

UNIVERSIDADE FEDERAL FLUMINENSE
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
TELECOMUNICAÇÕES

RODRIGO PIZETTI MARTINS

PREVISÃO PROBABILÍSTICA DE ENERGIA SOLAR UTILIZANDO REDES
NEURAS COM INFERÊNCIA BAYESIANA E MISTURA DE GAUSSIANAS

Niterói – RJ

2018

RODRIGO PIZETTI MARTINS

PREVISÃO PROBABILÍSTICA DE ENERGIA SOLAR UTILIZANDO REDES
NEURAS COM INFERÊNCIA BAYESIANA E MISTURA DE GAUSSIANAS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense como requisito parcial para a Obtenção do Grau de Mestre em Engenharia Elétrica e de Telecomunicações.

Área de concentração: Sistemas de Energia Elétrica.

Orientador: Prof. VITOR HUGO FERREIRA
Coorientador: Prof. THALES TERROLA E LOPES

Niterói – RJ
(março/2018)

Ficha catalográfica automática - SDC/BEE

M379p Martins, Rodrigo Pizetti
Previsão Probabilística de Energia Solar Utilizando Redes
Neurais com Inferência Bayesiana e Mistura de Gaussianas /
Rodrigo Pizetti Martins ; Vitor Hugo Ferreira, orientador ;
Thales Terrola e Lopes, coorientador. Niterói, 2018.
80 f. : il.

Dissertação (mestrado)-Universidade Federal Fluminense,
Niterói, 2018.

1. Rede Neural Artificial. 2. Inferência Bayesiana. 3.
Mistura de Gaussianas. 4. Energia Solar Fotovoltaica. 5.
Produção intelectual. I. Título II. Ferreira, Vitor Hugo,
orientador. III. Lopes, Thales Terrola e, coorientador. IV.
Universidade Federal Fluminense. Escola de Engenharia.

CDD -

RODRIGO PIZETTI MARTINS

PREVISÃO PROBABILÍSTICA DE ENERGIA SOLAR UTILIZANDO REDES
NEURAIS COM INFERÊNCIA BAYESIANA E MISTURA DE GAUSSIANAS

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense como requisito parcial para a Obtenção do Grau de Mestre em Engenharia Elétrica e de Telecomunicações.

Área de concentração: Sistemas de Energia Elétrica.

Aprovada em 06 de março de 2018.

Prof. Dr. Vitor Hugo Ferreira - Orientador
Universidade Federal Fluminense

Prof. Dr. Thales Terrola e Lopes - Coorientador
Universidade Federal Fluminense

Prof. Dr. Márcio Zamboti Fortes
Universidade Federal Fluminense

Prof. Dr. Bruno Henriques Dias
Universidade Federal de Juiz de Fora

Niterói – RJ
(março/2018)

DEDICATÓRIA

Aos meus pais, pelo amor, respeito e atenção especial dispensados durante todos esses anos.

Aos meus irmãos e sobrinhos queridos, amigos sempre presentes em todos os momentos.

À minha esposa, por sua paciência e pelo seu amor, ajudando-me a enfrentar todos os desafios impostos.

Aos meus amigos, pela paciência e companheirismo nos corridos finais de semana.

AGRADECIMENTOS

A Deus,

A toda a minha família,

À minha esposa,

Aos meus amigos,

À Universidade Federal Fluminense,

Aos professores Vitor Hugo e Thales Terrola, pelo empenho, pela dedicação e pela força em todos os momentos,

Aos demais professores do mestrado, que contribuíram com a sua experiência e dedicação.

RESUMO

O avanço da geração distribuída de energia solar fotovoltaica na matriz energética traz à tona uma série de problemas relacionados à operação e à segurança das redes de distribuição de energia elétrica, mostrando a necessidade de um monitoramento constante da produção de energia, visto a característica intermitente dessa fonte.

A previsão da produção de energia elétrica por sistemas solares fotovoltaicos a partir de modelos determinísticos, além de trabalhosa e complexa, visto o grande número de modelos necessários para a previsão e o grande número de parâmetros a serem especificados, não traz um dos aspectos mais importantes da análise da decisão que é a modelagem da incerteza, impossibilitando uma tomada de decisão segura e precisa.

Neste trabalho são abordados dois modelos neurais probabilísticos para previsão da produção de energia elétrica. O primeiro modelo, mais simples, realiza uma abordagem a partir de um único previsor. O segundo modelo, mais complexo que o primeiro, faz uma abordagem a partir da *clusterização* dos dados através de uma mistura de gaussianas.

Para dar confiabilidade aos resultados, as análises foram realizadas a partir de uma base de dados de sistemas solares fotovoltaicos reais conectados à rede. As previsões foram avaliadas através de métricas apropriadas para a previsão probabilística e apresentaram bons resultados nos dois casos, com baixos índices de erros, tendo o modelo de mistura de gaussianas alcançado melhor resultado que o modelo de previsor único.

Palavras-chave: previsão da produção de energia solar; previsão probabilística; previsão de energia.

ABSTRACT

The advance of distributed generation of solar photovoltaic energy in the energy matrix brings to light a series of problems related to the operation and safety of electricity distribution networks, showing the need for constant monitoring of energy production, given the intermittent characteristic from this source.

The prediction of the production of electric energy by photovoltaic solar systems from deterministic models, in addition to being laborious and complex, because of the large number of models required for the forecast and the large number of parameters to be specified, does not show one of the most analysis of the decision, which is the modeling of uncertainty, which prevents a secure and precise decision-making.

In this work two probabilistic neural network models for the prediction of solar energy are discussed. The first model takes an approach from a single predictor. The second model takes an approach of grouping the data through a mixture of Gaussian.

The analyzes were carried out from a database of real solar photovoltaic systems connected to the grid to increase to the results. The predictions were evaluated using appropriate metrics for probabilistic forecasting and presented satisfactory results in both cases, with low error rates. As expected, the Gaussian mixture-based model was better than the single-fader model.

Keywords: solar energy forecasting; probabilistic forecasting; energy forecasting.

LISTA DE FIGURAS

Figura 1: Componentes da radiação solar incidindo sobre um painel solar [10]	19
Figura 2: Efeito causado pela variação da irradiância solar sobre a curva característica I-V [10]	23
Figura 3: Efeito causado pela variação da temperatura na curva característica I-V [10]	23
Figura 4: Curva de eficiência do inversor – <i>CEC</i> – <i>California Electrical Code</i> [17]	24
Figura 5: Curva da potência versus eficiência e curva de redução de temperatura [18]	25
Figura 6: Diagrama esquemático de um neurônio [22]	27
Figura 7: Função tangente hiperbólica para diferentes valores de ganho [22].....	29
Figura 8: Função logística para diferentes valores de ganho [22]	30
Figura 9: Rede neural <i>feedforward</i> com múltiplas camadas e múltiplas saídas [22].....	31
Figura 10: Estrutura do <i>MLP</i> utilizado para classificação de padrões [22].....	35
Figura 11: Estrutura do <i>MLP</i> utilizado para regressão [22].....	42

LISTA DE TABELAS

Tabela 1: Variáveis medidas	58
Tabela 2: Resultados do <i>pinball</i> – Modelo Único	61
Tabela 3: Resultados das simulações – Zonas 1, 2 e 3 – Modelo Único	64
Tabela 4: Resultados das simulações – Tarefa 12 (junho) – Modelo Único.....	65
Tabela 5: Resultados das simulações – Tarefa 8 (fevereiro) – Modelo Único.....	65
Tabela 6: Resultados do Pinball – Modelo GMM.....	66
Tabela 7: Redução do pinball alcançada com o modelo GMM	67
Tabela 8: Resultados das simulações – Zonas 1, 2 e 3 – Modelo GMM.....	69
Tabela 9: Resultados das simulações – Tarefa 12 (junho) – Modelo GMM	69
Tabela 10: Resultados das simulações – Tarefa 8 (fevereiro) – Modelo GMM.....	69

LISTA DE GRÁFICOS

Gráfico 1: Curvas do pinball dos três primeiros colocados, do modelo Único e do <i>benchmark</i>	62
Gráfico 2: Classificação dos três primeiros colocados, do modelo Único e do <i>benchmark</i> ..	63
Gráfico 3: Curvas do pinball – Modelo Único	63
Gráfico 4: Curvas do pinball dos três primeiros colocados, do modelo GMM e do <i>benchmark</i>	67
Gráfico 5: Classificação dos três primeiros colocados, do modelo GMM e do <i>benchmark</i> ..	68
Gráfico 6: Previsão da produção de energia solar (setembro – zona 3).....	70
Gráfico 7: Previsão da produção de energia solar (fevereiro – zona 1).....	71
Gráfico 8: Previsão da produção de energia solar (junho – zona 3).....	71

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Problema	13
1.2	Objetivo	14
1.3	Revisão da Literatura	14
1.4	Produção Científica	16
1.5	Estrutura do Trabalho	17
2	PREVISÃO DE ENERGIA SOLAR	19
2.1	Modelos Físicos	19
2.1.1	Modelos de Transformação do Plano	20
2.1.2	Modelo de desempenho do arranjo	22
2.1.3	Modelo térmico	22
2.1.4	Modelo de desempenho de um inversor fotovoltaico	23
2.2	Modelos Baseados em Aprendizado de Máquina	25
2.2.1	Redes Neurais Artificiais	25
2.2.2	Previsão Probabilística	33
2.2.3	Treinamento Bayesiano Aplicado a <i>MLPs</i>	33
2.2.4	Modelos de <i>Clusterização</i>	45
3	METODOLOGIA PROPOSTA	52
3.1	Modelo Único	52
3.2	Modelo GMM	53
3.3	Critérios de Análise	55
3.3.1	Erro Relativo Percentual Médio	55
3.3.2	Erro Relativo	55
3.3.3	Média ponderada	55
3.3.4	Função de Perda de <i>Pinball</i>	56
4	APRESENTAÇÃO E ANÁLISE DOS RESULTADOS	57
4.1	Dados da Competição	57
4.1.1	Base de dados medida	57
4.1.2	Sistemas fotovoltaicos	58
4.1.3	Classificação	58
4.2	Seleção da base de dados	60
4.3	Resultados	60
4.3.1	Modelo Único	61

4.3.2	Modelo GMM.....	65
5	CONCLUSÃO E TRABALHOS FUTUROS	72
	BIBLIOGRAFIA.....	73
	APÊNDICE	76

1 INTRODUÇÃO

O avanço das redes inteligentes permite a inserção de novas tecnologias e serviços em sistemas de energia para tornar as redes elétricas mais confiáveis, eficientes e seguras. O aumento da geração distribuída a partir de fontes renováveis de energia pode trazer uma série de benefícios tanto para o setor elétrico quanto para o consumidor, como a redução das perdas elétricas por causa da proximidade com a carga e a redução dos custos de investimentos pela concessionária.

A energia solar fotovoltaica vem sendo amplamente adotada por suas facilidades e comodidades, como a facilidade de instalação, o baixo índice de manutenção, tecnologia amplamente difundida e a abundância do recurso energético. Entretanto, o comportamento estocástico de tais fontes de energia introduz questões desafiadoras no projeto de ferramentas e técnicas avançadas para o ótimo funcionamento e controle das redes de distribuição.

A previsão da produção de energia é uma ferramenta fundamental para a utilização eficiente dos recursos energéticos, o equilíbrio entre a geração e a demanda, a avaliação contínua de desempenho dos sistemas, o despacho econômico, a tomada de decisões no mercado de energia, a segurança do sistema, dentre outros benefícios.

Na literatura há uma série de modelos determinísticos e probabilísticos para previsão da energia solar. Em geral, essas modelagens envolvem o uso de variáveis meteorológicas, como radiação solar, temperatura ambiente, nebulosidade, umidade do ar, velocidade do vento e outras. Segundo [1], tais modelos devem ser desenvolvidos para caracterizar, avaliar e, sempre que possível, quantificar a incerteza. Desta forma, a conscientização para a incerteza é aumentada e os usuários passam a ter os meios necessários para quantificar e prever o risco.

1.1 Problema

A previsão da produção de energia elétrica a partir da energia solar fotovoltaica usando análise determinística exige o conhecimento aprofundado de diversos modelos matemáticos e físicos. Obviamente, existem métodos simplificados que não retornam bons

resultados, e, portanto, sequer devem ser considerados. Para um resultado satisfatório, o estudo deve levar em consideração o levantamento e a análise dos dados meteorológicos no local do projeto, a aplicação ou não do modelo de transformação do plano horizontal para o plano inclinado, o modelo de desempenho dos principais componentes do sistema e as condições de contorno específicas de cada projeto.

Nas modelagens mais sofisticadas, o grande número de parâmetros e de métricas distintas, além de dificultar a análise, pode ainda não retornar bons resultados. Os modelos de transformação do plano inclinado (até mesmo o mais sofisticado deles) não conseguem modelar a componente difusa da radiação solar. Apesar da existência de uma série de modelos para o painel solar e para o inversor fotovoltaico, a grande maioria não modela as especificidades de cada equipamento. Os poucos que modelam esses parâmetros necessitam de dados empíricos, o que dificulta a previsão precisa da produção de energia. Por fim, por serem determinísticos, esses modelos não modelam incertezas de nenhuma natureza.

1.2 Objetivo

O objetivo deste trabalho consiste em modelar uma rede neural artificial (RNA) para previsão probabilística da produção de energia elétrica a partir de sistemas solares fotovoltaicos.

A rede neural artificial é um modelo matemático desenvolvido com o intuito de simular o funcionamento do cérebro humano que apresenta uma série de vantagens na resolução de problemas, como autoaprendizado, capacidade de generalização, não linearidade e outros. A vantagem imediata que a RNA traz para o problema exposto é a capacidade de realizar um mapeamento dos dados, apresentados a partir de um conjunto de pares entrada-saída, sem que haja a necessidade de modelar matematicamente o problema [2]. A rede neural fornece a previsão pontual da produção de energia elétrica por um sistema solar fotovoltaico. As previsões probabilísticas fornecem informações sobre a incerteza associada aos erros de previsão. Apesar desta incerteza poder ser expressa de diversas formas, neste trabalho ela será expressa por um conjunto de quantis, em decorrência da base de dados adotada, disponibilizada na competição de previsão de energia *Global Energy Forecasting Competition 2014 (GEFCom 2014)* [3], apresentar os seus resultados com base nas previsões probabilísticas de 99 percentis.

1.3 Revisão da Literatura

A natureza estocástica da produção de energia solar fotovoltaica tem motivado a pesquisa por modelos de previsão para apoio à operação de tais sistemas. Em [4], é

apresentado um método baseado na inferência bayesiana que consiste em realizar uma previsão de curto prazo da potência ativa produzida por um sistema fotovoltaico a partir de uma estimativa do índice de claridade horária. Esse método leva em consideração a dependência da radiação solar terrestre com algumas variáveis atmosféricas, como a cobertura de nuvem e a umidade. O modelo em análise parte da premissa de que quando o sistema fotovoltaico está equipado com um rastreador do ponto de potência máxima (*MPPT*), existe uma relação analítica entre a potência P_{PV_t} na hora t e o índice de claridade K_t na referida hora. A combinação de técnicas probabilísticas, como inferência bayesiana e simulação de Monte Carlo, permite fornecer a função de densidade de probabilidade preditiva da potência do gerador fotovoltaico, o que é muito útil para o ótimo funcionamento e controle das redes inteligentes do futuro. No geral, as previsões apresentaram bons resultados, com erro relativo absoluto médio entre 14,5% no inverno e 18% no outono. Foram observados também que os valores reais da potência ativa horária estão sempre compreendidos entre os valores do 5º percentil e do 95º percentil, e que a relação não linear entre o índice de claridade e a saída de energia fotovoltaica pode refletir em erros não negligenciáveis nas distribuições previstas da energia fotovoltaica.

Em [5], é apresentado um multimodelo probabilístico que combina o método de inferência bayesiana com a regressão de quantis como preditores individuais. Nesse método, é utilizada uma função de distribuição de probabilidade beta para modelar a aleatoriedade da potência fotovoltaica (a premissa para a potência é a mesma do estudo [4]), tendo sido o modelo avaliado através de um grande número de aplicações numéricas com base em dados reais de irradiação. Os resultados apresentados, em geral, foram bons, porém, em alguns momentos, foi possível observar uma ligeira dispersão nas previsões resultantes. Significativamente, foram obtidas melhores previsões para o próximo dia, enquanto foram obtidos resultados menos significativos em caso de previsão da próxima hora.

Em [6], é proposto um modelo de análise de regressão linear múltipla para gerar previsões probabilísticas de produção de energia. Este trabalho também faz uso da base de dados *GEFCom* 2014. Nesse modelo, os dados são criteriosamente analisados, sendo selecionadas apenas as variáveis mais relevantes para a análise. Das variáveis selecionadas, duas delas foram utilizadas para criar, cada uma, uma função de segunda ordem, e, uma delas, para criar uma função de terceira ordem. Essas três funções foram então utilizadas para gerar os dados de três novas variáveis. As três novas variáveis substituíram as duas que lhes deram origem. O modelo proposto utiliza a técnica de análise de variância (ANOVA) [7], construída em um programa estatístico. Os resultados mostram

que a análise foi realizada apenas para um mês (setembro de 2013), tendo sido obtidos bons resultados com o modelo proposto.

Em [8], é apresentado um modelo estocástico para fornecer previsões probabilísticas de curto prazo para a energia solar fotovoltaica. O modelo proposto é inspirado na física e incorpora explicitamente a estocasticidade devido às nuvens usando diferentes parâmetros que abordam a atenuação da potência. Com base no comportamento estatístico dos parâmetros, propõe-se um simples processo de mudança no regime entre as três classes de sol (aberto, nublado e parcialmente nublado). Então, as previsões probabilísticas da potência solar fotovoltaica são obtidas identificando o atual regime usando medições da potência solar fotovoltaica e assumindo a persistência neste regime. O método de previsão proposto, segundo os resultados apresentados, supera uma variedade de modelos de referência que produzem previsões pontuais e probabilísticas, sendo, porém, o seu uso mais adequado para horizontes mais curtos, inferiores a 4 horas, uma vez que a persistência na condição climática é assumida.

O estudo proposto em [9] apresenta um modelo para previsão probabilística de curto prazo não paramétrica para geração de energia renovável, com aplicação à energia solar. Para isso, é utilizada uma máquina de aprendizado extremo (em inglês *ELM*), um modelo de regressão que pode gerar previsões de ponto e de quantis, dependendo do critério de treinamento. Os parâmetros para a camada oculta são selecionados aleatoriamente enquanto eles não precisam ser ajustados através de um processo de treinamento, e os pesos de saída são calculados analiticamente de uma vez sem passar por um longo treinamento iterativo. Para descrever as densidades preditivas, são emitidos 18 quantis com níveis nominais de 0,05 a 0,95, com incrementos de 0,05, com exceção da mediana (0,5). As análises são realizadas a partir de bases de dados reais de dois locais com diferentes regiões climáticas. Para validar a abordagem para a aplicação da energia solar, o modelo proposto foi avaliado contra quatro *benchmarks* relevantes, que podem ser consultados no referido trabalho. Para a grande maioria das análises realizadas, os resultados do modelo proposto foram melhores que os dos demais modelos.

1.4 Produção Científica

- Martins, R.; Ferreira, V.; Lopes, T., “Rede Neural Artificial para Previsão Probabilística de Energia Solar”, *em fase de submissão*.

Neste trabalho são abordados dois modelos neurais probabilísticos para previsão da produção de energia solar, usando como um dos parâmetros de entrada a radiação solar global horizontal. O primeiro modelo realiza uma abordagem a partir de um único previsor. O segundo modelo faz uma abordagem a partir da *clusterização* dos

dados através de uma mistura de gaussianas. As análises foram realizadas a partir de uma base de dados de sistemas solares fotovoltaicos reais conectados à rede. As previsões foram avaliadas através de métricas apropriadas para a previsão probabilística e apresentaram bons resultados nos dois casos, com baixos índices de erros.

- Martins, R.; Ferreira, V.; Lopes, T., “Rede Neural Artificial para Previsão Probabilística da Potência de Saída de Sistemas Fotovoltaicos”, *VII Simpósio Brasileiro de Sistemas Elétricos*, 12-16 Maio 2018.

Aborda o uso de uma rede neural artificial para previsão probabilística de energia solar, usando a radiação solar global horizontal como a principal variável de entrada, e a potência na saída do inversor como a variável de saída. A precisão das séries geradas pelas redes neurais é avaliada pela função de perdas de *pinball*, método de avaliação mais indicado atualmente para previsão probabilística.

- Martins, R.; Ferreira, V.; Lopes, T., “Rede Neural Artificial para Previsão da Potência de Saída de Sistemas Fotovoltaicos Conectados à Rede”, *VI Simpósio Brasileiro de Sistemas Elétricos*, 22-25 Maio 2016.

Aborda o uso de uma rede neural artificial para previsão da produção de energia solar, usando a radiação solar no plano do arranjo como a principal variável de entrada e a potência na saída do inversor como variável de saída. Paralelamente, são avaliados diferentes algoritmos de treinamento, de modo a determinar qual o melhor modelo de rede a ser aplicada. Os resultados obtidos mostram que as redes neurais avaliadas apresentam bom desempenho para a solução do problema proposto, que pode ser comprovado pelo baixo erro de estimação.

1.5 Estrutura do Trabalho

O capítulo 2 apresenta a previsão da produção de energia solar sob a ótica dos modelos de aprendizado de máquina, mostrando as redes neurais artificiais, a previsão probabilística, o treinamento bayesiano e os modelos de *clusterização*.

O capítulo 3 apresenta a metodologia proposta, que consiste na junção dos diferentes modelos de aprendizado de máquina apresentados, de modo a realizar a previsão probabilística da produção de energia elétrica a partir de sistemas solares fotovoltaicos. Para tal, são apresentados dois modelos neurais, onde um é composto de um único previsor, e o outro é composto de um conjunto de previsores, representados por uma mistura de gaussianas.

O capítulo 4 aborda a apresentação e a análise dos resultados, fazendo uma análise crítica sobre os principais resultados.

O capítulo 5 finaliza o presente trabalho mostrando a conclusão e as propostas para trabalhos futuros.

2 PREVISÃO DE ENERGIA SOLAR

Neste capítulo são apresentados os modelos físicos e os modelos baseados em aprendizado de máquina. A abordagem realizada na seção 2.1 tem por objetivo mostrar, de maneira bem resumida, as limitações dos modelos físicos.

2.1 Modelos Físicos

A radiação solar que atinge a superfície terrestre é constituída por uma componente direta, que passa pela atmosfera e atinge diretamente a superfície terrestre, por uma componente difusa, que atinge a superfície após sofrer espalhamento na atmosfera terrestre em todas as direções, e por uma componente refletida (albedo), que é originada pela reflexão do entorno quando a superfície está inclinada em relação à superfície horizontal. A Figura 1 mostra as componentes da radiação solar incidindo sobre um painel inclinado.

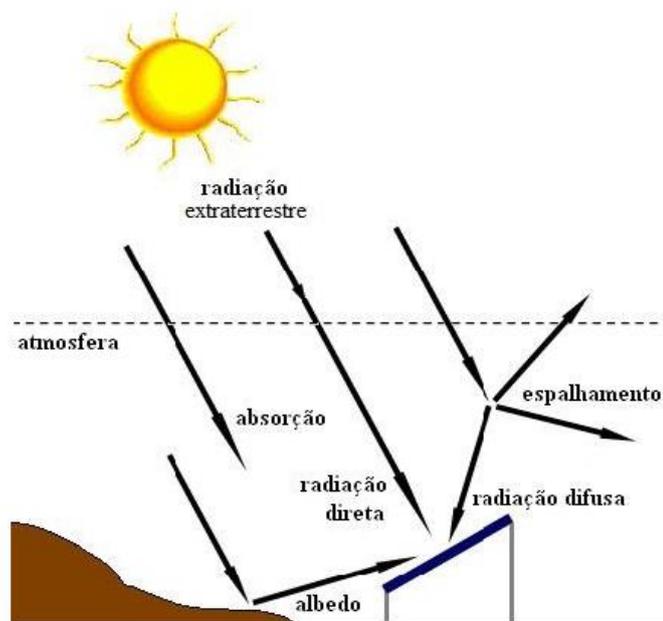


Figura 1: Componentes da radiação solar incidindo sobre um painel solar [10]

A medição dessa radiação é feita, em geral, na superfície horizontal, em que são medidas a irradiação solar global horizontal, a irradiação solar difusa, a irradiação normal

direta e outras. E essa medição é feita na superfície horizontal por essa ser a base de transformação para qualquer superfície inclinada. A medição da radiação solar em superfícies inclinadas até ocorre, porém, é incomum. E, quando isso ocorre, na maioria das vezes, a medição é feita com equipamentos de baixa precisão. É importante destacar que, ainda que seja possível realizar uma medição de precisão em uma superfície inclinada, essa medição só terá utilidade para essa superfície. Diferentemente, a medição na superfície horizontal tem uma série de utilidades: elaboração de atlas solarimétricos, criação de uma base de dados histórica, utilização em qualquer superfície inclinada e orientada, e outras.

