

ESTUDO COMPARATIVO ENTRE DADOS DE RADIAÇÃO SOLAR ESTIMADOS VIA SATÉLITE E MEDIDOS POR ESTAÇÕES METEOROLÓGICAS NO SERTÃO DA PARAÍBA

Baldoino Sonildo da Nóbrega – baldoino.nobrega@ifpb.edu.br

Raphael Henrique Falcão de Melo – raphael.melo@ifpb.edu.br

Maria Cassiana Pereira Gonçalves – maria.cassiana@academico.ifpb.edu.br

Instituto Federal de Educação Ciência e Tecnologia da Paraíba

Waleria Guerreiro Lima – waleria.lima@faculadeguararapes.edu.br

Centro Universitário dos Guararapes, Programa de Pós-Graduação em Inovação e Desenvolvimento

Resumo. A disponibilidade de energia pode ser determinada por diferentes metodologias, sendo as mais empregadas aquelas estimadas por satélites ou através do uso de dados de irradiação obtida de redes de estações de superfície. Diversas variáveis podem interferir na confiabilidade dos dados fornecidos pelos diferentes mecanismos de medição. Assim, este estudo teve como objetivo comparar os dados de Irradiação Global Horizontal de três estações meteorológicas automáticas localizadas no Sertão do Estado da Paraíba, com os dados disponíveis de estimativas de satélite do banco de dados do National Solar Radiation Database (NSRDB) da base de dados da NASA. Para realizar as comparações, foram utilizados métodos estatísticos como o Coeficiente de correlação de Pearson, o erro médio, a raiz do erro quadrático médio, e teste de hipótese, teste t, para avaliar as semelhanças entre os dados de cada localidade. Os resultados apresentaram boa correlação entre as bases de dados com coeficiente de Pearson de 0,90 para Cajazeiras, 0,88 para São Gonçalo e 0,91 para Patos. O Viés foi positivo em todas as comparações, o que demonstra uma superestimação dos dados do NSRDB. Com a aplicação do teste t, verificou-se que não existem evidências suficientes para rejeitar a hipótese que as médias são iguais. Observou-se, também, que os perfis das curvas diárias das estações do Instituto Nacional de Meteorologia (INMET) e NSRDB são qualitativamente similares, com apenas uma pequena dispersão nos horários de alta radiação solar. Já para a estação da Agência Executiva de Gestão das Águas (AESAs), as curvas possuem maior discrepância ao longo do dia.

Palavras-chave: Irradiação global horizontal, Métodos estatísticos, Estimativas

1. INTRODUÇÃO

A preocupação das nações com os aspectos climáticos futuros, a possibilidade de esgotamento das fontes de energias não renováveis e consequências ambientais causadas por estas são alguns fatores que despertaram a necessidade da humanidade para utilização das energias ditas renováveis, em substituição à matriz fóssil. Nos últimos anos, a energia proveniente do sol tem ganhado destaque dentre as demais fontes de energias renováveis e o aproveitamento dessa, tanto como fonte de calor quanto de luz, tem sido considerada uma das alternativas energéticas mais promissoras para prover a energia necessária ao desenvolvimento humano (Pinho e Galdino, 2014). No entanto, a variabilidade do recurso solar tem sido o fator de maior incerteza na previsão de performance e viabilidade financeira de um projeto de planta solar.

Identificar a taxa e o intervalo de tempo da radiação solar local são parâmetros iniciais e importantes para o desenvolvimento de projetos que visam a captação e a conversão da energia solar. Na opinião de Pinho et al. (2008), um sistema fotovoltaico, por exemplo, requer dados de medições em locais mais próximos possíveis daquele onde se pretende implantar o sistema. Com um histórico dessas medidas, pode-se viabilizar a instalação desses sistemas em uma determinada região, garantindo o máximo aproveitamento do recurso ao longo do ano.

As informações sobre radiação solar podem ser obtidas a partir de diferentes fontes de dados, tais como medições em solo por piranômetros ou células de referência ou derivada de satélite (Espinar, 2009). A medida direta através de radiômetros colocados na superfície é a forma mais simples para conhecer a disponibilidade do recurso solar que incide na superfície da Terra (Martins et al., 2004). Esses fornecem medidas pontuais com grande acurácia, entretanto, nem sempre se dispõe de séries históricas de dados pontuais coletados em superfície para todas as localidades desejadas, uma vez que esses instrumentos de coleta possuem elevado custo, principalmente para instalação em grandes extensões territoriais (Martins et al., 2004), e quando o interstício de tempo é longo, tende a ocorrer perda de informações devido a manutenção, calibração e falhas na alimentação e controle de qualidade.

Uma alternativa a falta de dados e séries históricas de radiação solar para uma determinada localidade é a obtenção de estimativas a partir da utilização de modelos computacionais simples que utilizam dados atmosféricos obtidos por satélite. Para Martins et al. (2004) a utilização de modelos de satélite para estimar o potencial de utilização da energia solar é uma excelente alternativa para exploração dessa fonte limpa e renovável de energia, principalmente em países de grande extensão territorial. As vantagens dos dados obtidos por satélite são a resolução espacial, dados de longo prazo,

baixos custos e mínimo de falhas (Mitrushi et al., 2016). Medidas coletadas por piranômetros e satélites diferenciam-se principalmente na área de amostragem, sendo que os piranômetros são amostradores pontuais e os satélites são amostradores de grandes áreas (Botan e Correa, 2010).

