

ANÁLISE ESTATÍSTICA DE BANCOS DE DADOS DE RADIAÇÃO SOLAR E TEMPERATURA PARA APLICAÇÃO EM SISTEMA FOTOVOLTAICO

Lucas Ramos Tavares (UNICAMP) - lucasrtavares92@gmail.com

Tatiane Silva Costa (UNICAMP) - tati_costa@rocketmail.com

Karen Barbosa de Melo (UNICAMP) - karen.bmelo@gmail.com

Marcelo Gradella Villalva (Instituição - a informar) - villalva@g.unicamp.br

Resumo:

A energia solar fotovoltaica tem sido explorada como uma solução para atender a crescente demanda por eletricidade através de uma fonte renovável de energia. Simulações otimizadas de sistemas fotovoltaicos requerem a utilização de dados solarimétricos e climáticos acurados. Este artigo tem como objetivo comparar a diferença na simulação de sistemas fotovoltaicos utilizando os dados de radiação solar e temperatura fornecidos pelas bases de dados Nasa SSE e Meteonorm, com os dados medidos por uma estação solarimétrica real instalada em Campinas. As análises foram realizadas por meio de métodos estatísticas e simulações de um sistema fotovoltaico para comparação de energia gerada, utilizando cada uma das bases de dados estudadas. Os resultados comparativos da radiação apresentam valores próximos entre Nasa SSE e Meteonorm, com uma pequena vantagem dos dados da Nasa onde as medidas de dispersão e os quartis são mais próximo de zero, assim como a média. Porém, as diferenças são mais pontuais ocorrem na análise de temperatura, em relação ao desvio padrão e o rMBE, assim como os erros máximos e mínimos. O trabalho conclui que os dados da Nasa SSE é o que se aproxima mais dos dados reais, apesar de apresentar diferença de geração de energia de aproximadamente 12 kWh/ano.

Palavras-chave: *Energia solar fotovoltaica, Dados solarimétricos, Nasa SSE , Meteonorm, Métodos estatísticos*

Área temática: *Radiação Solar*

Subárea temática: *Recursos Solares e Meteorologia da Radiação Solar*

ANÁLISE ESTATÍSTICA DE BANCOS DE DADOS DE RADIAÇÃO SOLAR E TEMPERATURA PARA APLICAÇÃO EM SISTEMA FOTOVOLTAICO

Lucas R. Tavares – lucasrtavares92@gmail.com

Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica

Tatiane S. Costa – tati_costa@rocketmail.com

Karen B. Melo – karen.bmelo@gmail.com

Marcelo G. Villalva – villalva@g.unicamp.br

Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação

Área: 1. Radiação solar – 1.1. Recursos solares e meteorologia da radiação solar

Resumo. A energia solar fotovoltaica tem sido explorada como uma solução para atender a crescente demanda por eletricidade através de uma fonte renovável de energia. Simulações otimizadas de sistemas fotovoltaicos requerem a utilização de dados solarimétricos e climáticos acurados. Este artigo tem como objetivo comparar a diferença na simulação de sistemas fotovoltaicos utilizando os dados de radiação solar e temperatura fornecidos pelas bases de dados Nasa SSE e Meteonorm, com os dados medidos por uma estação solarimétrica real instalada em Campinas. As análises foram realizadas por meio de métodos estatísticos e simulações de um sistema fotovoltaico para comparação de energia gerada, utilizando cada uma das bases de dados estudadas. Os resultados comparativos da radiação apresentam valores próximos entre Nasa SSE e Meteonorm, com uma pequena vantagem dos dados da Nasa onde as medidas de dispersão e os quartis são mais próximo de zero, assim como a média. Porém, as diferenças são mais pontuais ocorrem na análise de temperatura, em relação ao desvio padrão e o rMBE, assim como os erros máximos e mínimos. O trabalho conclui que os dados da Nasa SSE é o que se aproxima mais dos dados reais, apesar de apresentar diferença de geração de energia de aproximadamente 12 kWh/ano.

Palavras-chave: Energia solar fotovoltaica, Dados solarimétricos, Nasa SSE, Meteonorm, Métodos estatísticos.

