



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS  
DEPARTAMENTO DE ENERGIA NUCLEAR  
COMISSÃO NACIONAL DE ENERGIA NUCLEAR  
CENTRO REGIONAL DE CIÊNCIAS NUCLEARES DO NORDESTE  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS ENERGÉTICAS E  
NUCLEARES

EDSON RICARDO CALADO SABINO

**PREVISÃO DE RADIAÇÃO SOLAR E TEMPERATURA AMBIENTE  
VOLTADA PARA AUXILIAR A OPERAÇÃO DE USINAS FOTOVOLTAICAS**

Recife

2019

EDSON RICARDO CALADO SABINO

**PREVISÃO DE RADIAÇÃO SOLAR E TEMPERATURA AMBIENTE VOLTADA  
PARA AUXILIAR A OPERAÇÃO DE USINAS FOTOVOLTAICAS**

Tese submetida ao Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares para obtenção do título de Doutor em Tecnologias Energéticas e Nucleares.

**Área de concentração:** Fontes Renováveis de Energia.

**Orientadora:** Prof<sup>ª</sup>. Dr<sup>ª</sup>. Olga de Castro Vilela.

**Co-orientador:** Prof<sup>º</sup>. Dr. Alexandre Carlos Araújo da Costa.

Recife

2019

Catálogo na fonte  
Bibliotecário Carlos Moura, CRB-4 / 1502

S116m Sabino, Edson Ricardo Calado.  
Previsão de radiação solar e temperatura ambiente voltada para auxiliar a operação de usina fotovoltaicas / Edson Ricardo Calado Sabino – Recife, 2019.

112 f., il., tabs., abrev. e siglas.

Orientadora: Profa. Dra. Olga de Castro Vilela.

Coorientador: Prof. Dr. Alexandre Carlos Araújo da Costa.

Tese (Doutorado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares, 2019.

Inclui referências e apêndices.

1. Engenharia de energia solar. 2. Previsão de radiação solar. 3. Previsão de temperatura. 4. Imagens de satélite. 5. Combinação de previsão. I. Vilela, Olga de Castro (Orientadora). II. Costa, Alexandre Carlos Araújo da (Coorientador). III. Título.

621.47 CDD (22. ed.)

UFPE (BDEN 2019-37)

EDSON RICARDO CALADO SABINO

**PREVISÃO DE RADIAÇÃO SOLAR E TEMPERATURA AMBIENTE VOLTADA  
PARA AUXILIAR A OPERAÇÃO DE USINAS FOTOVOLTAICAS**

Tese apresentada ao Programa de Pós-Graduação em Tecnologias Energéticas e Nucleares da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Tecnologias Energéticas e Nucleares.

Aprovada em: 31/07/2019.

**BANCA EXAMINADORA**

---

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Olga de Castro Vilela (Orientadora)  
Universidade Federal de Pernambuco

---

Prof<sup>º</sup>. Dr. Alexandre Carlos Araújo da Costa (Coorientador)  
Universidade Federal de Pernambuco

---

Prof<sup>º</sup>. Dr. Fernando Roberto de Andrade Lima (Examinador Interno)  
Universidade Federal de Pernambuco

---

Prof<sup>º</sup>. Dr. Luis Arturo Gómez Malagón (Examinador Externo)  
Universidade de Pernambuco

---

Prof<sup>º</sup>. Dr. Tsang Ing Ren (Examinador Externo)  
Universidade Federal de Pernambuco

---

Dr. Thiago Luiz do Vale Silva (Examinador Externo)  
Agência Pernambucana de Águas e Clima

## AGRADECIMENTOS

Gostaria de agradecer a todos que contribuíram de forma direta ou indireta para que essa tese pudesse ser concluída.

Em primeiro lugar agradeço a Deus, por todas as oportunidades a mim concedidas, as quais me permitiram cursar este Doutorado.

Aos meus pais, que sempre primaram e incentivaram minha educação;

A minha Esposa, Betânia e a meus filhos, Vinícius e Lorena, pelo incentivo inicial, pelo apoio diário, pelos conselhos e pela abnegação de momentos de lazer em prol desta causa.

Aos meus orientadores, Olga Vilela e Alexandre Costa, pela solicitude, pela humanidade, pela expertise no tema, pelas injeções de ânimo, pelas valiosas análises e contribuições e pelas incontáveis reuniões que tivemos (incluindo finais de semana).

Aos alunos que contribuíram diretamente nesta pesquisa, desenvolvendo partes ou emitindo sugestões de melhorias, sem as quais essa tese não teria atingido os resultados almejados, dentre eles Leonardo Petribu, Heitor, Gabriel, Valentin, Emerson e Júlia.

Aos amigos do DEN e do CER, com os quais cursei as disciplinas, cujas amizades, trocas de experiências e apoios tornaram a caminhada mais amena.

Aos amigos da SE RCD, por todo o apoio e presteza, sempre incentivando e se disponibilizando em ajudar.

Ao IFPE *Campus* Pesqueira, em especial aos professores de Eletrotécnica, por todo o incentivo dado, sem o qual não seria possível concluir este trabalho.

A meus parentes, sogra, irmãos, tios, tias e primos, pelas palavras de encorajamento proferidas durante toda essa jornada.

A todos, os meus sinceros agradecimentos!

## RESUMO

A radiação solar e a temperatura do módulo são os fatores que mais influenciam a potência gerada por uma central Fotovoltaica. Este trabalho propõe uma metodologia para previsão de irradiância global horizontal (GHI) e temperatura em escala horária, visando atender aos horizontes de previsão de interesse para a programação de geração do Operador Nacional do Sistema (um dia à frente com *time-step* horário). A metodologia de previsão de GHI é composta essencialmente pela combinação de duas previsões: a primeira sendo baseada em dados de radiômetro, os quais apresentam boa precisão. A segunda baseia-se em imagens de satélite, das quais se busca extrair informações de um domínio espacial amplo, buscando assim aumentar a abrangência espacial e, conseqüentemente, a temporal. Pelo fato de a série temporal de entrada de GHI ter sido construída preservando-se apenas os períodos diurnos (de horas de incidência solar), os horizontes de previsão variaram de 1 até 12 passos à frente, os quais permitem cobrir a previsão de todos os horários da janela útil de radiação do dia seguinte. Já as previsões de temperatura são baseadas em dados de temperatura medidos na superfície e realizadas por modelos univariados de fácil implementação para horizontes de 1h a 24h à frente. Os sítios escolhidos (Brasília, Pesqueira e Petrolina) possuem climas variados, possibilitando o teste da metodologia em diferentes condições. São utilizados dois anos de dados horários de temperatura e três anos de GHI medidos em terra e imagens dos satélites GOES 12 e GOES 13 dos anos de 2013 a 2015, empregando a validação cruzada. Foram utilizados modelos lineares (Autorregressivos) para a temperatura e não lineares (Redes Neurais) para a GHI. Para o pré-tratamento (redução da dimensionalidade) das imagens de satélite, foram utilizadas análise de correlação para seleção dos pixels das imagens e análise de componentes principais. Os modelos de previsão de temperatura alcançaram erros médios de 0,6°C a 1,6°C, superando em todos os horizontes os modelos de persistência (modelo de referência). Os modelos individuais de previsão de GHI também superaram a persistência e modelos de outros autores em quase todos os horizontes. A combinação das previsões trouxe aumento no *Improvement* (melhora com relação ao modelo de Persistência) e no *Skill* de Taylor (estatístico calculado com base na correlação e nos desvios-padrão) de até 3 e 8 pontos percentuais, respectivamente. Além disso, foram alcançados resultados de previsão via imagens de satélite satisfatórios para horizontes além daqueles considerados pela literatura como favoráveis à sua utilização. Uma segunda contribuição do trabalho foi a proposição e validação de um modelo individual de previsão via imagens de satélite, o qual não utiliza dados de radiômetro nem mesmo na calibração, sendo uma alternativa para lugares onde não haja disponibilidade de dados medidos em estações radiométricas. Seus resultados conseguiram se aproximar dos obtidos por modelos treinados com dados de radiômetro.

**Palavras-chave:** Previsão de Radiação Solar. Previsão de Temperatura. Imagens de Satélite. Combinação de Previsão.

## ABSTRACT

Solar radiation and module's temperature are the power generated most influence factors by a photovoltaic plant. This work proposes a solar global horizontal irradiance (GHI) and hourly ambient temperature forecasting, aiming to meet forecast horizons of the generation dispatch. The GHI forecasting methodology is mainly composed of the two forecasting combination: the first one based on radiometer data, which have good accuracy. The second one based on satellite images, from which it is intended to extract information from a broad spatial domain, thus seeking to increase the spatial and, consequently, the temporal range. Because the GHI time series of entry was constructed preserving only the daytime periods (of solar incidence hours), forecast horizons varied from 1 to 12 steps ahead, which allow covering the forecast of all hours of the next day's radiation window. The temperature predictions are based on the ground measured temperature data and performed by univariate models of easy implementation for horizons from 1h to 24h ahead. The chosen sites (Brasília, Pesqueira, and Petrolina) have varied climates, allowing the establishment of the methodology in several conditions. Two years of hourly temperature data and three years (2013 to 2015) of radiometer measurements and satellites images of the GOES 12 and GOES 13, using cross-validation, are gone. Linear models (Autoregressive) and nonlinear models (Neural Networks) were used for temperature and GHI, respectively. For the pre-treatment (dimensionality reduction) of the satellite images, the selection of pixels by correlation, and the principal components analysis was used. The temperature prediction models reached RMSE from  $0.6^{\circ}\text{C}$  to  $1.6^{\circ}\text{C}$ , surpassing all persistence models. Individual GHI prediction models have also outperformed persistence and other authors' models in almost every horizon. Forecasts combination caused an increase in Improvement and Taylor Skill up to 3 and 8 percentage points, respectively. Also, interesting prediction results were achieved via satellite images for horizons beyond those considered in the literature as favorable to their use. The second contribution of this work was the proposition and validation of an individual satellite imagery prediction model, which does not use radiometer data in calibration, being an alternative to places where there is no availability of measured data in radiometric stations. Their results were able to approximate those obtained by trained models with radiometer data.

**Keywords:** Solar Radiation Forecasting. Temperature Forecasting. Satellite Images. Forecasting Combination.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Divisão das faixas de horizontes de previsão adotadas no presente trabalho. ....	18
Figura 2 – Relação entre as estratégias de previsão. ....	18
Figura 3 – Representação de um neurônio artificial.....	22
Figura 4 – Algumas das mais utilizadas funções de ativação em ANN.....	23
Figura 5 – Exemplo de rede neural artificial do tipo MLP.....	24
Figura 6 – Metodologia proposta para previsão de GHI e Temperatura ambiente com vistas a auxiliar previsões de potência de uma central fotovoltaica.....	49
Figura 7 – Distribuição geográfica e disponibilidade do recurso solar nos locais estudados. .	50
Figura 8 – Duas formas de converter dados da base minuto a minuto para a base horária.....	52
Figura 9 – Imagem da América do Sul do canal visível do Satélite GOES 13 do dia 05 de outubro de 2015 às 12h00 (horário de Brasília). ....	53
Figura 10 – Rosa dos ventos traçada com dados medidos de direção do vento da estação SONDA Petrolina. ....	54
Figura 11 – Tamanho e localização do recorte da imagem de satélite utilizado neste trabalho. ....	55
Figura 12 – Variação esperada dos componentes usados no cálculo do <i>Improvement</i> (Imp) na condição de céu nublado, onde os índices <i>m</i> e <i>p</i> dizem respeito ao modelo em estudo e ao modelo de persistência, respectivamente.....	57
Figura 13 – Variáveis sazonais anuais utilizadas como entrada para melhoria do modelo baseado em ANN.....	59
Figura 14 – Modelo de Previsão de GHI, chamado TER, que utiliza variáveis sazonais e amostras da série observada de GHI.....	60
Figura 15 – Metodologia para obtenção das previsões de GHI baseadas em imagens de satélite. Os processos que estão antes da terceira seta azul representam a estimativa de GHI. Daí em diante inicia a metodologia de previsão propriamente dita. ....	61
Figura 16 – Mapeamento dos pixels de Brasília selecionados por correlação para serem entradas dos modelos (ANN) dos horizontes: a) 1 passo à frente; b) 10 passos à frente. 62	
Figura 17 – Percentual acumulado de variância explicada (do conjunto inicial de variáveis) em função do número de componentes principais utilizado para a série de 10% de pixels usados no horizonte 1 passo a frente em Petrolina (2013 e 2014). ....	63
Figura 18 – Esquema usado para treinamento dos Modelos Estatísticos 2, baseados em imagens de satélite. A letra S representa uma variável sazonal. CP representa uma componente principal, <i>t</i> representa o tempo e <i>k</i> o horizonte de previsão.....	64

Figura 19 – Representação das entradas utilizadas para a realização da combinação de previsões de GHI. ....	66
Figura 20 – Coeficientes de autocorrelação parcial da série de temperatura horária de 2014 em Petrolina.....	67
Figura 21 – Erro dos diversos modelos autorregressivos de ordens variando entre 2 e 30 para dados de Petrolina.....	68
Figura 22 – Erros médios (RMSE) cometidos pelos 24 modelos de previsão de temperatura treinados com dados de Petrolina. ....	69
Figura 23 – <i>Improvement</i> em relação à Persistência dos 24 modelos de previsão de temperatura treinados com dados de Petrolina. ....	69
Figura 24 – Comparação entre os modelos TER (com variáveis sazonais) e os modelos TER-SVS (sem variáveis sazonais) para Pesqueira-PE. ....	71
Figura 25 – Estatísticos dos modelos de previsão baseados em dados de radiômetro de Brasília, Pesqueira e Petrolina. ....	72
Figura 26 – Precipitação média mensal de Pesqueira, Petrolina e Brasília, segundo dados da Normal Climatológica de 1981-2010, do INMET.....	74
Figura 27 – Insolação média mensal de Pesqueira, Petrolina e Brasília, segundo dados da Normal Climatológica de 1981-2010, do INMET.....	74
Figura 28 – <i>Improvement</i> (baseado no RMSE) calculado para os modelos proposto neste trabalho (TER BRB, TER PES, TER PTR e modelos de outros autores baseado em dados medidos por radiômetro. ....	75
Figura 29 – Estatísticos dos modelos de previsão (SAT) baseadas em imagens de Brasília, Pesqueira e Petrolina. ....	77
Figura 30 – Variação do Cosseno do ângulo zenital ao longo do ano para Pesqueira e Brasília. ....	78
Figura 31 – <i>Improvement</i> dos modelos aqui propostos de previsão via satélite (SAT) e do modelo de Perez, também baseado em imagens de satélite. ....	79
Figura 32 – Comparação entre tipos de modelos: resultados dos modelos baseados em dados de radiômetro (TER), dados de satélite e radiômetro (SAT) e somente em dados de satélite (PIX) para Petrolina. ....	80
Figura 33 – Comparação entre locais: resultados dos modelos baseados somente dados de satélite (PIX) para Brasília, Pesqueira e Petrolina.....	81
Figura 34 – Comparação entre o melhor modelo individual e a combinação de previsão para Petrolina.....	83

Figura 35 – Séries observada e prevista pela combinação de previsão para Petrolina (horizontes 1, 5 e 10 passos à frente). .....	85
Figura 36 – Dispersão entre a série observada e série prevista para Petrolina (horizontes 1, 5 e 10 passos à frente). .....	86
Figura 37 – Coeficientes da combinação MLR dos diversos horizontes para Petrolina. ....	87
Figura 38 – Skill de Taylor dos modelos SAT, TER e combinação MLR para Petrolina. ....	88
Figura 39 – (a) Correlação e (b) Razão entre os desvios-padrão da série do modelo em relação às observações. ....	89
Figura 40 – Erros médios (RMSE) cometidos pelos 24 modelos de previsão de temperatura treinados com dados de Pesqueira. ....	107
Figura 41 – Improvement em relação à Persistência dos 24 modelos de previsão de temperatura treinados com dados de Pesqueira. ....	107
Figura 42 – Erros médios (RMSE) cometidos pelos 24 modelos de previsão de temperatura treinados com dados de Brasília. ....	108
Figura 43 – Improvement em relação à Persistência dos 24 modelos de previsão de temperatura treinados com dados de Brasília. ....	108
Figura 44 – Comparação entre os modelos TER (com variáveis sazonais) e os modelos TER-SVS (sem variáveis sazonais) para Brasília. ....	109
Figura 45 – Comparação entre os modelos TER (com variáveis sazonais) e os modelos TER-SVS (sem variáveis sazonais) para Petrolina. ....	110
Figura 46 – Percentual acumulado de variância explicada (do conjunto inicial de variáveis) em função do número de componentes principais utilizado para a série de 10% de pixels usados no horizonte 1 passo a frente em Brasília (2013 e 2014). ....	111
Figura 47 – Percentual acumulado de variância explicada (do conjunto inicial de variáveis) em função do número de componentes principais utilizado para a série de 10% de pixels usados no horizonte 1 passo a frente em Pesqueira (2013 e 2014). ....	112

## LISTA DE TABELAS

Tabela 1 – Estatísticos utilizados para auxiliar a comparação entre modelos de previsão. ....	26
Tabela 2 – Modelos de Previsão de Temperatura Ambiente Baseados em Séries Temporais.	41
Tabela 3 – Aspectos relevantes de alguns dos principais trabalhos sobre previsão de GHI horária para o curto e médio prazo. ....	45
Tabela 4 – Características dos sítios estudados no presente trabalho.....	49
Tabela 5 – Especificações do Sensor <i>Imager</i> pertencente aos Satélites GOES 12 e GOES 1353	
Tabela 6 – Estatísticos de comparação das séries temporais de estimativas de GHI do pixel central das imagens de satélite com dados de radiômetro dos sítios Brasília, Pesqueira e Petrolina.....	55
Tabela 7 – Percentual acumulado de variância explicada por oito componentes principais para os diversos horizontes da série de Petrolina de 2013 e 2014.....	63
Tabela 8 – Improvement dos modelos TER (com variáveis sazonais) e TER-SVS (sem variáveis sazonais) para Pesqueira-PE. ....	71
Tabela 9 – Estatísticos das previsões baseadas em radiômetro (TER) relativas aos três sítios. ....	72
Tabela 10 – Descrição de modelos de previsão baseados em dados de radiômetros de outros autores.....	76
Tabela 11 – Estatísticos das previsões baseadas em satélite relativas aos três sítios. ....	77
Tabela 12 – Estatísticos dos modelos de radiômetro (TER), de satélite (SAT) e baseados no pixel central (PIX) de Petrolina. ....	81
Tabela 13 – Estatísticos dos modelos individuais (de satélite e de terra) e da combinação para Petrolina. A diferença de <i>Improvement</i> é calculada em relação ao modelo de radiômetro. ....	84
Tabela 14 – Comparação da Combinação de previsões com modelos de outros autores. ....	91
Tabela 15 – Improvement dos modelos TER (com variáveis sazonais) e modelos TER-SVS (sem variáveis sazonais) para Brasília.....	109
Tabela 16 – Improvement dos modelos TER (com variáveis sazonais) e modelos TER-SVS (sem variáveis sazonais) para Petrolina. ....	110
Tabela 17 – Percentual acumulado de variância explicada por oito componentes principais para os diversos horizontes da série de Brasília de 2013 e 2014.....	111
Tabela 18 – Percentual acumulado de variância explicada por oito componentes principais para os diversos horizontes da série de Pesqueira de 2013 e 2014. ....	112

## LISTA DE ABREVIATURAS E SIGLAS

AIC	Critério de Informação de Akaike ( <i>Akaike Information Criterion</i> )
ANEEL	Agência Nacional de Energia Elétrica
ANFIS	Sistemas de Inferência Fuzzy Baseado em Redes Adaptativas
ANN	Rede Neural Artificial ( <i>Artificial Neural Network</i> )
AR	Autorregressivo
ARIMA	Autorregressivo Integrado de Média Móvel ( <i>Autoregressive Integrated Moving Average</i> )
ARMA	Autorregressivo com Média Móvel ( <i>Autoregressive Moving Average</i> )
BRB	Brasília
CCI	Índice de Cobertura de Núvens ( <i>Cloud Cover Index</i> )
ECMWF	Centro Europeu de Previsões do Tempo de Médio Alcance ( <i>European Centre for Medium-Range Weather Forecasts</i> )
EOF	Funções Ortogonais Empíricas ( <i>Empirical Orthogonal Function</i> )
FV	Fotovoltaico
GHI	Irradiância Global Horizontal ( <i>Global Horizontal Irradiance</i> )
IV	Infravermelho
Kcs	Índices de Céu Claro
Kt	Índice de Claridade
MA	Média Móvel ( <i>Moving Average</i> )
MAE	Erro Absoluto Médio ( <i>Mean Absolute Error</i> )
MAPE	Erro Absoluto Médio Percentual ( <i>Mean Absolute Percentual Error</i> )
MLP	Perceptron Muticamadas ( <i>Multilayer Perceptron</i> )
MLR	Regressão Linear Múltipla ( <i>Multiple Linear Regression</i> )
MSE	Erro Médio Quadrático ( <i>Mean Squared Error</i> )
NWP	Previsão Numérica do Tempo ( <i>Numerical Weather Prediction</i> )
ONS	Operador Nacional do Sistema
PCA	Análise de Componentes Principais ( <i>Principal Components Analysis</i> )
PES	Pesqueira
PTR	Petrolina
RMSD	Raiz do Erro Médio Quadrático das Anomalias ( <i>Root Mean Squared Deviation</i> )
RMSE	Raiz do Erro Médio Quadrático ( <i>Root Mean Squared Error</i> )
SARIMA	Autorregressivo Sazonal Integrado de Média Móvel ( <i>Seasonal Autoregressive Integrated Moving Average</i> )
SEP	Sistema Elétrico de Potência
SONDA	Sistema de Organização Nacional de Dados Ambientais
SVR	Regressão por Vetor de Suporte ( <i>Support Vector Regression</i> )
TSI	Imageador de Céu ( <i>Total Sky Imager</i> )
UV	Ultravioleta
VIS	Visível
WRF	Modelo de Pesquisa e Previsão do Tempo ( <i>Weather Research and Forecasting</i> )

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>14</b>
<b>2</b>	<b>CONCEITOS PRELIMINARES .....</b>	<b>17</b>
<b>2.1</b>	<b>Previsão .....</b>	<b>17</b>
2.1.1	Horizontes de Previsão Relacionados à Radiação Solar.....	17
2.1.2	Estratégias de Previsão .....	18
2.1.2.1	<i>Estratégia de Previsão Recursiva .....</i>	<i>19</i>
2.1.2.2	<i>Estratégia de Previsão Direta.....</i>	<i>19</i>
2.1.2.3	<i>Estratégia de Previsão DirRec.....</i>	<i>20</i>
<b>2.2</b>	<b>Modelos Autorregressivos.....</b>	<b>20</b>
<b>2.3</b>	<b>Redes Neurais.....</b>	<b>22</b>
<b>2.4</b>	<b>Análise de Componentes Principais.....</b>	<b>24</b>
<b>2.5</b>	<b>Métricas de Avaliação dos Modelos de Previsão .....</b>	<b>25</b>
<b>3</b>	<b>REVISÃO DE LITERATURA .....</b>	<b>27</b>
<b>3.1</b>	<b>Estimativas de Irradiância Solar a partir de Imagens de Satélite .....</b>	<b>27</b>
<b>3.2</b>	<b>Previsão de Radiação Solar .....</b>	<b>31</b>
<b>3.3</b>	<b>Previsão de Temperatura.....</b>	<b>38</b>
<b>3.4</b>	<b>Combinação de Previsões aplicadas a Radiação Solar .....</b>	<b>41</b>
<b>3.5</b>	<b>Análise Crítica sobre o Estado da Arte de Previsão de Radiação Solar Global Horária em Curto e Médio Prazo.....</b>	<b>43</b>
<b>4</b>	<b>MATERIAL E MÉTODOS .....</b>	<b>47</b>
<b>4.1</b>	<b>Metodologia Proposta de Previsão de Irradiância e de Temperatura.....</b>	<b>47</b>
<b>4.2</b>	<b>Base de Dados .....</b>	<b>49</b>
4.2.1	Dados de Radiação Solar .....	51
4.2.2	Imagens de Satélite .....	52
4.2.3	Temperatura Ambiente .....	56
<b>4.3</b>	<b>Modelo de Referência.....</b>	<b>56</b>
<b>4.4</b>	<b>Modelos de Previsão Propostos .....</b>	<b>58</b>
4.4.1	Previsão Baseada em Dados de Radiômetro.....	58
4.4.2	Previsão Baseada em Imagens de Satélite .....	60

4.4.3	Previsão de Temperatura.....	65
4.4.4	Combinação de Previsões .....	65
<b>5</b>	<b>RESULTADOS E DISCUSSÃO.....</b>	<b>67</b>
<b>5.1</b>	<b>Modelo de Previsão de Temperatura.....</b>	<b>67</b>
<b>5.2</b>	<b>Modelo de Previsão de Radiação Solar baseado em Dados de Radiômetro.....</b>	<b>70</b>
<b>5.3</b>	<b>Modelo de Previsão de Radiação Solar baseado em Imagens de Satélite.....</b>	<b>76</b>
5.3.1	Modelo de satélite treinado com informações de Radiômetro como alvo (SAT)....	76
5.3.2	Modelo de satélite treinado com informações do pixel central como alvo (PIX)....	79
<b>5.4</b>	<b>Combinação de Previsões de Radiação Solar .....</b>	<b>82</b>
5.4.1	Análise Estatística da Complementariedade das Entradas da Combinação de Previsões .....	87
<b>5.5</b>	<b>Comparação dos Resultados Obtidos com a Literatura .....</b>	<b>89</b>
<b>6</b>	<b>CONCLUSÃO.....</b>	<b>93</b>
<b>7</b>	<b>PERSPECTIVAS .....</b>	<b>95</b>
	<b>REFERÊNCIAS .....</b>	<b>96</b>
	<b>APÊNDICE A – RESULTADOS PREVISÃO DE TEMPERATURA.....</b>	<b>107</b>
	<b>APÊNDICE B – INFLUÊNCIA DAS VARIÁVEIS SAZONAIS NAS PREVISÕES .....</b>	<b>109</b>
	<b>APÊNDICE C – RESULTADOS DAS ANÁLISES DE COMPONENTES PRINCIPAIS.....</b>	<b>111</b>

## 1 INTRODUÇÃO

Devido às dificuldades da exploração de novas usinas hidrelétricas de grande porte (TOMALSQUIM, 2017) e a preocupação com a dependência dos ciclos pluviais nas bacias hidrográficas que alimentam os reservatórios de tais usinas, cresce a busca por fontes alternativas para suprir o aumento da demanda por energia elétrica no Brasil. Nesse sentido, as fontes renováveis, em especial a eólica e a solar, vêm em vertiginoso crescimento de sua participação na matriz elétrica brasileira. Em 2016, as centrais fotovoltaicas representavam apenas 0,01% da capacidade de geração de energia elétrica no país. Já em 2019, segundo dados da ANEEL (Agência Nacional de Energia Elétrica) e da ABSOLAR (Associação Brasileira de Energia Solar Fotovoltaica), as centrais fotovoltaicas representam 1,2% da geração, com uma capacidade de 2074 MW, que, somando-se à micro e à mini-geração distribuída, totalizam 2759 MW. A expectativa realista (visto que esta capacidade já se encontra contratada) é que, em 2022, a potência instalada somente da geração fotovoltaica centralizada seja de 3700 MW, um crescimento de 78% em três anos (ABSOLAR, 2019).

Em alguns países europeus, a produção de energia fotovoltaica (FV) já atinge 30% da produção geral em dias de céu claro (PELLAND *et al.*, 2013). Na Alemanha, a produção FV já excede 40% da demanda elétrica nestes dias (PEREZ *et al.*, 2013).

Tendo em vista essa penetração crescente da energia fotovoltaica (fonte intermitente) no Sistema Elétrico de Potência (SEP) no Brasil, cresce a preocupação do Operador Nacional do Sistema (ONS) com o despacho das fontes em geral, posto que a intermitência (em escala horária) na saída das centrais fotovoltaicas pode acarretar dificuldades para a programação de fontes como, por exemplo, a fonte térmica (SOUSA, 2013).

Em 2008, devido à intermitência, o limite de inserção de energia renovável na rede elétrica da França foi estabelecido por ordem ministerial em 30% da potência instantânea (VOYANT *et al.*, 2014).

Com respeito à intermitência, existem basicamente três maneiras de abordar o problema: a primeira abordagem é o armazenamento de energia (LIU *et al.*, 2013), que se mostra como uma promessa para o futuro, porém ainda economicamente inviável para a realidade atual; a segunda abordagem busca tentar suavizar o problema através da integração geográfica das fontes baseada no efeito chamado de *Spatial Smoothing* (HOFF e PEREZ, 2010; UE *et al.*, 2011). A terceira abordagem baseia-se em previsões confiáveis da produção das centrais intermitentes em curto e médio prazo. Atualmente, esta abordagem apresenta-se como a solução mais imediata para o problema (COSTA *et al.*, 2008; MARTÍN *et al.*, 2010;

GIEBEL *et al.*, 2011), sendo, portanto, a abordagem adotada neste trabalho. A utilização combinada dessas abordagens também é possível. A associação de previsão de energia solar e armazenamento garante a disponibilidade de energia para a rede elétrica de distribuição. O armazenamento de energia absorve as fortes flutuações e excessos de geração e preenche as falhas de geração (VOYANT *et al.*, 2014).

O horizonte de previsão de maior interesse para o gestor da rede é o escrito como  $h+24$ , que representa 24 previsões hora a hora realizadas antes das 18h (VOYANT *et al.*, 2014). Nos Estados Unidos, o Operador Nacional do Sistema, chamado de CAISO (*California Independent System Operator*), utiliza duas previsões: a previsão um dia à frente e a previsão hora à frente. A previsão um dia à frente deve ser fornecida às 05h30 do dia anterior. Dessa forma, as previsões são fornecidas de 18,5h até 42,5h (18,5+24) antes. A previsão hora à frente deve ser fornecida com 105 minutos de antecedência e deve ter um alcance de 7h à frente. Dessa forma a previsão é na realidade de 1,75 a 8,75 horas à frente (PELLAND *et al.*, 2013).

Dentre os fatores ambientais que influenciam na potência de saída de uma central fotovoltaica, a irradiância solar incidente é o fator preponderante. Devido a isso, é grande o interesse do ONS em prevê-la. Diversos trabalhos foram publicados com respeito à previsão de radiação solar. Os trabalhos de DIAGNE *et al.* (2013), de VOYANT *et al.* (2017) e de INMAN *et al.* (2013) fazem uma profunda revisão dos modelos mais empregados nessa área.

Depois da irradiância, a temperatura do painel solar é a principal variável que influencia a potência FV (PELLAND *et al.*, 2013). Existem modelos simples que permitem obter a temperatura do painel com base na temperatura ambiente e na irradiância incidente. (GREEN, 1982; DUFFIE e BECKMAN, 2013).

O objetivo geral deste trabalho é desenvolver uma metodologia de previsão de irradiância global horária e um modelo de previsão de temperatura ambiente em curto e médio prazo (com horizonte de um dia à frente e *time-step* horário) capazes de auxiliar as previsões de potência de uma central fotovoltaica. Tais previsões devem fornecer aos atores envolvidos na operação do SEP subsídios em tempo hábil para minimizar os problemas de intermitência das fontes, os quais podem acarretar, entre outras consequências, a necessidade de maiores reservas de potência girante (ORTEGA-VAZQUEZ e KIRSCHEN, 2009), permitindo dessa forma a penetração de um percentual maior de geração fotovoltaica na matriz elétrica nacional.

Os objetivos específicos são:

1. Validar modelos de previsão univariados de temperatura ambiente média horária de simples calibração e operação;
2. Validar modelos de previsão de Irradiância Global Horizontal (GHI) com cada uma das características citadas:
  - a. Dados de radiômetro como entradas e que são calibrados tendo como referência dados de radiômetro.
  - b. Imagens de satélite como entradas e que são calibrados tendo como referência dados de radiômetro.
  - c. Imagens de satélite como entradas e que são calibrados tendo como referência GHI estimada pelas próprias imagens no pixel que contém o local em estudo.
3. Analisar o desempenho dos modelos de previsão de GHI horária em três locais (Brasília, Pesqueira e Petrolina), que possuem regimes de nebulosidade diferentes;
4. Avaliar a possibilidade de utilização de imagens de satélite em previsões para além das 6h à frente;
5. Realizar combinação de previsões de GHI de modelos que utilizam informações de naturezas diferentes (provindas de satélite e de radiômetro) a fim de obter uma melhoria na acurácia das mesmas.

O texto desta tese está estruturado da seguinte forma: na Seção 2 são apresentados alguns conceitos teóricos necessários ao entendimento do trabalho. Na Seção 3 é realizada uma revisão de literatura a respeito dos temas pertinentes a este trabalho: estimativas de irradiância a partir de imagens, previsão de radiação solar, previsão de temperatura e uma análise sobre as lacunas existentes em tal literatura. A Seção 4 apresenta as bases de dados de temperatura, irradiância e imagens de satélite e a metodologia proposta de previsão de GHI, bem como o modelo de previsão de temperatura ambiente. Na Seção 5 são analisados os resultados de cada modelo de previsão e da combinação deles, além de realizar a comparação com resultados de outros autores da literatura. Na Seção 6 é realizada a conclusão do trabalho e na Seção 7 são apresentadas as perspectivas de trabalhos futuros.

## 2 CONCEITOS PRELIMINARES

### 2.1 Previsão

Modelos de previsão são utilizados para extrapolar o comportamento do passado e do presente para o futuro. De forma concisa, a previsão (*forecast*) é o prognóstico de algum evento ou eventos futuros (MONTGOMERY *et al.*, 2015).

Do ponto de vista estatístico, previsão é o processo de estimativa em situações de incerteza. Do ponto de vista meteorológico, a previsão do tempo é a aplicação da ciência e tecnologia para fazer uma descrição detalhada de ocorrências futuras esperadas na atmosfera num certo local baseados em certos parâmetros atmosféricos e na habilidade e experiência de um meteorologista (QUADROS, 2005).

Os diversos tipos de procedimentos de previsão podem ser classificados como (CHATFIELD, 2013):

- Subjetivos: São geralmente baseados no julgamento por parte de especialistas da área. O método mais conhecido é o Método Delphi.
- Univariados: As previsões podem ser inteiramente baseadas em observações passadas de uma mesma série temporal, utilizando um modelo aplicado aos dados para extrapolação.
- Multivariados: As previsões são realizadas levando em consideração as observações de outras variáveis. Como exemplo, têm-se os modelos de regressão.

#### 2.1.1 Horizontes de Previsão Relacionados à Radiação Solar

Problemas de previsão normalmente são classificados de acordo com o seu alcance temporal. Denomina-se *horizonte de previsão* o número de períodos futuros para os quais devem ser produzidas previsões. Denomina-se ainda *intervalo de previsão* (ou *time-step*) a frequência com que as previsões são realizadas (MONTGOMERY *et al.*, 2015). Por exemplo, o ONS exige que as empresas geradoras entreguem previsões com horizonte de um dia à frente com *time-step* de meia hora.

Segundo INMAN *et al.* (2013), os horizontes de previsão estão associados com: a regulação (horizonte de segundos a minutos), o despacho de energia (horizonte de horas a dias), o planejamento da matriz elétrica (horizonte de anos). Não existe uma unanimidade em relação aos limites de cada horizonte.

Do ponto de vista da energia solar, horizonte de *curtíssimo prazo* é definido como o período de tempo de segundos a minutos, ao qual os meteorologistas comumente se referem como previsão imediata (nowcasting) (MCCANDLESS *et al.*, 2014). O horizonte de *curto prazo* se estende até 6 horas à frente. Já o *médio prazo* vai de 6 horas a 6 dias à frente (PEREZ *et al.*, 2010). Por fim, o *longo prazo* vai de meses a anos. A Figura 1 define as faixas de horizontes de previsão adotadas neste trabalho.

**Figura 1 – Divisão das faixas de horizontes de previsão adotadas no presente trabalho.**



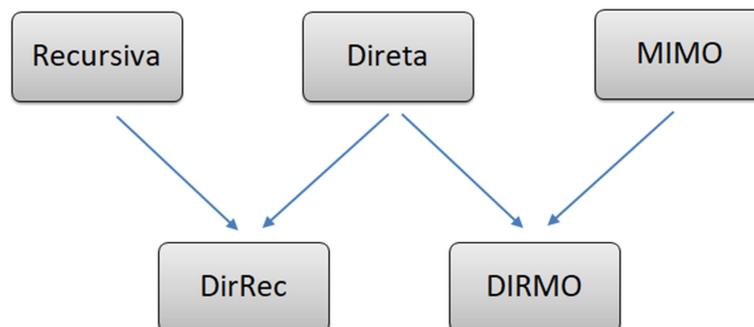
Fonte: O autor.

### 2.1.2 Estratégias de Previsão

Segundo TAIEB *et al.* (2012), a maioria dos pesquisadores de previsão dedica-se aos mais diversos aspectos a respeito de procedimentos de previsão, tais como seleção de modelo, eliminação da sazonalidade, combinação de previsão, entre outros. Entretanto ainda são poucos os trabalhos que investigam a geração de previsão multipassos à frente (KLINE e ZHANG, 2004).

Previsão  $H$  passos à frente consiste em prever os próximos  $H$  valores ( $y_{n+1}, \dots, y_{n+H}$ ) da série temporal ( $y_1, y_2, \dots, y_n$ ), onde  $H$  é o horizonte de previsão. As cinco estratégias de previsão mais difundidas na literatura são: Recursiva, Direta, DirReC (junção da Direta e Recursiva), MIMO (múltiplas entradas e múltiplas saídas) e DIRMO (junção da Direta e da MIMO) (TAIEB *et al.*, 2012). A Figura 2 mostra as cinco estratégias e as relações entre si.

**Figura 2 – Relação entre as estratégias de previsão.**



Fonte: (TAIEB *et al.*, 2012) (adaptado).

DE GIORGI *et al.* (2016) testaram o desempenho de três modelos de previsão de Potência de saída FV para vários horizontes até 24h combinados com diferentes estratégias de previsão multipassos à frente (Direta, Recursiva, DirRec). Com relação às estratégias, os autores concluíram que a estratégia DirRec obteve o melhor desempenho, enquanto que a estratégia Recursiva obteve o pior. O presente trabalho se limitará a discutir as três primeiras estratégias, Recursiva, Direta e DirRec, mesmo porque ainda não há um consenso entre os trabalhos publicados a respeito de qual das cinco estratégias é a que possui melhor desempenho (TAIEB *et al.*, 2012).