Vários estudos na literatura (Pereira et al., 2006; Pereira et al., 2017; Marciotto et al., 2008; Botan e Correa, 2010; Boyo e Adeyemi, 2011; Mitrushi et al., 2016; Lima et al., 2016) têm realizados comparações entre estimativas de dados modelados através de satélites hospedados em bases com valores medidos por sensores, seja para avaliar apenas as semelhanças entre esses dados ou para validar algum modelo. Os estudos em sua maioria se ocupam em investigar a qualidade e a credibilidade entre as bases de dados.

O “Atlas Brasileiro de Energia Solar” (Pereira et al., 2006 e Pereira et al., 2017) utiliza valores de irradiação obtidos com o modelo BRASIL-SR, os quais são validados por meio de comparação com valores medidos em superfície em estações distribuídas pelo território brasileiro. Os autores fazem uso de algumas métricas estatísticas como o coeficiente de Correlação de Pearson, análise de Viés e da Raiz do Erro Quadrático Médio.

Marciotto et al. (2008) realizaram uma comparação entre medidas de insolação obtidas na Plataforma Micro meteorológica do Instituto de Astronomia, Geofísica e Ciência Atmosféricas (PM-IAG/USP) e as estimativas do satélite GOES com a finalidade de avaliar a representatividade das medidas isoladas da plataforma micro meteorológica (PM-IAG) para Região Metropolitana de São Paulo (RMSP). O período utilizado foi o mês de junho de 2004. Os resultados são boas concordâncias entre os valores observados em superfície e os estimados por satélite, indicando que as observações na PM-IAG podem ser utilizadas para representar o campo de insolação na RMSP.

Botan e Correa (2010) apresentaram a comparação entre dados de radiação solar inferidos do satélite GOES com informações de superfície fornecidas por uma plataforma coletora de dados (PCD) instalada no campus da Universidade Federal de Itajubá. Fizeram uma análise do ajuste das informações remotas às verdades terrestres, de modo a validar mapas regionais da disponibilidade de radiação solar em superfície. Os resultados foram significativos e possibilitaram a geração dos primeiros mapas de disponibilidade mensal de radiação solar na região.

Já Boyo e Adeyemi (2011) compararam dados de radiação solar estimados por satélites disponíveis em NSRDB com os medidos em terra por Piranômetro obtidos pela Agência Meteorológica da Nigéria, para cinco localidades. Os autores propuseram o uso do teste de hipótese Komolgorov Smirnov, o qual analisa a similaridade entre os conjuntos de dados e não exige qualquer pressuposto sobre a distribuição de dados.

Mitrushi et al. (2016) realizaram comparação de dados de irradiação global horizontal fornecidos pelo NSRDB, administrado pela NASA (*National Aeronautics and Space Administration*) com dados terrestres disponíveis para 6 cidade cobertas pela rede meteorológica da Albânia, para um período de 7 anos. Os autores fizeram uso do teste de hipótese, teste t para verificar se há diferenças entre médias e obtiveram que apenas a cidade de Tirana difere estatisticamente dos dados fornecidos pelo NSRDB. Concluíram também, a partir do erro médio (Viés), que os dados das estações meteorológicas são menores que as estimativas fornecidas pela base de dados NSRDB.

Um outro estudo realizado por Lima et al. (2016) comparou dados de irradiação global horizontal (GHI) dessa mesma base de dados (NSRDB) com dados coletados por estações meteorológicas automatizadas (EMA) instaladas em cinco cidades do estado de Sergipe e pertencentes ao Instituto Nacional de Meteorologia – INMET, para o período dos meses de outubro a dezembro de 2016. Os autores utilizaram os índices estatísticos: erro médio relativo e a raiz do erro quadrático médio relativo para realizar as análises. Concluem que os dados de GHI do NSRDB são superestimados em relação aos dados medidos pelas estações do INMET para a maioria das localidades analisadas.

Com relação aos trabalhos acima citados, observa-se que a comparação entre bases de dados de radiação solar é feita rotineiramente em diversas localidades, isso ocorre porque há uma necessidade de se buscar dados de irradiação confiáveis, assim como, obter avaliações apropriadas para reduzir o risco na escolha do local de instalação de sistemas que dependam do recurso solar (Sengupta et al., 2015). Dados precisos de radiação solar ajudam a garantir melhores inferências nos resultados de viabilidade financeiras e diminuem os erros nas estimativas propostas para geração de energia, por exemplo, de sistemas fotovoltaicos ou coletores solares para aquecimento. Dessa forma, é relevante para uma determinada região, quantificar as semelhanças ou as diferenças entre as duas séries de dados disponíveis.

Neste sentido, este estudo propõe comparações entre valores de GHI obtidos por Estações Meteorológicas Automáticas disponibilizados pelo INMET e pela Agência Executiva de Gestão das Águas (AESAs) com os dados obtidos via satélite pelo NREL (National Renewable Energy Laboratory) e disponibilizados no banco de dados NSRDB para três municípios do Sertão Paraibano, Cajazeiras, São Gonçalo e Patos, com o objetivo de avaliar as semelhanças ou diferenças estatísticas entre essas bases de dados.

