



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

USO DE APRENDIZADO DE MÁQUINA PARA GERAÇÃO DE SÉRIES TEMPORAIS DE IRRADIAÇÃO SOLAR NA SUPERFÍCIE

Mariane Souza Guarachi

Relatório de Iniciação Científica do
programa PIBIC, orientada pelo Dr.
André Rodrigues Gonçalves e Dr.
Fernando Ramos Martins.



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

USO DE APRENDIZADO DE MÁQUINA PARA GERAÇÃO DE SÉRIES TEMPORAIS DE IRRADIAÇÃO SOLAR NA SUPERFÍCIE

Mariane Souza Guarachi

Relatório de Iniciação Científica do
programa PIBIC, orientada pelo Dr.
André Rodrigues Gonçalves e Dr.
Fernando Ramos Martins.

URL do documento original:

INPE
São José dos Campos
2021

AGRADECIMENTOS

A Deus, por ter me dado saúde e proporcionado condições para que eu conquistasse meus sonhos e vencesse os obstáculos encontrados ao longo da realização deste estudo.

Aos meus pais e irmãos por todo apoio, amor e carinho, além de sempre me incentivarem a ir atrás do que eu realmente desejo.

Aos meus amigos que acreditaram em mim e que contribuíram de alguma forma para que eu desenvolvesse este projeto.

Ao Kauê, pelo companheirismo ao longo dessa jornada e por sempre me encorajar nas minhas escolhas.

Ao auxílio financeiro do Instituto Nacional de Pesquisas Espaciais (INPE) e o Centro de Pesquisa e Desenvolvimento Científico e Tecnológico (CNPq), contemplada pelo PIBIC.

Em especial gostaria de agradecer meu orientador André Rodrigues Gonçalves e o meu coorientador Fernando Ramos Martins pelos ensinamentos passados, pela paciência no meu processo de aprendizagem, permitindo que eu desempenhasse o meu melhor, por todos os conselhos compartilhados, e principalmente pela amizade.

RESUMO

É de suma importância ter disponível séries temporais de irradiância solar de um determinado local na superfície da Terra para aplicação em áreas científicas e tecnológicas. Em um caso mais específico, como a geração de eletricidade pela energia solar são necessários dados detalhados de irradiância solar do local escolhido para modelar o desempenho de uma planta fotovoltaica. Frequentemente, estas séries temporais são obtidas pelas medições realizadas no solo ou são derivadas de imagens de satélites. No entanto, as medições de radiação solar podem apresentar descontinuidades no registro dos dados, que duram de minutos a dias, o que geralmente é causado pelo mau funcionamento dos instrumentos de medições, manutenção preventiva e corretiva deficientes e ausência de energia nos equipamentos. Esta falta de dados é prejudicial para o estudo de viabilidade de uma planta ou sistema fotovoltaico em um determinado local, pois a ausência de dados acarreta no crescimento das incertezas das avaliações sobre o potencial técnico disponível no local. Portanto, nesses casos se faz necessária a utilização de séries temporais geradas sinteticamente ou matematicamente para preencher as lacunas de dados das medições. Modelos estatísticos para a geração de séries temporais podem utilizar diversas metodologias, incluindo o aprendizado de máquina com a utilização de redes neurais artificiais capazes de acumular conhecimentos a partir de exemplos e repetições. Este trabalho propõe o desenvolvimento e emprego de metodologias de aprendizado de máquinas para a previsão de dados de irradiação solar por meio de séries temporais. A pesquisa foi iniciada pelo ex-bolsista Rafael Mariano Juvêncio, posteriormente substituído pela estudante Mariane Souza Guarachi em maio de 2021. Neste período foi realizado o levantamento bibliográfico do conhecimento publicado sobre a geração de séries temporais e capacitação no uso da linguagem Python. Em continuidade serão utilizados diferentes conjuntos de preditores e topologias de redes neurais com o intuito de investigar o desempenho no preenchimento de ausência de dados em séries temporais.

Palavras-chave: Energia Solar. Sensoriamento Remoto. Modelagem Numérica.