2.1.1 Modelos de Transformação do Plano

Os modelos de transformação do plano transformam a irradiação solar incidente em uma superfície horizontal para qualquer superfície inclinada. Existe uma série de modelos de transformação do plano, uns mais simples, outros mais sofisticados. O grande número de modelos se deve basicamente à modelagem da componente difusa. Destacam-se: modelo do céu isotrópico, modelo de Hay-Davies, modelo de Reindl, modelo de Muneer, modelo de Perez e outros [11], [12].

Nas subseções a seguir são apresentados de maneira resumida dois modelos de transformação do plano: o modelo isotrópico do céu, o mais simples, e o modelo de Perez, o mais sofisticado. Para a leitura deles, considere que $I_{h,b}$ é a componente da irradiância direta normal, R_b é a componente da irradiância direta que incide no plano do arranjo, $I_{h,d}$ é a irradiância solar global difusa horizontal, I_h é a irradiância solar global horizontal, β é a inclinação do arranjo em relação à superfície horizontal e ρ é o albedo.

2.1.1.1 Modelo Isotrópico do Céu

O modelo isotrópico do céu é o mais simples dos modelos da radiação difusa, sendo a base sobre a qual os modelos mais precisos são construídos. Este modelo assume que a radiação difusa da cúpula do céu é uniforme em todo o céu e que a reflexão na superfície é difusa [11]. A irradiação total na superfície inclinada I_T é dada pela soma das parcelas das componentes direta, difusa e refletida, de acordo com a expressão (2-1):

$$I_T = \underbrace{I_{h,b} R_b}_{\text{Direta}} + \underbrace{I_{h,d} \left(\frac{1 + \cos \beta}{2} \right)}_{\text{Difusa}} + \underbrace{I_h \rho \left(\frac{1 - \cos \beta}{2} \right)}_{\text{Refletida}} \quad (2-1)$$

2.1.1.2 Modelo de Perez

O modelo de Perez apresenta uma análise completa para a radiação difusa, a partir da análise detalhada das componentes difusa isotrópica, brilho circunsolar e brilho

horizontal, usando coeficientes empiricamente derivados [11]. A irradiação difusa que vem da região ao redor do disco solar é chamada de brilho circunsolar. O aumento na intensidade da radiação difusa próximo ao horizonte, decorrente da radiação incidente que percorre uma trajetória maior na atmosfera dessa região e das múltiplas reflexões existentes na atmosfera terrestre, é chamada de brilho horizontal. O restante da radiação difusa uniformemente distribuída na porção restante do hemisfério celeste recebe o nome de isotrópica [13]. A irradiação total na superfície inclinada I_T é dada pela soma das parcelas das componentes direta, difusa e refletida, de acordo com a expressão (2-2):

$$I_T = \underbrace{I_{h,b} R_b}_{\text{Direta}} + \underbrace{I_{h,d} \left[(1 - F_1) \left(\frac{1 + \cos \beta}{2} \right) + F_1 \frac{a}{b} + F_2 \text{sen} \beta \right]}_{\text{Difusa}} + \underbrace{I_h \rho \left(\frac{1 - \cos \beta}{2} \right)}_{\text{Refletida}} \quad (2-2)$$

Na equação (2-2), $a = \max(0, \cos \theta)$, em que θ é o ângulo de incidência da irradiância direta em relação à superfície horizontal, e $b = \max(\cos 85^\circ, \cos \theta_z)$, em que θ_z é o ângulo zenital, e ρ é o albedo. Analisando primeiramente a componente difusa, os coeficientes do brilho circunsolar F_1 e do brilho horizontal F_2 podem ser obtidos a partir das seguintes expressões:

$$F_1 = \max \left[0, (f_{11}(\varepsilon) + f_{12}(\varepsilon) \cdot \Delta + f_{13}(\varepsilon) \cdot \theta_z) \right] \quad (2-3)$$

$$F_2 = f_{21}(\varepsilon) + f_{22}(\varepsilon) \cdot \Delta + f_{23}(\varepsilon) \cdot \theta_z \quad (2-4)$$

$$\varepsilon = \frac{(I_{h,d} + I_n) / I_{h,d} + 1,041 \theta_z^3}{1 + 1,041 \theta_z^3} \quad (2-5)$$

$$\Delta = m \frac{I_{h,d}}{I_{on}} \quad (2-6)$$

$$I_n = \frac{I_h - I_{h,d}}{\cos(\theta_z)} \quad (2-7)$$

Nas equações (2-3) e (2-4), os coeficientes f_{ij} , derivados com base na análise de dados empíricos de localidades específicas, são dados em função das categorias do índice de claridade do céu. Os coeficientes do brilho circunsolar e horizontal F_1 e F_2 dependem dos parâmetros de condição do céu: do ângulo zenital θ_z , do índice de claridade ε dado pela equação (2-5), do brilho Δ dado pela equação (2-6), da irradiância direta normal I_n , da

irradiância solar global horizontal I_h , da irradiância solar global difusa horizontal $I_{h,d}$ e da irradiância normal direta extraterrestre I_{on} .

2.1.2 Modelo de desempenho do arranjo

Existe uma série de modelos analíticos para verificar o desempenho de um arranjo solar fotovoltaico. O modelo selecionado para servir de referência para a obtenção das variáveis que mais influenciam no funcionamento do gerador fotovoltaico apresenta um dos estudos mais detalhados acerca do gerador fotovoltaico.

Baseado empiricamente, o modelo de desempenho do arranjo alcança a sua versatilidade e precisão por utilizar equações derivadas de características individuais das células solares. A versatilidade e a precisão do modelo foram demonstradas para módulos solares de placas planas de todas as tecnologias, para módulos concentradores, bem como para grandes arranjos de módulos [14] [15].

O modelo considera os efeitos elétricos, térmicos, espectrais solares e ópticos para módulos fotovoltaicos. As variáveis que definem a condição de funcionamento são irradiância, temperatura da célula, massa de ar absoluta e ângulo de incidência solar no arranjo. A corrente de curto circuito, a corrente de potência máxima, a tensão de potência máxima e a tensão de circuito aberto fornecem os parâmetros necessários para o cálculo do fator de preenchimento, da potência máxima e da eficiência. A premissa fundamental desse modelo de desempenho é que a corrente de potência máxima, a tensão de potência máxima e a tensão de circuito aberto de uma célula, módulo ou arranjo são parâmetros bem comportados e previsíveis quando descritos em função da corrente de curto circuito.

A Figura 2 mostra o efeito causado pela variação da irradiação solar sobre a corrente e a tensão de um painel solar.

A Figura 3 mostra o efeito da variação da temperatura sobre a corrente e a tensão. O aumento da temperatura reduz a tensão e aumenta a corrente do gerador fotovoltaico. A redução da temperatura provoca o efeito inverso.

2.1.3 Modelo térmico

A temperatura é um dos parâmetros que mais afeta o funcionamento do sistema fotovoltaico como um todo, afetando módulos e inversores no caso dos sistemas conectados à rede. O modelo térmico simplificado para os arranjos fotovoltaicos de placas planas apresenta um estudo completo sobre a temperatura nas superfícies dos módulos, sendo usado para calcular a temperatura da célula a partir de poucos dados de entrada, como a radiação solar no plano do arranjo, a temperatura ambiente, a velocidade do vento, a altura

do anemômetro, a altura do arranjo em relação ao solo e a temperatura INOCT (*Installed Nominal Operating Cell Temperature*) [16].

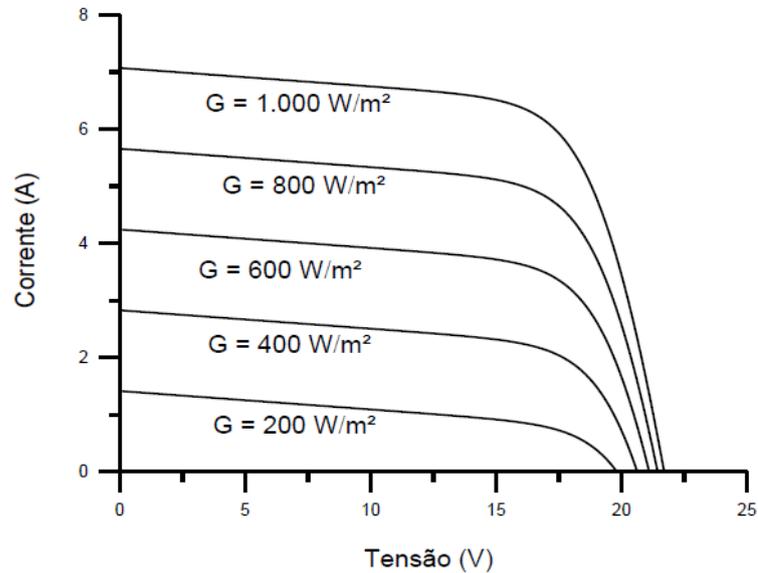


Figura 2: Efeito causado pela variação da irradiância solar sobre a curva característica I-V [10]

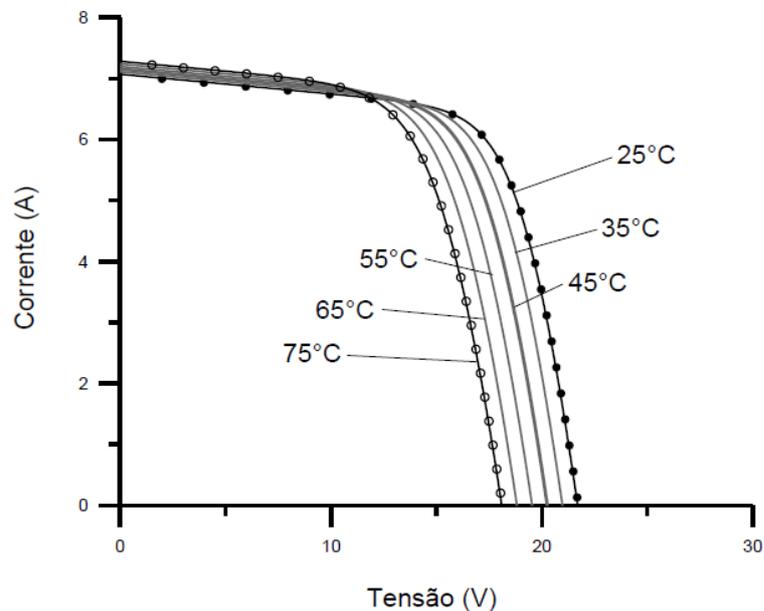


Figura 3: Efeito causado pela variação da temperatura na curva característica I-V [10]

2.1.4 Modelo de desempenho de um inversor fotovoltaico

O inversor fotovoltaico tem a função de converter a energia em corrente contínua gerada pelos módulos solares em corrente alternada para injeção na rede de distribuição de energia elétrica. Durante esse processo, parte dessa energia é injetada na rede, parte é consumida pelo próprio inversor e parte é perdida através de calor. A eficiência de um

inversor solar fotovoltaico indica a quantidade de energia em corrente contínua que é convertida em corrente alternada. A fórmula de eficiência geral é dada pela expressão (2-8):

$$\eta_{inv} = \frac{P_{CA}}{P_{CC}} \quad (2-8)$$

Na equação (2-8), P_{CA} é saída de energia em corrente alternada e P_{CC} é entrada de energia. A eficiência do inversor depende da carga do inversor. A Figura 4 mostra a curva característica de eficiência do inversor para três diferentes níveis de tensão.

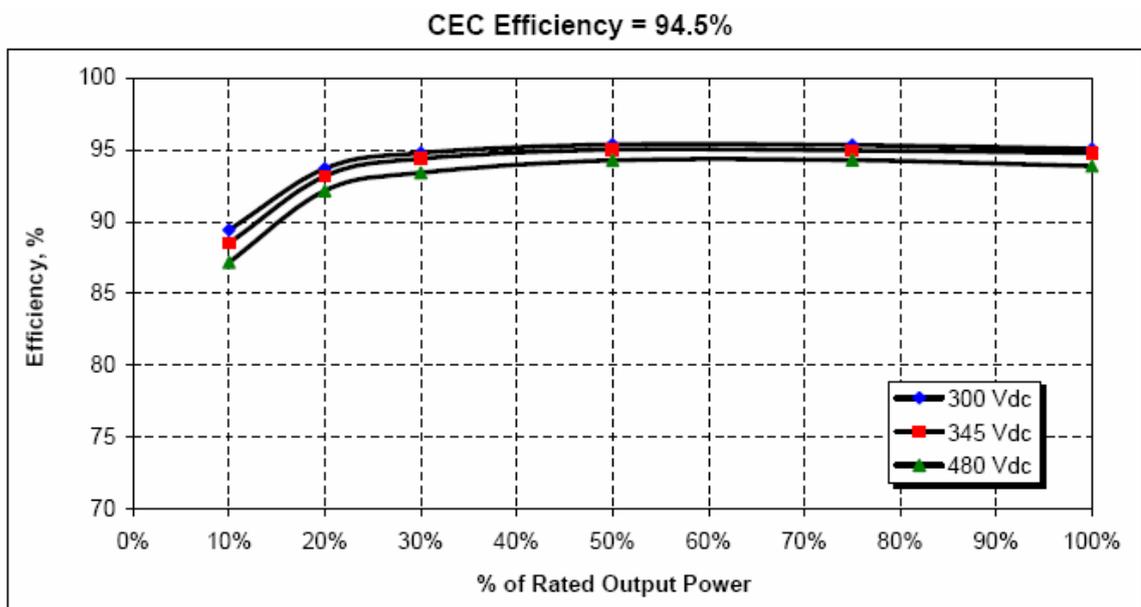


Figura 4: Curva de eficiência do inversor – CEC – California Electrical Code [17]

A eficiência do inversor pode ser avaliada através dos critérios estabelecidos por instituições confiáveis como a *Joint Research Center* da Comissão Europeia ou a *California Energy Commission (CEC)*, conhecidos, respectivamente, como eficiência europeia e eficiência CEC, dadas pelas expressões (2-9) e (2-10):

$$\eta_{Eu} = 0,03\eta_{5\%} + 0,06\eta_{10\%} + 0,13\eta_{20\%} + 0,10\eta_{30\%} + 0,48\eta_{50\%} + 0,20\eta_{100\%} \quad (2-9)$$

$$\eta_{CEC} = 0,04\eta_{10\%} + 0,05\eta_{20\%} + 0,12\eta_{30\%} + 0,21\eta_{50\%} + 0,53\eta_{75\%} + 0,05\eta_{100\%} \quad (2-10)$$

A temperatura afeta o funcionamento do inversor fotovoltaico podendo ocasionar, dependendo do fabricante, a diminuição gradativa da potência em decorrência do aumento da temperatura, reduzir a potência para um patamar fixo ou até mesmo desligar o inversor

até que a temperatura seja reduzida. A Figura 5 mostra a curva de eficiência para um inversor fotovoltaico de bateria que pode operar conectado à rede.

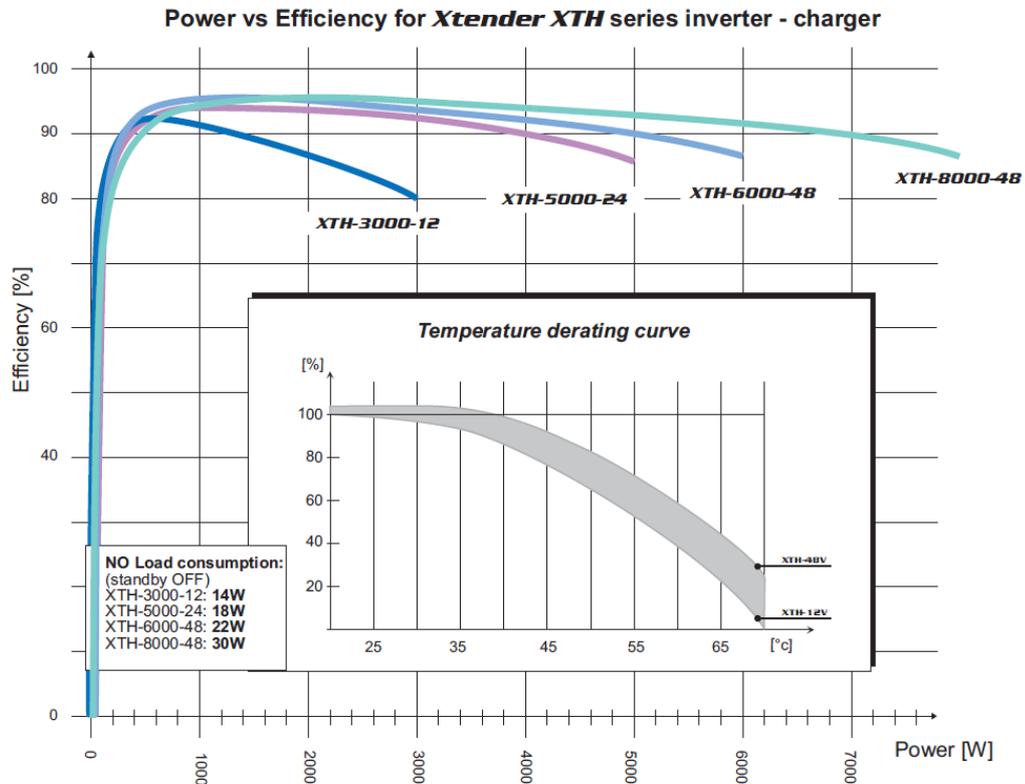


Figura 5: Curva da potência versus eficiência e curva de redução de temperatura [18]

2.2 Modelos Baseados em Aprendizado de Máquina

2.2.1 Redes Neurais Artificiais

A rede neural artificial é um modelo matemático desenvolvido com o intuito de simular o funcionamento do cérebro humano. Segundo [19], o cérebro é um computador altamente complexo, não linear e paralelo, que tem a capacidade de organizar seus neurônios de forma a realizar certos processamentos – por exemplo, reconhecimento de padrões, percepção e controle motor – muito mais rapidamente que o mais rápido computador digital existente.

O cérebro realiza rotineiramente tarefas de reconhecimento perceptivo, como a identificação de um rosto familiar em uma cena não familiar, em aproximadamente 100-200 milissegundos, diferentemente dos computadores convencionais que levam dias para executar tarefas mais simples [19]. Outra característica invejável do cérebro é a habilidade de desenvolver suas próprias regras através da experiência.

As redes neurais artificiais se destacam por uma série de vantagens, sendo as mais importantes destacadas a seguir [2]:

- Não linearidade: considerando o caso mais comum, em que a função de ativação dos neurônios da camada oculta é não linear, o modelo neural resultante da interconexão destas unidades mais simples apresenta considerável grau de não linearidade. Entretanto, esta vantajosa característica pode ser prejudicial na presença de dados ruidosos.
- Mapeamento entrada-saída: a partir de um conjunto de pares entrada-saída, as redes neurais realizam um mapeamento destes dados, sem que haja a necessidade de modelar matematicamente o problema.
- Adaptabilidade: estes modelos apresentam elevada capacidade de adaptação em virtude de mudanças nas condições do ambiente para o qual a rede foi treinada para operar. Para tal, basta treinar a rede novamente, incluindo no conjunto de treinamento os padrões referentes às novas condições operacionais.
- Implementação simples: a implementação desses modelos é extremamente simples quando comparada com o grau de complexidade dos modelos que podem ser gerados.

Segundo [20], a pesquisa da rede neural artificial pode ser dividida em duas áreas de investigação. A primeira área, o problema direto, consiste em empregar técnicas de computação e engenharia para modelar o cérebro humano. A segunda área, o problema inverso, consiste em simular estruturas biológicas com o objetivo de criar sistemas computacionais ou de engenharia.

2.2.1.1 Estrutura

O funcionamento da rede neural artificial está baseado no funcionamento do sistema nervoso central, mais especificamente no funcionamento do neurônio real ou biológico. Segundo [21], o neurônio artificial é uma estrutura que procura simular a forma, o comportamento e as funções de um neurônio biológico. Analogamente ao neurônio biológico, no neurônio artificial, as entradas representam os dendritos e os pesos sinápticos representam as sinapses neuronais. Os estímulos captados pelas entradas são processados pela função de soma e influenciados pelo *bias*. A função de transferência representa o limiar de disparo do neurônio biológico.

A Figura 6 apresenta um diagrama esquemático de um neurônio, cuja saída c é dada pela equação (2-11):

$$c = \phi \left(\sum_{i=1}^n \omega_i x_i + b \right) \quad (2-11)$$

Na equação (2-11), c representa a saída do neurônio, $\underline{\omega} \in \mathbb{R}^n$, $\underline{\omega} = [\omega_1 \ \dots \ \omega_n]^t$, o conjunto de pesos sinápticos que ligam as entradas $\underline{x} \in \mathbb{R}^n$, $\underline{x} = [x_1 \ \dots \ x_n]^t$ a este neurônio. Estas entradas podem ser oriundas da saída de outro neurônio ou da própria camada de entrada. O parâmetro $b \in \mathbb{R}$ representa o *bias* associado e $\phi(\cdot): \mathbb{R} \rightarrow \mathbb{R}$ a função de ativação do neurônio.

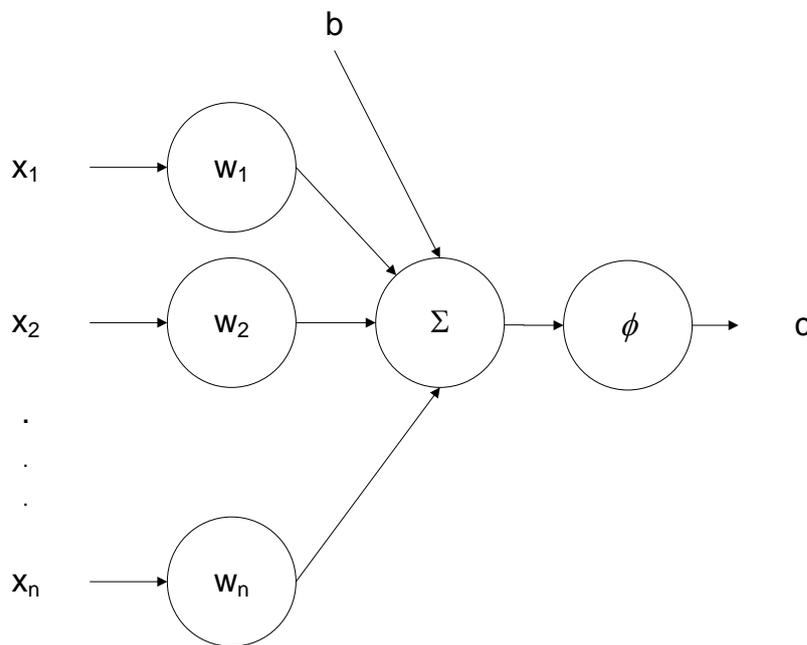


Figura 6: Diagrama esquemático de um neurônio [22]

2.2.1.2 Arquitetura

A arquitetura de uma rede neural pode ser de uma única camada ou de múltiplas camadas, do tipo recorrente ou não recorrente. Segundo [19], a estruturação dos neurônios que forma a arquitetura da rede está intimamente ligada com o algoritmo de aprendizagem usado para treinar a rede, sendo possível identificar três classes de arquiteturas distintas:

a) Rede alimentada adiante com uma camada

A rede alimentada adiante com uma camada (ou rede *feedforward* de uma camada) é uma rede do tipo não recorrente, que possui uma única camada de neurônios que é a própria camada de saída. Nas redes não recorrentes, as saídas dos neurônios de uma determinada camada se conectam apenas com as entradas dos neurônios da camada seguinte, sem a ocorrência de realimentação.

