



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

ELSON ANTONIO NUNES JUNIOR

**Information Theoretic Learning para Previsão
de Vazões Diárias e seu Impacto na Formação
do PLD Horário**

NITERÓI

2022

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

ELSON ANTONIO NUNES JUNIOR

**Information Theoretic Learning para Previsão de Vazões
Diárias e seu Impacto na Formação do PLD Horário**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica e de Telecomunicações. Área de concentração: Sistemas de Energia Elétrica.

Orientador:

Vitor Hugo Ferreira

Coorientador:

André da Costa Pinho

NITERÓI

2022

Ficha catalográfica automática - SDC/BEE
Gerada com informações fornecidas pelo autor

N972i Nunes junior, Elson Antonio
Information Theoretic Learning para Previsão de Vazões
Diárias e seu Impacto na Formação do PLD Horário / Elson
Antonio Nunes junior ; Vitor Hugo Ferreira, orientador ;
André da Costa Pinho, coorientador. Niterói, 2022.
101 f. : il.

Dissertação (mestrado)-Universidade Federal Fluminense,
Niterói, 2022.

DOI: <http://dx.doi.org/10.22409/PPGEET.2022.m.14650412730>

1. Redes neurais artificiais. 2. Information theoretic
learning. 3. Previsão de vazões. 4. Preço spot. 5.
Produção intelectual. I. Ferreira, Vitor Hugo, orientador.
II. Costa Pinho, André da, coorientador. III. Universidade
Federal Fluminense. Escola de Engenharia. IV. Título.

CDD -

ELSON ANTONIO NUNES JUNIOR

Information Theoretic Learning para Previsão de Vazões Diárias e seu Impacto na
Formação do PLD Horário

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Telecomunicações da Universidade Federal Fluminense, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica e de Telecomunicações. Área de concentração: Sistemas de Energia Elétrica.

BANCA EXAMINADORA

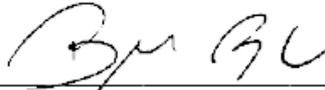


Assinado de forma digital por VITOR
HUGO FERREIRA
vfferreira@id.uff.br:01277679673
Dados: 2022.07.28 13:13:33 -03'00'

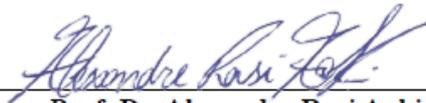
Prof. Dr. Vitor Hugo Ferreira - Orientador
Universidade Federal Fluminense - UFF



Prof. Dr. André da Costa Pinho - Coorientador
Universidade Federal Fluminense - UFF



Prof. Dr. Bruno Soares Moreira Cesar Borba
Universidade Federal Fluminense - UFF



Prof. Dr. Alexandre Rasi Aoki
Universidade Federal do Paraná - UFPR

Niterói

Julho de 2022

A Deus, minha família e amigos

Agradecimentos

Agradeço a Deus por todas as graças alcançadas e por ter me guiado não apenas no período do mestrado, mas por toda a minha vida.

Agradeço à minha família por todo amor, carinho e incentivo, bens que me dão forças para encarar todos os desafios existentes.

Agradeço à minha companheira para toda vida, Bianca, pelo amor e apoio mútuo, tão importantes em todos os passos que damos juntos. Também sou grato por sua família – que, na verdade, torna-se minha também – por todo o carinho e apoio dado.

Agradeço aos amigos que a vida me trouxe, na UFF ou fora dela, que me ajudaram direta ou indiretamente, mesmo com pequenos momentos de descontração em meio à rotina de estudos.

Agradeço ao professor Vitor Hugo Ferreira pelos papéis de orientador no mestrado e graduação e tutor nos tempos de PET-Elétrica, fundamentais no meu desenvolvimento acadêmico e profissional.

Agradeço ao professor André da Costa Pinho pela orientação dada durante o curso de mestrado, buscando sempre a melhoria do trabalho desenvolvido em minha dissertação.

Agradeço aos meus colegas de trabalho da Enel pela paciência e apoio com conceitos teóricos aprendidos e pela permissão de uso dos programas computacionais utilizados neste trabalho.

Agradeço também a todos aqueles que, mesmo que inconscientemente, tiveram participação nesta importante etapa da minha vida.

Resumo

Como a previsão de vazões afluentes a hidrelétricas é uma das informações de entrada do processo de programação da operação do Sistema Interligado Nacional (SIN), é importante que as previsões geradas sejam cada vez mais assertivas de modo que as saídas desse processo estejam cada vez mais condizentes com as reais condições do sistema. O uso de Redes Neurais Artificiais (RNAs) para previsão de afluências pode ser considerada como uma alternativa aos métodos preditivos oficiais adotados pelo Operador Nacional do Sistema Elétrico (ONS) e, nessa linha, os conceitos de Information Theoretic Learning (ITL) garantem análise estatística mais robusta aplicada na etapa de aprendizagem da rede. O objetivo deste trabalho é a modelagem de um previsor de afluências diárias a partir do uso de técnicas de RNAs e dados de entrada hidrológicos e meteorológicos. No treino da rede, foi utilizado o Critério de Maximização da Correntropia (MCC) – conceito oriundo do framework de ITL – com o objetivo de aumentar a similaridade entre amostras calculadas e reais, conseqüentemente reduzindo o erro de previsão. As saídas do modelo são incorporadas ao programa de otimização de despacho hidrotérmico de curtíssimo prazo, a fim de se analisar o impacto no preço spot do mercado de energia elétrica. Para efeito comparativo, foi desenvolvida outra rede neural, porém sem o uso do ITL, além de um previsor ingênuo. Na análise de exatidão das afluências previstas, os resultados mostraram que o previsor principal se sobressaiu aos outros dois modelos e, no horizonte do primeiro dia de previsão, foi melhor ou equivalente aos modelos utilizados oficialmente no setor em 63,9% das hidrelétricas avaliadas. Na etapa de avaliação de preço, a curva de variação do preço, referente à alternativa principal apresentou ótima aderência à referência oficial, acompanhando movimentos de subida e descida de maior volatilidade. Ao final, é adotada uma curva de preços referente ao cenário de vazões reais, provendo comparação na qual se verificou que, em média, os preços relativos ao previsor proposto foram mais próximos dos valores de afluências reais do que as curvas de referência oficiais.

Palavras-chaves: Redes Neurais Artificiais, Information Theoretic Learning, Correntropia, Previsão de Vazões, Preço *Spot*, PLD.

Abstract

As the streamflow forecast is one of the inputs of the system operation programming process, it is important that the forecasts become more accurate in such a way that the outputs of that process are increasingly aligned with the real conditions of the electrical system. The use of Artificial Neural Networks for streamflow forecast can be considered as an alternative to the official forecasters applied by the Transmission System Operator in Brazil (ONS) and, in this perspective, the concepts of Information Theoretic Learning provide robustness for statistical analysis during training step. The goal of this paper is the development of a daily streamflow forecast model with artificial neural networks techniques and hydrological and meteorological input data. In training step, it was adopted the Maximum Correntropy Criterion – a ITL concept – to maximize the similarity between calculated and real outputs, reducing prediction errors consequently. The outputs of this new model were used as input for the very short-term hydrothermal dispatch software to analyze the impact in the spot price of the Brazilian energy market. For comparative analysis, two other models were developed: an ANN without the application of ITL and a naïve alternative. In the accuracy analysis of predicted flows, the results showed that the main predictor overcame the other two models and, in the first day of forecast, it was equivalent or more accurate than the models officially used in the energy sector for 63,9% of the investigated hydro power plants. The outputs regarding the spot price indicated a similar behavior of the main model's curve compared to the official reference, following upward and downward movements with more volatility. At the end, a price curve referring to the application of the realized streamflow is adopted, providing the following conclusion: the prices related to the proposed forecaster were closer to the realized streamflow scenario than the official curves.

Keywords: Artificial Neural Networks, Information Theoretic Learning, Correntropy, Flow Forecast, Spot Price, PLD.

Lista de Figuras

FIGURA 1 - VAZÃO AFLUENTE DE BARRA GRANDE E JIRAU. FONTE: ELABORADO DE [3].	2
FIGURA 2 - LISTA DE USINAS TÉRMICAS ORDENADAS POR CVU. FONTE: [8].	8
FIGURA 3 - FLUXOGRAMA DE FORMAÇÃO E USO DO PLD. FONTE: ADAPTADO DE [10].	9
FIGURA 4 - DILEMA DO OPERADOR. FONTE: [13].	10
FIGURA 5 - GRÁFICO DE CUSTO POR ARMAZENAMENTO FINAL. FONTE: ADAPTADO DE [14].	11
FIGURA 6 - REPRESENTAÇÃO ESPACIAL ADOTADA NO NEWAVE. FONTE: [10].	12
FIGURA 7 - ÁRVORE DE CENÁRIOS HIDROLÓGICOS DO DECOMP. FONTE: [20].	14
FIGURA 8 - PLD DO SUDESTE/CENTRO-OESTE NO DIA 31/12/2020. FONTE: ELABORADO DE [26].	16
FIGURA 9 - ESQUEMA DE UTILIZAÇÃO DO MODELO DESSEM. FONTE: ADAPTADO DE [28].	17
FIGURA 10 - ARQUIVO DADVAZ. FONTE: [33].	19
FIGURA 11 - DIAGRAMA DE BLOCOS DO SISTEMA NERVOSO. FONTE: ADAPTADO DE [36].	23
FIGURA 12 - NEURÔNIO BIOLÓGICO. FONTE: [37].	23
FIGURA 13 - NEURÔNIO ARTIFICIAL. FONTE: ADAPTADO DE [36].	23
FIGURA 14 - FUNÇÕES DE ATIVAÇÃO LIMIAR, LIMIAR POR PARTES E SIGMOIDE.	25
FIGURA 15 - ARQUITETURAS DE REDES NEURAIS ARTIFICIAIS.	26
FIGURA 16 - MAPA DE PRECIPITAÇÃO OBSERVADA DIÁRIA. FONTE: [69].	44
FIGURA 17 - ALGORITMO DE PROCESSAMENTO DOS DADOS DE PRECIPITAÇÃO	45
FIGURA 18 - HISTOGRAMAS DE CORRELAÇÕES ENTRE VAZÕES EM T-N E T, COM T = 1, 2, ..., 10.	47
FIGURA 19 - EVOLUÇÃO DOS ERROS DE TREINO E VALIDAÇÃO DE ALGUMAS UHES.	49
FIGURA 20 - FLUXOGRAMA DO PREVISOR DE VAZÕES PROPOSTO.	50
FIGURA 21 - FLUXOGRAMA DOS TESTES DE HIPÓTESES DO TIPO T-PAREADO UNILATERAL À DIREITA	55
FIGURA 22 - FLUXOGRAMA DA METODOLOGIA APLICADA.	55
FIGURA 23 - <i>BOXPLOT</i> COM ERRO DE PREVISÃO DO PLD DO SUBMERCADO SUDESTE/CENTRO-OESTE.	59
FIGURA 24 - <i>BOXPLOT</i> COM ERRO DE PREVISÃO DO PLD DO SUBMERCADO SUL.	60
FIGURA 25 - <i>BOXPLOT</i> COM ERRO DE PREVISÃO DO PLD DO SUBMERCADO NORDESTE.	61
FIGURA 26 - <i>BOXPLOT</i> COM ERRO DE PREVISÃO DO PLD DO SUBMERCADO NORTE.	62
FIGURA 27 - CURVAS DE PLD PARA O DIA 05/03/2022.	63
FIGURA 28 - CURVAS DE PLD PARA O DIA 12/03/2022.	63
FIGURA 29 - CURVAS DE PLD PARA O DIA 19/03/2022.	64
FIGURA 30 - CURVAS DE PLD PARA O DIA 26/03/2022.	64
FIGURA 31 - CURVAS DE PLD PARA O DIA 02/04/2022.	65
FIGURA 32 - CURVAS DE PLD PARA O DIA 09/04/2022.	65
FIGURA 33 - CURVAS DE PLD PARA O DIA 16/04/2022.	66
FIGURA 34 - CURVAS DE PLD PARA O DIA 23/04/2022.	66
FIGURA 35 - CURVAS DE PLD PARA O DIA 30/04/2022.	67
FIGURA 36 - CURVAS DE PLD PARA O DIA 04/06/2022.	67
FIGURA 37 - CURVAS DE PLD PARA O DIA 11/06/2022.	68
FIGURA 38 - CURVAS DE PLD PARA O DIA 18/06/2022.	68
FIGURA 39 - CURVAS DE PLD PARA O DIA 25/06/2022.	69
FIGURA 40 - CURVAS DE PLD PARA O DIA 05/03/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	71