USE OF MACHINE LEARNING FOR GENERATING TIME SERIES OF SOLAR SURFACE IRRADIATION

ABSTRACT

It is of paramount importance to have available time series of solar irradiance of a given location on the Earth's surface for application in scientific and technological areas. In a more specific case, such as electricity generation from solar energy, detailed solar irradiance data from the chosen location are needed to model the performance of a photovoltaic plant. Often, these time series are obtained from measurements taken on the ground or are derived from satellite images. However, measurements of solar radiation can present discontinuities in data recording, which last from minutes to days, which is usually caused by the malfunction of measuring instruments, deficient preventive and corrective maintenance and lack of energy in the equipment. This lack of data is detrimental to the feasibility study of a plant or photovoltaic system at a particular location, since the absence of data entails the growth of uncertainties reviews of the technical potential available on site. Therefore, in these cases it is necessary to use synthetically or mathematically generated time series to fill the gaps in the measurement data. Statistical models for the generation of time series can use several methodologies, including machine learning with the use of artificial neural networks capable of accumulating knowledge from examples and repetitions. This paper proposes the development and employment of machine learning methodologies for the prediction of solar irradiation data through time series. The research was initiated by former scholarship holder Rafael Mariano Juvêncio, later replaced by student Mariane Souza Guarachi in May 2021. During this period, a bibliographic survey of published knowledge about the generation of time series and training in the use of the Python language was carried out. In continuity, different sets of predictors and neural network topologies will be used in order to investigate the performance in filling the absence of data in time series.

Keywords: Solar Energy. Remote sensing. Numerical Modeling.

LISTA DE FIGURAS

	<u>Pág.</u>
Figura 1 – Fluxograma dos próximos passos a serem realizados no projeto.....	4
Figura 2 – Representação do Ângulo de Zênite, de Azimute e do Ângulo Solar.	8
Figura 3 – Representação da declinação solar.	9

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO.....	1
1.1 Objetivos.....	2
2 METODOLOGIA.....	4
2.1 Modelagem Numérica.....	4
2.2 Dados Observados.....	5
3 DESENVOLVIMENTO.....	7
3.1 Radiação Solar.....	7
3.2 Modelagem e Séries Temporais.....	10
3.3 Redes Neurais Artificiais.....	11
4 CONCLUSÃO.....	13
REFERÊNCIAS BIBLIOGRÁFICAS.....	14

1 INTRODUÇÃO

Atualmente tem-se o conhecimento que o consumo e a geração de eletricidade através de fontes não renováveis ocasiona danos ao meio ambiente, como o aquecimento global e as mudanças climáticas, provocados pelo aumento da temperatura do planeta e pela emissão de gases do efeito estufa (FIORIN et al., 2011). Por isso, se faz cada vez mais necessário a alteração para uma matriz energética baseada em fontes renováveis (FIORIN et al., 2011). Além da questão ambiental é importante diversificar a matriz para garantir maior desempenho e segurança na distribuição de eletricidade para o país (FIORIN et al., 2011).

Uma alternativa para as questões levantadas anteriormente é a utilização da energia solar para geração de eletricidade. No Brasil, a utilização do recurso solar para geração de energia enfrenta vários desafios, mesmo que seja um país tropical, como o escasso conhecimento sobre a distribuição do recurso no espaço e no tempo, alto custo dos equipamentos, lacunas de dados e grande influência do clima (FIORIN et al., 2011). Com isso, baseado nas questões ambientais e nos desafios encontrados surge a necessidade de se desenvolver uma pesquisa para criação de conhecimentos e dados sobre a irradiação solar na superfície.

O recurso solar apresenta intermitências, tornando necessário o uso de séries de alta resolução para simular com mais confiabilidade o desempenho dos sistemas de conversão de energia. As séries temporais de irradiância solar disponíveis influenciam diretamente na tomada de decisão sobre a implementação de uma planta fotovoltaica para geração de energia em um determinado local (POLO et al., 2011). Logo, é de suma importância que se tenha dados detalhados de irradiância solar (POLO et al., 2011).

A obtenção destas séries temporais ocorre a partir de medições realizadas no solo ou são derivadas de imagens de satélites (POLO et al., 2011). No entanto, as medições de radiação solar podem apresentar discontinuidades no registro dos dados, que duram de minutos a dias, o que geralmente é causado pelo mau funcionamento dos instrumentos de medições, manutenção preventiva e

corretiva deficientes e ausência de energia nos equipamentos (SCHWANDT et al., 2014).

