pequenos intervalos do índice de claridade. Uma limitação dos modelos de um parâmetro é a dificuldade em reproduzir o alto desvio padrão associado à medição da radiação difusa nesses pequenos intervalos de k_t . Uma forma de modelar essa nuvem de pontos é mediante o acréscimo de mais parâmetros na entrada dos modelos de estimativa da difusa. Autores como Skartveit e Olseth (1987) sugerem que um dos parâmetros importantes na análise é a elevação solar (α). Reindl et al. (1990) agrega, além da elevação solar, a temperatura ambiente e umidade relativa do ar. Ridley et al. (2010) sugere a elevação solar, a hora solar aparente, o índice de claridade diário e um parâmetro de persistência do índice de claridade.

O presente estudo analisa modelos de um parâmetro e multiparâmetros para separação da irradiância global horizontal em sua componente difusa. O estudo analisa os modelos aplicados a três localidades no semiárido do estado da Paraíba, Região Nordeste do Brasil, e utiliza o diagrama de Taylor (Taylor, 2001) para escolha do melhor modelo para a região. O Nordeste brasileiro apresenta a média mais elevada de radiação solar do país e concentra uma alta quantidade de projetos de geração fotovoltaica

centralizada, sendo responsável por aproximadamente 70% da geração elétrica de origem fotovoltaica no Brasil em 2017 (BEN, 2018). A decomposição da irradiância global é parte essencial do cálculo para estimativa de produção de energia de centrais fotovoltaicas. Portanto, é bastante relevante a aplicação do presente estudo para a região.

BASE DE DADOS UTILIZADAS E PROCEDIMENTO DE QUALIDADE

Os dados utilizados compreendem medições de irradiância global horizontal (I_g), irradiância difusa horizontal (com esferas de sombreamento) (I_d), irradiância direta normal (I_{bn}), velocidade do vento, direção do vento, temperatura, umidade relativa e pressão atmosférica. Foram utilizados piranômetros classe 1 modelo CMP6 da Kipp&Zonen para medição de I_g e I_d (com esferas de sombreamento) e um pireliômetro classe 1 modelo CHP1 da Kipp&Zonen para medição de I_{bn} . O rastreador solar utilizado foi o Solys2 da Kipp&Zonen. As estações solarimétricas foram instaladas nos municípios de São João do Rio do Peixe, Coremas e Patos, próximos entre si, todos no estado da Paraíba. A estação foi instalada conforme requisitos da Empresa de Pesquisa Energética (EPE), com sensores solarimétricos classe 1 e em local livre de obstáculos que possam causar sombreamento, apresentando, portanto, boas práticas de medições. A disponibilidade de dados é de setembro de 2012 a dezembro de 2018, sendo a medição realizada a cada segundo e armazenada no *datalogger* em médias integralizadas de minuto em minuto. Os dados foram integralizados em médias horárias de forma a serem utilizados nos modelos de estimativa da irradiância difusa, já que esses modelos foram elaborados considerando uma base horária.

A natureza estatística dos modelos aqui utilizados faz com que seus resultados sejam altamente dependentes da qualidade dos dados observacionais. A fim de eliminar os erros e incertezas na entrada dos modelos, os dados observacionais foram submetidos a um procedimento de garantia de qualidade, tal como em Petribú et al. (2017). O procedimento consiste em três blocos de testes: i) testes globais, que avaliam a série temporal como um todo, além de identificar inconformidades na data e hora de registro de medições; ii) testes físicos, que avaliam a coerência física das medições realizadas e iii) testes comparativos, que avaliam a coerência entre as medições de I_g , I_{bn} e I_d . Os testes globais visam identificar as inconsistências no registro do instante de medição das variáveis pelo *datalogger*. Primeiramente, verifica-se o ordenamento cronológico dos registros de data e hora, ordenando (juntamente com as medições associadas) aqueles que estejam em posições cronologicamente equivocadas. Em seguida, busca-se por registros de data e hora idênticos ao longo da série temporal, eliminando aqueles que possuem igualdade em registro e diferença nas medições associadas. Caso registros idênticos possuam conjuntos de medições idênticos, mantêm-se um deles, anulando os demais. Por último, buscam-se por lacunas entre registros consecutivos na série temporal, preenchendo-as com *Not a Number (NaN)*. As Tabelas 1 e 2 apresentam um resumo dos testes físicos e comparativos. Os dados anômalos recebem *NaN*; caso um dado seja considerado suspeito em dois testes distintos, o mesmo é considerado anômalo.

Dados fora dos limites para k_r , k_d e k_n (k) são anômalos.	$0 < k < 1,1$	
Testes propostos pela <i>Baseline Surface Radiation Network (BSRN)</i> : os dados que desrespeitarem os limites para dados fisicamente impossíveis (i) são considerados anômalos, enquanto aqueles que desrespeitarem os limites para dados extremamente raros (ii) são considerados suspeitos. I_{oeff} é a irradiância extraterrestre incidente em um plano normal.	i)	$-4 \leq I_g \leq I_{oeff} 1,5 \cos(\theta_Z)^{1,2} + 100$ $-4 \leq I_{bn} \leq I_{oeff}$ $-4 \leq I_d \leq I_{oeff} 0,95 \cos(\theta_Z)^{1,2} + 50$
	ii)	$-2 \leq I_g \leq I_{oeff} 1,2 \cos(\theta_Z)^{1,2} + 50$ $-2 \leq I_{bn} \leq I_{oeff} 0,95 \cos(\theta_Z)^{0,2} + 10$ $-2 \leq I_d \leq I_{oeff} 0,75 \cos(\theta_Z)^{1,2} + 30$
Modelo de céu claro (<i>cc</i>) de Rigollier et al. (2000): dados que desrespeitarem os limites são considerados dados anômalos. Foi utilizado um índice de turbidez de Link de 2,5, conforme Ruiz-Arias et al. (2010).	$I_g \leq I_{g,cc} + 1,1 I_{g,cc}$	
	$I_{bn} \leq I_{bn,cc} + 1,1 I_{bn,cc}$	
	$I_d \geq I_{d,cc} - 0,2 I_{d,cc}$	

Tabela 1: Teste físicos.