É importante ressaltar que a Região do Sertão da Paraíba está recebendo grandes investimentos em energia solar, por exemplo, a construção do Complexo de geração de energia solar fotovoltaica de Coremas-PB, o qual contará com três usinas denominadas Coremas I, Coremas II e Coremas III, com 93 MWp (Megawatts-pico) em capacidade de geração, com investimentos de até R\$ 426 milhões, sendo que Coremas I e II já estão em fase de construção, cada uma com 31,12 MWp e com entrega prevista para o final de 2017. Já Coremas III está com previsão de entrega apenas para o final de 2018. Coremas I foi vencedor do 1º Leilão de energia solar realizado no Brasil em 2014, Coremas II no 1º leilão de 2015 e Coremas III no 2º Leilão de 2015. Além das usinas de Coremas, há também já outorgados os projetos das usinas de Angico I e Malta com capacidade 27 MWp cada, situados no município de Malta pertencente também ao Sertão Paraibano.

2. MATERIAL E MÉTODO

2.1. Definição da área de estudo

Este estudo foi realizado no Sertão da Paraíba, região Nordeste do Brasil, com dados de Irradiação Global Horizontal (GHI) nos municípios de Cajazeiras, São Gonçalo e Patos (Tab. 1). Tais localidades foram as escolhidas por possuírem estações meteorológicas automáticas de dados do INMET (Patos e São Gonçalo) e da AESA (Cajazeiras) (Fig. 1).

Tabela 1: Descrição das Cidades

Região	Cidade	Latitude	Longitude	Altitude
Sertão da Paraíba	Cajazeiras	-6,890153	-38,544097	295 m
	São Gonçalo	-6,8357777	-38,311583	237 m
	Patos	-7,079836	-37,272817	264 m



Figura 1: Localização espacial das estações meteorológicas automáticas de dados do estudo

Fonte: Adaptado do INMET

2.2. Coleta dos dados e período de estudo

Para realizar a coleta de informações das estações meteorológicas automáticas do solo, foram solicitados os dados de Irradiação global horizontal (GHI) ao INMET e a AESA. A coleta dos dados de estimativas de satélite para os municípios em estudo foram obtidas no site <https://maps.nrel.gov/nsrdb> disponibilizados pelo *National Renewable Energy Laboratory-NREL* no banco de dados NSRDB. Uma análise preliminar foi realizada para 3 anos (2013 a 2015) das informações de Irradiação global horizontal por hora para as duas base de dados (informações terrestre e valores estimados). Nesta análise preliminar o ano de 2014 foi o que apresentou o menor número de inconsistências de medição nas estações automáticas. Portanto, o ano de 2014 foi o escolhido para realizar as análise de diferenças entre as bases de dados de estimativas de satélites e medidas observadas no solo nos municípios em estudo.

2.3. Tratamento de dados

As informações de todas as bases de dados referentes ao ano de 2014 foram submetidas a três procedimentos: (i) análise de valores quando fisicamente impossível; (ii) análise de valores quando o evento é extremamente raro; (iii) se uma das bases de dados contivesse em uma linha erros ou falta de dados era excluído essa linha das duas bases de dados para se manter a convergência entre dias e horas das matrizes. Após o procedimento de correção de dados foram calculadas as médias para cada dia do ano e as médias horárias anual para proceder a comparação das curvas geradas pelas informações terrestres e as estimativas do banco de dados NSRDB. Após o tratamento, a quantidade de dias com dados aceitáveis para os municípios de Cajazeiras, São Gonçalo e Patos foram, 304, 353, 363 respectivamente.

2.4. Análise de dados

Na literatura vários métodos estatísticos são utilizados para comparar e avaliar os modelos de estimativas de radiação solar (Sengupta et al., 2015). Para verificar as diferenças entre as bases de dados do NSRDB e das estações meteorológicas da AESA e do INMET foram utilizados os seguintes métodos: o Erro Médio ou Viés (BIAS em Inglês), o Erro Absoluto Médio (MAE em Inglês), a Raiz do erro quadrático médio (RMSE em Inglês) e o coeficiente de correlação de Pearson (r). Essas medidas foram calculadas a partir da média de GHI para cada dia do ano em cada município em estudo. De

forma complementar realizou-se um teste de hipótese de diferenças entre média (teste t) para verificar se há diferenças entre as médias dos conjuntos de dados durante o ano de 2014. Em termos de compreensão dos métodos estatísticos utilizados neste trabalho temos que o Viés Eq. (1) mede a tendência de o modelo superestimar ou subestimar os valores modelados em relação aos observados, também chamado de erro sistemático, e se apresentar valor positivo, indica que os valores estimados excedem os valores medidos (Segunpta et al., 2015). O Erro Absoluto Médio Eq. (2) é menos afetado por pontos extremos que o Viés e considerado preciso e robusto como medida da habilidade de modelos numéricos em reproduzir a realidade. A Raiz do erro quadrático médio Eq.(3) é comumente usada para expressar a acurácia dos resultados numéricos com a vantagem de apresentar valores do erro nas mesmas dimensões da variável analisada. Quanto menor o valor dos erros, definidos anteriormente, melhor será a correlação entre as duas bases de dados. O coeficiente de Pearson (r) mede o grau de intensidade da correlação entre duas variáveis, e ainda, o sentido dessa correlação (positivo ou negativo) e quanto mais próximo de 1 e -1 maior será a correlação. As equações (1), (2), (3) e (4) representam o Erro médio (Viés), o Erro absoluto médio (MAE), a raiz do erro quadrático médio (RMSE) e o coeficiente de Pearson (r), respectivamente.