A ausência de dados provoca um aumento das incertezas envolvidas na modelagem do potencial energético disponível e com isso compromete economicamente a projeção de instalação de uma planta ou sistema fotovoltaico para um determinado local, pois os investimentos necessários são altos para serem realizados em meio a grandes incertezas.

Por isso, de acordo com a natureza dos dados ausentes se faz necessária a utilização de séries temporais geradas sinteticamente ou matematicamente para preencher as lacunas de dados das medições (POLO et al., 2011). A geração de séries temporais pode ser realizada a partir de modelos estatísticos utilizando várias metodologias diferentes, inclusive o aprendizado de máquina com a utilização de redes neurais artificiais capazes de acumular conhecimentos a partir de exemplos e repetições. Este trabalho propõe o desenvolvimento e emprego de metodologias de aprendizado de máquinas para a previsão de dados de irradiação solar por meio de séries temporais.

Os modelos estatísticos preveem e estimam os dados de irradiação solar que incidem na atmosfera em um determinado local de interesse por meio de equações empíricas, os dados são obtidos pelas imagens de satélites ou pelas medições feitas no solo (FIORIN et al., 2011). Os modelos de Redes Neurais Artificiais contemplam um modelo estatístico, onde são compostos por unidades simples, chamadas de neurônios, os quais processam as funções matemáticas utilizadas (FIORIN et al., 2011). Logo, as Redes Neurais Artificiais possuem o potencial de auto-organização, generalização e processamento temporal das informações (FIORIN et al., 2011).

1.1 Objetivos

A regulamentação do setor elétrico estabelece a obrigatoriedade de aquisição de um ano de dados para participação de leilões de venda de energia proveniente de fonte solar. No entanto, apenas um ano de medições não é suficiente para análise de variabilidade interanual do recurso solar em uma

área para produção de energia e nem suficiente para atrair possíveis investidores em razão das incertezas e riscos associados.

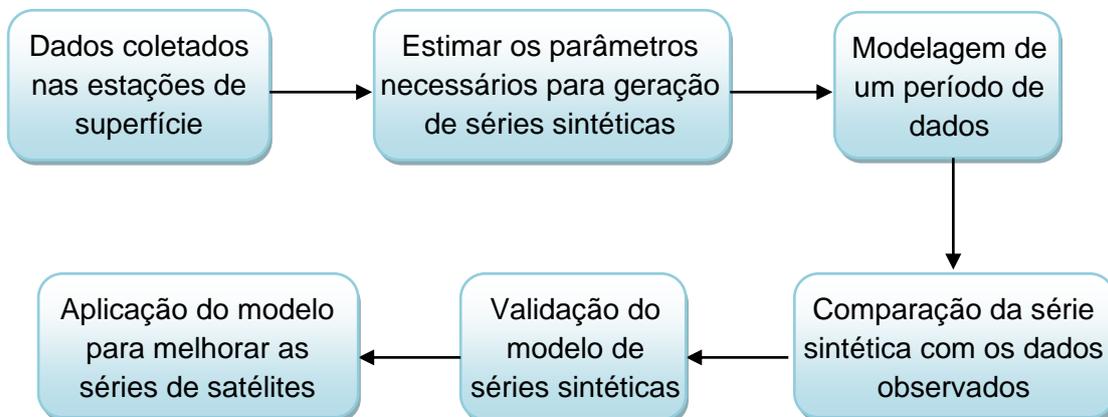
Para que possa ter maior confiabilidade no potencial disponível de energia solar, é desejável analisar os índices de irradiação solar por períodos mais longos, de até dez anos. Nesse período é possível ter uma ideia de fenômenos e variabilidades climáticas que podem ocorrer em determinada área para que seja viável a análise de irradiação solar.

Portanto, esse trabalho tem como objetivo o desenvolvimento de metodologias de aprendizado de máquina para obtenção de estimativas de irradiação solar por meio de geração de séries temporais de dados a partir de fontes distintas: com base nas informações disponibilizadas por (i) modelos numéricos de transferência radiativa na atmosfera e/ou previsão do tempo e (ii) em conjunto de dados observados em estações automáticas meteorológicas e radiométricas. A série artificial de dados possibilitará calcular a variância no potencial de energia solar no local de interesse.