<p><i>Tracker off</i>: caso as quatro condições sejam satisfeitas no mesmo instante de tempo, o dado é considerado anômalo.</p>	$\frac{I_d > 50 \text{ W/m}^2 \text{ e } \theta_z < 75^\circ \text{ e } I_d + I_{bn} \cos(\theta_z)}{I_{g,cc}} > 0,85 \text{ e } \frac{I_d}{I_g} > 0,85$
<p>Coerência: os dados são considerados suspeitos para cada condição especificada.</p>	<p>i) $\left \frac{I_d + I_{bn} \cdot \cos(\theta_z) - I_g}{I_g} \right \leq 0,15 \text{ se } 75^\circ < \theta_z < 93^\circ \text{ e } I_d + I_{bn} \cos(\theta_z) > 50 \text{ W/m}^2$</p> <p>ii) $\left \frac{I_d + I_{bn} \cdot \cos(\theta_z) - I_g}{I_g} \right \leq 0,08 \text{ se } \theta_z < 75^\circ \text{ e } I_d + I_{bn} \cos(\theta_z) > 50 \text{ W/m}^2$</p> <p>iii) $k_d < 1,05 \text{ para } \theta_z < 75^\circ \text{ e } I_g > 50 \text{ W/m}^2$</p> <p>iv) $k_d < 1,1 \text{ para } 75^\circ < \theta_z < 93^\circ \text{ e } I_g > 50 \text{ W/m}^2$</p> <p>v) $k_d \geq 0,9 \text{ para } k_t < 0,2$</p> <p>vi) $k_d \leq 0,8 \text{ para } k_t > 0,6$</p>
<p>Envelope</p>	<p>-</p>

Tabela 2: Testes comparativos.

O teste envelope destaca uma tendência na dispersão $k_d \times k_t$, buscando para tal, traçar curvas envelopes para os limites superiores e inferiores da dispersão. Conforme proposto em Younes et al. (2005), o eixo do índice de claridade foi dividido em 10 intervalos, cada intervalo comportando um subconjunto da dispersão. Para cada intervalo, são calculados a média (\bar{k}_d) e o desvio padrão (σ_{k_d}) da razão difusa e são definidas barras superiores e inferiores, através dos valores de, respectivamente, $\bar{k}_d + 2\sigma_{k_d}$ e $\bar{k}_d - 2\sigma_{k_d}$. Para traçar a curva envelope, aplicou-se médias móveis ao conjunto de dados formados pelas barras superiores e inferiores. A média móvel suaviza os pontos que delineam os limites superiores e inferiores. Ajusta-se, então, um polinômio de grau 8, utilizando o método dos mínimos quadrados, aos pontos resultantes da média móvel.

METODOLOGIA

Foram selecionados três modelos de um parâmetro, o de Orgill e Hollands (1976), um dos pioneiros na modelagem do índice de claridade em diferentes intervalos, o de Erbs et al. (1982), por ser largamente utilizado em *softwares* desenvolvidos para simulações de energia solar fotovoltaica, e o de Chandrasekaran e Kumar (1994), por ter feito uma modelagem baseada em dados tropicais; e quatro multiparâmetros bastante citados na literatura, o de Reindl et al. (1990), Perez et al. (1992), Skartveit et al. (1998) e Ridley et al. (2010). Os modelos foram utilizados, em geral, na sua formulação original.

Para a estimativa da irradiância difusa, Orgill e Hollands (1976), Erbs et al. (1982) e Chandrasekaran e Kumar (1994) propõem modelos baseados em equações da razão difusa (k_d) em função de um único parâmetro, o índice de claridade (k_t), para diferentes intervalos de k_t . No primeiro modelo (Orgill e Hollands, 1976), foram utilizados dados de um aeroporto em Toronto, Canadá, de setembro de 1967 a agosto de 1971. A partir do agrupamento do índice de claridade em intervalos de 0,05, foi aplicada a média aos correspondentes valores da razão difusa (I_g/I_d). Já o segundo modelo (Erbs et al., 1982) teve como base medições locais de quatro cidades dos Estados Unidos (Livermore, Raleigh, Fort Hood e Maynard), fazendo o uso de médias horárias, diárias e mensais da irradiância global horizontal. Por fim, o terceiro modelo (Chandrasekaran e Kumar, 1994) utiliza medições realizadas em Madra, na Índia (clima tropical), de 1983 a 1987, para obter as equações de k_d em função de k_t .

No modelo de Reindl et al. (1990), foram empregados dados dos Estados Unidos (Albany, Nova Iorque e Cape Canaveral, Flórida) e Europa (Copenhague, Hamburg, Valentia e Oslo). Com os dados, os autores analisaram regressões com até 28 potenciais variáveis regressoras de forma a analisar os melhores parâmetros para estimativa da razão difusa nos locais em questão. Os quatro melhores parâmetros para representar os desvios na razão horária difusa foram: o índice de claridade, a elevação solar (α), a temperatura ambiente (T , em °C) e a umidade relativa do ar (ϕ , em percentual).

Maxwell (1987) modela a irradiância direta normal (I_{bn}) utilizando equações baseadas em princípios físicos. Para a entrada do modelo, utiliza-se k_t , a massa de ar (AM) e um limite (k_{nc}) calculado a partir do modelo de céu claro de Bird e Hulstrom (1981) para a transmitância da irradiância direta normal, k_n (I_{bn}/I_{oeff}) – aqui, I_{oeff} faz menção à irradiância extraterrestre incidente em um plano normal a radiação para a hora considerada. Perez et al. (1992) propuseram uma correção ao modelo de Maxwell (1987) para considerar os seguintes parâmetros que dependem de condições de insolação: o ângulo zenital (θ_z), um índice de claridade dependente da massa de ar (Kt'), um índice de estabilidade ($\Delta Kt'$) e a água precipitável (W). Essa correção é feita utilizando-se um coeficiente $X(Kt', \theta_z, W, \Delta Kt')$, aplicado na saída do modelo de Maxwell, conforme Equação 1; o coeficiente X é obtido diretamente a partir da busca em uma matriz $6 \times 6 \times 5 \times 7$ baseada em intervalos definidos dos parâmetros de X . As Equações de 2 a 4 apresentam as expressões para Kt' e $\Delta Kt'$; os subscritos $i+1$ e $i-1$ da Equação 3 referem-se ao instante posterior e anterior ao índice i . Caso Kt'_{i+1} ou Kt'_{i-1} seja ausente, utiliza-se a Equação 4. Devido a não convergência do modelo de Perez et al. (1992) para valores de ângulo zenital (θ_z) maiores que 85° (valores de baixo k_t), utilizou-se o modelo de Reindl et al. (1990) no intervalo para $k_t < 0,3$. Para obter a irradiância difusa, utilizou-se a relação entre as irradiâncias, conforme Equação 5 abaixo.