#### 2.1.2.1 Estratégia de Previsão Recursiva

Também chamada de Previsão Iterada ou Multiestágios; um único modelo é treinado para executar previsões de um passo à frente. Em seguida, o mesmo modelo é utilizado para as previsões seguintes, sendo que a previsão anterior passa a ser entrada do modelo e assim sucessivamente (SORJAMAA e LENDASSE, 2006). Um exemplo dado por COSTA (2014) da estratégia recursiva para três passos à frente e que considera os quatro últimos valores como preditores é dado pela Equação (1):

$$\begin{aligned}\hat{y}(t+1) &= f(y(t), y(t-1), y(t-2), y(t-3)) \\ \hat{y}(t+2) &= f(\hat{y}(t+1), y(t), y(t-1), y(t-2)) \\ \hat{y}(t+3) &= f(\hat{y}(t+2), \hat{y}(t+1), y(t), y(t-1))\end{aligned}\tag{1}$$

A grande crítica que se faz à estratégia recursiva é que a mesma é sensível ao acúmulo de erros conforme aumenta o horizonte de previsão (TAIEB *et al.*, 2012).

#### 2.1.2.2 Estratégia de Previsão Direta

Também chamada de Estratégia Independente, consiste em prever cada horizonte independente dos outros. Isso implica que essa estratégia não utiliza nenhum valor previsto para computar as previsões. Sendo assim, são treinados vários modelos (representados por  $f_H$  na Equação (2)), um para cada horizonte (TAIEB *et al.*, 2012).

Como exemplo, segue uma representação da estratégia de previsão direta (Equação (2)), que considera os quatro últimos valores como preditores (COSTA, 2014):

$$\begin{aligned}
\hat{y}(t+1) &= f_1(y(t), y(t-1), y(t-2), y(t-3)) \\
\hat{y}(t+2) &= f_2(y(t), y(t-1), y(t-2), y(t-3)) \\
\hat{y}(t+3) &= f_3(y(t), y(t-1), y(t-2), y(t-3))
\end{aligned} \tag{2}$$

Embora se tenha aumentado a complexidade do modelo (devido ao treinamento de três modelos,  $f_1$ ,  $f_2$  e  $f_3$ ), esta estratégia apresenta como vantagem em relação à Recursiva a eliminação do erro acumulativo.

### 2.1.2.3 Estratégia de Previsão DirRec

Trata-se de uma estratégia que combina as principais características de duas outras estratégias: a Direta e a Recursiva. “A estratégia DirRec calcula as previsões com modelos diferentes para cada horizonte (como a estratégia direta) e, a cada passo de tempo, amplia o conjunto de entradas, adicionando variáveis correspondentes às previsões do passo anterior” (TAIEB *et al.*, 2012).

Como exemplo, segue uma representação da estratégia de previsão DirRec (Equação (3)), que, a partir do instante  $t+2$ , utiliza como preditores valores previstos (além dos quatro últimos valores observados) (COSTA, 2014):

$$\begin{aligned}
\hat{y}(t+1) &= f_1(y(t), y(t-1), y(t-2), y(t-3)) \\
\hat{y}(t+2) &= f_2(\hat{y}(t+1), y(t), y(t-1), y(t-2), y(t-3)) \\
\hat{y}(t+3) &= f_3(\hat{y}(t+2), \hat{y}(t+1), y(t), y(t-1), y(t-2), y(t-3))
\end{aligned} \tag{3}$$

Conforme testada com dados observacionais nos trabalhos de DE GIORGI *et al.* (2016) e SORJAMAA e LENDASSE (2006), o desempenho da estratégia DirRec mostrou-se superior às estratégias Direta e Recursiva.

## 2.2 Modelos Autorregressivos

Séries observacionais que descrevem a evolução de um fenômeno físico com o tempo podem ser classificadas em dois grandes grupos: as séries temporais estacionárias e as séries temporais não estacionárias (INMAN *et al.*, 2013). Neste sentido, caso o processo estocástico

$[I(t)]$  na Equação (4) seja estacionário, as propriedades estatísticas de qualquer sequência finita  $I(1), I(2), \dots, I(m)$  de amostras de  $[I(t)]$  são semelhantes às da sequência  $I(1+h), I(2+h), \dots, I(m+h)$ , sendo  $h$  um número inteiro qualquer (NETTO, 2015).

Na década de 1920, foram publicadas duas abordagens para séries temporais que as tratam considerando sua natureza estocástica: autorregressão linear (AR) (YULE, 1921) e média móvel (MA) (YULE, 1927). Com respeito aos modelos AR, BOX e JENKINS (1976) os aplicam como previsores lineares para séries temporais. Por sua relativa simplicidade, os modelos AR têm sido muito utilizados para previsão em séries temporais de radiação solar. A ideia dos modelos AR baseia-se na extrapolação para um futuro próximo por meio de uma regressão linear sobre uma janela estreita formada por observações de um passado recente, incluindo-se o momento atual. Desloca-se essa janela por toda a série temporal para que se ajustem os parâmetros da função linear de forma adaptativa à dinâmica da variável observada (COSTA, 2005).

A origem do nome autorregressivo advém do fato de que o valor da série em um dado instante  $t+k$  depende de alguns (finitos) valores prévios do processo (até o instante  $t$ ) e do valor atual (no instante  $t$ ) do ruído inerente ao processo (INMAN *et al.*, 2013). O modelo AR fundamenta-se na hipótese de que as sucessivas observações de radiação  $I$  são fortemente correlacionadas entre si:

$$I(t+k) = \mu + b_0 I(t) + b_1 I(t-1) + \dots + b_{p-1} I(t-p+1) + e(t+k) \quad (4)$$

$$e(t+k) = I(t+k) - \hat{I}(t+k) \quad (5)$$

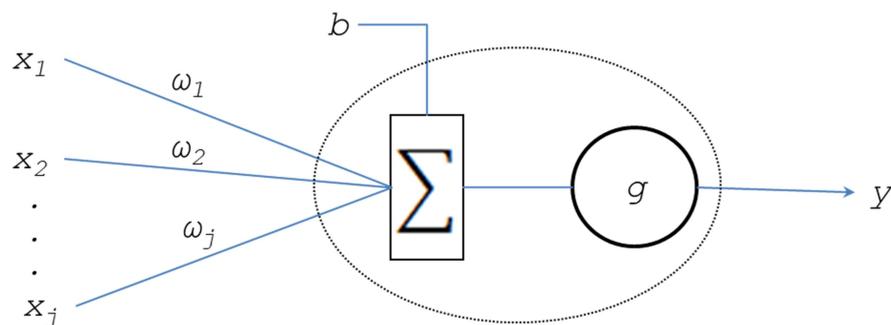
onde  $I(t+k)$  diz respeito à observação;  $\hat{I}(t+k)$  diz respeito à previsão para o instante  $t+k$ ;  $k$  é o horizonte de previsão;  $\mu$  diz respeito ao valor médio ou *bias* (viés) em caso de processo estacionário (em caso de processo não estacionário,  $\mu$  não se preserva constante com o tempo);  $p$  indica a ordem do processo autorregressivo;  $b_i$  ( $i=1 \dots p-1$ ) são os parâmetros do modelo a serem ajustados, por exemplo, de forma adaptativa (COSTA, 2005);  $e(t+k)$  (mostrado na Equação (5)) assume-se, tipicamente, como um ruído branco ( $[e(t)]$  é uma sequência de variáveis aleatórias independentes com função de densidade de probabilidade Gaussiana para a qual se tem média nula e variância  $\sigma_e^2$ ) (NETTO, 2015).

### 2.3 Redes Neurais

Redes Neurais (*Artificial Neural Network*, ANN) são modelos matemáticos inspirados na estrutura neural humana e que adquirem conhecimento através da experiência. Por possuírem característica não linear, são muitas vezes aplicadas com sucesso na solução de problemas que englobam relações não lineares complexas entre as variáveis.

O neurônio é a unidade básica de uma ANN. Semelhante ao correspondente biológico, os neurônios artificiais podem ser conectados uns aos outros para formar uma rede (INMAN *et al.*, 2013). A estrutura típica de um neurônio artificial é mostrada na Figura 3, onde  $x_1, x_2, \dots, x_j$  representam as entradas,  $\omega_1, \omega_2, \dots, \omega_j$  são os pesos sinápticos,  $b$  representa o *bias*,  $g$  é a função de ativação e  $y$  representa a saída.

**Figura 3 – Representação de um neurônio artificial.**



Fonte: (GALLEGO *et al.*, 2013) (adaptado).

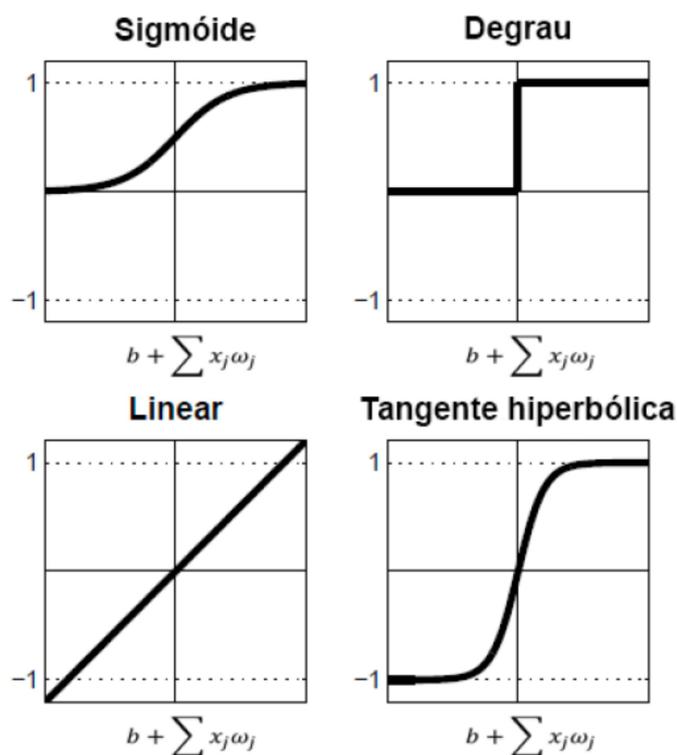
Cada entrada é multiplicada pelo seu respectivo peso sináptico e então entregue ao neurônio via elos de conexão. Em seguida, é realizada a soma dessas entradas com o *bias*. Posteriormente, esse resultado é introduzido em uma função de ativação  $g$ , onde a saída do neurônio é então gerada. O *bias* causa o aumento ou diminuição da entrada líquida da função de ativação  $g$  (HAYKIN, 1994). Pode-se modelar matematicamente o neurônio da Figura 3 através da Equação (6).

$$y = g\left(b + \sum_{i=1}^N \omega_i x_i\right) \quad (6)$$

A função de ativação  $g$  tem a tarefa de limitar a amplitude da saída do neurônio a um intervalo finito. É ela que tem o poder de atribuir ou não à rede neural a característica de não

linearidade, devendo ser escolhida de acordo com a aplicação da rede. Não é necessário que seja a mesma função para todos os neurônios. Algumas das principais funções de ativação são ilustradas na Figura 4.

**Figura 4 – Algumas das mais utilizadas funções de ativação em ANN.**

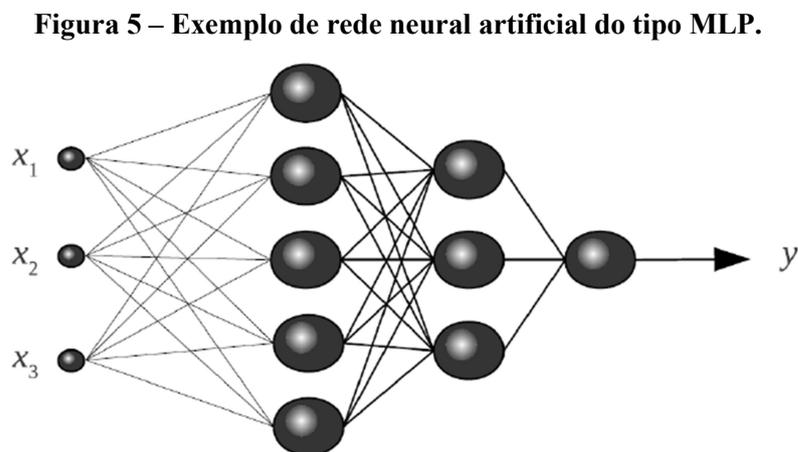


Fonte: (GALLEGO *et al.*, 2013) (adaptado).

Uma ANN é organizada de forma que os neurônios estejam interconectados e separados por camadas. A disposição dos neurônios nas camadas e o padrão de conexão entre estas definem a arquitetura da ANN. A Figura 5 ilustra um exemplo de ANN contendo doze neurônios, uma camada de entrada, duas camadas ocultas e uma camada de saída.

A característica marcante das ANN é a sua habilidade para aprender a partir de informações do seu ambiente, melhorando dessa forma o seu desempenho. Uma ANN aprende acerca de seu ambiente através de um processo iterativo de ajustes de seus pesos e níveis de *bias*. Denomina-se *algoritmo de aprendizagem* um conjunto preestabelecido de regras para a solução de um problema de aprendizagem. O treinamento (ajuste dos pesos) pode ocorrer de forma supervisionada (quando é feita a comparação com base nas saídas de referência para cada entrada, a fim de minimizar o erro) ou de forma não supervisionada (quando apenas os dados de entrada são utilizados) (HAYKIN, 1994).

A mais popular rede neural é a estrutura *Multi Layer Perceptron* (MLP), que consiste de uma camada de entrada, uma ou mais intermediárias e uma de saída (DIAGNE *et al.*, 2013). Um exemplo de MLP com duas camadas intermediárias (ocultas) é mostrado na Figura 5.



Fonte: (GALLEGO *et al.*, 2013).

## 2.4 Análise de Componentes Principais

A Análise de Componentes Principais (*Principal Components Analysis* - PCA) ou de Funções Ortogonais Empíricas (*Empirical Orthogonal Function* - EOF) é uma técnica bastante utilizada nas áreas de meteorologia, climatologia e oceanografia para compressão de dados e redução da dimensionalidade de um conjunto de variáveis regressoras (SMITH, 2002; MONAHAN *et al.*, 2009). De fato, a PCA é uma das técnicas da estatística multivariada mais empregadas para a eliminação de redundâncias entre as variáveis de um conjunto de variáveis regressoras. Consiste em projetar um conjunto de variáveis sobre uma base canônica sendo cada componente (de tal base) uma combinação linear de todas as variáveis originais. Após a projeção, ordenam-se todas as componentes (da nova base) em função da quantidade relativa de variância da componente com respeito à variância total do conjunto de variáveis originais (VARELLA, 2008). Logo, eliminam-se as componentes associadas às menores variâncias, preservando-se unicamente as componentes principais, isto é, preservando-se apenas as componentes que respondem pela maior parte da variância total do conjunto de variáveis originais (GUTIÉRREZ *et al.*, 2004).

Considerando-se o conjunto original de variáveis regressoras como aquele formado por uma série temporal de mapas de irradiância (gerados a partir de imagens de satélite) sobre

uma dada região, deve-se ter em mente que as variáveis regressoras costumam transportar em tal caso bastante informação linearmente redundante por conta da distribuição homogênea da irradiância sobre o espaço (pixels vizinhos). Associando-se cada *pixel* do mapa de irradiância a uma diferente variável regressora definida no domínio do tempo, deve-se buscar eliminar a redundância entre *pixels* (entre variáveis) antes do emprego de tais mapas como entrada a qualquer modelo, dando-se a eliminação através do processo chamado de “redução da dimensionalidade” do conjunto de variáveis regressoras. Ao se empregar a PCA para a citada redução da dimensionalidade no caso em questão, o primeiro passo a ser dado é o cálculo da matriz de covariância entre as variáveis originais (SMITH, 2002). Logo, calculam-se os autovetores e correspondentes autovalores da matriz de covariância, tendo-se em mente que cada autovetor representa uma componente (função empírica ortogonal) em uma base canônica e seu correspondente autovalor indica a quantidade de variância associada ao autovetor (quanto menor o autovalor, tanto menor é a variância associada ao respectivo autovetor). Ordenando-se, agora, os autovetores em função de seus correspondentes autovalores, é possível eliminar, um a um, os autovetores que menos contribuem a explicar a variância total do conjunto de variáveis regressoras. O truncamento desta matriz de autovetores é realizado de tal forma que os autovetores eliminados representam, tipicamente, apenas 5% da variância total do conjunto, isto é, mesmo com a eliminação de componentes, preservam-se aproximadamente 95% da variância total (GUTIÉRREZ *et al.*, 2004). Vale mencionar que, nos casos de variáveis homogeneamente distribuídas no espaço, a redução da dimensionalidade costuma ser da ordem de 10 para 1.

## 2.5 Métricas de Avaliação dos Modelos de Previsão

Algumas métricas estatísticas costumam ser utilizadas para análise de desempenho dos modelos de previsão, as quais estão descritas na Tabela 1.

O *Skill* de Taylor é um estatístico que reúne informação de outros três parâmetros estatísticos: a correlação, a razão entre os desvios-padrões do modelo e da observação e, por consequência, o erro (RMSD), já que os três são relacionados por uma equação (TAYLOR, 2001). Por ser um estatístico com maior riqueza de informações, o mesmo é muito útil na comparação (classificação) entre modelos submetidos às mesmas condições. Já para comparar com modelos de outros autores, os quais são submetidos a condições (variabilidade dos recursos) diferentes, é mais plausível realizar essa comparação através do *Improvement*. Por

esse motivo, foi adotado neste trabalho a apresentação de gráficos casados, um com o *Improvement* e, logo em seguida, outro com o *Skill*.

**Tabela 1 – Estatísticos utilizados para auxiliar a comparação entre modelos de previsão.**

ESTATÍSTICO	DESCRIÇÃO
$\sigma_x = \left[ \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right]^{1/2}$	Desvio Padrão de uma variável x qualquer, onde $\bar{x}$ é a média de x.
$RMSE = \left[ \frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i)^2 \right]^{1/2}$	Raiz do Erro Quadrático Médio entre o valor estimado (previsto) $\hat{x}$ e o valor observado x.
$nRMSE = \frac{RMSE}{\bar{x}} \times 100$	RMSE normalizado, onde $\bar{x}$ é a média de x.
$RMSD = \left[ \frac{1}{N} \sum_{i=1}^N [(\hat{x}_i - \bar{\hat{x}}) - (x_i - \bar{x})]^2 \right]^{1/2}$	Raiz do Erro Quadrático Médio dos desvios. $\bar{\hat{x}}$ representa o valor médio da série prevista.
$nRMSD = \frac{RMSD}{\bar{x}} \times 100$	RMSD normalizado, onde $\bar{x}$ é a média de x.
$Imp = 1 - \frac{RMSD_m}{RMSD_p}$	<i>Improvement</i> (ou <i>Forecasting Skill</i> ), onde os índices <i>m</i> e <i>p</i> são relativos ao modelo em questão e à persistência, respectivamente.
$R = \frac{\frac{1}{N} \sum_{i=1}^N (x_i + \bar{x})(y_i + \bar{y})}{\sigma_x \sigma_y}$	Coefficiente de Correlação entre as variáveis x e y.
$S = \frac{4(1 + R)^4}{\left(\frac{\sigma_m}{\sigma_o} + \frac{\sigma_o}{\sigma_m}\right)^2 (1 + R_0)^4}$	<i>Skill</i> de TAYLOR (2001), onde $R_0$ é a máxima correlação atingível e $\sigma_o$ e $\sigma_m$ são os desvios-padrões das séries observada e modelada, respectivamente.

Fonte: O autor.

Além dos estatísticos apresentados na Tabela 1, é comum o uso do MAE (*Mean Absolute Error*), que é a média das magnitudes dos erros de previsão. Já o RMSE dá mais ênfase aos valores de erro mais altos (PELLAND *et al.*, 2013). MADSEN *et al.* (2005) apud (PELLAND *et al.*, 2013) afirma que grandes erros são desproporcionalmente custosos, de forma que o uso do RMSE reflete melhor os erros de previsão para operadores de sistemas.

### 3 REVISÃO DE LITERATURA

Para aumentar a confiabilidade do Sistema Elétrico de Potência (SEP), são necessárias ferramentas de previsão de potência FV acuradas para diversos horizontes de tempo (INMAN *et al.*, 2013), os quais estão associados com a regulação da central (horizonte de segundos a minutos), o despacho de energia (horizonte de horas a dias) e os estudos de viabilidade e planejamento da matriz elétrica (horizonte de anos). Entre as aplicações citadas, este trabalho se centra no despacho da energia elétrica visando à proposição de uma nova metodologia de previsão de geração de uma central fotovoltaica cuja intenção é a de captar informações de boa acurácia (séries temporais observacionais de radiação solar) combinando com uma informação a respeito de um domínio espacial amplo associado às imagens de satélite.

As duas variáveis que mais influenciam a potência de saída de uma usina FV são a irradiância e a temperatura. Devido a isso, neste capítulo serão abordados os principais trabalhos que tratam da previsão de tais variáveis. Como neste trabalho serão utilizadas imagens de satélite, a Seção 3.1 é dedicada a trabalhos que realizam a conversão de níveis de cinza das imagens em irradiância. Já a Seção 3.4 faz uma explanação sobre as publicações de combinação de previsões voltadas para a área de energia solar.

#### 3.1 Estimativas de Irradiância Solar a partir de Imagens de Satélite

Dados de radiação geralmente não são disponíveis para todas as localidades. Mesmos sendo disponíveis, os dados podem ser de qualidade questionável (SAFI *et al.*, 2002).

Dados de radiação solar oriundos de sensores terrestres foram empregados para estimar radiação em locais vizinhos. Para isso, foram utilizadas a extrapolação e interpolação em diferentes distâncias. Os erros dessas estimativas foram comparados com os erros de estimativas realizadas por modelos baseados em imagens de satélite. Chegou-se à conclusão de que as estimativas por extrapolação e interpolação só obtiveram menor erro do que as estimativas de satélite para distâncias inferiores a 60 km (MARTINS e PEREIRA, 2011). Portanto, o uso de modelos físicos de transferência radiativa alimentados por dados de satélites representam a melhor solução para a estimativa da disponibilidade de energia solar em locais distantes das estações radiométricas (MARTINS e PEREIRA, 2011).

As imagens de satélite são capazes de fornecer um banco de dados de irradiância de 10 anos ou mais, bem como dados em quase tempo real (BLANC *et al.*, 2017).

"A construção da base de dados de radiação solar é muito útil para aplicações de energia solar, ambientais, agrícolas e outras. A utilização de dados de redes de satélites é uma forma econômica e eficaz de estimar a radiação solar e de construir uma base de dados solar" (ŞENKAL e KULELI, 2009).

Existem basicamente dois métodos de estimativa de radiação via imagens de satélite: um é pela obtenção do índice de nuvens (modelos empíricos e semiempíricos) e outro é pela modelagem da atmosfera através da transferência radiativa (modelos físicos) (BLANC *et al.*, 2017).

As incertezas dos métodos de estimativa de irradiância via imagens de satélite são da ordem de 15% a 20% para valores horários. Estes valores são também as incertezas de partida das previsões de irradiância via imagens (BLANC *et al.*, 2017).

As estimativas de irradiância global horizontal (GHI) baseadas em imagens de satélite oriundas do programa HelioClim-3 são fornecidas com resolução temporal de 15 minutos e RMSE de 25% em média (DAMBREVILLE *et al.*, 2014).

Existem duas grandes dificuldades quando se quer comparar medidas de satélite com medidas terrestres. A primeira é em relação à correta localização na imagem de satélite do local contendo o piranômetro. A segunda está relacionada com a diferença entre as duas técnicas de medidas. Medidas por satélite são fotografias instantâneas tiradas sob um pequeno ângulo sólido de visão. Já as medidas de terra são integrações temporais de medidas realizadas utilizando um grande ângulo sólido de visão (INMAN *et al.*, 2013).

Em relação às previsões de radiação solar obtidas a partir de imagens de satélite, quase todos os erros presentes no estado inicial (entrada) serão mantidos nas etapas de tempo previstas (saídas), as quais dificilmente terão incertezas menores do que as entradas, podendo apenas conseguir alguns efeitos de suavização que reduzam os erros de forma discreta (BLANC *et al.*, 2017).

TARPLEY (1979) propôs uma metodologia para estimativa da radiação solar na superfície terrestre baseada em imagens de satélite GOES. Trata-se de um modelo empírico que visa relacionar, além de outras variáveis de entrada, o brilho observado nas imagens e a nebulosidade com a radiação solar horária ( $\text{Wh/m}^2$ ) incidente na superfície terrestre. Para derivar tal regressão, foi escolhida uma região onde se dispunha de imagens de satélite e medidas de piranômetros. Foi definido um índice de nuvens dado por  $n = (0,5N_2 + N_3)/N$ , onde  $N$  é o número total de *pixels* da área estudada,  $N_2$  e  $N_3$  são o número de *pixels* com condições parcialmente nublada e com condições nubladas, respectivamente. Foram desenvolvidas três equações, uma para cada condição de cobertura de nuvem (céu claro,

parcialmente nublado e nublado), às quais são classificadas de acordo com o valor de  $n$ . Tais equações obtiveram correlações entre seus resultados e os valores observados de 0,94, 0,77 e 0,70 respectivamente para as três condições citadas anteriormente (MOURA, 1993). O erro encontrado por Tarpley (1979) foi de 10% na base diária.

GAUTIER *et al.* (1980) desenvolveram um modelo físico para estimação de radiação solar instantânea incidente na superfície a partir de medidas de brilho de imagens provenientes do satélite GOES, o qual pode ser aplicado tanto para condições de céu claro quanto para céu nublado. Esse mesmo modelo foi estudado por Arai e Almeida (1982), realizando estimativas de radiação solar instantânea para São José dos Campos-SP. Das quatro estimativas realizadas em horários diferentes do mesmo dia nublado, três atingiram diferenças menores ou iguais a 15% quando comparadas às observações do piranômetro. O modelo leva em consideração o espalhamento de Rayleigh e a absorção por vapor de água, mas o efeito das nuvens é o principal modulador da radiação solar.

RAPHAEL e HAY (1984) testaram o modelo estatístico proposto por HAY e HANSON (1978) para estimação de irradiância média horária na superfície  $GHI$ , a qual pode ser calculada pela Equação (7):

$$GHI = I_0(a - bR) \quad (7)$$

onde  $I_0$  é a irradiância no topo da atmosfera,  $R$  é a reflectância medida pelo satélite GOES no canal visível e  $a$  e  $b$  são coeficientes ajustáveis pela regressão. Esse modelo foi testado em ilhas no Oceano Atlântico Tropical. Quando comparados com medidas de piranômetros, os resultados apresentaram erros de 22% na estimativa horária e de 8% na estimativa diária (MOURA, 1993).

ZARZALEJO *et al.* (2005) investigaram um modelo híbrido para estimar irradiância baseado em imagens de satélite da Espanha com resolução de 7 km x 7 km e dados observacionais de piranômetros na superfície. Das imagens, são derivados índices de nuvens, os quais são utilizados em três métodos para transformá-los em irradiância solar. Essa abordagem está baseada em redes neurais e lógica difusa. Os autores concluíram que os modelos baseados em redes neurais obtiveram melhores resultados comparados ao modelo Heliosat I.

ŞENKAL e KULELI (2009) usaram ANN para estimar radiação solar (soma diária média mensal) a partir de imagens de satélite. Foram usados dados de 12 estações, sendo 9 para treinamento do modelo e 3 para testes. Os dados de entrada foram latitude, longitude,

altitude, mês, radiação difusa média e de radiação direta média. O RMS alcançado variou de 0,07% a 4,91%.

No Brasil, os institutos de pesquisa têm estudado com mais afinco dois modelos físicos para estimativa de irradiância solar baseada em informações de satélites: o modelo Brasil-SR e o modelo GL (MARTINS *et al.*, 2004).

O modelo GL trata-se de um modelo físico derivado do modelo de Gautier. Em sua versão atual, a GL1.2, são utilizados como entradas, além da reflectância visível e da cobertura relativa de nuvens “C”, ambas derivadas das imagens do satélite GOES, a água precipitável, a quantidade de ozônio e de dióxido de carbono. Então, o espectro solar é dividido em apenas duas largas bandas: o visível juntamente com o ultravioleta (UV + VIS) e o Infravermelho solar (IV). O modelo GL1.2 considera ainda três premissas: (1) Na estratosfera, onde o ozônio é abundante, existe apenas uma atenuação em relação à banda (UV+VIS), não existindo dispersão para esses comprimentos de onda; (2) Na troposfera, a banda (UV+VIS) não sofre atenuação, sofrendo apenas dispersão; (3) A absorção da banda IV é associada à presença de dióxido de carbono, vapor de água e gotas e cristais em nuvens. Além disso, o modelo GL 1.2 não considera a presença de aerossol na atmosfera (CEBALLOS e BOTTINO, 2004). Apesar dessas considerações, esse modelo obteve erros menores que 10% na média mensal em regiões com baixa quantidade de aerossóis.

O modelo Brasil-SR é um modelo físico para obtenção de estimativas da radiação solar incidente na superfície que resolve de forma aproximada a equação de transferência radiativa. Tal resolução aproximada baseia-se na aproximação de “Dois fluxos”. É assumido ainda que a radiação solar refletida que chega ao topo da atmosfera está linearmente distribuída entre duas condições: Céu Claro e Céu Extremamente Nublado (MARTINS *et al.*, 2004).

A equação principal do modelo Brasil-SR que permite calcular a estimativa de irradiância a partir do nível de cinza do pixel é dada por:

$$GHI = I_0 \cdot [(\tau_{clr} - \tau_{cld}) \cdot (1 - CCI) + \tau_{cld}] \quad (8)$$

onde  $I_0$  é a irradiância extraterrestre efetiva no topo da atmosfera.

Para obtenção das estimativas de radiação solar, é necessário realizar duas tarefas principais:

- O cálculo do índice de cobertura de nuvens (*Cloud Cover Index, CCI*) do pixel da imagem;

- A determinação das transmitâncias de céu claro ( $\tau_{clr}$ ) e de céu nublado ( $\tau_{cld}$ ).

“O modelo assume que a cobertura de nuvens é o principal fator de modulação do fluxo de radiação solar na superfície e que a transmitância atmosférica em qualquer condição de nebulosidade pode ser determinada pela relação linear entre a transmitância em condições de céu claro e em condições de céu totalmente encoberto” (MARTINS, 2001).

Para determinar o índice de cobertura de nuvens, é necessário conhecer os níveis de cinza máximo (céu nublado) e mínimo (céu claro) de cada pixel em cada hora do dia. Tais valores são encontrados fazendo uma busca separada por hora ao longo de um ano. Dessa forma, o CCI pode ser calculado pela Equação (9), onde  $L_r$  é o nível de cinza atual do pixel,  $L_{clr}$  e  $L_{cld}$  são os níveis de cinza de céu claro e de céu nublado, respectivamente.

$$CCI = \frac{L_r - L_{clr}}{L_{cld} - L_{clr}} \quad (9)$$

A segunda etapa do Brasil-SR consiste na determinação dos valores de transmitância atmosférica de céu claro e céu nublado. Por definição, transmitância representa a fração de radiação incidente que consegue atravessar um meio físico. No caso da radiação solar, a transmitância atmosférica ( $\tau$ ) é dada pela razão entre a irradiância solar global que chega na superfície horizontal e a irradiância no topo da atmosfera. Como citado anteriormente, no modelo Brasil-SR, a transmitância é estimada através de modelos físicos (transferência radiativa). Porém, em lugares onde se tem GHI medida em solo, a transmitância pode ser obtida através do simples cálculo da razão mencionada anteriormente. De posse dos valores de  $\tau$  calculados ao longo de todos os anos, detém-se os valores mínimo ( $\tau_{cld}$ ) e máximo ( $\tau_{clr}$ ) para cada hora e para cada mês. Utilizando os valores de CCI,  $\tau_{cld}$  e  $\tau_{clr}$  na Equação (8), é feita então a conversão de nível de cinza para GHI.

MARTINS *et al.* (2004) testaram o modelo Brasil-SR cuja validação para quaisquer condições de nebulosidade apresentou RMSE da ordem de 8% e MBE da ordem de 6% na base diária.

### 3.2 Previsão de Radiação Solar

Na literatura, existem vários modelos de previsão de GHI, a qual é a mais importante componente da radiação para sistemas sem rastreamento (LARSON e COIMBRA, 2018). Os primeiros métodos de previsão (regressivos), chamados de pré-estocásticos, usavam técnicas

de regressão comuns e tentavam apenas estabelecer relações matemáticas a partir de médias de longo prazo e valores de estado estacionário, resultando em modelos essencialmente estáticos que descrevem apenas mudanças sazonais e diurnas. Não consideravam a natureza estocástica dos dados de radiação (INMAN *et al.*, 2013). Esses modelos também desprezavam a natureza correlacionada entre tais dados, o que é uma característica marcante na radiação solar. Neles se encaixam os modelos de correlação e modelos baseados em médias de longo prazo.

Segundo MONJOLY *et al.* (2017), os modelos de previsão solar são divididos de acordo com o tipo de dado de entrada, os quais também determinam o horizonte de previsão.

Previsões via TSI (*Total Sky Imager*) tem alcance da ordem de 10-30min à frente. Já as imagens de satélite tem alcance de até 5h (PEREZ *et al.*, 2010). Devido a baixa resolução temporal e espacial, as previsões via Imagens de satélite são menos acuradas do que as de TSI para a intra-hora. Porém, a maior cobertura espacial abrangida pelas imagens de satélite favorece a projeção do movimento do campo das nuvens para horizontes mais à frente (PELLAND *et al.*, 2013).

Métodos de imagens de satélite são uma alternativa quando comparados aos dispendiosos métodos de sensores de piranômetros terrestres. Além disso, permitem previsões em lugares onde nenhum outro tipo de dado está disponível. Estes métodos sofrem da limitação de horizontes de tempo, podendo apenas ser aplicados do intra-hora ao intra-dia (INMAN *et al.*, 2013).

Métodos de previsão de irradiância por imagens de satélite geralmente atingem rRMSD de 10% para horizontes de 30 min e vão a 26% para horizontes de 6h à frente (HAMMER *et al.*, 2001).

O modelo baseado em vetor de movimento possui apenas a variabilidade espacial, pois o mesmo considera que o campo espacial de nuvens no instante  $t+k$  corresponde ao campo de nuvens do instante  $t$ , só que transladado (BLANC *et al.*, 2017).

Para previsões 1h à frente, as previsões via vetores de movimento e a persistência se equivalem, provavelmente devido aos erros de paralaxe e de navegação (localização) do satélite (PELLAND *et al.*, 2013).

Existem atualmente duas grandes linhas de pesquisa no tocante aos modelos de previsão aplicados à geração fotovoltaica (MARTÍN *et al.*, 2010): a) modelos físicos; b) modelos estatísticos. Os modelos físicos são baseados em equações dinâmicas (leis físicas de conservação) que descrevem o comportamento da atmosfera (KALNAY *et al.*, 1996; SIMMONS e GIBSON, 2000; SKAMAROCK *et al.*, 2008). Já, os modelos estatísticos são

baseados na construção de relações empíricas entre as variáveis predictoras e o preditando. Tais modelos (estatísticos) subdividem-se ainda em: modelos baseados em séries temporais (BOX, 1976), modelos baseados em imagens de satélite (PEREZ *et al.*, 2010) e modelos de *downscaling* estatístico. Estes últimos, empregados para o aumento da resolução espacial da saída dos modelos físicos (WILBY e WIGLEY, 1997).

Em relação aos modelos físicos, as equações não possuem solução única devido às suas não linearidades. Por isso, recorre-se às soluções por aproximação numérica (MARTÍN *et al.*, 2010). São os chamados modelos NWP (*Numerical Weather Prediction*). Os erros destes modelos variam significativamente e dependem, entre outros aspectos, do clima e dinâmica da radiação solar no local em estudo (LORENZ *et al.*, 2009). A respeito da acurácia dos NWPs, MARTÍN *et al.* (2010) indicam que, para um caso particular, os erros médios para a previsão horária de GHI variam entre 20,8% e 31,7% para as primeiras 24 horas e de 21,3% a 36,8% no segundo dia. LIMA (2015) avaliou o desempenho de alguns modelos de previsão de irradiância no Nordeste do Brasil, encontrando um RMSE de 40% nas previsões de GHI horária do modelo numérico WRF (*Weather Research and Forecasting*).

Os modelos estatísticos são baseados na construção de relações empíricas entre preditores e preditando. Em relação aos modelos baseados em séries temporais, diversas abordagens vêm sendo estudadas de forma a alcançar resultados mais acurados. As abordagens mais recorrentes nesse sentido são os modelos Autorregressivos, as Redes Neurais e os modelos baseados em Lógica Difusa, conforme enfatizados nos trabalhos de COSTA (2005) e de DIAGNE *et al.* (2013).

Uma das maneiras de se melhorar as previsões de radiação solar é através da combinação de previsões. Isso vem sendo bastante utilizado em eólica. Já em solar, CHAOUACHI *et al.* (2010) obtiveram melhorias nas previsões simplesmente realizando uma média das previsões de três modelos distintos (PELLAND *et al.*, 2013).

Em algumas situações, é conveniente se fazer uma previsão de radiação em uma área ao invés de previsões pontuais. Previsões de área de produção de energia de sistemas fotovoltaicos são importantes para os operadores de sistemas elétricos de potência. As previsões de área podem ser realizadas usando dados de todos os sistemas instalados em uma determinada área ou usando uma técnica de *upscaling*, que consiste na representação da potência FV total de todos os sistemas instalados em uma dada área usando dados de um subconjunto representativo daqueles sistemas. Quanto maior a área, melhores são os resultados. Os erros atingidos em previsões por área são geralmente menores do que os erros

de previsões pontuais, com reduções de RMSE da ordem de até 67% (PELLAND *et al.*, 2013).

LEWIS (1982) apud (YADAV e CHANDEL, 2014) classificou a acurácia das previsões de radiação solar utilizando o MAPE (*Mean Absolute Percentual Error*), de forma que as previsões de alta acurácia teriam  $MAPE < 10\%$ , as previsões de boa acurácia possuiriam  $10\% < MAPE < 20\%$ , as previsões de acurácia ruins teriam  $20\% < MAPE < 50\%$  e as previsões de péssima acurácia teriam  $MAPE > 50\%$ .