2 METODOLOGIA

O trabalho realizado compreende o período de substituição do bolsista de quase três meses, onde foi feita uma revisão bibliográfica, por meio de pesquisas em periódicos (Google Scholar, Science Direct e CAPES) e pesquisas em busca de artigos científicos, teses ou dissertações. Esta revisão bibliográfica possuiu o objetivo de familiarização do bolsista com o tema e com alguns conceitos fundamentais e essenciais para se iniciar o trabalho. Com isso, foi realizado um levantamento bibliográfico do conhecimento publicado sobre a geração de séries temporais e capacitação no uso da linguagem Python. Em continuidade serão utilizados diferentes conjuntos de preditores e topologias de redes neurais com o intuito de investigar o desempenho no preenchimento de ausência de dados em séries temporais. O fluxograma abaixo demonstra os próximos passos e o que será realizado.

Figura 1 – Fluxograma dos próximos passos a serem realizados no projeto.



Fonte: Produção da autora.

2.1 Modelagem Numérica

O presente estudo fará uso de redes neurais artificiais alimentadas por estimativas de irradiação solar na superfície produzidos pelo modelo de transferência radiativa BRASIL-SR e dados meteorológicos observados em estações da rede SONDA. O modelo BRASIL-SR simula os processos físicos

de atenuação da radiação solar na atmosfera utilizando dados observados por satélite.

As estimativas produzidas pelo modelo possuem resolução temporal igual às imagens de satélite (entre 30 minutos e 1 hora). As redes neurais artificiais serão utilizadas para refinar as estimativas produzidas pelos modelos numéricos mencionados anteriormente produzindo estimativas de irradiação solar incidente na superfície com resolução de 10 minutos ou menos. Considerando essa uma resolução temporal adequada para a simulação de operação de plantas de geração fotovoltaica.

A proposta tem como ponto de partida a aplicação de Redes Neurais Artificiais (RNA) amplamente adotadas para simulação de processos não-lineares. A RNA é uma ferramenta capaz de armazenar conhecimento a partir de treinamentos (exemplos) e pode ser utilizada em diversas áreas, como modelos de previsão, reconhecimento de padrões e outras aplicações.

A utilização desta metodologia de aprendizagem de máquina é resultado da capacidade de auto-organização, generalização e processamento temporal que permite solucionar problemas de alta complexidade (FIORIN et al., 2011). A organização de uma RNA é dependente da complexidade do problema a ser resolvido, variando desde problemas mais simples (linearmente separáveis) até problemas mais complexos (não-lineares).

2.2 Dados Observados

O Sistema de Organização Nacional de Dados Ambientais (SONDA) é um projeto do Instituto Nacional de Pesquisas Espaciais (INPE) que visa a implementação de infraestrutura, tanto física quanto de recursos humanos, para aprimoramento na base de dados coletados em superfície indispensáveis para o levantamento do potencial de recurso eólico e solar, este que é obtido através do modelo de satélite, BRASIL-SR (SONDA, 2021).

A base de dados da Rede SONDA será utilizada no treinamento da RNA para duas localidades com climatologias distintas: Petrolina (PE) e São Martinho da

Serra (RS). O principal motivo da escolha dessas duas localidades é devido as particularidades climáticas que apresentam. Petrolina se localiza na região Nordeste do Brasil, região quente e seca, com elevada incidência de radiação solar na superfície. A precipitação é baixa ao longo de todo o ano assim como a variabilidade sazonal da irradiação solar na superfície. São Martinho da Serra está localizada no Sul do país, bem na região central do Rio Grande do Sul. Nesta região, há uma variabilidade sazonal bastante elevada na incidência de radiação solar na superfície conforme indicado no Atlas Brasileiro de Energia Solar (PEREIRA et al., 2017).

3 DESENVOLVIMENTO

Este tópico foi elaborado para detalhamento e apresentação da revisão bibliográfica realizada pelo bolsista durante o período que compreende sua iniciação científica.