$$I_{bn,Perez} = I_{bn,Maxwell} X(Kt', \theta_z, W, \Delta Kt') \quad (1)$$

$$Kt' = \frac{k_t}{1,031 \exp(-1,4/(0,9 + 9,4/AM)) + 0,1} \quad (2)$$

$$\Delta Kt' = 0,5(|Kt' - Kt'_{i+1}|) + (|Kt' - Kt'_{i-1}|) \quad (3)$$

$$\Delta Kt' = |Kt' - Kt'_{i\pm 1}| \quad (4)$$

$$I_d = I_g - I_{bn} \cos \theta_z \quad (5)$$

Skartveit et al. (1998) propuseram um modelo não linear composto por extensas equações para cada um dos intervalos do índice de claridade. O modelo é baseado em três principais parâmetros: a elevação solar (α , em graus), o índice de claridade (k_t) e o índice de variabilidade horária (σ_3). Os autores sugerem o parâmetro σ_3 para indicar a presença de nuvens não homogêneas (variáveis). Esse parâmetro depende do índice de céu claro ρ (calculado a partir de k_t , conforme Equação 6), e é definido nas Equações 7 e 8, esta última utilizada nos horários de nascer e pôr-do-sol.

$$\rho = \frac{k_t}{0,83 - 0,56e^{(-0,06\alpha)}} \quad (6)$$

$$\sigma_3 = \{[(\rho - \rho_{t-1})^2 + (\rho - \rho_{t+1})^2]/2\}^{0,5} \quad (7)$$

$$\sigma_3 = |\rho - \rho_{t\pm 1}| \quad (8)$$

Na prática, como σ_3 é calculado a partir de três valores consecutivos da série de índice de céu claro, conforme mostra a Equação 7, para cada lacuna existente nessa série, três são geradas na série de σ_3 . Para contornar esse problema, os autores sugerem as expressões para σ_3 , mostradas nas Equações 9 e 10, utilizadas no caso em que o parâmetro σ_3 retorna uma lacuna.

$$\sigma_3 = 0,021 + 0,397\rho - 0,231\rho^2 - 0,13e^{\left\{-\left[\frac{(\rho-0,931)}{0,134}\right]^2\right\}^{0,834}} \quad \text{para } \rho \leq 1,04 \quad (9)$$

$$\sigma_3 = 0,12 + 0,65(\rho - 1,04) \quad \text{para } \rho \geq 1,04 \quad (10)$$

De acordo com o índice de variabilidade σ_3 , os autores subdividem a série temporal em um período de “horas invariáveis” para o qual $\sigma_3 \approx 0$ e “horas variáveis” para o qual $\sigma_3 > 0$. Para o primeiro período, são propostas equações que dependem do índice de claridade e da elevação solar. Já para o segundo período, σ_3 indica que deve ser adicionado um termo Δ – que depende de k_t , α e σ_3 – às equações propostas para “horas invariáveis”, a fim de considerar os efeitos de nuvens não homogêneas. O valor de σ_3 considerado próximo de zero foi 0,001.

Ridley et al. (2010) verificaram que os modelos propostos até então não se adequavam a regiões como a Austrália, pois foram obtidos por meio de dados dos Estados Unidos e Europa. Eles propuseram, então, um modelo não linear e multiparâmetros que, segundo os autores, pode ser usado nos

hemisférios norte e sul. O modelo engloba o índice de claridade horário (k_t), índice de claridade diário (K_t), a elevação solar (α , em radianos), a hora solar aparente (HS) e uma medida da persistência (ψ) de k_t como parâmetros. O horário solar aparente deriva diretamente do cálculo do ângulo horário, sendo a hora solar aparente utilizada na unidade de horas nesse modelo. A persistência (ψ) citada leva em conta as médias dos índices de claridade anterior e posterior a dado intervalo t , conforme mostra a Equação 11.

$$\psi = \begin{cases} \frac{k_{t-1} + k_{t+1}}{2} & \text{pôr - do - sol} < t < \text{nascido - do - sol} \\ k_{t+1} & t = \text{pôr - do - sol} \\ k_{t-1} & t = \text{nascido - do - sol} \end{cases} \quad (11)$$

De maneira similar ao cálculo de σ_3 no trabalho de Skartveit et al. (1998), o cálculo do parâmetro da persistência (ψ) pode resultar em mais lacunas do que as presentes na série de k_t , devido a média feita para os índices anterior e posterior. Para contornar esse problema, propõe-se uma pequena alteração na função persistência: se os dois índices (anterior e posterior a um intervalo t) são ausentes (NaN), a persistência recebe NaN ; contudo, se somente um dos índices anterior ou posterior for ausente, e o índice do intervalo t for um valor válido, é realizada a média entre o índice do intervalo t e o índice não ausente.

A Tabela 3 mostra as Equações propostas por cada modelo, com exceção do modelo de Skartveit et al. (1998) e Maxwell (1987), que, devido às extensas equações propostas, decidiu-se omiti-las do presente trabalho, podendo ser consultadas nos artigos citados.

Modelos	Equações
Orgill e Hollands (1976)	$k_d = 1,0 - 0,249k_t$ para $k_t < 0,35$ $k_d = 1,557 - 1,84k_t$ para $0,35 \leq k_t \leq 0,75$ $k_d = 0,177$ para $k_t > 0,75$
Chandrasekaran e Kumar (1994)	$k_d = 1,0086 - 0,178k_t$ para $k_t \leq 0,24$ $k_d = 0,9686 + 0,1325k_t + 1,4183k_t^2 - 10,1860k_t^3 + 8,3733k_t^4$ para $0,24 < k_t \leq 0,8$ $k_d = 0,197$ para $k_t > 0,8$
Erbs et al. (1982)	$k_d = 1,0 - 0,09k_t$ para $k_t \leq 0,22$ $k_d = 0,9511 - 0,1604k_t + 4,388k_t^2 - 16,638k_t^3 + 12,336k_t^4$ para $0,22 < k_t \leq 0,8$ $k_d = 0,165$ para $k_t \geq 0,8$
Perez et al. (1992), baseado no modelo de Maxwell (1987).	$W = \exp(0,07T_{dp} - 0,075)$ $T_{dp} = \frac{B_1 \left[\ln\left(\frac{\phi}{100}\right) + \frac{A_1 T}{B_1 + T} \right]}{A_1 - \ln\left(\frac{\phi}{100}\right) - \frac{A_1 T}{B_1 + T}}$ $A_1 = 17,625$ e $B_1 = 243,04$ °C T é a temperatura ambiente em °C e ϕ , a umidade relativa do ar, em %
Reindl et al. (1990)	$k_d = 1,000 - 0,232k_t + 0,0239\sin(\alpha) - 0,000682T_a + 0,0195\phi$ para $0 \leq k_t \leq 0,3$ e $k_d \leq 1,0$ $k_d = 1,329 - 1,716k_t + 0,267\sin(\alpha) - 0,00357T_a + 0,106\phi$ para $0,3 < k_t < 0,78$ e $k_d \leq 0,97$ e $k_d \geq 0,1$ $k_d = 0,426 - 0,256k_t + 0,267\sin(\alpha) - 0,00349T_a + 0,07344\phi$ para $0,78 \leq k_t$ e $k_d \leq 0,1$
Ridley et al. (2010)	$k_d = \frac{1}{1 + \exp(-5,38 + 6,63k_t + 0,06HS - 0,007\alpha + 1,75K_t + 1,31\psi)}$

Tabela 3: Equações propostas pelos diferentes modelos.