FIGURA 41 - CURVAS DE PLD PARA O DIA 12/03/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	71
FIGURA 42 - CURVAS DE PLD PARA O DIA 19/03/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	72
FIGURA 43 - CURVAS DE PLD PARA O DIA 26/03/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	72
FIGURA 44 - CURVAS DE PLD PARA O DIA 02/04/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	73
FIGURA 45 - CURVAS DE PLD PARA O DIA 09/04/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	73
FIGURA 46 - CURVAS DE PLD PARA O DIA 16/04/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	74
FIGURA 47 - CURVAS DE PLD PARA O DIA 23/04/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	74
FIGURA 48 - CURVAS DE PLD PARA O DIA 30/04/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	75
FIGURA 49 - CURVAS DE PLD PARA O DIA 04/06/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	75
FIGURA 50 - CURVAS DE PLD PARA O DIA 11/06/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	76
FIGURA 51 - CURVAS DE PLD PARA O DIA 18/06/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	76
FIGURA 52 - CURVAS DE PLD PARA O DIA 25/06/2022, CONSIDERANDO PREÇOS DO CASO DE VAZÃO REALIZADA.	77

Lista de Tabelas

TABELA 1 - CURTOSE E ASSIMETRIA DAS DISTRIBUIÇÕES DE ERROS DE PREVISÃO DE VAZÃO PARA ALGUMAS USINAS DO SIN	4
TABELA 2 - INFORMAÇÕES DOS DADOS DE ENTRADA TEMPORAIS	46
TABELA 3 - CONFIGURAÇÕES DO MODELO RNA-MCC	51
TABELA 4 - AVALIAÇÃO DO NÚMERO DE NEURÔNIOS OCULTOS.....	56
TABELA 5 - RESULTADOS DOS TESTES DE HIPÓTESES: ONS E RNA-MCC.....	57
TABELA 6 - RESULTADOS DOS TESTES DE HIPÓTESES: ONS E RNA-MSE	57
TABELA 7 - RESULTADOS DOS TESTES DE HIPÓTESES: RNA-MCC E RNA-MSE.....	58
TABELA 8 - RESULTADOS DOS TESTES DE HIPÓTESES: RNA-MCC E MODELO INGÊNUO.....	58
TABELA 9 - NMAE DOS PLDS DOS MODELOS EM RELAÇÃO AOS VALORES OFICIAIS DA CCEE	69
TABELA 10 - NMAE DOS PLDS DA RNA-MCC E DA CCEE EM RELAÇÃO AOS PREÇOS DO CENÁRIO DE VAZÕES REALIZADAS	77

Lista de Abreviaturas e Siglas

ANEEL	Agência Nacional de Energia Elétrica
CCEE	Câmara de Comercialização de Energia Elétrica
CEPEL	Centro de Pesquisas de Energia Elétrica
CIM	<i>Correntropy Induced Metric</i>
CMO	Custo Marginal de Operação
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos
CVU	Custo Variável Unitário
ENA	Energia Natural Afluente
FCF	Função de Custo Futuro
FCI	Função de Custo Imediato
INMET	Instituto Nacional de Meteorologia
INPE	Instituto Nacional de Pesquisas Espaciais
IP	<i>Information Potential</i>

ITL	<i>Information Theoretic Learning</i>
MAE	<i>Mean Absolute Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MCC	<i>Maximum Correntropy Criterion</i>
MCP	Mercado de Curto Prazo
MEE	<i>Minimum Entropy Error</i>
MILP	Programação Linear Inteira Mista
ML	<i>Machine Learning</i>
MLP	<i>Multilayer Perceptron</i>
MME	Ministério de Minas e Energia
MSE	<i>Mean Squared Error</i>
NMAE	<i>Normalized Mean Absolute Error</i>
ONS	Operador Nacional do Sistema Elétrico
PDDE	Programação Dinâmica Dual Estocástica
PDF	<i>Probability Density Function</i>

PLD	Preço de Liquidação das Diferenças
REE	Reservatório Equivalente de Energia
RNA	Rede Neural Artificial
SIN	Sistema Interligado Nacional
TEO	Tarifa de Energia de Otimização
TI	Teoria da Informação
UHE	Usina Hidrelétrica

Sumário

Capítulo 1 - Introdução.....	1
1.1 Objetivos	5
1.2 Estrutura do Trabalho	5
Capítulo 2 - Formação do PLD Horário.....	7
2.1 Preço de Liquidação das Diferenças	7
2.2 Cadeia de Modelos Encadeados.....	9
2.2.1 NEWAVE.....	10
2.2.2 DECOMP	13
2.2.3 DESSEM	15
2.3 Previsão de Vazões.....	18
2.3.1 Relação com Modelagem de Curtíssimo Prazo	18
2.3.2 Modelos Utilizados pelo ONS.....	20
Capítulo 3 - Conceitos de RNA e ITL.....	22
3.1 Redes Neurais Artificiais.....	22
3.1.1 Conceitos Básicos de RNAs	22
3.1.2 Redes do Tipo Perceptron	27
3.2 Information Theoretic Learning	29
3.2.1 Conceitos Básicos de ITL.....	29
3.2.2 Critério de Máxima Correntropia	34
Capítulo 4 - Revisão da Literatura	37
Capítulo 5 - Materiais e Método.....	43
5.1 Descrição dos Dados de Entrada e Saída.....	43
5.2 Modelagem do Previsor	46
5.3 Metodologia	51
Capítulo 6 - Resultados e Discussão.....	56
6.1 Resultados quanto à Previsão de Afluências	56
6.2 Resultados quanto à Formação do PLD	59
Capítulo 7 - Conclusão	78
7.1 Trabalhos Futuros	81
Bibliografia.....	83

Capítulo 1 - Introdução

Apesar da recente expansão de fontes intermitentes, as usinas hidrelétricas (UHEs) permanecem como a principal fonte de energia no Brasil. Ilustrando tal informação, no ano de 2021, a parcela hidrelétrica correspondia a 62,3% de toda a geração do Sistema Interligado Nacional (SIN), quase três vezes o percentual da segunda colocada, a geração térmica, com 22,1% e mais de cinco vezes a terceira colocada, a geração eólica, com 12% [1]. Diante da relevância da produção hidrelétrica para a matriz de geração de energia elétrica brasileira, é importante o bom entendimento acerca das características desse tipo de geração, no que diz respeito à atividade de programação da operação do SIN.

Dentre tais características, uma das mais marcantes é a estocasticidade, uma vez que a geração hidrelétrica depende de fatores que não podem ser previstos com total exatidão. Um desses fatores é a precipitação de chuvas que, mesmo com a utilização de modelos matemáticos previsores, ainda traz grande incerteza em relação à magnitude da chuva e à localização geográfica de onde ocorre o evento meteorológico. A combinação de quantidade e posição, aliás, é extremamente determinante no impacto da precipitação no volume de água disponível para hidrelétricas, dado que um evento de chuva torrencial não aumenta os níveis de reservatórios se a mesma não ocorrer sobre a área das bacias onde estão localizadas as usinas. Nessa perspectiva, como a informação de precipitação impacta diretamente no fluxo de água que chega às hidrelétricas, o caráter probabilístico é passado para a chamada vazão natural afluente, definida em [2] como a vazão provida pela própria natureza.

Levando em conta as dimensões continentais do país, existem diferentes situações climatológicas que coexistem e impactam umas nas outras no território brasileiro, gerando condições hidrológicas concomitantes como, por exemplo, seca na região Nordeste, chuva em abundância no Sul, assim como o inverso pode acontecer. Esse cenário dinâmico pode alavancar a incerteza na determinação das vazões afluentes, dependendo da bacia e reservatório considerados. A Figura 1 [3] mostra um exemplo de diferença de comportamento de afluências ao longo do tempo, apresentando a comparação entre as vazões afluentes, no ano de 2021, dos reservatórios de Barra Grande e Jirau, usinas localizadas nas bacias hidrográficas dos rios Uruguai e Amazonas, respectivamente.

Desconsiderando a comparação da ordem de grandeza das duas curvas de vazões, verifica-se que, enquanto a curva de afluência da primeira possui comportamento mais instável, com movimentos acentuados de subida e descida, a vazão de Jirau é marcada por características mais previsíveis, com tendência de alta no primeiro semestre e descida na segunda parte do ano.

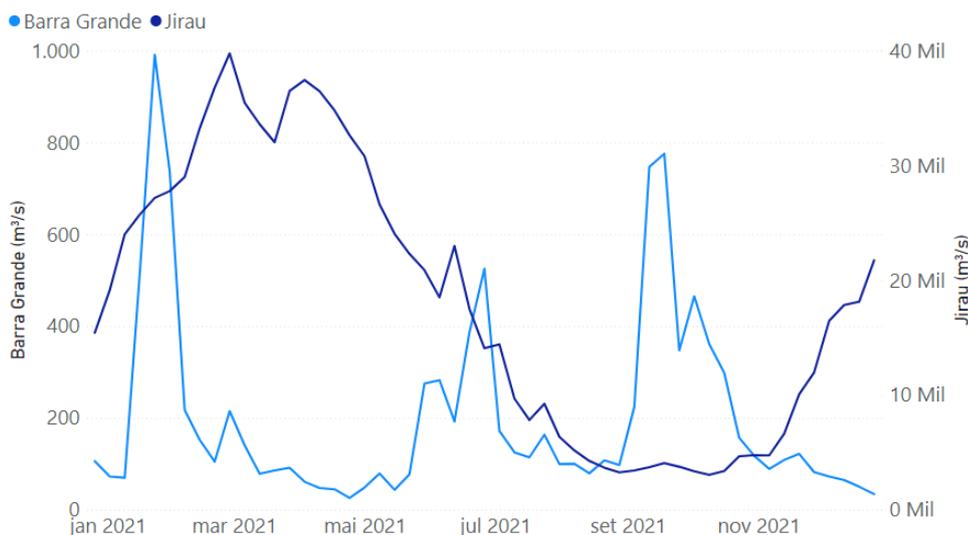


Figura 1- Vazão afluente de Barra Grande e Jirau. Fonte: Elaborado de [3].

Na prática, a vazão afluente é adotada como dado de entrada da programação da operação do SIN, atividade de responsabilidade do Operador Nacional do Sistema Elétrico (ONS). Nesse processo, o ONS adota três modelos computacionais com execução encadeada que visam a definição do despacho de usinas hidrelétricas e térmicas a custo operacional reduzido. Tratando especificamente do terceiro e último modelo na sequência encadeada, nomeado DESSEM e que será mais detalhado neste trabalho, a previsão de afluências é utilizada em resolução temporal diária, fornecendo ao programa perspectiva da situação hidrológica no país. A previsão de vazões, por sua vez, é determinada também pelo ONS, gerada a partir da adoção de modelos preditivos com diferentes algoritmos que tentam minimizar o valor do erro de previsão [4].

Diante desse cenário, pode-se inferir que, se a previsão feita é mais exata, ou seja, mais próxima do valor a ser realizado, mais as saídas do DESSEM se tornam alinhadas frente às reais condições do sistema e, assim, a programação da operação do SIN é também mais adequada, o que é importante, por exemplo, do ponto de vista da redução de encargos

decorrentes do despacho fora da ordem de mérito (aquele não definido na programação da operação e que serve para suprir a demanda existente).

O mesmo raciocínio pode ser utilizado para o processo de formação do preço *spot* do mercado de energia elétrica no Brasil, uma vez que a Câmara de Comercialização de Energia Elétrica – entidade que integra geradores, distribuidores, comercializadores e consumidores de energia no âmbito financeiro – utiliza a mesma sequência de programas encadeados mencionados anteriormente para determinar o valor do Preço de Liquidação das Diferenças (PLD) que vai nortear as operações no mercado de curto prazo (MCP). Dessa forma, quanto mais exata é a afluência prevista, mais o PLD é condizente com as condições do SIN [5].

A relação entre o PLD e a previsão de vazões (assim como a previsão de outros dados de entrada do DESSEM, como a demanda de energia) é importante para agentes comercializadores. Uma empresa que faz uso de um modelo predictor de vazões com resultados satisfatórios pode tentar antecipar o PLD a ser publicado pela CCEE, simulando a execução do deck de entrada oficial, posteriormente divulgado da entidade. Com base no PLD previsto, a empresa vai planejar sua estratégia de compra e venda de contratos de energia. Por exemplo, se o preço *spot* previsto for alto, a comercializadora pode ter como objetivo a formação de balanço positivo no MCP a partir da compra de contratos negociados a preços menores, fazendo com que a liquidação financeira seja maior do que o custo das compras executadas.

Visando a geração de previsões de afluências mais exatas, modelos de *Machine Learning* (ML) podem ser alternativas complementares aos programas utilizados no setor elétrico, baseados, por exemplo, em cálculos autorregressivos. Tomando o histórico de vazões de uma determinada usina como uma série temporal, é possível desenvolver um modelo com técnicas de Redes Neurais Artificiais, o que pode ser uma boa saída para lidar com o possível comportamento instável das vazões, como no caso de Barra Grande mostrado na Figura 1.