Esse tipo de rede é indicado para a solução de problemas linearmente separáveis como filtragem linear e classificação de padrões [23].

b) Redes alimentadas diretamente com múltiplas camadas

Também conhecida como rede *feedforward* de múltiplas camadas, essa rede possui uma ou mais camadas ocultas ou escondidas, que a torna capaz de extrair estatísticas de ordem elevada [19]. Mais informações a respeito desse tipo de rede são dadas na seção 2.2.1.4.

c) Redes recorrentes

As redes recorrentes representam as redes que possuem, pelo menos, um laço de realimentação, ou seja, que possuem conexões ligando um neurônio de uma determinada camada a outro neurônio de uma camada anterior ou a outro neurônio de uma mesma camada.

2.2.1.3 Função de ativação

Segundo [19], existem diversos tipos de funções de ativação. Algumas são apresentadas a seguir:

- a) Função de limiar: restringe a saída do neurônio em valores binários [0,1], sendo representada por:

$$\phi(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases} \quad (2-12)$$

- b) Função linear por partes: representada por:

$$\phi(x) = \begin{cases} 1 & \text{se } x \geq 1/2 \\ x & \text{se } \frac{1}{2} > x > -1/2 \\ 0 & \text{se } x \leq -1/2 \end{cases} \quad (2-13)$$

- c) Função sigmoideal logística: é definida como uma função crescente com balanceamento adequado entre o comportamento linear e não linear que assume valores de ativação no intervalo [0,1], sendo a inclinação da função sigmoide determinada pelo parâmetro a . Matematicamente:

$$\phi(x) = \frac{1}{1 + e^{-ax}} \quad (2-14)$$

- d) Função sigmoïdal tangente hiperbólica: a função tangente hiperbólica preserva a forma sigmoïdal da função logística, mas assume valores positivos e negativos. Essa função é representada por:

$$\phi(x) = \tanh(ax) \quad (2-15)$$

A função logística assume sempre valores positivos, enquanto a função tangente hiperbólica pode assumir valores positivos e negativos. Todavia, essas funções podem provocar a saturação dos dados de saída. A função linear não provoca o efeito de saturação como as funções sigmoïdes e hiperbólicas, sendo uma das funções mais utilizadas na camada de saída.

A Figura 7 apresenta o gráfico da função tangente hiperbólica para diferentes valores de ganho a . Em geral, este ganho a é feito igual a 1, permitindo que a inclinação da região linear seja calculada ao longo do processo de treinamento do modelo por meio da estimativa dos pesos sinápticos e do *bias*.

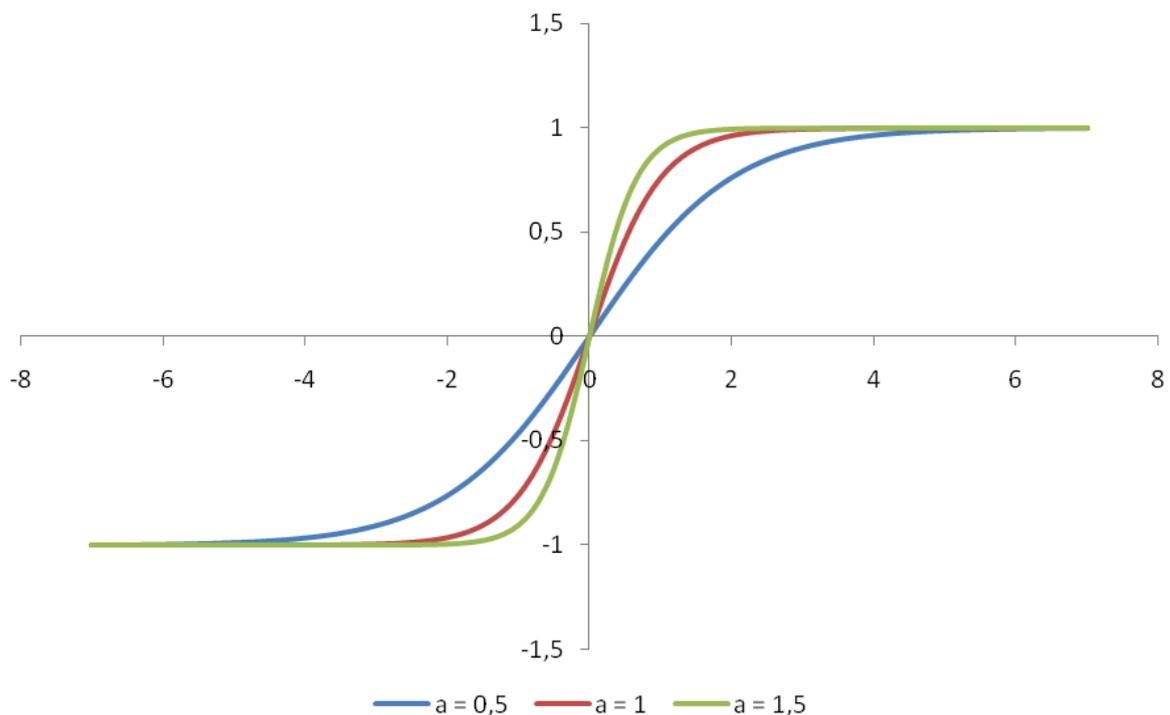


Figura 7: Função tangente hiperbólica para diferentes valores de ganho [22]

A Figura 8 apresenta o gráfico da função logística para diferentes valores de ganho a . Analogamente à função tangente hiperbólica, na função logística, o ganho é feito igual a 1

e a inclinação da região linear também é calculada ao longo do processo de treinamento do modelo por meio do cálculo dos pesos sinápticos e do *bias*.

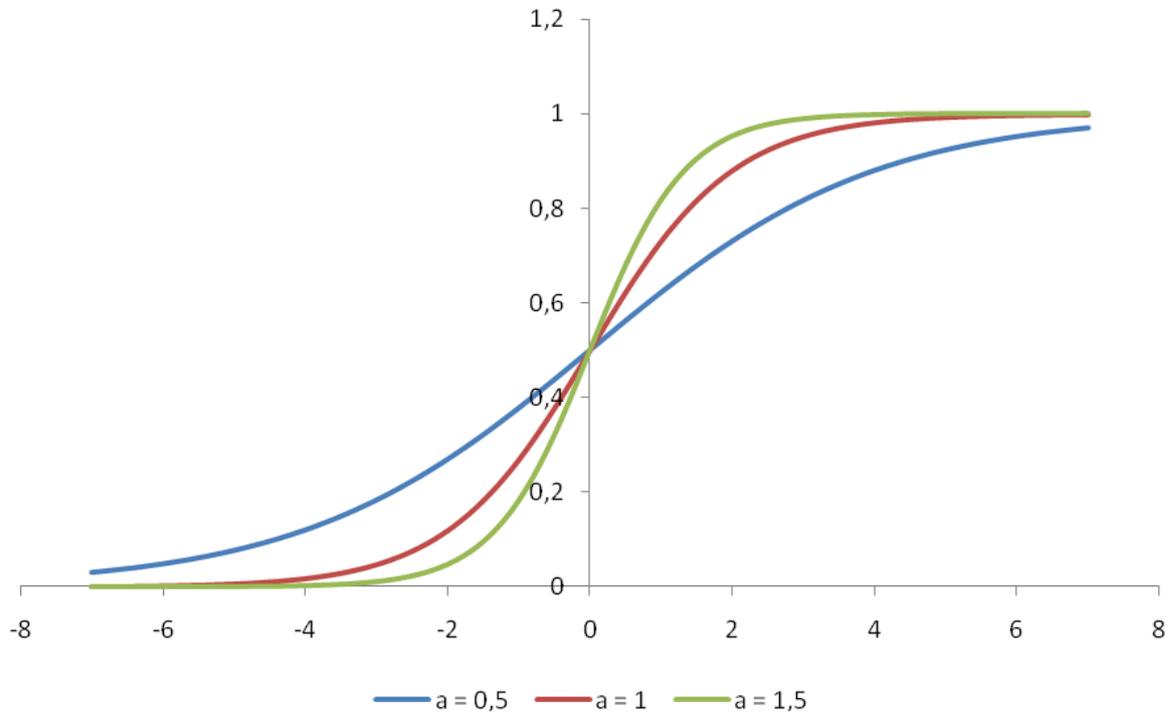


Figura 8: Função logística para diferentes valores de ganho [22]

2.2.1.4 Redes Perceptron de Múltiplas Camadas (MLPs)

A rede neural artificial do tipo perceptron de múltiplas camadas é uma rede baseada no perceptron, constituída por uma camada de entrada, uma ou mais camadas escondidas e uma camada de saída. Com exceção da camada de entrada, todas as demais camadas são formadas por neurônios.

A rede perceptron de múltiplas camadas é uma rede alimentada adiante ou progressiva (*feedforward*). Nela, o sinal de entrada se propaga adiante, camada a camada. A rede pode ser completamente conectada, quando cada nó se conecta a todos os nós da camada adjacente, ou parcialmente conectada, quando houver um ou mais nós não conectados, conforme pode ser visto na Figura 9.

Como dito na seção 2.2.1.2, a rede de uma única camada é indicada para a solução de problemas linearmente separáveis. A solução de problemas não linearmente separáveis é alcançada com o uso de redes com uma ou mais camadas escondidas. De acordo com o teorema da aproximação universal [24], uma rede com uma camada escondida pode implementar qualquer função contínua. Para tal, a estrutura deve apresentar ao menos uma camada oculta contendo neurônios com função de ativação contínua, não constante,

limitada, e pelo menos uma saída representando a aproximação de $F(\underline{x})$ gerada pelo modelo.

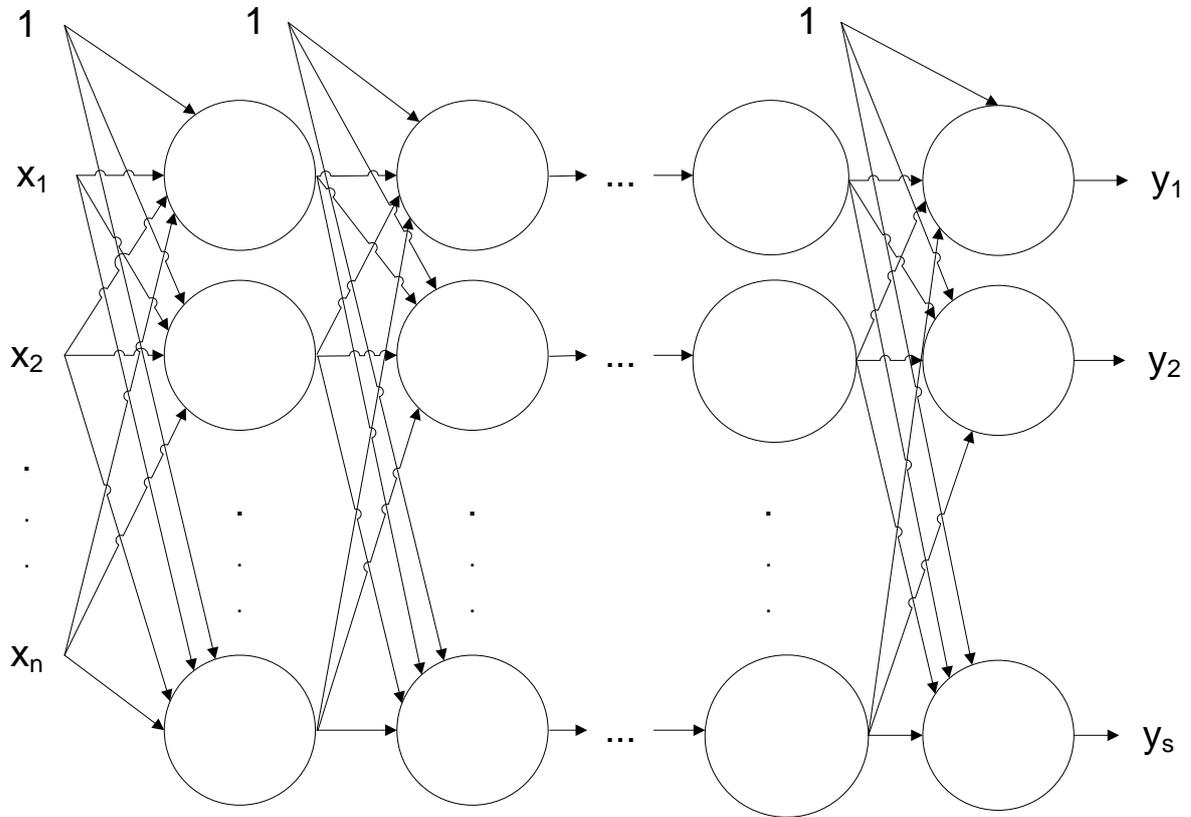


Figura 9: Rede neural *feedforward* com múltiplas camadas e múltiplas saídas [22]

A k -ésima saída y_k deste tipo de modelo, contendo m neurônios na camada oculta, é dada pela equação (2-16):

$$y_k = f_k(\underline{x}, \underline{w}) = \phi_{saída} \left[\sum_{j=1}^m \varpi_{kj} \phi_{oculta} \left(\sum_{i=1}^n \varpi_{ji} x_i + b_j \right) + b_k \right] \quad (2-16)$$

Na equação (2-16), $\underline{w}_k \in \mathbb{R}^m$, $\underline{w}_k = [\varpi_{k1} \ \dots \ \varpi_{km}]^t$, representa os pesos que ligam os neurônios da camada oculta ao k -ésimo neurônio de saída, $\underline{w}_j \in \mathbb{R}^n$, $\underline{w}_j = [\varpi_{j1} \ \dots \ \varpi_{jn}]^t$, constituído pelos pesos que ligam as entradas ao j -ésimo neurônio da camada escondida, $b_j \in \mathbb{R}$ o *bias* deste neurônio, $b_k \in \mathbb{R}$ o *bias* do k -ésimo neurônio de saída, $\phi_{oculta}(\cdot): \mathbb{R} \rightarrow \mathbb{R}$ a função de ativação dos neurônios da camada oculta e $\phi_{saída}(\cdot): \mathbb{R} \rightarrow \mathbb{R}$ a função de ativação dos neurônios da camada de saída. Assim, o vetor

$\underline{w} \in \mathbb{R}^M$, $\underline{w} = [\underline{w}_s^t \quad \underline{w}_1^t \quad \dots \quad \underline{w}_j^t \quad b \quad b_1 \quad \dots \quad b_j]^t$ apresenta um total de $M = m(n + s + 1) + s$ parâmetros livres, onde s representa o número de saídas do *MLP* e m o número de neurônios na camada escondida.

Em geral, os *MLPs* utilizam nas camadas ocultas funções de ativação do tipo sigmoidal, apresentadas na seção 2.2.1.3. Na camada de saída, as funções são especificadas, geralmente, em função da natureza do trabalho. Para problemas de regressão, o tipo mais usual é a função de ativação linear por ser uma função que não limita os valores de saída. Já em problemas de classificação, são comumente utilizadas funções de ativação do tipo sigmoidal.

Conhecida a estrutura geral do *MLP*, são necessários algoritmos para especificação dos seus parâmetros, como definição das entradas e do número de neurônios na camada escondida, como também para estimação dos pesos e dos *bias*.

O número de neurônios a ser adotado em cada camada é, em geral, definido empiricamente, uma vez que depende de uma série de fatores ligados ao problema que está sendo estudado, como número de exemplos de treinamento, quantidade de ruídos presentes nos exemplos, complexidade da função a ser aprendida e outros. Em alguns casos, o uso de duas ou mais camadas ocultas pode facilitar o treinamento da rede neural. Segundo o teorema da aproximação universal, modelos com uma única camada oculta possuem capacidade de aproximar qualquer função contínua com precisão arbitrária. Visando reduzir o número de parâmetros a serem especificados, o uso de uma única camada oculta é suficiente.

Na presença de dados ruidosos, o elevado grau de não linearidade disponibilizado pelos *MLPs* permite que além da função geradora dos dados também sejam modelados traços específicos do conjunto de dados disponível, o que compromete o desempenho do modelo para novos padrões. Evitar o ajuste excessivo dos dados de treinamento (*overfitting*) constitui uma das principais tarefas no desenvolvimento de modelos neurais.

O *overfitting* ocorre quando a rede gera superfícies de decisão mais complexas que as necessárias para a solução do problema, memorizando os padrões de treinamento em vez de retirar dos dados as características gerais que possibilitam a generalização ou o reconhecimento de padrões ainda não visualizados. O processo inverso, ou seja, o *underfitting* ocorre quando o problema necessita de uma superfície de decisão mais complexa que a modelada pela rede. Nesse caso, a rede gasta mais tempo buscando uma solução ótima para o problema. Para evitar a ocorrência de *overfitting* e *underfitting* é necessário obter uma topologia de rede que tenha uma boa capacidade de generalização e que seja capaz de modelar os dados do conjunto de treinamento sem modelar o ruído contido neles.

O *overfitting* também pode ser evitado a partir da estimação do erro de generalização durante o processo de treinamento. Para isso, basta dividir os dados em dois conjuntos: treinamento e validação. O conjunto de treinamento deve ser utilizado para o ajuste dos pesos e o conjunto de validação para estimar a capacidade de generalização da rede durante o processo de aprendizagem. A ideia consiste em parar o treinamento da rede quando o erro de validação começar a subir, ou seja, quando a rede começar a modelar o ruído presente nos dados.

2.2.2 Previsão Probabilística

A previsão probabilística consiste em determinar a previsão de uma variável aleatória e indicar a incerteza associada à mesma. Esta pode ser expressa por medidas probabilísticas tais como a função densidade de probabilidade, a função de distribuição acumulada, um conjunto de quantis ou momentos da distribuição [25].

As abordagens podem ser feitas por métodos paramétricos e não paramétricos. As redes neurais artificiais estão inseridas nos métodos não paramétricos e são capazes de lidar com problemas de elevada dimensionalidade e modelar diferentes tipos de relações.

Em [26], são apresentadas três técnicas para cálculo de intervalos de confiança, a saber: saída de erro, reamostragem dos erros e regressão multilinear adaptada para redes neurais.

2.2.3 Treinamento Bayesiano Aplicado a MLPs

2.2.3.1 Classificação

A classificação de padrões consiste, basicamente, em classificar os padrões em um número de categorias ou classes. Considerando um problema de classificação de padrões entre C classes mutuamente exclusivas, seja o conjunto $D = \{X, Y\}$, tal que $X = \{\underline{x}_1, \dots, \underline{x}_N\}$ representa o conjunto de entrada, $Y = \{\underline{d}_1, \dots, \underline{d}_N\}$ representa o conjunto de saída, $\underline{x} \in \mathbb{R}^n$, $\underline{x} = [x_1, \dots, x_n]^t$, $\underline{d} \in [0,1]^C$ e, para os padrões pertencentes à classe C_j , $\underline{x}_i \in C_j$, $\underline{d}_{ij} = 1$, $\underline{d}_{ik} = 0$, $k = 1, \dots, C$, $k \neq j$ e $\sum_{k=1}^C \underline{d}_{ik} = 1$. Essas relações mostram que a saída desejada \underline{d}_i é um vetor binário de dimensão C formado por um único elemento ativo ($\underline{d}_{ij} = 1$) que se relaciona à classe C_j à qual o padrão \underline{x}_i pertence, sendo todos os demais elementos do vetor inativos ($\underline{d}_{ik} = 0$). Como as classes são mutuamente exclusivas, ou seja, cada padrão de entrada \underline{x}_i só pode pertencer a uma única classe, $\sum_{k=1}^C \underline{d}_{ik} = 1$. Neste contexto, o objetivo do treinamento do modelo sob a ótica *bayesiana* consiste na determinação do vetor de parâmetros $\underline{w} \in \mathbb{R}^M$ que maximize a probabilidade a posteriori, $p(\underline{w}|Y, X)$, de acordo com a expressão (2-17) [22]:

$$p(\underline{w}|Y, X) = \frac{p(Y|X, \underline{w}) p(\underline{w})}{p(Y|X)} \quad (2-17)$$

Na equação (2-17), o termo $p(Y|X) = \int p(Y|X, \underline{w}) p(\underline{w}) d\underline{w}$ é um fator de normalização, que garante que $\int p(\underline{w}|Y, X) d\underline{w} = 1$. A determinação da probabilidade a posteriori $p(\underline{w}|Y, X)$ do vetor \underline{w} fica condicionada, portanto, ao conhecimento da distribuição de probabilidade a priori $p(\underline{w})$ e da função de verossimilhança $p(Y|X, \underline{w})$.

No caso do conhecimento prévio sobre a solução ser vago ou inexistente, uma escolha razoável para a distribuição de probabilidade a priori $p(\underline{w})$, considerando os modelos que apresentam componentes do vetor \underline{w} com pequena magnitude e reproduzem mapeamentos suaves, pode ser concebida pela distribuição gaussiana com vetor média nulo e matriz de covariância $\alpha^{-1} \underline{I}$, onde $\alpha \in \mathbb{R}^+$ é o hiperparâmetro cujo valor é assumido, a princípio, constante e conhecido, e $\underline{I} \in \mathbb{R}^M \times \mathbb{R}^M$ é a matriz identidade, sendo essa distribuição dada pela expressão (2-18) [2] [22]:

$$p(\underline{w}) = \frac{1}{Z_{\underline{w}}(\alpha)} e^{-\left(\frac{\alpha}{2} \|\underline{w}\|^2\right)} \quad (2-18)$$

$$Z_{\underline{w}}(\alpha) = \left(\frac{2\pi}{\alpha}\right)^{\frac{M}{2}}$$

A escolha da distribuição gaussiana para $p(\underline{w})$, dada pela expressão (2-18), simplifica as análises subsequentes, dando origem a um algoritmo analítico e iterativo para o cálculo de \underline{w} [2]. O uso de distribuições que refletem completo desconhecimento sobre o comportamento de \underline{w} , apesar de viável, pode acarretar em métodos de solução dispendiosos, não abordados, inclusive, por esse trabalho [27].

A função de verossimilhança $p(Y|X, \underline{w})$ está relacionada com a probabilidade dos padrões pertencerem a cada uma das classes. Para a sua determinação, seja y_k a k -ésima saída do *MLP*, dada pela expressão (2-16), representando a probabilidade de um determinado padrão \underline{x}_i pertencer à classe C_k , conforme mostra a expressão (2-19):

$$p(\underline{x}_i \in C_k | \underline{x}_i) = f_k(\underline{x}_i, \underline{w}) \quad (2-19)$$

Para que todas C saídas do *MLP* representem uma medida de probabilidade, considerando que as classes são mutuamente exclusivas, $0 \leq y_k \leq 1$ e $\sum_{k=1}^C f_k(\underline{x}_i, \underline{w}) = 1$

$$p(\underline{d}_i | \underline{x}_i, \underline{w}) = \prod_{k=1}^C [f_k(\underline{x}_i, \underline{w})]^{d_{ik}} \quad (2-22)$$

Supondo que os padrões de saída são independentes e identicamente distribuídos, a probabilidade de ocorrência do conjunto de padrões de saída Y dado o conjunto de padrões de entrada X e o vetor de parâmetros \underline{w} , $p(Y|X, \underline{w})$, pode ser escrita conforme a expressão (2-23):

$$p(Y|X, \underline{w}) = \prod_{i=1}^N \prod_{k=1}^C [f_k(\underline{x}_i, \underline{w})]^{d_{ik}} \quad (2-23)$$

Aplicando as equações (2-18) e (2-23) na regra de Bayes, dada pela expressão (2-17), é possível calcular a probabilidade a posteriori $p(\underline{w}|X, Y)$ de \underline{w} , dado o conjunto de saídas desejadas Y e o conjunto de entrada X , de acordo com a expressão (2-24) [22]:

$$p(\underline{w}|X, Y) = \frac{1}{Z_S} e^{-S_C(\underline{w})}$$

$$Z_S = \int e^{-S_C(\underline{w})} d\underline{w} \quad (2-24)$$

$$S_C(\underline{w}) = \sum_{i=1}^N \sum_{k=1}^C d_{ik} \ln[f_k(\underline{x}_i, \underline{w})] + \frac{\alpha}{2} \sum_{j=1}^M w_j^2 = E_D^C(\underline{w}) + \frac{\alpha}{2} E_C(\underline{w})$$

Visto que Z_S é uma constante de normalização, para obtenção do conjunto mais provável de pesos \underline{w} é necessária a obtenção do vetor de pesos que minimiza o funcional $S_C(\underline{w})$ [22]. O funcional $S_C(\underline{w})$, dado pela equação (2-24), é composto por duas parcelas. A primeira parcela, $E_D^C(\underline{w})$, é o risco empírico, que representa a capacidade do modelo classificar corretamente os padrões disponíveis. A segunda parcela, $E_C(\underline{w})$, relacionada com a inserção de conhecimento prévio através da probabilidade a priori $p(\underline{w})$, na teoria da regularização representa um funcional regularizador $E_C[f(\underline{x}, \underline{w})]$, conhecido como decaimento de pesos (*weight decay*) [2]. A utilização deste funcional favorece modelos apresentando componentes do vetor \underline{w} com pequena magnitude, buscando gerar mapeamentos suaves [28]. Desta forma, maximizar a probabilidade a posteriori de \underline{w} é equivalente à aplicação do regularizador de decaimento de pesos.