Para comparar os modelos, foi utilizado o diagrama de Taylor (Taylor, 2001), pois ele é uma ferramenta útil para análise estatística gráfica. Os estatísticos associados ao diagrama são o desvio padrão, a correlação e o *skill score* (SS4). No diagrama mostrado na Figura 3, o desvio padrão é

representado pela distância radial, partindo-se da origem, enquanto que a correlação é vista na posição azimutal. O *skill score* é representado pelas linhas tracejadas em vermelho, partindo-se do ponto da observação; ele indica o desempenho geral do modelo, já que leva em conta tanto a correlação quanto o desvio padrão do modelo, conforme Equação 12, onde r é o coeficiente de correlação entre o modelo e a observação e σ_m , o desvio do modelo.

$$SS4 = \frac{(1 + r)^4}{4 (\sigma_m + 1/\sigma_m)^2} \quad (12)$$

Utiliza-se, também, a raiz do erro quadrático médio (RMSD) normalizado pela média, o erro médio absoluto (MBE) normalizado pela média e a razão entre os desvios (σ_m/σ_o), onde σ_o é o desvio da observação. A razão entre os desvios está relacionada à amplitude de variação das séries observada e modelada, enquanto que o coeficiente de correlação à estrutura de fase e frequência entre as séries.

RESULTADOS

Todos os testes realizados no procedimento de garantia de qualidade resultaram nos percentuais apresentados na Tabela 4. Os dados não testados são aqueles com ângulo de elevação menor ou igual a 7° . Dados ausentes estão relacionados àqueles não adquiridos pelo sistema de aquisição de dados, em conjunto com os associados à inconformidades na data de registro. Os resultados dos testes físicos e comparativos são apresentados nos percentuais de dados anômalos para I_g , I_d e I_{bn} . Os resultados evidenciam que os dados observacionais são de boa qualidade, haja vista que acumulam menos de 10% de lacunas (ausentes + anômalos) para a maioria dos parâmetros – a exceção da difusa para Patos, que retorna um percentual ligeiramente maior que 10%, o que, no entanto, não inviabiliza a utilização de tais dados.

Dados	São João do Rio do Peixe	Coremas	Patos
Não testados	8,33%	8,33%	8,33%
Ausentes	4,63%	3,33%	7,68%
Anômalos I_g	1,26%	1,41%	1,61%
Anômalos I_d	3,96%	4,95%	4,17%
Anômalos I_{bn}	1,92%	1,55%	2,37%

Tabela 4: Resultados do procedimento de qualidade.

As dispersões de k_d em função de k_t são apresentadas para a estação de São João do Rio do Peixe, separando os modelos de um parâmetro (Figura 1) dos multiparâmetros (Figura 2). As estações de Coremas e Patos apresentaram dispersões similares as de São João do Rio do Peixe. Pode-se observar através das Figuras 1 e 2 que a dispersão dos dados medidos (pontos em cinza) apresenta uma alta concentração de valores elevados de k_t , com correspondentes frações difusa baixas, indício de uma irradiância direta normal elevada e céu mais claro. Já para baixos valores de k_t , nota-se frações difusa elevadas, com grande concentração de dados na região $1 < k_d < 1,1$. A dispersão apresenta, também, uma forma peculiar na região com $k_t < 0,4$ e $k_d < 0,8$, que corresponde em sua maioria aos dados com ângulos de elevação solar menores que 7° (horários de nascer e pôr-do-sol).

Na Figura 2, os dados observacionais estão representados pelos pontos em cinza, enquanto os modelos de Reindl et al. (1990), Perez et al. (1992), Skartveit et al. (1998) e Ridley et al. (2010), pelos pontos em rosa, laranja, verde e vermelho, respectivamente. Nos modelos multiparâmetros, alguns pontos estimados saíram fora da nuvem de pontos $k_d \times k_t$ medida. No modelo de Skartveit et al. (1998), é observada uma maior dispersão nos pontos para $k_t > 0,5$ devido ao termo $\Delta(k_t, \alpha, \sigma_3)$ de variabilidade adicionado. O deslocamento apresentado no modelo de Reindl et al. (1990) para $k_t < 0,3$ e $k_t > 0,8$ é característico das equações propostas pelos autores para os intervalos de k_t , conforme Tabela 3. Vale salientar que a região com $k_t > 0,58$ dos dados medidos (pontos em cinza) concentra mais de 50% dos dados observacionais para as três estações.