$$VIÉS = \frac{1}{n} \sum_{i=1}^n (e_i - o_i) \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i - o_i| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i - o_i)^2} \quad (3)$$

$$r = \frac{n \sum_{i=1}^n e_i o_i - (\sum e_i)(\sum o_i)}{\sqrt{n \sum o_i^2 - (\sum o_i)^2} \cdot \sqrt{n \sum e_i^2 - (\sum e_i)^2}} \quad (4)$$

Onde e_i representa os valores estimados, o_i representa os valores observados e n é a quantidade de amostras.

3. RESULTADOS

Na análise exploratória dos dados sobre a média diária para o ano de 2014 mostrada em Tab. 2, observa-se que a média máxima diária de GHI ocorreu na Cidade de Cajazeiras (618,5 W/m²) seguida por Patos (616,9 W/m²) e São Gonçalo (612,3 W/m²) para o conjunto de dados NSRDB. Já para o conjunto de dados terrestre a média máxima ocorreu em Patos (611,2), seguido por São Gonçalo (604,5) e Cajazeiras (574,6). Em relação a mínima, os dados terrestres de Patos e São Gonçalo, onde as estações são do INMET, nota-se que os sensores pontuais no solo são mais sensíveis as baixas irradiações que os valores estimados, enquanto que Cajazeiras tem praticamente a mesma mínima entre os dois conjuntos de dados. Ao comparar a média diária observa-se uma divergência maior entre as bases de dados de Cajazeiras (diferença de 61,2 W/m²). Para as outras localidades os valores são bem próximos, mas com os dados do NSRDB acima das informações terrestres. O desvio padrão das informações de GHI dentro de cada conjunto de dados varia aproximadamente entre 80 W/m² e 93 W/m² para cada localidade em estudo. Os coeficientes de variação, demonstram que a média possui uma moderada representatividade em cada conjunto, geralmente, varia entre 0,17 e 0,21 o que também explica uma baixa variação de GHI durante o ano.

Tabela 2 – Análise descritiva dos valores relativos a média diária de GHI (W/m²) para o ano de 2014

Cidades	Cajazeiras		São Gonçalo		Patos	
	NSRDB	AESA	NSRDB	INMET	NSRDB	INMET
Métricas / Base de dados						
Máxima	618,5	574,6	612,3	604,5	616,9	611,2
Mínima	189,4	184,7	151,4	169,4	151,2	198,3
Média	477	415,8	481,27	473,2	490,9	473,8
Desvio Padrão	93,24	86,79	89,14	80,44	90,61	87,85
Coeficiente de Variação	0,19	0,21	0,18	0,17	0,18	0,18
Quartil 1	424,8	361,1	429,2	423,6	439,7	422,7
Quartil 2	484,7	420,4	487,5	484,9	504,2	491,3
Quartil 3	552,4	482,9	552,4	534,5	566,2	542,3

Ao realizar um filtro em cada conjunto de dados sobre a proporção de dias que atinge os valores de GHI entre 1000 W/m² e 1105 W/m², obteve-se os dados demonstrados em Tab. 3. Há uma divergência acentuada entre as bases de dados de Cajazeiras. Enquanto NSRDB estima 33,5% dos dias com GHI atingindo em um determinado momento 1000 W/m², a estação da AESA só encontra 2%, o que pode demonstrar uma possível descalibração no sensor ou até falha humana nos registros, visto que a 32 km deste local, a estação do INMET em São Gonçalo encontra uma proporção de 27,3% dos dias para o mesmo período. A Região de Patos é a que possui maior proporção para o mesmo período nas duas bases de dados (45,6% - NSRDB e 34,5% - INMET). Ao analisar qual o valor máximo possível de GHI registrado para o período de 2014, encontra-se 1104 W/m² para Estação do INMET em São Gonçalo e 1080 W/m² para NSRDB Cajazeiras.

Tabela 3 – Proporção de dias com GHI entre 1000 W/m² e 1105 W/m² para o período 2014 e valor máximo registrado para um determinado momento do dia

Localidade	Base de dados	% Dias (1000 < GHI < 1105)	Valor máximo registrado
Cajazeiras-PB	NSRDB	33,5%	1080 W/m ²
	AESA	2,0%	1071 W/m ²
São Gonçalo-PB	NSRDB	37,4%	1057 W/m ²
	INMET	27,3%	1104 W/m ²
Patos-PB	NSRDB	45,6%	1058 W/m ²
	INMET	34,5%	1040 W/m ²