3.1 Radiação Solar

A Terra possui o Sol como sua principal fonte de fornecimento de energia (GUARNIERI, 2006). A energia provinda do sol incide na Terra por meio de feixes luminosos, os quais possuem uma faixa contínua de comprimentos de onda em todo o espectro eletromagnético, de acordo com a lei de Planck (GUARNIERI, 2006). O espectro eletromagnético é dividido em raios gama, raios-X, ultravioleta, visível, infravermelho, micro-ondas e ondas de rádio.

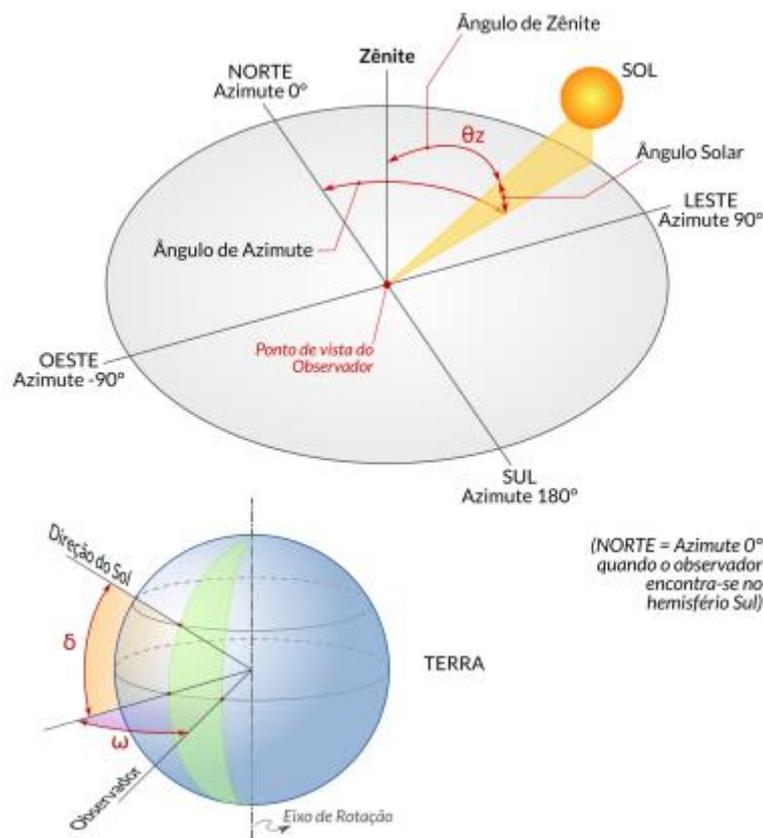
A radiação solar representa a principal força motriz para que ocorram na Terra e na atmosfera os processos químicos, dinâmicos e térmicos (MARTINS et al., 2004). Da energia emitida pelo Sol, cerca de 49% representa a radiação infravermelha, 43% a radiação de luz visível, 7% a radiação ultravioleta e menos de 1% representa os raios gama, raios-X e ondas de rádio (FIORIN et al., 2011).

Ao incidir na atmosfera a radiação pode ser espalhada ou absorvida, e aproximadamente 25% dessa radiação entra em contato com a superfície sem sofrer influência da atmosfera e os outros 75% são absorvidos ou espalhados de volta ao espaço ou em direção a Terra (FIORIN et al., 2011). O fluxo radiante é a taxa com que ocorre a transferência de energia por radiação eletromagnética, sendo definido em unidades de energia por unidade de tempo (Joules por segundo - $J.s^{-1}$ ou Watts - W) (GUARNIERI, 2006). Já o fluxo radiante que passa por uma unidade de área é caracterizado pela densidade de fluxo radiante, também conhecida como irradiância, expressa em Watts por metro quadrado ($W.m^{-2}$) (FIORIN et al., 2011; MARTINS et al., 2004).

A radiação solar incide na superfície plana com diferentes ângulos de inclinação (GUARNIERI, 2006). O ângulo zenital solar (θ_z) é definido como o

ângulo entre a vertical da superfície (normal) e a direção da radiação no mesmo local da Terra ao Sol, podendo variar de 0° a 90° (PEREIRA et al., 2017). O zênite local é o ponto que forma um ângulo de 90° com a superfície terrestre, ou seja, quando o Sol se localiza no horizonte (FIORIN et al., 2011; PEREIRA et al., 2017). Já o ângulo azimutal solar é definido como o ângulo formado entre a projeção do Sol no plano horizontal e o meridiano do observador, podendo variar de -180° a $+180^\circ$ (PEREIRA et al., 2017).