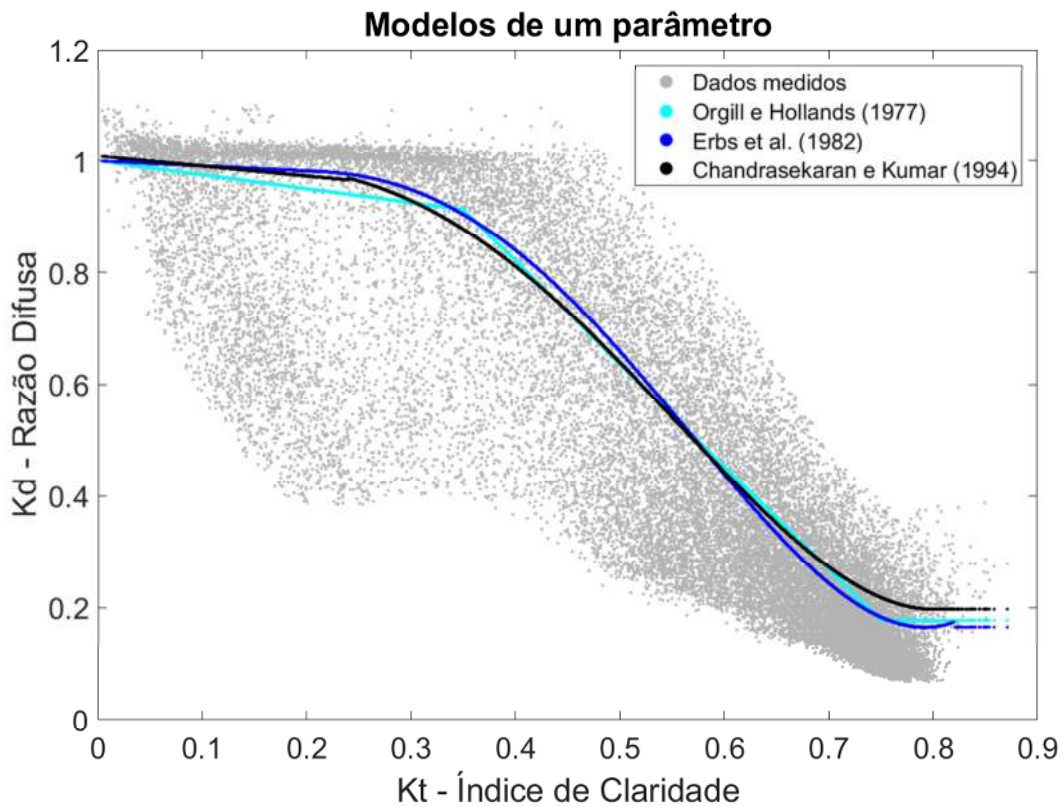


Figura 1: Estimativa da razão difusa pelos modelos de um parâmetro para São João do Rio do Peixe.

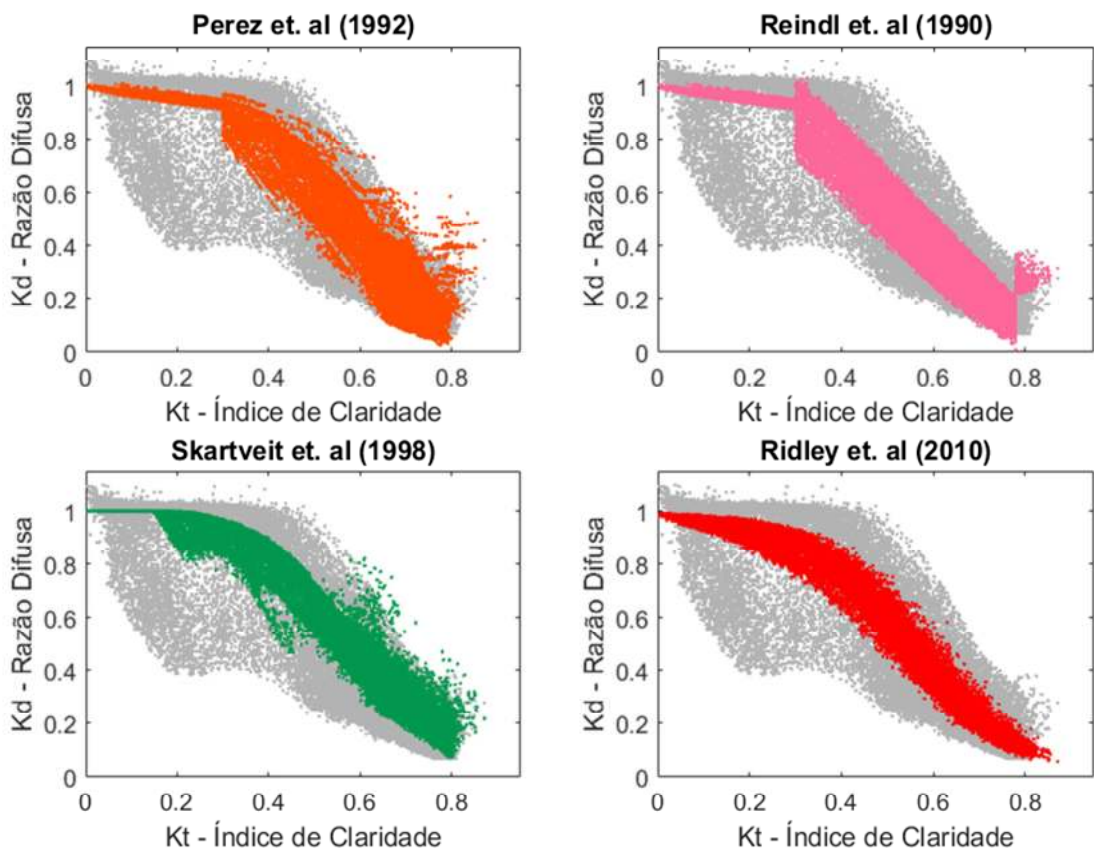
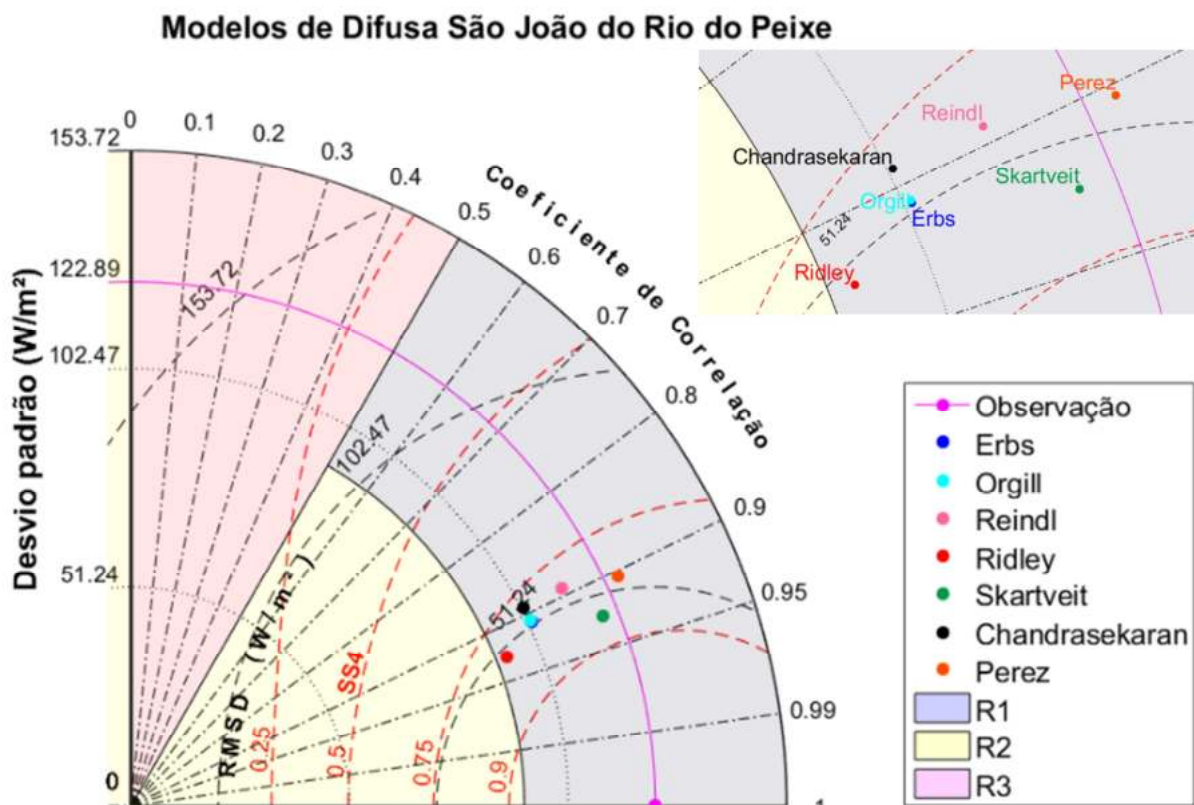