Na utilização de um modelo preditivo, uma análise necessária é relativa à distribuição de erros gerados no processo de previsão, indicando o desempenho da estrutura arquitetada. Nessa linha, a Tabela 1 apresenta valores de curtose e assimetria das distribuições de erros absolutos das previsões de vazões diárias de 10 hidrelétricas, feitas pelo ONS e considerando a previsão para um dia à frente. Nesse estudo, foram consideradas as previsões feitas de janeiro a julho de 2021. Quanto mais positivo o

coeficiente de curtose, mais a distribuição é pontiaguda, enquanto valores mais baixos representam o achatamento da função. Em relação à assimetria, valores positivos indicam que a cauda da direita é mais longa na distribuição, ao passo que o esticamento da cauda da esquerda está ligado a números negativos. Dos dados apresentados, é possível verificar que os valores de curtose e assimetria diferem daqueles pertencentes à curva gaussiana, podendo-se inferir, então, que as distribuições geradas possuem formato gaussiano não simétrico.

De forma convencional, o ajuste dos parâmetros de uma RNA é feito a partir de funções objetivo mais adequadas para a análise de distribuições de erro gaussianas, como o MSE (erro quadrático médio ou, em inglês, *mean squared error*). Como forma de otimizar a análise de distribuição não gaussianas, os conceitos de *Information Theoretic Learning* (ITL) podem ser aplicados, uma vez que estes mudam a visão sobre o formato da curva de erro, isto é, tenta-se minimizar o conteúdo informativo da mesma e não a sua variância (o que é feito ao se adotar o MSE) [6].

Tabela 1 - Curtose e assimetria das distribuições de erros de previsão de vazão para algumas usinas do SIN

Usina	Região Hidrográfica	Curtose	Assimetria
Mauá	Paraná	8,50	-2,62
Itá	Uruguai	6,97	-1,62
Ernestina	Atlântico Sul	7,55	2,07
Sobragi	Atlântico Sudeste	8,73	-1,98
Irapé	Atlântico Leste	4,57	-2,11
Queimado	São Francisco	4,53	-1,51
Boa Esperança	Parnaíba	0,57	0,39
Guaporé	Amazonas	7,22	-1,91
Serra da Mesa	Tocantins	4,04	-1,87
Manso	Paraguai	1,49	-0,58
Distribuição Gaussiana		0	0

1.1 Objetivos

O objetivo desse trabalho consiste na avaliação da aplicação do *framework* de *Information Theoretic Learning* no âmbito do cenário brasileiro de previsão hidrológica, com subsequente análise de impacto na formação do preço do mercado de curto prazo de energia elétrica no Brasil. Para tal, propõe-se o desenvolvimento de um modelo predictor de vazões diárias baseado em Redes Neurais Artificiais, com o uso dos conceitos de ITL, adotando-se dados históricos de aflúências e precipitação em diferentes regiões hidrográficas do país. Com a aplicação desse modelo para as hidrelétricas contempladas na programação da operação do SIN, os resultados gerados são utilizados como dados de entrada do DESSEM, de forma que sejam também produzidos valores de PLD horário com a posterior execução do programa. São avaliados os desempenhos tanto no âmbito hidrológico – com a comparação às previsões feitas pelo ONS e análise de exatidão das previsões em relação às aflúências realizadas – quanto no ponto de formação do preço *spot* horário, buscando-se o impacto das previsões de vazões no cálculo do PLD.

Como forma de avaliação do modelo proposto, o trabalho também apresenta outros dois predictores: uma rede neural com o MSE como função de erro aplicada e uma estrutura que gera previsões baseadas nas médias históricas de aflúências das UHEs, denominado modelo ingênuo.

Ao final, é feita uma análise comparativa, em termos de preço, entre o predictor proposto, o conjunto de predictores do ONS e o caso em que as vazões realizadas são inseridas como *inputs* do DESSEM (ou seja, o caso em que os preços calculados estão totalmente alinhados com o cenário real de aflúências).

1.2 Estrutura do Trabalho

O trabalho foi definido em capítulos da seguinte forma. O capítulo 1 introduz o problema a ser explorado ao longo deste trabalho. O capítulo 2 apresenta os conceitos referentes ao processo de formação do preço *spot* do mercado de energia elétrica no Brasil, com a descrição dos custos envolvidos e os modelos computacionais adotados, tanto para o cálculo do PLD, quanto para a previsão de vazões. No capítulo 3, é exposta a fundamentação teórica das Redes Neurais Artificiais e de *Information Theoretic Learning*, conceitos aplicados neste trabalho. O capítulo 4 consiste no mapeamento de produções

bibliográficas que serviram de base para este trabalho. O capítulo 5 descreve o projeto proposto, com o levantamento de dados utilizados, modelagem de estruturas e metodologia aplicada. O capítulo 6 apresenta os resultados obtidos, divididos conforme as análises hidrológicas e de formação do PLD. Por fim, o capítulo 7 pontua as conclusões tiradas do trabalho, com a sugestão de melhorias para projetos futuros.

Capítulo 2 - Formação do PLD Horário

Neste capítulo, é abordado o processo de formação do Preço de Liquidação das Diferenças em resolução horária, explicando-se sua utilização e pontuando ferramentas computacionais utilizadas no seu cálculo.

2.1 Preço de Liquidação das Diferenças

Antes do escopo financeiro do PLD, deve-se entender o significado por trás de seu valor e, para isso, é interessante a análise do Custo Marginal de Operação (CMO) gerado a partir da definição da programação operativa feita pelo ONS. O CMO é definido como o custo de produção de 1 MW necessário para atender determinado acréscimo de carga do mesmo montante. Ele equivale ao Custo Variável Unitário (CVU) – referente ao valor necessário para cobrir os gastos operacionais de um empreendimento gerador [7] – de uma determinada térmica designada para a geração do acréscimo de energia. Essa usina é determinada de acordo com o propósito de minimização dos custos operativos do sistema no atendimento à demanda, ou seja, primeiro são despachadas aquelas com gastos operativos mais baixos, seguindo a pilha de termelétricas ordenadas por CVU (como o exemplo mostrado na Figura 2 [8]), até que geração total seja equivalente à carga. Assim, determinada térmica na pilha torna-se responsável pela produção do MW extra.

Nessa ótica, o CMO é um possível indicador das condições encontradas no sistema e de como o ONS abordou o problema de despacho a custo minimizado. Por exemplo, no cenário de valor de CMO alto, muitas usinas térmicas são despachadas, o que pode estar relacionado a situação de níveis baixos dos reservatórios de hidrelétricas e anomalias negativas nos mapas de previsão de precipitação. Na perspectiva contrária, o Custo Marginal de Operação mais baixo pode representar a maior utilização da geração hidrelétrica na condição de níveis de reservatórios mais altos, fazendo com que menor número de termelétricas gerem por Ordem de Mérito (geração definida pelos modelos de otimização do despacho hidrotérmico).

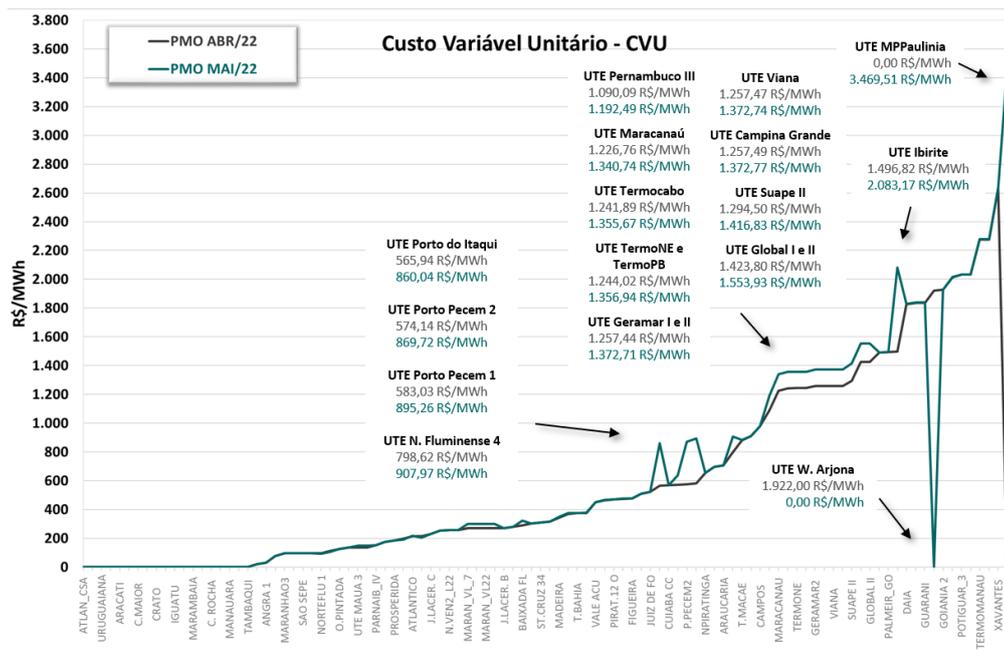


Figura 2 - Lista de usinas térmicas ordenadas por CVU. Fonte: [8].

Se o CMO é adotado no âmbito operativo, o PLD está relacionado ao escopo financeiro. Este é adotado para a valoração de exposições positivas ou negativas, decorrentes de operações de compra e venda de energia, no mercado de curto prazo, cuja contabilização é coordenada pela CCEE. Embora publicados por diferentes instituições, o CMO e o PLD são diretamente ligados, uma vez que a CCEE utiliza os mesmos modelos matemáticos usados pelo ONS para definição da programação da operação, com algumas diferenças entre os processos. Diariamente, o Operador encaminha à Câmara de Comercialização, após a execução do DESSEM, os decks de arquivos de entrada do modelo de curtíssimo prazo, além dos resultados obtidos. Posteriormente, a CCEE realiza tratamento dos arquivos enviados, retirando o efeito de restrições elétricas internas aos submercados de energia (regiões com fronteiras definidas segundo fluxos de energia e restrições elétricas), de forma que o valor de PLD seja o mesmo em todos os pontos de consumo de um mesmo submercado. Terminado o tratamento, a entidade executa o DESSEM, gerando valores de preço *spot* em resolução horária para o dia seguinte. Antes da publicação oficial, a CCEE ainda aplica limites mínimo, máximo horário e máximo estrutural (todos definidos pela Agência Nacional de Energia Elétrica – ANEEL) aos valores de PLD. O primeiro limite é referente ao maior valor entre a Tarifa de Energia de Otimização (TEO) – custo para cobrir operação e manutenção incrementais de hidrelétricas [9] – da usina de Itaipu e a TEO das demais hidrelétricas do SIN. O limite horário

determina o teto de preço para qualquer horário do dia ao longo de um ano completo, enquanto o caso máximo estrutural limita o PLD médio de um dia, também para o horizonte de um ano [10].

Além da aplicação no mercado de curto prazo, o Preço de Liquidação das Diferenças é adotado em outras atividades gerenciadas pela CCEE, como na valoração de encargos gerados, na apuração de eventuais penalidades de agentes, entre outras. A Figura 3 [10] apresenta o fluxograma de formação e aplicação do PLD.

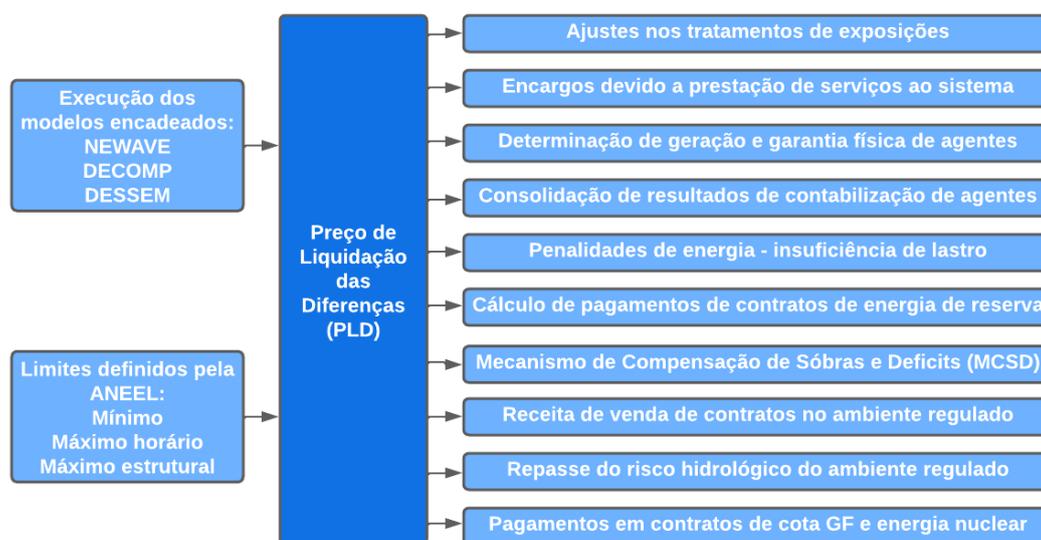


Figura 3 - Fluxograma de formação e uso do PLD. Fonte: Adaptado de [10].