Ainda sobre as comparações e correlações entre as bases de dados, a Fig. 2 apresenta os gráficos de dispersão entre a média diária para os valores estimados no NSRDB e os valores observados para estações do INMET e AESA. Foram correlacionados 304 valores de cada banco de dados para Cajazeiras, 353 para São Gonçalo e 363 para Patos. Pode-se verificar uma boa correlação entre os dados estimados e observados, com um coeficiente de Pearson (r) entre 0,88 e 0,91 para as localidades em estudo. Valores estes, próximos aos resultados dos trabalhos de Baton e Correa (2010) e Pereira et al. (2017). Ao analisar os valores do RSME na Tab. 4, nota-se baixos valores para as estações do INMET em São Gonçalo (8,9%) e Patos (8,6%), inclusive convergindo com valores de referência para o Brasil citados por Segunpta et al. (2015) que indica valores próximo ou inferiores a 13%. Vale salientar que esses valores também estão bem próximos dos valores de RSME do Atlas Brasileiro de Energia Solar 2017 (2ª edição), que indica para região Nordeste um valor de 8,3%. Para Cajazeira, mesmo com coeficiente de correlação da ordem de $r = 0,90$, deve-se atentar para um RMSE (16,1%) considerado elevado, se comparado os valores de referência do atlas e de Sengupta et al. (2015), fato que pode demonstrar uma maior dispersão dos pontos em torno da diagonal central que é a melhor estimativa.

Tabela 4 - Valores de comparação das médias diárias de irradiação global horizontal para cada município em estudo.

Cidade/Métricas	r	Viés (W/m ²)	Viés%	RSME (W/m ²)	RSME (%)	MAE (W/m ²)
Cajazeiras-PB	0,90	61,2	14,7%	66,9	16,1%	63,7
São Gonçalo-PB	0,88	8,060	1,7%	42,120	8,9%	31,590
Patos-PB	0,91	17,10	6,26%	40,61	8,6%	29,66

A Fig. 2 apresenta também os histogramas dos Viés para cada localidade e mostra que há variação em maiores proporções entre 0 W/m² e 40 W/m² para São Gonçalo (65%) e Patos (68%) indicando um Viés positivo, isto é, uma tendência dos dados NSRDB superestimar a GHI, fato encontrado também no estudo de Mitrusi et al. (2016). Porém, quando comparados os valores em porcentagem (Tab. 4) dessas localidades com os de referência (Valor 8%) Sengupta et al. (2015) observa-se que os valores estão dentro de um limite aceitável para essa região. Já em Cajazeiras, a variação em maiores proporções acontece entre 20 W/m² e 100 W/m² (60%) o que mostra um Viés um pouco elevado e positivo em relação as outras localidades, o que evidencia-se desvios maiores entre as bases de dados.