Figura 2: Estimativa da razão difusa pelos modelos multiparâmetros para São João do Rio do Peixe.

As Figuras 3, 4 e 5 apresentam os diagramas de Taylor para as séries de irradiância difusa modeladas e medida das três estações. O ponto em magenta representa a observação e a linha que parte dele representa o desvio padrão da observação. Quanto mais próximo da linha em magenta, melhor os modelos representam o desvio padrão da série de dados observacionais. Quanto mais próximo da observação, em relação a posição azimutal, maior correlação tem o modelo com as medições. As regiões R1, R2 e R3 são indicativos de desempenho dos modelos. A região R1 indica modelos com desvio padrão e correlação razoáveis; R2 representa modelos com correlações razoáveis e desvios distantes do desvio dos dados observacionais; por fim, a região R3 indica os piores modelos, com correlações muito baixas.

Os diagramas das três localidades apresentam resultados semelhantes. Os modelos de um parâmetro apresentaram resultados bastante similares entre si para as três localidades, exibindo um valor alto de correlação, porém desvio padrão distante da amostra. Dentre os modelos multiparâmetros, o de Reindl et al. (1990) tem desempenho inferior aos demais e o modelo de Ridley et al. (2010) exibiu valores de desvio padrão distantes das observações. De forma geral, o modelo de Skartveit et al. (1998) apresentou o melhor desempenho, com a melhor correlação e SS4. Isso se deve a uma boa representação da região com $k_t > 0,58$ por esse modelo, já que essa região concentra mais da metade dos dados medidos. Para a cidade de Coremas, o modelo de Reindl et al. (1990) reproduziu melhor o desvio padrão da série observada, apesar disso, o modelo de Skartveit et al. (1998) ainda pode ser apontado como o mais representativo (considerando os demais estatísticos).



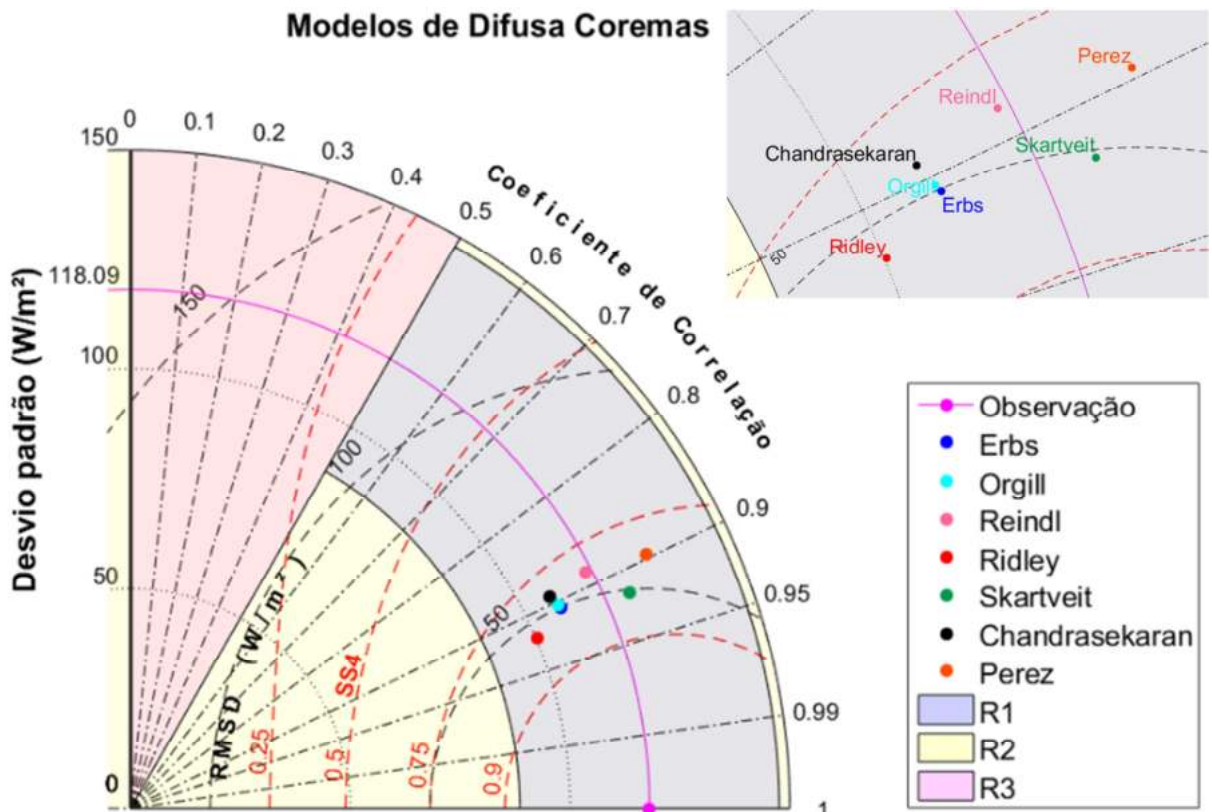


Figura 4: Diagrama de Taylor para Coremas.