A aplicação de inferência *bayesiana* foi limitada até aqui à estimativa do vetor de parâmetros \underline{w} , desconhecido à priori, supondo que o valor do hiperparâmetro α era

conhecido. Assim, como \underline{w} e α são desconhecidos, a probabilidade *a posterior* $p(\underline{w}|X, Y)$ deve ser obtida através da integração da probabilidade *a posteriori* $p(\underline{w}, \alpha|X, Y)$ sobre o espaço dos hiperparâmetros, ou seja:

$$p(\underline{w}|X, Y) = \int p(\underline{w}, \alpha|X, Y) d\alpha = \int p(\underline{w}|\alpha, X, Y) p(\alpha|X, Y) d\alpha \quad (2-25)$$

A estimativa do hiperparâmetro α pode ser obtida a partir do estudo proposto em [29], conhecido como aproximação da evidência [28]. Esse estudo parte do princípio de que a probabilidade a priori $p(\alpha|X, Y)$ apresenta pouca dispersão em torno do valor mais provável de α^* , permitindo a simplificação da equação da distribuição de probabilidade a posteriori $p(\underline{w}|X, Y)$, dada pela equação (2-25), como mostra a expressão (2-26):

$$p(\underline{w}|X, Y) \simeq \int p(\underline{w}|\alpha^*, X, Y) \int p(\alpha|X, Y) d\alpha = p(\underline{w}|\alpha^*, X, Y) \quad (2-26)$$

Para determinar os valores mais prováveis para os hiperparâmetros, a probabilidade a posteriori $p(\alpha|X, Y)$ deve ser maximizada através da regra de Bayes, de onde se obtém a expressão (2-27):

$$p(\alpha|X, Y) = \frac{p(X, Y|\alpha)p(\alpha)}{p(X, Y)} \quad (2-27)$$

Na equação (2-27), a distribuição de probabilidade $p(\alpha)$ mostra que é necessária a inserção de algum conhecimento prévio acerca do hiperparâmetro α . Entretanto, na maioria das vezes, o conhecimento que se tem acerca desses parâmetros é vago ou inexistente. Desta forma, a probabilidade de distribuição $p(\alpha)$ deve ser escolhida de modo que a distribuição seja insensível a valores específicos de α . Como $p(X, Y)$ independe dos hiperparâmetros, a maximização da probabilidade a posteriori $p(\alpha|X, Y)$ é obtida pela maximização da probabilidade $p(X, Y|\alpha)$. Visto que α está relacionado apenas com a probabilidade a priori do vetor \underline{w} , tem-se as seguintes expressões para $p(X, Y|\alpha)$ [2]:

$$p(X, Y|\alpha) = p(Y, X|\alpha) \int p(Y, X, \underline{w}|\alpha) d\underline{w} = \int p(Y|X, \underline{w}, \alpha) p(X, \underline{w}|\alpha) d\underline{w} \quad (2-28)$$

$$p(X, Y|\alpha) = \int p(Y|X, \underline{w}) p(\underline{w}) d\underline{w} \quad (2-29)$$

Aplicando as expressões (2-18) e (2-23) na expressão (2-29), obtém-se a expressão (2-30):

$$p(Y|\alpha) = \frac{Z_S(\alpha)}{Z_{\underline{w}}(\alpha)} \quad (2-30)$$

Na expressão (2-30), $Z_S(\alpha)$ é dado pela expressão (2-24) e $Z_{\underline{w}}(\alpha)$ é dado pela expressão (2-18). Considerando uma aproximação *gaussiana* da distribuição de probabilidade *a posteriori* $p(\underline{w}|X, Y)$ de \underline{w} , o funcional $Z_S(\alpha)$ passa a ser dado pela expressão (2-31) [22]:

$$Z_S(\alpha) = e^{-S_C(\underline{w}^*)} (2\pi)^{\frac{M}{2}} \left\{ \det [\alpha \underline{I}] \right\}^{-\frac{1}{2}} \quad (2-31)$$

Na equação (2-31), \underline{w}^* é o vetor de parâmetros que minimiza o funcional $S_C(\underline{w})$ dado pela expressão (2-24), $\underline{I} \in \mathbb{R}^M \times \mathbb{R}^M$ é a matriz identidade e $\det [\underline{Z}]$ é o operador que representa o determinante de \underline{Z} . O logaritmo neperiano da expressão (2-30), após a substituição da expressão $Z_{\underline{w}}(\alpha)$, dada pela expressão (2-18), é dado pela expressão (2-32) [22]:

$$\ln p(X, Y|\alpha) = -\frac{\alpha}{2} \sum_{j=1}^M (w_j^*)^2 - \frac{1}{2} \ln \left\{ \det [\alpha \underline{I}] \right\} + \frac{M}{2} \ln \alpha - \frac{N}{2} \ln(2\pi) \quad (2-32)$$

Para maximizar o logaritmo neperiano da evidência $p(X, Y|\alpha)$ em relação à α , é necessário obter uma estimativa para \underline{w}^* . No treinamento de *MLPs*, métodos de segunda ordem, como *Levenberg-Marquardt* [24], utilizam a cada iteração uma aproximação quadrática do funcional $S_C(\underline{w})$ em torno do ponto \underline{w}^* . A maximização de $\ln p(X, Y|\alpha)$ pode ser obtida, analogamente, por um método numérico realizado após a obtenção de uma estimativa para \underline{w}^* , considerando um dado valor de α , ou seja, após a convergência do método numérico utilizado para a minimização de $S_C(\underline{w})$, dado pela expressão (2-24), por meio das expressões (2-33) e (2-34) [2]:

$$\gamma(l+1) = M - \text{trace} \left\{ [\alpha(l) \underline{I}]^{-1} \right\} \quad (2-33)$$

$$\alpha(l+1) = \frac{\gamma(l+1)}{\sum_{j=1}^M [w_j(l)]^2} \quad (2-34)$$

Na equação (2-33), $\gamma(l+1)$ é o número efetivo de parâmetros e o operador $trace(Z)$ se refere ao traço da matriz \underline{Z} .

A escolha da probabilidade *a priori* $p(\underline{w})$ dada pela equação (2-18), que conduz ao funcional $S_c(\underline{w})$ apresentado na expressão (2-24), apresenta inconsistência em relação às propriedades de escalonamento dos mapeamentos a serem realizados pelos *MLPs*. Quando transformações lineares são aplicadas às entradas e/ou à saída, funcionais consistentes devem dar origem a modelos similares àqueles treinados com o conjunto original de dados, a menos de transformações aplicadas aos seus parâmetros. Esta propriedade garante que este tipo de funcional não favoreça de forma arbitrária um modelo em detrimento do outro, visto que ambos são equivalentes. Um funcional regularizador como o apresentado na equação (2-18), que aborda de maneira equânime todo o conjunto de parâmetros \underline{w} , não satisfaz esta propriedade [24], sendo necessária a escolha de outro tipo de funcional, ou seja, outra distribuição *a priori* $p(\underline{w})$ [2].

A escolha de uma única distribuição de probabilidade $p(\underline{w})$ para todo o conjunto de pesos pressupõe comportamento semelhante para todos os parâmetros, não considerando a função específica de cada um. Ao contrário desta premissa, é esperado que alguns grupos de pesos, como os que ligam as entradas à camada oculta, apresentem comportamento distinto dos grupos que ligam a camada oculta à saída. A escolha de diferentes distribuições *gaussianas* para a especificação da distribuição *a priori*, todas com vetor média nulo e diferindo somente nas matrizes de covariância para distintos grupos de pesos, de modo que reflita o conhecimento prévio sobre o problema, além de representar de maneira mais fidedigna o comportamento do problema, conduz a funcionais regularizadores consistentes, diferentemente da especificação dada pela expressão (2-24).

Para esta nova distribuição de probabilidade *a priori* $p(\underline{w})$, seja g o número de conjuntos nos quais os pesos e *bias* são agrupados, $\underline{w}_i \in \mathbb{R}^{M_i}$, $\underline{w}_i = [w_{i1} \ w_{i2} \ \dots \ w_{iM_i}]^t$, o vetor contendo os M_i elementos do i -ésimo grupo, e α_i o hiperparâmetro associado. A probabilidade *a priori* $p(\underline{w}_i)$ associada ao conjunto de pesos representado pelo vetor \underline{w}_i é dada pela expressão (2-35) [22]:

$$p(\underline{w}_i) = \frac{1}{\left(\frac{2\pi}{\alpha_i}\right)^{\frac{M_i}{2}}} e^{-\frac{1}{2}\alpha_i \|\underline{w}_i\|^2} \quad (2-35)$$

A distribuição a priori $p(\underline{w})$ para todo o conjunto de pesos $\underline{w} \in \mathbb{R}^M$, $\underline{w} = [w_1 \ w_2 \ \dots \ w_g]^t$, $M = \sum_{i=1}^g M_i$, considerando a independência entre os grupos de pesos, é dada pela expressão (2-36) [22]:

$$p(\underline{w}) = \prod_{i=1}^g p(\underline{w}_i) = \frac{1}{\prod_{i=1}^g \left(\frac{2\pi}{\alpha_i}\right)^{\frac{M_i}{2}}} e^{-\frac{1}{2}\sum_{i=1}^g \alpha_i \|\underline{w}_i\|^2} \quad (2-36)$$

Analogamente ao desenvolvimento anterior, aplicando as expressões (2-23) e (2-36) na regra de Bayes, dada pela expressão (2-17), a maximização da probabilidade a posteriori $p(\underline{w}|X, Y)$ é obtida através da minimização do funcional $S_c(\underline{w})$, dado pela expressão (2-37) [22]:

$$S_c(\underline{w}) = \sum_{i=1}^N \sum_{k=1}^C d_{ik} \ln[f_k(\underline{x}_i, \underline{w})] + \frac{1}{2} \sum_{i=1}^g \left(\alpha_i \sum_{j=1}^{M_i} w_{ij}^2 \right) \quad (2-37)$$

Sendo $\underline{\alpha} \in \mathbb{R}^g$, $\underline{\alpha} = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_g]^t$, maximizando o logaritmo neperiano da evidência $p(Y|\alpha)$, dado pela equação (2-32), em relação à α , é possível obter as seguintes expressões para estimativa dos hiperparâmetros:

$$\gamma_i(l+1) = M_i - \text{trace} \left\{ \left[\alpha(l) \underline{I}_{\alpha_i} \right]^{-1} \right\} \quad (2-38)$$

$$\alpha_i(l+1) = \frac{\gamma_i(l+1)}{\|\underline{w}_i(l)\|^2} \quad (2-39)$$

$$\gamma(l+1) = \sum_{i=1}^g \gamma_i(l+1) \quad (2-40)$$

Nas expressões acima, $\underline{I}_{\alpha_i} \in \mathbb{R}^{M_i} \times \mathbb{R}^{M_i}$ é a matriz identidade de dimensão M_i , com γ_i representando o número efetivo de parâmetros associado ao i -ésimo grupo de pesos \underline{w}_i e γ representando o número efetivo de parâmetros (pesos e *bias*) de todo o modelo.

O algoritmo de treinamento *bayesiano* de *MLPs* pode ser resumido da seguinte forma [2]:

1. Faça $l = 0$.
2. Inicialize o vetor de parâmetros $\underline{w}(l)$ e o vetor de hiperparâmetros $\alpha_i(l)$.
3. Utilizando alguma técnica de otimização (por exemplo, um algoritmo de retropropagação do erro), atualize o vetor de parâmetros $\underline{w}(l+1)$ através da minimização do funcional $S(\underline{w})$ dado pela expressão (2-37).
4. Se o critério de parada for atendido, encerre o algoritmo. Do contrário, vá para o passo 5.
5. Atualize os hiperparâmetros $\alpha_i(l+1)$ utilizando as expressões (2-38) e (2-39).
6. Faça $l = l + 1$ e retorne ao passo 3.

2.2.3.2 Regressão

De forma análoga à classificação, o objetivo do treinamento do *MLP* na regressão reside na determinação do vetor de parâmetros $\underline{w} \in \mathbb{R}^n$ que maximize a probabilidade a posteriori dada pela expressão (2-17). Definida uma estrutura para o *MLP*, dado o conjunto de padrões de entrada e saída $D = \{X, Y\}$, tal que $X = \{\underline{x}_1, \dots, \underline{x}_N\}$, $Y = \{d_1, \dots, d_N\}$, $\underline{x} \in \mathbb{R}^n$ e $\underline{x} = [x_1, \dots, x_n]^t$, utilizando a mesma especificação de probabilidade *a priori* para \underline{w} dada pela expressão (2-36), resta especificar a probabilidade de ocorrência do conjunto de saídas desejadas a partir das entradas e dos parâmetros do modelo, ou seja, $p(Y|X, \underline{w})$, para este novo contexto onde $d \in \mathbb{R}$. Para tanto, considere o modelo com ruído aditivo dado pela expressão (2-41) [22]:

$$d = F(\underline{x}) + \zeta \tag{2-41}$$

A saída desejada d , dada pela expressão (2-41), pode ser representada pela combinação de uma função determinística e desconhecida do vetor de entrada \underline{x} , $F(\underline{x}): \mathbb{R}^n \rightarrow \mathbb{R}$, com uma variável aleatória $\zeta \in \mathbb{R}$ que pode representar diversos fatores, tais como erros de medição, ruídos e dinâmicas não modeladas por $F(\underline{x})$. A função determinística e desconhecida $F(\underline{x})$ será modelada por meio do *MLP* cujo vetor de parâmetros \underline{w} deve ser estimado. Matematicamente,

$$d_k = f(\underline{x}_k, \underline{w}) + \zeta \quad (2-42)$$

Na expressão (2-42), $f(\underline{x}, \underline{w}): \mathbb{R}^n \rightarrow \mathbb{R}$ é a saída do *MLP* apresentado na Figura 11, dada pela expressão (2-43), em que $\phi_{saída}(z) = z$ e $\phi_{oculta}(z) = \phi_{hiper}(z)$:

$$f(\underline{x}, \underline{w}) = \phi_{saída} \left[\sum_{j=1}^m \bar{w}_{kj} \phi_{oculta} \left(\sum_{l=1}^n \bar{w}_{jl} x_l + b_j \right) + b_k \right] \quad (2-43)$$

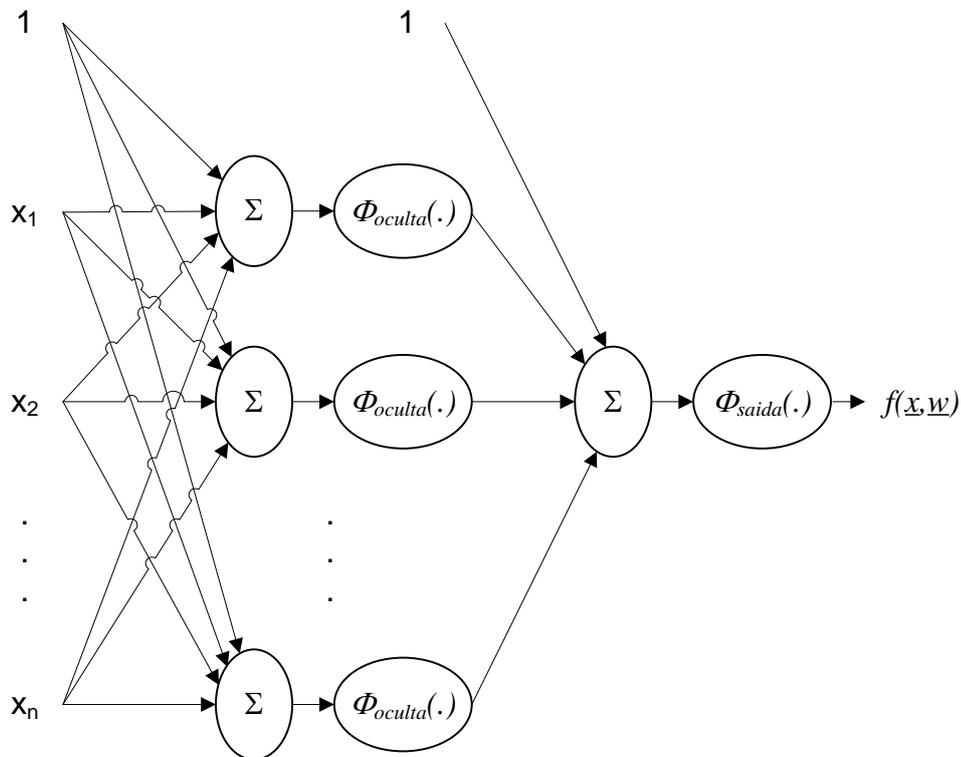


Figura 11: Estrutura do *MLP* utilizado para regressão [22]

A definição de $p(Y|X, \underline{w})$ depende da premissa adotada para a variável aleatória $\zeta \in \mathbb{R}$. Admitindo que ζ possua distribuição gaussiana com média nula e variância β^{-1} , a probabilidade de ocorrência de uma saída específica d_k , dado o vetor de entrada \underline{x}_k e o vetor de parâmetros \underline{w} , é dada pela expressão (2-44):

$$p(d_k | \underline{x}_k, \underline{w}) = \frac{e^{\left\{ -\frac{\beta}{2} [d_k - f(\underline{x}_k, \underline{w})]^2 \right\}}}{\int e^{\left\{ -\frac{\beta}{2} [d_k - f(\underline{x}_k, \underline{w})]^2 \right\}} dd_k} \quad (2-44)$$

Partindo do pressuposto que os padrões do conjunto de treinamento foram obtidos de maneira independente a partir desta distribuição, a probabilidade de ocorrência do conjunto de saídas desejadas a partir das entradas e dos parâmetros do modelo, ou seja, $p(Y|X, \underline{w})$, também conhecida como função de verossimilhança, pode ser estabelecida pela expressão (2-45) [2], em que $\beta \in \mathbb{R}^+$ é um hiperparâmetro que, neste ponto, é entendido como sendo conhecido.

$$p(Y|X, \underline{w}) = \frac{1}{Z_Y(\beta)} e^{\left\{-\frac{\beta}{2} \sum_{k=1}^N [d_k - f(\underline{x}_k, \underline{w})]^2\right\}} \quad (2-45)$$

$$Z_Y(\beta) = \left(\frac{2\pi}{\beta}\right)^{\frac{N}{2}}$$

Utilizando as expressões (2-36) e (2-45) na regra de Bayes dada pela expressão (2-17), a probabilidade a *posteriori* de \underline{w} pode ser escrita, resultando nas expressões (2-46) e (2-47) [2] [22]:

$$p(\underline{w}|X, Y) = \frac{1}{Z_S} e^{-S_r(\underline{w})} \quad (2-46)$$

$$Z_S = \int e^{-S_r(\underline{w})} d\underline{w}$$

$$S_r(\underline{w}) = \frac{\beta}{2} \sum_{k=1}^N [d_k - f(\underline{x}_k, \underline{w})]^2 + \frac{1}{2} \sum_{i=1}^g \alpha_i \left(\sum_{j=1}^M w_{ij}^2 \right) = \frac{\beta}{2} E_D^r(\underline{w}) + \frac{1}{2} E_C(\underline{\alpha}, \underline{w}) \quad (2-47)$$

Analogamente ao problema de classificação, o valor mais provável para \underline{w} é obtido pela minimização do funcional $S_r(\underline{w})$ dado pela expressão (2-47). Nesta expressão, $E_D^r(\underline{w})$ representa o ajuste do modelo aos dados disponíveis por meio da soma do erro quadrático e $E_C(\underline{\alpha}, \underline{w})$ controla a complexidade do modelo estimado, evitando o ajuste excessivo dos dados (evitando o *overfitting*).

Os hiperparâmetros β e $\underline{\alpha} \in \mathbb{R}^g$ também devem ser estimados a partir dos dados. Mantendo a analogia para o problema de classificação, a maximização da evidência para os hiperparâmetros dá origem a um processo iterativo que deve ser acoplado ao método numérico utilizado para estimação de \underline{w} . O algoritmo de treinamento *bayesiano* de *MLPs* para problemas de regressão pode então ser resumido da seguinte forma [2]:

1. Faça $l = 0$.
2. Inicialize o vetor de parâmetros $\underline{w}(l)$ e os hiperparâmetros $\underline{\alpha}(l)$ e $\underline{\beta}(l)$.
3. Utilizando alguma técnica de otimização, atualize o vetor de parâmetros $\underline{w}(l + 1)$ através da minimização do funcional $S_r(\underline{w})$ dado pela expressão (2-47).
4. Se o critério de parada for atendido, encerre o algoritmo. Do contrário, vá para o passo 5.
5. Atualize os hiperparâmetros $\alpha_i(l + 1)$ e $\beta(l + 1)$ utilizando as expressões:

$$\underline{A}_r(\underline{w})\Big|_{\underline{w}=\underline{w}(l)} = \beta(l) \underline{H}(\underline{w})\Big|_{\underline{w}=\underline{w}(l)} + \underline{\alpha}(l) \underline{I} \quad (2-48)$$

$$\underline{B}_i(l) = \left[\underline{A}_r(\underline{w})\Big|_{\underline{w}=\underline{w}(l)} \right]^{-1} \underline{I}_i \quad (2-49)$$

$$\gamma_i(l + 1) = M_i - \text{trace} \{ \underline{B}_i(l) \} \quad (2-50)$$

$$\alpha_i(l + 1) = \frac{\gamma_i(l + 1)}{\|\underline{w}_i(l)\|^2} \quad (2-51)$$

$$\gamma(l + 1) = \sum_{i=1}^g \gamma_i(l + 1) \quad (2-52)$$

$$\beta(l + 1) = \frac{N - \gamma(l + 1)}{\sum_{k=1}^N \{d_k - f[\underline{x}_k, \underline{w}(l)]\}^2} \quad (2-53)$$

onde $\underline{H}(\underline{w})\Big|_{\underline{w}=\underline{w}(l)}$ representa a matriz *hessiana* com as segundas derivadas de $E_D^r(\underline{w})$, calculada por meio de retropropagação do erro para a iteração corrente ($\underline{w} = \underline{w}(l)$) [28]; e $\underline{I}_i \in \mathbb{R}^M \times \mathbb{R}^M$ sendo uma matriz diagonal nula, com elementos iguais a um apenas para as componentes do vetor \underline{w} pertencentes ao i -ésimo grupo de pesos.

6. Faça $l = l + 1$ e retorne ao passo 3.

A expressão (2-42) em conjunto com a especificação da função de verossimilhança gaussiana, dada pela expressão (2-45), mostra que a saída $f(\underline{x}_k, \underline{w}): \mathbb{R}^n \rightarrow \mathbb{R}$ do MLP apresentado na Figura 11 representa o valor esperado da saída desejada \underline{d}_k relacionada ao padrão de entrada \underline{x}_k . A variância σ_k^2 associada pode ser obtida pela expressão [28]:

$$\sigma_k^2 = \frac{1}{\beta^*} + \left[\underline{g} \Big|_{\underline{w}=\underline{w}^*} \right]^t \left[\underline{H} \Big|_{\underline{w}=\underline{w}^*} \right]^{-1} \underline{g} \Big|_{\underline{w}=\underline{w}^*} \quad (2-54)$$

Na expressão (2-54), β^* representa a variância da distribuição a priori obtida ao final do algoritmo de treinamento, $\underline{g} \Big|_{\underline{w}=\underline{w}^*}$ o vetor gradiente e $\underline{H} \Big|_{\underline{w}=\underline{w}^*}$ a matriz hessiana do funcional $S_r(\underline{w})$ dado pela expressão (2-47), com o vetor gradiente e a matriz hessiana calculados para o vetor de pesos \underline{w}^* obtido ao final do algoritmo de treinamento. A variância σ_k^2 da expressão (2-54) pode ser utilizada em conjunto com a saída $f(\underline{x}_k, \underline{w}): \mathbb{R}^n \rightarrow \mathbb{R}$ do MLP da Figura 11, relacionada com o valor esperado para a saída desejada \underline{d}_k , para a construção dos intervalos de confiança para as estimativas.