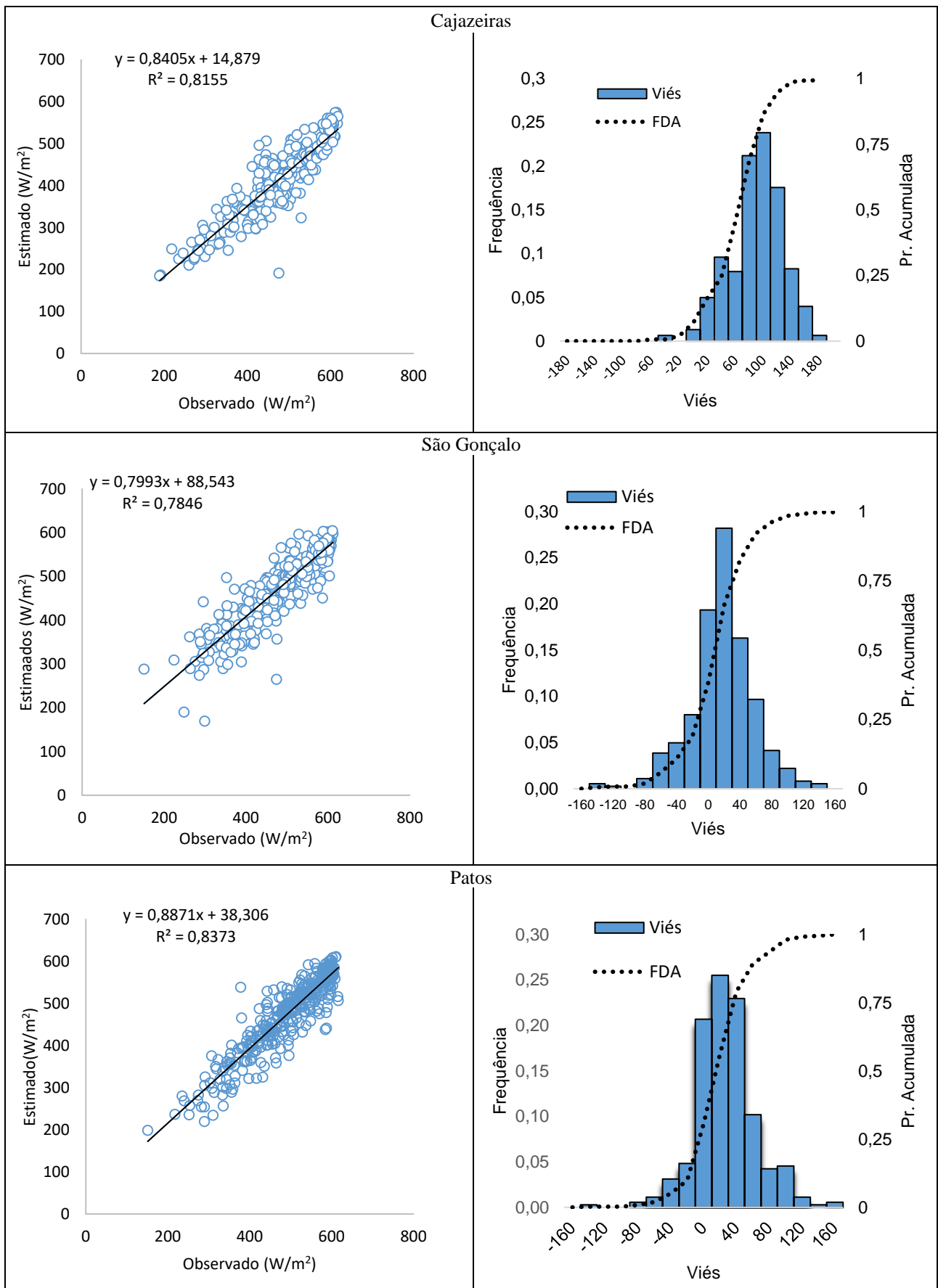


Figura 2: Diagrama de dispersão entre valores estimados e observados para as médias diárias de GHI no período de 2014 e histogramas do Viés para o mesmo período no Sertão Paraibano.

Para uma análise horária da GHI, resolveu-se calcular a média para cada hora do dia no período 2014. Dessa forma, ao analisar a Fig. 3, foi possível compreender o comportamento das curvas dos conjuntos de dados durante o dia e em seguida aplicar um teste de hipóteses (teste t) para averiguar se existiam diferenças estatisticamente significativas entre as médias das bases de dados para cada localidade. Antes de realizar o teste de hipótese, efetuou-se, primeiro, a análise descritiva dos dados entre às 06h:00 e 18h:00 para NSRDB e as estações do INMET e AESA, as quais mostraram que os maiores valores de GHI ocorrem entre 11 horas da manhã até as 13 horas da tarde em todas as localidades, sendo que a média de GHI nesse horário em Cajazeiras varia entre 773,3 W/m² - AESA e 833,3W/m²-NSRDB, São Gonçalo 795,16 W/m²-INMET e 842,53W/m²-NSRDB e Patos 833,3 W/m²-INMET e 866,4 W/m²-NSRDB.

É perceptível na Fig. 3 que, geralmente, as diferenças entre as curvas do NSRDB e as estações do INMET ocorrem nos horários de pico de radiação solar. Nos demais horários há uma semelhança entre as curvas. Para Cajazeiras (estação AESA), essa situação não ocorre. Nesse local, durante todo o dia as curvas horárias sempre mantêm uma diferença considerável, sendo mais elevado nos horários de maior radiação. A curva para os dados do NSRDB geralmente é superior à curva para dados terrestres o que converge com os resultados de Lima et al. (2016) e Mitrusi et al. (2016).