Figura 2 – Representação do Ângulo de Zênite, de Azimute e do Ângulo Solar.

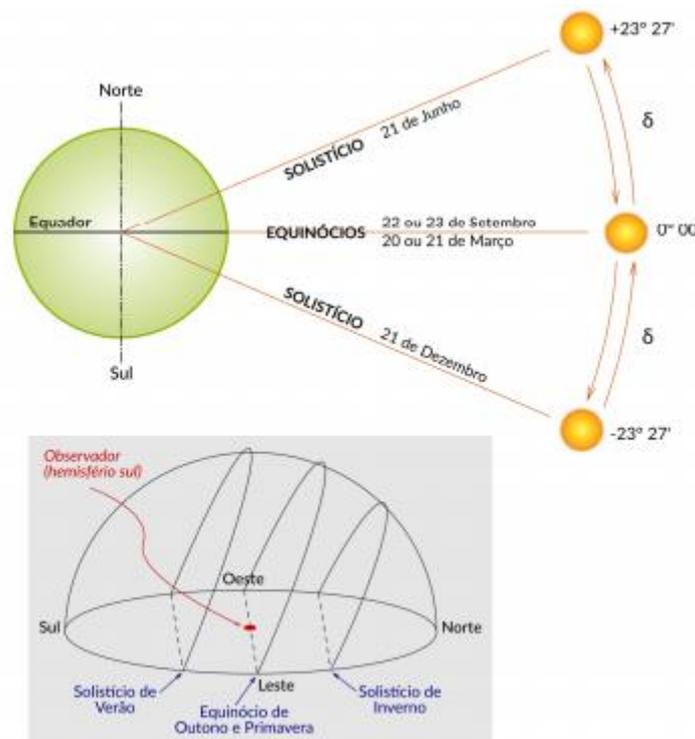


Fonte: Pereira et al. (2017).

O ângulo horário solar (ω) é definido como o deslocamento angular do movimento aparente do Sol devido à rotação da Terra, podendo variar entre -180° e $+180^\circ$ (PEREIRA et al., 2017). A cada uma hora o ângulo horário solar varia 15° , logo, como convenção utilizamos o ângulo positivo na parte da

manhã e negativo a tarde, e ao meio dia é considerado zero (PEREIRA et al., 2017). A declinação solar (δ) é definida como o ângulo formado pela linha de radiação que liga o Sol a Terra e a inclinação do plano equatorial da Terra, podendo variar entre $-23^{\circ} 27'$ e $+23^{\circ} 27'$ no período de um ano (PEREIRA et al., 2017).

Figura 3 – Representação da declinação solar.



Fonte: Pereira et al. (2017).

A separação das componentes da radiação solar incidente depende de todos os processos radiativos que ocorreram até que ela chegasse na Terra (FIORIN et al., 2011). Com isso, a irradiância difusa é aquela que incide na superfície horizontal resultante do espalhamento ocasionado pelas nuvens, material particulado na atmosfera, entre outros (FIORIN et al., 2011). A irradiância direta é a que incide na superfície horizontal, mas não sofre nenhum processo de espalhamento (FIORIN et al., 2011). Já a irradiância direta normal é a irradiância que possui a sua incidência em um ângulo de 90° com a superfície e

a irradiância global é a soma da irradiância direta com a irradiância difusa (FIORIN et al., 2011).

3.2 Modelagem e Séries Temporais

A disponibilidade de energia solar incidente na superfície terrestre pode ser avaliada através de uma rede de radiômetros e de interpolações das medidas de radiação (GUARNIERI, 2006). Porém, para que esta avaliação seja confiável é necessário que se tenha muitos radiômetros e que eles possuam pouco distanciamento entre si (GUARNIERI, 2006). Logo, uma proposta mais interessante para a obtenção dessas estimativas de radiação é a utilização de modelos computacionais (GUARNIERI, 2006).

Os modelos computacionais utilizados para estimar a radiação solar incidente são divididos em modelos estatísticos e modelos físicos (GUARNIERI, 2006). Os modelos estatísticos são definidos pelo emprego de expressões empíricas que estimam a radiação incidente em um determinado local na Terra em detrimento das propriedades do sistema Terra-Atmosfera obtidas pelos instrumentos no solo ou pelos satélites (FIORIN et al., 2011).