2.2.4 Modelos de Clusterização

2.2.4.1 Clusterização

A *clusterização* é uma técnica de mineração de dados que consiste em organizar os dados de um conjunto em grupos (*clusters*), a partir da maior similaridade existente entre os dados de um mesmo grupo que os de outro, com base em algum critério predeterminado [30]. A *clusterização* é uma técnica na qual nenhuma suposição é feita a respeito dos grupos. Portanto, não conta com classes predefinidas e exemplos de treinamento de classes rotuladas, realizando uma forma de aprendizado não supervisionado [31].

A principal vantagem desse método é que ele permite descrever de forma mais eficiente as características particulares de cada grupo identificado, apresentando um panorama detalhado acerca dos dados. Isso permite desenvolver novos esquemas de classificação e descobrir correlações entre os atributos dos dados que até então eram desconhecidas.

Seja $X = \{X_1, X_2, \dots, X_n\}$ o conjunto de n objetos, onde $X_i \in \mathbb{R}^p$ é um vetor de p medidas reais que dimensionam as características do objeto. Supondo que os dados do conjunto X devem ser clusterizados em k grupos distintos $C = \{C_1, C_2, \dots, C_k\}$, as seguintes condições devem ser respeitadas [32]:

$$a) C_1 \cup C_2 \cup \dots \cup C_k = X$$

- b) $C_i \neq \emptyset, \forall i, 1 \leq i \leq k$
 c) $C_i \cap C_j = \emptyset, \forall i \neq j, 1 \leq i \leq k, 1 \leq j \leq k$

Das condições acima apresentadas, é possível afirmar que cada um dos objetos disponíveis deve pertencer a um único grupo e que cada grupo deve ter pelo menos um objeto.

Em [33], são citados alguns critérios que devem ser observados quando da aplicação de qualquer método de *clusterização*. Os três mais importantes deles, segundo [30], são apresentados a seguir: definição da forma de representação do conjunto de dados de entrada a serem agrupados, definição de uma medida adequada de semelhança entre os dados e definição da técnica de *clusterização* a ser utilizada para a construção dos clusters.

2.2.4.1.1 Representação do conjunto de dados de entrada

Na equação (2-55), a matriz $X_{n \times p}$ representa a estrutura de dados comumente utilizada para armazenar os padrões de entrada. Nessa matriz, as linhas representam cada um dos objetos a serem clusterizados e as colunas representam os atributos ou características de cada objeto.

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{np} \end{bmatrix} \quad (2-55)$$

As características de um padrão podem ser qualitativas ou quantitativas, e devem ser alocadas em função dos seguintes tipos de variáveis: binárias, nominais, ordinais, escaladas em intervalos, escaladas em proporção e outros.

Apesar de não existir uma regra a ser seguida que permita definir os padrões e as características mais apropriadas para uma dada situação, uma escolha equivocada da forma de representação de cada padrão pode levar à construção de um agrupamento complexo, de difícil visualização.

2.2.4.1.2 Medida de similaridade

Os relacionamentos entre os objetos podem ser descritos por uma matriz contendo uma medida de dissimilaridade ou de proximidade entre cada par de objetos [34]. Na matriz de dissimilaridade D , definida pela equação (2-56), cada elemento da matriz representa a distância entre pares de objetos.

$$D = \begin{bmatrix} 0 & d(1,2) & d(1,3) & \dots & d(1,p) \\ d(2,1) & 0 & d(2,3) & \dots & d(2,p) \\ d(3,1) & d(3,2) & 0 & \dots & d(3,p) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d(n,1) & d(n,2) & d(n,3) & \dots & 0 \end{bmatrix} \quad (2-56)$$

Uma vez que um conjunto de dados pode ser interpretado como um conjunto de pontos em um espaço k -dimensional, o conceito de similaridade pode ser traduzido como a distância entre dois pontos [35]. Assim, quanto menor a distância $d(i,j)$, maior a similaridade, e os objetos i e j ficam no mesmo grupo. Por outro lado, quanto maior a distância $d(i,j)$, menor a similaridade, e os referidos objetos ficam em grupos distintos.

Visto que a similaridade é dada pela distância entre dois pontos, quando um atributo assume valores em um intervalo relativamente grande, sobrepujando os valores dos demais atributos, o resultado da *clusterização* pode ser prejudicado. Por exemplo: se uma aplicação tem dois atributos A e B, e A varia entre 1 e 1000, e B varia entre 1 e 10, então o atributo A exercerá a sua influência sobre o atributo B. Uma forma de contornar esse problema consiste na normalização dos padrões de entrada.

Há diversas métricas de similaridade na literatura, a saber: Minkowski, Manhattan, Hamming, Euclidiana, Mahalanobis e outras. As expressões (2-57), (2-58) e (2-59) apresentam, respectivamente, as distâncias de Minkowski, Hamming e Euclidiana:

$$d(i,j) = \left(\sum_{l=1}^p |x_{il} - x_{jl}|^q \right)^{1/q} \quad (2-57)$$

$$d(i,j) = \sum_{l=1}^p |x_{il} - x_{jl}| \quad (2-58)$$

$$d(i,j) = \sqrt{\sum_{l=1}^p |x_{il} - x_{jl}|^2} \quad (2-59)$$

Uma função de distância deve ser definida de tal forma que obedeça as seguintes propriedades:

- a) $D(x,x) = 0$
- b) $D(x,y) = D(y,x)$

Esta propriedade afirma que a distância entre dois pontos não varia, independente do ponto a partir do qual ela é medida. Tem-se, assim, uma matriz simétrica.

$$c) D(x, y) \leq D(x, z) + D(z, y)$$

Esta propriedade especifica que a menor distância entre dois pontos é uma reta.

2.2.4.2 K-Means

O *k-means* é um dos algoritmos de aprendizagem não supervisionada mais utilizado em problemas de agrupamento. A ideia desse algoritmo consiste em classificar um determinado conjunto de dados a partir dos seus próprios valores, separando-os em k grupos distintos, definidos inicialmente à priori. Para isso, ele busca minimizar a distância dos pontos a um conjunto de k centros de forma iterativa.

Seja $C = \{C_1, C_2, \dots, C_k\}$ o conjunto de k centros ou centróides. A distância entre um ponto p_i e um conjunto de clusters, dada por $d(p_i, C)$, é definida como sendo a distância do ponto ao centro mais próximo dele [34]. A função a ser minimizada é dada por:

$$d(p_i, C) = \frac{1}{n} \sum_{i=1}^n d(p_i, C)^2 \quad (2-60)$$

No *k-means*, primeiramente, devem ser escolhidos, aleatoriamente, os k centróides dos grupos. Em seguida, cada ponto do conjunto de dados deve ser associado ao centróide mais próximo. O valor do centróide de cada grupo deve ser calculado pelo valor médio dos pontos associados ao referido grupo. O processo termina quando os centróides param de se modificar, ou após um número limitado de iterações especificado pelo usuário.

O algoritmo *k-means*, uma vez escolhido k , divide um conjunto de n objetos em k grupos, tal que a similaridade *intracluster* seja alta e a similaridade *intercluster* seja baixa. O *k-means* tenta determinar k partições que minimizem a função do erro quadrado [35].

Segundo [31], o *k-means* é um algoritmo de fácil implementação, simplicidade, eficiência e sucesso empírico. É eficiente, geralmente convergindo em poucas iterações para uma configuração estável, na qual nenhum elemento está designado para um *cluster* cujo centro não lhe seja o mais próximo. Esse algoritmo apresenta bom desempenho quando os clusters são densos e compactos e bem separados uns dos outros [32].

O algoritmo *k-means* tende a convergir para um mínimo da função definida. Essa condição enfatiza a questão da homogeneidade e ignora a importante questão da boa separação dos grupos. Isto pode causar uma má separação dos conjuntos no caso de uma má inicialização dos centros. Outro ponto que pode afetar a qualidade dos resultados é a escolha do número de grupos. Um número pequeno demais pode causar a união de dois ou mais grupos, e um número grande demais pode provocar a segmentação de um grupo em mais grupos distintos [34].

O *k-means* é sensível a ruídos, visto que um pequeno número de dados ruidosos pode influenciar os valores médios dos grupos. Além disso, não é adequado para grupos com formas convexas ou grupos de tamanhos muito diferentes.

O algoritmo de *clusterização k-means* pode ser resumido da seguinte forma:

1. Escolher k centróides aleatoriamente
2. Associar cada ponto ao centróide mais próximo
3. Recalcular os novos centróides de cada grupo

O centróide de cada grupo é calculado pela distância média dos pontos associados a cada grupo.

4. Repetir os passos 3-4 até a convergência

Atualmente, há muitas variações disponíveis do método *k-means*. Em geral, as diferenças estão na seleção das k médias iniciais, no cálculo da similaridade ou na estratégia para calcular a média dos clusters [35]. O modelo *x-means* apresenta uma proposta que permite calcular o valor dos k centróides automaticamente.

2.2.4.3 X-Means

O *x-means* [36] apresenta uma proposta que permite solucionar um dos grandes problemas do tradicional modelo de agrupamento *k-means*, que consiste na necessidade de especificar o número dos k clusters a serem agrupados.

No *x-means*, o usuário deve especificar o número máximo de clusters k_{max} possível para o conjunto de dados, para que o algoritmo encontre, de forma eficiente, o melhor valor de k dentro do intervalo especificado.

O algoritmo consiste nas duas seguintes operações repetidas até a conclusão.

1. Melhorar os parâmetros:
Consiste no uso do *k-means* convencional para convergência.
2. Melhorar a estrutura:
Descobre se e onde os novos centróides devem aparecer.
3. Se $k > k_{max}$ pare e relate o modelo de melhor pontuação encontrado; senão vá para 1.

Percorrendo um loop de 1 a k_{max} , a ideia consiste em determinar os k centroides. A operação de melhoria da estrutura começa dividindo cada centróide em dois filhos, que são movidos a uma distância proporcional ao tamanho da região em direções opostas ao longo de um vetor escolhido aleatoriamente. Em seguida, é executado um *k-means* local ($k = 2$) para cada par de filhos. Este é o local em que os filhos estão lutando entre si pelos pontos na região dos pais. Neste ponto, um teste de seleção do modelo é realizado em todos os pares de filhos. Em cada caso, o teste pergunta se há evidências de que os dois filhos estão modelando a estrutura real ou se o pai original modelará a distribuição igualmente bem. De

acordo com o resultado do teste, o pai ou seus filhos são mortos. A esperança é que os centroides que já possuem um conjunto de pontos que formam um *cluster* na verdadeira distribuição subjacente não serão modificados por esse processo (ou seja, eles vão sobreviver a seus filhos). Por outro lado, as regiões do espaço que não são bem representadas pelos centroides atuais receberão mais atenção aumentando o número de centroides neles.

Considere as seguintes informações para a abordagem do critério de seleção de modelos, em que i é o índice do centróide mais próximo para o i -ésimo ponto de dados, $\mu_{(i)}$ é o centróide associado ao i -ésimo ponto durante a iteração, D é o conjunto de dados de entrada, $D_i \subseteq D$ é o conjunto de pontos que tem $\mu_{(i)}$ como centróide, $R = |D|$, $R_i = |D_i|$, M é a dimensão e a matriz de covariância gaussiana dada por $\sum \text{diag}(\sigma^2)$.

O Critério de Informação Bayesiano (*BIC*, em inglês) [36] ou Critério Bayesiano de Schwarz é um critério para a seleção de um modelo entre um conjunto finito de modelos, baseado na verossimilhança, e é dado pela seguinte equação:

$$BIC(M_j) = \hat{l}_j(D) - \frac{p_j}{2} \cdot \log R \quad (2-61)$$

Na equação (2-61), M_j é o conjunto de modelos com diferentes valores de k ou o número de atributos; $\hat{l}_j(D)$ é a probabilidade de registro dos dados de acordo com o j -ésimo modelo e tomado no ponto de máxima verossimilhança, p_j é o número de parâmetros em M_j e R é o número de pontos.

A estimativa de máxima verossimilhança para a variância, assumindo uma gaussiana esférica idêntica, pode ser obtida pela equação (2-62). As probabilidades pontuais podem ser obtidas pela equação (2-63) e a probabilidade de registro dos dados pode ser obtida pela equação (2-64).

$$\hat{\sigma}^2 = \frac{1}{R - K} \sum_i (x_i - \mu_{(i)})^2 \quad (2-62)$$

$$\hat{P}(x_i) = \frac{R_{(i)}}{R} \cdot \frac{1}{\sqrt{2\pi\hat{\sigma}^M}} e^{-\frac{1}{2\hat{\sigma}^2} \|x_i - \mu_{(i)}\|^2} \quad (2-63)$$

$$l(D) = \log \prod_i P(x_i) = \sum_i \left(\log \frac{1}{\sqrt{2\pi\hat{\sigma}^M}} - \frac{1}{2\hat{\sigma}^2} \|x_i - \mu_{(i)}\|^2 + \log \frac{R_{(i)}}{R} \right) \quad (2-64)$$

Seja $1 \leq n \leq k$. Considerando apenas o conjunto de pontos D_n que pertence ao centróide n e as estimativas de máxima verossimilhança, tem-se:

$$\hat{l}(D_n) = -\frac{R_n}{R} \log(2\pi) - \frac{R_n \cdot M}{2} \log(\hat{\sigma}^2) - \frac{R_n - K}{2} + R_n \log R_n - R_n \log R \quad (2-65)$$

2.2.4.4 Mistura de Gaussianas

O modelo de mistura de gaussianas, em inglês *Gaussian mixture model (GMM)*, é um modelo probabilístico, de aprendizagem não supervisionada, que consiste em representar subpopulações normalmente distribuídas dentro de uma população geral.

A distribuição de mistura de gaussianas pode ser escrita como uma superposição linear de gaussianas da seguinte forma [24]:

$$p(x) = \sum_{k=1}^K \pi_k \mathcal{N}(x | \mu_k, \Sigma_k) \quad (2-66)$$

Na equação (2-66), $\pi_k \in [0,1]$, com $k \in (1,2, \dots, K)$ e $\sum_{k=1}^K \pi_k = 1$. Esta formulação pode ser utilizada em problemas onde a densidade de probabilidade de x não segue uma distribuição normal. Os diversos parâmetros listados na equação (2-66) podem ser estimados utilizando técnicas estatísticas como máxima verossimilhança [24]. Neste trabalho, estes parâmetros serão estimados por modelos neurais conforme será apresentado no capítulo 3.

3 METODOLOGIA PROPOSTA

Para a metodologia proposta são apresentados dois modelos neurais probabilísticos para previsão de energia solar. O primeiro modelo, chamado de modelo Único, tem esse nome por apresentar apenas um único previsor. O segundo modelo, modelo GMM, é assim chamado por ser um modelo de misturas de gaussianas.

Os dois modelos a serem apresentados utilizam basicamente a mesma metodologia para especificação e treinamento das redes neurais envolvidas. Trata-se de uma rede de arquitetura do tipo perceptron de múltiplas camadas, composta por uma única camada oculta, cujo processo de aprendizagem ocorre de maneira supervisionada utilizando o treinamento bayesiano apresentado na seção 2.2.3. Foi utilizado o ambiente MATLAB para as simulações.

3.1 Modelo Único

No modelo Único, os dados do conjunto de entrada-saída não são clusterizados. Os parâmetros da rede foram especificados da seguinte forma: 25 neurônios na camada oculta e função de saída linear para os modelos de regressão (seção 2.2.3.2). Visto que o treinamento bayesiano controla automaticamente a complexidade do modelo, o número de neurônios deve ser especificado de forma a evitar o *underfitting* visto que o problema do *overfitting* é abordado diretamente pelo algoritmo de treinamento. Esta estratégia de previsão pode ser resumida pelos passos a seguir:

1. Leitura dos dados de entrada

O conjunto de dados é dividido, de acordo com as regras da competição, em dados de treinamento e dados de previsão.

2. Filtragem dos dados

Nessa etapa é feita uma filtragem nos dados de treinamento em função dos parâmetros de radiação solar, sendo retirados os dados cuja radiação é nula. Em seguida, as entradas constantes também são retiradas do conjunto de entradas.

3. Etapa de treinamento

Antes de serem apresentados à rede neural, os dados de treinamento são normalizados no intervalo entre $[-1,1]$. Em seguida, é iniciado o treinamento bayesiano, com base no modelo de regressão apresentado, usando uma rede probabilística que considera que a variável de saída segue uma distribuição gaussiana, cuja média e desvio padrão será projetada pela rede do tipo *MLP*.

4. Etapa de previsão

Os dados de previsão são normalizados e apresentados, um a um, à rede neural probabilística com base no modelo de regressão. Ao final do processo, são obtidos o valor estimado e a variância estimada desnormalizados para essa entrada.

5. Cálculo dos percentis

Para cada par estimado – média μ e variância σ^2 previstas – devem ser calculados 99 percentis. O primeiro passo consiste em determinar o intervalo de confiança da gaussiana, obtido em função do par estimado. Esse intervalo deve ser dividido em 100 partes iguais. Os percentis 0 e 100 não devem ser calculados, apenas os 99 interiores. Cada percentil é calculado através da gaussiana inversa em função do par estimado. Por exemplo, o percentil 98 (p_{98}) é dado por:

$$p_{98} = \mu + 2,06\sigma \quad (3-1)$$

3.2 Modelo GMM

No modelo GMM, inicialmente a série de produção de energia solar fotovoltaica é clusterizada utilizando o método x-means. Esta abordagem se fez necessária uma vez que esta variável não segue uma distribuição normal, conforme será apresentado no capítulo 4. O uso do método x-means se deveu pela sua característica automática, sendo necessária somente a definição do número máximo de clusters. Neste trabalho, esse valor foi definido como 24, ou seja, supondo no limite a definição de um grupo para cada hora do dia.

Uma vez clusterizada a série de produção de energia solar fotovoltaica, o valor de energia é transformado em probabilidade de pertencer a um dado grupo. Exemplificando, considere que a série tenha sido clusterizada em 5 grupos. Suponha ainda que o registro de produção de 25% da capacidade do parque tenha sido atribuído ao *cluster* 1. Assim, a saída desejada para este padrão, que no modelo único seria 0,25, para o classificador será substituída por um vetor de cinco posições (número de clusters) representando a

probabilidade do padrão pertencer a cada um dos clusters. Para o exemplo em questão, a saída seria [1 0 0 0], ou seja, o padrão tem 100% de chance de pertencer ao *cluster* 1.

O modelo de classificação (seção 2.2.3.1) então passa a ter o papel de estimação dos parâmetros $\pi_k \in (1, 2, \dots, K)$ da expressão (2-66). Os parâmetros (média e variância) de cada uma das gaussianas apresentadas nesta expressão são previstos por modelos de regressão. Para tanto, os padrões entrada-saída de cada *cluster* são isolados, sendo estimados modelos neurais para regressão (seção 2.2.3.2) para cada cluster, tendo como saída desejada a produção do parque solar.

Os parâmetros da rede foram especificados da seguinte forma: 25 neurônios na camada oculta, função de saída linear para os modelos de regressão, limites para a saída da rede entre 0 e 100% e número máximo de centros para agrupamento igual a 24. A metodologia pode ser resumida da forma que segue:

1. Leitura dos dados de entrada

O conjunto de dados é dividido, de acordo com as regras da competição, em dados de treinamento e dados de previsão.

2. Filtragem dos dados

Nessa etapa é feita uma filtragem nos dados de treinamento em função dos dois parâmetros de radiação solar, sendo retirados os dados cuja radiação é nula. Em seguida, as entradas constantes também são retiradas do conjunto de entradas.

3. Etapa de *clusterização*

A *clusterização* é feita sobre os dados de saída do treinamento. Para cada número de centros definidos, será criado um agrupamento com base na distância euclidiana usando o *k-means* e calculado o índice *BIC* do respectivo agrupamento usando o *x-means*. O *BIC* de maior valor retorna o melhor agrupamento.

4. Etapa de treinamento

Nesta etapa, os dados de treinamento são separados por classes de acordo com o agrupamento obtido através dos modelos *k-means/x-means*. Cada classe é treinada a partir do modelo de regressão. A classificação do agrupamento, obtida através dos modelos *k-means/x-means*, é transformada em probabilidade através da codificação 1-de-N para ser treinada a partir do modelo de classificação.

5. Etapa de previsão

Inicialmente, para cada padrão de previsão, é prevista a probabilidade desse padrão pertencer a cada uma das classes, usando, para isso, o modelo de classificação. A seguir, para esse mesmo padrão, são calculados a média

prevista e a variância prevista a partir do modelo de regressão. A previsão média é obtida a partir do produto matricial entre as probabilidades das classes e a média prevista para cada cluster.

6. Cálculo dos percentis

Neste modelo, os percentis são calculados em função de uma mistura de gaussianas, a partir do par de vetores saída estimada e variância estimada. O intervalo de confiança é obtido a partir do par de vetores apresentado, observando-se os limites inferior e superior para o intervalo. O intervalo de confiança deve ser dividido em 100 partes iguais. Os percentis 0 e 100 não devem ser calculados, apenas os 99 interiores. Cada percentil é calculado através da gaussiana inversa em função do par de vetores, da média e da variância.

3.3 Critérios de Análise

Os resultados das simulações serão avaliados de acordo com as métricas apresentadas nesta seção. Para a classificação do modelo na competição, a precisão das séries geradas pelas redes neurais será avaliada a partir da função de perdas de *pinball*.

3.3.1 Erro Relativo Percentual Médio

O erro relativo E_p calcula a relação entre a diferença do valor medido z_M e o valor estimado z_E por um valor base z_B , de acordo com a equação (3-2):

$$E_p = \frac{1}{N} \sum_{i=1}^N \frac{|z_{M_i} - z_{E_i}|}{z_B} \times 100 \quad (3-2)$$

3.3.2 Erro Relativo

O erro relativo percentual permite a comparação entre um valor aproximado V e um valor exato V_E , de acordo com a equação (3-3):

$$E_R = \frac{|V - V_E|}{V_E} \quad (3-3)$$

3.3.3 Média ponderada

A média ponderada consiste na relação entre a soma dos valores x_i influenciados pelos seus respectivos pesos p_i , divididos pela soma dos pesos p_i , de acordo com a seguinte expressão:

$$\bar{X}_p = \frac{\sum_{i=1}^k x_i p_i}{\sum_{i=1}^k p_i} \quad (3-4)$$

3.3.4 Função de Perda de *Pinball*

A função de perda de *pinball* retorna um valor que serve para avaliar a precisão de um modelo de previsão probabilística de quantis. Quanto menor for a perda de *pinball*, mais precisa será a previsão de quantis.

Pelas regras da competição *GEFCom 2014* [3] (base de dados utilizada neste trabalho), as precisões das séries geradas serão avaliadas com base nas previsões probabilísticas de 99 percentis. Os percentis 0 e 100 são considerados, respectivamente, $-\infty$ e $+\infty$.

Os quantis são pontos que dividem a amostra em intervalos regulares a partir da função de distribuição acumulada de uma variável aleatória. A divisão da amostra em 100 quantis é chamada de percentis [37]. A função de perda de *pinball* é dada pela seguinte equação [3]:

$$L(q_a, y) = \begin{cases} \left(1 - \frac{a}{100}\right)(q_a - y), & \text{se } y < q_a \\ \frac{a}{100}(y - q_a), & \text{se } y \geq q_a \end{cases} \quad (3-5)$$

Na equação (3-5), q_a é o quantil estimado, y é o valor real e a é o índice do percentil no intervalo entre [1,99].

4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Nesta seção são abordadas as análises e os resultados das simulações executadas para o modelo Único e para o modelo GMM.