1. INTRODUÇÃO

O consumo de energia elétrica se propaga com maior intensidade, principalmente desde a Revolução Industrial. A modernização demanda um fornecimento de energia contínuo, com garantias de qualidade e eficiência dos recursos utilizados. O desenvolvimento das sociedades, observado por parâmetros socioeconômicos, apresenta uma relação intrínseca ao crescimento da demanda de energia, trazendo à tona discussões sobre os impactos em diversas áreas do conhecimento, desde a economia até os efeitos ao meio ambiente (Raimundo, 2008). Para suprir toda a necessidade energética das populações, aumenta-se, principalmente, a exploração de combustíveis fósseis. Porém, devido às preocupações ambientais globais, o interesse em fontes renováveis de energia tem se expandido. Uma das fontes de renováveis de energia mais promissoras é a energia solar fotovoltaica, denominada dessa forma por causa do efeito fotovoltaico que consiste na geração de energia através de fótons de luz que penetram materiais semicondutores provocando uma corrente elétrica.

O conhecimento dos níveis de irradiância solar que atingem a superfície terrestre é um pré-requisito para várias aplicações da energia solar na indústria, na agricultura e em outras atividades relevantes sob o ponto de vista socioeconômico. Para se ter esse conhecimento é necessário mapear o comportamento dos dados disponíveis na superfície terrestre. Esse procedimento em todo território nacional é inviável, pois exige incentivos tecnológicos que nem sempre são atendidos, além da mão de obra especializada para instalação e manutenção dos instrumentos de medição de irradiância solar (Varela, 2007). O Brasil, por situar-se na região tropical próximo a linha do Equador, recebe grande incidência de radiação solar durante todo o ano ao longo de seu território e sua localização também resulta em uma menor variabilidade interanual (Pereira, 2006).

O contínuo progresso no país em relação ao incentivo de fontes de geração alternativa (micro e minigeração distribuída, leilões de energia e projetos de P&D) vem crescendo. Além disso, há o crescimento da demanda sobre informações de medições de recurso solar, pois a quantidade de estações solarimétricas no Brasil ainda não é suficiente, e das poucas existentes, a disponibilidade dos dados brutos sofrem restrição (Soares, 2016) e (EPE, 2017).

A implantação desses sistemas requer dimensionamento adequado e estudos prévios dos recursos disponíveis no local de instalação. São realizadas diversas análises e simulações para prever a geração de energia elétrica, balanço energético e estudos econômicos para comprovar a viabilidade do sistema. Para a modelagem dos sistemas e precisão dos cálculos é necessário que sejam utilizados dados precisos que simulem algo próximo da realidade. Entre as informações que implicam na qualidade dos resultados das maiores dessas simulações estão os dados de radiação e temperatura, painel solar e dos outros componentes (inversores, controladores de cargas, baterias, entre outros) (da Silva, 2018).

Para a realização de um dimensionamento de sistemas de geração fotovoltaica otimizado, a análise de dados climáticos está entre os mais importantes parâmetros necessários. Dessa forma, o objetivo desse artigo é comparar as

informações contidas em três bancos de dados sobre irradiância solar e temperatura da cidade de Campinas no estado de São Paulo. Os dados são reportados pelo site da Nasa Surface meteorology and Solar Energy (SSE), Meteornom 7.1 e uma base de dados coletados por uma estação localizada na cidade. Em seguida, os três bancos de dados são aplicados em uma simulação de sistema fotovoltaico para verificar a geração de energia.

2. RADIAÇÃO SOLAR

A energia solar pode ser definida como a energia eletromagnética emitida pelo sol, que se propaga no espaço em velocidades altíssimas. No vácuo sua velocidade é da ordem de 300.000 km/s. A radiação solar incidente no topo da atmosfera terrestre é chamada de radiação solar extraterrestre, 97% dos quais está confinado à faixa espectral de 290 a 3000 nm e é chamado de radiação solar (ou às vezes de ondas curtas). Parte da radiação solar extraterrestre penetra a atmosfera até a superfície da Terra, enquanto parte dela é espalhada e /ou absorvida pelas moléculas de gás, partículas de aerossol, gotículas de nuvens e cristais de nuvem na atmosfera. Para uma temperatura de 300 K, 99,99% da energia da radiação terrestre tem um comprimento de onda superior a 3.000 nm e cerca de 99% superior a 5.000 nm. Para temperaturas mais baixas, o espectro é deslocado para comprimentos de onda mais longos (Langlo, 2013), (WMO, 2008) e (Vieira, 2010).

Os dados de radiação solar podem ser disponibilizados como irradiância ou irradiação. A irradiância é uma medida instantânea de potência, cuja unidade é W/m², e a irradiação é uma medida de energia, cuja unidade é Wh/m² (Kitayama, 2019).