A tecnologia de sensoriamento remoto por meio de satélites artificiais proporcionou a obtenção de imagens de satélites, as quais são importantes para a definição da cobertura de nuvens e outros parâmetros atmosféricos (MARTINS et al., 2004). Já os modelos físicos são definidos pelo emprego de métodos matemáticos que simulam as condições atmosféricas e resolvem a equação de transferência radioativa (FIORIN et al., 2011). Além disso, a credibilidade do modelo físico está relacionada com a confiabilidade das informações disponíveis sobre as condições atmosféricas (FIORIN et al., 2011).

Independente do modelo utilizado, possuir séries temporais de irradiância solar de um local em específico com dados confiáveis é imprescindível para modelar e prever o desempenho de uma planta fotovoltaica (POLO et al., 2011). Portanto, quando não se é possível ter todos os dados de irradiação disponíveis, devido a alguma falha nas medições ou outros motivos descritos anteriormente, uma possível solução é empregar o modelo estatístico de

aprendizado de máquina para geração de séries temporais de irradiação solar sintéticas na superfície.

3.3 Redes Neurais Artificiais

As Redes Neurais Artificiais são consideradas uma ferramenta estatística e possuem em sua composição neurônios, que a partir dos dados de entrada fazem a computação de funções matemáticas não lineares (GUARNIERI, 2006). Os neurônios, também chamados de unidades de processamento devido a sua principal função, são colocados em paralelo formando camadas que são interligadas entre si a partir de conexões (sinapses) associadas a pesos sinápticos (GUARNIERI, 2006).

Os conhecimentos adquiridos pela RNA são armazenados pelos pesos sinápticos, por meio da ponderação das entradas de cada um dos neurônios (GUARNIERI, 2006). A princípio, para que se utilize a RNA é necessário que ela seja treinada para aprender os processos requeridos na resolução do problema, e para isso os pesos sinápticos são modificados até que a resolução seja aceitável (GUARNIERI, 2006). Logo, após o treinamento os pesos sinápticos são fixados, onde a partir de dados de entrada disponíveis pode-se estimar os dados de saída (GUARNIERI, 2006).

Os treinamentos são divididos em supervisionados e não-supervisionados. O treinamento supervisionado consiste na utilização de um vetor de entrada para calcular o vetor de saída e comparar o resultado obtido com o vetor alvo, logo, os pesos sinápticos são controlados para que os vetores de saída coincidam com o alvo (GUARNIERI, 2006). O ajuste dos pesos é realizado após a apresentação de todos os exemplos de dados de treinamento para a RNA (treinamento por lote) ou a cada iteração (treinamento sequencial) (GUARNIERI, 2006).

As RNAs possuem diferentes arquiteturas, e a escolha de qual utilizar está diretamente ligado com o problema a ser resolvido (FIORIN et al., 2011). O que varia de uma arquitetura para outra é a quantidade de neurônios e de camadas, o tipo de topologia da RNA e o tipo de conexão entre os neurônios

(FIORIN et al., 2011). As RNAs podem possuir apenas uma camada, ou seja, apenas um neurônio entre a camada de entrada e a de saída, ou podem possuir multicamadas com um conjunto de neurônios, uma camada de entrada e saída e uma ou mais camadas intermediárias (FIORIN et al., 2011).

Os neurônios possuem conexões que são classificadas em *feedback* (cíclica), onde a saída de um neurônio é a entrada de outro pertencentes a mesma ou a distintas camadas, e em *feedforward* (acíclica), onde a saída de um neurônio condiz com a entrada de outro neurônio pertencente a uma camada posterior (FIORIN et al., 2011). As conexões do tipo cíclica ou recorrente são aquelas que possuem realimentação da saída na entrada. Logo, as RNAs podem ser divididas em *Perceptrons* de Múltiplas Camadas (MLP), Processamento Temporal, *Self-Organizing*, *Perceptron* e *Adaline*, Memórias Matriciais, entre outras (FIORIN et al., 2011).