4.1 Dados da Competição

A base de dados utilizada para a elaboração deste trabalho foi disponibilizada na *Global Energy Forecasting Competition 2014* [5] (*GEFCom 2014*), uma competição de previsão de demanda de energia.

A base de dados é formada por dados meteorológicos e de produção de energia solar de três localidades distintas da Austrália, identificadas apenas por zona 1, zona 2 e zona 3. A única informação que se tem sobre as zonas 1, 2 e 3 é a altitude, que são de, respectivamente, 595 m, 602 m e 951 m [3]. Informações acerca dos sistemas solares fotovoltaicos instalados em cada zona também são fornecidas, e serão apresentadas adiante.

Na competição, são dadas 15 tarefas, que correspondem às previsões da produção de energia solar a serem realizadas para os últimos 15 meses de medição dos dados. A tarefa 1 consiste na previsão da produção de energia a ser realizada para o mês de abril de 2013 para as três zonas, a tarefa 2 para o mês de maio de 2013 para as três zonas, e, assim, sucessivamente. Todavia, de acordo com as regras da competição, para a classificação dos resultados são consideradas apenas as previsões realizadas para as 12 últimas tarefas.

4.1.1 Base de dados medida

Os dados em análise, apresentados em valores horários, foram medidos de 01 de abril de 2012 a 30 de junho de 2014, totalizando 27 meses de medição ou 19.704 horas de medição por zona [3].

A série de dados em análise é composta por doze variáveis identificadas numericamente, conforme mostra a Tabela 1.

Tabela 1: Variáveis medidas

Variável	Número de Identificação
Coluna total de água líquida	78
Coluna total de água sólida	79
Pressão superficial	134
Umidade relativa a 1000 mbar	157
Cobertura total da nuvem	164
Componente u do vento a 10 metros	165
Componente v do vento a 10 metros	166
Temperatura do ar a 2 metros de altura	167
Radiação solar na superfície	169
Radiação térmica na superfície	175
Radiação solar líquida no topo da atmosfera	178
Precipitação total	228

A série traz ainda duas variáveis não numeradas: data e hora. A variável hora foi modelada em função do seno de modo a representar a curva de radiação solar para um dia claro.

4.1.2 Sistemas fotovoltaicos

Em cada uma das zonas 1, 2 e 3, há um sistema solar fotovoltaico instalado, cuja configuração é distinta das dos demais sistemas.

O sistema fotovoltaico instalado na zona 1 é composto por 8 painéis solares da marca Solarfun, modelo SF 160-24-1M195, de potência nominal de 195 watts de pico, totalizando 1.560 watts de pico de potência instalada. O arranjo da zona 1 encontra-se inclinado em 36° e orientado em 38° para o norte geográfico no sentido horário.

O sistema fotovoltaico instalado na zona 2 é composto por 26 painéis solares da marca Suntech, modelo STP190S-24/Ad+, de potência nominal de 190 watts de pico. Totalizando 4.940 watts de pico de potência instalada, esse arranjo encontra-se inclinado em 35° e orientado em 327° para o norte geográfico no sentido horário.

O arranjo do sistema fotovoltaico instalado na zona 3 é composto por 20 módulos solares da marca Suntech, modelo STP200-18/ud, de potência nominal de 200 watts de pico, com potência instalada de 4.000 watts de pico. O arranjo instalado na zona 3 apresenta inclinação de 21° e orientação de 31° para o norte geográfico no sentido horário.

4.1.3 Classificação

Conforme dito anteriormente, na competição, a classificação é feita com base nos resultados apresentados para as 12 últimas tarefas, sendo cada tarefa composta pelas previsões da produção de energia para as zonas 1, 2 e 3. A precisão das séries geradas

pelas redes neurais será avaliada a partir da função de perdas de *pinball*, cuja formulação matemática foi apresentada no item 3.3. O resultado a ser apresentado em cada tarefa corresponde ao valor médio do *pinball* apurado para as três zonas [3].

A *GEFCom* apresenta uma metodologia para a classificação final da competição. É importante destacar que essa metodologia é necessária para verificar a classificação final dos modelos discutidos neste trabalho em relação aos modelos apresentados pelas equipes na competição. Essa metodologia permite também pontuar as entradas faltantes e erradas. Porém, a *GEFCom* não destaca em quais condições isso ocorre. Toda a classificação é feita seguindo os valores de referência (*benchmark*) disponibilizados pela *GEFCom*.

A metodologia apresentada é constituída pelos seguintes passos:

Passo 0: Começa com as pontuações (*pinball*) coletadas ao longo da competição, destacando as entradas faltantes em azul e as entradas erradas em amarelo.

Passo 1: É criada uma matriz de classificação para cada entrada válida na tarefa correspondente. Ou seja, os modelos são classificados em função dos valores de referência (*benchmark*). Para cada equipe (modelo), é calculada a segunda pior posição entre as entradas válidas. As pontuações das entradas azul e amarela são permutadas usando a sua segunda pior posição. Além disso, são aplicadas as seguintes regras:

- (1) Para uma tarefa determinada, se a segunda pior posição for maior do que o número de entradas válidas, sua classificação permutada será igual ao número de entradas válidas (entradas dos participantes na tarefa determinada).
- (2) Se todas as entradas válidas correspondentes a segunda pior posição desta equipe estiverem acima do *benchmark* (forem melhores que o *benchmark*), sua classificação permutada será igual ou melhor que a posição do *benchmark*.
- (3) Se algumas das entradas válidas correspondentes a sua segunda pior posição são piores do que o *benchmark*, as entradas faltantes e erradas são tratadas de formas diferentes. Para uma entrada azul, sua classificação será igual a segunda pior posição, podendo estar abaixo do *benchmark*. Para uma entrada amarela, sua classificação permutada será igual ou melhor que a posição do *benchmark*.

Passo 2: Com as posições permutadas, pode-se permutar as pontuações. Uma pontuação permutada para a equipe *i* na tarefa *j* com classificação permutada *r* é a mesma que a pontuação de uma equipe com posição *r* na tarefa *j*.

Passo 3: A classificação é definida para cada entrada como a porcentagem pela qual ela supera o *benchmark*. Para dar preferência às equipes que melhoram suas metodologias ao longo do caminho, foi atribuído um peso linearmente crescente as 12 tarefas de avaliação. A última tarefa, em relação à primeira tarefa, é ponderada em 12 vezes. Esse método de ponderação reduz o impacto de entradas faltantes e erradas para as primeiras tarefas. Para

cada tarefa, foi calculado o erro relativo entre o valor do *pinball* calculado e o valor do *pinball* do *benchmark*. A classificação da equipe é a soma ponderada de suas classificações ao longo das 12 tarefas. Esta classificação mostra, por alto, o quanto uma equipe melhora no *benchmark*.

4.2 Seleção da base de dados

Para a execução de todas as simulações, foi adotada a base de dados disponibilizada pela *GEFCom 2014*, sendo utilizados os dados das 12 variáveis de entrada existentes apenas para os períodos relacionados à incidência de radiação solar. Ou seja, todos os dados relacionados aos períodos em que a radiação solar é nula foram descartados, não sendo apresentados à rede neural, permitindo melhorar o desempenho e reduzir o tempo de processamento da rede.

Em estudos anteriores, percebeu-se que a informação referente a hora atua positivamente ajudando a rede neural na previsão da produção de energia solar. Desta forma, esse parâmetro foi modelado em função do seno de modo a representar a curva de radiação solar para um dia claro, a partir da hora solar ($0 \leq h \leq 23$), de acordo com a seguinte equação:

$$H(h) = \begin{cases} \text{sen} \left(2\pi \frac{(h-5)}{12} \right), & 7 \leq h \leq 18 \\ 0, & h < 7 \text{ ou } h > 18 \end{cases} \quad (4-1)$$

Os dados de entrada e de saída, apresentados à rede neural, foram normalizados no intervalo entre [-1,1], em função dos limites mínimos e máximos identificados em cada conjunto de dados, de acordo com a seguinte equação:

$$y = \frac{(y_{max} - y_{min}) \times (x - x_{min})}{(x_{max} - x_{min})} + y_{min} \quad (4-2)$$

Na equação (4-2), y_{max} é igual a 1, y_{min} é igual a -1, x é o valor a ser convertido, x_{min} é o menor valor da série a ser convertida e x_{max} é o maior valor da série a ser convertida.

4.3 Resultados

Nos subitens a seguir são apresentados os resultados e as análises realizadas para o modelo Único e para o modelo GMM.

4.3.1 Modelo Único

Seguindo as regras da competição, foram executadas as simulações correspondentes para a obtenção dos valores do *pinball* para as 12 tarefas solicitadas. Na Tabela 2 são apresentados os resultados obtidos para o modelo Único, sendo apresentados também, para efeito de comparação, os valores do *benchmark*. Comparando esses dados, é possível observar que o modelo Único apresenta resultados satisfatórios, com valores muito abaixo dos valores do *benchmark*, obtendo êxito e conquistando a 14ª colocação na competição, conforme mostrado no Apêndice.

Tabela 2: Resultados do *pinball* – Modelo Único

Tarefas	<i>Pinball</i>				
	<i>Benchmark</i>	Modelo Único Zonas 1, 2 e 3	Modelo Único Zona 1	Modelo Único Zona 2	Modelo Único Zona 3
1	0,0331	0,0156	0,0164	0,0164	0,0141
2	0,0388	0,0208	0,0164	0,0201	0,0260
3	0,0359	0,0175	0,0163	0,0184	0,0176
4	0,0361	0,0197	0,0185	0,0220	0,0185
5	0,0479	0,0194	0,0179	0,0218	0,0184
6	0,0357	0,0193	0,0155	0,0252	0,0171
7	0,0421	0,0197	0,0155	0,0281	0,0155
8	0,0399	0,0165	0,0154	0,0189	0,0150
9	0,0435	0,0180	0,0173	0,0191	0,0175
10	0,0377	0,0158	0,0144	0,0175	0,0153
11	0,0320	0,0117	0,0113	0,0131	0,0106
12	0,0285	0,0155	0,0158	0,0157	0,0150

O Gráfico 1 mostra a curva do *pinball* dos três primeiros colocados, do modelo Único e do *benchmark*. Analisando o gráfico, é possível observar que as curvas do *pinball* das três equipes mais bem colocadas apresentam, praticamente, o mesmo perfil, sendo possível afirmar que os valores do *pinball* estão muito próximos. Mostra ainda que a curva do modelo Único apresenta perfil semelhante ao dos melhores colocados em oito tarefas, evidenciando que a rede neural probabilística do modelo Único está operando adequadamente.

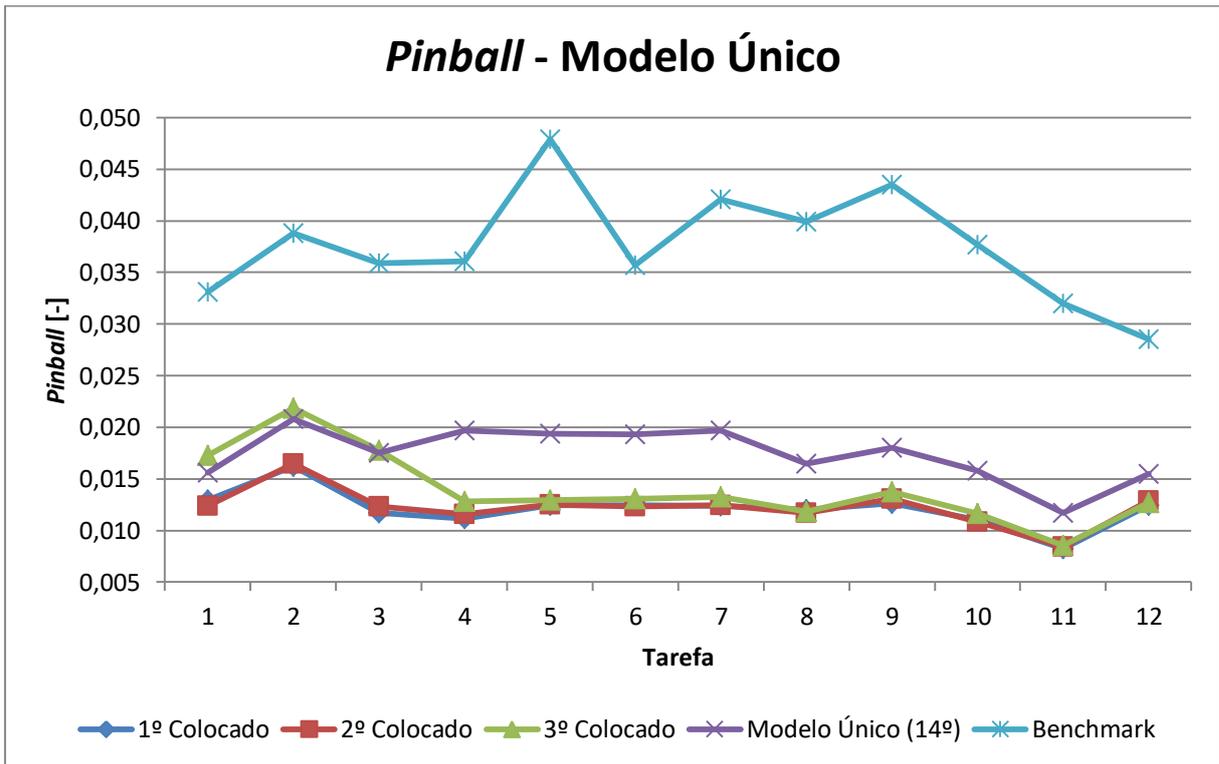


Gráfico 1: Curvas do *pinball* dos três primeiros colocados, do modelo Único e do *benchmark*

O Gráfico 2 apresenta as classificações dos três primeiros colocados, do modelo Único e do *benchmark* durante a competição. É possível observar que as três melhores equipes ocupam as três primeiras posições na maior parte do tempo, apresentando um perfil de curva praticamente constante. O modelo Único, 14º colocado na classificação final, fica oscilando entre a 10ª e a 17ª posição ao longo da competição. Observe que há muitas equipes entre o modelo Único e o terceiro colocado. Porém, conforme mostrado no Gráfico 1, além da curva do modelo Único acompanhar o perfil da curva das três melhores equipes, a diferença entre os valores do *pinball* dessas equipes para o modelo Único é muito pequena, permitindo afirmar que as equipes apresentam, estatisticamente, o mesmo resultado.

O Gráfico 3 comprova que a rede neural está operando adequadamente, pois os perfis das curvas do *pinball* para as três zonas, na maioria dos casos, são semelhantes, ou seja, apresentam um comportamento padrão, variando apenas a taxa dos erros. É possível afirmar também que, no geral, os sistemas solares fotovoltaicos e as condições a que estão submetidos apresentam, praticamente, o mesmo comportamento.

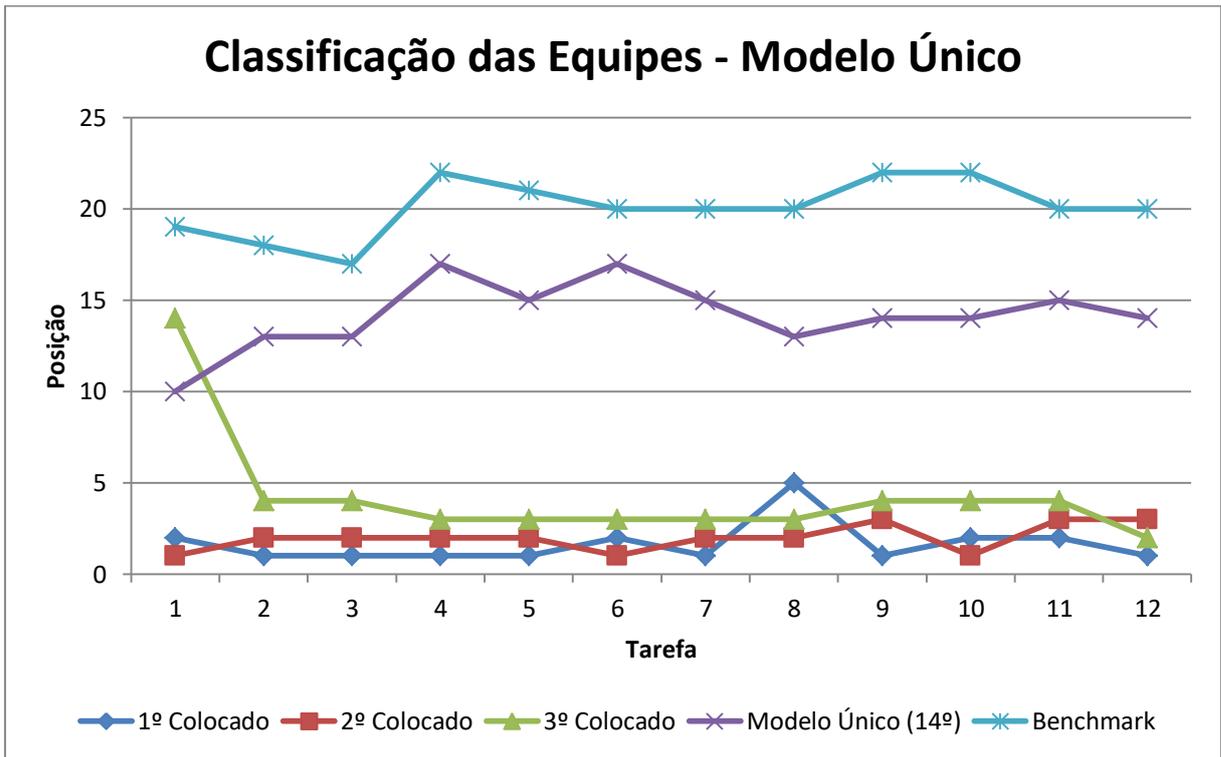


Gráfico 2: Classificação dos três primeiros colocados, do modelo Único e do *benchmark*

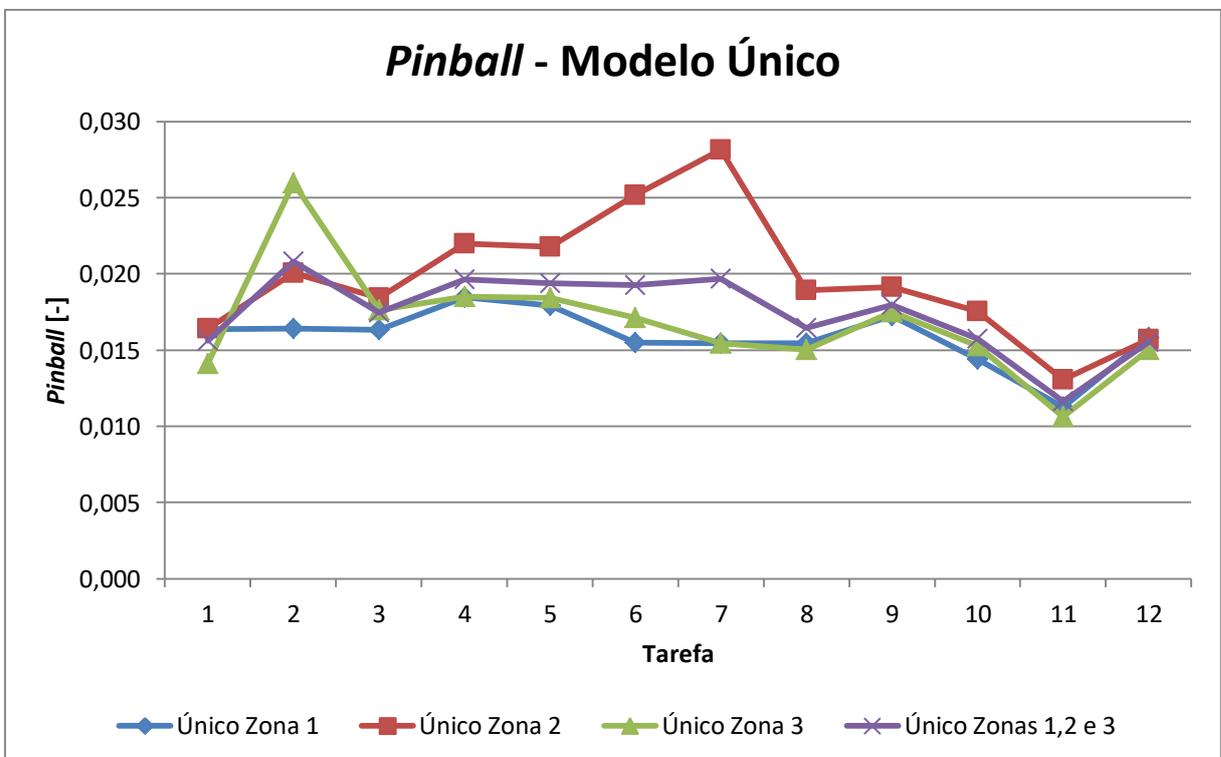


Gráfico 3: Curvas do *pinball* – Modelo Único

Uma análise mais aprofundada dos resultados das simulações é apresentada na Tabela 3. No caso em análise, o baixo erro relativo de previsão mostra que a rede está conseguindo estimar com precisão a grande maioria das previsões. De modo equivalente, o baixo valor do erro sistemático médio mostra que a série estimada apresenta pouca tendência para a subestimação ou superestimação dos resultados. O erro relativo foi usado para calcular o erro entre o montante de energia medido e o montante de energia estimado, sem levar em consideração a taxa de assertividade da rede neural a cada previsão. O erro relativo mais elevado nos meses de junho, julho e agosto pode ser decorrente de uma série de fatores. Nesse período, em decorrência da declinação solar, os raios do Sol incidem mais inclinadamente na superfície terrestre, fornecendo menos energia radiante, provocando o inverno. Assim, esse erro poderia estar associado a uma grande instabilidade momentânea não representada nos dados utilizados para o treinamento da rede neural. Analisando o Gráfico 1, é possível constatar que os modelos das equipes selecionadas apresentam o mesmo comportamento para a tarefa 2, referente ao mês de agosto, mostrando que, a princípio, o problema não é de modelagem. Apesar de ser de extrema importância averiguar a origem do erro, não há elementos suficientemente disponíveis para realizar essa análise, uma vez que as zonas 1, 2 e 3 são desconhecidas.

Tabela 3: Resultados das simulações – Zonas 1, 2 e 3 – Modelo Único

Tarefa	Mês	Energia Total Medida [kWh]	Energia Total Estimada [kWh]	Erro Relativo [%]	Erro Relativo de Previsão [%]	Erro Sistemático Médio [kWh]
1	Julho	275	307	11,64	3,79	1,43
2	Agosto	348	390	12,08	4,99	1,88
3	Setembro	424	410	3,14	4,38	-0,62
4	Outubro	537	534	0,43	4,89	-0,10
5	Novembro	505	504	0,19	4,71	-0,04
6	Dezembro	545	534	1,90	4,56	-0,46
7	Janeiro	514	539	4,82	4,69	1,11
8	Fevereiro	417	422	1,18	3,81	0,24
9	Março	379	378	0,17	4,36	-0,03
10	Abril	325	334	2,80	3,69	0,42
11	Maior	330	320	2,83	2,74	-0,42
12	Junho	217	244	12,62	3,86	1,27

A Tabela 4 mostra os resultados das simulações separadamente para as zonas 1, 2 e 3 para o mês de junho. Nesse mês, o erro relativo é mais acentuado para as zonas 1 e 3. Observe que a zona 2 apresenta erro relativo da ordem de 6%, a zona 1, da ordem de 19%,

e a zona 3, da ordem de 14%. Todos os demais erros são baixos (inferiores a 5%) e comprovam o bom funcionamento da rede neural.

Tabela 4: Resultados das simulações – Tarefa 12 (junho) – Modelo Único

Zona	Energia Medida [kWh]	Energia Estimada [kWh]	Erro Relativo [%]	Erro Relativo de Previsão [%]	Erro Sistemático Médio [kWh]
1	65	78	18,75	3,84	1,70
2	73	77	5,93	4,00	0,60
3	79	89	13,72	3,75	1,50

A Tabela 5 apresenta os resultados separadamente para as três zonas para o mês de fevereiro. Observa-se que o erro relativo em todas as zonas é baixo e equilibrado.