Parte da radiação extraterrestre sofre absorção na camada da ionosfera por moléculas de nitrogênio, oxigênio, ozônio e outros gases. Portanto, a atmosfera atua como um filtro, de forma que a irradiância sofre atenuação e espalhamento causado pelos gases e partículas suspensas (Kitayama, 2019). A fração da irradiância que atinge a superfície terrestre sem sofrer nenhuma interferência de partículas de poeira e gases é chamada de irradiância direta normal (Direct Normal Irradiance, DNI). A parte da irradiância que sofre espalhamento é chamada de irradiância difusa (Diffuse Horizontal Irradiance, DHI). A soma das irradiâncias direta e difusa resulta na irradiância global. Portanto, pode-se definir a irradiância global horizontal como a energia emitida pelo sol que atinge a superfície terrestre. A irradiância é um dos fatores que mais influenciam na produção de energia dos sistemas fotovoltaicos, quanto maior a irradiância sobre os painéis, maior a corrente gerada (Kitayama, 2019).

3. MÉTODOS ESTATÍSTICOS APLICADOS NA ANÁLISE DE DADOS

Para análise comparativa entre as bases de dados estudadas, foram utilizados métodos estatísticos que visam avaliar o grau de variabilidade dos dados em torno da média, para isso é utilizado as medidas de dispersão como desvio médio, variância e desvio padrão. Além disso, o conjunto de dados médios e medianas são incluídos para representar o valor central do conjunto total de dados. A seguir são apresentados os métodos aplicados.

3.1 Médias

A média é dada pela soma de todos os elementos que estão sendo considerados, dividido pelo número de elementos n como mostrado na Eq. (1).

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

3.2 Variância e desvio padrão amostral

Variância e desvio padrão amostral são medidas de dispersão de uma amostra, servem pra dar uma ideia do quão distante da média a amostra está (Gurgel, 1958). O desvio padrão amostral é dado pela Eq.(2) e a variância amostral é o resultado obtido pelo desvio padrão ao quadrado.

$$dp(x) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2)$$

3.3 Mediana

A mediana é uma medida de posição que separa o conjunto de dados ordenados exatamente na metade, deixando 50% dos valores abaixo e 50% acima.

3.4 Quartis

Os quartis são medidas de posição que separam o conjunto de dados ordenados em 4 partes iguais. O primeiro quartil Q1 indica o ponto de separação onde 25% da amostra está abaixo desse valor, o segundo quartil Q2 é a mediana, com 50% acima e abaixo e o terceiro quartil Q3 é o ponto onde 75% da amostra está abaixo (Moretin, 2017).

3.5 Boxplot

O gráfico Boxplot ajuda a visualizar como se portam os dados em relação a simetria e dispersão. Na sua construção é considerado o eixo y do gráfico o valor dos dados, primeiro é traçado um retângulo onde a parte superior é o Q3 e a inferior Q1, com um traço na mediana (Q2), em seguida é criada uma linha do Q3 até o limite superior (valor máximo amostrado) e do Q1 até o limite inferior. Os pontos discrepantes, chamados *outliers* são valores que ficam abaixo do valor $Q1 - 1.5(Q3 - Q1)$ ou acima de $Q3 + 1.5(Q3 - Q1)$ (Moretin, 2017). Como mostrado na Fig. 1.

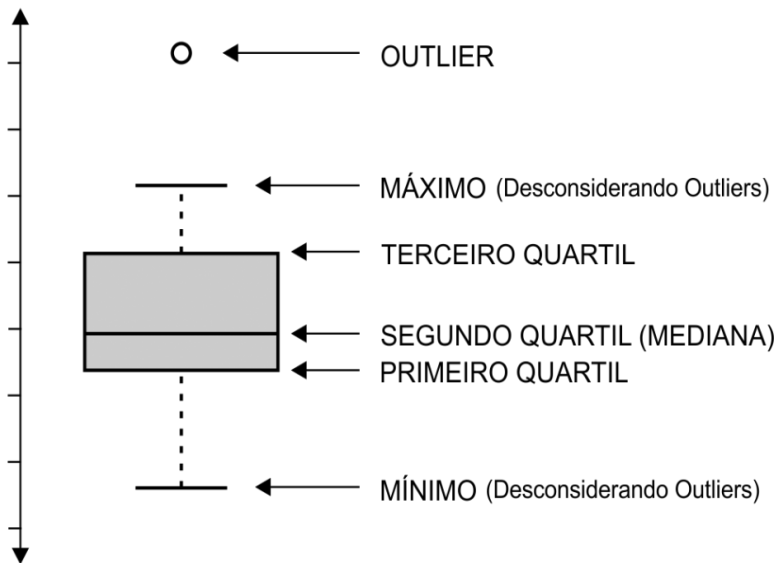


Figura 1 – Informações contidas no Boxplot. Fonte: ABG¹, 2018.