SFETSOS e COONICK (2000) conseguiram uma melhoria de 71% e 74% em relação à persistência para previsão de GHI horária uma hora à frente usando ANN com algoritmos de treinamento retropropagação e Levenberg-Marquardt, respectivamente, sendo que o último contou com a adição de variáveis exógenas. Os dados utilizados são de 63 dias com predominância de céu claro.

SAFI *et al.* (2002) afirmaram que a tendência sazonal da radiação solar pode ser extraída de 3 formas: a) transformar em  $K_t$ ; b) subtrair o primeiro harmônico da série de Fourier e c) subtrair de  $I_{0h}$  (GHI no topo da atmosfera) o valor de GHI (componente perdida). A fim de comparar, os autores realizaram previsões via modelo MA(2) (*Moving Average*) de valores diários de  $K_t$  e de componente perdida, obtendo nRMSE de 16% e 28% respectivamente para horizonte de um passo à frente.

CHAABENE e BEN AMMAR (2008) propuseram uma previsão de irradiância ( $W/m^2$ ) utilizando técnicas de redes Neuro-fuzzy e ARMA (*Autoregressive Moving Average*) associados a Filtros de Kalman. O primeiro passo é a previsão do padrão (valores máximos e mínimos) de Irradiância e temperatura para o dia a ser previsto. Em seguida faz-se uma previsão de curtíssimo prazo (5 min à frente) utilizando tais padrões e dados medidos instantâneos. Os resultados alcançaram nRMSEs inferiores 10% para a irradiância.

REIKARD (2009) testou várias ferramentas (Regressão, ARIMA, Função de Transferência, ANN e modelos híbridos) aplicadas à previsão de irradiância para horizontes de 5 minutos até 4 horas à frente. Os modelos foram testados em seis diferentes sítios dos EUA. Para horizontes de 1h a 4h à frente, obteve-se uma melhoria em relação à persistência de 29% a 8,5% respectivamente para o sítio de Clark, Nevada. Já o MAE foi de 19,6% a 30,9%. O modelo que alcançou melhores resultados nesses horizontes foi o ARIMA, em contraposição a outros trabalhos, que apontam a ANN como melhor ferramenta.

MARTÍN *et al.* (2010) propuseram um modelo de previsão solar de uma metade diária (correspondente a 6 horas de radiação solar) com *time-step* de 6 horas e 6 horizontes de previsão à frente, totalizando 3 dias. Foram utilizados como entradas dados de Índice de

Claridade (Kt) ou de Componente Perdida (diferença entre a irradiância solar horizontal no topo da atmosfera e a GHI que chega à superfície terrestre). Os modelos empregados foram Autorregressivos, Redes Neurais com retropropagação e Sistemas de Inferência Fuzzy baseado em redes adaptativas (ANFIS). O melhor modelo foi o que utilizou redes neurais. Os erros (rRMSE) variaram de 20% para o primeiro horizonte até 30,3% para o último horizonte. Já a melhoria variou de 4,84% no primeiro horizonte até 25,58% no último horizonte.

MELLIT e PAVAN (2010) propuseram um modelo que previsse hora a hora a irradiância solar do dia seguinte utilizando como entradas apenas a irradiância solar média diária e a temperatura média diária. Usando ANN com 24 neurônios na camada de saída (um para cada hora do dia), conseguiu-se uma correlação na faixa de 98-99% para dias de céu claro e de 94-96% para dias nublados. Já os RMSE ficaram entre 13% a 67% para dias de céu claro e 54% a 85% para dias nublados.

PEREZ *et al.* (2010) realizaram previsões de curto (até 6h) e médio (1 a 7 dias) prazos de GHI horária utilizando dois modelos, um que usou como entrada imagens de satélite e outro que usou saídas de NWP, respectivamente. O primeiro método utiliza índice de céu claro e vetor de movimento das nuvens. Foram realizadas previsões para diversos sítios, desde sítios desérticos (Desert Rock) até sítios úmidos (Penn State). Para o sítio Fort Peck, o nRMSE foi de 26,3% (1h) até 41,1%(6h). Já as previsões por NWP obtiveram nRMSE de 40,6% (1 dia à frente) a 48,1% (7 dias à frente).

MARQUEZ e COIMBRA (2011) propuseram um modelo de previsão de GHI e DNI (*Direct Normal Irradiance*) que utiliza como entradas variáveis geométricas ( $\cos(\theta_z)$ ), variável temporal (hora angular) e previsões NWP de temperatura, ponto de orvalho, umidade, velocidade do vento, entre outras. As saídas do modelo são previsões horárias de GHI e DNI até 6 dias à frente. Foi aplicado o teste gamma para seleção de variáveis. A base de dados utilizada era composta de 13 meses de dados. O melhor modelo de previsão de GHI atingiu um nRMSE de 17,7% e um  $R^2=0,94$  para o mesmo dia e um RMSE de aproximadamente 21% para o dia seguinte.

WANG *et al.* (2011) utilizaram ANN MLP (*Multilayer Perceptron*) com retropropagação para previsão de irradiância solar para uma hora à frente. Os dados de entrada consistem das últimas 24 observações de irradiância solar horária, as quais passaram por um processo de normalização que possui um intervalo de [0,1; 0,9]. Segundo os autores, diferentemente do intervalo usual [0; 1], este intervalo garante a convergência das redes. O autor observou que após a 24ª época de treinamento, o MSE, que vinha diminuindo, cresceu levemente (*over fitting*). Os resultados alcançaram um nRMSE de 1,19% e um  $R^2$  de 0,996.

FU e CHENG (2013) estudaram previsão de irradiância solar para horizontes de 5 a 15 minutos à frente utilizando características extraídas de TSI como número de pixels nublados, brilho médio, média e variância do número de pixels nublados, etc. O modelo produziu como saídas valores de  $K_t$ , os quais posteriormente foram transformados em GHI. O modelo de previsão proposto é baseado em técnicas de regressão. O RMSE variou de 24,6% a 27,5% para horizontes de 5 a 15 minutos, respectivamente.

VOYANT *et al.* (2014) realizaram um estudo sobre previsão de radiação solar [ $\text{Wh/m}^2$ ] utilizando como entrada 13 anos de radiação global no plano horizontal na região de Córsega, França. Como as centrais FV normalmente não são construídas sem inclinação (*tilt*), foi utilizado um método para projetar a radiação horizontal global para o plano inclinado da central. Os dados foram transformados em Índices de Céu Claro ( $K_{cs}$ ), de forma a obter uma nova série estacionária. Como  $K_{cs}$  é a razão da radiação medida pela radiação de céu claro, um pequeno erro proveniente do modelo de céu claro adotado é inserido na série temporal. Inicialmente foram realizadas previsões com modelos simples como persistência, ARMA e MLP. As previsões resultantes dos dois últimos modelos atingiram nRMSE semelhantes, da ordem de 40% para um horizonte de 24h. Em relação ao modelo ARMA e tendo em vista a alta variabilidade da radiação solar, os autores, baseados em um trabalho anterior, escolheram uma ordem muito baixa (AR(1)), além de não utilizar o conceito de MA (*Moving Averagem*). Além disso, optaram por modificar a série, gerando uma série de entrada composta pelos valores de radiação solar da mesma hora desejada de todos os dias. Foram utilizadas as técnicas de Informação Mútua para realizar seleção de variáveis exógenas a serem incorporadas ao sistema híbrido. Foram também utilizadas algumas regras bayesianas para determinar a melhor composição entre os modelos estudados, de forma a produzir um modelo híbrido com previsões mais precisas, que alcançou um nRMSE de 36% para as 24 previsões à frente na base horária. Os autores concluíram que o uso das regras de Bayes não é interessante.

DONG *et al.* (2014) propuseram um modelo híbrido não recursivo de previsão baseado em imagens de satélite e dados observacionais de irradiância solar. Das imagens, foram derivadas estimativas de índices de cobertura de nuvens, formando assim uma série temporal estimada de tais índices. Utilizando essa série, foram aplicadas técnicas para prever índice de cobertura de nuvens. O passo seguinte foi transformar o índice de cobertura de nuvens previsto (futuro) em irradiância solar prevista. Já, os dados observacionais foram usados para treinar e validar as redes neurais. Os erros (nRMSE) do modelo híbrido foram calculados mês a mês e variaram de 28 a 43%, dependendo do mês.

DAMBREVILLE *et al.* (2014) propuseram utilizar em horizontes de curto prazo (até uma hora à frente) um modelo AR que utiliza entradas de naturezas distintas: um dos conjuntos de entradas trata-se de índices de céu claro do presente e do passado obtidos de pixels da vizinhança via satélite. O outro conjunto de entradas constitui-se de observações presentes e passadas medidas por piranômetro. Dessa forma foi considerada a variabilidade espacial e temporal. A saída foi o índice de céu claro no tempo  $t+k$ . Os resultados alcançaram um rRMSE de 28% (1 hora à frente), uma correlação de 0,909 e uma melhoria de 19% em termos do RMSE comparados ao modelo AR com apenas dados de piranômetro como entrada.

MCCANDLESS *et al.* (2014) treinaram modelos ANN, Persistência e autorregressivos para horizontes de 30 min até 3 horas com *time-step* de 15 minutos usando dois anos de dados, os quais foram divididos em três condições de céu: nublado, parcialmente nublado e céu claro. Foram treinados tanto modelos específicos para cada condição de céu quanto modelos gerais (qualquer condição), onde se observou que estes últimos obtiveram erros praticamente iguais aos primeiros. A melhoria dos modelos gerais para horizontes de 1h até 3h foi de até 13% em relação à persistência. Já para céu claro, a persistência mostrou-se melhor do que os demais.

LIMA (2015) realizou um estudo no nordeste brasileiro que comparou previsões de irradiância de modelos numéricos de mesoescala (WRF) com resolução temporal de 6 horas e horizonte de 30 horas de previsão com modelos baseados em ANN e MLR (*Multiple Linear Regression*), que tinham como preditores as próprias saídas do modelo WRF. Enquanto as previsões do modelo WRF atingiram RMSE de 37% a 45% (dependendo do sítio), os modelos refinados por MLR atingiram RMSE de 29% a 33%. Já os modelos refinados via ANN obtiveram RMSE de 22% a 30%.

AGUIAR *et al.* (2016) propuseram um modelo de previsão de GHI horária para horizontes de 1h a 6h à frente utilizando dados medidos em terra (Ilhas Canárias, Espanha) e dados exógenos provindo de satélite (GHI) e previsões de Índice de Cobertura de Nuvens e radiação solar provindos do ECMWF (*European Centre for Medium-Range Weather Forecasts*). Foram realizadas comparações entre os modelos Persistência, ANN baseada em observações (modelo chamado de NN), saídas do ECMWF e dados de satélite (chamados de SAT), como também diversas combinações desses modelos. O melhor resultado foi alcançado pelo modelo combinado NN+SAT+ECMWF (nRMSE=24% para h=1h e nRMSE=34% para h=6h), o qual obteve uma melhoria máxima (h=6h à frente) de 3,5% em relação ao modelo NN, o que não é muito significativo tendo em vista o aumento da complexidade do modelo.

MONJOLY *et al.* (2017) propuseram uma abordagem para previsão de radiação solar uma hora à frente que, após transformar radiação em índice de céu claro, aplicaram métodos de decomposição multiescala, obtendo dois sinais, onde o sinal de escala de tempo pequena foi previsto usando ANN e o sinal de escala de tempo grande foi previsto usando AR. Os dois sinais previstos foram então agregados e posteriormente transformados em radiação solar. O método de decomposição que obteve melhor resultados foi a decomposição por Wavelet. Os erros (RMSE) alcançados pelo melhor modelo proposto foi de 7,86%.

(BAE *et al.*, 2017) realizaram previsão de irradiância solar uma hora à frente utilizando *Support Vector Machine* (SVM) e tendo como entradas temperatura, umidade, janela de brilho solar, duração do brilho solar, cobertura de nuvem, tipo de nuvem, precipitação e velocidade do vento. Foi também abordada uma técnica de dimensionamento de sistema de armazenamento de energia baseada na análise do erro cometido pela previsão. Os dados foram agrupados segundo técnica de clusterização. Com dados não clusterizados, os melhores resultados obtidos com SVM alcançaram um RMSE= 58 W/m<sup>2</sup> e uma correlação de 0.95. Já com dados clusterizados, os melhores resultados (divididos por tipo de dados, ensolarados, parcialmente nublado e nublado) foram: correlação de 0.96, 0.94 e 0.91 e RMSE de 49, 62 e 57 W/m<sup>2</sup>, o que evidenciou um baixo ganho pelo uso da clusterização.

QING e NIU (2018) propuseram um modelo de previsão que utiliza como entradas previsões NWP de temperatura, ponto de orvalho, umidade, visibilidade, velocidade do vento, entre outras, cujas saídas foram previsões horárias de GHI até um dia à frente. A base de dados consiste em 30 meses de dados coletados na ilha de Cabo Verde. A ferramenta utilizada para previsão foram as redes de memória de curto prazo (*Long Short-Term Memory*, LSTM). O *Improvement* (ou *Forecasting Skill*) geral para todos os horizontes do modelo LSTM foi de 30,6%. Já para as ANN com retropropagação, o *Improvement* obtido foi de 15,25%.

### 3.3 Previsão de Temperatura

As principais variáveis que influenciam a potência FV são a irradiância no plano do coletor e a temperatura do painel fotovoltaico (PELLAND *et al.*, 2013).

Existem vários modelos de previsão de temperatura ambiente. A grande maioria provém de ferramentas de previsão de tempo numérica (NWP), as quais são calculadas para pontos específicos de uma malha que cobre uma vasta região, possuindo uma baixa resolução espacial. Além disso, para horizontes de curto prazo, modelos baseados em séries temporais

tendem a ter melhores resultados comparados ao NWP. A maioria das ferramentas possuem apenas grandezas médias, máximas e mínimas diárias. Poucos modelos de previsão de temperatura ambiente oferecem saídas com intervalos de tempos horários.

Vários modelos de previsão de temperatura baseados em séries temporais de dados observacionais foram propostos na literatura. A Tabela 2 traz um resumo das principais características de alguns desses modelos.

MCCUTCHAN (1976) utilizou expansão por série de Fourier para prever a variação diurna da temperatura em terras montanhosas do Sudeste da Califórnia. A expressão permite calcular a temperatura para qualquer instante de tempo do dia. Os valores de suas constantes dependem da insolação, da hora do dia, da classe de tempo sinótica e da altitude. Os resultados para horizontes de tempo de 2h a 11h em avanço alcançaram faixas de RMSE variando de 1,3°C a 4,7°C e coeficientes de correlação em torno de 0,90.

TASADDUQ *et al.* (2002) utilizaram um modelo baseado em redes neurais para previsão da temperatura média horária com um horizonte específico de 24h à frente, isto é, previsão da temperatura no dia seguinte e na mesma hora. A entrada do modelo é unicamente a temperatura da hora de interesse. O RMSE alcançado pela previsão foi de 1.75°C.

ABDEL-AAL (2004) propôs uma abordagem de previsão de temperatura baseado no conceito de redes abduativas, as quais são mais simples e mais automatizadas em relação aos demais métodos de entrada-saída. As entradas dos modelos eram: temperaturas horárias das 24 horas do dia ( $d-1$ ), temperatura máxima e mínima do dia ( $d-1$ ) e previsões da temperatura máxima e mínima do dia ( $d$ ) ao qual se quer prever. As saídas do modelo são temperaturas horárias 24h à frente com *time step* horário. Foram testadas tanto estratégias diretas quanto recursivas. Os resultados obtiveram MAE médio ao longo dos 24 horizontes de previsão de 1,06 °C para previsões diretas sem a utilização de previsões de temperaturas máxima e mínima.

SMITH *et al.* (2006) propuseram uma metodologia de previsão de temperatura do ar em intervalos horários de 1h a 12h à frente baseados em Redes Neurais (ANN). As entradas do modelo são temperatura e variáveis sazonais determinísticas. Tal metodologia visa à melhoria da acurácia das previsões baseada em quatro ações: aumento do número de padrões de inicialização de treinamento, inclusão das variáveis com informação sazonal, extensão da duração dos dados a priori e variação do número de camadas intermediárias. É mostrado que o mesmo modelo quando iniciado com diferentes pesos iniciais aleatórios produzem erros de previsão diferentes. Portanto, sugere-se que os desenvolvedores treinem suas redes com 30 diferentes pesos iniciais à fim de estabelecer os melhores parâmetros. Em relação à variação

do tamanho do conjunto de dados para treinamento, não foi encontrada relação direta da mesma com o MAE. Já a inclusão de variáveis exógenas promoveu apenas discreta melhoria no MAE, que passou de 1,51°C para 1,48°C. O aumento da ordem do modelo provocou uma melhoria no MAE. Considerando uma janela de 24 observações passadas, obteve-se o melhor MAE (1,38°C). A última melhoria foi devido à variação de nós intermediários da ANN, a qual obteve melhor resultado com 120 nós e (já considerando as melhorias anteriores) alcançou um MAE de 1,35°C.

HAYATI e MOHEBI (2007) realizaram previsões de temperatura em curto prazo (um dia à frente com *time step* de 3h) utilizando redes neurais. Foram utilizados 10 anos de observações. As entradas do modelo são dados meteorológicos (temperatura de bulbo seco e temperatura de bulbo úmido, velocidade do vento, pressão, umidade relativa do ar, brilho solar e radiação solar). O trabalho obteve um MAE de 1,8°C.

CHAABENE e BEN AMMAR (2008) realizaram previsão de irradiância e temperatura utilizando técnicas de redes Neuro-fuzzy e ARMA associados a Filtros de Kalman. O primeiro passo foi a previsão dos valores máximos e mínimos de Irradiância e temperatura para o dia a ser previsto. Em seguida faz-se uma previsão de curtíssimo prazo (com *time-steps* de 5 min) utilizando tais previsões e dados medidos instantâneos. Os resultados apresentaram nRMSEs entre 0,87% e 2,28% (dependendo do dia) para a temperatura.

CHEVALIER *et al.* (2011) propuseram um método de previsão de temperatura baseado em SVR (*Support Vector Regression*). Com objetivo de reduzir o número de padrões (linhas de dados) de treinamento de um grandioso conjunto de dados e, conseqüentemente o tempo de treinamento do modelo, foi proposto um método para reduzir o conjunto de treinamento. A previsão foi realizada para horizontes de 1h a 12h, com MAE variando de 0,5°C a 2,3°C quando apenas dados dos invernos foram utilizados e de 0,5°C a 1,9°C quando os dados usados foram do ano todo. Para dados apenas dos invernos, o método SVR foi discretamente melhor que a ANN (SMITH *et al.*, 2006). Já para dados de todo o ano, o método baseado em SVR superou a ANN em apenas 4 dos 20 horizontes.

Tabela 2 – Modelos de Previsão de Temperatura Ambiente Baseados em Séries Temporais

Assunto	Autor	Modelos usados	Horizontes	Erros	Quantidade de dados	Local	OBS
Previsão temperatura para H=24h	Tasadduq et al. (2002)	ANN	24h	1,75°C (RMSE)	1 ano (trein.) e 3 anos(valid.)	Arábia Saudita	Univariado
Previsão temperatura para H=12h com time step horário.	Mccutchan (1976)	Expansão de Séries de Fourier	2-11h	1,3 a 4,7°C (RMSE)	131 dias	Montanhas do Sul da Califórnia	Univariado
Previsão de temperatura média mensal para H=1 mês	Asamoah-Boaheng (2014)	SARIMA	média mensal	0,51°C (RMSE)	33 anos	Ashanti, Gana. Clima equatorial	Univariado
Previsão de temperatura horária média para H=1dia com time step de 3h	Hayati e Mohebi (2007)	ANN, MLP	1-24h	1,8°C (MAE)	10 anos	Kermanshah, Iran	Multivariado. Entradas: temp, umid., pressão, vel. Vento, visibilidade,....
Previsão de temperatura horária média para H=24h com time step de 1h	Abdel-Aal (2004)	Redes Abdutiva	1-24h	1,06 °C (MAE)	7 anos	Seattle, USA	Entradas: Temperatura, Tmax e Tmin e previsão de Max e Min
Previsão de temperatura H=12 com time step horário.	Chevalier et al. (2011)	Support Vector Regression	1h a 12h	0,5°C a 1,9°C (MAE)	de 1997 a 2005.	Geórgia, EUA	Univariado
Previsão de temperatura horária para H=12 com time step horário	Smith et al. (2006)	ANN	1h a 12h	1,35°C (MAE)	4 anos	Geórgia, EUA	Entradas: Temperatura e Variável Sazonal
Previsão de Temperatura intatânea para 5 min à frente	Chaabene e Ben Ammar (2008)	ARMA, Redes Neuro-fuzzy e Filtros de Kalman	5 min	entre 0,87% e 2,28% (nRMSEs)	20 dias (treinar) e 1 dia (teste)	Norte da Tunísia	Varição do erro de acordo com o dia. Utiliza dados dos 20 últimos dias

Fonte: O autor.

Em seu trabalho, ASAMOAH-BOAHENG (2014) avaliou modelos para previsão da temperatura média mensal do ar na superfície. Foram usados dados do período de 1950 a 1994 do Serviço Meteorológico Estadual da Turquia. De acordo com sua análise, o modelo que melhor se ajustou aos dados foi o SARIMA (2,1,1)(1,1,2)(12) (*Seasonal Autoregressive Integrated Moving Average*). O modelo referido obteve previsões com um passo à frente e com RMSE de 0,51°C (média mensal).

### 3.4 Combinação de Previsões aplicadas a Radiação Solar

Em diversas áreas, a busca pela melhoria da acurácia das técnicas de previsão tem sido largamente pesquisada. Em geração de energia, previsões mais acuradas podem significar economia de recursos e diminuição de problemas de despacho na geração. As técnicas de previsão vêm sendo aperfeiçoadas e outras novas vêm sendo desenvolvidas. Como exemplo,

há a combinação de previsão, que se utiliza de previsões originadas de diferentes técnicas para produzir uma previsão com erro menor do que as técnicas individualmente poderiam oferecer. Uma das maneiras de se melhorar as previsões de radiação solar é através da combinação de previsões (PELLAND *et al.*, 2013). Isso vem sendo bastante utilizado em eólica.

Na literatura, o termo combinação de previsão é utilizado como um conceito muito abrangente, muitas vezes confundido com hibridização de modelos. No presente trabalho, combinação de previsão consiste em utilizar como entradas previsões geradas por outros modelos, diferentemente da hibridização, que utiliza como entradas variáveis de naturezas diversas ou subprodutos de modelos de previsão. Modelos de previsão que decompõem um sinal de entrada, realizam a previsão de tais componentes separadamente e em seguida remontam o sinal no futuro também não são considerados neste trabalho como sendo combinação.

Várias são as técnicas de combinação de previsões. Elas vão desde as mais simples, como técnicas lineares (MLR, média aritmética e método da variância mínima) até técnicas mais sofisticadas como combinação através de redes neurais. No trabalho de KHOTANZAD *et al.* (2000) os autores investigaram 8 diferentes técnicas de combinação de previsão: média, mínimos quadrados recursivos, lógica difusa, ANN *feedforward*, ANN funcional, abordagem de espaço de temperatura, algoritmo de programação linear de Karmarkar e mistura adaptativa de especialistas locais. Eles concluíram que a combinação realizada por ANN obteve o melhor desempenho.

Na ótica de combinação de previsão anteriormente exposta, poucos são os trabalhos na literatura que se propuseram a realizar tal tarefa com radiação solar ou potência fotovoltaica.

CHAOUACHI *et al.* (2010) treinaram 4 tipos de redes neurais para prever geração de potência fotovoltaica 24h à frente, sendo a quarta uma combinação das saídas das outras três. A combinação de previsão foi realizada de forma simples através da média de tais previsões. Os valores obtidos para o MAPE variaram entre 2,78% e 4,68% dependendo da estação do ano. Vale ressaltar que esses índices levam em consideração os horários da noite. Trata-se de um modelo com muitas entradas, como temperaturas máximas, mínimas e média, umidade, pressão, cobertura de nuvens, duração do brilho solar e as últimas 24 medidas de radiação solar.

PEREZ *et al.* (2013) realizaram uma comparação entre modelos NWP de previsão de radiação solar para diversos sítios do Canadá, Estados Unidos e Europa. Foi utilizado como referência o modelo de Persistência Medido, que é o valor atual projetado para o futuro

levando em consideração as mudanças de geometria solar. Os modelos de previsão baseados no ECMWF de maneira geral obtiveram melhores resultados comparados aos modelos baseados em GFS (*Global Forecast System*). Os modelos de mesoescala GEM (*Global Environmental Multiscale*) obtiveram desempenhos parecidos com os do ECMWF. Além disso, descobriu-se que a média simples das saídas dos modelos tende a ter um desempenho melhor do que os modelos individuais. No Canadá, os nRMSE dos modelos GEM, ECMWF e ECMWF+GEM foram de 33%, 32% e 31%, respectivamente. Nos EUA, o rRMSE (composto, de todos os sítios) apresentaram valores de 32%, 33% e 32%.

YAGLI *et al.* (2018) utilizaram um modelo de combinação de previsão de geração FV para o dia seguinte à frente, com *time-step* horário. Foi usado um ano de dados, os quais foram de simulação de doze usinas FV na região da Califórnia. Tais dados foram normalizados pela potência total da usina. As previsões foram geradas usando métodos físicos e de séries temporais. Previsões individuais geradas por cinco modelos diferentes são combinadas para eliminar as desvantagens dos métodos individuais. Duas estratégias de combinação foram utilizadas: a média aritmética e a minimização do critério de informação Akaike (AIC). A combinação AIC obteve os melhores resultados. O nRMSE (média geral de todos os horizontes) da combinação variou de 18,97% a 33,94%, a depender da usina em questão. Já o *Forecasting Skill* (melhoria em relação ao modelo de referência) variou de 24% a 33%.

É importante salientar que os trabalhos encontrados que tratam de combinação estrita de previsões foram aplicados a séries de potência e não a séries de GHI, com exceção do trabalho de PEREZ *et al.* (2013). Apesar de se configurar como combinação estrita, o trabalho de PEREZ *et al.* (2013) utilizou uma ferramenta pouco inteligente (média simples) para combinar. Além disso, foram utilizados modelos de uma mesma natureza de dados (NWP) para compor as entradas da combinação.

### **3.5 Análise Crítica sobre o Estado da Arte de Previsão de Radiação Solar Global Horária em Curto e Médio Prazo**

Vários modelos de previsão de radiação solar foram estudados na literatura, que vão de horizontes desde o curto prazo (intra-hora), passando pelo médio prazo até o longo prazo (dias ou meses à frente). Para o despacho de carga do ONS, os horizontes de 24h à frente com *time step* horário são os de maior interesse (VOYANT *et al.*, 2014). Segundo a literatura, modelos baseados em dados terrestres em geral preservam sua acurácia para horizontes

temporais de até algumas horas à frente. Já para horizontes de 02h a 06h à frente, os modelos de previsão mais indicados são os que utilizam como entradas imagens de satélite. Acima de 06h até dias à frente, são indicados modelos baseados em saídas de NWP (DIAGNE *et al.*, 2013). Segundo PEREZ *et al.* (2010), a autocorrelação do índice de céu claro cai rapidamente com o passar de poucas horas. Por isso, para previsões além desse horizonte, recomenda-se a incorporação de saídas de NWP. Porém poucos são os trabalhos que possuem modelos de previsão que cubram todo esse horizonte (curto e médio prazo).

A Tabela 3 traz informações a respeito de alguns aspectos relevantes dos principais trabalhos sobre previsão de GHI horária publicados recentemente e/ou de grande relevância. O aspecto da estratégia de previsão recursiva corresponde a quando um único modelo é treinado para executar previsões de um passo à frente e o mesmo modelo é utilizado para as previsões seguintes, sendo que a previsão anterior passa a ser entrada do modelo e assim sucessivamente. A grande crítica que se faz à estratégia recursiva é que a mesma é sensível à acumulação de erros conforme aumenta o horizonte de previsão (TAIEB *et al.*, 2012).

A combinação relatada na Tabela 3 faz menção ao tipo de combinação estrita de previsões (Seção 3.4), onde saídas de modelos de previsões distintos são combinadas visando agregar as qualidades de dois tipos de modelos:

- Os baseados em séries temporais de dados de radiômetro, dos quais se espera contar com a melhor precisão dos dados de entrada, esperando que estes contribuam mais nas previsões de curto prazo;
- Os baseados em imagens de satélite com um domínio espacial expandido (amplo), de forma a integrar informações ao modelo a respeito das nuvens mais longínquas, acreditando que essa medida possa ampliar o domínio temporal das previsões por imagens de satélite para além dos horizontes comumente relatados na literatura (até 6 horas).

O uso de variáveis determinísticas tem por razão inserir no modelo informações de fases e frequências a respeito do comportamento não estocástico da variável a ser prevista. Segundo SMITH *et al.* (2006), a inserção dessas informações através de uma ou duas variáveis pode não ser suficiente para que a rede consiga extrair todas as informações necessárias. No caso de radiação solar, alguns trabalhos (MARQUEZ e COIMBRA, 2011) utilizaram variáveis determinísticas como o cosseno do ângulo zenital  $\cos(\theta_z)$ . Porém o  $\cos(\theta_z)$  carrega em si informações da sazonalidade diária e da sazonalidade anual. Seguindo a lógica empregada em SMITH *et al.* (2006), tais informações precisariam ser separadas em mais variáveis para que a rede neural captasse-as de forma mais eficaz.

O último aspecto relatado na tabela trata do uso de dados medidos em solo para calibração dos modelos, isto é, eles são utilizados como alvo (valor de referência) para cálculo do erro e, conseqüentemente, do ajuste dos parâmetros do referido modelo.

**Tabela 3 – Aspectos relevantes de alguns dos principais trabalhos sobre previsão de GHI horária para o curto e médio prazo.**

(Autor, Ano)	(Comb (Proposta), 2019)	(PIX (Proposta), 2019)	(Qing e Niu, 2018)	(Bae et al., 2017)	(Monjoly et al., 2017)	(Aguilar et al., 2016)	(Lima, 2015)	(Dong et al., 2014)	(Perez, 2013)	(Marquez e Coimbra, 2011)	(Perez et al., 2010) Satélite	(Mellit e Pavan, 2010)	(Reikard, 2009)
Horizonte de previsão ≥ 6h	X	X	X	-	-	-	X	X	X	X	-	X	-
Entradas	Radiômetro	X	-	-	X	X	X	-	-	-	-	X	X
	Satélite	X	X	-	-	X	-	X	-	-	X	-	-
	NWP	-	-	X	-	X	X	-	X	X	-	-	-
	Não usa Exógenas	X	X	-	-	-	-	-	-	-	X	-	X
	Determinísticas	X	X	X	-	-	-	-	-	-	X	-	X
Modelo não Recursivo	X	X	X	-	-	X	X	-	X	X	X	X	
Combinação de Previsões	X	-	-	-	-	-	-	-	X	-	-	-	
Aplicado ao Brasil	X	X	-	-	-	-	X	-	-	-	-	-	
Não usa dados de Radiômetro na calibração	-	X	-	-	-	-	-	-	-	-	-	-	

Fonte: O autor.

Observando a Tabela 3, além dos trabalhos propostos neste trabalho, apenas dois trabalhos (MELLIT e PAVAN, 2010; DONG *et al.*, 2014) não usaram saídas de NWP e se arriscaram a realizar previsões além das 6h à frente. Porém, somente o de DONG *et al.* (2014) também utilizou imagens de satélite como dados de entrada.

Tendo em vista a escassez de métodos que previsão de GHI horária para horizontes de até 24h à frente, é proposto neste trabalho uma nova metodologia de previsão univariada e não recursiva que realiza uma combinação de saídas de dois modelos de previsão baseados em ANN, tendo como entradas dados de radiômetros, imagens amplas de satélite e oito variáveis sazonais com características *fuzzy*.

Além disso, foi testada uma metodologia que permite calibrar modelos de previsão de curto e médio prazo baseados em imagens de satélite proposta para locais que não disponham de dados medidos por radiômetro.

Haja vista que uma importante aplicação da previsão de GHI é auxiliar as previsões de potência FV e, sabendo que para tal tarefa, a temperatura prevista é essencial no processo de

conversão de GHI em Potência FV, é proposto um modelo simples, univariado e não recursivo de previsão de temperatura ambiente baseado em modelos autorregressivos.

## 4 MATERIAL E MÉTODOS

### 4.1 Metodologia Proposta de Previsão de Irradiância e de Temperatura

Uma das tarefas iniciais em processos de previsão é a seleção das variáveis de entrada. Para modelos baseados em ANN, não existe um procedimento sistemático que possa ser seguido. A Função de Autocorrelação e a Função de Autocorrelação Parcial são boas ferramentas para seleção de variáveis quando se usam modelos lineares. Já quando usamos modelos não lineares, como ANN, esses estatísticos dão apenas uma boa indicação (SFETSOS e COONICK, 2000).

A metodologia proposta está baseada nas previsões das duas variáveis mais importantes para a descrição do comportamento da potência de saída da central fotovoltaica (MERAL e DINÇER, 2011): a radiação solar e a temperatura ambiente, estando as previsões de radiação solar baseadas tanto em observações de estações solarimétricas quanto imagens de satélite e as previsões de temperatura baseadas em observações de estações meteorológicas.

O principal diferencial da metodologia aqui proposta reside em não utilizar hibridização (junção de entradas), mas sim uma combinação (junção de saídas) das previsões de irradiância oriundas de dois modelos, um alimentado por imagens de satélite (das quais se pretende inferir conhecimento sobre a transmitância da atmosfera considerando-se um domínio espacial amplo, e, por consequência, também um domínio temporal suficientemente amplo) e outro alimentado por dados observacionais da irradiância em superfície (das quais se pretende inferir conhecimento sobre a dinâmica rápida da irradiância na vizinhança do local em que se encontra instalada a estação solarimétrica). Metodologia semelhante já foi utilizada por COSTA (2005) para previsões eólicas, com a diferença de que o conjunto de dados empregados para representar o maior domínio espacial proveio de modelos NWP em lugar de imagens de satélite.

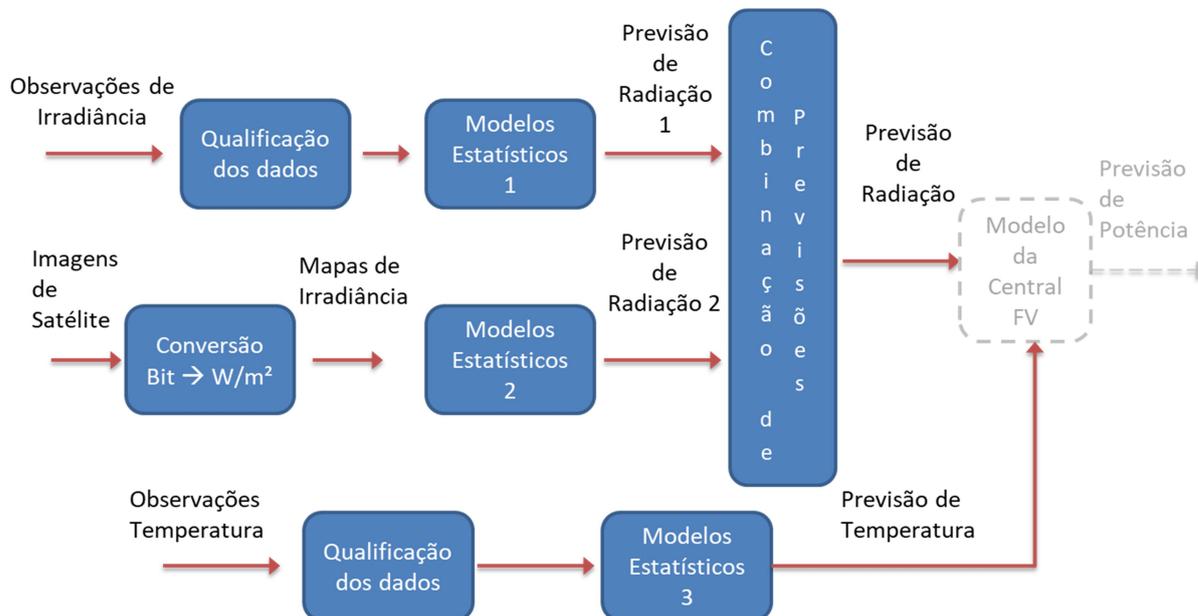
No presente trabalho, as imagens de satélite foram preferidas em relação a dados NWP (numéricos) por possuírem maior disponibilidade, pois modelos NWP de mesoescala, os quais tem resolução temporal adequada (horária), não são facilmente generalizáveis para qualquer parte do globo. Por terem alto custo computacional, tais modelos de mesoescala não estão disponíveis para todos os sítios. Além disso, tratando-se da estimação de radiação solar, estimativas de GHI oriundas de imagens de satélite apresentam melhores resultados comparadas às estimativas fornecidas por saídas de modelos numéricos. Um exemplo disso é

encontrado no trabalho de CAMARGO *et al.* (2015). Acredita-se assim que uma série de dados de entrada mais bem correlacionada com a série a ser prevista traria maiores benefícios para a calibração dos modelos de previsão.

Embora medidas de radiômetro possuam uma maior acurácia, comparadas às estimativas de GHI via imagens de satélite, os modelos de previsão baseados exclusivamente em séries temporais de dados de radiômetro acabam perdendo a informação espacial fornecida pelo satélite, informação essa que poderia melhorar os resultados (BLANC *et al.*, 2017). A Figura 6 sintetiza a metodologia proposta para este trabalho, na qual imagens de satélite, após passarem por devido tratamento, serão convertidas em uma série temporal de matrizes (mapas) de irradiância solar incidente na superfície. Em seguida, ocorre uma redução da dimensionalidade dessas matrizes através da aplicação de técnicas de seleção de pixel por correlação e de análise de componentes principais (WILKS, 2011). Essa nova série temporal de matrizes serve como entrada para os Modelos Estatísticos 2, que (após treinamento e validação) fornecerão previsão de irradiância para o ponto (*pixel*) escolhido sobre o mapa. Por outro lado, os dados de irradiância observados na superfície (após avaliação de sua qualidade) alimentarão os Modelos Estatísticos 1, que fornecerão uma outra previsão de irradiância solar para o mesmo local (*pixel*) referido anteriormente. Essas duas previsões serão combinadas através de redes neurais, resultando em uma previsão final de irradiância que seja mais acurada que as duas previsões anteriores individualmente.