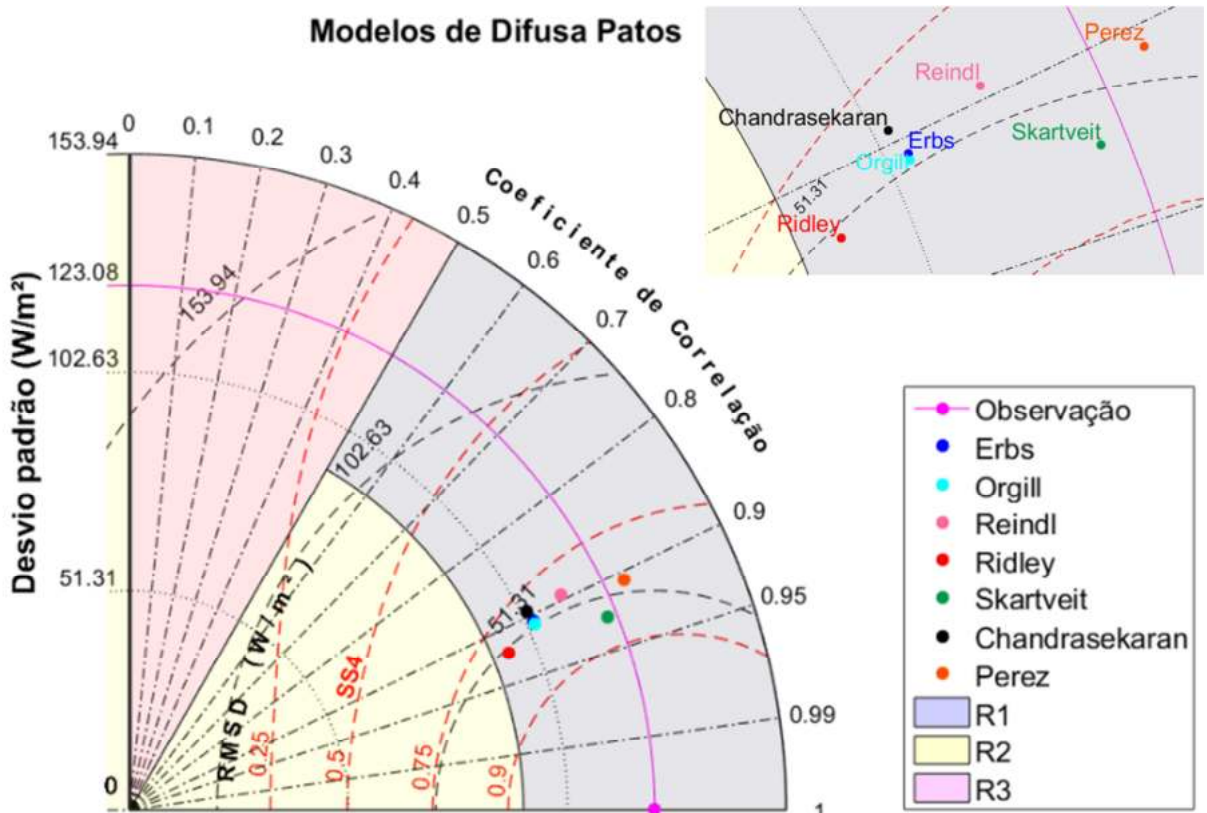


Figura 5: Diagrama de Taylor para Patos.

A Tabela 5, 6 e 7 apresentam os estatísticos analisados: o desvio padrão de todas as séries; a razão entre os desvios (σ_m/σ_o), relacionado à amplitude de variação das séries modeladas em relação à observada; a correlação (r), relacionada à estrutura de fase e frequência; o SS4, um índice de desempenho geral do modelo, considerando tanto a correlação quanto o desvio padrão; a raiz do erro quadrático médio (RMSD) normalizado pela média; e, por fim, o erro médio absoluto (MBE), também normalizado pela média. O modelo de Skartveit et al. (1998) apresenta melhores razão entre desvios, *skill score* e correlação (os principais estatísticos) para São João do Rio do Peixe e Patos. Para Coremas, Skartveit et al. (1998) apresenta o melhor *skill score*, mas possui um desvio maior, superestimando a série observada; ainda assim, como possui o melhor SS4, foi considerado o melhor modelo também para Coremas.

São João do Rio do Peixe						
Modelos	Desvio Padrão (W/m ²)	σ_m/σ_o	r	SS4	RMSD (%)	MBE (%)
Observação	122,90	1,00	1,00	1,00	0,00	0,00
Orgill e Hollands (1976)	103,33	0,84	0,91	0,80	0,36	-0,10
Erbs et al. (1982)	103,22	0,84	0,91	0,80	0,35	-0,07
Reindl et al. (1990)	112,99	0,92	0,89	0,80	0,37	-0,07
Perez et al. (1992)	126,16	1,03	0,90	0,82	0,37	-0,09
Chandrasekaran e Kumar (1994)	103,05	0,84	0,89	0,78	0,39	-0,13
Skartveit et al. (1998)	119,14	0,97	0,93	0,86	0,34	-0,14
Ridley et al. (2010)	94,84	0,77	0,93	0,81	0,34	0,08

Tabela 5: Estatísticos dos modelos para São João do Rio do Peixe.

Coremas						
Modelos	Desvio Padrão (W/m ²)	σ_m/σ_o	r	SS4	RMSD (%)	MBE (%)
Observação	118,09	1,00	1,00	1,00	0,00	0,00
Orgill e Hollands (1976)	107,78	0,91	0,90	0,81	0,36	-0,14
Erbs et al. (1982)	108,11	0,92	0,91	0,82	0,35	-0,12
Reindl et al. (1990)	116,81	0,99	0,89	0,79	0,38	-0,11
Perez et al. (1992)	130,78	1,11	0,90	0,80	0,41	-0,15
Chandrasekaran e Kumar (1994)	106,97	0,91	0,89	0,79	0,39	-0,17
Skartveit et al. (1998)	123,69	1,05	0,92	0,84	0,38	-0,20
Ridley et al. (2010)	100,23	0,85	0,92	0,83	0,30	0,03

Tabela 6: Estatísticos dos modelos para Coremas.

Patos						
Modelos	Desvio Padrão (W/m ²)	σ_m/σ_o	R	SS4	RMSD (%)	MBE (%)
Observação	123,08	1,00	1,00	1,00	0,00	0,00
Orgill e Hollands (1976)	104,59	0,85	0,91	0,81	0,34	-0,07
Erbs et al. (1982)	104,67	0,85	0,91	0,80	0,34	-0,05
Reindl et al. (1990)	113,11	0,92	0,90	0,80	0,35	-0,06
Perez et al. (1992)	127,72	1,04	0,91	0,82	0,36	-0,09
Chandrasekaran e Kumar (1994)	104,00	0,84	0,90	0,78	0,37	-0,11
Skartveit et al. (1998)	120,68	0,98	0,93	0,86	0,32	-0,13
Ridley et al. (2010)	96,26	0,78	0,92	0,81	0,33	0,10

Tabela 7: Estatísticos dos modelos para Patos.