2.2 Cadeia de Modelos Encadeados

Como pode ser visto na Figura 3, o Preço de Liquidação das Diferenças é resultante da execução de três modelos encadeados, ou seja, programas executados sequencialmente, onde o sucessor recebe, como dado de entrada, um arquivo de saída do antecessor. O primeiro da cadeia é o NEWAVE, executado uma vez ao mês pela CCEE. O DECOMP é o próximo modelo da sequência, cuja execução é semanal. Por fim, o DESSEM gera os valores de preço *spot* horário, diariamente.

Os três programas foram desenvolvidos pelo Centro de Pesquisas de Energia Elétrica (CEPEL) [11], sendo o modelo DESSEM implementado de forma oficial pelo ONS em janeiro de 2020 e pela CCEE no ano seguinte, dando origem ao chamado PLD horário, devido à resolução temporal do programa.

2.2.1 NEWAVE

Como é o primeiro modelo na sequência encadeada, o NEWAVE é o programa com maior horizonte de estudo, considerado como médio prazo. No planejamento da operação do SIN feito pelo ONS, o horizonte é de 5 anos, discretizado em estágios mensais. Basicamente, o modelo tenta otimizar o despacho de usinas hidrelétricas e termelétricas ao se minimizar o custo total de operação calculado durante o período completo de estudo.

Com a grande relevância da produção hidrelétrica na matriz energética brasileira, a forma como a água é utilizada ao longo do tempo ganha importância na medida que são avaliadas as consequências de se optar pela geração hidrelétrica no início do horizonte ou em estágios posteriores, via acoplamento temporal [12]. Por exemplo, no caso de uso desse tipo de geração no período inicial, os níveis de reservatórios são menores em estágios futuros, sendo necessário, se a hidrologia futura não for boa, o despacho maior de termelétricas para atendimento da carga. O estudo também leva em consideração a estocasticidade mencionada na introdução deste trabalho, com a aplicação de cenários hidrológicos probabilísticos. Essa análise pode ser vista na Figura 4 [13] de forma simplificada, apresentando o chamado “dilema do operador”.

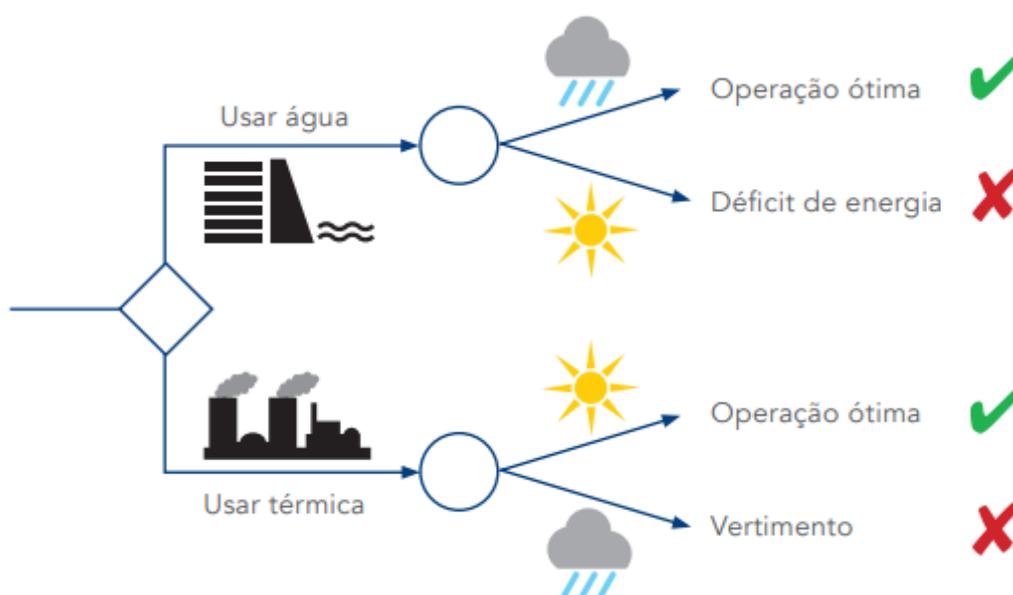


Figura 4 - Dilema do operador. Fonte: [13].

As possibilidades de tomada de decisão quanto ao uso da água e, consequentemente, ao acionamento de térmicas ao longo do horizonte são variáveis que impactam diretamente na função objetivo de custo total de operação, que se divide em duas parcelas: a Função de Custo Imediato (FCI) e a Função de Custo Futuro (FCF). A primeira consiste no cálculo do gasto operacional para o instante de tempo corrente, enquanto a segunda contempla os períodos à frente.

Quando relacionadas ao nível de reservatórios de hidrelétricas, as funções de custo têm comportamentos distintos. Por exemplo, com o aumento da água armazenada, a FCI também aumenta, uma vez que o cenário indica o encarecimento de gastos por conta do despacho maior de termelétricas. Já no caso da FCF, a elevação do nível dos reservatórios acarreta a redução de custos operacionais, pois mais água estaria disponível para geração hidrelétrica em estágios futuros [14].

Segundo [15], na otimização do despacho buscada pela execução do NEWAVE, a minimização do custo total de operação é igual ao mínimo valor de soma das parcelas FCI e FCF, o que ocorre quando as derivadas das duas funções, em relação ao nível de reservatório (operação denominada como valor da água), são iguais. Tal raciocínio pode ser visto na Figura 5 [14].

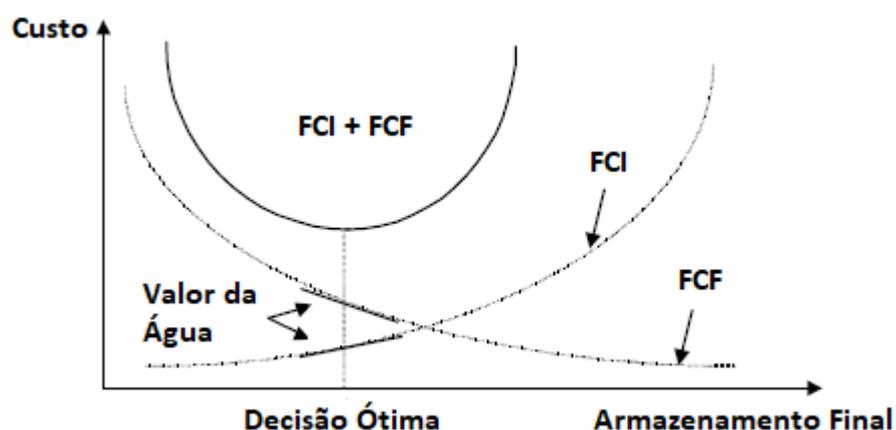


Figura 5 - Gráfico de custo por armazenamento final. Fonte: Adaptado de [14].

No NEWAVE, a definição da operação é definida com a aplicação da Programação Dinâmica Dual Estocástica (PDDE), com o trabalho de geração de múltiplos cenários probabilísticos de Energia Natural Afluyente (ENA) – definida como a energia elétrica que pode ser gerada a partir da vazão natural de um rio [16] – obtidos via série histórica de afluências ou modelo autorregressivo periódico.

Considerando o número de cenários de ENA, a quantidade de estágios mensais por todo o horizonte e o número de reservatórios de hidrelétricas presentes no SIN, a quantidade de variáveis seria enormemente elevada, o que inviabilizaria o processamento do NEWAVE. Como forma de evitar esse problema, chamado de maldição da dimensionalidade, o modelo utiliza Reservatórios Equivalentes de Energia (REEs), que são representações que agregam os reservatórios reais pelos 4 subsistemas existentes (Sudeste/Centro-Oeste, Sul, Nordeste e Norte). O modelo de despacho hidrotérmico de médio prazo trabalha atualmente com 12 REEs: Sudeste, Madeira, Teles Pires, Itaipu, Paraná, Paranapanema, Sul, Iguaçu, Nordeste, Norte, Belo Monte e Manaus. Outra simplificação é feita na modelagem das linhas de transmissão que escoam energia entre os subsistemas, a partir da adoção de conexões de intercâmbio equivalentes. A disposição de subsistemas, REEs e intercâmbios de energia pode ser analisada na Figura 6 [10].

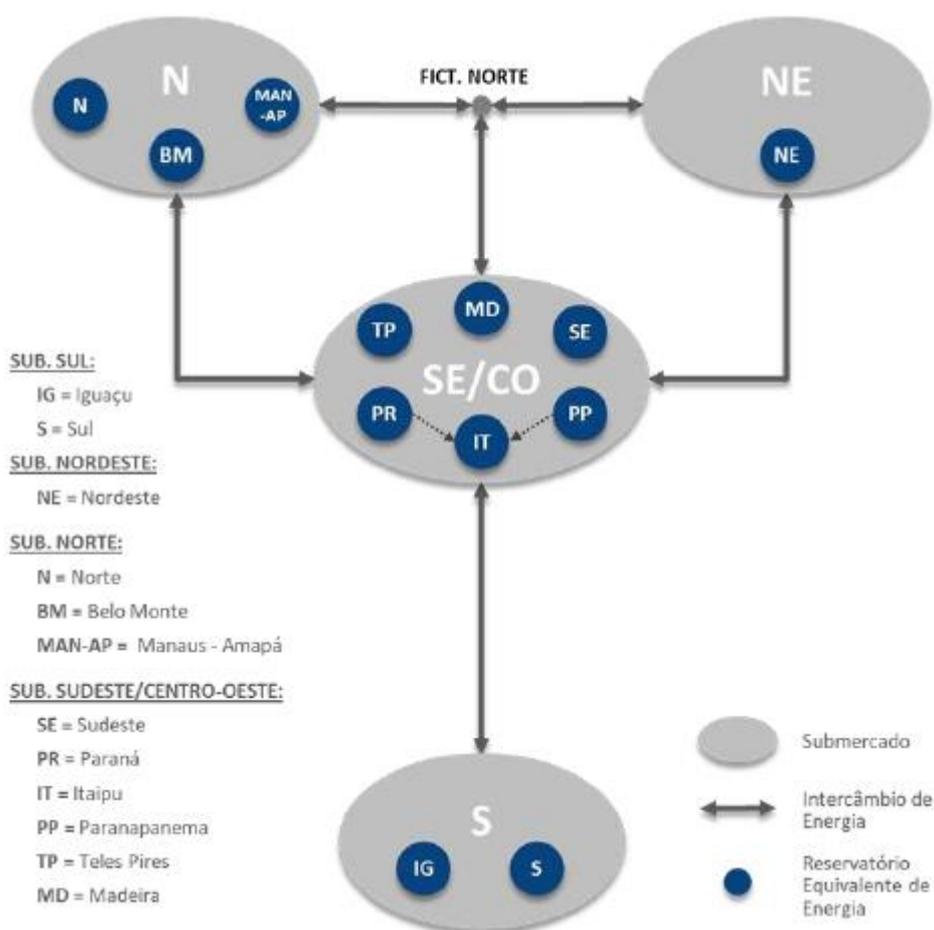


Figura 6 - Representação espacial adotada no NEWAVE. Fonte: [10].

Segundo [17], a otimização feita pelo NEWAVE é alcançada por uma sequência de passos. Após a definição dos volumes iniciais de cada REE, são escolhidos cenários aleatórios de ENA para todo o horizonte de estudo. Na etapa seguinte, é feita a simulação *forward*, baseada no cálculo das configurações do sistema (volumes armazenados, despacho hidrotérmico, intercâmbio entre subsistemas, entre outras) e na formação da FCF até o final do horizonte de estudo. Posteriormente, é aplicada a simulação *backward*, definindo a operação no sentido inverso com a busca pelos caminhos referentes aos menores custos futuros, definindo a FCF com maiores detalhes. Ao final, o processo iterativo é interrompido no caso de atendimento do critério de parada. Caso contrário, as iterações continuam até que seja o critério seja atendido.

Após a determinação do despacho hidrotérmico otimizado no médio prazo, são geradas as saídas do NEWAVE, entre elas, arquivo contendo a informação da função de custo futuro.

2.2.2 DECOMP

O DECOMP é o modelo de otimização do despacho hidrotérmico adotado para a programação da operação do SIN considerando o período de curto prazo. O ONS utiliza o programa com o horizonte de dois meses, sendo o primeiro mês discretizado em semanas. Seu acoplamento com o NEWAVE ocorre através do arquivo da FCF exportado do modelo de médio prazo.