O passo seguinte (em pontilhado na Figura 6) seria transformar a previsão final de irradiância em previsão de potência da central FV. Para realizar tal transformação, é necessário saber a temperatura no tempo  $t+k$ . Visando ser capaz de fornecer todos os insumos de tal transformação, será utilizado um modelo de estimação da saída da central FV baseado na previsão final de irradiância e na previsão da temperatura do gerador FV, que por sua vez depende da temperatura ambiente. Dessa forma, é necessário treinar também um modelo de previsão de temperatura (Modelos Estatísticos 3), o qual terá como entradas dados observacionais da temperatura e sua saída será a previsão da temperatura ambiente horária na superfície para 24h à frente.

**Figura 6 – Metodologia proposta para previsão de GHI e Temperatura ambiente com vistas a auxiliar previsões de potência de uma central fotovoltaica.**



Fonte: O autor.

## 4.2 Base de Dados

Para este estudo, três sítios foram estudados: Brasília (BRB), Pesqueira (PES) e Petrolina (PTR). Os locais selecionados possuem dados medidos de GHI confiáveis na base de minuto-a-minuto. Algumas características dessas estações estão descritas na Tabela 4.

**Tabela 4 – Características dos sítios estudados no presente trabalho.**

Local	Latitude	Longitude	Altitude (m)	Nebulosidade Média Anual <sup>1</sup> (décimos)	Precipitação anual <sup>1</sup> (mm)	Clima <sup>2</sup>
Brasília	-15,60083	-47,71306	1023	0,6	1477	Tropical de Altitude
Pesqueira	-8,3666333	-36,68065	650	0,6	591	Tropical de Savana <sup>3</sup>
Petrolina	-9,0689	-40,3197	387	0,5	482	Semiárido Quente Estépico

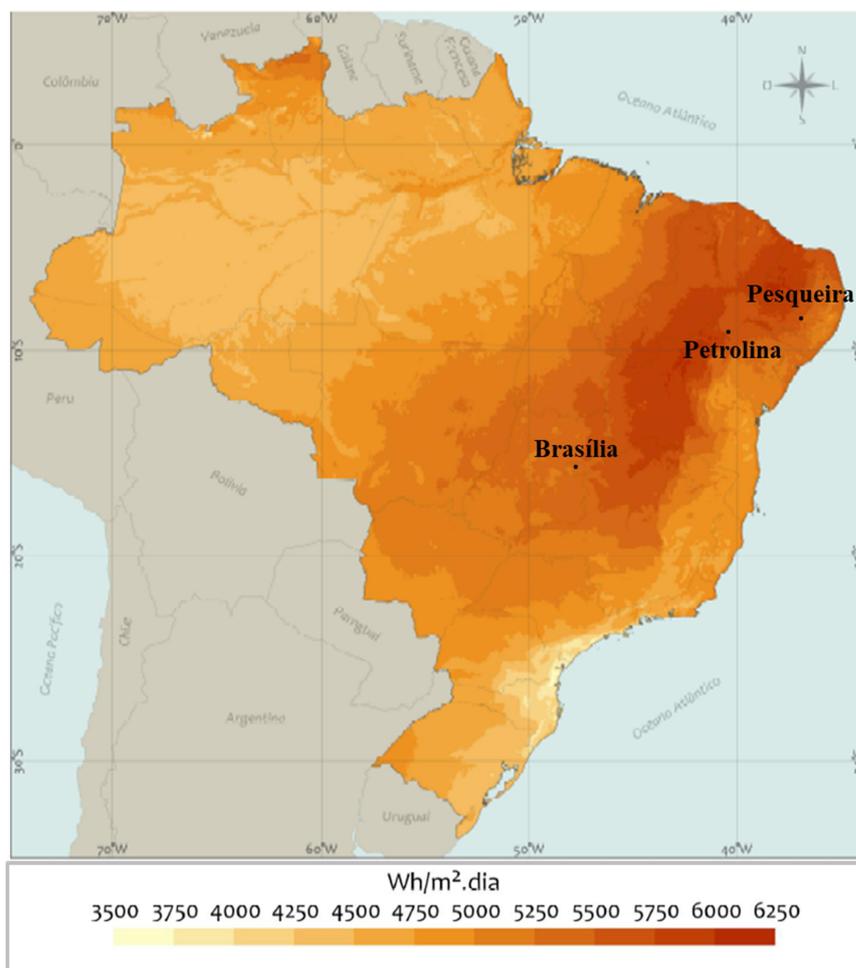
Fonte: O autor.

<sup>1</sup> Fonte: <http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas>

<sup>2</sup> Fonte: <http://sonda.ccst.inpe.br/infos/index.html>

<sup>3</sup> Fonte: <http://www.weatherbase.com/weather/weather-summary.php3?s=598082&cityname=Cana%E3%2C+Pernambuco%2C+Brazil&units=>

**Figura 7 – Distribuição geográfica e disponibilidade do recurso solar nos locais estudados.**



Fonte: (PEREIRA *et al.*, 2017) (adaptado).

As estações de Petrolina e Brasília são de responsabilidade do INPE e são estações de referência pertencentes à Rede SONDA (Sistema de Organização Nacional de Dados Ambientais) e à rede *Baseline Surface Radiation Network* (BSRN). Já a estação Pesqueira é de responsabilidade do Instituto Federal de Pernambuco, possuindo uma base de dados de GHI de boa qualidade (baixa incidência de valores anômalos e buracos).

Os sítios estudados estão distribuídos geograficamente conforme Figura 7. Brasília está localizada na região central do país, no Distrito Federal, sujeita a uma disponibilidade de radiação solar média anual de 5,25 KWh/m<sup>2</sup>.dia. Pesqueira está situada na mesorregião do Agreste pernambucano, no Vale do Ipojuca, apresentando uma radiação média anual de 5,4 KWh/m<sup>2</sup>.dia. Petrolina encontra-se no Sertão de Pernambuco, na divisa com o estado da Bahia, apresentando uma radiação média anual de 5,8 KWh/m<sup>2</sup>.dia (PEREIRA *et al.*, 2017).

#### 4.2.1 Dados de Radiação Solar

Neste estudo foram utilizados dados medidos por radiômetros das três estações descritas na Tabela 4. Tratam-se de dados na base minuto a minuto correspondentes aos anos de 2013, 2014 e 2015. Apenas os dados contidos na janela útil do dia foram utilizados. Tal janela compreende os dados do período do dia no qual o ângulo zenital  $\theta_z > 80^\circ$ . As razões para esta escolha são:

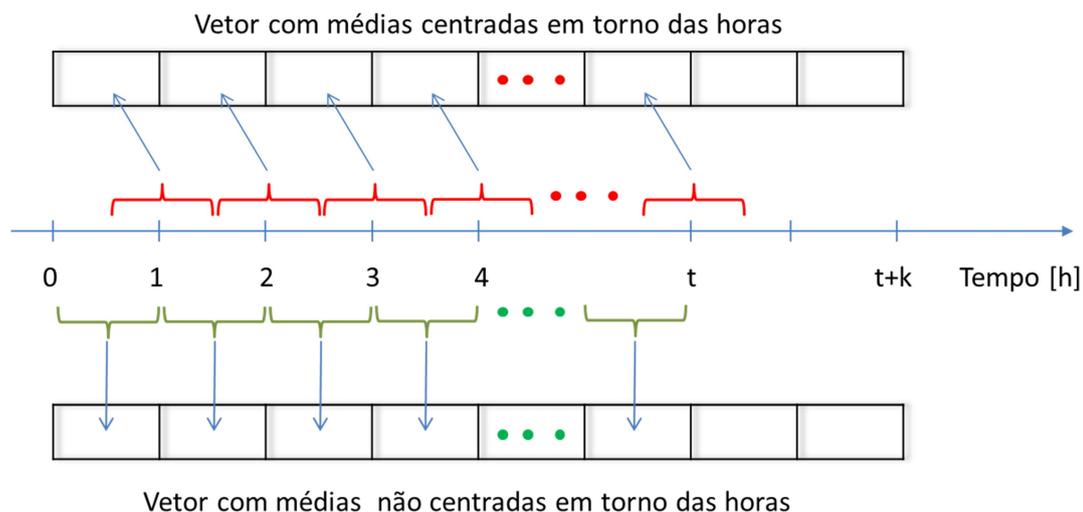
- Os sensores de radiação solar global em geral só conseguem medir com precisão dados com ângulos de elevação solar superiores a  $10^\circ$ ;
- Valores de radiação solar fora dessa janela não são representativos do ponto de vista da geração de energia fotovoltaica;
- Em relação à outra base de dados de GHI, obtida a partir de imagens de satélite, tais estimativas apresentam alta incerteza para baixos ângulos de altitude solar.
- Caso os dados da noite e do começo da manhã não fossem excluídos da série de calibração, os períodos longos de irradiância zero intercalados com irradiância positiva trariam dificuldade para as ANN absorverem tal dinâmica.

Nesse sentido de recorte da série temporal, embora seja bem entendido que a radiação solar durante a noite é zero, todos os modelos em geral são desenvolvidos para prever apenas as horas do dia. Com isso, é possível testar a capacidade dos modelos de Inteligência Artificial de lidar com descontinuidades do conjunto de dados (SFETSOS e COONICK, 2000).

Em uma aplicação de previsão semelhante a deste presente trabalho, MONJOLY *et al.* (2017) realizaram o processo de filtragem dos dados que removeu os dados com  $\theta_z > 80^\circ$  (ângulo de elevação  $\alpha < 10^\circ$ ). A radiação nesses horários é muito baixa e insignificante para a geração de eletricidade (MONJOLY *et al.*, 2017).

Os dados de radiação solar passaram ainda por uma garantia de qualidade contendo filtros físicos e estatísticos, os quais estão descritos em PETRIBU *et al.* (2017). Em seguida os dados foram transformados para a base horária. Vale destacar que a passagem para a base horária foi realizada de duas maneiras: através de médias centradas e através de médias não-centradas (Figura 8).

**Figura 8 – Duas formas de converter dados da base minuto a minuto para a base horária**



Fonte: O autor.

As médias horárias centradas no minuto zero de cada hora (isto é, 07h00, 08h00, 09h00,...) tendem a ser mais correlacionadas com os valores instantâneos correspondentes aos minutos zero dessas mesmas horas. Essa é uma característica importante quando se trabalha com modelos que utilizam imagens de satélite como entrada, as quais dizem respeito àquele instante específico no qual tal imagem foi gerada, isto é, nos instantes 07h00, 08h00, 09h00,... Os dados centrados são utilizados como alvo apenas na calibração desses modelos. Já quando se trabalha com modelos de previsão os quais utilizam observações médias horárias como entrada, não se deve trabalhar com médias centradas, pois, de acordo com a chave vermelha mais a direita da Figura 8, o valor médio horário referente ao instante  $t$  estaria utilizando informação *a posteriori* (da meia hora futura) na entrada do modelo, o que fere uma premissa básica de previsão. Para esses modelos, é indicado o uso de dados não centrados.

Seguindo a teoria da validação cruzada (MICHAELSEN, 1987), os dados horários foram divididos em dois subconjuntos, um de calibração (2013 e 2014) e outro de validação (2015).

#### 4.2.2 Imagens de Satélite

As imagens de satélite utilizadas são oriundas dos satélites geoestacionários GOES 12 (Janeiro a Julho de 2013) e do GOES 13 (Agosto de 2013 a Dezembro de 2015). Tais satélites fazem parte do projeto GOES (*Geostationary Operational Environmental Satellite*), administrados pelo NOAA (*National Oceanic and Atmospheric Administration*). As características do sensor Imager do GOES estão mostradas na Tabela 5.

Para a obtenção das estimativas de GHI, foram utilizadas as imagens do canal visível (Figura 9), as quais foram obtidas junto ao Centro de Previsão de Tempo e Estudos Climáticos (CPTEC/INPE), que, apesar de serem geradas pelo satélite GOES com resolução espacial de 1 Km<sup>2</sup>, são disponibilizadas pelo referido centro com resolução espacial de 4 Km<sup>2</sup>. Por levarem 26 minutos para a geração de cada imagem (Tabela 5), os satélites GOES 12 e GOES13 disponibilizavam imagens a cada 30 minutos. A título de informação, o satélite GOES 16, posto em operação em 18 de dezembro de 2017, leva 9 minutos para gerar cada imagem, sendo então capaz de prover imagens a cada 10 minutos.

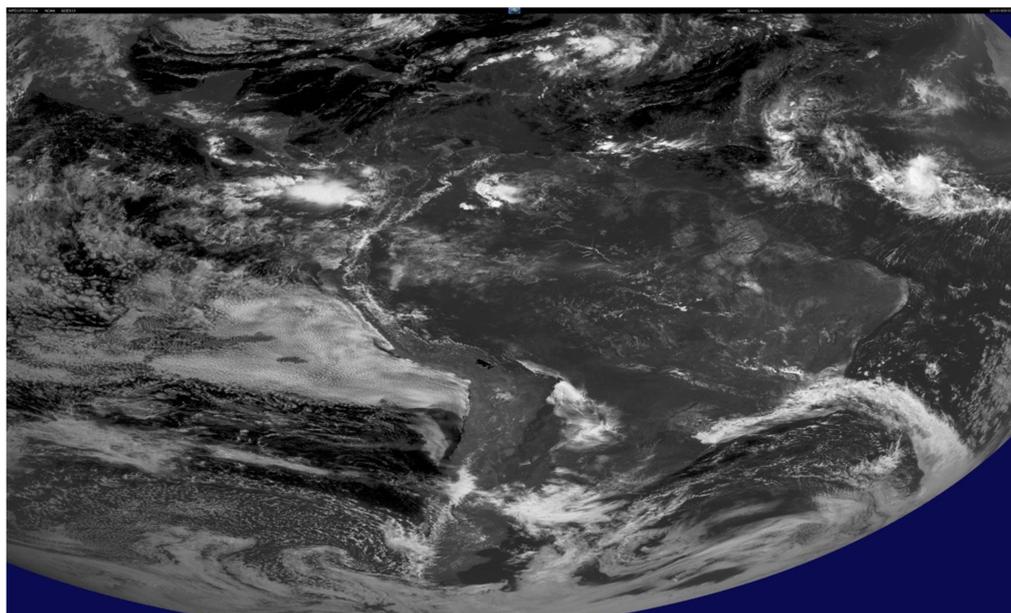
A resolução radiométrica do canal visível é de 8 bits, o que proporciona uma variação do nível de cinza de cada pixel de 0 (cor preta) a 255 (cor branca) (BARROS *et al.*, 2018).

**Tabela 5 – Especificações do Sensor *Imager* pertencente aos Satélites GOES 12 e GOES 13**

Características	Canal				
	1. Visível	2. Ondas curtas	3. Umidade	4. IR 1	5. IR 2
Comprimento de onda (µm)	0,55 – 0,75	3,8 – 4,0	6,5 – 7,0	10,2 – 11,2	11,5 – 12,5
Resolução espacial	1 km <sup>2</sup>	4 km <sup>2</sup>	8 km <sup>2</sup>	4 km <sup>2</sup>	4 km <sup>2</sup>
Acurácia do sistema	Canal Infravermelho: Menor ou igual a 1 K Canal Visível: 5% da Irradiância de cena				
Resolução temporal	Disco terrestre inteiro em até 26 minutos				

Fonte: (BARROS *et al.*, 2018) (adaptado).

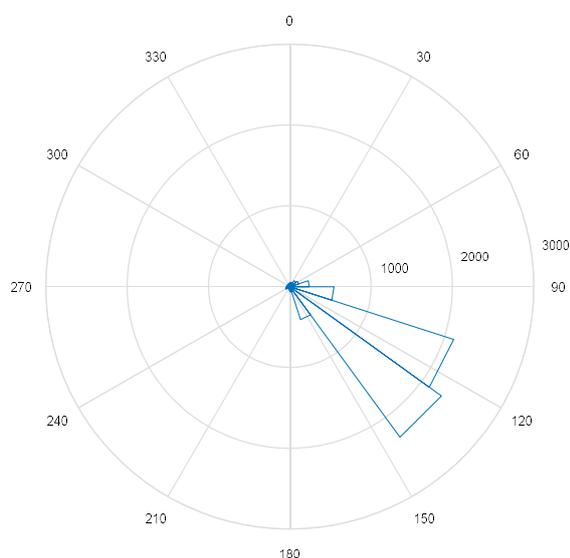
**Figura 9 – Imagem da América do Sul do canal visível do Satélite GOES 13 do dia 05 de outubro de 2015 às 12h00 (horário de Brasília).**



Fonte: (DSA/INPE, 2015).

Para se determinar o tamanho da imagem com a qual se trabalharia (recorte da imagem da América do Sul), tomou-se como base a estação de Petrolina. Traçou-se a rosa dos ventos dos dados medidos, de onde se observou que os ventos em Petrolina possuem direção predominantemente sudeste (Figura 10). Decidiu-se trabalhar com uma imagem de satélite composta por 131x131 pixels de 4 km cada um, formando assim um quadrado de 524 km de lado centrado no pixel que contém a estação meteorológica em estudo (Figura 11). Considerando que a diagonal do centro da imagem até o canto inferior direito possui 370 km e considerando uma velocidade média das nuvens em 30 km/h, uma nuvem levaria em torno de doze horas para se deslocar do canto ao centro da imagem, tempo esse compatível com o máximo horizonte de previsão a ser utilizado neste trabalho. LARSON e COIMBRA (2018) realizaram um estudo que embasou a escolha do tamanho (64x64 pixels) das imagens usadas para a previsão de potência fotovoltaica. Porém seu horizonte de previsão era de até seis horas à frente. Isso faz parecer existir uma proporção coerente entre o seu tamanho de imagem calculado e o tamanho calculado pelo presente trabalho.

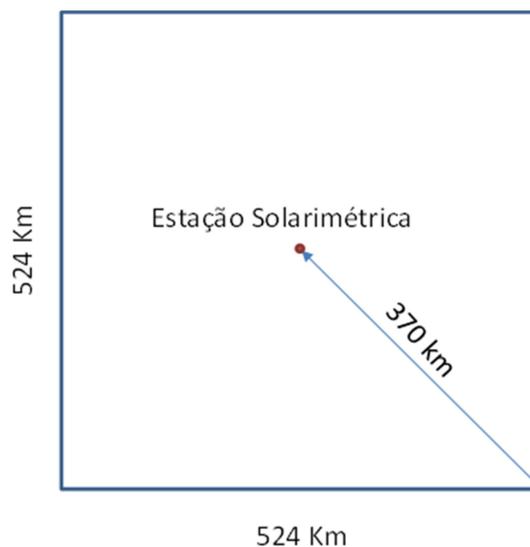
**Figura 10 – Rosa dos ventos traçada com dados medidos de direção do vento da estação SONDA Petrolina.**



Percebe-se a predominância de ventos na direção sudeste (sentido sudeste para o centro).

Fonte: O autor.

**Figura 11 – Tamanho e localização do recorte da imagem de satélite utilizado neste trabalho.**



Fonte: O autor.

Para cada ano, montou-se uma série temporal de imagens de satélites recortadas e centradas no ponto de interesse. A identificação do pixel correto que continha a estação meteorológica foi realizada através de correlação entre as séries de estimativas de GHI de cada pixel e a série de GHI medida em solo, conforme estratégia semelhante à utilizada por AGUIAR *et al.* (2016).

Após a aplicação do modelo Brasil-SR, descrito na seção 3.1, em cada pixel das imagens recortadas, a série temporal de imagens se converteu em série temporal de (mapas) radiação. Com isso encerrou-se o processo de conversão de nível de cinza em radiação solar ( $W/m^2$ ). Os parâmetros estatísticos do pixel de interesse usados para avaliar a acurácia do processo de estimativas de irradiância via Brasil-SR estão descritos na Tabela 6. Estes valores foram considerados aceitáveis, visto que na literatura, os erros são da ordem de 25% para valores horários de GHI ( $W/m^2$ ) (DAMBREVILLE *et al.*, 2014).

**Tabela 6 – Estatísticos de comparação das séries temporais de estimativas de GHI do pixel central das imagens de satélite com dados de radiômetro dos sítios Brasília, Pesqueira e Petrolina.**

	BRB				PES				PTR			
	2013	2014	2015	Média	2013	2014	2015	Média	2013	2014	2015	Média
Correlação	0,88	0,92	0,92	0,91	0,88	0,91	0,91	0,90	0,88	0,93	0,92	0,91
nRMSE	29,20	25,19	24,24	26,21	32,48	31,33	26,15	29,99	27,59	20,67	23,10	23,79

Fonte: O autor.

### 4.2.3 Temperatura Ambiente

A base de dados de Petrolina, contendo dois anos de medidas de temperatura, foi escolhida devido ao interesse relacionado à previsão de geração de energia elétrica de uma central fotovoltaica, a qual se encontra em fase de implantação. Os dados de temperatura de Petrolina são oriundos da estação meteorológica pertencente à rede SONDA, mantida pelo INPE, cuja localização está descrita na Tabela 4. Trata-se de uma estação meteorológica com medidas de temperatura minuto a minuto. Por conveniência, os dados foram transformados para a base horária.

Os algoritmos aqui utilizados em modelos autorregressivos não funcionam em séries temporais com dados faltantes. Por isso, adotou-se a seguinte estratégia:

- Para um dado horário faltante único, este foi estimado a partir de interpolação simples, através da média aritmética entre as duas medidas vizinhas.
- Para um dado horário faltante não isolado, isto é, com outros dados faltantes vizinhos, o dia associado a tais dados foi excluído da série.

O treinamento e a validação dos modelos aqui testados seguem a filosofia da validação cruzada, onde o conjunto de dados é subdividido em dois conjuntos mutuamente exclusivos (MICHAELSEN, 1987). A calibração foi feita usando dados de 2014, enquanto que a validação foi realizada com os dados de 2015.

## 4.3 Modelo de Referência

Para se avaliar o desempenho de um modelo de previsão, é comum entre muitos autores (COSTA, 2005; MONTGOMERY *et al.*, 2012) a utilização de um outro modelo (de referência), para assim comparar os resultados de ambos com as observações. A comparação via modelo de referência se justifica pelo fato de que quando se compara, por exemplo, o RMSE encontrado por um modelo de um autor com o RMSE encontrado nos trabalhos de outro autor, tal comparação não é adequada (justa) devido aos experimentos terem sido realizados sob condições (local, época do ano, altitude, variabilidade do recurso, entre outras) diferentes (INMAN *et al.*, 2013). O modelo de referência comumente utilizado é o Modelo de Persistência, que toma como premissa a ideia de que um parâmetro mantém seu valor após um intervalo de tempo  $\Delta t$ , tempo esse chamado de horizonte de previsão (INMAN *et al.*, 2013). É definido pela Equação (10) como:

$$\hat{y}_{t+k} = \hat{y}_t \quad (10)$$

onde  $\hat{y}_{t+k}$  representa a previsão  $k$  passos à frente.

Falando especificamente em séries temporais de GHI média horária, modelos de Persistência tendem a ter um pior desempenho do que se fossem utilizados em séries de Kt. Isso se deve ao comportamento determinístico da radiação, que tem uma maior variabilidade comparada ao Kt.

Outra particularidade referente à Persistência em radiação solar horária tem relação com a condição de céu. É de se esperar que, em dias de céu nublado, o RMSE de um modelo de previsão (RMSE<sub>m</sub>) tenda a aumentar, como apresentado na literatura (MELLIT e PAVAN, 2010; DIAGNE *et al.*, 2013; BAE *et al.*, 2017). Isso, por si só, já faria com que o *Improvement* (também chamado de *Forecasting Skill*, melhoria em relação ao modelo de Persistência) tendesse a diminuir. Levando em consideração que, em dias de céu nublado (este raciocínio não se aplica a dias parcialmente nublados), onde a faixa de variação da radiação solar horária é menor, podemos inferir que o RMSE da persistência (RMSE<sub>p</sub>) tenda a diminuir, pois o sinal tende a ficar mais persistente, pois a faixa de variação da GHI diminui. Essa é uma ajuda extra sistemática que tende a degradar ainda mais o *Improvement* e que não tem relação com o modelo de previsão, mas sim com o modelo de referência que está sendo adotado. Olhando a Figura 12, podemos entender melhor as explicações citadas.

Figura 12 – Variação esperada dos componentes usados no cálculo do *Improvement* (Imp) na condição de céu nublado, onde os índices  $m$  e  $p$  dizem respeito ao modelo em estudo e ao modelo de persistência, respectivamente.

$$Imp = 1 - \frac{RMSE_m}{RMSE_{pe}}$$

$\uparrow$   
 $\downarrow$

Tendências dos  
erros em condição  
de céu nublado  
(comparadas ao céu  
claro)

Fonte: O autor.

Na literatura é possível encontrar modelos de Persistência modificados, os quais são corrigidos considerando-se a geometria Sol-Terra (SFETSOS e COONICK, 2000) ou pela razão entre os valores nos instantes  $t$  e  $t+k$  do modelo de céu claro (LARSON e COIMBRA, 2018). A Persistência modificada utilizada por SFETSOS e COONICK (2000) apresenta resultados 17% melhores do que a Persistência comum. Porém, com o intuito de facilitar a

comparação com modelos de outros autores, será adotada neste trabalho a persistência comum.

#### 4.4 Modelos de Previsão Propostos

Nesta seção são detalhados separadamente os modelos de previsão de GHI baseados em dados medidos em superfície (Modelos Estatísticos 1), imagens de satélite (Modelos Estatísticos 2), previsão de temperatura (Modelos Estatísticos 3) e a combinação de previsões, conforme ilustrados na Figura 6. Com respeito aos Modelos Estatísticos 2, será ainda apresentado um modelo alternativo, denominado de PIX, o qual não utiliza dados de radiômetros nem mesmo durante a calibração. Esse modelo alternativo visa fornecer previsões de GHI para lugares onde não se tenham nenhum tipo de dado medido em superfície.

##### 4.4.1 Previsão Baseada em Dados de Radiômetro

Para a tarefa de previsão de GHI baseada em dados de superfície, modelos baseados em redes neurais artificiais (ANN) são propostos. A estrutura de ANN utilizada neste trabalho foi a Perceptron Multicamada (MLP), a qual tem sido aplicada com sucesso em previsão de irradiância solar (MELLIT e PAVAN, 2010; WANG *et al.*, 2011; VOYANT *et al.*, 2014). A função de ativação adotada foi a Tangente-Sigmóide e o algoritmo de aprendizagem para treinamento da rede foi o de Levenberg-Marquardt. Foram utilizadas 50 épocas. As redes testadas possuem estruturas com até três camadas ocultas e, em cada uma delas, até oito neurônios, respeitando a característica triangular (a camada oculta  $i$  tem que ter número de neurônios maior ou igual à camada  $i+1$ ). Foram calibrados modelos não-recursivos, isto é, um modelo para cada horizonte, de 1 até 12 passos à frente.

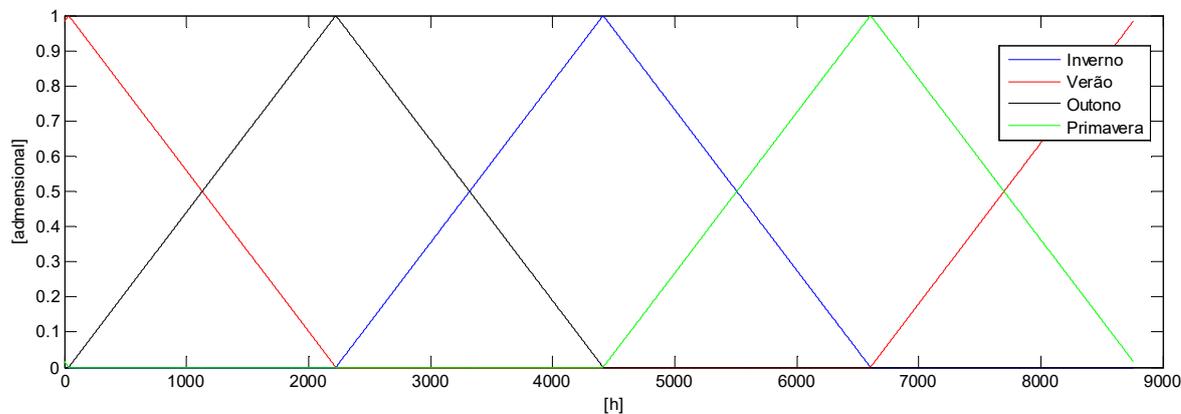
Para a tarefa de treinar as várias combinações de ANN, foi utilizado o cluster Papalégua de alto desempenho pertencente ao CER-UFPE, com 22 CPUs AMD Opteron 6000 series, 352 cores e 384 Gb de RAM compartilhada. Para cada rede, são realizadas 200 inicializações e o modelo com melhor desempenho baseado no Skill de TAYLOR (2001) é então escolhido.

Duas classes de modelos foram propostas, classificadas de acordo com as variáveis de entrada:

- TER-SVS: modelos SVS (sem variáveis sazonais) que possuem como entradas valores  $y$  de GHI medidos em terra no tempo presente,  $y(t)$ , e até 15 valores passados,  $y(t-1)$ ,  $y(t-2)$ ,...,  $y(t-15)$ , tendo assim o modelo uma ordem máxima de  $p=16$ .
- TER: modelos semelhantes ao TER-SVS, acrescentando as 8 variáveis sazonais.

A primeira classe de modelos possui como entradas apenas o valor presente e, no máximo, 15 valores passados da série de GHI. É comum que, além da variável endógena, as variáveis data e hora sejam utilizadas como entradas de modelos baseados em ANN. Segundo SMITH *et al.* (2006), a inserção dessas duas variáveis não é suficiente para que a ANN capte satisfatoriamente as informações de sazonalidade da variável em estudo. Para contornar tal problema, cada uma dessas variáveis (data e hora) foi dividida em quatro outras variáveis determinísticas, totalizando oito, as quais são construídas utilizando conceitos de lógica difusa, com formatos triangulares e com amplitudes variando entre 0 e 1.

**Figura 13 – Variáveis sazonais anuais utilizadas como entrada para melhoria do modelo baseado em ANN.**

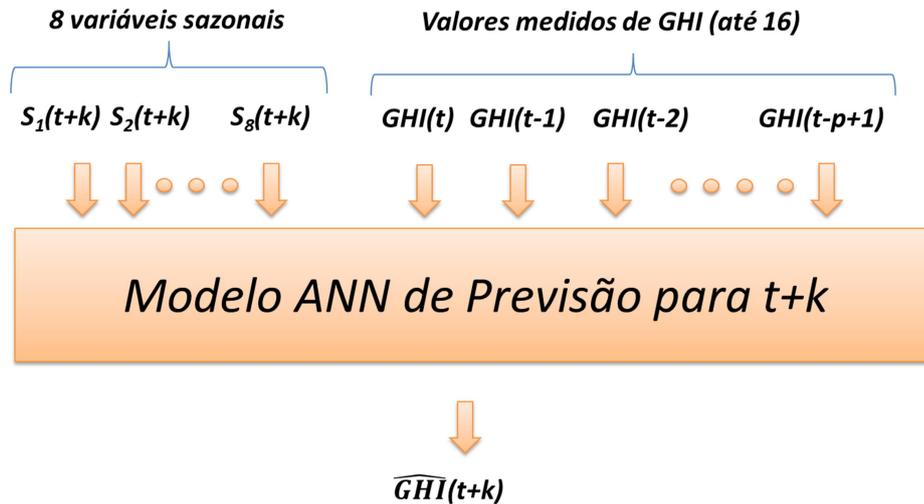


Fonte: O autor.

As oito variáveis sazonais foram incluídas como entradas da rede, sendo 4 representando a sazonalidade das estações (Primavera, Verão, Outono e Inverno, ilustradas na Figura 13) e 4 representando a sazonalidade diária (Manhã, Meio-Dia, Tarde e Meia-Noite). As 4 variáveis sazonais das estações possuem mesma frequência (anual), diferenciando-se entre si pela fase. Já as outras 4 variáveis possuem frequência diária e também diferenciando-se pelo ângulo. A este novo modelo de ANN, contendo como entradas valores de GHI e as variáveis sazonais, foi dado o nome de Modelo TER, o qual está esboçado na Figura 14. Por

serem determinísticas, podem-se utilizar os valores do futuro das variáveis sazonais, isto é, valores em  $t+k$ .

**Figura 14 – Modelo de Previsão de GHI, chamado TER, que utiliza variáveis sazonais e amostras da série observada de GHI.**



Fonte: O autor.

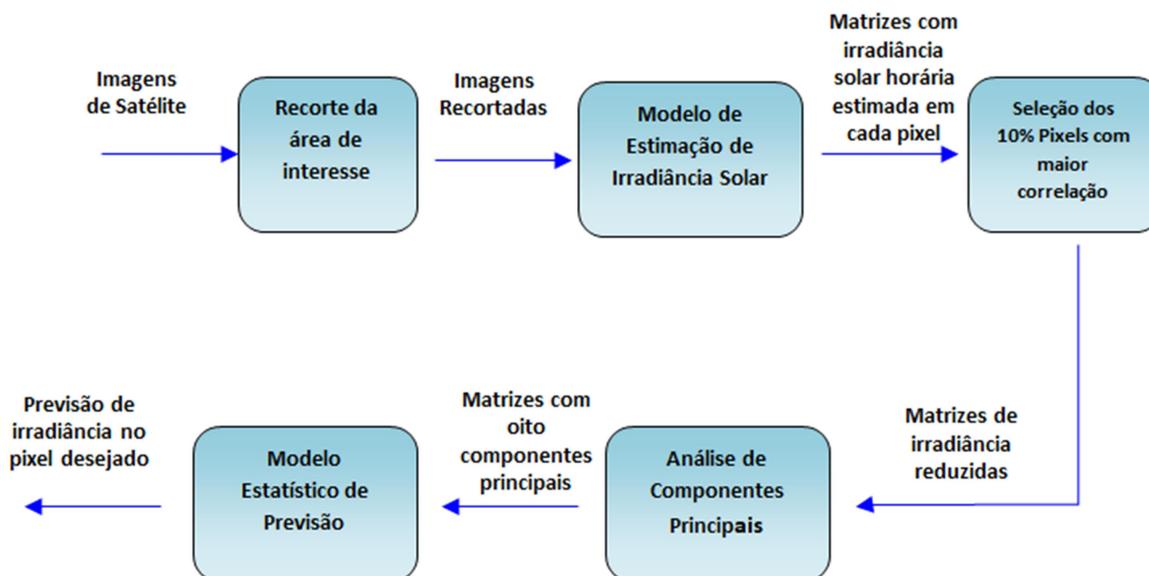
O emprego de tais variáveis sazonais com característica difusa na previsão de GHI horária, embora já utilizado em outra área (Temperatura), não foi encontrado na literatura, sendo o trabalho de SABINO *et al.* (2018) uma exceção.

#### 4.4.2 Previsão Baseada em Imagens de Satélite

Trabalhar com imagens de satélite remete a um grande esforço computacional, pois normalmente há uma grande massa de dados a serem processados. Isso exige um tratamento prévio para adequar os dados (imagens) ao problema em questão. A Figura 15 resume os processos empregados no presente trabalho, que vão desde o recorte das imagens até a previsão de GHI no ponto de interesse.

**Figura 15 – Metodologia para obtenção das previsões de GHI baseadas em imagens de satélite.**

**Os processos que estão antes da terceira seta azul representam a estimativa de GHI. Daí em diante inicia a metodologia de previsão propriamente dita.**



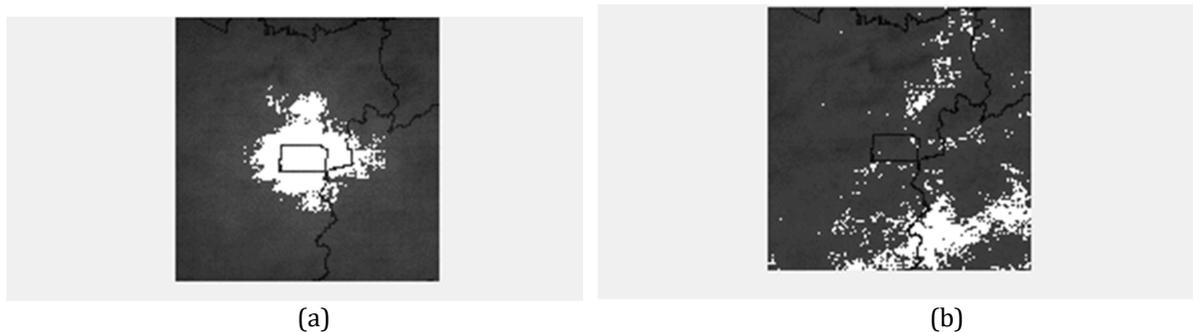
Fonte: O autor

Realizada a conversão de nível de cinza em GHI, descrita na Seção 3.1, cada mapa contém 17161 pixels, o que é uma quantidade inviável de variáveis de entrada a ser utilizada em uma rede neural. Para reduzir a dimensionalidade deste conjunto, aplicaram-se duas estratégias:

- A seleção dos pixels via correlação (selecionando os 10% dos pixels mais bem correlacionados com a série de GHI defasada  $k$  deslocamentos à frente);
- A análise de componentes principais, para eliminar as informações redundantes desse conjunto resultante.

O método de seleção de pixels adotado neste trabalho já foi testado em trabalhos anteriores (DAMBREVILLE *et al.*, 2014; AGUIAR *et al.*, 2016). Os pixels que apresentam maior correlação com o pixel de interesse com  $k$  lags (deslocamentos no tempo) à frente são escolhidos. Para cada horizonte de previsão, um conjunto de pixels diferentes é selecionado. A Figura 16 ilustra o mapeamento dos pixels das imagens de Brasília selecionados (na cor branca) por maior correlação absoluta para os horizontes 1 e 10 passos à frente.