3.6 rMBE e rRMSE

Também serão analisados o erro médio relativo (rMBE – *relative Mean Bias Error*) e a raiz do erro quadrático médio relativo (rRMSE – *relative Root Mean Square Deviation*) (Lima, 2016). Esses dois métodos apresentam a diferença entre os bancos de dados da Nasa SSE e da Meteonorm em comparação com os dados reais de campinas.

O rMBE é dada pela Eq.(3) e o rRMSE pela Eq.(4).

$$rMBE(\%) = \frac{100}{n} \sum_{i=1}^n \left| \frac{A_i - C_i}{A_i} \right| \quad (3)$$

$$rRMSE(\%) = \frac{100}{n} \sqrt{\sum_{i=1}^n \left(\frac{A_i - C_i}{A_i} \right)^2} \quad (4)$$

Onde será utilizado A_i como os valores da Nasa e Meteonorm e C_i os valores medidos dos dados reais de Campinas, n é a quantidade de amostras.

4. RESULTADOS E DISCUSSÃO

Para o estudo comparativo entre os bancos de dados solarimétricos, foram utilizados os dados de temperatura e irradiação global horizontal de campinas obtidos a partir da Nasa SSE e da Meteonorm através do software PVsyst, e os dados reais são da estação solarimétrica da Usina Solar Fotovoltaica de Tanquinho.

¹ Disponível em: <http://www.abgconsultoria.com.br/blog/boxplot-como-interpretar>

O tratamento dos dados e a análise dos indicadores estatísticos são gerados no Rstudio que é um software livre com uma linguagem de programação voltada para gráficos e cálculos estatísticos.

Os dados coletados do PVSyst foram extraídos de hora em hora, enquanto os dados reais são fornecidos de minuto a minuto (de março a novembro) de 2016. Ambos os dados foram organizados na mesma unidade de medida, e para efeitos visuais de gráfico foram geradas as médias diárias e mensais dos três banco de dados. É avaliado as diferenças entre os valores medidos utilizando dados horários, e a média horária dos dados reais.

Os resultados obtidos dos bancos de dados são apresentados por uma análise gráfica das médias mensais de radiação (Fig. 4) e temperatura na (Fig. 5) entre as três bases de dados, em que é possível observar certa proximidade entre os valores e principalmente da média anual dos dados.