Na prática, o DECOMP é executado pelo ONS semanalmente, mais precisamente no dia anterior ao primeiro dia da semana operativa seguinte que se inicia no sábado e vai até a sexta-feira. Antes de executar o modelo, o Operador Nacional do Sistema Elétrico faz a atualização dos arquivos de entrada do programa ao longo da semana, sendo necessário o envio de informações por parte dos agentes envolvidos na programação da operação, com prazo de envio até o dia útil que antecede a elaboração do deck de entrada do modelo [18]. Dentre as informações enviadas pelos agentes de geração, estão o cronograma de manutenções de unidades geradoras, armazenamentos previstos para o início da semana operativa, inflexibilidade de geração (valor mínimo que possa ser gerado), CVU por patamar de carga, restrições operativas, entre outras. Agentes de importação e exportação também devem informar dados ao ONS, tais como a disponibilidade líquida de suprimento e a inflexibilidade da interconexão, ambas por patamar de carga [19].

Para a resolução do problema do despacho hidrotérmico a custo minimizado, o DECOMP faz uso do algoritmo da PDDE, também utilizado no programa antecessor. Uma diferença na metodologia aplicada ao modelo de curto prazo é o nível de detalhamento do parque hidrelétrico. Se os REEs são adotados no NEWAVE, o DECOMP conta com a representação individual de usinas, cada uma caracterizada por uma lista de configurações: enchimento de volume morto, vazão deplecionada mínima, produtividade variável, tempo de viagem da vazão defluente dos aproveitamentos hidrelétricos, entre outras [20].

Outra diferença entre as metodologias aplicadas nos dois modelos apresentados é referente à representação estocástica da hidrologia do sistema. Segundo [21], o NEWAVE trabalha com combinações de cenários puramente probabilísticos, enquanto o DECOMP adota a chamada árvore de cenários, que pode ser vista na Figura 7 [20]. Nessa representação, as semanas do primeiro mês possuem dados de vazões determinísticos, enquanto o restante do horizonte contém a abertura de possibilidades estocásticas.

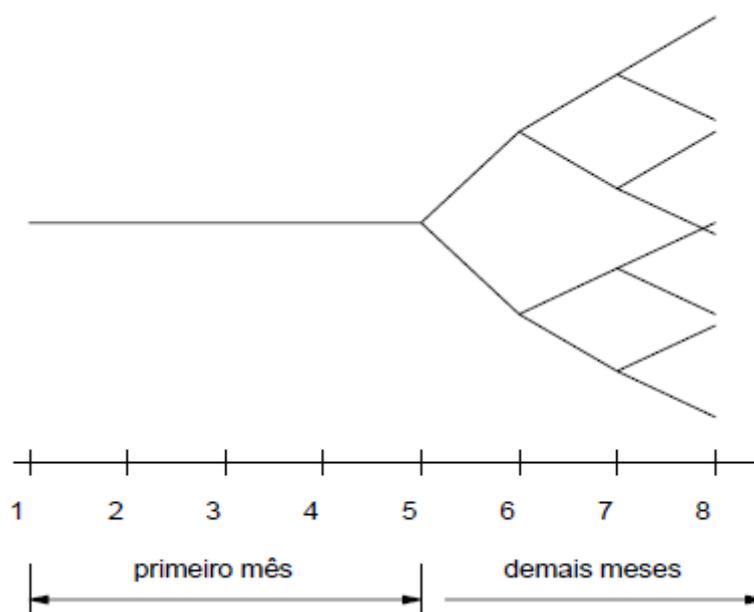


Figura 7 - Árvore de cenários hidrológicos do DECOMP. Fonte: [20].

Embora a árvore de cenários possua uma porção determinística em sua composição, essa parcela é determinada através de previsões de vazões naturais afluentes às usinas hidrelétricas do SIN, discretizadas em base semanal. As previsões são feitas pelo ONS (exceto para algumas usinas onde o próprio agente gerador fornece a informação) a partir do uso de um conjunto de modelos previsores, sendo a escolha do modelo a ser aplicado

dependente da localização do aproveitamento hidrelétrico em determinada bacia hidrográfica. Após a adoção de aflúências previstas, os cenários estocásticos sucessores da parte determinística são gerados por outro modelo matemático desenvolvido pelo CEPEL, chamado de GEVAZP. Segundo [22], o programa trabalha com a produção de cenários artificiais e estocásticos de vazões e energia, modelando a aflúência de um determinado mês como a combinação linear dos meses anteriores, além do uso de uma componente randômica.

Considerando a aplicação da PDDE, a sequência de passos de execução do DECOMP é semelhante ao que ocorre com o NEWAVE, com iterações de simulações *forward* e *backward* repetidas até que o critério de parada seja atendido. Assim como ocorre com o modelo de médio prazo, o DECOMP gera um arquivo de saída contendo a informação da FCF.

2.2.3 DESSEM

A recente adoção do DESSEM pelas entidades ONS e CCEE faz parte de um movimento de modernização do setor elétrico brasileiro, embasado na aprovação no Senado do Projeto de Lei 232/2016 e na Consulta Pública 033/2017 do Ministério de Minas e Energia (MME) [23] [24]. Com a mudança da resolução temporal do CMO e PLD, antes tratados em patamares de carga do DECOMP e depois em estágios semi-horários via DESSEM, enseja-se o aprimoramento na representação de custos operativos do sistema através da definição da programação da operação mais alinhada com as reais condições do sistema, reduzindo encargos de custos operacionais decorrentes de geração despachada fora da ordem de mérito de custo. A alteração também é justificada por outros fatores que têm modificado a estrutura do SIN, exigindo maiores complexidade e detalhamento na coordenação da operação do sistema, tais como o aumento da participação de fontes solar e eólica na matriz energética, novos eixos geradores a partir da geração distribuída, gerenciamento de carga pelo lado da demanda, entre outros. A Figura 8 mostra um exemplo comparativo entre os valores de preço *spot* gerados dos modelos DECOMP e DESSEM, no submercado Sudeste/Centro-Oeste e para o dia 31 de dezembro de 2020. É possível ver uma maior modulação na curva originada do segundo programa, explicada pelo estudo em menor resolução temporal.

Último na cadeia de modelos encadeados, o DESSEM é a plataforma de previsão considerada de curtíssimo prazo: horizonte de até duas semanas, sendo o primeiro dia dividido em estágios de 30 minutos e, para os dias seguintes, são utilizados períodos correspondentes aos níveis de patamares de carga [25]. Assim como ocorre entre NEWAVE e DECOMP, o programa é encadeado ao modelo de curto prazo através do arquivo de FCF adotado como seu dado de entrada.

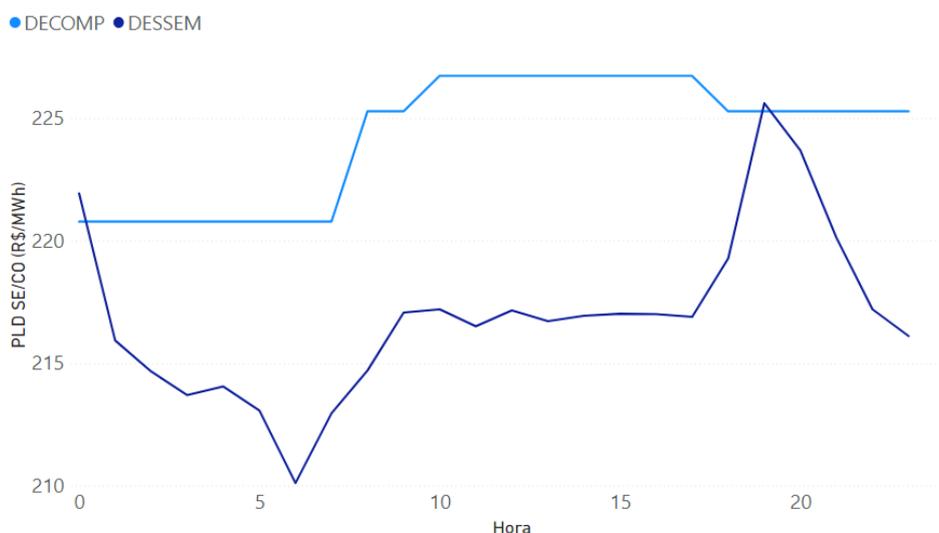


Figura 8 - PLD do Sudeste/Centro-Oeste no dia 31/12/2020. Fonte: Elaborado de [26].

Assim como no caso do DECOMP, o ONS conta com o envio de informações de cunho operativo dos agentes envolvidos para a montagem do deck de entrada do DESSEM. Segundo [27], o Operador Nacional do Sistema Elétrico executa o programa diariamente até às 16h00 do dia anterior ao dia a ser programado, com posterior envio à CCEE dos arquivos de entrada e resultados obtidos. Um conjunto de informações é obtido da execução do modelo: relação de termelétricas a serem despachadas por ordem de mérito de custo, proposta de geração termelétrica e de usinas intermitentes, proposta de intercâmbios entre subsistemas, entre outras [4].

Com o horizonte de estudo menor quando comparado aos outros modelos, o DESSEM garante maior nível de detalhamento das configurações do SIN. As usinas conseguem ser detalhadas ao nível de suas unidades geradoras, as produtibilidades variáveis de hidrelétricas em função das alturas de queda são detalhadas e o balanço de água nos reservatórios é projetado de forma acurada através de tempos de viagem de afluências e curvas de propagação. Em relação às usinas termelétricas, são consideradas as

restrições de *unit commitment* de unidades geradoras, além de operações a ciclo combinado. Já a rede elétrica é modelada a partir de fluxo de potência linearizado [28].

Outra diferença do DESSEM para os outros programas da cadeia encadeada é a forma de definição da otimização do despacho hidrotérmico. Enquanto NEWAVE e DECOMP adotam abordagens probabilísticas a partir da PDDE, o modelo de curtíssimo prazo ataca o problema com técnicas de Programação Linear Inteira-Mista (MILP). Nessa metodologia, é aplicado o método de resolução *branch and bound*, no qual são eliminados os cenários operativos oriundos de hipóteses inviáveis, garantindo eficiência na procura pelo caso ótimo [29]. No caso de inexistência de cenário otimizado que respeite todas as restrições operativas, o cenário final será aquele que tiver o menor número de violações, uma vez que estas são penalizadas de forma contundente na função objetivo a ser minimizada [28]. Na prática, se a solução encontrada não for viável, o ONS faz o relaxamento de restrições, como, por exemplo, a desconsideração do *unit commitment* de usinas termelétricas, realizando nova execução do modelo [4]. A Figura 9 [28] apresenta o processo de utilização do DESSEM. Após a definição do despacho hidrotérmico a custo de operação minimizado, são gerados um conjunto de arquivos de saída. Dentre esses, o arquivo PDO_SIST apresenta alguns resultados obtidos para os subsistemas, tais como o Custo Marginal de Operação, geração hidrelétrica e termelétrica, energia armazenada em reservatórios, entre outros.

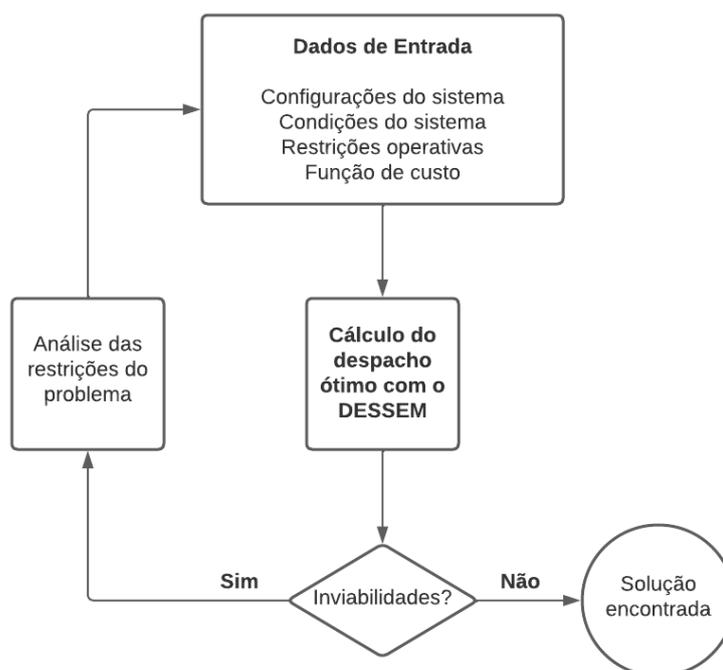


Figura 9 - Esquema de utilização do modelo DESSEM. Fonte: Adaptado de [28].

2.3 Previsão de Vazões

As próximas seções do capítulo abordam a forma como as aflúências previstas são utilizadas no programa de otimização de curtíssimo prazo, além dos programas adotados pelo ONS na formação dessas previsões.