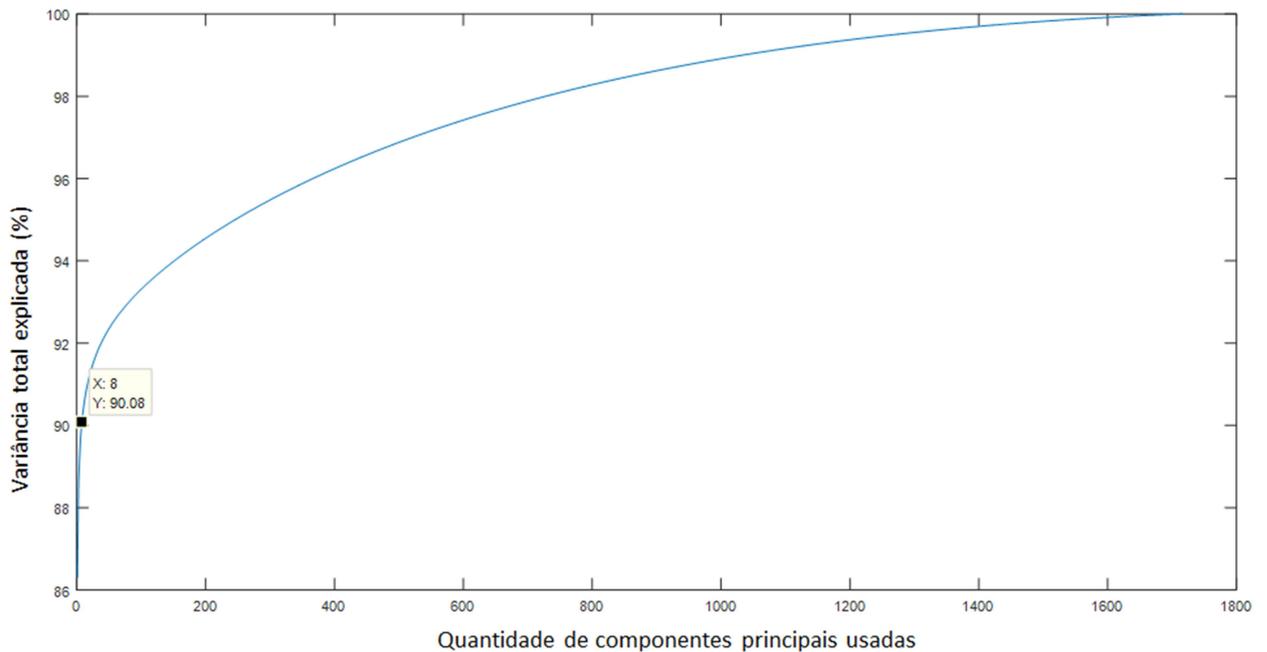
**Figura 16 – Mapeamento dos pixels de Brasília selecionados por correlação para serem entradas dos modelos (ANN) dos horizontes: a) 1 passo à frente; b) 10 passos à frente.**



Fonte: O autor

Após a seleção dos pixels, o número de variáveis de entrada caiu de 17161 (131x131) para 10% desse valor, ou seja, 1716 variáveis de entrada. Esta ainda é uma quantidade alta de entradas. Para reduzir ainda mais a dimensionalidade desse problema, foi aplicada a técnica de Análise de Componentes Principais (WILKS, 2011), que visa a eliminação de redundâncias presentes nas imagens de satélite. Tal eliminação de redundâncias se justifica por conta da forte correlação entre irradiâncias sobre pixels vizinhos (KLEISSL, 2013) e da dimensão das imagens de satélite, tipicamente compostas por milhares de pixels, acarretando dificuldades para seu processamento. As componentes principais são organizadas de forma decrescente em relação ao percentual de variância explicada do conjunto inicial de variáveis. O gráfico com a variância explicada acumulada em função do número de componentes principais utilizadas é mostrado na Figura 17. A partir de análise de gráficos semelhantes ao da Figura 17 para os diversos horizontes, foi escolhido o número de 8 componentes principais para se trabalhar como entradas das ANN, uma vez que esse número de componentes, no pior caso, é capaz de explicar 87,5% da variância do conjunto original (1716 variáveis). A Tabela 7 sintetiza a quantidade de variância explicada por 8 componentes principais para as variáveis de Petrolina em 2013 e 2014. Esses estudos foram realizados para os demais sítios, os quais apresentaram resultados semelhantes (Apêndice C).

**Figura 17 – Percentual acumulado de variância explicada (do conjunto inicial de variáveis) em função do número de componentes principais utilizado para a série de 10% de pixels usados no horizonte 1 passo a frente em Petrolina (2013 e 2014).**



Fonte: O autor.

**Tabela 7 – Percentual acumulado de variância explicada por oito componentes principais para os diversos horizontes da série de Petrolina de 2013 e 2014.**

	Horizonte (h)											
	1	2	3	4	5	6	7	8	9	10	11	12
Variância explicada (%)	90,08	89,74	87,57	88,42	89,56	89,66	89,8	89,79	88,03	87,93	88,75	89,41

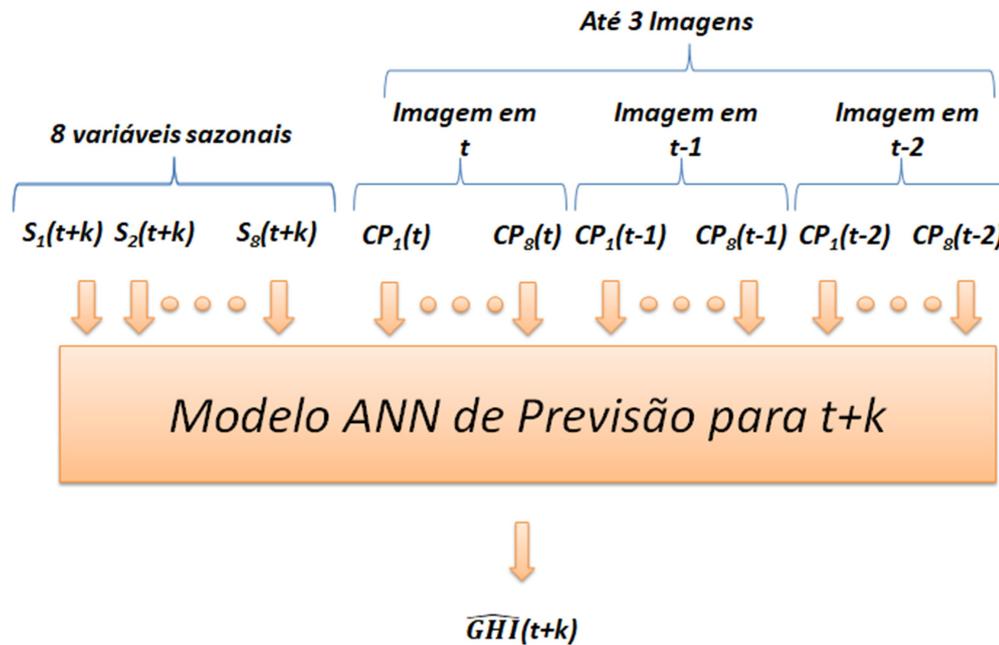
Fonte: O autor.

O processo de previsão propriamente dito é realizado utilizando redes neurais treinadas do tipo MLP.

Assim como no caso dos modelos usados nas previsões de dados de radiômetro, foi calibrado um modelo ANN para cada horizonte de previsão utilizando o conjunto de dados de calibração. A estrutura de ANN utilizada neste trabalho foi a Perceptron Multicamada (MLP), a qual tem sido aplicada com sucesso em previsão de irradiância solar (MELLIT e PAVAN, 2010; VOYANT *et al.*, 2017). A função de ativação adotada foi a Tangente-Sigmóide e o algoritmo de aprendizagem para treinamento da rede foi o de Levenberg-Marquardt. As redes

testadas possuem estruturas com até três camadas ocultas e, em cada uma delas, até oito neurônios, respeitando a característica triangular. Para a tarefa de treinar as várias combinações de ANN, foi usado um cluster de forma semelhante à utilizada na Seção 4.4.1. Em princípio, foram usados como entrada da ANN apenas os dados das 8 componentes principais referentes aos anos de 2013 e 2014 incluindo também as variáveis sazonais (no instante  $t+k$ ) descritas no modelo de dados de radiômetro. Tais modelos foram denominados de Modelos SAT (em alusão às imagens de satélite). Foram treinados os diversos modelos com ordem menor ou igual a 3 (ou seja, usando informações de até três imagens: a atual, no tempo  $t$ , e as duas anteriores, nos tempos  $t-1$  e  $t-2$ ) que atendiam as arquiteturas de rede descritas a pouco. Um diagrama ilustrado na Figura 18 esquematiza as entradas e saídas desse modelo.

**Figura 18 – Esquema usado para treinamento dos Modelos Estatísticos 2, baseados em imagens de satélite. A letra S representa uma variável sazonal. CP representa uma componente principal,  $t$  representa o tempo e  $k$  o horizonte de previsão**



Fonte: O autor

Uma variante do modelo SAT, denominada aqui de PIX, também foi testada. A principal diferença está nos valores-alvos do período de treinamento das redes neurais. O modelo PIX não utiliza dados medidos em terra (radiômetro). Ele os substitui por valores de GHI estimados via Brasil-SR no pixel central no instante  $t+k$ . Dessa forma, esse modelo

poderá fornecer previsões de GHI em sítios onde não haja estações meteorológicas, usando para isso apenas imagens de satélite.

#### 4.4.3 Previsão de Temperatura

De modo a conseguir uma ferramenta de previsão de simples operação e que pudesse ser calibrada para a maioria dos sítios onde se encontram as centrais fotovoltaicas, optou-se por desenvolver um modelo univariado, tendo como entrada apenas a série observacional de temperatura. Tendo em vista sua simplicidade e sua vasta utilização em previsão de séries temporais, o modelo autorregressivo (AR) foi escolhido para esta tarefa. Além disso, as previsões devem possuir *time step* horário e um horizonte final H de 24h à frente.

A seleção da ordem do modelo AR é realizada baseada na análise da função de autocorrelação parcial da série de temperatura, a qual indica quais *lags k* carregam em si uma quantidade de correlação significativa entre a amostra no instante de tempo  $t$  e a amostra no instante de tempo  $t-k$ . Além disso, serão considerados os conhecimentos físicos a respeito da periodicidade diária da temperatura. Para tornar a análise ainda mais substancial, serão testados modelos AR variando da ordem  $p=2$  até a ordem  $p=30$ . Portanto, a seleção final da ordem  $p$  dos modelos AR leva em consideração a função de autocorrelação parcial, a periodicidade diária da temperatura, o RMSE dos diversos modelos testados e a minimização da quantidade de variáveis de entrada.

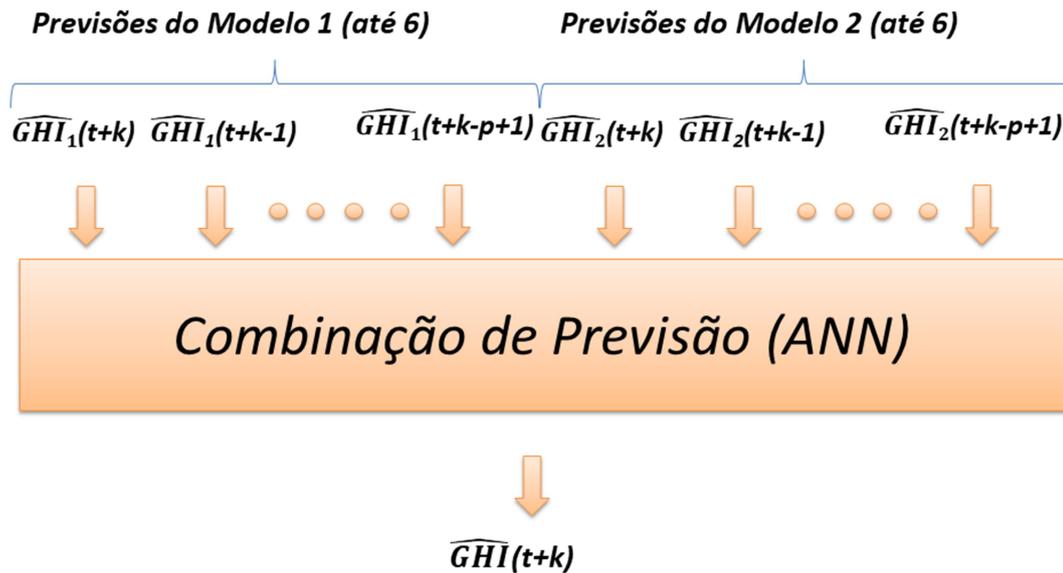
Realizada a seleção da ordem do modelo, os parâmetros do modelo AR serão então estimados através da regressão baseada no método dos mínimos quadrados (MONTGOMERY *et al.*, 2012). Para cada horizonte de tempo será calibrado um modelo AR, isto é, trata-se de modelos de previsão não recursivos.

#### 4.4.4 Combinação de Previsões

Para a tarefa de combinar as previsões foi escolhida a ANN. Trata-se de um modelo já utilizado com sucesso em outras áreas da ciência para realização de combinações de previsão (KHOTANZAD *et al.*, 2000; COSTA, 2005; RENDON e MENEZES, 2016). Novamente foi utilizada uma arquitetura de rede semelhante à utilizada nas Seções 4.4.1 e 4.4.2. Após alguns testes preliminares, a ordem do modelo (tamanho da janela temporal a ser utilizada como entrada) que obteve melhor resultado foi de  $p \leq 6$ , a depender do horizonte de previsão. Foram três camadas ocultas com número máximo de neurônios igual 8, também respeitando a

característica triangular. O diagrama da Figura 19 ilustra as entradas e saídas da combinação para um determinado horizonte de previsão  $k$ .

**Figura 19 – Representação das entradas utilizadas para a realização da combinação de previsões de GHI.**



Fonte: O autor

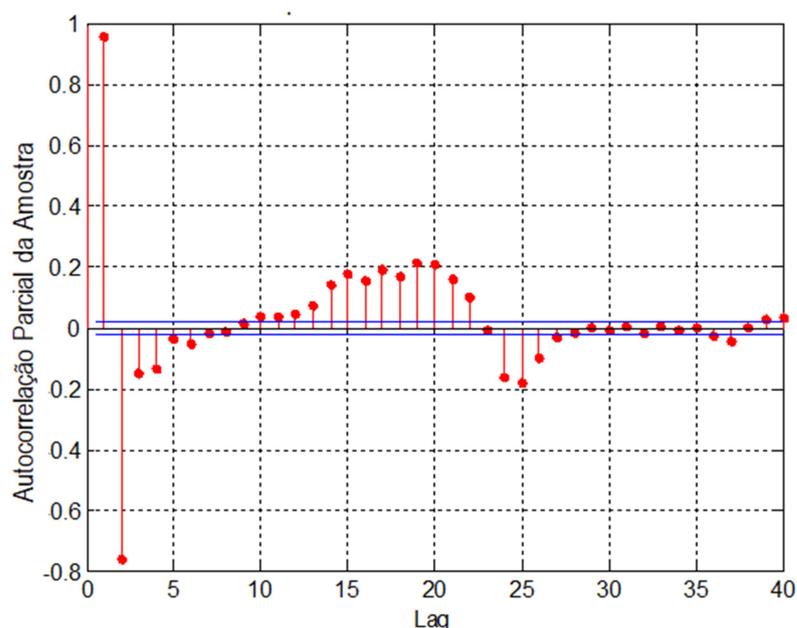
Para a calibração dos modelos de combinação, foram utilizadas as saídas dos modelos treinados SAT e TER calculadas com dados de 2013 e 2014. Já a validação foi feita com as saídas dos modelos SAT e TER calculadas com dados de 2015. Os valores alvos do modelo de combinação durante a calibração são observações calculadas como médias centradas pelo fato de os operadores de Sistema necessitarem de previsões correspondentes aos horários específicos 07h00, 08h00, e assim por diante.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Modelo de Previsão de Temperatura

A primeira ação destinada à escolha da ordem do AR foi a análise da função de autocorrelação parcial, ilustrada na Figura 20. Os valores que se encontram fora da faixa entre as linhas azuis são considerados significativamente diferentes de zero. Nota-se a presença de vários *lags* (deslocamentos no tempo) significativos nas proximidades do *lag* 24.

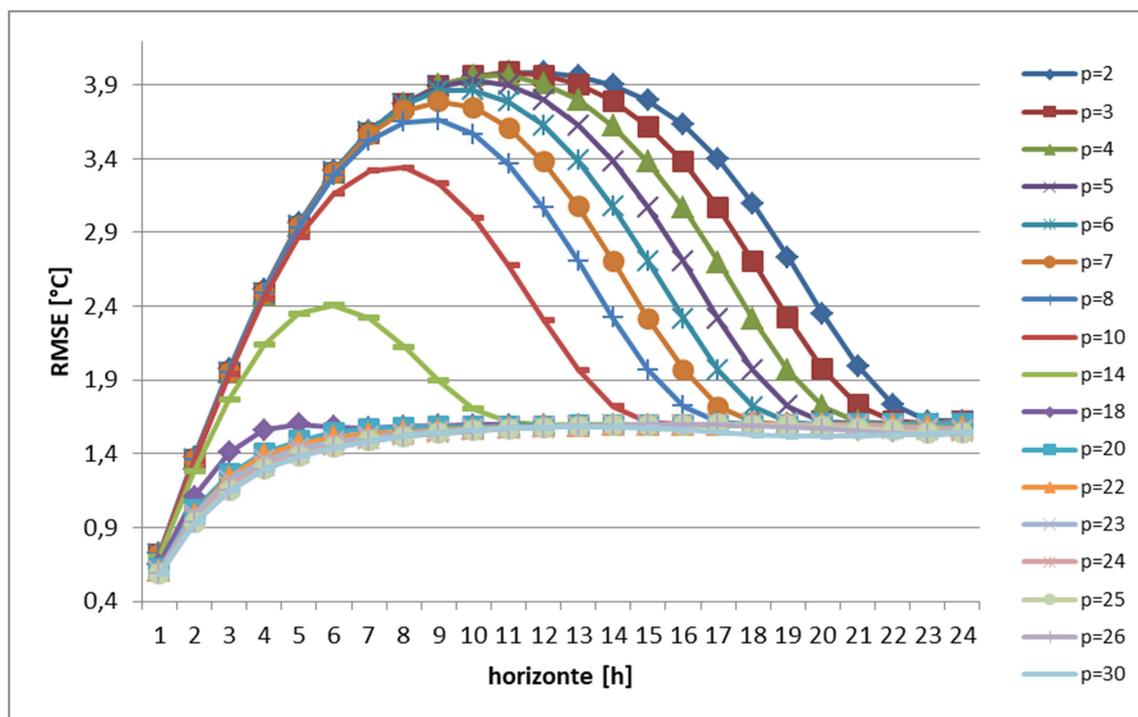
**Figura 20 – Coeficientes de autocorrelação parcial da série de temperatura horária de 2014 em Petrolina.**



As linhas em azul delimitam a região dentro da qual os lags são considerados significativamente iguais a zero (CHATFIELD, 2013). Fonte: O autor.

Para dar maior suporte ao processo, foram testados modelos AR de várias ordens  $p$  a fim de identificar qual a ordem com melhor desempenho nos diversos horizontes de tempo. A Figura 21 mostra que modelos com ordens elevadas apresentaram melhor desempenho para a base de dados de Petrolina. Porém, para valores de  $p > 20$ , as reduções no RMSE foram discretas. Com base na minimização da quantidade de parâmetros, no conhecimento físico da periodicidade diária da temperatura, nos valores da função de autocorrelação parcial (Figura 20) e nos erros provenientes de cada modelo (Figura 21), conclui-se que a ordem do modelo AR a ser escolhida é de  $p=24$ .

**Figura 21 – Erro dos diversos modelos autorregressivos de ordens variando entre 2 e 30 para dados de Petrolina.**



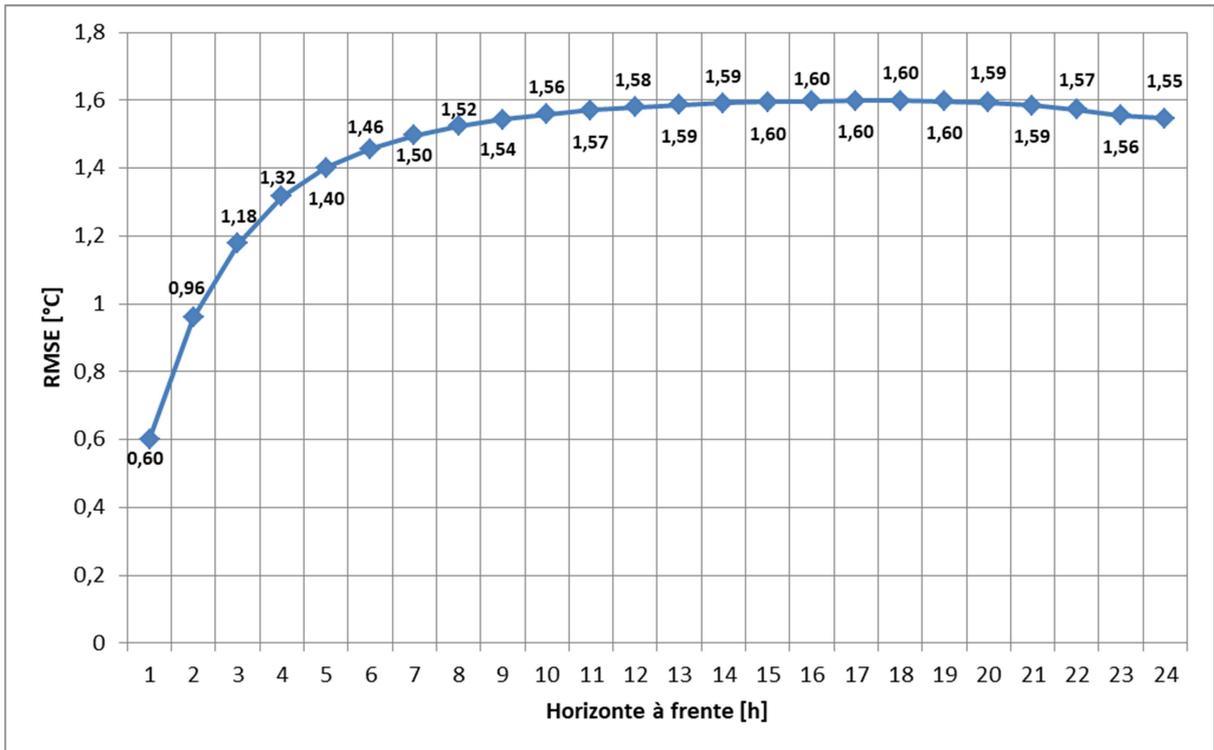
Fonte: O autor.

A alta ordem do modelo autorregressivo para a temperatura destoa de outras variáveis, como por exemplo, velocidade de vento e irradiância, que tipicamente são descritas por processos autorregressivos de ordem 2 (BROWN *et al.*, 1984; NFAOUI *et al.*, 1996). Uma possível explicação para tal característica reside no fato de que a temperatura é uma variável relativamente bem comportada, com alta inércia e que possui um comportamento periódico bem definido, diferentemente do vento, por exemplo.

Em seu trabalho, SMITH *et al.* (2006) testou, dentre outras coisas, o tamanho (ordem do modelo) da janela de valores passados da variável que pudesse produzir melhor desempenho de suas previsões de temperatura via ANN e obteve o melhor tamanho de 24 observações passadas. CHEVALIER *et al.* (2011) também utilizou o mesmo tamanho de janela.

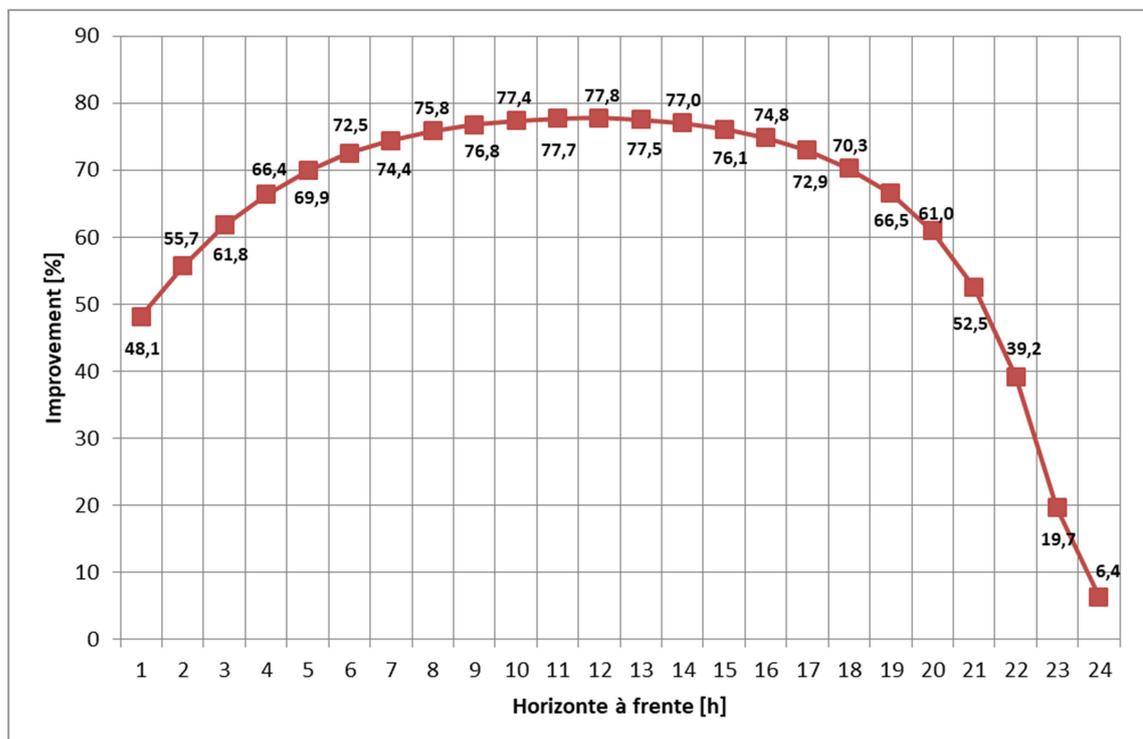
Após treinamento com os dados de Petrolina, os valores de temperatura previstos pelos modelos autorregressivos de previsão de temperatura (um para cada horizonte) foram comparados com os dados observados do conjunto de validação. Os resultados dessa comparação estão representados na Figura 22. Já a comparação dos modelos AR em relação à Persistência é ilustrada na Figura 23.

**Figura 22 – Erros médios (RMSE) cometidos pelos 24 modelos de previsão de temperatura treinados com dados de Petrolina.**



Fonte: O autor.

**Figura 23 – *Improvement* em relação à Persistência dos 24 modelos de previsão de temperatura treinados com dados de Petrolina.**



Fonte: O autor.

Percebe-se pelos resultados que os modelos de previsão de temperatura aqui proposto pode ser considerado adequado para o objetivo ao qual foi pensado, já que obteve erros (RMSE) da ordem de 0,6°C a 1,6°C, dependendo do horizonte. BLANC *et al.* (2017) consideraram que uma incerteza de 2°C de modelos de previsão de temperatura é aceitável, visto que os modelos de conversão não são tão sensíveis em relação à temperatura do ar.

SABINO *et al.* (2018) realizaram uma análise de sensibilidade com modelos semelhantes e obtiveram que erros de 1,59°C provocariam um erro de 1,17% na saída do modelo de conversão radiação→potência de Green/Sandstrom (SANDSTROM, 1967). Portanto, os modelos de previsão de temperatura aqui obtidos podem ser considerados de boa acurácia quando comparados com outros modelos já propostos na literatura, conforme erros mostrados na Tabela 2.

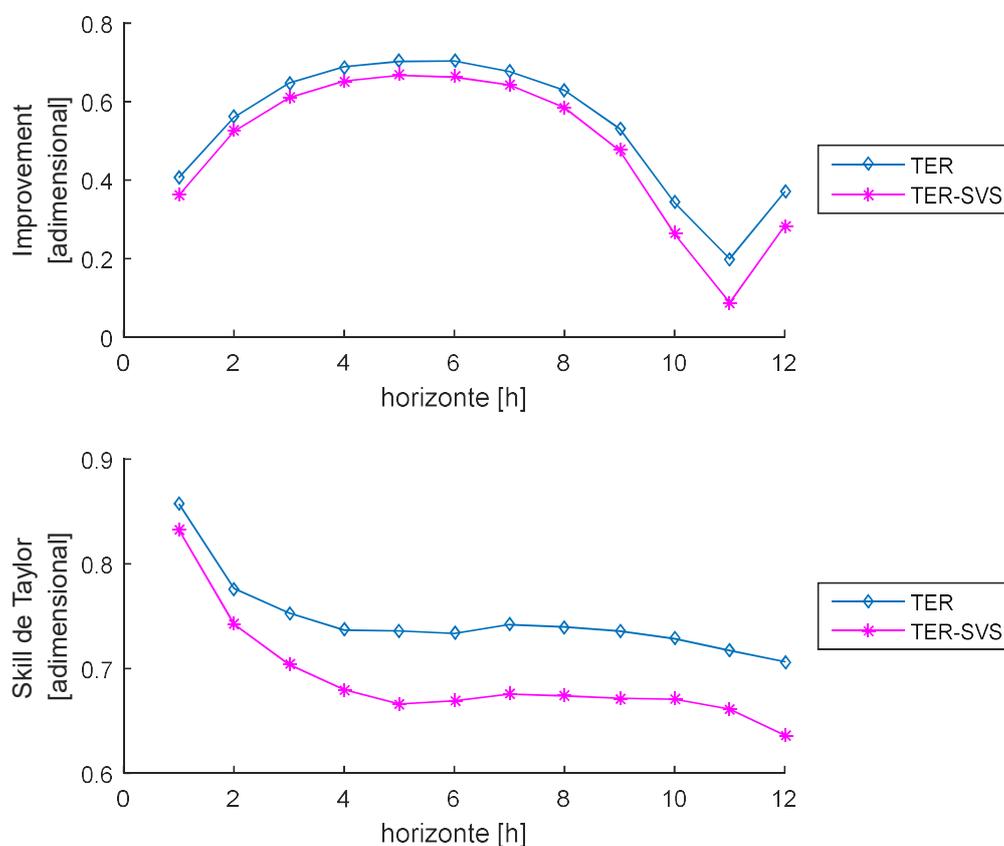
Os resultados dessa mesma metodologia aplicada a Pesqueira e Brasília apresentaram comportamento similar e encontram-se no Apêndice A.

## **5.2 Modelo de Previsão de Radiação Solar baseado em Dados de Radiômetro**

Depois de treinados, os doze modelos foram testados com o conjunto de validação dos dados. O primeiro objetivo foi de avaliar a contribuição da inserção de variáveis sazonais para a melhoria dos resultados de previsão de GHI. Na Figura 24, estão representados os resultados dos modelos de Pesqueira treinados com e sem o uso de variáveis sazonais, ou seja, os modelos TER e TER-SVS, respectivamente. Foi evidenciado que a utilização das variáveis sazonais provoca melhorias em todos os horizontes de previsão. Percebe-se ainda que tais melhorias tendem a se acentuar com o aumento do horizonte.

As variáveis sazonais produziram em Pesqueira incrementos de até onze pontos percentuais no *Improvement*, conforme pode ser verificado na Tabela 8. Os resultados para os demais sítios foram similares (Apêndice B).

**Figura 24 – Comparação entre os modelos TER (com variáveis sazonais) e os modelos TER-SVS (sem variáveis sazonais) para Pesqueira-PE.**



Fonte: O autor.

**Tabela 8 – Improvement dos modelos TER (com variáveis sazonais) e TER-SVS (sem variáveis sazonais) para Pesqueira-PE.**

Modelo	Horizonte [h]											
	1	2	3	4	5	6	7	8	9	10	11	12
TER	40,63%	56,03%	64,74%	68,88%	70,25%	70,34%	67,68%	62,90%	53,12%	34,23%	19,97%	36,98%
TER-SVS	36,10%	52,45%	60,98%	65,25%	66,70%	66,23%	64,25%	58,46%	47,56%	26,27%	8,77%	28,37%

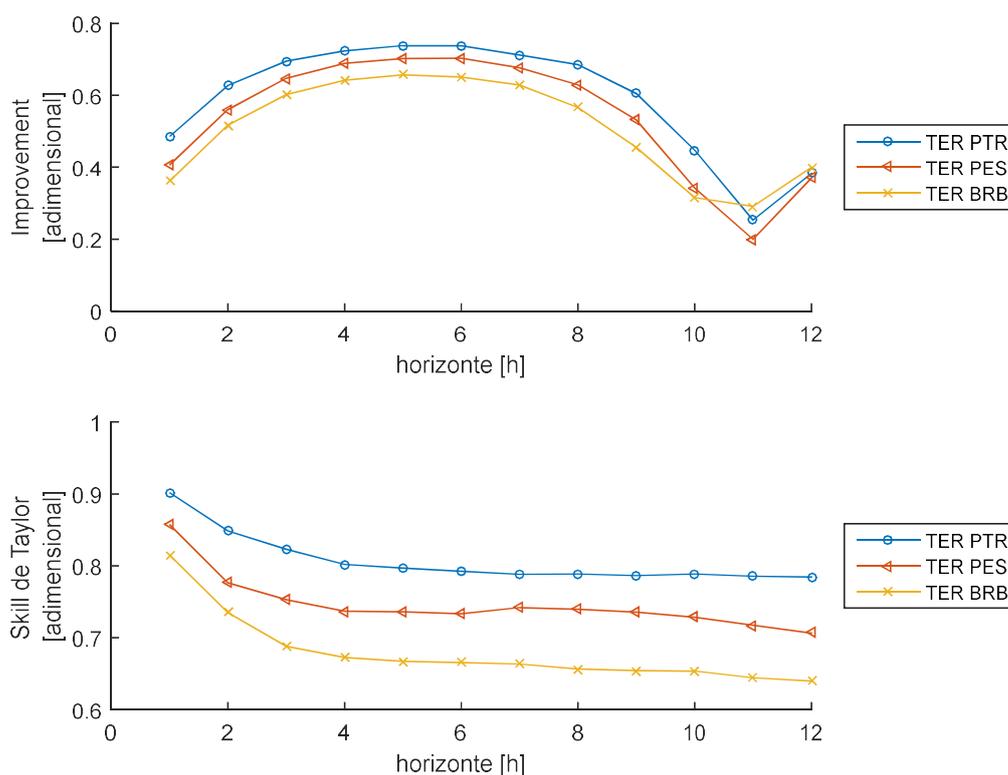
Fonte: O autor.

Em um trabalho semelhante, SABINO *et al.* (2017) realizaram previsão de índice de claridade para Petrolina e chegaram à conclusão que ao incluírem as variáveis sazonais, todos os modelos de previsão apresentaram melhorias, havendo, por exemplo, uma redução no nRMSD de 20,78% para 17,89% no horizonte uma hora à frente.

A comparação entre os resultados das previsões baseadas em dados de radiômetro dos três sítios em estudos é mostrada na Figura 25. A partir dela, é possível inferir preliminarmente uma relação entre a acurácia das previsões e sua respectiva localização. Em

ordem decrescente de acurácia estão as previsões de Petrolina, Pesqueira e Brasília. A respeito desse tema, o exercício de comparação tem mostrado que a acurácia das previsões depende da região, variando conforme a localidade, oscilando, por exemplo, de 20 a 35% na Espanha e de 40% a 60% na Europa Central (PELLAND *et al.*, 2013).

**Figura 25 – Estatísticos dos modelos de previsão baseados em dados de radiômetro de Brasília, Pesqueira e Petrolina.**



Fonte: O autor.

**Tabela 9 – Estatísticos das previsões baseadas em radiômetro (TER) relativas aos três sítios.**

		Horizonte											
		1	2	3	4	5	6	7	8	9	10	11	12
Correlação	PTR	0,95	0,92	0,91	0,90	0,89	0,89	0,89	0,89	0,89	0,89	0,89	0,89
	PES	0,93	0,89	0,88	0,87	0,86	0,87	0,86	0,87	0,87	0,87	0,87	0,86
	BRB	0,90	0,86	0,83	0,82	0,82	0,81	0,82	0,81	0,81	0,81	0,81	0,80
nRMSE	PTR	0,17	0,21	0,23	0,24	0,24	0,25	0,26	0,25	0,26	0,26	0,26	0,26
	PES	0,22	0,28	0,29	0,30	0,31	0,31	0,31	0,30	0,30	0,31	0,31	0,32
	BRB	0,24	0,29	0,31	0,32	0,33	0,33	0,32	0,33	0,33	0,33	0,33	0,34
Improve- ment	PTR	0,49	0,63	0,70	0,72	0,74	0,74	0,71	0,69	0,60	0,45	0,25	0,38
	PES	0,41	0,56	0,65	0,69	0,70	0,70	0,68	0,63	0,53	0,34	0,20	0,37
	BRB	0,36	0,52	0,60	0,64	0,66	0,65	0,63	0,57	0,45	0,32	0,29	0,40

Fonte: O autor.

A acurácia das previsões de radiação também possui relação com a quantidade de nuvens (condição de céu) no local. Os modelos treinados em outros trabalhos (MELLIT e PAVAN, 2010; DIAGNE et al., 2013; BAE et al., 2017) apresentam desempenho inferior em condições de céu nublado e parcialmente nublado. Por essa razão, foi realizada uma investigação nos dados de nebulosidade dos locais citados. Analisando a nebulosidade de cada local na Tabela 4, Petrolina (onde os modelos de previsão obtiveram melhor desempenho) possui o menor valor, com 0,5 décimos, o que caracteriza uma menor presença de nuvens no local. Brasília e Pesqueira aparecem empatadas com nebulosidade média anual de 0,6 décimos.