Tabela 5: Resultados das simulações – Tarefa 8 (fevereiro) – Modelo Único

Zona	Energia Medida [kWh]	Energia Estimada [kWh]	Erro Relativo [%]	Erro Relativo de Previsão [%]	Erro Sistemático Médio [kWh]
1	125	130	4,25	3,29	0,79
2	150	141	5,51	4,63	-1,23
3	142	150	5,53	3,52	1,17

4.3.2 Modelo GMM

O modelo GMM considera a aplicação de um conjunto de gaussianas para classificar os dados a serem apresentados para a rede neural probabilística. Assim, seguindo as regras da competição, as simulações foram executadas para as 12 tarefas solicitadas.

A Tabela 6 apresenta os resultados do *pinball* para o modelo GMM e para o *benchmark*. Como se esperava, o modelo GMM também apresenta resultados satisfatórios para o *pinball*, com valores muito abaixo dos valores do *benchmark*, apresentando, inclusive, uma sutil melhora em relação ao modelo Único, da ordem de 10%, aproximadamente.

Tabela 6: Resultados do *Pinball* – Modelo GMM

Tarefas	<i>Pinball</i>				
	<i>Benchmark</i>	Modelo GMM Zonas 1, 2 e 3	Modelo GMM Zona 1	Modelo GMM Zona 2	Modelo GMM Zona 3
1	0,0331	0,0140	0,0145	0,0159	0,0116
2	0,0388	0,0186	0,0141	0,0180	0,0239
3	0,0359	0,0158	0,0144	0,0176	0,0154
4	0,0361	0,0174	0,0166	0,0194	0,0162
5	0,0479	0,0168	0,0157	0,0181	0,0166
6	0,0357	0,0176	0,0168	0,0209	0,0152
7	0,0421	0,0165	0,0128	0,0238	0,0129
8	0,0399	0,0140	0,0131	0,0151	0,0140
9	0,0435	0,0157	0,0145	0,0163	0,0163
10	0,0377	0,0131	0,0123	0,0141	0,0130
11	0,0320	0,0107	0,0104	0,0107	0,0109
12	0,0285	0,0142	0,0132	0,0143	0,0150

Com essa melhora, o modelo GMM logrou êxito na classificação, conquistando a 9ª posição na competição. Com isso, o modelo Único cai uma posição, passando de 14º para 15º colocado. As tabelas com os valores do *pinball* de todos os participantes e as classificações podem ser consultadas no Apêndice.

O Gráfico 4 mostra a curva do *pinball* dos três primeiros colocados, do modelo GMM e do *benchmark*. A análise do gráfico permite confirmar a melhora da taxa de erro através da aproximação da curva do modelo GMM em relação aos participantes mais bem colocados, mostrando também que nas três primeiras tarefas os resultados do modelo GMM superaram os resultados do terceiro colocado. A Tabela 7 compara os resultados do *pinball* do modelo Único com o modelo GMM, mostrando a redução obtida com o uso deste último modelo. Em alguns casos, a redução chega a 17%. Na média, fica em torno de 12%.

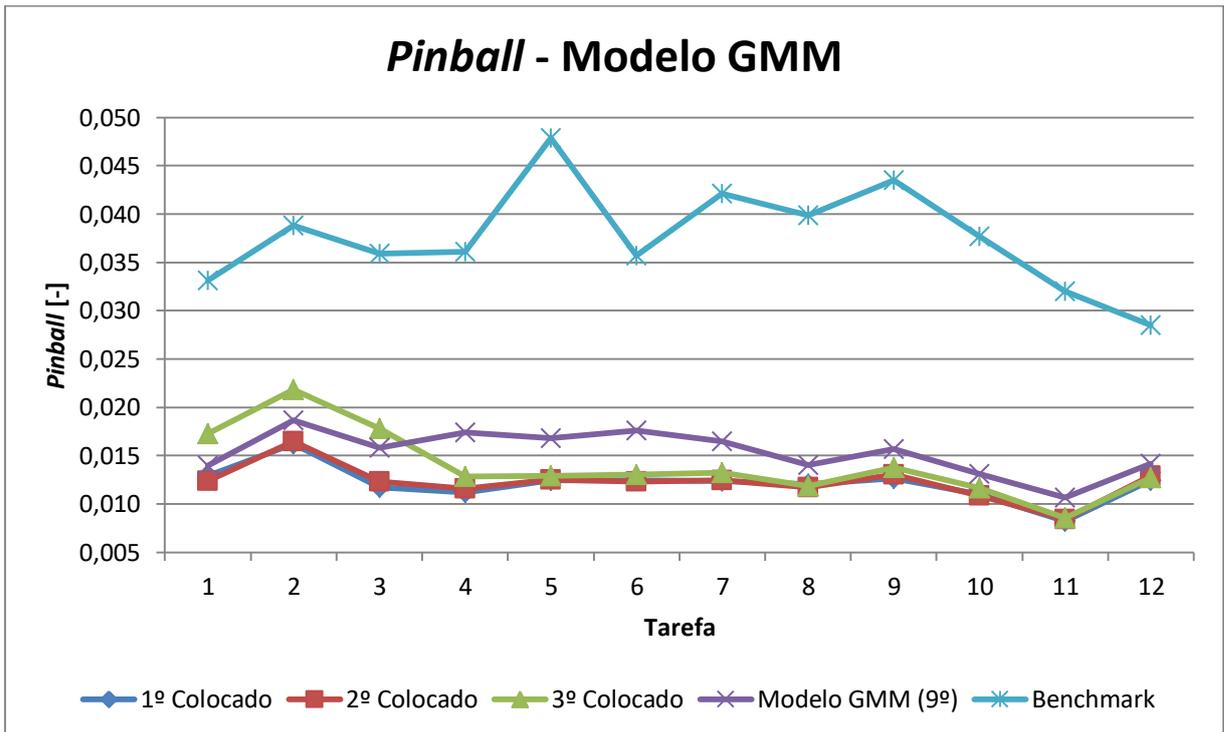


Gráfico 4: Curvas do *pinball* dos três primeiros colocados, do modelo GMM e do *benchmark*

Tabela 7: Redução do *pinball* alcançada com o modelo GMM

Tarefa	Mês	<i>Pinball</i>		
		Modelo Único	Modelo GMM	Redução
1	Julho	0,0156	0,0140	10%
2	Agosto	0,0208	0,0186	10%
3	Setembro	0,0175	0,0158	10%
4	Outubro	0,0197	0,0174	12%
5	Novembro	0,0194	0,0168	13%
6	Dezembro	0,0193	0,0176	9%
7	Janeiro	0,0197	0,0165	16%
8	Fevereiro	0,0165	0,0140	15%
9	Março	0,0180	0,0157	13%
10	Abril	0,0158	0,0131	17%
11	Maio	0,0117	0,0107	9%
12	Junho	0,0155	0,0142	9%

O Gráfico 5 apresenta as classificações das equipes durante a competição. É possível observar que as três melhores equipes ocupam as três primeiras posições na maior parte do tempo, apresentando um perfil de curva praticamente constante. O modelo GMM, 9º colocado na classificação final, fica oscilando entre a 5ª e a 15ª posição ao longo da competição. Observe que há muitas equipes entre o modelo GMM e o terceiro colocado. Porém, conforme mostrado no Gráfico 4, além da curva do modelo GMM acompanhar o perfil da curva das três melhores equipes, a diferença entre os valores do *pinball* dessas

equipes para o modelo GMM é muito pequena, permitindo afirmar que as equipes apresentam, estatisticamente, o mesmo resultado.

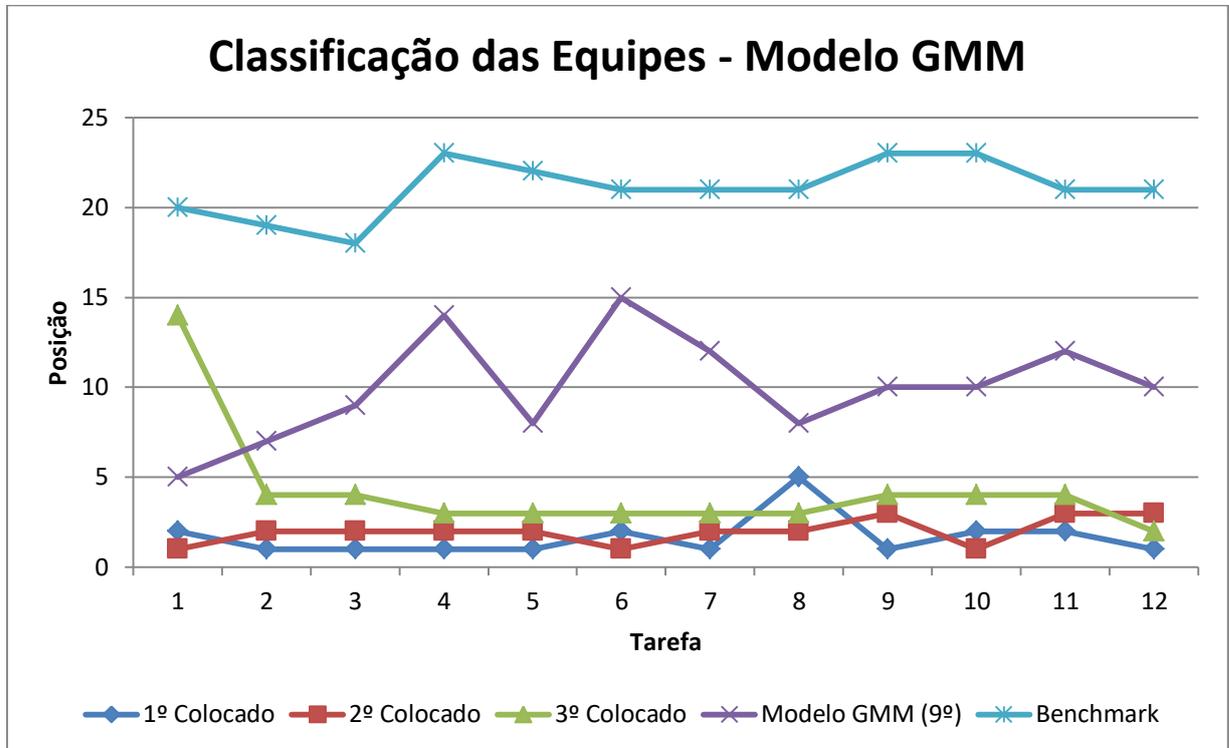


Gráfico 5: Classificação dos três primeiros colocados, do modelo GMM e do *benchmark*

A análise aplicada ao modelo Único é agora aplicada ao modelo GMM, e os resultados obtidos são apresentados na Tabela 8. Com relação ao erro sistemático médio, que verifica se a rede neural está gerando séries subestimadas ou superestimadas, o resultado não teve alterações significativas, mostrando que a rede neural modelada está funcionando bem nesse aspecto. O baixo erro relativo de previsão mostra que o modelo está conseguindo estimar com precisão a grande maioria das previsões. Comparando com o erro relativo de previsão do modelo Único, é possível observar uma melhora nas previsões em, praticamente, todos os meses, exceto no mês de maio. Para o erro relativo, observa-se, mais uma vez, que os piores resultados ocorrem para os meses de junho, julho e agosto. Analisando o Gráfico 4, é possível constatar que os modelos das equipes selecionadas apresentam o mesmo comportamento para a tarefa 2, referente ao mês de agosto, mostrando que, a princípio, o problema não é de modelagem. Como não há informações suficientes que permitam averiguar o porquê desse erro, não será possível realizar essa análise.

Tabela 8: Resultados das simulações – Zonas 1, 2 e 3 – Modelo GMM

Tarefa	Zona	Energia Medida [kWh]	Energia Estimada [kWh]	Erro Relativo [%]	Erro Relativo de Previsão [%]	Erro Sistemático Médio [kWh]
1	Julho	275	304	10,83	3,60	1,33
2	Agosto	348	386	10,94	4,93	1,71
3	Setembro	424	410	3,25	4,33	-0,64
4	Outubro	537	536	0,14	4,59	-0,03
5	Novembro	505	510	1,14	4,55	0,27
6	Dezembro	545	537	1,47	4,37	-0,36
7	Janeiro	514	539	4,77	4,17	1,10
8	Fevereiro	417	420	0,80	3,67	0,17
9	Março	379	377	0,37	4,21	-0,06
10	Abril	325	331	1,86	3,50	0,28
11	Maior	330	327	0,94	2,79	-0,14
12	Junho	217	246	13,70	3,75	1,38

Analisando isoladamente os resultados das zonas 1, 2 e 3 para a tarefa 12, apresentados na Tabela 9, é possível verificar mais uma vez que o erro relativo das zonas 1 e 3 impactam o resultado final da referida tarefa. Fazendo a mesma análise para os resultados da tarefa 8, apresentados na Tabela 10, é possível confirmar que os resultados desta tarefa atestam a eficiência do modelo GMM.

Tabela 9: Resultados das simulações – Tarefa 12 (junho) – Modelo GMM

Zona	Energia Medida [kWh]	Energia Estimada [kWh]	Erro Relativo [%]	Erro Relativo de Previsão [%]	Erro Sistemático Médio [kWh]
1	65	78	19,81	3,59	1,80
2	73	77	5,36	3,82	0,54
3	79	91	16,36	3,85	1,79

Tabela 10: Resultados das simulações – Tarefa 8 (fevereiro) – Modelo GMM

Zona	Energia Medida [kWh]	Energia Estimada [kWh]	Erro Relativo [%]	Erro Relativo de Previsão [%]	Erro Sistemático Médio [kWh]
1	125	127	1,68	3,18	0,31
2	150	146	2,72	4,33	-0,61
3	142	148	3,73	3,50	0,79

Os gráficos a seguir mostram as previsões de energia solar realizadas. As curvas pontilhadas mostram os percentis p_1 e p_{99} dos modelos Único (cor vermelha) e GMM (cor verde). As curvas sólidas mostram o valor verificado (valor real medido, cor preta) e os valores previstos pelos modelos Único (cor vermelha) e GMM (cor verde). O Gráfico 6

mostra o resultado da previsão realizada para a zona 3, para três dias do mês de setembro. Considerando a previsão para um dia claro, é possível constatar que a rede neural probabilística apresenta bons resultados.

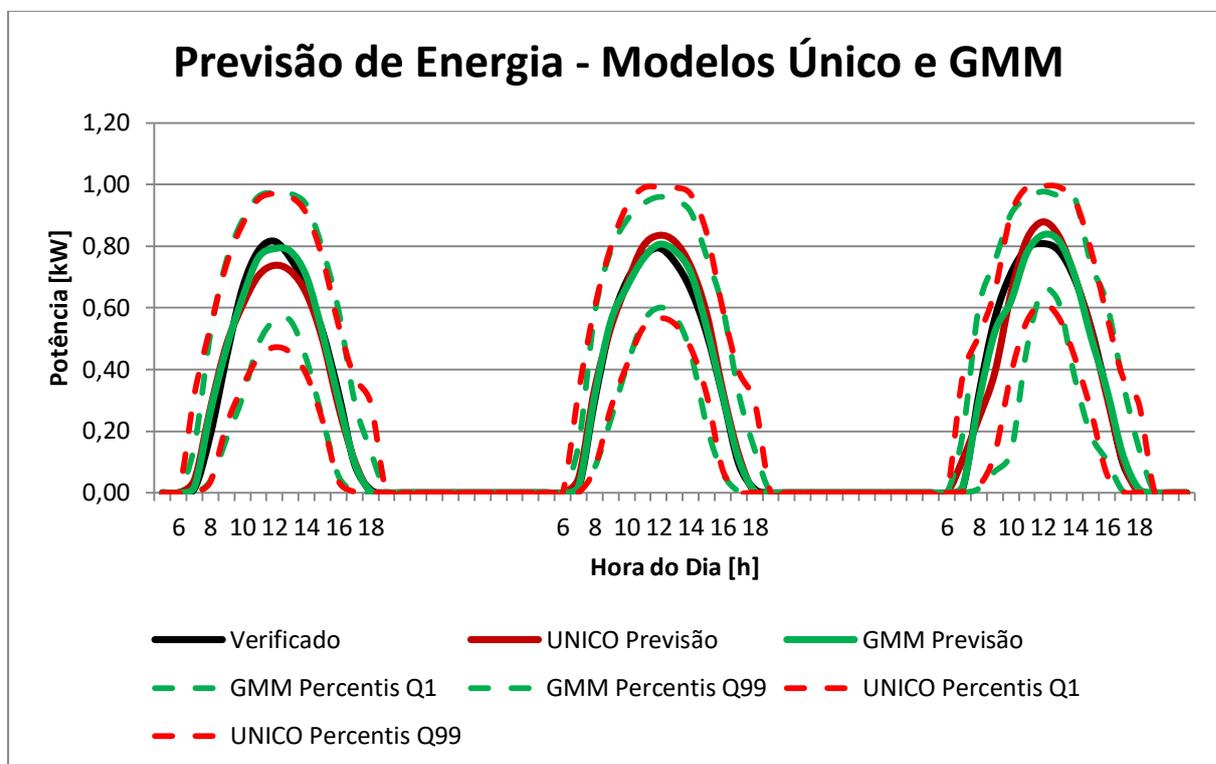


Gráfico 6: Previsão da produção de energia solar (setembro – zona 3)

O Gráfico 7 mostra as previsões da zona 1 para três dias de fevereiro. Mais uma vez é possível observar a ocorrência de dias claros. Para esses dias, as curvas sobrepostas mostram que as previsões foram praticamente precisas. Já para o terceiro dia, é possível observar, a princípio, a ocorrência de nebulosidade, que também ocorre para as previsões da zona 3 realizadas para três dias de junho, mostradas no Gráfico 8. Essa baixa produção de energia poderia estar associada de fato a um período instável, de nebulosidade, e o problema, nesse caso, estaria nos dados que foram apresentados à rede neural, visto que as curvas do *pinball* de, praticamente, todas as equipes apresentam o mesmo comportamento. Poderia ser um problema decorrente de qualidade da medição ou decorrente da falta de variáveis que pudessem permitir à rede neural modelar o problema a ser solucionado. Por exemplo, o estudo apresentado em [38] mostra que a rede neural obtém excelentes resultados para dados de irradiação solar medidos na superfície do arranjo fotovoltaico. Como os dados apresentados à rede foram medidos na superfície horizontal, pode ser que as demais variáveis repassadas à rede não fossem suficientes para permitir a modelagem do problema.

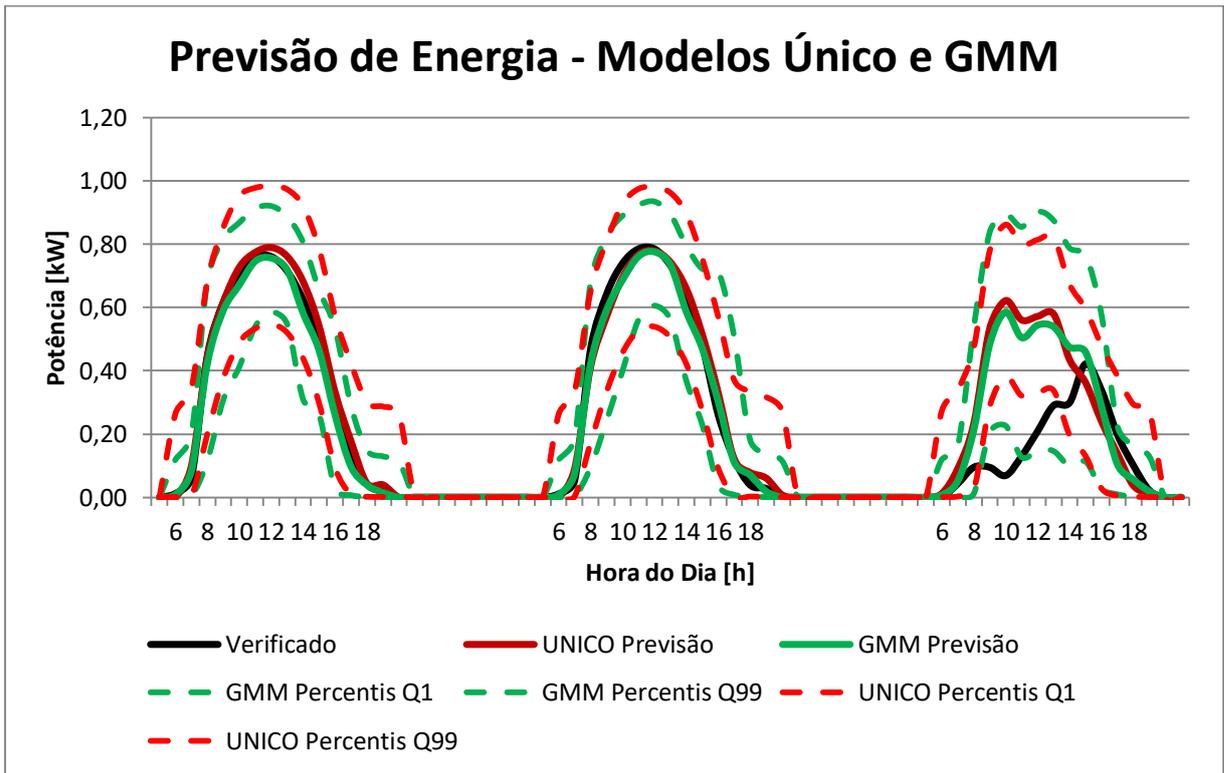


Gráfico 7: Previsão da produção de energia solar (fevereiro – zona 1)

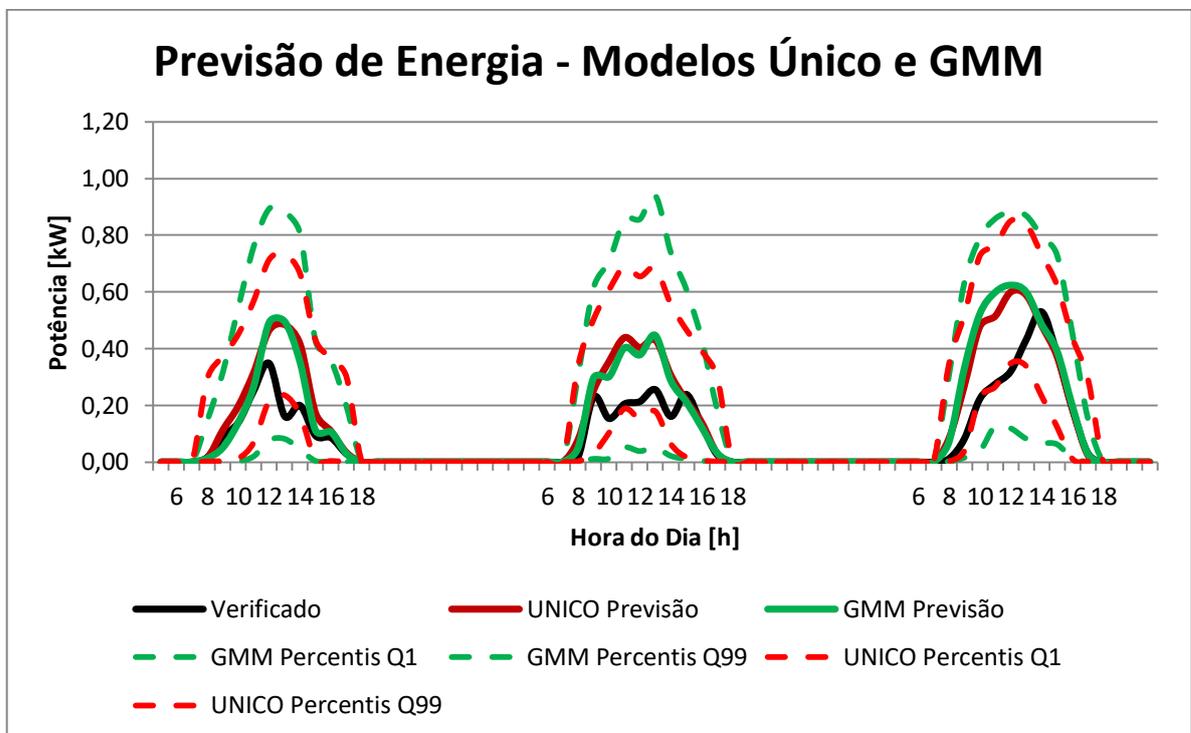


Gráfico 8: Previsão da produção de energia solar (junho – zona 3)

5 CONCLUSÃO E TRABALHOS FUTUROS

O presente trabalho apresentou dois modelos neurais probabilísticos para a previsão da produção de energia elétrica a partir de sistemas solares fotovoltaicos, visando a substituição dos tradicionais modelos determinísticos de previsão.