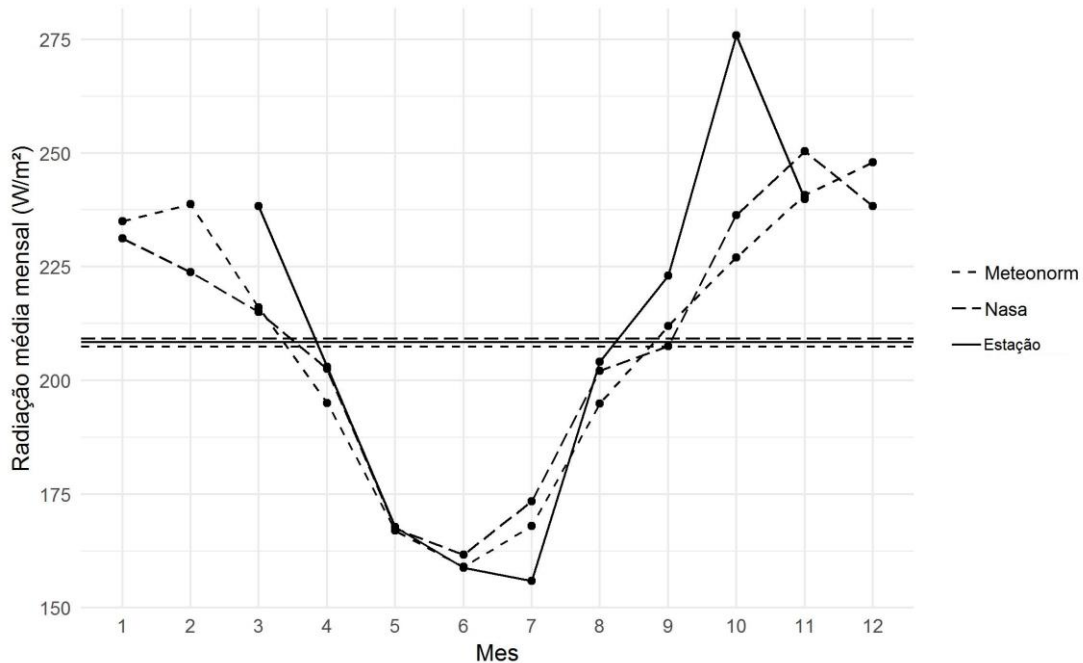


Figura 4 – Médias mensais de Radiação. Fonte: Próprio autor.

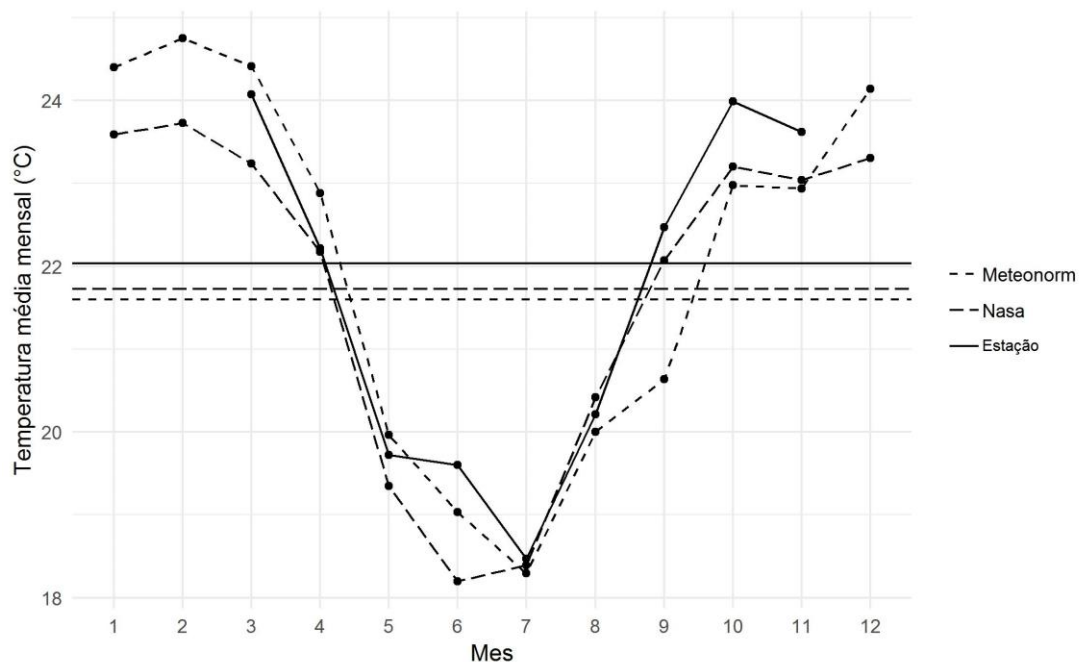


Figura 5 – Médias mensais de Temperatura. Fonte: Próprio autor.

Observa-se pelos Boxplots as diferenças dos dados diários nas Fig. 6 e 7 que as bases medianas em cima do mesmo ponto próximo a 0. A diferença do cálculo de dispersão é vista nas Tab. 1 e 2, em que aparecem as diferenças horárias, que podem ser causadas pela proximidade dos dados em relação ao tempo.