De acordo com o trabalho de Fiorin et al. (2011), pode-se perceber que a utilização de um modelo empírico como a RNA *Perceptrons* de Múltiplas Camadas (MLP) se mostrou eficaz para estimar os parâmetros relacionados com a incidência de irradiação na superfície e modelar parâmetros instantâneos de radiação derivados dos dados de cobertura de nuvens. Além disso, as RNAs demonstraram melhor desempenho nas previsões do que os modelos que utilizam regressão linear múltipla, em razão da RNA ser eficiente na exploração das não-linearidades das análises realizadas e das informações dos diferentes parâmetros de entrada. As redes MLP se destacam por serem o tipo mais simples capaz de representar qualquer função contínua, como processos físicos, por isso são muito utilizadas em geociências (processos naturais são tipicamente contínuos).

4 CONCLUSÃO

Devido a substituição do bolsista em maio de 2021 o cronograma de atividades a serem realizadas foi adaptado para que pudesse suprir as necessidades da nova bolsista. Com isso, nesse primeiro momento foi concluída a etapa inicial da pesquisa, a revisão bibliográfica.

A partir da revisão bibliográfica pude compreender melhor os conceitos envolvidos na geração de energia a partir de fontes renováveis, como a energia solar e sobre os processos de incidência de radiação solar na superfície, além da interação da radiação com a atmosfera.

Na literatura, o aprendizado de máquina para geração de séries temporais de radiação solar sintéticas se mostrou eficiente e capaz de fazer simulações e previsões dos dados de radiação solar. Além disso, ficou evidente a importância desses dados gerados sinteticamente para a modelagem do desempenho de uma planta fotovoltaica.

Portanto, os próximos passos a serem cumpridos são realizar a capacitação e desenvolvimento na linguagem de programação Python para a geração de séries de radiação solar sintéticas, verificar e validar o modelo computacional e por fim aplicar este modelo para melhorar a qualidade das séries obtidas pelos satélites.

REFERÊNCIAS BIBLIOGRÁFICAS

FIORIN, Daniel V. et al. Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares. **Revista Brasileira de Ensino de Física**, v. 33, n. 1, p. 1309-2-1309-20, 2011. Disponível em: <https://www.scielo.br/j/rbef/i/2011.v33n1/>. Acesso em: 6 jul. 2021.

GUARNIERI, Ricardo André. **Emprego de redes neurais artificiais e regressão linear múltipla no refinamento das previsões de radiação solar do modelo ETA**. 2006. 171 f. Dissertação (Mestrado) - Curso de Meteorologia, INPE, São José dos Campos, 2006. Disponível em: <http://mtc-m16b.sid.inpe.br/col/cptec.inpe.br/walmeida/2003/04.25.17.13/doc/mirrorget.cgi?languagebutton=pt-BR&metadatarepository=sid.inpe.br/MTC-m13@80/2006/05.30.18.50.03&index=0&choice=full>. Acesso em: 18 jul. 2021.

PEREIRA, E. B.; MARTINS, F. R.; GONÇALVES, A. R.; COSTA, R. S.; LIMA, F. L.; RÜTHER, R.; ABREU, S. L.; TIEPOLO, G. M.; PEREIRA, S. V.; SOUZA, J. G. **Atlas brasileiro de energia solar**. 2.ed. São José dos Campos: INPE, 2017. 80p. Disponível em: <http://doi.org/10.34024/978851700089>. Acesso em: 10 jul. 2021.

POLO, J. et al. A simple approach to the synthetic generation of solar irradiance time series with high temporal resolution. **Solar Energy**, v. 85, n. 5, p. 1164-1170, 2011. Disponível em: <https://doi.org/10.1016/j.solener.2011.03.011>. Acesso em: 6 jul. 2021.

SONDA. **SISTEMA DE ORGANIZAÇÃO NACIONAL DE DADOS AMBIENTAIS**. Disponível em: <http://sonda.ccst.inpe.br>. Acesso em: 15 jul. 2021.

SCHWANDT, Marko et al. Development and test of gap filling procedures for solar radiation data of the Indian SRRA measurement network. **Energy Procedia**, v. 57, p. 1100-1109, 2014. Disponível em: <https://doi.org/10.1016/j.egypro.2014.10.096>. Acesso em: 6 jul. 2021.