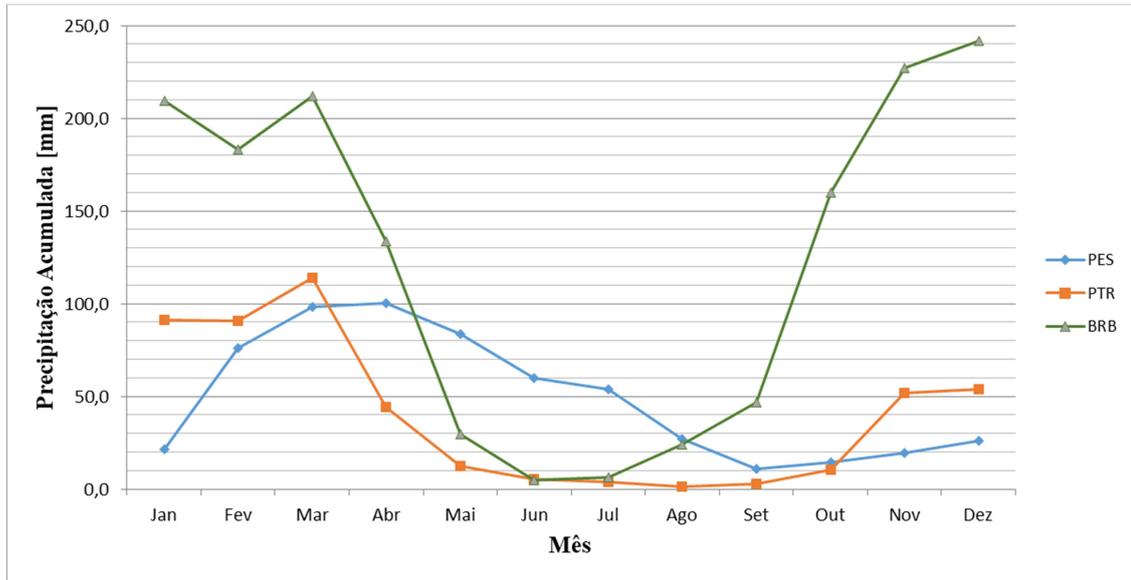
Os dados de nebulosidade fornecidos pelo INMET não são medidos por instrumentos, mas sim estimados através da observação de um funcionário treinado para tal, o que pode gerar imprecisão. Já os dados de precipitação e insolação, medidos por equipamentos, podem fornecer informação indireta a respeito do comportamento das nuvens nesses locais. Olhando por esse ângulo, Petrolina possui uma precipitação média anual acumulada de 482 mm, seguida por Pesqueira, com 591 mm e Brasília, com 1477 mm de chuvas por ano. A distribuição ao longo do ano da precipitação e da insolação pode ser observada no gráfico da Figura 26 e Figura 27. A partir deles é possível tecer os seguintes comentários:

- Petrolina é a localidade com menor nebulosidade dentre as três. Além disso, o período de chuva é bem concentrado (três meses). Em vários meses, a precipitação é bem próxima de zero.
- Embora possua a mesma nebulosidade média anual que Pesqueira, Brasília possui uma precipitação muito acima e mais bem distribuída ao longo dos meses. Seu período de chuvas ao longo do ano é extenso (sete meses).
- A insolação em Brasília é mais sensível à nebulosidade em comparação com Petrolina e Pesqueira, sugerindo que os sistemas meteorológicos em Brasília tenham nuvens com maior desenvolvimento vertical do que Pesqueira e Petrolina.

Diante do exposto, os dados de nebulosidade, insolação e precipitação sugerem que Brasília possui uma maior distribuição de nebulosidade ao longo do ano dentre os três sítios estudados. Além disto, os sistemas meteorológicos que causam precipitações são mais convectivos que na Região Nordeste, apresentando nuvens com desenvolvimento vertical mais profundo, portanto com maior influência na insolação. Essas são as prováveis razões para que Brasília também fosse o sítio com previsões menos acuradas, visto que a insolação é mais sensível ao tipo de nuvem. Os dados climatológicos sugerem que a ordem crescente dos

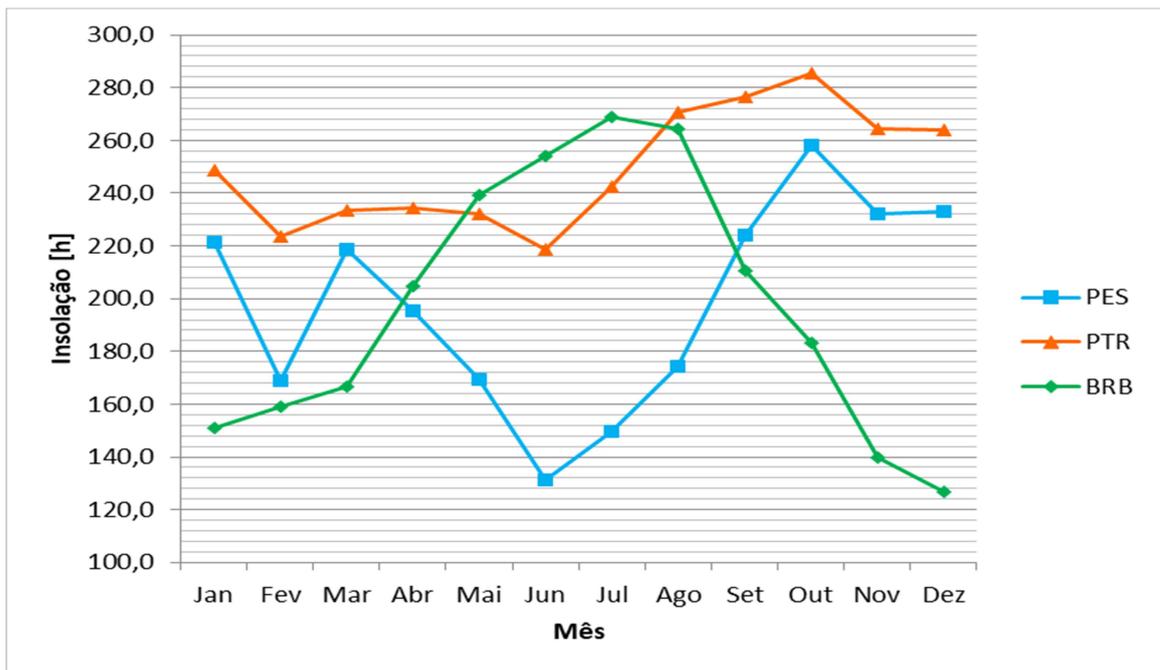
locais com maior quantidade de meses com nuvens é Petrolina, Pesqueira e Brasília, que é a ordem inversa em relação à acurácia das previsões.

**Figura 26 – Precipitação média mensal de Pesqueira, Petrolina e Brasília, segundo dados da Normal Climatológica de 1981-2010, do INMET.**



Fonte: O autor.

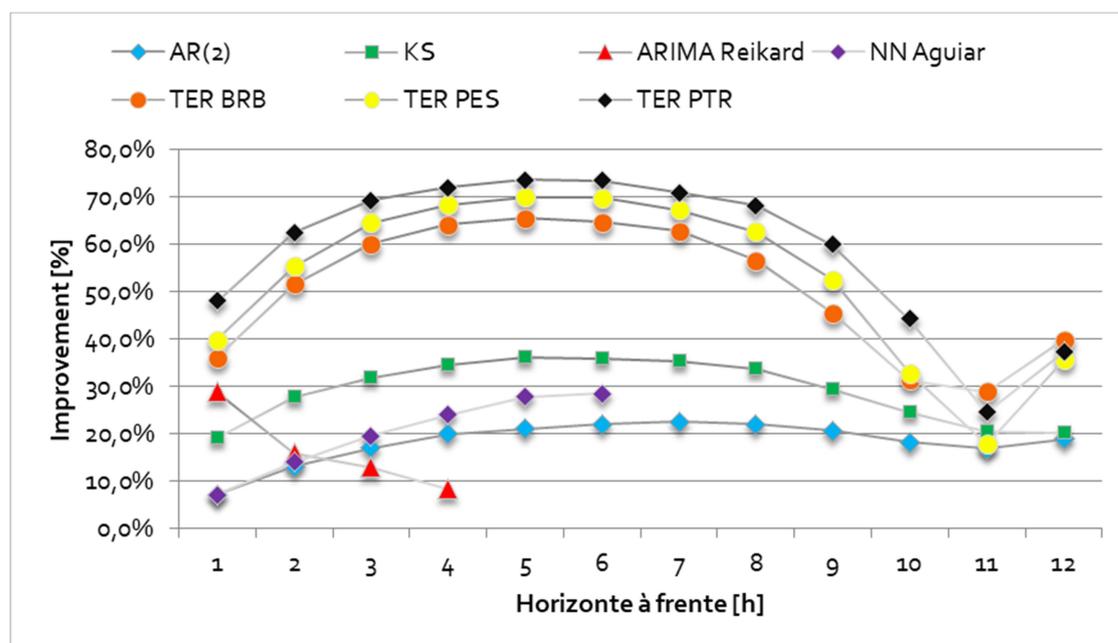
**Figura 27 – Insolação média mensal de Pesqueira, Petrolina e Brasília, segundo dados da Normal Climatológica de 1981-2010, do INMET.**



Fonte: O autor.

Mesmo Brasília apresentando pior desempenho entre as três localidades, ainda sim seus resultados são considerados satisfatórios, visto que atingiram *Improvement* variando de 29% a 66%. Alguns estatísticos dos modelos para os três sítios estão descritos na Tabela 9. Comparados com outros autores da literatura que usaram apenas dados de radiômetros, os modelos propostos neste trabalho foram superiores em praticamente todos os horizontes. Na Figura 28 é realizada uma comparação com *Improvements* de trabalhos de outros autores, os quais estão com suas características resumidas na Tabela 10. Vale salientar que os modelos KS (desenvolvidos por SABINO *et al.* (2017) e que possuem metodologia semelhante aos dos modelos TER, porém usando como entradas Índice de claridade ao invés de GHI) apresentaram previsões com acurácias menores em relação aos modelos TER. Esse fato evidencia preliminarmente que os modelos baseados em ANN responderam melhor quando as entradas foram dados de GHI. Em resumo, os modelos propostos neste trabalho se mostraram superiores aos dos demais autores em praticamente todos os horizontes, aumentando o *Improvement* em até mais de 50 pontos percentuais.

**Figura 28 – Improvement (baseado no RMSE) calculado para os modelos proposto neste trabalho (TER BRB, TER PES, TER PTR e modelos de outros autores baseado em dados medidos por radiômetro.**



Fonte: O autor.

**Tabela 10 – Descrição de modelos de previsão baseados em dados de radiômetros de outros autores.**

Nome do Autor	Local	Descrição
<b>Modelo</b>		
<b>ARIMA</b> (REIKARD, 2009)	Clark, Nevada, USA	Baseado em ARIMA e dados observacionais (GHI)
<b>NN Aguiar</b> (AGUIAR et al., 2016)	Ilhas Canárias, Espanha	Baseado em ANN e dados observacionais (GHI)
<b>AR(2)</b> (SABINO et al., 2017)	Petrolina, Brasil	Baseado em AR e Índice de Claridade (Kt)
<b>KS</b> (SABINO et al., 2017)	Petrolina, Brasil	Baseado em ANN e Índice de Claridade (Kt)

Fonte: O autor.

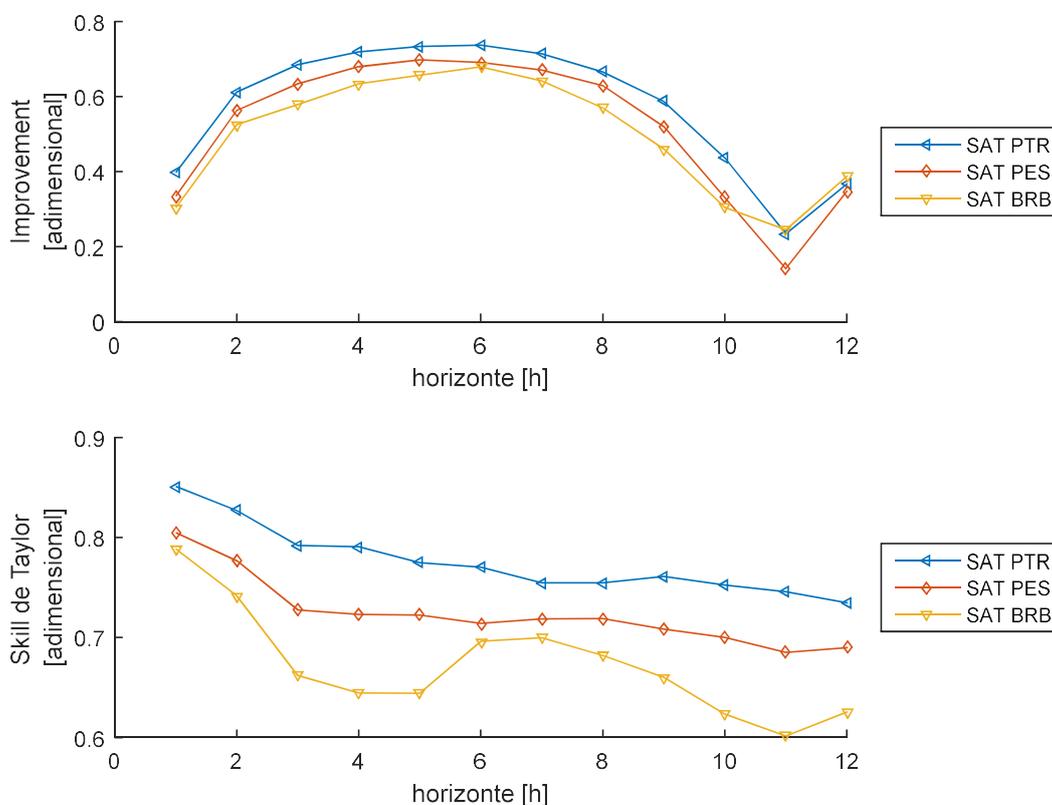
### 5.3 Modelo de Previsão de Radiação Solar baseado em Imagens de Satélite

Nesta seção serão apresentados os resultados de duas classes de modelos de previsão de radiação solar que usam como entrada, além das variáveis determinísticas sazonais, informações de imagens de satélite. A diferença entre as duas está nos dados que são utilizados como alvo pelas ANN em seu treinamento. A primeira classe, denominada SAT, é treinada para reproduzir no futuro o comportamento do seu alvo, que são as medidas de radiômetro, realizadas em solo. Já a segunda classe, denominada PIX, é treinada para reproduzir os valores de GHI estimados (via imagens de satélite) do pixel central. O principal objetivo desse último modelo é poder realizar previsões para lugares onde não tenham medidas de GHI oriundas de estações meteorológicas.

#### 5.3.1 Modelo de satélite treinado com informações de Radiômetro como alvo (SAT)

Os resultados da validação dos modelos SAT para os três sítios são mostrados na Figura 29. Nela é possível observar que a ordem decrescente de acurácia dos modelos foi a mesma dos modelos baseados em dados de radiômetro, ou seja, Petrolina, Pesqueira e Brasília. As maiores diferenças entre os modelos de Petrolina e Brasília atingiram 14 e 8 pontos percentuais no Improvement e no nRMSE respectivamente, conforme Tabela 11.

**Figura 29 – Estatísticos dos modelos de previsão (SAT) baseadas em imagens de Brasília, Pesqueira e Petrolina.**



Fonte: O autor.

**Tabela 11 – Estatísticos das previsões baseadas em satélite relativas aos três sítios.**

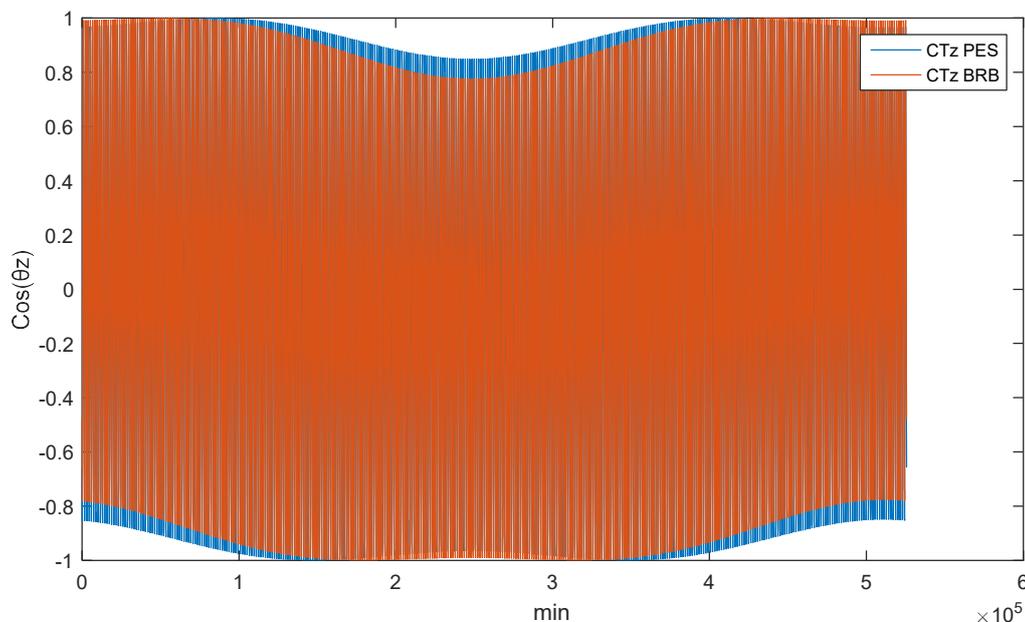
		Horizonte											
		1	2	3	4	5	6	7	8	9	10	11	12
Correlação	PTR	0,93	0,91	0,90	0,89	0,89	0,90	0,89	0,88	0,88	0,88	0,88	0,87
	PES	0,91	0,90	0,87	0,86	0,86	0,86	0,86	0,86	0,86	0,86	0,84	0,84
	BRB	0,89	0,86	0,81	0,81	0,81	0,84	0,84	0,83	0,82	0,80	0,77	0,79
nRMSE	PTR	0,22	0,24	0,26	0,27	0,28	0,29	0,30	0,30	0,28	0,28	0,27	0,28
	PES	0,27	0,30	0,31	0,32	0,34	0,34	0,34	0,32	0,35	0,33	0,32	0,31
	BRB	0,26	0,29	0,34	0,34	0,34	0,34	0,34	0,33	0,33	0,33	0,34	0,34
Improve- ment	PTR	0,40	0,61	0,69	0,72	0,73	0,74	0,71	0,67	0,59	0,44	0,23	0,37
	PES	0,33	0,56	0,63	0,68	0,70	0,69	0,67	0,63	0,52	0,33	0,14	0,35
	BRB	0,30	0,53	0,58	0,63	0,66	0,68	0,64	0,57	0,46	0,30	0,24	0,39

Fonte: O autor.

Vale notar que, em relação ao *Improvement*, Brasília apresenta desempenho inferior às demais, com exceção dos horizontes 11 e 12. Isso provavelmente pode ter relação com a diferença entre as faixas de valores do cosseno do ângulo zenital. Como o  $\cos(\theta_z)$  leva em consideração a latitude, o valor maior de latitude de Brasília provoca uma leve diminuição na faixa de variação do  $\cos(\theta_z)$  (Figura 30) e, conseqüentemente, na hora na qual a série temporal

é recortada. Tal fato fez com que alguns dias possuíssem apenas 10 horas de janela válida. O resultado disso é que no décimo primeiro horizonte há uma passagem para o valor correspondente do dia seguinte e esse fato influenciou o desempenho dos modelos.

**Figura 30 – Variação do Cosseno do ângulo zenital ao longo do ano para Pesqueira e Brasília.**

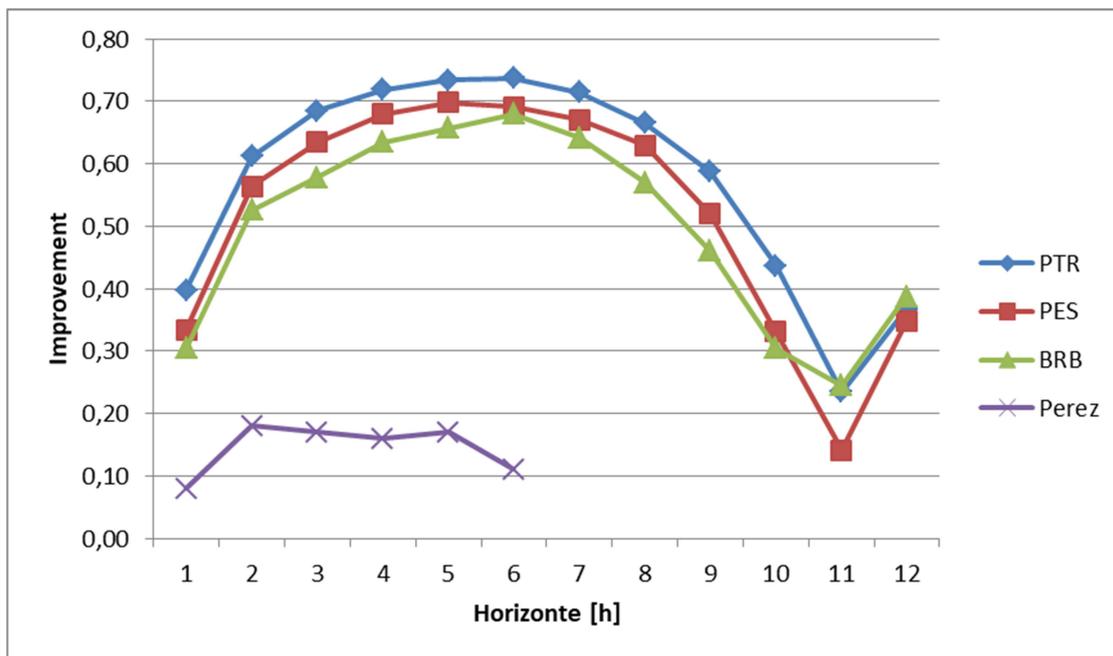


Fonte: O autor.

Considerando que os erros de partida dos modelos de previsão, isto é, os erros das entradas, eram altos (Tabela 6), da ordem de 23% a 26% em média para os três sítios, os resultados encontrados na saída atingiram erros de 22% a 35% (Tabela 11), o que aparenta ser pertinente. "Os desempenhos dos modelos estatísticos de séries temporais baseados em satélites são intrinsecamente limitados pelas incertezas das próprias estimativas baseadas em satélite" (BLANC *et al.*, 2017). Já GENSLER *et al.* (2016) são mais otimistas nessa questão, afirmando que modelos estatísticos ou de aprendizagem de máquinas podem abstrair erros inerentes ao processo de aquisição das medidas.

Os resultados dos modelos de previsão via satélite aqui propostos apresentaram bons resultados comparados aos modelos de outros autores. Na Figura 31 é possível realizar uma comparação com o modelo proposto por (PEREZ *et al.* (2010)), o qual também é baseado em imagens de satélite, e atingiu *Improvements* variando de 8% a 18% para horizontes de 1h a 6h à frente. O RMSE do modelo de Perez atingiu de 26% (1h) até 41%(6h) para o sítio de Fort Peck, Estados Unidos.

**Figura 31 – Improvement dos modelos aqui propostos de previsão via satélite (SAT) e do modelo de Perez, também baseado em imagens de satélite.**

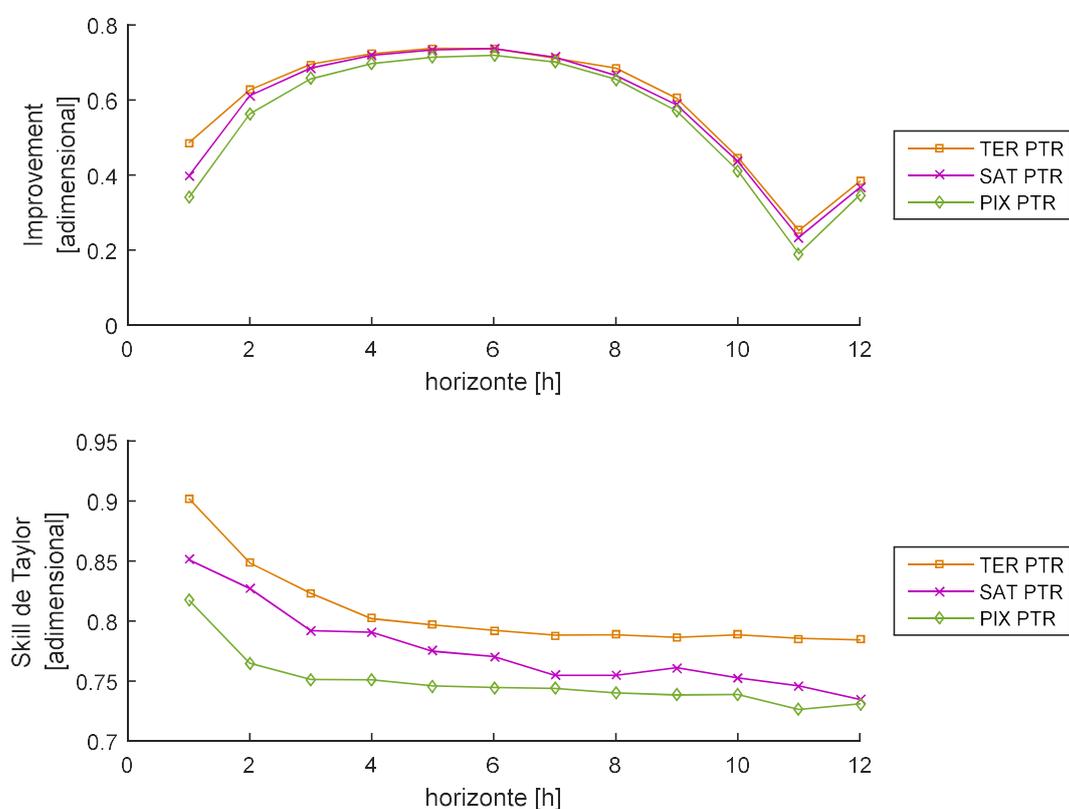


Fonte: O autor.

### 5.3.2 Modelo de satélite treinado com informações do pixel central como alvo (PIX)

O modelo PIX foi treinado com os dados estimados oriundos do pixel central das imagens de satélite como sendo alvo. As entradas do modelo foram idênticas às do modelo SAT, isto é, componentes principais das imagens e variáveis sazonais. Já no período de validação, as previsões foram comparadas com dados de radiômetro, a fim de testar sua acurácia. Os resultados dos modelos PIX de Petrolina são comparados com os modelos de radiômetro (TER) e de imagens de satélite (SAT) na Figura 32. Nela é possível observar que, apesar de estar com acurácia sempre abaixo dos demais modelos, o modelo PIX conseguiu atingir níveis de *improvement* e de *Skill Taylor* muito próximos aos dos modelos que necessitam de dados de radiômetros para serem treinados. A diferença média entre os *nRMSEs* do modelo PIX e do modelo TER é da ordem de 4 pontos percentuais, conforme Tabela 12. Esse é um importante resultado que viabiliza uma previsão de acurácia aceitável para lugares que não disponham de medida de irradiância. Além disso, as entradas desse modelo são imagens de satélite, as quais podem fornecer séries temporais de mais de 10 anos de dados.

**Figura 32 – Comparação entre tipos de modelos: resultados dos modelos baseados em dados de radiômetro (TER), dados de satélite e radiômetro (SAT) e somente em dados de satélite (PIX) para Petrolina.**



Fonte: O autor.

O *improvement* do PIX em Petrolina variou de 34% a 72% (Tabela 12) nos horizontes de 1h a 6h à frente, enquanto no modelo de Perez variou de 8% a 18% (Figura 31). É importante salientar que Perez utilizou dados medidos em terra na calibração de seu modelo, enquanto que o modelo PIX proposto utiliza apenas dados de satélite. Perez transformou a irradiância em índice de céu claro (variável mais persistente), o que favorece o seu modelo de Persistência e desfavorece o *Improvement* (melhora em relação à Persistência). Essa pode ser uma das razões para a discrepância entre os resultados de seu modelo e do modelo PIX.

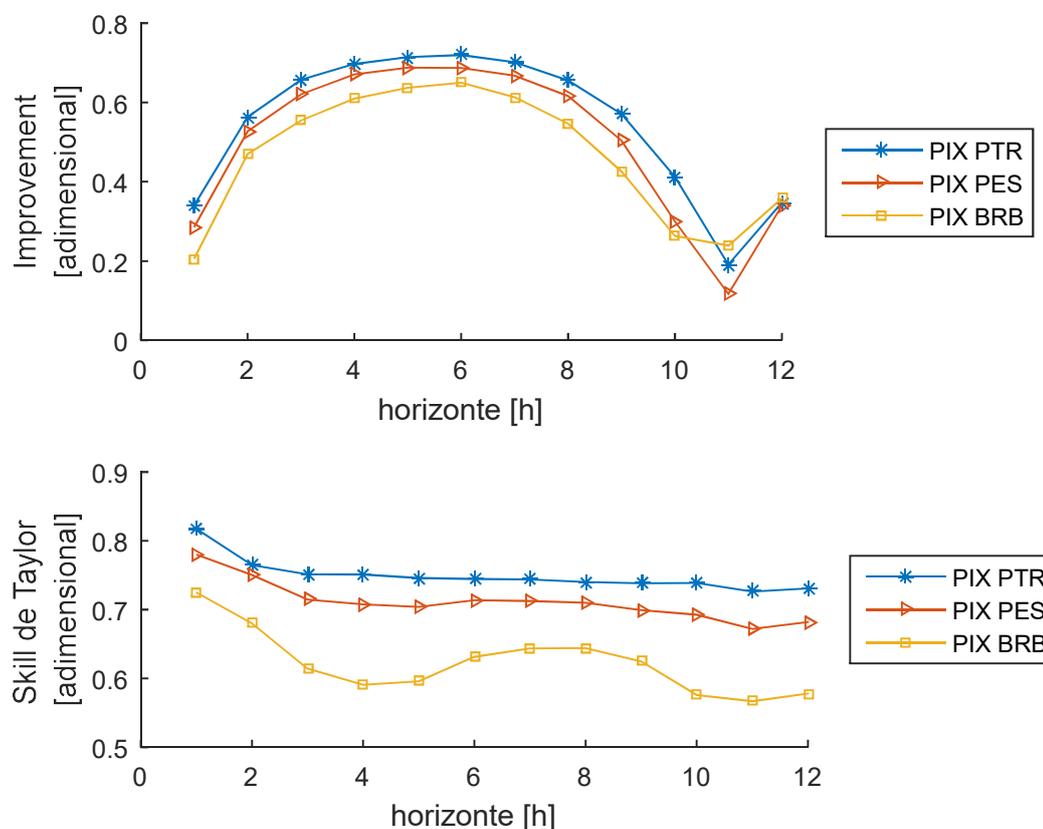
**Tabela 12 – Estatísticos dos modelos de radiômetro (TER), de satélite (SAT) e baseados no pixel central (PIX) de Petrolina.**

		Horizonte											
		1	2	3	4	5	6	7	8	9	10	11	12
Correlação	TER	0,95	0,92	0,91	0,90	0,89	0,89	0,89	0,89	0,89	0,89	0,89	0,89
	SAT	0,93	0,91	0,90	0,89	0,89	0,90	0,89	0,88	0,88	0,88	0,88	0,87
	PIX	0,91	0,89	0,88	0,87	0,87	0,88	0,87	0,87	0,87	0,87	0,86	0,86
nRMSE	TER	0,17	0,21	0,23	0,24	0,24	0,25	0,26	0,25	0,26	0,26	0,26	0,26
	SAT	0,22	0,24	0,26	0,27	0,28	0,29	0,30	0,30	0,28	0,28	0,27	0,28
	PIX	0,22	0,25	0,27	0,28	0,28	0,29	0,30	0,30	0,29	0,28	0,27	0,28
Improve- ment	TER	0,49	0,63	0,70	0,72	0,74	0,74	0,71	0,69	0,60	0,45	0,25	0,38
	SAT	0,40	0,61	0,69	0,72	0,73	0,74	0,71	0,67	0,59	0,44	0,23	0,37
	PIX	0,34	0,56	0,66	0,70	0,71	0,72	0,70	0,66	0,57	0,41	0,19	0,35

Fonte: O autor.

Em relação ao aspecto da localização, o modelo PIX seguiu a tendência dos modelos de terra (TER) e de satélite (SAT). Petrolina, Pesqueira e Brasília (nessa ordem) atingiram os melhores índices de acurácia de suas previsões, conforme mostrado na Figura 33.

**Figura 33 – Comparação entre locais: resultados dos modelos baseados somente dados de satélite (PIX) para Brasília, Pesqueira e Petrolina.**



Fonte: O autor.

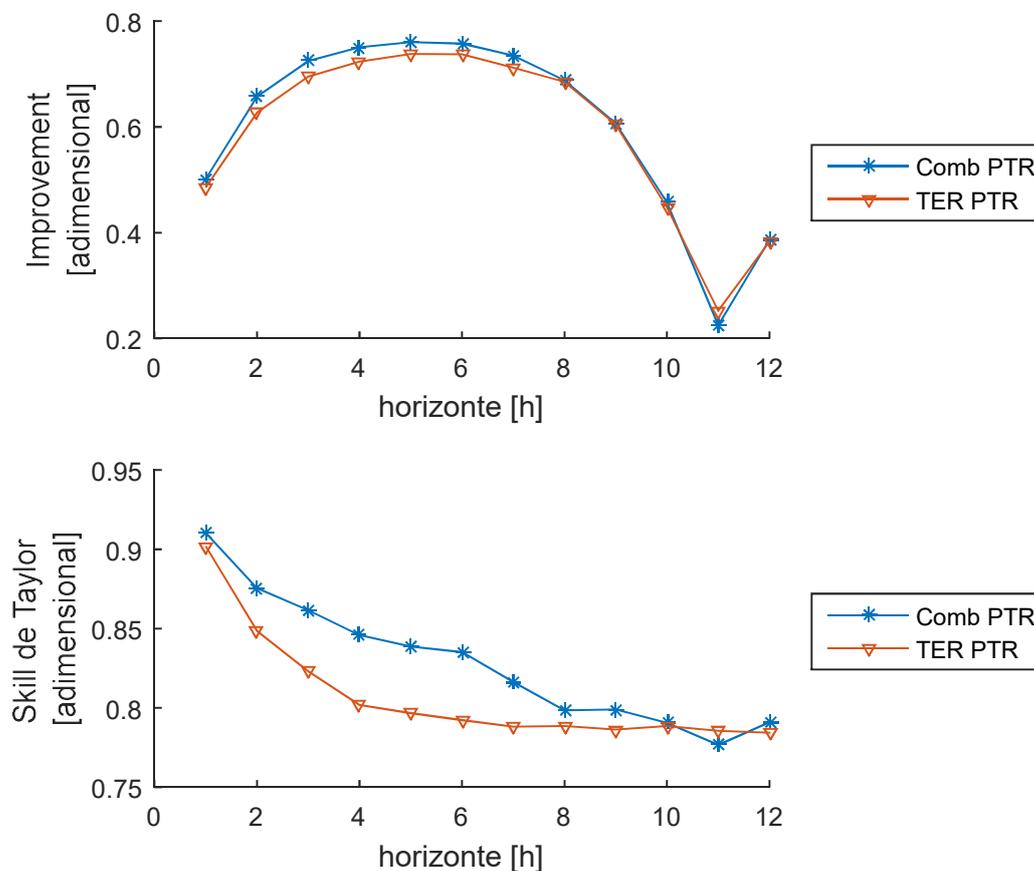
Levando em conta que o modelo PIX é uma ótima alternativa para previsões de GHI em lugares que não dispõem de dados de radiômetro, é interessante realizar a comparação deste com modelos de previsão baseados em dados de NWP (que também podem ser utilizados para tais situações). O modelo de Petrolina alcançou um *Improvement* médio de 55%, para 12 passos à frente. Já o modelo de QING e NIU (2018), que utiliza como entradas previsões de NWP de variáveis meteorológicas, apresentou um *Improvement* de 30,6% para horizontes de até um dia à frente.

Em relação ao erro, LIMA (2015) testou para o nordeste brasileiro tanto previsões de GHI do modelo numérico WRF quanto um refinamento desse modelo através de ANN e encontrou RMSE variando (de acordo com o sítio) de 37% a 45% e de 22% a 32%, respectivamente. O modelo PIX proposto neste trabalho atingiu RMSE médio de 28% para Petrolina.

#### **5.4 Combinação de Previsões de Radiação Solar**

Os resultados da combinação de previsões foram comparados ao modelo TER (baseados em dados medidos em terra), considerado o melhor modelo individual. Na Figura 34 pode-se verificar a superioridade dos modelos de combinação em relação ao modelo individual na maioria dos horizontes de previsão.

**Figura 34 – Comparação entre o melhor modelo individual e a combinação de previsão para Petrolina.**



Fonte: O autor.

Na Tabela 13 pode-se realizar uma comparação quantitativa entre os resultados dos dois modelos individuais e a combinação. Os melhores resultados estão destacados em vermelho. Percebe-se que o uso da combinação provocou no *Improvement* um ganho maior entre os horizontes 2 a 7, que apresentaram em média um aumento de 2,5 pontos percentuais. Já no horizonte 11 houve redução significativa do *Improvement*. Uma análise do *Skill* de Taylor mostra que houve um aumento médio de 3,7 pontos percentuais nos horizontes de 2 a 7 (comparados ao modelo de terra). Nos outros horizontes o aumento foi de 2,1 pontos percentuais em média, e em apenas um horizonte, o horizonte 11, onde houve uma diminuição do *Skill* com o uso da combinação.