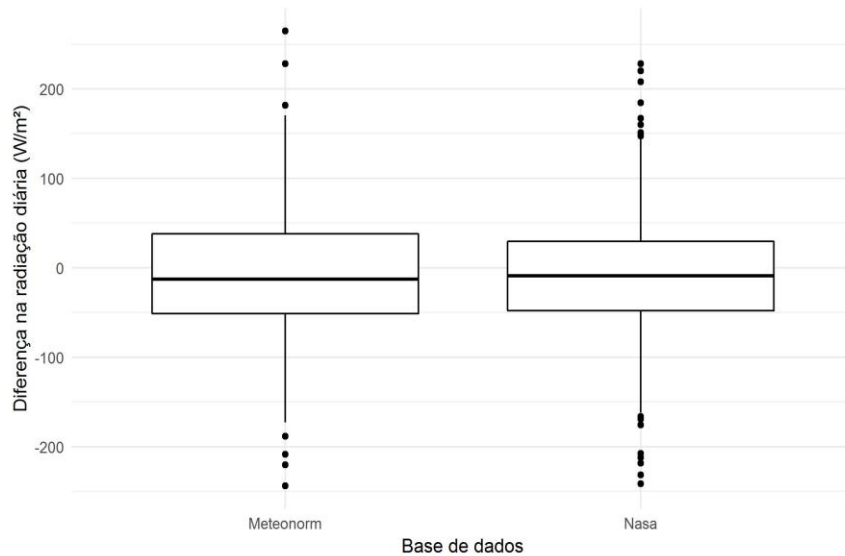


Figura 6 – Boxplot das diferenças entre a radiação considerada por dia nas bases de dados.
Fonte: Próprio autor.

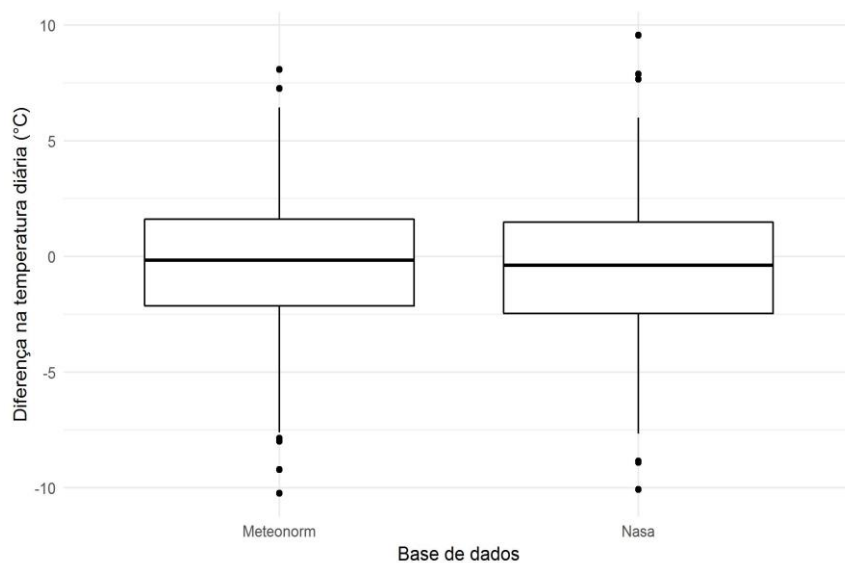


Figura 7 – Boxplot das diferenças entre as temperaturas considerada por dia nas bases de dados.
Fonte: Próprio autor.

As diferenças entre os bancos de dados são aproximadamente simétricas e centralizadas próximas a zero, o que justifica a paridade entre as médias anuais, mas em ambos os casos resultam em diferenças nas estatísticas dos erros, além de diferenças nas comparações entre os máximos que é de 36,595 e mínimos de -2,236 nos resultados. O desvio padrão entre dados da Nasa/Dados reais e Meteonorm/Dados reais é de 0,69613. Todos os resultados das medidas de radiação entre os bancos de dados são exibidos na Tab. 1. Em relação a temperatura, a Tab. 2 mostra nos indicadores diferença sem significância nos resultados do rMBE, mínimo e desvio padrão. Porém, as diferenças passam a ser significativas ao analisar os máximos, mediana e média.

Tabela 1 – Estatísticas das diferenças das medidas de Radiação.

Fonte: Próprio autor.

Estatísticas	Nasa SSE – Dados Reais	Meteonorm – Dados Reais
Desvio Padrão	79.76283	79.0667
rMBE	33.34091%	37.85059%
rRMSE	3.743523%	3.831132%
Mínimo	-241.394	-243.63
1º Quartil	-48.123	-51.36
Mediana	-9.189	-12.96
Média	-5.645	-9.74
3º Quartil	29.402	38.09
Máximo	228.025	264.62

Tabela 2 – Estatísticas das diferenças das medidas de Temperatura.

Estatísticas	Nasa SSE – Dados Reais	Meteonorm – Dados Reais
Desvio Padrão	3.042434	3.06441
rMBE	12.1911%	12.13335%
rRMSE	0.988235%	0.9960121%
Mínimo	-10.0726	-10.2274
1º Quartil	-2.4594	-2.1458
Mediana	-0.4755	-0.1696
Média	-0.47555	-0.3565
3º Quartil	1.4748	1.6055
Máximo	9.5534	8.0834

Os dados de radiação apresentam valores próximos entre Nasa SSE e Meteonorm, com uma pequena vantagem dos dados da Nasa SSE onde as medidas de dispersão e os quartis são mais próximo de zero, assim como a média. Contudo, as diferenças são mais preocupantes do que as de temperatura, os desvios padrão e os rMBE estão altos, assim como os erros máximos e mínimos.