Os resultados das simulações, realizadas com dados medidos de sistemas reais, mostram que os modelos Único e GMM apresentaram, de maneira geral, resultados satisfatórios para as previsões, tendo o modelo GMM logrado êxito ao ficar entre os 10 melhores na competição.

Mais simples, o modelo Único, que não utiliza *clusterização*, alcançou a 15ª colocação na competição, apresentando baixos índices de erros relacionados à assertividade da rede e à tendência da rede neural em subestimar ou superestimar os resultados, obtendo valores de *pinball* muito melhores do que o *benchmark*. Mais sofisticado, o modelo GMM classifica os dados usando misturas de gaussianas, evitando a premissa de distribuição gaussiana para a variável de saída. O modelo GMM apresentou, na maioria dos casos, índices de erros inferiores aos do modelo Único, obtendo, em relação a este modelo, uma redução de 12% no valor médio do *pinball*, o que lhe garantiu a 9ª colocação na competição.

O fato da série de dados não ter sido analisada como deveria, por exemplo, sem estudos para processamento de *outliers*, pode ter prejudicado o aprendizado pela rede neural, influenciando negativamente alguns dos resultados apresentados. Como proposta para trabalhos futuros, propõe-se a execução de novas simulações usando os modelos e a base de dados apresentados, considerando uma análise criteriosa para os dados de entrada, de modo que apenas as variáveis realmente necessárias para a resolução do problema sejam apresentadas à rede neural. Também é vislumbrado o uso dos modelos probabilísticos propostos para pré-processamento de *outliers*. Paralelamente, uma análise a partir das localidades dos sistemas fotovoltaicos poderia por fim à questão da previsão para os meses de junho, julho e agosto.

BIBLIOGRAFIA

- [1] FREI, C. M. *Probabilistic Forecasts of Precipitation Using Quantiles*. Ruprecht-Karls-Universität Heidelberg. [S.l.]. 2012.
- [2] FERREIRA, V. H. *Desenvolvimento de Modelos Neurais Autônomos para Previsão de Carga Elétrica*. Tese (Doutorado) - Universidade Federal do Rio de Janeiro. Rio de Janeiro, p. 302. 2008.
- [3] HONG, T. et al. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. *International Journal of Forecasting*, n. 32, p. 896-913, 2016.
- [4] BRACALE, A. et al. A Bayesian Method for Short-Term Probabilistic Forecasting of Photovoltaic Generation in Smart Grid Operation and Control. *Energies*, n. 6, p. 733-747, 2013.
- [5] BRACALE, A.; CARPINELLI, G.; FALCO, P. A New Ensemble Probabilistic Method for Short-Term Photovoltaic Power Forecasting. *InTech*. Nápoles, 2016. p.140-159 DOI: 10.5772/65993
- [6] ABUELLA, M.; CHOWDHURY, B. Solar Power Probabilistic Forecasting by Using Multiple Linear Regression Analysis. *IEEE SoutheastCon 2015*. Fort Lauderdale: [s.n.]. 2015.
- [7] ADEJUMO, A. O.; SULEIMAN, E. A.; OKAGBUE, H. I. Exploration of Solar Radiation Data From Three Geo-Political Zones in Nigeria. *Data in Brief*, n. 13, p. 60-68, 2017.
- [8] RAMAKRISHNA, R.; SCAGLIONE, A.; VITTAL, V. A Stochastic Model for Short-Term Probabilistic Forecast of Solar Photo-Voltaic Power, *arXiv:1706.05445v2*, n. 1, 2017.
- [9] GOLESTANEH, F.; PINSON, P.; GOOI, H. B. Very Short-Term Nonparametric Probabilistic Forecasting of Renewable Energy Generation - With Application to Solar Energy. *IEEE Transactions On Power Systems*, v. 31, n. 5, p. 3850-3863, 2016.
- [10] CEPEL CENTRO DE PESQUISAS DE ENERGIA ELÉTRICA - CRESESB CENTRO DE REFERÊNCIA PARA ENERGIA SOLAR E EÓLICA SÉRGIO DE SALVO BRITO. *Manual de Engenharia para Sistemas Fotovoltaicos*. Rio de Janeiro. 2014.
- [11] LOUTZENHISER, P. G. et al. Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. *Solar Energy*, n. 81, p. 254-267, 2007.
- [12] KALOGIROU, S. *Engenharia de Energia Solar: Processos e Sistemas*. 1ª. ed. Rio de

Janeiro: Elsevier Brasil, 2017.

- [13] SCOLAR, J. *Estimativa da irradiação total sobre uma superfície inclinada a partir da irradiação global na horizontal*. Tese (Doutorado) - Universidade Estadual Paulista Júlio de Mesquita Filho. Botucatu. 2003.
- [14] KING, D. L.; BOYSON, W. E.; KRATOCHVIL, J. A. Photovoltaic Array Performance Model. *Sandia National Laboratories*, SAND2004-3535. New Mexico. 2004.
- [15] KING, D. L.; BOYSON, W. E.; KRATOCHVIL, J. A. Field Experience with a New Performance Characterization Procedure for Photovoltaic Arrays. *2nd World Conference and Exhibition on PV Solar Energy Conversion*. Vienna: [s.n.]. 1998.
- [16] FUENTES, M. K. A Simplified Thermal Model for Flat-Plate Photovoltaic Arrays. *Sandia National Laboratories*, SAND85-0330. Albuquerque. 1987.
- [17] SANDIA NATIONAL LABORATORIES. Performance Model for Grid-Connected Photovoltaic Inverters. *Sandia National Laboratories*. California. 2007.
- [18] STUDER INNOTECH (cortesia), 2014.
- [19] HAYKIN, S. *Redes Neurais - Princípios e Práticas*. São Paulo: Bookman, 2008.
- [20] HUDSON, D. L.; COHEN, M. E. *Neural Networks and Artificial Intelligence for Biomedical Engineering*. New York: IEEE Press, 2000.
- [21] GUYON, I. Neural networks and applications tutorial. *Physics Reports*, v. 207, n. 3, p. 215-259, 1991.
- [22] LOPES, D. S. *Modelos Neurais Autônomos para Classificação e Localização de Falhas em Linhas de Transmissão*. Dissertação (Mestrado) - Universidade Federal Fluminense. Niterói. 2017.
- [23] SILVA, I. N.; SPATTI, D. H.; FLAUZINO, A. *Redes neurais artificiais para engenharia e ciências aplicadas*. São Paulo: Artliber, 2010.
- [24] BISHOP, C. M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- [25] MOREIRA, R. F. C. *Previsão Probabilística dos Preços de Energia Elétrica do Mercado Ibérico de Eletricidade*. Dissertação (Mestrado) - Universidade do Porto. Porto. 2015.
- [26] SILVA, A. P. A.; MOULIN, L. Confidence Intervals for Neural Network Based Short-Term Load Forecasting. *IEEE Transactions On Power Systems*, p. 1191-1196, 2000.
- [27] LEE, H. K. H. A Noninformative Prior for Neural Networks. *Machine Learning*, v. 50, n. 1-2, p. 197-212, Jan. 2003.
- [28] BISHOP, C. M. *Neural Networks for Pattern Recognition*. New York: Oxford University Press, 1995.

- [29] MACKAY, D. J. C. *Bayesian Methods for Adaptive Models*. Tese (Doutorado) - California Institute of Technology. Pasadena, California, USA. 1992.
- [30] LACHI, R. L.; ROCHA, H. V. Aspectos básicos de clustering: conceitos e técnicas. *Universidade Estadual de Campinas*, IC-05-003. Campinas. 2005.
- [31] CASSIANO, K. M. *Análise de Séries Temporais Usando Análise Espectral Singular (SSA) e Clusterização de Suas Componentes Baseada em Densidade*. Tese (Doutorado) - Pontifícia Universidade Católica. Rio de Janeiro. 2014.
- [32] CARLANTONIO, L. M. *Novas Metodologias para Clusterização de Dados*. Dissertação (Mestrado) - Universidade Federal do Rio de Janeiro. Rio de Janeiro. 2001.
- [33] JAIN, A.; DUBES, R. *Algorithms for Clustering Data*. New Jersey: Prentice-Hall, 1988.
- [34] LINDEN, R. Técnicas de Agrupamento. *Revista de Sistemas de Informação da FSMA*, n. 4, p. 18-36, 2009.
- [35] GOLDSCHMITT, R.; PASSOS, E. *Data mining: um guia Prático*. Rio de Janeiro: Campus Elsevier, 2005.
- [36] PELLEGG, D.; MOORE, A. W. X-means: Extending K-means with Efficient Estimation of the Number of Clusters. *International Conference on Machine Learning. Proceedings of the Seventeenth International Conference on Machine Learning.*, p. 727-734, 2000.
- [37] LOKAD. Lokad. Disponível em: <<https://www.lokad.com/pinball-loss-function-definition>>. Acesso em: 20 novembro 2017.
- [38] MARTINS, R. P.; FERREIRA, V. H.; LOPES, T. T. Rede Neural Artificial para Previsão da Potência de Saída de Sistemas Fotovoltaicos Conectados à Rede. *VI Simpósio Brasileiro de Sistemas Elétricos*, Natal-RN, 22-25 Maio 2016.

APÊNDICE

Classificação das equipes

A Tabela 11 apresenta os valores do *pinball* de cada equipe para cada tarefa. Os campos não preenchidos, destacados pelas cores azul e amarela representam, respectivamente, os valores faltantes e com erro, conforme os critérios da *GEFCom 2014* [3] apresentados no item 4.1.3. A linha 3 da referida tabela apresenta o *benchmark* ou valores de referência. O *benchmark* representa os valores máximos que podem ser alcançados para que a estimativa seja válida. Valores superiores ao *benchmark* apresentam uma estimativa ruim.

Os resultados das simulações Modelo Único e Modelo GMM também são apresentados na Tabela 11, uma vez que são necessários para que os novos valores do *pinball* possam ser permutados devidamente. Na Tabela 12 são apresentadas as classificações parciais de cada tarefa e a segunda pior classificação de cada equipe. Com base nos critérios da *GEFCom 2014*, é possível preencher as posições faltantes. A primeira equipe listada, por exemplo, não tem dado válido para a tarefa 3. Observe que a segunda pior colocação dessa equipe é a posição 22. Entretanto, na tarefa 3, há apenas 18 valores válidos, sendo que um é o *benchmark*. Assim, de acordo com os critérios estabelecidos, essa equipe deve ocupar a posição 18. O mesmo ocorre para as duas últimas equipes.

Tabela 11: Valores do *pinball* apurados para as zonas 1, 2 e 3

Nome	Tarefas											
	1	2	3	4	5	6	7	8	9	10	11	12
Arkadiy Strelnikov	0,0289	0,0256		0,0355	0,0352	0,0417	0,0322	0,0320	0,0303	0,0287	0,0223	0,0372
Azhar Ahmed Mohammed				0,0189	0,0193	0,0151	0,0151	0,0153	0,0164	0,0142	0,0106	0,0145
Benchmark - Solar	0,0331	0,0388	0,0359	0,0361	0,0479	0,0357	0,0421	0,0399	0,0435	0,0377	0,0320	0,0285
C3 Green Team	0,0173			0,0128	0,0129	0,0131	0,0132	0,0119	0,0137	0,0117	0,0086	0,0127
E.S. Mangalova		0,0177	0,0160	0,0143	0,0171	0,0167	0,0150		0,0198	0,0168	0,0112	0,0153
Gang-gang	0,0129	0,0162	0,0117	0,0112	0,0125	0,0124	0,0124	0,0120	0,0126	0,0110	0,0082	0,0124
Giuseppe C.	0,0131	0,0171	0,0178	0,0129	0,0131	0,0134	0,0134	0,0116	0,0143	0,0115	0,0091	0,0130
Humboldt State University Solar Falcons	0,0179		0,0168	0,0160	0,0233	0,0200	0,0726	0,0214	0,0244	0,0227		0,0201
Manuel Oviedo de la Fuente	0,0206	0,0249	0,0197	0,0174	0,0189	0,0174	0,0207	0,0391	0,0209	0,0179	0,0138	0,0190
Mohamed Abuella	0,0151	0,0201	0,0148	0,0153	0,0173	0,0163	0,0275	0,0166	0,0164	0,0146	0,0104	0,0144
Power Team (SAOR)				0,0236	0,0209	0,0189		0,0292	0,0194	0,0186		0,0184
RDSsol	0,0138	0,0187	0,0165	0,0205	0,0224	0,0231	0,0208	0,0137	0,0148	0,0122	0,0098	0,0137
San/Saini	0,0194	0,0231	0,0176	0,0174	0,0180	0,0175	0,0178	0,0172	0,0181	0,0153	0,0119	0,0174
T-morning	0,0151	0,0186	0,0140	0,0142	0,0152	0,0147	0,0148	0,0139	0,0151	0,0127	0,0093	0,0137
THU_EILAB#6	0,0210	0,0184	0,0146	0,0147	0,0158		0,0157	0,0142	0,0151	0,0126	0,0098	0,0132
UT_Argonne	0,0145	0,0198	0,0155	0,0144	0,0176	0,0154	0,0141	0,0120	0,0129	0,0125	0,0082	0,0131
Xiaorong (Iris) Sun	0,0158	0,0207		0,0285	0,0178	0,0168	0,0188		0,0198	0,0173	0,0108	
Yao Zhang	0,0159	0,0172	0,0149	0,0136	0,0143	0,0141	0,0156	0,0148	0,0145	0,0127	0,0102	0,0138
Ying Chen	0,0156	0,0194	0,0146	0,0163		0,0167	0,0162	0,0157	0,0167	0,0140	0,0108	0,0143
dmlab	0,0124	0,0165	0,0123	0,0116	0,0125	0,0123	0,0125	0,0117	0,0131	0,0109	0,0084	0,0129
pat1	0,0151	0,0218	0,0174	0,0188	0,0246	0,0178	0,0177	0,0157	0,0159	0,0160	0,0149	0,0189
Modelo Único	0,0156	0,0208	0,0175	0,0197	0,0194	0,0193	0,0197	0,0165	0,0180	0,0158	0,0117	0,0155
Modelo GMM	0,0140	0,0186	0,0158	0,0174	0,0168	0,0176	0,0165	0,0140	0,0157	0,0131	0,0107	0,0142

Tabela 12: Classificação das equipes por tarefa e 2ª pior posição

Nome	Tarefas												2ª Pior Posição
	1	2	3	4	5	6	7	8	9	10	11	12	
Arkadiy Strelnikov	19	18	18	22	21	22	20	19	22	22	20	22	22
Azhar Ahmed Mohammed	15	15	15	17	15	7	8	11	12	12	11	13	15
Benchmark - Solar	20	19	18	23	22	21	21	21	23	23	21	21	23
C3 Green Team	14	4	4	3	3	3	3	3	4	4	4	2	4
E.S. Mangalova	17	5	10	7	9	11	7	17	18	17	15	14	17
Gang-gang	2	1	1	1	1	2	1	5	1	2	2	1	2
Giuseppe C.	3	3	16	4	4	4	4	1	5	3	5	4	5
Humboldt State University Solar Falcons	15	19	12	11	19	19	22	17	21	21	21	20	21
Manuel Oviedo de la Fuente	17	17	17	15	14	13	17	20	20	19	18	19	20
Mohamed Abuella	9	12	6	10	10	9	19	15	13	13	10	12	15
Power Team (SAOR)	20	19	18	20	17	17	20	18	17	20	20	17	20
RDSsol	4	9	11	19	18	20	18	6	7	5	8	7	19
San/Saini	16	16	15	13	13	14	14	16	16	14	17	16	16
T-morning	8	8	3	6	6	6	6	7	8	9	6	8	8
THU_EILAB#6	18	6	5	9	7	10	10	9	9	7	7	6	10
UT_Argonne	6	11	8	8	11	8	5	4	2	6	1	5	11
Xiaorong (Iris) Sun	12	13	18	21	12	12	15	19	19	18	13	19	19
Yao Zhang	13	4	7	5	5	5	9	10	6	8	9	9	10
Ying Chen	10	10	4	12	13	10	11	13	14	11	13	11	13
dmlab	1	2	2	2	2	1	2	2	3	1	3	3	3
pat1	7	15	13	16	20	16	13	12	11	16	19	18	19
Modelo Único	11	14	14	18	16	18	16	14	15	15	16	15	18
Modelo GMM	5	7	9	14	8	15	12	8	10	10	12	10	14

Tabela 13: Valores finais do pinball

Nome	Tarefas											
	1	2	3	4	5	6	7	8	9	10	11	12
Arkadiy Strelnikov	0,0289	0,0256	0,0359	0,0355	0,0352	0,0417	0,0322	0,0320	0,0303	0,0287	0,0223	0,0372
Azhar Ahmed Mohammed	0,0179	0,0218	0,0176	0,0189	0,0193	0,0151	0,0151	0,0153	0,0164	0,0142	0,0106	0,0145
Benchmark - Solar	0,0331	0,0388	0,0359	0,0361	0,0479	0,0357	0,0421	0,0399	0,0435	0,0377	0,0320	0,0285
C3 Green Team	0,0173	0,0172	0,0146	0,0128	0,0129	0,0131	0,0132	0,0119	0,0137	0,0117	0,0086	0,0127
E.S. Mangalova	0,0206	0,0177	0,0160	0,0143	0,0171	0,0167	0,0150	0,0214	0,0198	0,0168	0,0112	0,0153
Gang-gang	0,0129	0,0162	0,0117	0,0112	0,0125	0,0124	0,0124	0,0120	0,0126	0,0110	0,0082	0,0124
Giuseppe C.	0,0131	0,0171	0,0178	0,0129	0,0131	0,0134	0,0134	0,0116	0,0143	0,0115	0,0091	0,0130
Humboldt State University Solar Falcons	0,0179	0,0388	0,0168	0,0160	0,0233	0,0200	0,0726	0,0214	0,0244	0,0227	0,0320	0,0201
Manuel Oviedo de la Fuente	0,0206	0,0249	0,0197	0,0174	0,0189	0,0174	0,0207	0,0391	0,0209	0,0179	0,0138	0,0190
Mohamed Abuella	0,0151	0,0201	0,0148	0,0153	0,0173	0,0163	0,0275	0,0166	0,0164	0,0146	0,0104	0,0144
Power Team (SAOR)	0,0331	0,0388	0,0359	0,0236	0,0209	0,0189	0,0322	0,0292	0,0194	0,0186	0,0223	0,0184
RDSsol	0,0138	0,0187	0,0165	0,0205	0,0224	0,0231	0,0208	0,0137	0,0148	0,0122	0,0098	0,0137
San/Saini	0,0194	0,0231	0,0176	0,0174	0,0180	0,0175	0,0178	0,0172	0,0181	0,0153	0,0119	0,0174
T-morning	0,0151	0,0186	0,0140	0,0142	0,0152	0,0147	0,0148	0,0139	0,0151	0,0127	0,0093	0,0137
THU_EILAB#6	0,0210	0,0184	0,0146	0,0147	0,0158	0,0167	0,0157	0,0142	0,0151	0,0126	0,0098	0,0132
UT_Argonne	0,0145	0,0198	0,0155	0,0144	0,0176	0,0154	0,0141	0,0120	0,0129	0,0125	0,0082	0,0131
Xiaorong (Iris) Sun	0,0158	0,0207	0,0359	0,0285	0,0178	0,0168	0,0188	0,0320	0,0198	0,0173	0,0108	0,0190
Yao Zhang	0,0159	0,0172	0,0149	0,0136	0,0143	0,0141	0,0156	0,0148	0,0145	0,0127	0,0102	0,0138
Ying Chen	0,0156	0,0194	0,0146	0,0163	0,0180	0,0167	0,0162	0,0157	0,0167	0,0140	0,0108	0,0143
dmlab	0,0124	0,0165	0,0123	0,0116	0,0125	0,0123	0,0125	0,0117	0,0131	0,0109	0,0084	0,0129
pat1	0,0151	0,0218	0,0174	0,0188	0,0246	0,0178	0,0177	0,0157	0,0159	0,0160	0,0149	0,0189
Modelo Único	0,0156	0,0208	0,0175	0,0197	0,0194	0,0193	0,0197	0,0165	0,0180	0,0158	0,0117	0,0155
Modelo GMM	0,0140	0,0186	0,0158	0,0174	0,0168	0,0176	0,0165	0,0140	0,0157	0,0131	0,0107	0,0142

Tabela 14: Classificação final das equipes

Nome	Tarefas												Classif.	Posição Atual	Posição Anterior
	1	2	3	4	5	6	7	8	9	10	11	12			
Gang-gang	0,61	0,58	0,67	0,69	0,74	0,65	0,71	0,70	0,71	0,71	0,74	0,56	68,1%	1	1
dmlab	0,63	0,58	0,66	0,68	0,74	0,66	0,70	0,71	0,70	0,71	0,74	0,55	67,6%	2	2
C3 Green Team	0,48	0,56	0,59	0,64	0,73	0,63	0,69	0,70	0,68	0,69	0,73	0,55	66,1%	3	3
Giuseppe C.	0,61	0,56	0,50	0,64	0,73	0,63	0,68	0,71	0,67	0,70	0,72	0,54	65,4%	4	4
UT_Argonne	0,56	0,49	0,57	0,60	0,63	0,57	0,67	0,70	0,70	0,67	0,74	0,54	64,3%	5	5
T-morning	0,54	0,52	0,61	0,61	0,68	0,59	0,65	0,65	0,65	0,66	0,71	0,52	62,9%	6	6
Yao Zhang	0,52	0,56	0,58	0,62	0,70	0,60	0,63	0,63	0,67	0,66	0,68	0,52	62,6%	7	7
THU_EILAB#6	0,37	0,53	0,59	0,59	0,67	0,53	0,63	0,64	0,65	0,67	0,69	0,54	61,9%	8	8
Modelo GMM	0,58	0,52	0,56	0,52	0,65	0,51	0,61	0,65	0,64	0,65	0,67	0,50	59,9%	9	-
Ying Chen	0,53	0,50	0,59	0,55	0,62	0,53	0,61	0,61	0,62	0,63	0,66	0,50	59,1%	10	9
Azhar Ahmed Mohammed	0,46	0,44	0,51	0,48	0,60	0,58	0,64	0,62	0,62	0,62	0,67	0,49	58,6%	11	10
RDSsol	0,58	0,52	0,54	0,43	0,53	0,35	0,51	0,66	0,66	0,68	0,69	0,52	57,8%	12	11
Mohamed Abuella	0,54	0,48	0,59	0,58	0,64	0,54	0,35	0,59	0,62	0,61	0,67	0,49	56,7%	13	12
E.S. Mangalova	0,38	0,54	0,55	0,60	0,64	0,53	0,64	0,46	0,55	0,55	0,65	0,46	55,6%	14	13
Modelo Único	0,53	0,46	0,51	0,45	0,60	0,46	0,53	0,59	0,59	0,58	0,64	0,45	54,5%	15	14
San/Saini	0,41	0,41	0,51	0,52	0,62	0,51	0,58	0,57	0,58	0,59	0,63	0,39	54,4%	16	15
pat1	0,54	0,44	0,52	0,48	0,49	0,50	0,58	0,61	0,63	0,58	0,54	0,34	52,1%	17	16
Xiaorong (Iris) Sun	0,52	0,47	0,00	0,21	0,63	0,53	0,55	0,20	0,55	0,54	0,66	0,33	45,8%	18	17
Manuel Oviedo de la Fuente	0,38	0,36	0,45	0,52	0,61	0,51	0,51	0,02	0,52	0,52	0,57	0,33	44,3%	19	18
Power Team (SAOR)	0,00	0,00	0,00	0,35	0,56	0,47	0,23	0,27	0,55	0,51	0,30	0,35	36,5%	20	19
Humboldt State University Solar Falcons	0,46	0,00	0,53	0,56	0,51	0,44	-0,72	0,46	0,44	0,40	0,00	0,29	25,1%	21	20
Arkadiy Strelnikov	0,13	0,34	0,00	0,02	0,26	-0,17	0,23	0,20	0,30	0,24	0,30	-0,30	11,8%	22	21
Benchmark - Solar	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,0%	23	22