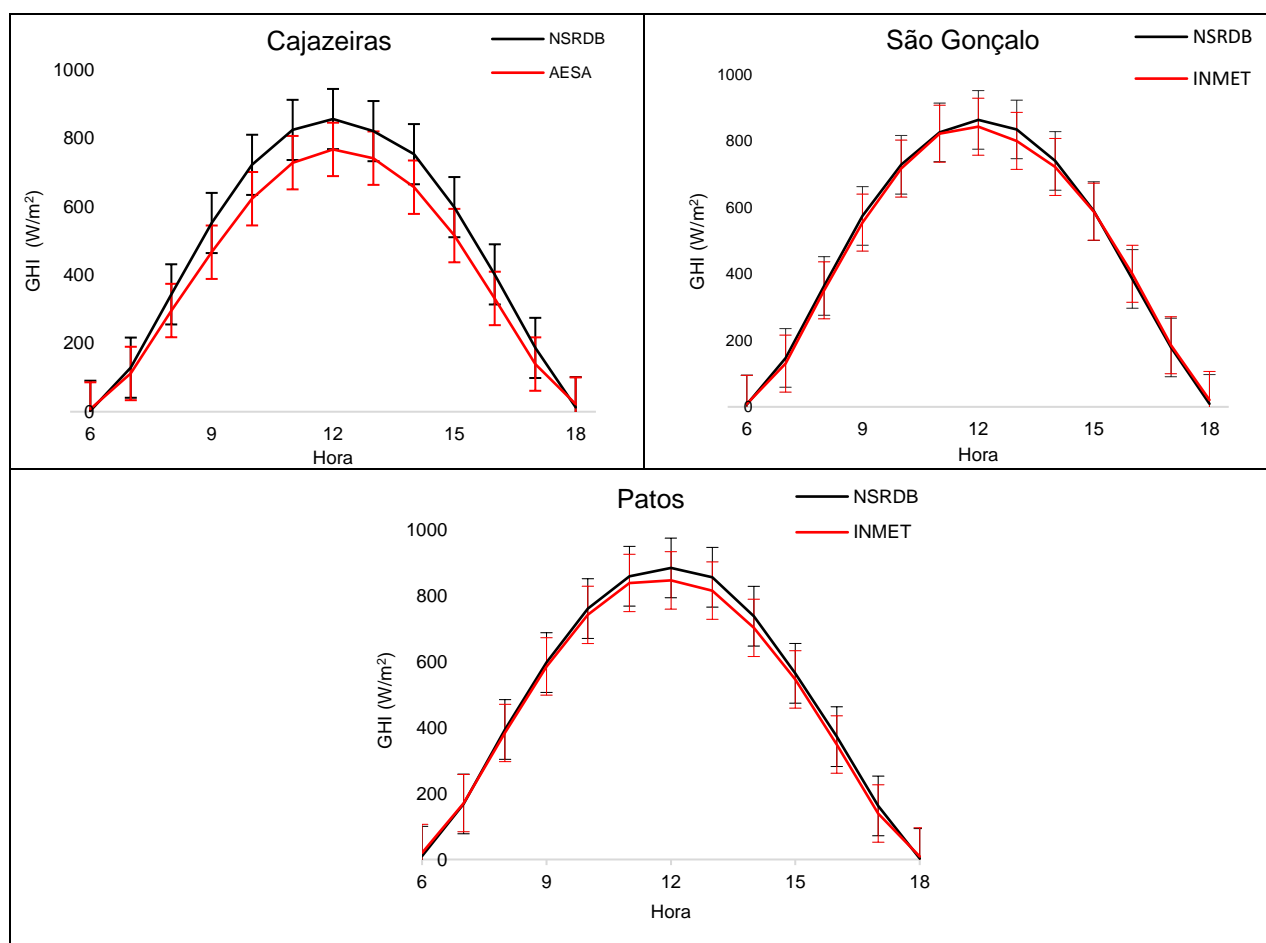


Figura 3: Médias e desvios horários do período de 2014 para Cajazeiras, São Gonçalo e Patos, no Sertão da Paraíba.

A máxima dentre todos os dados, como já mencionado (1104 W/m²), ocorreu às 13 horas do dia 27 de janeiro de 2014 e foi registrada pela estação do INMET em São Gonçalo e no mesmo dia o NSRDB registrou apenas 1064 W/m² para este local. Fato de destaque são os dados horários de nascer e pôr do sol, os quais são bastante inconsistentes e sujeitos a grandes flutuações, detectadas também no estudo de Marcioto et al. (2008). Mesmo assim manteve-se esses horários por não influenciarem diretamente nos resultados, uma vez que os valores são pequenos.

Para aplicar o teste de hipótese (teste t), que assume como hipótese nula (H_0): igualdade entre médias e hipótese alternativa H_1 : diferença entre médias, por ser um teste paramétrico, é necessário primeiro verificar os aspectos de normalidade para cada conjunto de dados. O teste de hipóteses aplicado para verificação da normalidade foi o Shapiro Wilk que gerou $p_valor > 0,05$ para todos os conjuntos de dados, isto é, os dados se comportam de forma a seguir um modelo de uma distribuição normal. A Tab. 5 apresenta os resultados do teste t e é possível identificar que as médias não se diferem estatisticamente em nenhuma localidade. Portanto, não há evidência para considerar que as bases de dados de estimativas de Satélite e observados no solo se diferem estatisticamente entre si. Porém, é intrigante neste estudo o caso particular da diferença existente entre os dados de Cajazeiras 61,2 W/m². Entende-se que supondo uma planta solar fotovoltaica de médio ou grande porte, essa diferença poderia causar um erro de acurácia nos resultados de geração de energia a depender de qual base de dados tivesse sido usada.

Tabela 5 – Teste t para igualdade de médias de irradiação global horizontal obtidas para Cajazeiras, São Gonçalo e Patos, no Sertão da Paraíba.

Cidade	Base de dados	Diferenças médias	Significância
Cajazeiras	NSRDB	61,2	0,608
	AESA-CZ		
São Gonçalo-PB	NSRDB	8,1	0,948
	INMET-SG		
Patos-PB	NSRDB	17,1	0,893
	INMET-PT		

*Significância de 5%

4. CONSIDERAÇÕES FINAIS

Neste estudo, foi possível comparar os dados de estimativa de GHI fornecidos pela NASA disponíveis no NSRDB com dados de medidas feitas por estação meteorológica terrestre no Sertão da Paraíba, Região Nordeste do Brasil. As avaliações estatísticas mostraram boa correlação entre as bases de dados, com Viés positivo em todas as localidades, indicando que as estimativas do NSRDB são superestimadas, no entanto, para as estações do INMET, o Viés e a raiz do erro quadrático médio são aceitáveis, pois converge com resultados da literatura, dessa forma, podendo considerar que as medidas são representativas.

Para estação da AESA, em Cajazeiras, as diferenças são maiores e perceptíveis nas curvas diárias (Fig. 3) e nas métricas estatísticas. Os valores nesta estação automática, em sua maioria, são menores que os valores obtidos nas estações do INMET. Como Cajazeiras está a apenas 32 km de São Gonçalo, a possibilidade de se obter valores próximo são elevados, visto que os locais estão dentro da mesma região semiárida da Paraíba. Neste sentido, pode-se levantar algumas hipóteses sobre as diferenças existentes na estação da AESA, como por exemplo, problemas ou limitações dos sensores, falhas na alimentação e controle de qualidade dos dados, dentre outras.