Para verificar a influência na geração em um gerador fotovoltaico, foi realizada uma simulação a partir do Matlab utilizando essas bases de dados com módulos de 265W da Canadian e o resultado apresenta os dados da Meteonorm com diferença de 22.239 kWh ao ano em relação aos dados reais. Apesar da diferença de 12.6068 kWh/ano, o banco de dados da Nasa SSE é que mais se aproxima dos dados reais de geração de energia. Os resultados podem ser melhor observados a partir da Fig. 8 e da Tab. 3.

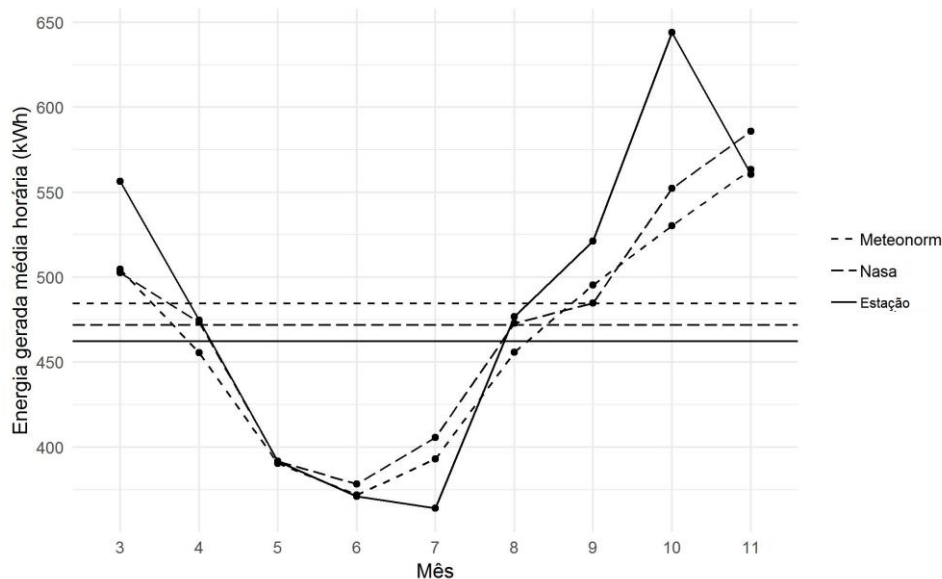


Figura 8 – Gráfico da geração mensal simulada com cada base de dados. Fonte: Próprio autor.

Tabela 3 – Geração média simulada. Fonte: Próprio autor.

Base de Dados	Média horária durante o ano (kWh)
Dados Reais	484.4814
Nasa SSE	471.8806
Meteonorm	462.2424

5. CONCLUSÃO

Ao comparar as bases de dados da Nasa SSE, Meteonorm e os dados reais da localizada de Campinas é visto uma proximidade em relação a média mensal e anual nas informações de temperatura e radiação, mas os dados se diferenciam muito quando a análise é observada entre os dados diários e mais ainda quando visto de hora em hora.

Apesar dessas diferenças nos dados meteorológicos ao ser feita a simulação para observar os dados de energia é visto que a geração de energia entre Nasa SSE foi muito próxima aos dados reais, ou seja, as diferenças pontuais dessas duas bases de dados não influenciaram radicalmente a geração de energia anual da base de dados. Porém, os dados da Meteonorm apresentam maior conservadorismo para a aplicação em projetos de sistemas fotovoltaicos.

Em trabalhos futuros será proposto métodos mais robustos para testes de igualdade, além de novas simulações para verificar suas influências em casos reais.

Agradecimentos

Este trabalho foi desenvolvido no âmbito do Programa de Pesquisa e Desenvolvimento do Setor Elétrico PD-00063-3032/2017 - PA3032: "Modelo de campus sustentável na Universidade de Campinas - Brasil: um laboratório vivo integrado para energia renovável, mobilidade elétrica, eficiência energética, monitoramento e gestão da demanda de energia", regulamentada pela Agência Nacional de Eletricidade (ANEEL), em parceria com a CPFL Brasil. Além do apoio e suporte da FAPESP (2016/08645-9), CAPES e CNPQ.