2.3.1 Relação com Modelagem de Curtíssimo Prazo

Embora o DESSEM adote metodologia determinística para a definição do despacho hidrotérmico ótimo, os arquivos de entrada do programa contêm previsões de diversas variáveis que detalham as condições do sistema, tais como a carga, gerações não simuladas (eólica e solar, por exemplo) e a vazão afluyente a usinas hidrelétricas. No caso das aflúências previstas, de acordo com [30], a informação é produzida pelo ONS a partir de um conjunto de modelos preditivos, que serão mais bem descritos na seção 2.3.2.

Antes da geração das vazões previstas, os agentes geradores devem enviar diariamente ao ONS informações dos aproveitamentos que subsidiam a consolidação de aflúências performadas, usadas nos processos preditivos. Além disso, devem enviar também as previsões de vazões naturais médias diárias, nos locais de interesse onde o Operador não aplique modelo preditivo [31].

Em relação ao DESSEM, o modelo de despacho hidrotérmico de curtíssimo prazo tem o arquivo DADVAZ como uma de suas entradas que, por sua vez, é preenchido com previsões de vazões médias diárias de hidrelétricas do SIN. Ao todo, 162 usinas são contempladas no arquivo, com a possível aplicação de três diferentes tipos de vazões: incremental, total e regularizada [28]. Atualmente, nos arquivos DADVAZ presentes nos decks publicados pelo ONS, tem-se optado pela aflúência incremental, conceitualmente definida em (1) [32]. Na equação, Q_{afl} é a vazão afluyente ao reservatório da usina, Q_{defmp} é a vazão defluyente dos reservatórios a montante, Q_{uso} é a aflúência relativa aos usos consuntivos da bacia e Q_{evp} a vazão referente à evaporação do reservatório. Todas as vazões são consideradas em m^3/s .

$$Q_{inc} = Q_{afl} - Q_{defmp} + Q_{uso} + Q_{evp} \quad (1)$$

A Figura 10 [33] apresenta uma amostra do arquivo DADVAZ, com parte do bloco de vazões contendo informações indentadas em seis colunas específicas. Observando da esquerda para a direita na imagem, estão os dados referentes ao número de identificação da hidrelétrica, seu nome, o *flag* para identificação do tipo de vazão utilizada, o dia do mês do estágio inicial, o dia do mês do estágio final e a vazão prevista. Adicionalmente, a letra “F” é usada para definir que o estágio inicial e final são iguais.

No arquivo DADVAZ, a quantidade de dias previstos é variável, mas com o horizonte de previsão sempre terminando na sexta-feira da semana operativa na qual é feita a execução do deck do DESSEM. O horizonte máximo é feito de 7 dias, no caso em que o primeiro dia previsto é sábado e o deck é divulgado na sexta-feira anterior. No caso do horizonte mínimo, tem-se apenas 1 dia projetado, a sexta-feira, com o deck divulgado na quinta-feira anterior.

```

Hr  Dd  Mm  Ano
XX  XX  XX  XXXX
00  01  05  2022
Dia inic(1-SAB...7-SEX); sem da FCF; n. semanas; pre-interesse
X X X X
2 1 1 0
VAZOES DIARIAS PARA CADA USINA (m3/s)
NUM      NOME      itp  DI HI M DF HF M      VAZAO
XXX XXXXXXXXXXXX  X    XXxXXxXxXXxXXxX    XXXXXXXXX
  1 CAMARGOS      1    01          F          78
  1 CAMARGOS      1    02          F          77
  1 CAMARGOS      1    03          F          76
  1 CAMARGOS      1    04          F          75
  1 CAMARGOS      1    06          F          74
  2 ITUTINGA      1    01          F           0
  4 FUNIL-GRANDE  1    01          F          98
  4 FUNIL-GRANDE  1    02          F          97
  4 FUNIL-GRANDE  1    03          F          96
  4 FUNIL-GRANDE  1    04          F          95
  4 FUNIL-GRANDE  1    06          F          96
 24 EMBORCACAO    1    01          F         186
 24 EMBORCACAO    1    02          F         184
 24 EMBORCACAO    1    03          F         182
 24 EMBORCACAO    1    04          F         180
 24 EMBORCACAO    1    05          F         179
 24 EMBORCACAO    1    06          F         178
 27 CAPIM BRANC1  1    01          F           1
 28 CAPIM BRANC2  1    01          F           2

```

Figura 10 - Arquivo DADVAZ. Fonte: [33].

2.3.2 Modelos Utilizados pelo ONS

Como mencionado em 2.3.1, o Operador Nacional do Sistema Elétrico utiliza um conjunto de modelos para a geração de previsões de vazões afluentes, sendo as afluências produzidas em bases mensal, semanal ou diária. Especificamente na perspectiva diária, o Operador conta com três principais ferramentas computacionais, além de alternativas, tais como a opção pela média das últimas vazões, extrapolação linear, regressão com base nas afluências de outros aproveitamentos e previsão feita pelo próprio agente gerador.

O principal modelo utilizado, isto é, aquele aplicado para a maioria das bacias do SIN é o SMAP, cujo horizonte de previsão vai até 12 dias à frente. Com análise de dados de vazões de precipitação, sua metodologia se baseia na transformação da chuva em vazão, adotando coeficientes representativos de precipitação e evapotranspiração e representações de quatro reservatórios lineares hipotéticos: reservatório do solo, de escoamento superficial, de escoamento subterrâneo e de planície. Os dados de entrada do SMAP são os totais diários de precipitação previstos e observados, os totais climatológicos diários de evapotranspiração potencial e as vazões médias diárias performadas. A escolha de aplicação do modelo para a maioria dos aproveitamentos hidrelétricos se dá por diversos motivos: possibilidade de uso para horizontes superiores a uma semana, facilidade de entendimento de sua metodologia, o potencial de aprimoramento vislumbrado pelo ONS, a adaptabilidade em relação a bacias de comportamentos hidrológicos diferentes, o código da ferramenta ser aberto, entre outros [34].

Outro modelo utilizado para previsão de afluências diárias é o PREVIVAZH. Segundo [25], a ferramenta gera vazões diárias previstas até 14 dias à frente a partir da desagregação da previsão semanal, usando a tendência inferida das últimas vazões performadas e a ponderação de séries sintéticas diárias de vazões naturais. A previsão semanal, por sua vez, é produzida pelo programa PREVIVAZ. A ferramenta conta com 94 tipos de modelos estatísticos que podem ser divididos em estacionários (por exemplo, média semanal, média anual, modelos autorregressivos AR e ARMA, entre outros), periódicos (modelos autorregressivos PAR e PARMA) e médias semanal e anual. Após estimação e validação dos parâmetros, é feita a previsão de afluências com o modelo que, dentre todas as opções, obteve o menor desvio padrão do erro de previsão. Diferentemente do modelo SMAP, o PREVIVAZH e o PREVIVAZ não utilizam dados de precipitação.

O terceiro modelo adotado é o CPINS, aplicado para previsão de afluências incrementais e naturais diárias do reservatório de Sobradinho para o horizonte de até 30 dias. Segundo [35], a metodologia consiste na estimação da vazão incremental em diversos trechos do rio, levando em consideração taxas de recessão ajustadas para diversos trechos do rio, além de valores arbitrados a partir das condições meteorológicas observadas e de informações hidrométricas dos principais afluentes do rio São Francisco.

Semanalmente, o ONS divulga um relatório do processo de consistência da previsão de vazões diárias, publicando arquivo em formato Excel com os modelos utilizados para cada uma das usinas hidrelétricas consideradas no processo de programação da operação do SIN, além das afluências previstas.

Capítulo 3 - Conceitos de RNA e ITL

Este capítulo aborda os conceitos de Redes Neurais Artificiais e *Information Theoretic Learning*, que foram os pilares na modelagem do previsor proposto neste trabalho.

3.1 Redes Neurais Artificiais

Técnicas de redes neurais são empregadas em modelos de *Machine Learning* (ML) para resolução de diferentes problemas, entre eles, a previsão de valores usando séries temporais, exercício proposto neste trabalho. As próximas subseções abordam os conceitos de RNAs que sustentaram a modelagem do previsor de afluências desenvolvido.

3.1.1 Conceitos Básicos de RNAs

Uma rede neural artificial pode ser compreendida como um processador massivo composto por unidades menores distribuídas de forma paralela. Essas unidades têm o potencial de ganho de conhecimento através de atividade experimental com determinado meio, com possível utilização posterior de tal informação. Como a palavra “neural” sugere, a estrutura pode ser comparada ao sistema nervoso de um organismo. No funcionamento biológico, tem-se a transmissão de dados em sentidos distintos, como exposto na representação em diagrama de blocos da Figura 11. As setas da esquerda para a direita sinalizam o envio de informações que chegam a uma resposta final, enquanto as setas da direita para a esquerda refletem a retroalimentação do sistema a partir de estímulo externo, com receptores e atuadores como estruturas que auxiliam a conversão de dados [36]. Na perspectiva da RNA, a transmissão de conteúdo informativo também é feita em dois sentidos, com a geração de saídas provenientes dos cálculos feitos no interior da estrutura e com a retroalimentação de dados que provocam a atualização de parâmetros internos.

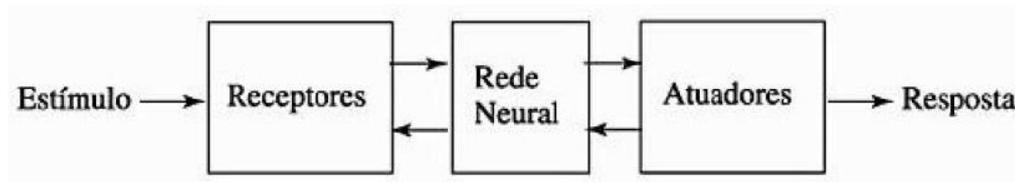


Figura 11 - Diagrama de blocos do sistema nervoso. Fonte: Adaptado de [36].

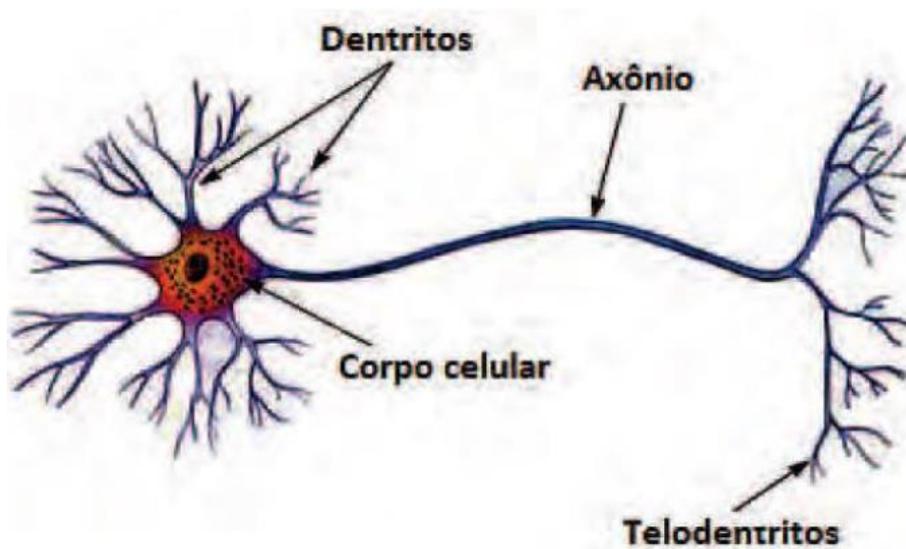


Figura 12 - Neurônio biológico. Fonte: [37].

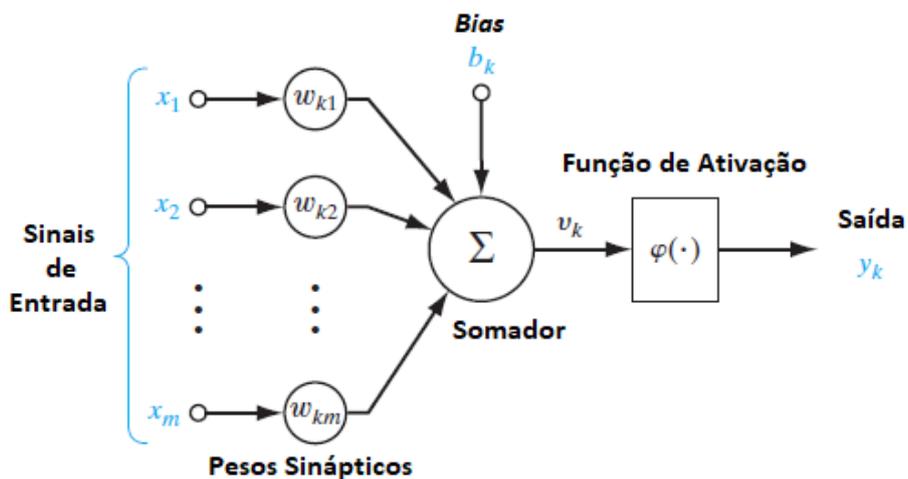


Figura 13 - Neurônio artificial. Fonte: Adaptado de [36].