**Tabela 13 – Estatísticos dos modelos individuais (de satélite e de terra) e da combinação para Petrolina. A diferença de *Improvement* é calculada em relação ao modelo de radiômetro.**

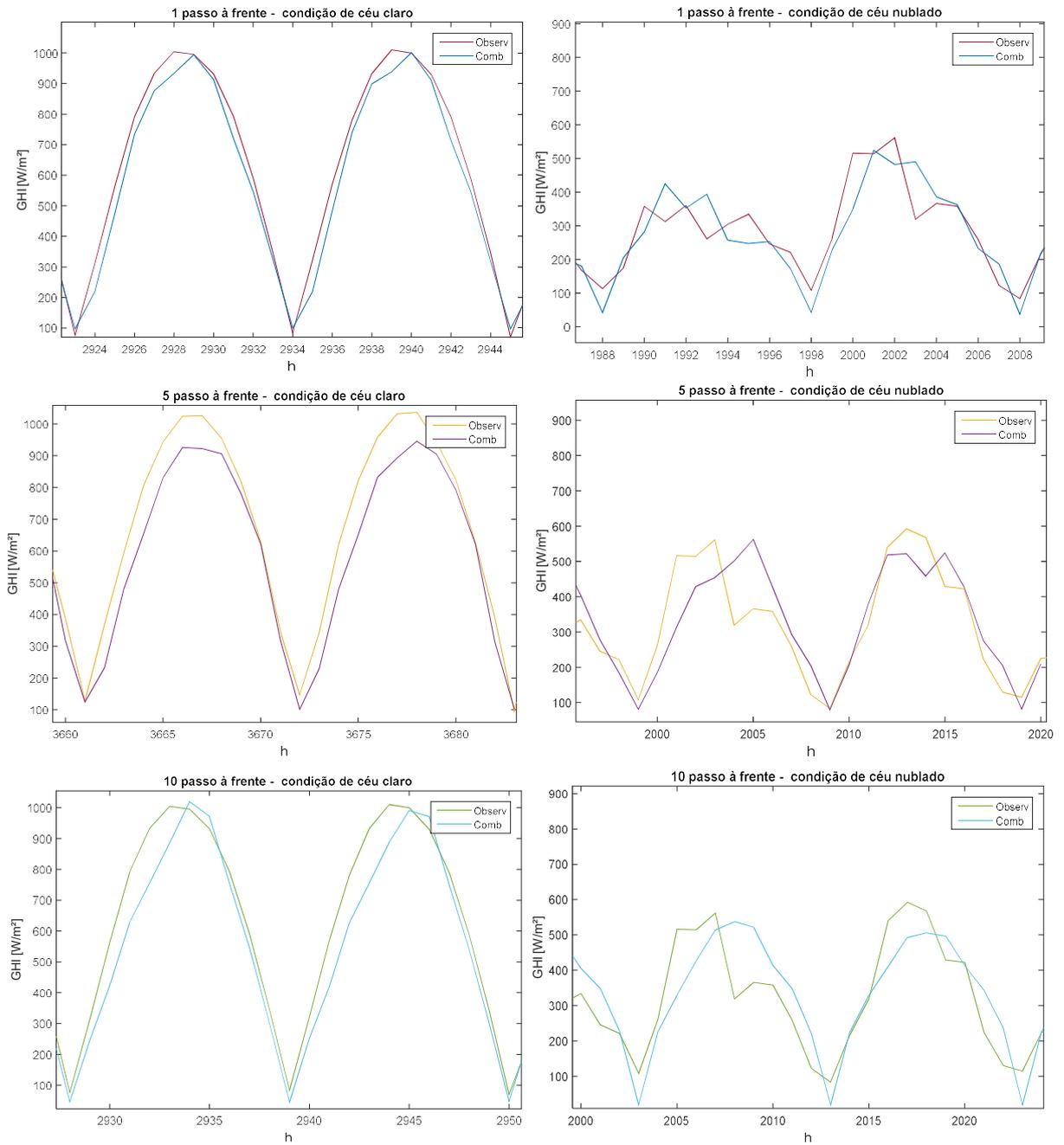
Modelo	Horizonte (h)											
	1	2	3	4	5	6	7	8	9	10	11	12
Improvement em relação à Persistência												
TER	48,55%	62,71%	69,53%	72,34%	73,82%	73,74%	71,20%	68,52%	60,47%	44,74%	25,25%	38,29%
SAT	39,74%	61,21%	68,53%	71,91%	73,37%	73,74%	71,46%	66,56%	58,71%	43,64%	23,45%	36,65%
COMB	50,16%	65,77%	72,47%	75,06%	76,04%	75,76%	73,49%	68,82%	60,70%	45,69%	22,48%	38,64%
Diferença de improvement tomando como base o improvement do modelo TER												
TER	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
SAT	-8,81%	-1,51%	-1,00%	-0,43%	-0,45%	-0,01%	0,26%	-1,96%	-1,76%	-1,09%	-1,80%	-1,64%
COMB	1,61%	3,06%	2,94%	2,72%	2,22%	2,02%	2,29%	0,30%	0,22%	0,95%	-2,77%	0,35%
Skill de Taylor												
TER	90,20%	84,84%	82,31%	80,19%	79,67%	79,23%	78,82%	78,86%	78,64%	78,86%	78,55%	78,43%
SAT	85,10%	82,70%	79,20%	79,06%	77,48%	77,04%	75,47%	75,48%	76,10%	75,25%	74,59%	73,47%
COMB	91,04%	87,55%	86,15%	84,59%	83,87%	83,52%	81,64%	79,85%	79,90%	79,06%	77,67%	79,10%
Diferença de Skill tomando como base o Skill do modelo TER												
TER	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
SAT	-5,10%	-2,14%	-3,11%	-1,13%	-2,19%	-2,19%	-3,35%	-3,38%	-2,54%	-3,61%	-3,96%	-4,96%
COMB	0,83%	2,71%	3,84%	4,40%	4,19%	4,29%	2,82%	0,99%	1,26%	0,20%	-0,88%	0,67%

Fonte: O autor.

Diante dos resultados, considera-se que o uso da combinação entre as previsões apresentaram êxito, melhorando os resultados das previsões de GHI em Petrolina. Na Figura 35 são mostradas as séries observadas e previstas para alguns horizontes  $k$  de Petrolina em períodos de céu claro e céu nublado. Os modelos se mostraram visualmente bem aderentes nas duas condições de céu e também nos diversos horizontes de previsão. Já a Figura 36 traz a dispersão entre as séries medida e prevista para os horizontes 1, 5 e 10 passos à frente. Nela é possível perceber um aumento da dispersão à medida que o horizonte de previsão aumenta, o que é o comportamento esperado, já que a tendência é que a acurácia das previsões tenda a se degradar com o aumento do horizonte (PELLAND *et al.*, 2013). Também é possível perceber (principalmente nos horizontes maiores) uma tendência de superestimação dos valores previstos, isto é, um *bias* (viés) positivo. É possível que, utilizando técnicas de correção de *bias*, se consiga diminuir esse efeito. Salienta-se que, em modelos de previsão, o erro cometido por *bias* são menos complicado de se corrigir comparados a erros na frequência e na fase.

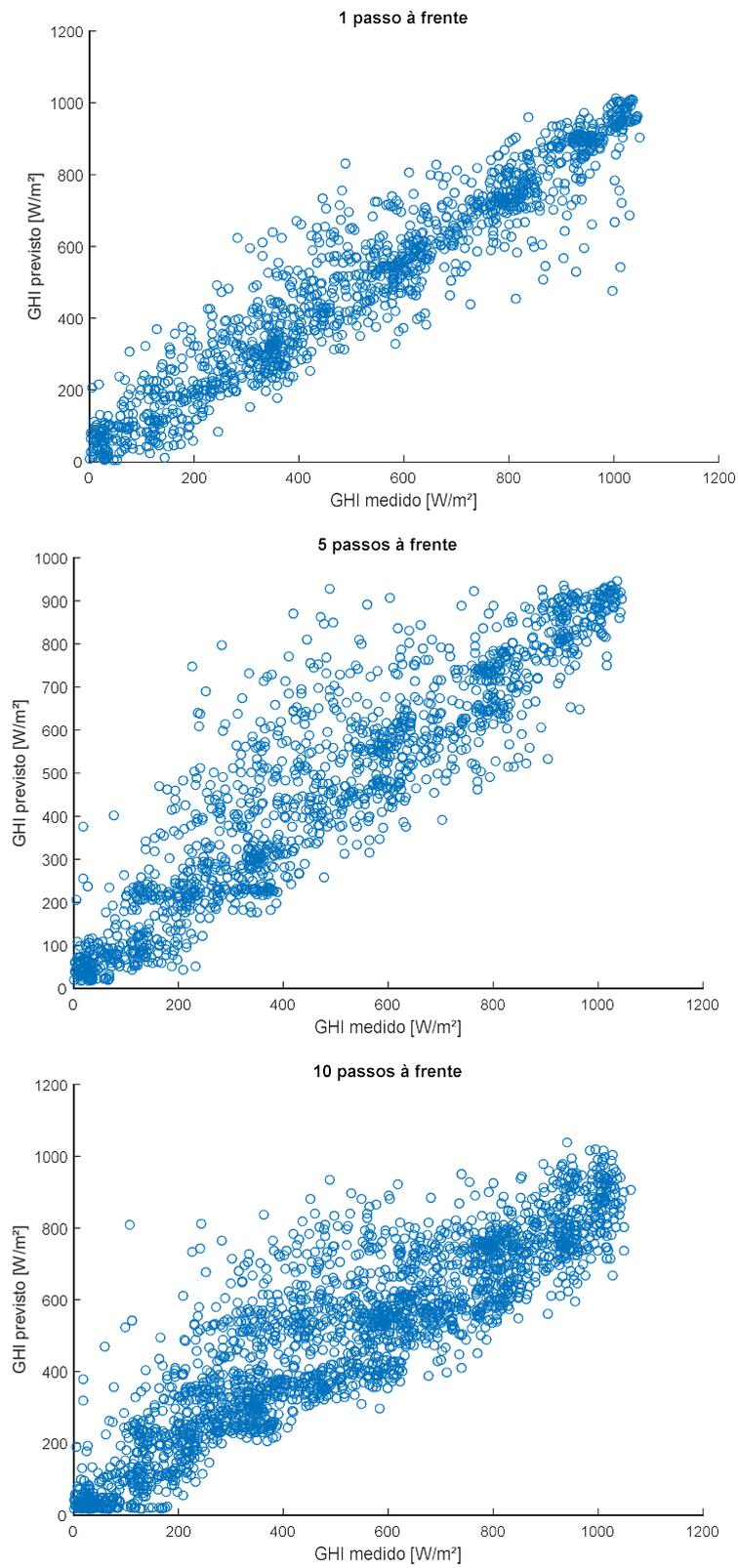
Analisando a correlação e o nRMSD médios, os quais não são afetados pelo *bias*, a combinação produziu uma melhoria discreta nos mesmos, da ordem de 1 e 0,3 pontos percentuais respectivamente, comparados ao modelo de radiômetro.

**Figura 35 – Séries observada e prevista pela combinação de previsão para Petrolina (horizontes 1, 5 e 10 passos à frente).**



Fonte: O autor.

**Figura 36 – Dispersão entre a série observada e série prevista para Petrolina (horizontes 1, 5 e 10 passos à frente).**



Fonte: O autor.

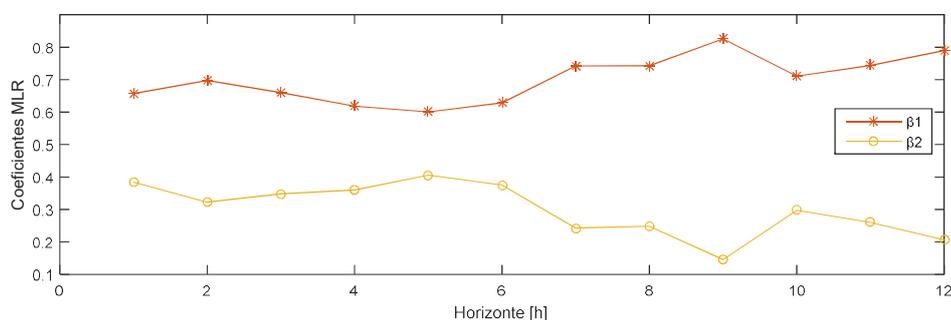
#### 5.4.1 Análise Estatística da Complementariedade das Entradas da Combinação de Previsões

Nesta seção, uma análise estatística é realizada com o objetivo de identificar aspectos que contribuíram para a melhoria das previsões usando combinação das previsões TER e SAT. A combinação descrita na Seção 3.4 foi realizada através de um modelo de redes neurais com até três camadas ocultas, o que torna a estrutura complexa e de difícil interpretação de seus parâmetros. Foi então proposta a utilização de uma ferramenta de combinação mais simples, a regressão multilinear (*Multiple Linear Regression - MLR*), e com menos entradas (pois não utiliza valores de previsões em instantes anteriores a  $t+k$ ). O objetivo de treinar tal combinação foi de poder analisar a contribuição de cada entrada no resultado final (saída combinada) de tais previsões.

O MLR utilizado é uma combinação linear das previsões SAT e TER multiplicadas pelos parâmetros  $\beta_1$  e  $\beta_2$ , respectivamente, conforme Equação (11). O parâmetro  $\beta_0$  representa o viés do modelo. Os valores dos parâmetros  $\beta_1$  e  $\beta_2$  para os 12 modelos de previsão (um para cada horizonte) estão ilustrados no gráfico da Figura 37. Percebe-se que os dois coeficientes possuem valores da mesma ordem de magnitude, o que elimina a hipótese de que a combinação pudesse estar desprezando uma das previsões de entrada e usando informação apenas da outra previsão.

$$MLR(t + k) = \beta_0 + \beta_1 \times SAT(t + k) + \beta_2 \times TER(t + k) \quad (11)$$

**Figura 37 – Coeficientes da combinação MLR dos diversos horizontes para Petrolina.**

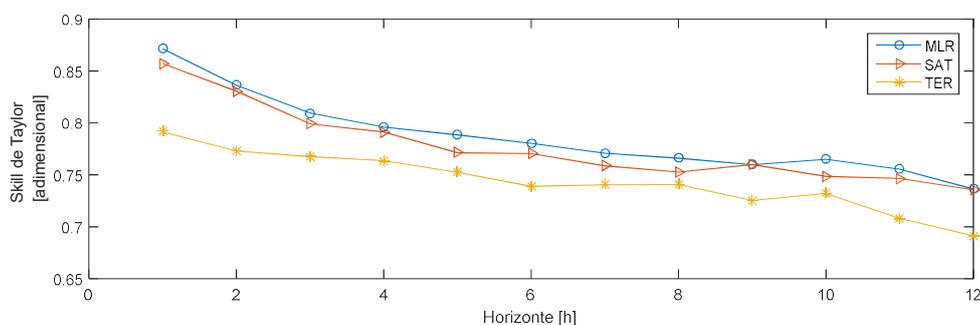


Fonte: O autor.

Analisando agora os estatísticos das três séries (TER, SAT e combinação MLR) em relação às observações, foram identificadas algumas características interessantes. A regressão MLR foi treinada com vistas a maximizar o *Skill*. Na Figura 38 observa-se que, em relação ao

*Skill*, a combinação MLR superou os modelos individuais em todos os horizontes. A respeito da correlação com as observações, a combinação MLR também foi superior aos modelos individuais. Nesse critério, o modelo SAT se destaca (Figura 39a) pelo fato de ter sido calibrado com dados horários centrados, em concordância com os dados de calibração da combinação (também centrados). Já em relação à razão entre os desvios-padrão das séries do modelo e da observação (*STD\_ratio*), o modelo TER alcançou melhor desempenho (Figura 39b). Embora a combinação MLR não tenha conseguido acompanhar o *STD\_ratio* do modelo TER, ela conseguiu superar nesse sentido o modelo SAT em 8 dos 12 horizontes. Dessa forma, conclui-se de forma preliminar que a combinação MLR conseguiu absorver características de correlação do modelo SAT e características de *STD\_ratio* do modelo TER.

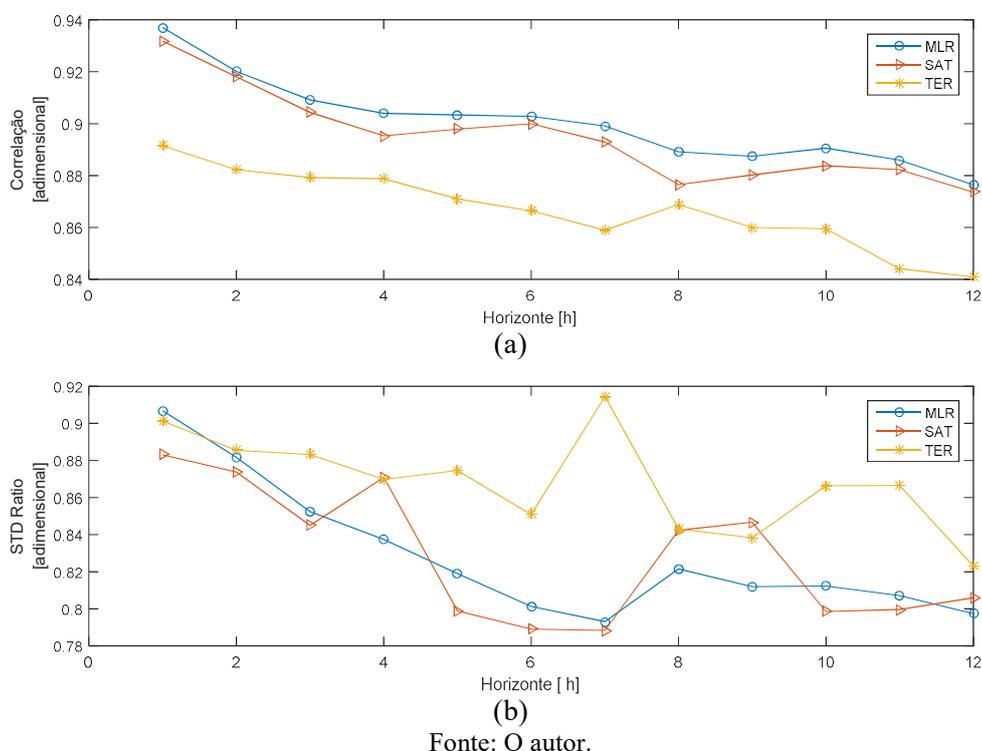
**Figura 38 – Skill de Taylor dos modelos SAT, TER e combinação MLR para Petrolina.**



Fonte: O autor.

As evidências estatísticas encontradas nesta seção sugerem uma complementariedade dos modelos SAT e TER na composição dessa combinação. As duas entradas mostraram-se importantes na composição da combinação aqui analisada, cada uma contribuindo com aspectos diferentes para melhorar os resultados finais.

**Figura 39 – (a) Correlação e (b) Razão entre os desvios-padrão da série do modelo em relação às observações.**



## 5.5 Comparação dos Resultados Obtidos com a Literatura

A combinação das previsões, comparada aos modelos individuais (Tabela 13), trouxe melhoria na acurácia das previsões, sendo considerada a melhor opção em quase todos os horizontes. A Tabela 14 traz um resumo de diversos trabalhos da literatura a respeito de previsão de GHI, bem como seus estatísticos. O objetivo é poder fazer uma comparação entre a combinação aqui proposta e os demais modelos. Tal tarefa não é trivial, visto que cada autor calcula seus estatísticos de forma não padronizada, o que dificulta as comparações. Alguns autores calculam os estatísticos como uma média geral, sem fazer separação por horizontes de previsão. Outros não calculam o Improvement, que é o estatístico mais indicado para se comparar trabalhos de diferentes sítios e, conseqüentemente, de diferentes dinâmicas das variáveis. Outros ainda calculam os estatísticos para conjuntos de dados divididos pela condição de céu (claro, parcialmente nublado e nublado). Para contornar essas dificuldades, os estatísticos da combinação proposta neste trabalho foram calculados de diversas formas: valores médios, valores separados por horizontes e para os estatísticos mais utilizados (Improvement, nRMSE, nRMSD e Correlação). Foram inseridas também informações

relevantes a respeito das condições e considerações realizadas como, por exemplo, a não realização de validação cruzada, o tamanho da base de dados, entre outros.

Vale salientar que o modelo de combinação de previsões proposto no presente trabalho foi comparado a trabalhos com modelos individuais de previsão, com exceção do último, que não é propriamente uma combinação estrita, mas sim um chaveamento de modelos de previsão (CRAWFORD e FRATANTONI, 2003).

Tendo em vista facilitar a comparação entre modelos, é dada uma sugestão de padronização de como e quais estatísticos deveriam ser calculados e apresentados. Além de gráficos, que facilitam a comparação visual, dever-se-ia apresentar tabelas conforme a Tabela 11, que permitiriam que outros autores pudessem comparar seus resultados.

**Tabela 14 – Comparação da Combinação de previsões com modelos de outros autores.**

Assunto	Tipo do modelo	Autor	Dados de entrada	Erro	Comparação com Combinação	Horizonte à frente	Local	OBS
Combinação de Previsões de GHI	ANN	Presente trabalho, 2019	Previsões de dados de radiômetro e de satélite	nRMSE%=[18,4; 27,9]; nRMSE% med= 24,58; FS= [22,2;75,7]; FS med= 59,4; nRMSD%=[18,0; 27,5]; nRMSD% med= 23,9; Corr=[88,1;95,4]; Corr med= 91,0;	-	12 passos à frente*	Petrolina, Brasil	Improvement calculado com respeito ao RMSE
Previsão de GHI	ANN	(Wang et al., 2011)	24 medidas previas e atual de GHI	MPE= 2,71% R <sup>2</sup> = 0,996	RMSE= 18,4% e R <sup>2</sup> =0,902	1h	Denver, Colorado.	Dados de entrada normalizados [0,1;0,9]; Interpolação linear para corrigir dados perdidos;
Previsão de GHI	ANN +NWP+SAT	(Aguiar et al., 2016)	Terrestre +NWP+ SAT	nRMSE=24%(1h) e 36%(6h)	nRMSE=18,4%(1h) e 23,5%(6h)	1h a 6h	Ilhas Canárias, Espanha	-
Previsão de Irradiância através de previsão de Índice de Cobertura de nuvens	ANN + ESSS (Iterativo)	(Dong et al., 2014)	Imagens de satélite	nRMSE calculado juntamente para todos os horizontes. Dividido apenas em meses. Variou entre 28% (melhor mês) e 43% (pior mês). 0,76< R <sup>2</sup> < 0,90	nRMSE med = 24,58% (ao longo de todo o ano). 0,77<R <sup>2</sup> <0,91	Das 07 às 19h. Processo iterativo	Cingapura	Modelo iterativo; previsão 1h a frente de índice de núvens. Em seguida é feita a conversão para GHI.
Review de Previsão de Radiação Solar		(Diagne et al., 2013)		rRMSE de 17% a 22% (céu claro). 33% a 44% (nublado)	nRMSE= de 18,4% a 27,9%	1h até 3 dias		Erros do trabalho de Kostylev e Pavlovski (2011)
Previsão de Radiação Solar	ANN MLP	(Mellit e Pavan, 2010)	GHI médio diário e Temperatura média diária	RMSE Geral (todos os horizontes) de 19% a 67% (céu claro) e de 54% a 85%(nublado)	nRMSEmed= 24,5% (todas as condições de céu)	1h a 24h	Trieste, Itália	O trabalho também calculou erros de Potência PV.
Previsão da próxima imagem e de GHI	Vetor de movimento; ANN; ACP	(Hammer et al., 2001)	Imagens de Satélite	nRMSE= 18% (1h) e 23%(2h)	nRMSE= 18% (1h) e 21%(2h)	30 min até 2h	Área no norte da Alemanha	Previsão de área. Um mês de dados (verão)
Modelo de previsão via Satélite	Vetor de movimento	(Perez et al., 2010)	Imagens de Satélite	nRMSE de 26% (1h) até 41%(6h) Improvement de 8% a 18%	nRMSE= de 18,4%(1h) a 23,5%(6h); Improvement de 48% a 75%(6h)	1h até 6h	EUA(sítio de Fort Peck)	Foi feito para outros sítios
Previsão de GHI	ANN, ARIMA e Funções de transferência	(Reikard, 2009)	Observações com alguns preenchimentos de valores modelados	Forecasting Skill de 29% a 8,5%	Improvement de 48% a 72%(4h)	1h até 4h	EUA, Clark, Nevada	Foi feito para outros sítios
Modelo Híbrido Previsão solar	Decomposição via Wavelet, EMD ou EEMD combinando com modelos de previsões AR e ANN	(Monjoly et al., 2017)	Dados medidos de GHI, os quais são transformados em Índice de céu claro	RMSE =7,86% Improvement = 72% (modelo WD-NN)	nRMSE= de 18,4%(1h); Improvement de 48%(1h)	1h	Ilha de Guadalupe	2 anos de dados

Previsão de irradiância e potência solar	SVM; clusterização	(Bae et al., 2017)	temperatura, umidade, hora do nascer e por do sol, heliofania relativa, cobertura de nuvem, tipo de nuvem, precipitação e velocidade do vento	correlação 0,96, 0,94 e 0,91; RMSE 49, 62 e 57W/m <sup>2</sup> ; MRE 3,1, 4,2 e 4,1% para céu claro, parcialmente nublado e nublado.	RMSE= 93W/m <sup>2</sup> (1h); Correlação 0,95;	1h	Coreia do Sul	26 meses de dados; usou dados de entrada do "futuro" medidos ao invés de previstos; não usou validação cruzada
Previsão de GHI horária um dia a frente	Redes LSTM (long short-term memory)	(Qing e Niu, 2018)	previsões NWP de temperatura, ponto de orvalho, umidade, visibilidade, velocidade do vento e resumo descritivo do tempo	Imp= 30,6% para LSTM e 15,6% quando usado ANN com retropropagação. (RMSE geral, não separados por horizontes)	Imp med= 60,4% (de 1 a 12 passos à frente)	Um dia à frente	Cabo Verde. (possui apenas 2 meses de chuva)	30 meses de dados. Não realizaram validação cruzada
Combinação (chaveamento) de previsão de GHI	modelos CARDS e autorregressivos acoplados	(Huang et al., 2013)	Dados medidos de GHI em solo	nRMSE de 16,5%	nRMSE= 17,7%	1h	Mildura, Austrália	1 ano de dados; Não é combinação estrita. É um chaveamento de Previsões.
Previsão de GHI e DNI	ANN, GA, Teste gamma para seleção de variáveis,	Marquez e Coimbra, 2011)	Previsões NWP de variáveis meteorológicas, cos(tetaz) e hora solar	nRMSE de 17,7% (mesmo dia) e 21% (dia seguinte) R <sup>2</sup> =0,94 (mesmo dia)	nRMSE%=[17,7; 27,5]; R <sup>2</sup> =0,83	6 dias	Califórnia, EUA	13 meses de dados;
Comparação de Previsões de GHI de modelos NWP	Modelos Numéricos NWP	Perez et al., 2013)	Previsões NWP de variáveis meteorológicas,	nRMSE=31% (24h à frente) Canadá	nRMSEmed= 24,5%	48h ou 7 dias, a depender do modelo	EUA, Canadá e Europa.	Realiza também a Combinação de 2 modelos (média).

Fonte: O autor.

De início, pode-se destacar que modelos que foram pensados exclusivamente para o curtíssimo prazo (1h à frente), como os modelos de BAE *et al.* (2017), de HUANG *et al.* (2013), de MONJOLY *et al.* (2017) e WANG *et al.* (2011), comparados à combinação foram mais bem sucedidos nas previsões para esse horizonte. Já em relação aos modelos pensados para horizontes de curto e médio prazo (REIKARD, 2009; MELLIT e PAVAN, 2010; PEREZ *et al.*, 2010; DIAGNE *et al.*, 2013; DONG *et al.*, 2014; AGUIAR *et al.*, 2016; QING e NIU, 2018), a combinação superou todos eles (com exceção do modelo de MARQUEZ e COIMBRA (2011)), conseguindo inclusive superar modelos com uma maior gama de variáveis de entrada (AGUIAR *et al.*, 2016) ou com maior complexidade (DONG *et al.*, 2014), ratificando o bom ajuste dos modelos propostos neste trabalho.

## 6 CONCLUSÃO

Foi desenvolvida uma metodologia de previsão de GHI horária para 24h à frente com *time step* horário, a qual consiste na combinação de previsões de dois outros modelos, um alimentado por imagens de satélite e outro alimentado por dados de radiômetro. Em cada etapa foram avaliados os resultados, que superaram a persistência em todos os horizontes, atingindo quase sempre *Improvements* com valores elevados. Cada um dos modelos foi comparado com resultados da literatura, mostrando que os modelos ajustados neste trabalho se sobressaíram na maioria dos horizontes.

A combinação de previsões possibilitou melhorar os resultados das previsões individuais, aumentando em média nos dez primeiros horizontes o *Improvement* e o *Skill* de Taylor (em relação ao modelo TER) em 1,78 e 5,26 pontos percentuais em média para os 10 primeiros horizontes. Nos dois últimos horizontes, a combinação não logrou êxito, apresentando desempenho inferior ao do modelo individual.

A retirada dos dados não úteis (noite, início da manhã e fim de tarde) das séries temporais realizada com base no valor do  $\cos(\theta_z)$  faz variar ao longo do ano o tamanho da janela útil do dia por conta da variação do próprio  $\cos(\theta_z)$  (Figura 30), dificultando assim o treinamento das redes neurais.

O uso das variáveis sazonais com características difusas (até então ainda não utilizadas para previsão de GHI horária) melhorou significativamente os resultados dos modelos, aumentando o *Improvement* em até 11 pontos percentuais (Tabela 8). Resultados com a mesma tendência já havia sido encontrado na previsão de Kt (SABINO *et al.*, 2017), porém com desempenho bem inferiores (Figura 28).

Os modelos de previsão de radiação foram testados em três sítios com características de nebulosidade distintas. Foi verificado que a qualidade das previsões depende da nebulosidade característica do local. Em locais com aparente regime de maior intensidade de nuvens, a previsão sofreu uma diminuição na sua acurácia.

Embora citado por algumas fontes da literatura que previsões de GHI através de imagens de satélite são eficientes para até 5 ou 6 horas à frente (PEREZ *et al.*, 2010), os resultados aqui obtidos divergem dessa afirmativa, mostrando que se pode obter boas previsões nos 12 passos à frente, lembrando que 12 passos com a série utilizada (excluindo-se os períodos das noites) podem representar previsões até um dia à frente.

Importante variável no processo de conversão de radiação para potência FV por modelos de desempenho de centrais, a temperatura foi prevista nesse trabalho utilizando um

modelo que utiliza apenas a série temporal de temperatura e cujo treinamento demanda pouquíssimo esforço computacional. Os erros cometidos por tal modelo são considerados baixos quando comparados aos modelos da literatura. Uma análise de sensibilidade realizada mostra que esses erros são aceitáveis para auxiliar as previsões de potência FV (SABINO *et al.*, 2018).

A comparação entre modelos dos diversos autores fica comprometida quando os estatísticos (ou o modo como eles são calculados) não são padronizados. É de fundamental importância que trabalhos sobre previsão disponibilizem tabelas (e não apenas gráficos) com estrutura semelhante à da Tabela 11, de forma a facilitar a comparação com outros trabalhos da literatura.

Uma importante contribuição deste trabalho, o modelo PIX trouxe uma nova alternativa de realização de previsões de GHI para locais que não dispõem de medidas de radiômetro. Os erros cometidos por esse modelo se aproximaram bastante dos modelos calibrados com dados de radiômetro.

A metodologia proposta neste trabalho serve como base para auxiliar previsões de potência de centrais FV, facilitando, do ponto de vista técnico, o aumento da inserção desse tipo de energia, contribuindo para uma matriz elétrica mais limpa e mais renovável.

## 7 PERSPECTIVAS

- Transformação dos modelos propostos em ferramentas de previsão operacional;
- Utilização de outros métodos (Seção 3.1) para estimar GHI a partir de imagens de satélite.
- Utilização de outras estratégias de previsão, como a DirRec, a MIMO ou a DIRMO para a previsão de GHI horária.
- Avaliar as implicações do uso de dados não centrados *versus* dados centrados como alvos durante a calibração do modelo de combinação de previsões.

## REFERÊNCIAS

ABDEL-AAL, R. Hourly temperature forecasting using abductive networks. **Engineering Applications of Artificial Intelligence**. v. 17, n. 5, p. 543-556, 2004.

ABSOLAR. Infográfico ABSOLAR. 2019. Disponível em: <http://absolar.org.br/infografico-absolar.html>. Acesso em 13/05/2019.

AGUIAR, L. M.; PEREIRA, B.; LAURET, P.; DÍAZ, F.; DAVID, M. Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting. **Renewable Energy**. v. 97, p. 599-610, 2016.

ARAI, N.; ALMEIDA, F. C. d. Estimativa da Radiação Solar Instantânea por Satélite para Céu Claro ou Encoberto. *In*: II Congresso Brasileiro de Meteorologia, 1982, Pelotas, RS. p. <http://www.cbmet.com/cbm-files/17-d25e5433e4041b9531cda6f41499e255.pdf>.

ASAMOAH-BOAHENG, M. Using SARIMA to Forecast Monthly Mean Surface Air Temperature in the Ashanti Region of Ghana. **International Journal of Statistics and Applications**. v. 4, n. 6, p. 292-298, 2014.

BAE, K. Y.; JANG, H. S.; SUNG, D. K. Hourly solar irradiance prediction based on support vector machine and its error analysis. **IEEE Transactions on Power Systems**. v. 32, n. 2, p. 935-945, 2017.

BARROS, H. F. d.; VILELA, O. d. C.; SABINO, E. R. C.; BRENNAND, L. J. P. Obtenção de Dados de Radiação Solar para Simulação de Central Fotovoltaica em Petrolina. 2018. Disponível em: <https://anaiscbens.emnuvens.com.br/cbens/article/view/711>.

BLANC, P.; REMUND, J.; VALLANCE, L. Short-term solar power forecasting based on satellite images. *In*: **Renewable Energy Forecasting**; ed.: Elsevier, v. p.179-198, 2017.

BOX, G. E. P. **Time series analysis : forecasting and control**. Rev. ed. ed. Oakland (Cal.): Holden-Day, 1976.

BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. ed.: Holden-Day, 1976.

BROWN, B. G.; KATZ, R. W.; MURPHY, A. H. Time series models to simulate and forecast wind speed and wind power. **Journal of Climate and Applied Meteorology**. v. 23, n. 8, p. 1184-1195, 1984.

CAMARGO, L. R.; ALTAMIRANO, M.; BELMONTE, S.; DORNER, W. Comparasiòn de Fuentes Satelitales, de Re-Análisis y Métodos Estadísticos para el Mapeo de la Radiaciòn Solar en el Valle de Lerma (Salta-Argentina). **ASADES**. v. 3., p. 11.31-11.42, 2015.

CEBALLOS, J. C.; BOTTINO, M. J. Estimativa de radiaçaõ solar por satélite: Desempenho do modelo operacional GL1. 2. **Anais do XIII Congresso Brasileiro de Meteorologia**. 2004.

CHAABENE, M.; BEN AMMAR, M. Neuro-fuzzy dynamic model with Kalman filter to forecast irradiance and temperature for solar energy systems. **Renewable Energy**. v. 33, n. 7, p. 1435-1443, 2008. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0960148107003035>.

CHAOUACHI, A.; KAMEL, R. M.; NAGASAKA, K. Neural network ensemble-based solar power generation short-term forecasting. **Journal of Advanced Computational Intelligence and Intelligent Informatics**. v. 14, n. 1, p. 69-75, 2010.

CHATFIELD, C. **The analysis of time series: an introduction**. ed.: CRC press, 2013.

CHEVALIER, R. F.; HOOGENBOOM, G.; MCCLENDON, R. W.; PAZ, J. A. Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks. **Neural Computing and Applications**. v. 20, n. 1, p. 151-159, 2011. DOI:10.1007/s00521-010-0363-y. Disponível em: <https://link.springer.com/article/10.1007%2Fs00521-010-0363-y>.

COSTA, A. C. A. **Matemathical/statistical and physical/meteorological models for short-term prediction of wind farms output**. 2005. 233 p. Departamento de Ingeniería Energética y Fluidomecánica, Universidad Politécnica de Madrid, 2005.

COSTA, A. C. A.; CRESPO, A.; NAVARRO, J.; LIZCANO, G.; MADSEN, H.; FEITOSA, E. A review on the young history of the wind power short-term prediction. **Renewable and Sustainable Energy Reviews**. v. 12, n. 6, p. 1725-1744, 2008.

COSTA, S. D. **Estratégias de Previsao Multipassosa Frente para Vazao Afluente em Bacias Hidrográficas de Diferentes Dinâmicas**. 2014. 122 p. (Doutorado). PPGEE, UFMG, 2014.

CRAWFORD, G. W.; FRATANTONI, M. C. Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices. **Real Estate Economics**. v. 31, n. 2, p. 223-243, 2003.

DAMBREVILLE, R.; BLANC, P.; CHANUSSOT, J.; BOLDO, D. Very short term forecasting of the Global Horizontal Irradiance using a spatio-temporal autoregressive model. **Renewable Energy**. v. 72, p. 291-300, 2014.

DE GIORGI, M. G.; MALVONI, M.; CONGEDO, P. M. Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine. **Energy**. v. 107, p. 360-373, 2016. DOI:<http://dx.doi.org/10.1016/j.energy.2016.04.020>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0360544216304261>.

DIAGNE, M.; DAVID, M.; LAURET, P.; BOLAND, J.; SCHMUTZ, N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. **Renewable and Sustainable Energy Reviews**. v. 27, p. 65-76, 2013.

DONG, Z.; YANG, D.; REINDL, T.; WALSH, W. M. Satellite image analysis and a hybrid esss/ann model to forecast solar irradiance in the tropics. **Energy Conversion and Management**. v. 79, p. 66-73, 2014.

DSA/INPE. Divisão de Satélites e Sistemas Ambientais. 2015. Disponível em: <http://satelite.cptec.inpe.br/pedidoweb/pedido.formulario.logic#>. Acesso em 28/03/2019.

DUFFIE, J. A.; BECKMAN, W. A. **Solar engineering of thermal processes**. ed.: Wiley New York, 2013.

FU, C. L.; CHENG, H. Y. Predicting solar irradiance with all-sky image features via regression. **Solar Energy**. v. 97, p. 537-550, 2013.

GALLEGO, C.; COSTA, A.; CUERVA, Á.; LANDBERG, L.; GREAVES, B.; COLLINS, J. A wavelet-based approach for large wind power ramp characterisation. **Wind Energy**. v. 16, n. 2, p. 257-278, 2013.

GAUTIER, C.; DIAK, G.; MASSE, S. A simple physical model to estimate incident solar radiation at the surface from GOES satellite data. **Journal of Applied Meteorology**. v. 19, n. 8, p. 1005-1012, 1980.

GENSLER, A.; HENZE, J.; SICK, B.; RAABE, N. Deep Learning for solar power forecasting — An approach using AutoEncoder and LSTM Neural Networks. *In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016, 2016.* p.002858-002865.

GIEBEL, G.; BROWNSWORD, R.; KARINIOTAKIS, G.; DENHARD, M.; DRAXL, C. The state-of-the-art in short-term prediction of wind power: A literature overview. 2011. Disponível em: <https://www.osti.gov/etdeweb/biblio/1011554>.

GREEN, M. A. Solar cells: operating principles, technology, and system applications. 1982.

GUTIÉRREZ, J. M.; COFIÑO, A.; CANO, R.; RODRÍGUEZ, M. A. Clustering methods for statistical downscaling in short-range weather forecasts. **Monthly Weather Review.** v. 132, n. 9, p. 2169-2183, 2004.

HAMMER, A.; HEINEMANN, D.; HOYER, C.; LORENZ, E. Satellite based short-term forecasting of solar irradiance—comparison of methods and error analysis. *In: The 2001 EUMETSAT meteorological satellite data user's conference, 2001, 2001.* p.677-684.

HAY, J. E.; HANSON, K. J. Satellite-based methodology for determining solar irradiance at the ocean surface during gate. **Bulletin of the American Meteorological Society.** v. 59, n. 11, p. 1549-1549, 1978.