CONCLUSÕES

Sete modelos para estimativa da irradiância difusa foram analisados para avaliar seu desempenho no sertão da Paraíba, com objetivo de melhorar a acurácia das estimativas de produção de energia de projetos de centrais fotovoltaicas. Skartveit et al. (1998) foi o modelo que melhor descreveu a série de irradiância difusa observada nos três locais analisados. Um dos modelos para estimativa da irradiância difusa mais utilizados nos *softwares* comerciais é o de Erbs et al. (1982), que apresenta uma razão entre desvios média para os três locais de 0,87, mostrando uma diferença considerável na amplitude da série de irradiância modelada em relação a observada. O impacto de modelos que melhor descrevam a razão difusa devem ser levados em consideração nas simulações de empreendimentos fotovoltaicos, sobretudo as de geração centralizada, já que esses modelos estão diretamente ligados ao cálculo da irradiância no plano inclinado e, conseqüentemente, à potência de saída do módulo (Hofmann e Seckmeyer, 2017).

Este trabalho pode servir como uma ferramenta útil nas estimativas de produção de energia de centrais fotovoltaicas, trazendo maior acurácia à simulação quando propõe modelos, dentre os disponíveis na literatura, que melhor descrevam a radiação difusa no sertão da Paraíba. Um estudo do impacto da seleção de diferentes modelos de separação da irradiância difusa (em comparação com a radiação difusa medida) nos modelos de transposição da radiação para o plano inclinado seria um segundo passo para avaliar o real impacto dos modelos de difusa nas simulações de energia de centrais solares fotovoltaicas.

REFERÊNCIAS

- Bird, R. E. e Hulstrom, R. L. (1981). A Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces, TR-642-761, 33 pp., Solar Energy Research Institute, Golden, CO.
- Chandrasekaran, J. e Kumar, S. (1994). Hourly diffuse fraction correlation at a tropical location. *Solar Energy*, 53: 505–10.
- Collares-Pereira, M. e Rabl, A. (1979). The Average Distribution of Solar Radiation-Correlations between Diffuse and Hemispherical and between Daily and Hourly Insolation Values. *Solar Energy*, 22, 155-164.
- EPE [Empresa de Pesquisa Energética] Balanço Energético Nacional (BEN) 2018: Ano base 2017, 2018. Disponível em <<http://epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2018>> [consulta: 26 de agosto de 2019]

- Erbs, J.A.; Klein, S.A. e Duffie, J.A. (1982). Estimation of the diffuse radiation fraction for hourly, daily and monthly average global radiation. *Solar Energy*, 18, 293–302.
- Liu, B. e Jordan, R. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar Energy* 1960, 4, 1–19.
- Hofmann, M. e Seckmeyer, G. (2017). A New Model for Estimating the Diffuse Fraction of Solar Irradiance for Photovoltaic System Simulations. *Energies*, MDPI, Open Access Journal, vol. 10(2), pages 1-21.
- Maxwell, E. L. (1987). A Quasi-Physical Model for Converting Hourly Global Horizontal to Direct Normal Insolation, Technical Report No. SERI/TR-215-3087, Golden, CO: Solar Energy Research Institute.
- Orgill, J.F. e Hollands, K.G.T. (1976). Correlation equation for hourly diffuse radiation on a horizontal surface. *Solar Energy*, 19, 357–359.
- Petribú, L. B.; Sabino, E.; Barros, H.; Costa, A.; Barbosa, E. e Vilela, O.C. (2017). Procedimento objetivo para a garantia de qualidade de dados de radiação solar. Apresentado em XL Reunión de Trabajo de la Asociación Argentina de Energías Renovables y Ambiente (ASADES), San Juan, Argentina.
- Perez, R. R.; Ineichen, P.; Maxwell, E. L.; Seals, R. D. e Zelenka, A. (1992). Dynamic global-to-direct irradiance conversion surface. *ASHRAE Transactions* 1992:354–69.
- Reindl, D.T.; Beckman, W.A. e Duffie, J.A. (1990). Diffuse fraction correlations. *Solar Energy*, 45, 1–7.
- Ridley, B.; Boland, J. e Laurent, P. (2010). Modelling of diffuse solar fraction with multiple predictors. *Renew. Energy* 2010, 35, 478–483.
- Rigollier, C.; Bauer, O. e Wald, L. (2000). On the clear sky model of the ESRA - European Solar Radiation Atlas with respect to the Heliosat method. *Solar Energy*, v. 68, n. 1, pp.33-48.
- Ruiz-Arias, J. A.; Alsamamra, H.; Tovar-Pescador, J. e Pozo-Vázquez, D. (2010). Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. *Energy Conversion and Management*, v. 51, n. 5, p. 881-893.
- Skartveit, A. e Olseth, J. A. (1987). A model for the diffuse use in daily fraction of hourly global radiation. *Solar Energy* 38, 271–274.
- Skartveit, A.; Olseth, J.A. e Tuft, M.E. (1998). An hourly diffuse fraction model with correction for variability and surface albedo. *Solar Energy*, 63(3):173–83.
- Taylor, K.E. (2001). Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 106, 7183-7192.
- Younes, S.; Claywell, R. e Muneer, T. (2005). Quality control of solar radiation data: Present status and proposed new approaches. *Energy*, v. 30, n. 9, p. 1533-1549, 2005.

COMPARATIVE STUDY OF MODELS FOR ESTIMATING DIFFUSE RADIATION IN THE SEMIARID OF PARAÍBA, BRAZILIAN NORTHEAST

ABSTRACT: Separation models for estimating diffuse horizontal irradiance from global horizontal irradiance are important for the simulation of the energy produced by solar generation systems. In this study, seven diffuse models were applied to the semiarid of Paraíba, in the Brazilian northeast. The models were divided into those of one parameter and multi parameter, generating to each one the dispersion of diffuse fraction (k_d) in function of the clearness index (k_t). The performance of the models were evaluated by means of graphical analysis, using the Taylor diagram, and statistical like the skill score, root mean square error, mean bias error and standard deviation ratio. The multi parameter model of Skartveit presented the best results.

Keywords: diffuse horizontal irradiance, clearness index, diffuse fraction, multi parameter and Taylor diagram.