Outra comparação que pode ser feita refere-se às unidades mínimas da rede neural artificial e do sistema nervoso: os neurônios. As células biológicas são altamente estimuláveis, processando e transmitindo a informação através de sinais eletroquímicos. Como apresentado na Figura 12 [37], o neurônio é composto por um corpo celular, onde

ocorre a síntese proteica; dendritos, que recebem estímulos oriundos de outros neurônios; o axônio, prolongamento que conduz os impulsos nervosos provenientes do corpo celular; e os telodendritos, ramificações capazes de estabelecer comunicação com dendritos de outros neurônios [38]. Por sua vez, a célula artificial também é definida por componentes responsáveis por receber a informação à montante do neurônio, processar os dados recebidos e transmitir os sinais gerados para as estruturas à jusante. Da Figura 13 [36], é possível elencar esses componentes, começando pelos elos de conexão que trazem sinais de entrada para respectivos pesos sinápticos (de valores positivos ou negativos). O somador é a peça intermediária responsável por agregar os dados ponderados dos elos antecessores (com a possível adoção de um elemento *bias* como sinal não externo ao neurônio). Por fim, a função de ativação é encarregada de filtrar o valor resultante do somador e gerar o sinal final do neurônio, que pode ser o *output* de uma RNA ou *input* de outra célula artificial à frente.

A partir da representação gráfica da Figura 13 pode ser formulada a equação matemática que relaciona os sinais de entrada x_j , os pesos sinápticos w_{kj} e o elemento *bias* b_k com a saída do neurônio y_k , como evidenciado em (2). A função de ativação $\varphi(\cdot)$ presente na equação é um parâmetro que deve ser configurado pelo usuário da rede neural, de forma que as saídas sejam otimizadas. Dentre as fórmulas mais utilizadas, vale destacar a função de limiar, função de limiar por partes e a função sigmoide (com o termo a sendo o parâmetro de inclinação da porção não linear da curva), expostas em (3), (4), (5) e na Figura 14.

$$y_k = \varphi\left(b_k + \sum_{j=1}^m w_{kj}x_j\right) \quad (2)$$

$$\varphi\left(b_k + \sum_{j=1}^m w_{kj}x_j\right) = \begin{cases} 1, & \text{se } b_k + \sum_{j=1}^m w_{kj}x_j \geq 0 \\ 0, & \text{se } b_k + \sum_{j=1}^m w_{kj}x_j < 0 \end{cases} \quad (3)$$

$$\varphi\left(b_k + \sum_{j=1}^m w_{kj}x_j\right) = \begin{cases} 1, & \text{se } b_k + \sum_{j=1}^m w_{kj}x_j \geq 0,5 \\ 0,5 + b_k + \sum_{j=1}^m w_{kj}x_j, & \text{se } -0,5 < b_k + \sum_{j=1}^m w_{kj}x_j < 0,5 \\ 0, & \text{se } b_k + \sum_{j=1}^m w_{kj}x_j \leq -0,5 \end{cases} \quad (4)$$

$$\varphi\left(b_k + \sum_{j=1}^m w_{kj}x_j\right) = \frac{1}{1 + \exp[-a(b_k + \sum_{j=1}^m w_{kj}x_j)]} \quad (5)$$

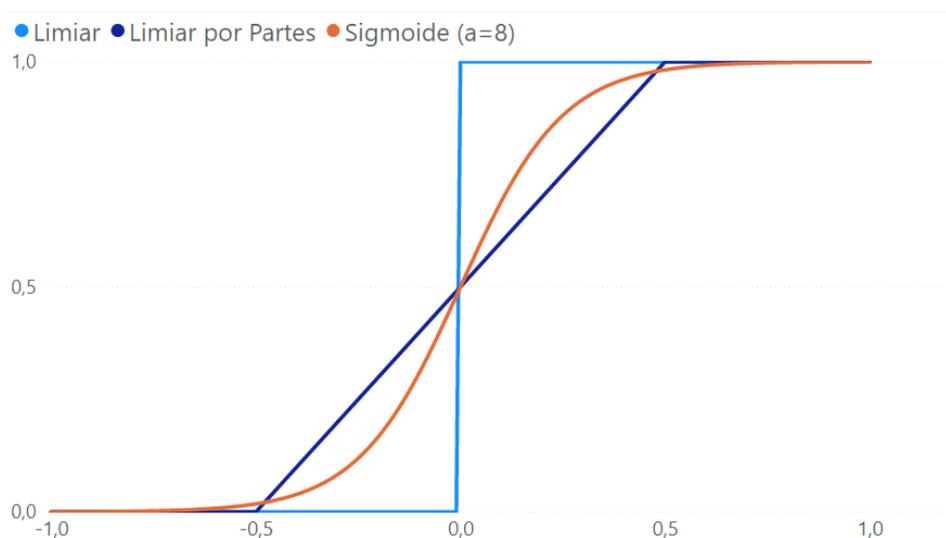


Figura 14 - Funções de ativação limiar, limiar por partes e sigmoide.

Outro ajuste a ser configurado pelo usuário da rede neural diz respeito à organização de seus neurônios. Segundo [36], existem três tipos de arquiteturas RNAs essencialmente distintas entre si. O primeiro modelo é a rede *feed-forward* de camada única (Figura 15a), opção onde cada neurônio é responsável por uma saída da rede. O segundo é o modelo *feed-forward* de múltiplas camadas (Figura 15b), alternativa na qual os neurônios de uma camada geram saídas para unidades de outra camada à frente. Nesse caso, a inserção de uma ou mais camadas ocultas (aquelas dispostas entre as camadas de entrada e saída) garante maior número de conexões à rede, fazendo com que ela consiga analisar dados estatísticos mais complexos, isto é, de ordem mais elevada. A rede neural recorrente é o terceiro tipo elencado (Figura 15c), com estrutura mais complexa que resulta

no profundo impacto na capacidade de aprendizado da rede. Neste modelo, os neurônios de uma camada podem gerar saídas para outros presentes na mesma camada ou em antecessoras.

Outra forma de se caracterizar uma RNA consiste na função destinada a ela. Nessa perspectiva, as redes neurais artificiais podem ser divididas em modelos de classificação e regressão. No primeiro caso, a rede é aplicada para definir uma linha de agrupamento das saídas geradas a partir de amostras de entrada utilizadas, ou seja, o modelo é treinado para o reconhecimento de padrões. No segundo caso, foco deste trabalho, a saída produzida não é clusterizada como a situação anterior, podendo assumir qualquer valor e o objetivo é aproximá-la da saída real. Tanto para classificação quanto para regressão, a estrutura funciona como uma função g dependente dos dados de entrada x e pesos w que, somada a um termo de erro ε , equivale a saída desejada y^{real} , assim como exposto em (6). Nessa perspectiva, a análise de impacto dos dados de entrada na saída real é uma boa prática da modelagem de RNAs, uma vez que a rede pode aumentar sua capacidade de extração de informação estatística se a magnitude da correlação entre *inputs* e *outputs* for alta.

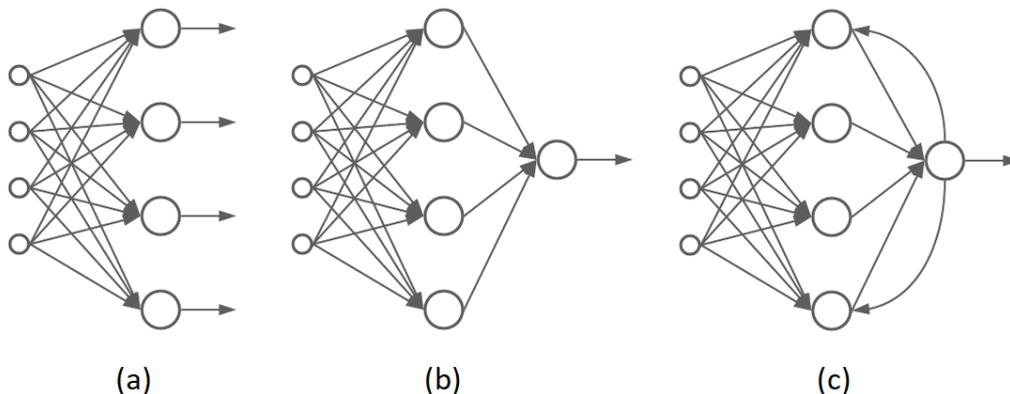


Figura 15 - Arquiteturas de redes neurais artificiais.

$$g(w, x) + \varepsilon = y^{real} \quad (6)$$

As redes neurais artificiais são tratadas como modelos de ML devido ao processo de aprendizado realizado pela estrutura, resultando na geração de saídas adequadas através da generalização de padrões analisados. De acordo com [36], certas propriedades caracterizam as RNAs e algumas delas merecem menção. A primeira é a não-linearidade a partir da utilização de neurônios não-lineares, o que é uma importante configuração no

estudo do comportamento de ambientes mais instáveis. A adaptabilidade é a segunda propriedade, considerando que redes neurais podem passar por consecutivos processos de aprendizagem, com consequente atualização de parâmetros internos, na medida que o meio estudado sofre alterações. Sobre esse ponto, o melhor cenário seria a calibragem do modelo de forma que ele seja adequadamente adaptado ao meio com perturbações significativas e, ao mesmo tempo, desconsidere alterações de menor relevância. Por fim, o terceiro traço destacado das RNAs é o potencial de extração de conhecimento dos dados recebidos. No caso da análise de uma série temporal, por exemplo, uma rede neural bem modelada pode assimilar características de sazonalidade, periodicidade e tendência.

3.1.2 Redes do Tipo Perceptron

A versão mais simplificada de uma rede neural trata-se do modelo *Perceptron*, arquitetado com apenas 1 neurônio com função de ativação limiar. Essa estrutura pode ser utilizada, por exemplo, como um classificador binário, cujo funcionamento correto é condicionado pela separação linear entre as classes de padrões estudadas.

Caso necessário o estudo de um meio não linear, a adoção do *Multilayer Perceptron* (MLP) torna-se adequada devido à sua maior complexidade e, conseqüentemente, sua maior capacidade de análise, em comparação ao *Perceptron*. O MLP tem a mesma estrutura mostrada na Figura 15b, com camadas de neurônios associadas de forma sequencial, sem conexões de retroalimentação. De acordo com [36], o modelo pode ser definido em três pontos principais. Primeiro, o fato de cada neurônio da rede possui função de ativação não linear e diferenciável. Além disso, o MLP contém uma ou mais camadas ocultas, ou seja, que não tem contato direto com os nós de entrada e saída da RNA. Por fim, o modelo exibe alto nível de conectividade determinada por seus pesos sinápticos.

O processo de aprendizado da rede neural artificial, etapa na qual ocorre a atualização dos parâmetros internos da estrutura, depende da adoção de uma metodologia devidamente ajustada que seja capaz de garantir a convergência das saídas geradas na direção do espectro de saídas desejadas, evitando casos de *overfitting* (problema de generalização causado pelo treino demasiado) e *underfitting* (má performance da rede, com erros consideráveis). O algoritmo de retropropagação é a opção mais adotada para a aprendizagem de MLPs. O método se baseia em processo iterativo, em que cada iteração,

chamada de época, é dividida em duas etapas. No primeiro passo, a rede recebe vetor de amostras de entrada que ocasiona a geração de vetor de saídas calculadas. No segundo passo, a partir do vetor de diferença entre saídas geradas e reais, propaga-se o sinal de erro na direção contrária, com a consequente atualização dos valores dos pesos sinápticos através do cálculo de gradientes locais.