HAYATI, M.; MOHEBI, Z. Application of artificial neural networks for temperature forecasting. **World Academy of Science, Engineering and Technology.** v. 28, n. 2, p. 275-279, 2007.

HAYKIN, S. S. **Neural networks : a comprehensive foundation.** ed. New York: Macmillan, 1994.

HOFF, T. E.; PEREZ, R. Quantifying PV power output variability. **Solar Energy.** v. 84, n. 10, p. 1782-1793, 2010.

HUANG, J.; KOROLKIEWICZ, M.; AGRAWAL, M.; BOLAND, J. Forecasting solar radiation on an hourly time scale using a Coupled AutoRegressive and Dynamical System (CARDS) model. **Solar Energy.** v. 87, p. 136-149, 2013.

INMAN, R. H.; PEDRO, H. T. C.; COIMBRA, C. F. M. Solar forecasting methods for renewable energy integration. **Progress in Energy and Combustion Science.** v. 39, n. 6, p. 535-576, 2013.

KALNAY, E.; KANAMITSU, M.; KISTLER, R.; COLLINS, W.; DEAVEN, D.; GANDIN, L.; IREDELL, M.; SAHA, S.; WHITE, G.; WOOLLEN, J. The NCEP/NCAR 40-year reanalysis project. **Bulletin of the American meteorological Society**. v. 77, n. 3, p. 437-471, 1996.

KHOTANZAD, A.; ELRAGAL, H.; LU, T. Combination of artificial neural-network forecasters for prediction of natural gas consumption. **IEEE Transactions on Neural Networks**. v. 11, n. 2, p. 464-473, 2000. DOI:10.1109/72.839015

KLEISSL, J. **Solar energy forecasting and resource assessment**. ed.: Academic Press, 2013.

KLINE, D. M.; ZHANG, G. Methods for multi-step time series forecasting with neural networks. **Neural networks in business forecasting**. p. 226-250, 2004.

LARSON, D. P.; COIMBRA, C. F. M. Direct Power Output Forecasts From Remote Sensing Image Processing. **Journal of Solar Energy Engineering**. v. 140, n. 2, p. 021011-021011-8, 2018. DOI:10.1115/1.4038983. Disponível em: <http://dx.doi.org/10.1115/1.4038983>.

LEWIS, C. D. **International and Business Forecasting Methods**. ed. London: Butterworths, 1982.

LIMA, F. J. L. d. **Previsão de Irradiação Solar no Nordeste do Brasil Empregando o Modelo WRF Ajustado por Redes Neurais Artificiais (RNAs)**. 2015. 251 p. (Doutorado). INPE, 2015.

LIU, J.; ZHANG, J. G.; YANG, Z.; LEMMON, J. P.; IMHOFF, C.; GRAFF, G. L.; LI, L.; HU, J.; WANG, C.; XIAO, J. Materials science and materials chemistry for large scale electrochemical energy storage: from transportation to electrical grid. **Advanced Functional Materials**. v. 23, n. 8, p. 929-946, 2013.

LORENZ, E.; REMUND, J.; MÜLLER, S. C.; TRAUNMÜLLER, W.; STEINMAURER, G.; POZO, D.; RUIZ-ARIAS, J. A.; FANEGO, V. L.; RAMIREZ, L.; ROMEO, M. G. Benchmarking of different approaches to forecast solar irradiance. *In: 24th European photovoltaic solar energy conference, Hamburg, Germany, 2009, 2009*. p.25.

MADSEN, H.; PINSON, P.; KARINIOTAKIS, G.; NIELSEN, H. A.; NIELSEN, T. S. Standardizing the performance evaluation of short-term wind power prediction models. **Wind engineering**. v. 29, n. 6, p. 475-489, 2005.

MARQUEZ, R.; COIMBRA, C. F. M. Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database. **Solar Energy**. v. 85, n. 5, p. 746-756, 2011. DOI:<https://doi.org/10.1016/j.solener.2011.01.007>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0038092X11000193>.

MARTÍN, L.; ZARZALEJO, L. F.; POLO, J.; NAVARRO, A.; MARCHANTE, R.; CONY, M. Prediction of global solar irradiance based on time series analysis: application to solar thermal power plants energy production planning. **Solar Energy**. v. 84, n. 10, p. 1772-1781, 2010.

MARTINS, F. **Influência do processo de determinação da cobertura de nuvens e dos aerossóis de queimada no modelo físico de radiação BRASIL-SR**. 2001. p. Tese de doutoramento, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 330pp, 2001.

MARTINS, F. R.; PEREIRA, E. B. Estudo comparativo da confiabilidade de estimativas de irradiação solar para o sudeste brasileiro obtidas a partir de dados de satélite e por interpolação/extrapolação de dados de superfície. **Revista Brasileira de Geofísica**. v. 29, n. 2, p. 265-276, 2011.

MARTINS, F. R.; PEREIRA, E. B.; ECHER, M. P. d. S. Levantamento dos recursos de energia solar no Brasil com o emprego de satélite geostacionário—o Projeto Swera. **Revista Brasileira de Ensino de Física**. v. 26, n. 2, p. 145-159, 2004.

MCCANDLESS, T. C.; HAUPT, S. E.; YOUNG, G. S. SHORT TERM SOLAR RADIATION FORECASTS USING WEATHER REGIME-DEPENDENT J3. 5 ARTIFICIAL INTELLIGENCE TECHNIQUES. *In*: 94th American Meteorological Society Annual Meeting, 2014. p.

MCCUTCHAN, M. H. Diagnosing and predicting surface temperature in mountainous terrain. **Monthly Weather Review**. v. 104, n. 8, p. 1044-1051, 1976.

MELLIT, A.; PAVAN, A. M. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. **Solar Energy**. v. 84, n. 5, p. 807-821, 2010.

MERAL, M. E.; DINÇER, F. A review of the factors affecting operation and efficiency of photovoltaic based electricity generation systems. **Renewable and Sustainable Energy Reviews**. v. 15, n. 5, p. 2176-2184, 2011.

MICHAELSEN, J. Cross-validation in statistical climate forecast models. **Journal of Climate and Applied Meteorology**. v. 26, n. 11, p. 1589-1600, 1987.

MONAHAN, A. H.; FYFE, J. C.; AMBAUM, M. H. P.; STEPHENSON, D. B.; NORTH, G. R. Empirical orthogonal functions: The medium is the message. **Journal of Climate**. v. 22, n. 24, p. 6501-6514, 2009.

MONJOLY, S.; ANDRÉ, M.; CALIF, R.; SOUBDHAN, T. Hourly forecasting of global solar radiation based on multiscale decomposition methods: A hybrid approach. **Energy**. v. 119, p. 288-298, 2017. DOI:<https://doi.org/10.1016/j.energy.2016.11.061>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0360544216316668>.

MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to time series analysis and forecasting**. ed.: John Wiley & Sons, 2015.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**. ed.: John Wiley & Sons, 2012.

MOURA, G. B. A. **Estimativa de radiação solar na Paraíba mediante satélite Meteosat: testes de dois modelos usuais**. 1993. 88 p. (Mestrado). Universidade Federal da Paraíba, 1993.

NETTO, M. A. C. A Previsão com a Metodologia de Box-Jenkins. 2015. Disponível em: <http://www.ie.ufjf.br/download/index.php>. Acesso em 14/11/2016.

NFAOUI, H.; BURET, J.; SAYIGH, A. Stochastic simulation of hourly average wind speed sequences in Tangiers (Morocco). **Solar Energy**. v. 56, n. 3, p. 301-314, 1996.

ORTEGA-VAZQUEZ, M. A.; KIRSCHEN, D. S. Estimating the spinning reserve requirements in systems with significant wind power generation penetration. **Power Systems, IEEE Transactions on**. v. 24, n. 1, p. 114-124, 2009.

PELLAND, S.; REMUND, J.; KLEISSL, J.; OZEKI, T.; DE BRABANDERE, K. Photovoltaic and solar forecasting: state of the art. **IEA PVPS, Task**. v. 14, p. 1-36, 2013.

PEREIRA, E. B.; MARTINS, F. R.; GONÇALVES, A. R.; COSTA, R. S.; LIMA, F. J. L.; RÜTHER, R.; ABREU, S. L.; TIEPOLO, G. M.; PEREIRA, S. V.; SOUZA, J. G. **Atlas Brasileiro de Energia Solar, INPE, São José dos Campos, Brazil**. 2ª ed., 2017.

PEREZ, R.; KIVALOV, S.; SCHLEMMER, J.; HEMKER, K.; RENNÉ, D.; HOFF, T. E. Validation of short and medium term operational solar radiation forecasts in the US. **Solar Energy**. v. 84, n. 12, p. 2161-2172, 2010.

PEREZ, R.; LORENZ, E.; PELLAND, S.; BEAUHARNOIS, M.; VAN KNOWE, G.; HEMKER, K.; HEINEMANN, D.; REMUND, J.; MÜLLER, S. C.; TRAUNMÜLLER, W.; STEINMAUER, G.; POZO, D.; RUIZ-ARIAS, J. A.; LARA-FANEGO, V.; RAMIREZ-SANTIGOSA, L.; GASTON-ROMERO, M.; POMARES, L. M. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. **Solar Energy**. v. 94, p. 305-326, 2013. DOI:<https://doi.org/10.1016/j.solener.2013.05.005>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0038092X13001886>.

PETRIBU, L.; SABINO, E.; BARROS, H.; COSTA, A.; BARBOSA, E.; VILEA, O. C. PROCEDIMENTO OBJETIVO PARA A GARANTIA DE QUALIDADE DE DADOS DE RADIAÇÃO SOLAR. **Avances en Energías Renovables y Medio Ambiente**. v. 21, p. 11.67-11.78, 2017.

QING, X.; NIU, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. **Energy**. v. 148, p. 461-468, 2018.

QUADROS, F. S. d. **Sistema de Divulgação de Dados Meteorológicos**. 2005. 105 p. (Doutorado). Centro de Ciências Tecnológicas da Terra e do Mar, Universidade do Vale do Itajaí, 2005.

RAPHAEL, C.; HAY, J. E. An assessment of models which use satellite data to estimate solar irradiance at the earth's surface. **Journal of Climate and Applied Meteorology**. v. 23, n. 5, p. 832-844, 1984.

REIKARD, G. Predicting solar radiation at high resolutions: A comparison of time series forecasts. **Solar Energy**. v. 83, n. 3, p. 342-349, 2009.

RENDON, J.; MENEZES, L. M. d. Structural Combination of Neural Network Models. *In*: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), 2016, 2016. p.406-413.

SABINO, E. R. C.; DE PETRIBÚ BRENNAND, L. J.; DE BARROS, H. F.; DE CASTRO VILELA, O.; DA COSTA, A. C. A. Previsão em Curto Prazo de Índice de Claridade utilizando Redes Neurais e Variáveis Sazonais. **Avances en Energías Renovables y Medio Ambiente**. v. 21, p. 11.39-11.48, 2017.

SABINO, E. R. C.; DE PETRIBÚ BRENNAND, L. J.; DE BARROS, H. F.; DE CASTRO VILELA, O.; DA COSTA, A. C. A. Previsão de Temperatura Ambiente Horária Voltada para Auxiliar Previsões de Geração de Potência Fotovoltaica. *In: VII Congresso Brasileiro de Energia Solar*, 2018, Gramado-RS, Brasil. p.  
<https://anaiscbens.emnuvens.com.br/cbens/article/view/659>.

SAFI, S.; ZEROUAL, A.; HASSANI, M. Prediction of global daily solar radiation using higher order statistics. **Renewable Energy**. v. 27, n. 4, p. 647-666, 2002.  
DOI:[https://doi.org/10.1016/S0960-1481\(01\)00153-7](https://doi.org/10.1016/S0960-1481(01)00153-7). Disponível em:  
<http://www.sciencedirect.com/science/article/pii/S0960148101001537>.

SANDSTROM, J. D. A method for predicting solar cell current-voltage curve characteristics as a function of incident solar intensity and cell temperature. **IEEE VOL. II**. v. 199, 1967.

ŞENKAL, O.; KULELI, T. Estimation of solar radiation over Turkey using artificial neural network and satellite data. **Applied Energy**. v. 86, n. 7, p. 1222-1228, 2009.  
DOI:<https://doi.org/10.1016/j.apenergy.2008.06.003>. Disponível em:  
<http://www.sciencedirect.com/science/article/pii/S0306261908001578>.

SFETSOS, A.; COONICK, A. Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques. **Solar Energy**. v. 68, n. 2, p. 169-178, 2000.

SIMMONS, A. J.; GIBSON, J. K. The ERA-40 Project Plan, ERA-40 Project Report Series No. 1, European Centre for Medium-Range Weather Forecasts. **Reading**. 2000.

SKAMAROCK, W.; KLEMP, J. B.; DUDHIA, J.; GILL, D.; BARKER, D.; DUDA, M.; HUANG, X.; WANG, W.; POWERS, J. **A description of the advanced research WRF version 3**. NCAR Technical Note, NCAR/TN\ u2013475? STR, 123 pp 2008.

SMITH, B. A.; MCCLENDON, R. W.; HOOGENBOOM, G. Improving air temperature prediction with artificial neural networks. **International Journal of Computational Intelligence**. v. 3, n. 3, p. 179-186, 2006.

SMITH, L. I. A tutorial on principal components analysis. **Cornell University, USA**. v. 51, p. 52, 2002.

SORJAMAA, A.; LENDASSE, A. Time series prediction using dirrec strategy. *In: European symposium on artificial neural networks*, 2006, Bruges, Belgium. p. 143–148.

SOUSA, J. A Importância da Previsão da Geração Elétrica de Base Eólica. *In: Brazil Windpower 2013 Conference and Exhibition, 2013, Brazil.* p.  
<http://www.brazilwindpower.org/archives/anuario/Sala%201%20-%2014h00%20-%20A%20Importancia%20da%20Previsao%20da%20Geracao%20Eletrica.pdf>. Acesso em: 18/01/2015.

TAIEB, S. B.; BONTEMPI, G.; ATIYA, A. F.; SORJAMAA, A. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. **Expert systems with applications.** v. 39, n. 8, p. 7067-7083, 2012.

TARPLEY, J. D. Estimating incident solar radiation at the surface from geostationary satellite data. **Journal of Applied Meteorology.** v. 18, n. 9, p. 1172-1181, 1979.

TASADDUQ, I.; REHMAN, S.; BUBSHAIT, K. Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia. **Renewable Energy.** v. 25, n. 4, p. 545-554, 2002.

TAYLOR, K. E. Summarizing multiple aspects of model performance in a single diagram. **Journal of Geophysical Research: Atmospheres (1984–2012).** v. 106, n. D7, p. 7183-7192, 2001.

TOMALSQUIM, M. **Integração das Fontes Renováveis Intermitentes na América Latina.** ed.: Corporación Andina de Fomento, 2017.

UE, Y.; HARA, R.; KITA, H.; SAITO, Y.; TAKITANI, K.; SAITO, M. Analysis of Spatial Smoothing Effect of Short-term Fluctuation of Global Solar Radiation based on the Data Including Quantization Error. **IEEEJ Transactions on Power and Energy.** v. 131, p. 413-420, 2011.

VARELLA, C. A. A. Análise de componentes principais. **Universidade Federal Rural do Rio de Janeiro, Disponível em: <http://www.ufrj.br/institutos/it/deng/varella/Downloads>.** Acessado em: 15/10/2015.

VOYANT, C.; DARRAS, C.; MUSELLI, M.; PAOLI, C.; NIVET, M.-L.; POGGI, P. Bayesian rules and stochastic models for high accuracy prediction of solar radiation. **Applied Energy.** v. 114, p. 218-226, 2014.

VOYANT, C.; NOTTON, G.; KALOGIROU, S.; NIVET, M.-L.; PAOLI, C.; MOTTE, F.; FOUILLOY, A. Machine learning methods for solar radiation forecasting: A review. **Renewable Energy.** v. 105, p. 569-582, 2017.

WANG, Z.; WANG, F.; SU, S. Solar irradiance short-term prediction model based on BP neural network. **Energy Procedia**. v. 12, p. 488-494, 2011.

WILBY, R. L.; WIGLEY, T. M. L. Downscaling general circulation model output: a review of methods and limitations. **Progress in Physical Geography**. v. 21, n. 4, p. 530-548, 1997.

WILKS, D. S. Statistical methods in the atmospheric sciences. 2011.

YADAV, A. K.; CHANDEL, S. S. Solar radiation prediction using Artificial Neural Network techniques: A review. **Renewable and Sustainable Energy Reviews**. v. 33, p. 772-781, 2014. DOI:<https://doi.org/10.1016/j.rser.2013.08.055>. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1364032113005959>.

YAGLI, G. M.; YANG, D.; SRINIVASAN, D. Using Combinational Methods for Forecast Improvement in PV Power Plants. *In*: 2018 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia), 2018, IEEE, 2018. p.540-545.

YULE, G. U. On the time-correlation problem, with especial reference to the variate-difference correlation method. **Journal of the Royal Statistical Society**. p. 497-537, 1921.

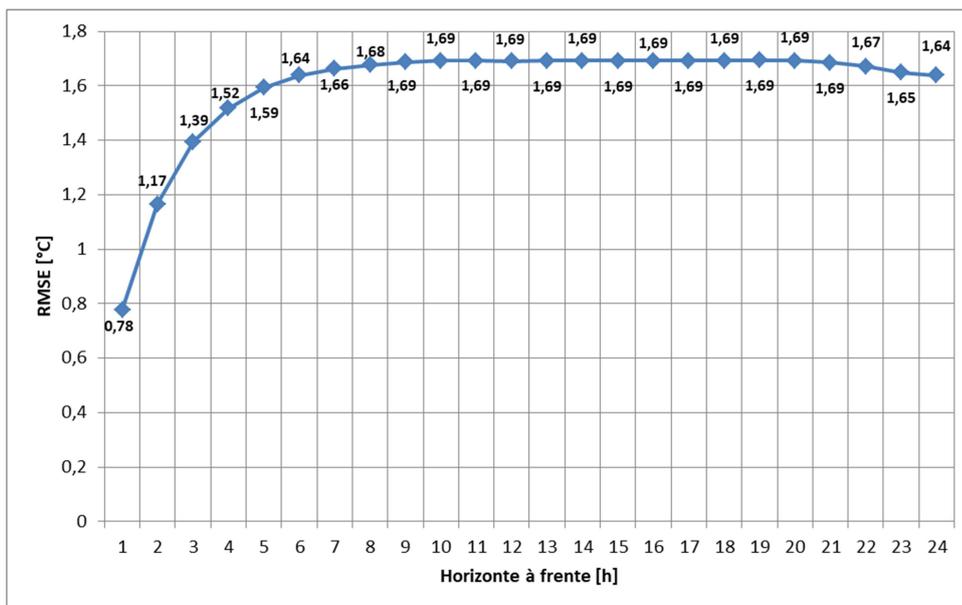
YULE, G. U. VII. On a method of investigating periodicities disturbed series, with special reference to Wolfer's sunspot numbers. **Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character**. v. 226, n. 636-646, p. 267-298, 1927.

ZARZALEJO, L. F.; RAMIREZ, L.; POLO, J. Artificial intelligence techniques applied to hourly global irradiance estimation from satellite-derived cloud index. **Energy**. v. 30, n. 9, p. 1685-1697, 2005.

## APÊNDICE A – RESULTADOS PREVISÃO DE TEMPERATURA

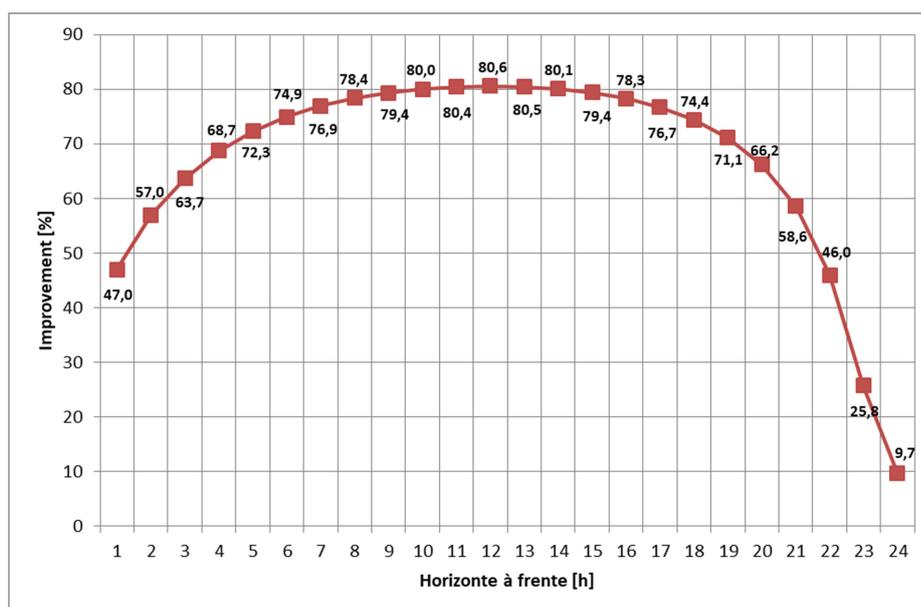
São apresentados os resultados dos modelos autorregressivos de previsão de temperatura horária para Pesqueira e Brasília.

**Figura 40 – Erros médios (RMSE) cometidos pelos 24 modelos de previsão de temperatura treinados com dados de Pesqueira.**



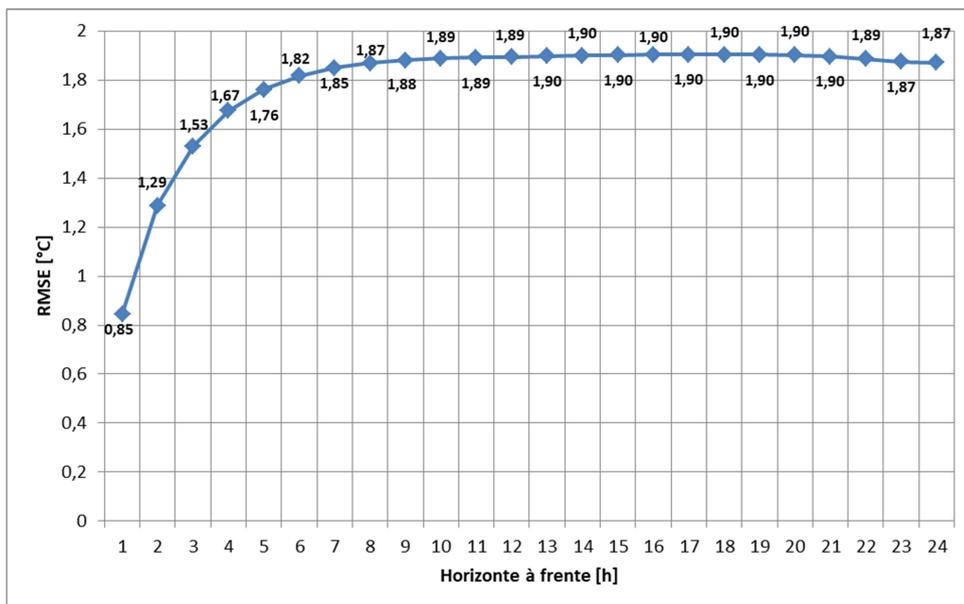
Fonte: O autor.

**Figura 41 – Improvement em relação à Persistência dos 24 modelos de previsão de temperatura treinados com dados de Pesqueira.**



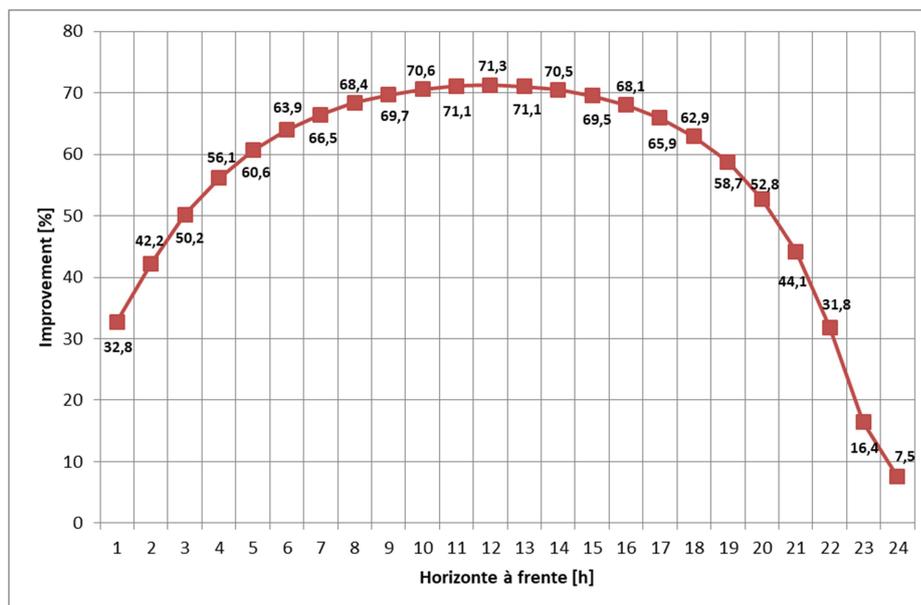
Fonte: O autor.

**Figura 42 – Erros médios (RMSE) cometidos pelos 24 modelos de previsão de temperatura treinados com dados de Brasília.**



Fonte: O autor.

**Figura 43 – Improvement em relação à Persistência dos 24 modelos de previsão de temperatura treinados com dados de Brasília.**

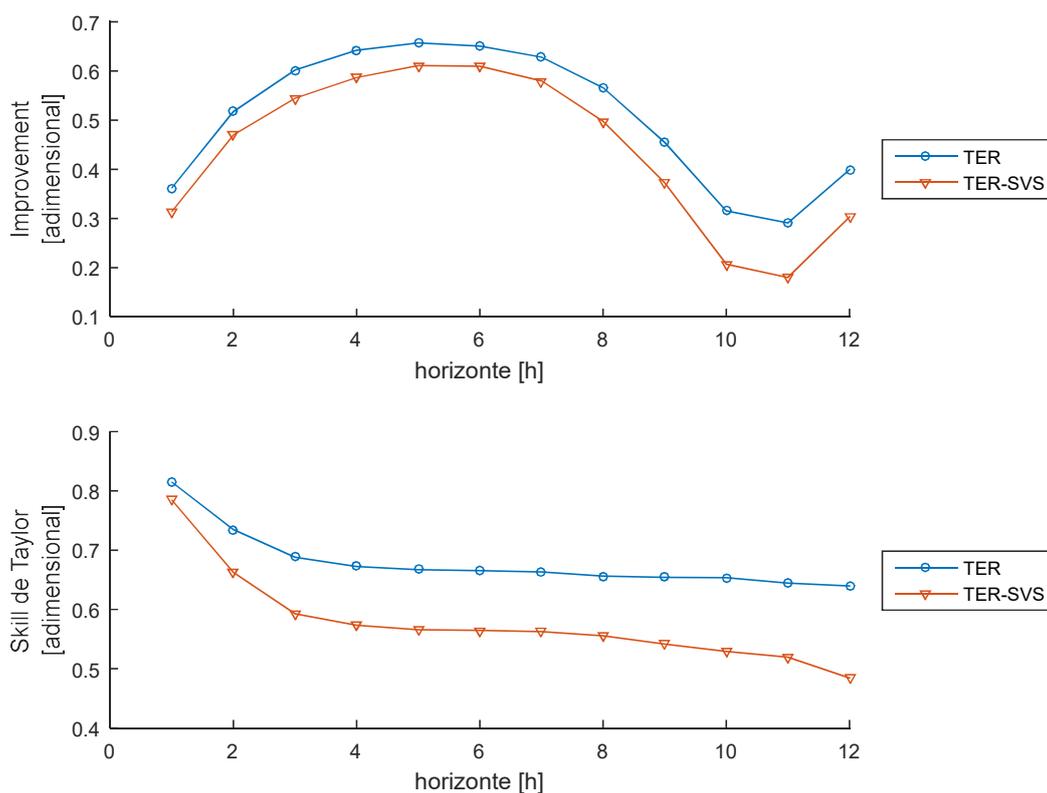


Fonte: O autor.

## APÊNDICE B – INFLUÊNCIA DAS VARIÁVEIS SAZONAIS NAS PREVISÕES

Resultados dos modelos TER (com variáveis sazonais) e TER-SVS (sem variáveis sazonais) em Brasília e Petrolina.

**Figura 44 – Comparação entre os modelos TER (com variáveis sazonais) e os modelos TER-SVS (sem variáveis sazonais) para Brasília.**



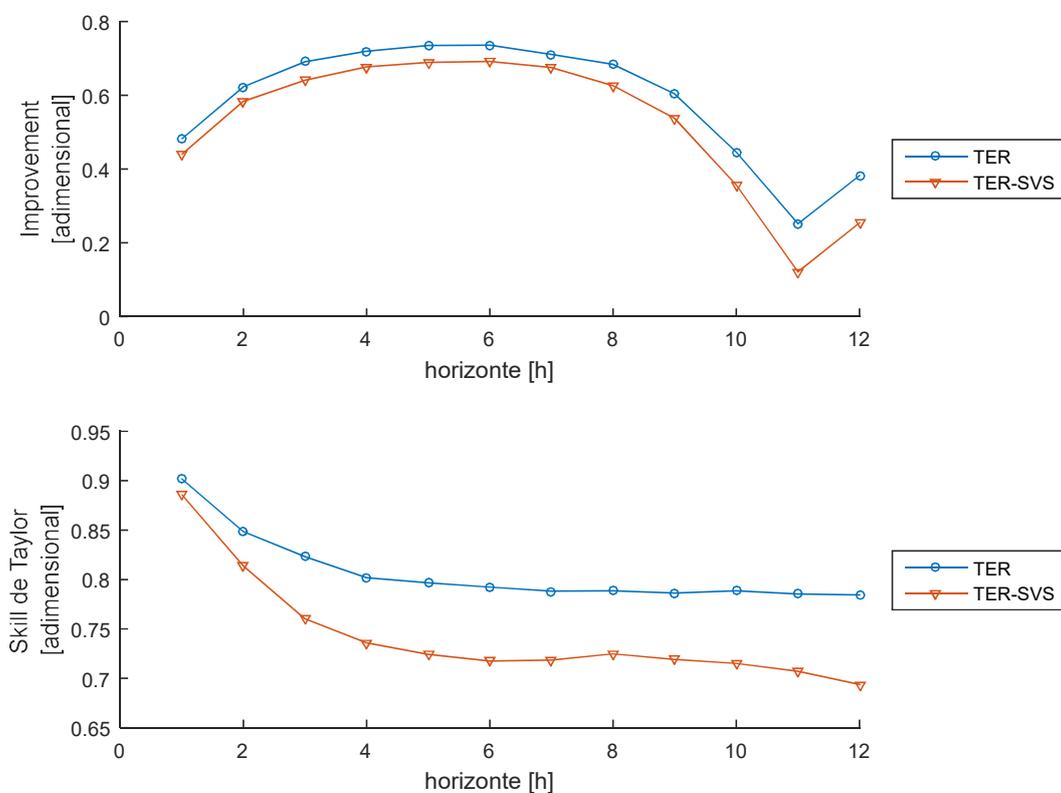
Fonte: O autor.

**Tabela 15 – Improvement dos modelos TER (com variáveis sazonais) e modelos TER-SVS (sem variáveis sazonais) para Brasília.**

Modelo	Horizonte [h]											
	1	2	3	4	5	6	7	8	9	10	11	12
TER	36,14%	51,72%	60,19%	64,21%	65,72%	65,07%	62,83%	56,61%	45,47%	31,55%	29,10%	39,92%
TER-SVS	31,34%	46,97%	54,45%	58,66%	61,08%	61,00%	58,00%	49,77%	37,27%	20,70%	18,00%	30,31%

Fonte: O autor.

**Figura 45 – Comparação entre os modelos TER (com variáveis sazonais) e os modelos TER-SVS (sem variáveis sazonais) para Petrolina.**



Fonte: O autor.

**Tabela 16 – Improvement dos modelos TER (com variáveis sazonais) e modelos TER-SVS (sem variáveis sazonais) para Petrolina.**

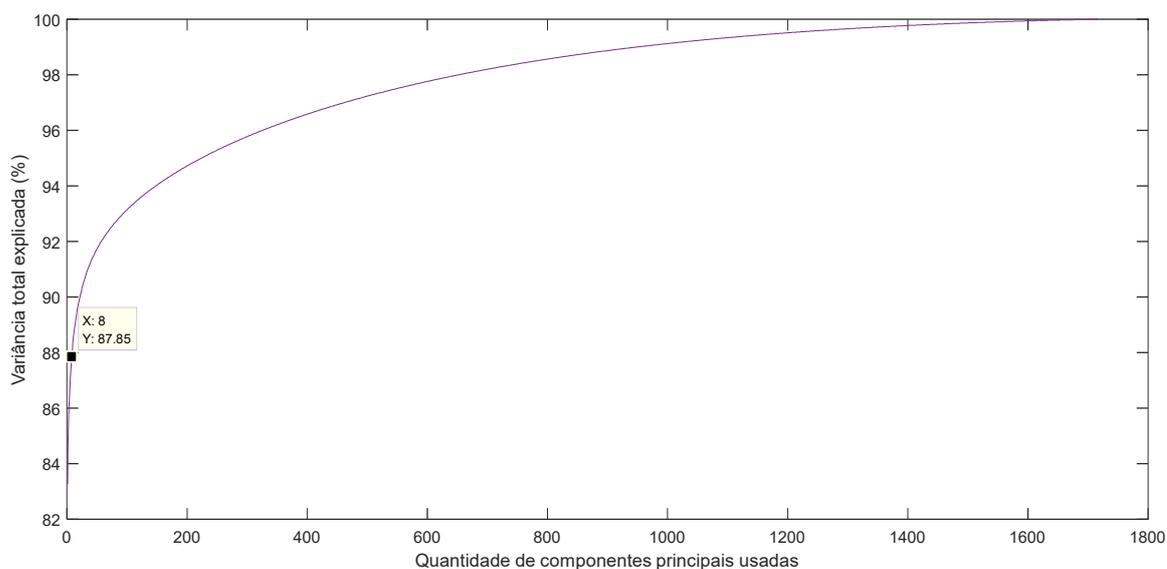
Modelo	Horizonte [h]											
	1	2	3	4	5	6	7	8	9	10	11	12
TER	48,14%	62,29%	69,11%	71,97%	73,53%	73,58%	71,08%	68,42%	60,41%	44,60%	25,16%	38,21%
TER-SVS	43,95%	58,38%	64,07%	67,72%	68,91%	69,22%	67,54%	62,61%	53,73%	35,65%	12,13%	25,40%

Fonte: O autor.

## APÊNDICE C – RESULTADOS DAS ANÁLISES DE COMPONENTES PRINCIPAIS

São apresentados os resultados das análises de componentes principais para Brasília e Pesqueira.

**Figura 46 – Percentual acumulado de variância explicada (do conjunto inicial de variáveis) em função do número de componentes principais utilizado para a série de 10% de pixels usados no horizonte 1 passo a frente em Brasília (2013 e 2014).**



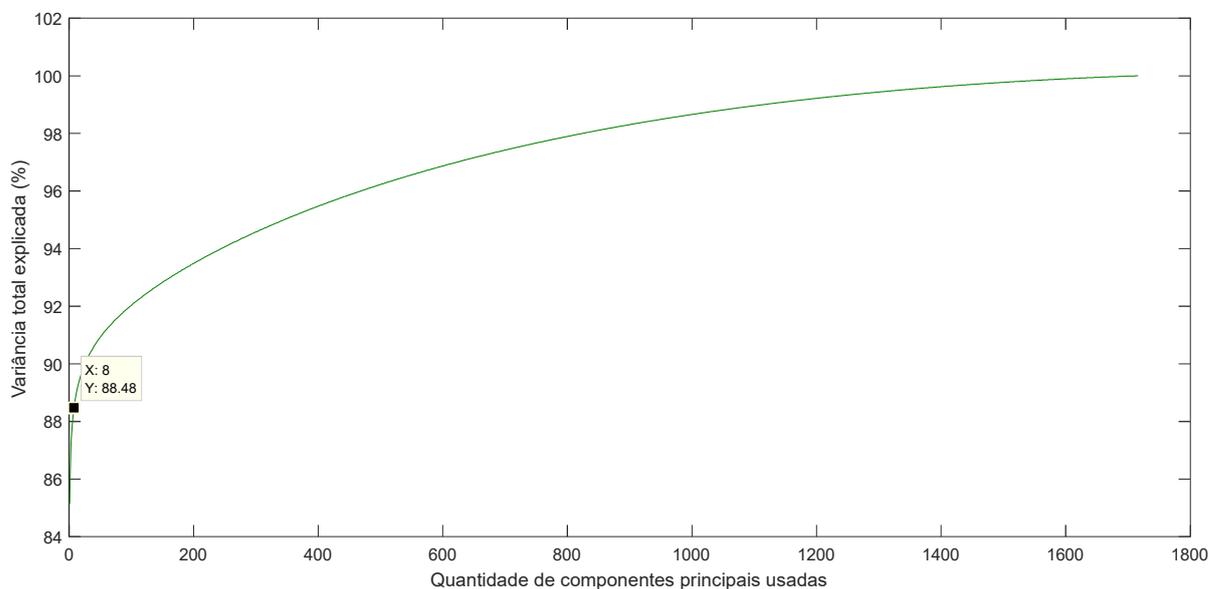
Fonte: O autor.

**Tabela 17 – Percentual acumulado de variância explicada por oito componentes principais para os diversos horizontes da série de Brasília de 2013 e 2014.**

	Horizonte											
	1	2	3	4	5	6	7	8	9	10	11	12
Variância explicada (%)	87,85	87,16	87,27	87,58	87,55	87,15	85,88	88,06	88,34	88,60	88,84	88,02

Fonte: O autor.

**Figura 47 – Percentual acumulado de variância explicada (do conjunto inicial de variáveis) em função do número de componentes principais utilizado para a série de 10% de pixels usados no horizonte 1 passo a frente em Pesqueira (2013 e 2014).**



Fonte: O autor.

**Tabela 18 – Percentual acumulado de variância explicada por oito componentes principais para os diversos horizontes da série de Pesqueira de 2013 e 2014.**

	Horizonte											
	1	2	3	4	5	6	7	8	9	10	11	12
Variância explicada (%)	88,48	88,25	87,53	88,32	89,23	90,00	89,26	87,19	88,03	89,00	89,92	89,39

Fonte: O autor.