REFERÊNCIAS

- Cantor, Guillermo Andres Rodriguez. "Influência dos fatores climáticos no desempenho de módulos fotovoltaicos em regiões de clima tropical." (2017).
- Da Silva, Francisco Raimundo, et al., 2008. Implantação da Estação Solarimétrica de Natal-RN. II Congresso Brasileiro de Energia Solar e III Conferência Regional Latino-Americana da ISES, Florianópolis, Brasil. 2008.
- Da Silva, Michelle Kitayama; Narvaez, Dante Inga; de Melo, Karen Barbosa; Costa, Tatiane Silva; de Siqueira, Thais Gama & Villalva, Marcelo Gradella (2018). Comparative Analysis of Meteorological Databases and Transposition Models Applied To Photovoltaic Systems. Proceedings XXII Congresso Brasileiro de Automática. Disponível em: <https://doi.org/10.20906/cps/cba2018-0672>
- EPE – de Pesquisa Energética, Empresa, 2014. "Instruções para Solicitação de Cadastramento e Habilitação Técnica com vistas à participação nos Leilões de Energia Elétrica." Rio de Janeiro: EPE (2014).
- Gurgel, J. T. A., Frederico Pimentel Gomes, and A. P. Trivelin. "Determinação da dependência aritmética e desvio padrão de quocientes de dependência e dependentes." Anais da Escola Superior de Agricultura Luiz de Queiroz 14 (1958): 31-42.
- Jarraud, M., 2008. Guide to meteorological instruments and methods of observation (WMO-No. 8). World Meteorological Organisation: Geneva, Switzerland.
- Khatib, Tamer, and Wilfried Elmenreich. "Novel simplified hourly energy flow models for photovoltaic power systems." Energy conversion and management 79 (2014): 441-448.
- Kitayama, Michelle. "Estudo de modelos matemáticos para análise da radiação solar e desenvolvimento de ferramenta para modelagem e simulação de sistemas fotovoltaicos" (2019).
- Lima, José Claudeni O., Flávio SC Bispo, and Douglas Bressan Riffel. "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 (2016).
- Lira, André Luiz de O., Bruno de Lima Soares, and Samira de Azevedo Santos, 2016. "Estação solarimétrica de referência–instalação, operação e manutenção." Revista Brasileira de Energia Solar 7.2 (2016): 132-139.
- Morettin, P. A., BUSSAB, W. O., 2017. Estatística básica. Editora Saraiva.
- Pereira, Enio Bueno, et al., 2006. Atlas brasileiro de energia solar. Vol. 1. São José dos Campos: Inpe, 2006.
- Team, Rstudio, (2015). Rstudio: integrated development for R. Rstudio, Inc., Boston, MA UR. Rstudio. Com 42 (2015): 14.
- Varela, Gustavo Hellstrom, 2007. "Medição e Avaliação dos dados Solarimétricos." Seminário Estudantil de Produção Acadêmica 10.1 (2007).
- Vieira, A. P. V., Abreu, A. R., Luiz, W., Costa, L., 2010. Piranômetro.

Zo, Il-Sung, et al., 2017. Baseline Surface Radiation Network (BSRN) quality control of solar radiation data on the Gangneung-Wonju National University radiation station. Asia-Pacific Journal of Atmospheric Sciences 53.1 (2017): 11-19.

STATISTICAL ANALYSIS OF SOLAR RADIATION AND TEMPERATURE DATABASES FOR PHOTOVOLTAIC SYSTEM APPLICATION

Abstract. Photovoltaic solar energy has been exploited as a solution to meet the growing demand for electricity through a renewable source of energy. Optimized simulations of photovoltaic systems require the use of accurate solarimetric and climate data. This paper aims to compare the difference in the simulation of photovoltaic systems using the solar radiation and temperature data provided by the Nasa SSE and Meteonorm databases, with the data measured by a real solarimetric station installed in Campinas. The analyses were performed using statistical methods and simulations of a photovoltaic system to compare the generated energy, using each of the studied databases. Comparative radiation results show close values between Nasa SSE and Meteonorm, with a slight advantage of NASA data where dispersion measurements and quartiles are closer to zero, as well as the mean. However, the most punctual differences occur in temperature analysis about standard deviation and rMBE, as well as the maximum and minimum errors. The paper concludes that NASA SSE data is the closest to the actual data, despite having a power generation difference of approximately 12 kWh/year.

Key words: Photovoltaic solar energy, solarimetric data, Nasa SSE, Meteonorm, Statistical methods.