Os resultados do teste t demonstraram que não há evidências para rejeitar a hipótese de igualdade das médias entre os conjuntos de dados dentre as três localidades pesquisadas, mas deve-se considerar que há uma diferença que desperta maior atenção na estação da AESA. Essa diferença ocorrida em Cajazeiras será informada a AESA, para que, se possível, possa verificar se existiu algum tipo de erro com os dados desse período. Uma avaliação mais aprofundada, com uma série histórica de GHI mais longa e identificação dos tipos de sensores utilizados em cada estação pode ser indicada como perspectivas para trabalhos futuros no intuito de dirimir as diferenças existentes entre Cajazeiras e São Gonçalo.

Agradecimentos

Os autores agradecem ao IFPB pelo auxílio financeiro em forma de bolsa para um estudante e taxa de bancada para o desenvolvimento do projeto de pesquisa via edital Interconecta 2017. A Agência Executiva de Gestão das Águas e ao Instituto Nacional de Meteorologia por terem disponibilizados os dados da série histórica de radiação solar.

REFERÊNCIAS

- Boyo, A. O.; Adeyemi, K.A, 2012. Analysis of Solar Radiation data From Satellite and Nigeria Meteorological Station, Journal of Renewable Energy Research (IJRER), v. 1, n. 4, p. 314-322.
- Botan J. S., Correa M. P. 2010. Elementos de uma climatologia da radiação solar para a região de Itajubá, MG. XVI Congresso Brasileiro de Meteorologia. Belém –PA.
- Espinar, B. et al., 2009. Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations. Solar Energy. Vol. 83, [s.i.]. pp. 118–125.
- Martins, F. R., Pereira, E B, Echer, M. P. S, 2004. Levantamento dos recursos de energia solar no Brasil com o emprego de satélite geoestacionário – o Projeto Swera. Revista Brasileira de Ensino de Física, v. 26, n. 2, p. 145-159.
- Marciotto, E.R.; Oliveira, A.P., Escobedo, J. F., 2008. Comparação entre a Radiação Solar Medida na Superfície e Estimada por Satélite na Região Metropolitana de São Paulo. II Congresso Brasileiro de Energia Solar e III Conferência Regional Latino-Americana da ISES – Florianópolis-SC.
- Mitrushi D. et al., 2016. A Comparative Study of Satellite Estimation for Solar Insolation in Albania with Ground Measurements. 9TH International Physics Conference of the Balkan.
- NREL - National Renewable Energy Laboratory, http://redc.nrel.gov/solar/glossary/gloss_d.html.
- Pinho, J. T.; et al., 2008. Sistemas híbridos - Soluções energéticas para a Amazônia. 1. ed. Brasília, Brasil: Ministério de Minas e Energia.

- Pinho, J. T.; Galdino, M. A., 2014. CRESESB - Centro de Referência para Energia Solar e Eólica Sérgio de Salvo Brito. Grupo de Trabalho de Energia Solar - GTES. Manual de engenharia para sistemas fotovoltaicos. Rio de Janeiro: CEPTEL.
- Pereira, E. B. et al., 1996. "Survey of the incident solar radiation in Brazil by use of meteosat satellite data". Solar Energy, Vol. 57, No. 2, p.125-132.
- Pereira, E. B. et al., 2006, "Atlas Brasileiro de Energia Solar". São José dos Campos: INPE, Brasil.
- Pereira, E. B.; et al., 2017, "Atlas Brasileiro de Energia Solar". 2 ed. São José dos Campos: INPE. 88 p.
- Sengupta, M. et al., 2015, "Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications - Technical Report NREL/TP-5D00-63112", National Renewable Energy Laboratory, Golden-CO, USA.
- Lima, J. C. O. et al., 2016. Comparação entre dados de radiação solar medidas por estações meteorológicas automáticas no estado de Sergipe com valores estimados a partir de imagens de satélite. Simpósio Sergipano de Energia Solar. Aracaju-SE.

COMPARATIVE STUDY OF SOLAR RADIATION DATA ESTIMATED BY SATELLITE AND MEASURED BY METEOROLOGICAL STATIONS IN PARAIBA SEMI ARID

Abstract. Energy availability can be determined by different methodologies, being the most used those estimated by satellites or by the use of irradiation data obtained from networks of surface stations. Several variables may interfere with the reliability of the data provided by the different measuring mechanisms. Thus, this study had the objective of comparing the Global Horizontal Irradiation data from three automatic meteorological stations located in semiarid of Paraíba State, with available satellite estimation data from the National Solar Radiation Database (NSRDB) data from NASA. To perform the comparisons, it was used statistical methods such as Pearson's correlation coefficient, the mean error, root mean square error and hypothesis test, *t* test, to evaluate the similarities between the data of each locality. The results demonstrated a good correlation between the data bases with Pearson coefficient of 0.90 for Cajazeiras city, 0.88 for São Gonçalo city and 0.91 for Patos city. The Bias of the data was positive in all comparisons, which demonstrates an overestimation of the NSRDB data. The *t*-test results showed that there is insufficient evidence to reject the hypothesis that the means are equal. It was also observed that the profiles of the daily curves of the National Institute of Meteorology (INMET) and NSRDB stations are qualitatively similar with only a dispersion during high solar radiation times. For the station of the Executive Agency of Water Management (AESAs), the curves have a greater discrepancy throughout the day.

Keywords: Global Horizontal Irradiation, Statistical methods, Estimates