A atualização dos pesos w_{ji} é apresentada na equação (7), no qual i e j equivalem, respectivamente, à sinapse e ao neurônio associados ao peso, n é a época corrente, η é a taxa de aprendizado e E corresponde ao erro obtido. Vale ressaltar o sinal negativo do gradiente ponderado pela taxa de aprendizado, mostrando que o valor é decrescido do peso sináptico no instante corrente, configurando-se a técnica de descida do gradiente [36]. Presentes na equação, a taxa de aprendizado, o número de épocas e a função de erro são parâmetros da rede estabelecidos por seu usuário e impactam diretamente no desempenho do modelo. No caso da taxa de aprendizado, um valor muito baixo pode tornar o processo de aprendizagem muito lento, ao passo que uma taxa muito alta pode causar instabilidade na etapa de treino [37]. Quanto ao número de épocas, o *underfitting* pode ser derivado de uma quantidade pequena de épocas, enquanto o *overfitting* pode ocorrer por conta de um número exagerado de épocas. Em relação à função de erro, utilizada como função objetivo a ser minimizada, a escolha de uma opção adequada pode ajudar na convergência do processo de aproximação entre saídas geradas e reais. Dentre as alternativas mais utilizadas nos casos de regressão, estão o MSE, o erro absoluto médio (em inglês, *mean absolute error* ou MAE) e o erro percentual absoluto médio (*mean absolute percentage error*, MAPE), apresentados em (8), (9) e (10), respectivamente, com N equivalente ao número de amostras utilizadas na rede neural.

$$w_{ji}(n + 1) = w_{ji}(n) - \eta \frac{\delta E(n)}{\delta w_{ji}(n)} \quad (7)$$

$$MSE(Y^{previsto}, Y^{real}) = \frac{1}{N} \sum_{k=1}^N (y_k^{previsto} - y_k^{real})^2 \quad (8)$$

$$MAE(Y^{previsto}, Y^{real}) = \frac{1}{N} \sum_{k=1}^N |y_k^{previsto} - y_k^{real}| \quad (9)$$

$$MAPE(Y^{previsto}, Y^{real}) = \frac{100}{N} \sum_{k=1}^N \left| \frac{y_k^{previsto} - y_k^{real}}{y_k^{real}} \right| \quad (10)$$

Além do ajuste de parâmetros citados, costuma-se aplicar o método de validação cruzada, que consiste na divisão de amostras de entrada (e suas respectivas saídas) em conjuntos menores destinados às etapas de treino, validação e teste. Na etapa de treino, o primeiro conjunto é utilizado para o processo de aprendizagem da rede, com a atualização dos valores dos pesos sinápticos a partir de algoritmo de treino aplicado. O segundo grupo é adotado para a validação da rede modelada no passo anterior, sem qualquer alteração dos parâmetros da RNA. Por fim, o último grupo é usado para teste da rede, com a avaliação final da estrutura. No caso de uma rede bem modelada, os erros das três etapas são próximos e baixos, indicando que, com a aplicação de novas entradas, as saídas geradas têm boas chances de serem próximas dos *outputs* esperados.

3.2 Information Theoretic Learning

Como visto na equação (6), o termo de erro é o que separa a saída gerada por uma rede neural da saída real e, assim, é desejável que esse fator seja reduzido para que o modelo melhore seu desempenho. Estudos em *Information Theoretic Learning*, foco das próximas subseções, apresentam uma abordagem não convencional para o tratamento desse erro.

3.2.1 Conceitos Básicos de ITL

Com o desenvolvimento expressivo de métodos de *Machine Learning* para resolução de problemas, estruturas cada vez mais robustas tomam lugar de modelos lineares convencionais. Tal robustez passa por dois pontos principais: o alto desempenho no processamento de dados e o uso de estatística de mais alta ordem para análise específica da informação estudada. No escopo deste segundo fator se encontram os conceitos de ITL, baseados na aplicação de noções estatísticas derivadas do campo da Teoria da Informação (TI) [39].

Uma vez que modelos de ML dependem da etapa de treinamento para ajuste de parâmetros internos, a escolha pela aplicação de técnicas de ITL passa pela substituição na forma de se analisar a distribuição do erro calculado durante esse processo. Nessa perspectiva, ao invés da minimização da variância da curva (efetuada através da aplicação da função de erro médio quadrático), deseja-se a redução de seu conteúdo informativo, a ponto de que parâmetros de um modelo treinado não possam mais extrair dados relevantes dos *inputs*. A justificativa para essa mudança está no fato de que, segundo [40], a aplicação do MSE é ótima apenas quando a distribuição de erro é gaussiana. Visando os casos de funções não gaussianas, conceitos de TI são aplicados para modelagem de funções de erro que possam substituir o MSE.

Os conceitos do campo da Teoria da Informação foram primeiramente aplicados para o projeto de sistemas de comunicações, gerando impacto em questões de eficiência e confiabilidade. Com base nos conceitos de TI, foi possível definir a melhor (mínima) informação possível de um dado transmitido e qual a máxima quantidade de conteúdo a ser transferida por um canal de comunicação específico. Entretanto, as noções de TI, por si só, não são suficientes no tratamento estatístico de funções de distribuição, necessário em modelos de ML. O *framework* de ITL, desenvolvido por José Carlos Príncipe [41], serve para suprir essa necessidade, usando a otimização de funções baseadas em entidades estatísticas do campo da Teoria da Informação para o ajuste de parâmetros de uma estrutura de *Machine Learning* [42].

A principal entidade estatística derivada da Teoria de Informação foi originada do trabalho escrito por Claude Elwood Shannon, em meados do século XX, no qual o autor discorre, no campo dos sistemas de comunicações, sobre fontes de dados de origens discreta e contínua, especificando processos de transmissão com e sem ruído. Em seu trabalho, Shannon apresenta a equação da entropia, mostrada em (11), definida como uma medida de informação, escolha ou incerteza [43]. Na fórmula da entropia H , $p_X(x)$ é a função massa de probabilidade (FMP) de uma variável aleatória X e K é um parâmetro multiplicativo definido por Shannon. De tal equação, é possível definir algumas propriedades como, por exemplo, os casos de entropias nula e máxima. A entropia nula refere-se ao caso em que uma variável aleatória é definida por um único evento de probabilidade unitária, ou seja, não há incerteza alguma sobre possíveis resultados. O cenário inverso ilustra o caso de entropia máxima, no qual existem múltiplas possibilidades

de eventos, todos com a mesma probabilidade de ocorrência, de uma variável aleatória (condição de FMP uniforme).

Em meados da década de 1950, Alfred Rényi formulou uma generalização matemática para a entropia de Shannon, apresentada em (12), comumente chamada de entropia de Rényi [44]. Segundo [39], enquanto a fórmula de Shannon era mais adotada na área de comunicações, o cálculo de Rényi ganhou destaque em outras áreas, como a teoria de códigos, mecânica quântica, sistemas de dinâmica caótica e como medida de diversidade em economia. Nos conceitos de ITL, destaca-se a entropia quadrática de Rényi, alcançada quando α é igual a 2, como pode ser visto em (13). A versão contínua dela é apresentada em (14). Essa versão quadrática ganha importância devido ao fato de ser uma função contínua, não-paramétrica e conveniente no cálculo envolvendo o método de estimação da função de densidade e probabilidade (em inglês, *probability density function* ou PDF), que será abordado a seguir [39].

$$H(x) = -K \sum_{x \in X} p_X(x) \log p_X(x) \quad (11)$$

$$H_{X\alpha}(x) = \frac{1}{1-\alpha} \log \sum_{x \in X} p_X^\alpha(x) \quad \text{com } \alpha > 0, \alpha \neq 1 \quad (12)$$

$$H_{X2}(x) = -\log \sum_{x \in X} p_X^2(x) \quad (13)$$

$$H_{X2}(x) = -\log \int_{-\infty}^{+\infty} f(x)^2 dx \quad (14)$$

Como visto na equação (13), a entropia de Rényi é resultado da soma dos quadrados das probabilidades de eventos de uma variável aleatória. Um obstáculo surge do fato de que, em casos práticos de *Machine Learning*, a PDF relativa aos dados de entrada não é conhecida. Uma maneira de contornar esse problema é através da adoção de um estimador da função de densidade e probabilidade, alternativa que apresenta diversos métodos de aplicação. Neste trabalho, o método não-paramétrico de estimação pela janela de Parzen será abordado. Segundo [45], a metodologia consiste no uso de uma função de

kernel que possa representar as contribuições de cada amostra de uma variável aleatória, sendo a PDF definida pela soma dessas contribuições, como visto em (15). Na equação, N é o número de amostras da variável aleatória $X = [x_1, x_2, \dots, x_N]$, $\phi(\cdot)$ é a função de kernel e h é o parâmetro de largura da janela. Em [45], Parzen mostrou que, se ϕ e h forem escolhidos de forma adequada, \hat{f} converge para a real função de densidade e probabilidade. No caso do kernel aplicado para um evento específico, a função deve estar centrada na amostra, de modo que a largura da janela considere minimamente a vizinhança do domínio ao redor da amostra analisada. As opções mais utilizadas como kernel são as funções uniforme e gaussiana, mostradas em (16) e (17), respectivamente. Dentre as duas, é preferível a utilização da segunda (como será feito neste trabalho), pois o comportamento assintótico da curva gaussiana permite que todos os pontos tenham contribuição na probabilidade de um ponto específico, mesmo que seja mínima. A partir dessa escolha, a equação (15) toma a forma em (18).

$$\hat{f}(x) = \frac{1}{N} \sum_{k=1}^N \phi(x - x_k, h) \quad (15)$$

$$\phi(x, x_k, h) = u(x, x_k, h) = \begin{cases} \frac{1}{h} & \text{se } x \in [x_k - \frac{h}{2}; x_k + \frac{h}{2}] \\ 0 & \text{caso contrário} \end{cases} \quad (16)$$

$$\phi(x - x_k, h) = G(x - x_k, h^2) = \frac{1}{h\sqrt{2\pi}} e^{-\frac{1}{2h^2}(x-x_k)^2} \quad (17)$$

$$\hat{f}(x) = \frac{1}{N} \sum_{k=1}^N G(x - x_k, h^2) = \frac{1}{Nh\sqrt{2\pi}} \sum_{k=1}^N e^{-\frac{1}{2h^2}(x-x_k)^2} \quad (18)$$

Uma vez definida a PDF de um conjunto de amostras de entrada a partir da aplicação do método da janela de Parzen com kernel gaussiano, é possível expandir a entropia quadrática de Rényi, juntando as equações (14) e (18), como observado em (19). Expandindo a equação, chega-se em (20). Rearrmando esta, tem-se (21).

$$H_{X_2}(x) = -\log \int_{-\infty}^{+\infty} \left[\frac{1}{N} \sum_{k=1}^N G(x - x_k, h^2) \right]^2 dx \quad (19)$$

$$H_{X_2}(x) = -\log \frac{1}{N^2} \int_{-\infty}^{+\infty} \left[\sum_{j=1}^N \sum_{k=1}^N G(x - x_k, h^2) G(x - x_j, h^2) \right] dx \quad (20)$$

$$H_{X_2}(x) = -\log \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N \int_{-\infty}^{+\infty} G(x - x_k, h^2) G(x - x_j, h^2) dx \quad (21)$$

Em (21), é formada a integral do produto de gaussianas. Segundo [41], esse cálculo resulta em uma gaussiana cuja média é igual a diferença entre os argumentos x_k e x_j e a variância é a soma dos dois valores originais. Assim, a equação toma a forma apresentada em (22), com o somatório de gaussianas das combinações das distâncias entre amostras de duas variáveis aleatórias.

$$H_{X_2}(x) = -\log \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N G(x_k - x_j, 2h^2) \quad (22)$$

Na perspectiva de redução do conteúdo informativo das amostras de erro de um modelo de *Machine Learning* através da redução da entropia, a minimização da equação (22) pode ser alternativa a ser aplicada no processo de aprendizagem da estrutura, onde as saídas geradas e reais substituiriam as variáveis x_k e x_j . Esse critério de ITL é chamado de erro de entropia mínima (em inglês, *minimum entropy error* ou MEE), com formulação final exposta em (23) [46]. Vale destacar que, neste caso, a minimização da equação da entropia equivale à maximização do logaritmando da fórmula, chamado de potencial de informação (em inglês, *information potential* ou IP).

$$\begin{aligned}
MEE &= \min \left(-\log \frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N G(x_k - x_j, 2h^2) \right) \\
&\equiv \max \left(\frac{1}{N^2} \sum_{j=1}^N \sum_{k=1}^N G(x_k - x_j, 2h^2) \right)
\end{aligned} \tag{23}$$

3.2.2 Critério de Máxima Correntropia

Dentre as metodologias para medida de semelhança entre duas variáveis, a correlação é certamente uma das mais aplicadas. Assim como o MSE, sua eficiência depende da condição de PDF gaussiana e linearidade. Para a aplicação em casos não lineares e com distribuições não gaussianas, uma função de correlação generalizada chamada correntropia é uma alternativa que contempla momentos estatísticos de mais alta ordem da PDF. A entidade estatística foi primeiramente definida em [47] na análise de apenas uma variável aleatória, com posterior expansão elaborada para o caso geral de duas variáveis aleatórias em [48].

A correntropia consiste na medida de similaridade generalizada entre duas variáveis aleatórias X e Y , definida como V na equação (24) [48]. Na fórmula, $f(\cdot)$ é uma função de kernel convenientemente escolhida como gaussiana e $E[\cdot]$ é o momento estatístico de valor esperado. O nome da função de correlação generalizada se deve à relação com a entropia quadrática de Rényi. Segundo [47], a média da correntropia converge assintoticamente para o valor de potencial de informação obtido do método da janela de Parzen com kernel gaussiano.

$$V(X, Y) = E[f(X - Y)] \tag{24}$$

Considerando a expansão da série da função de kernel gaussiana, a equação (24) pode ser reescrita como em (25). A equação (26) apresenta o termo da expansão quando n é igual a 1. Na fórmula, a parcela $E[\langle x, y \rangle]$ equivale ao momento conjunto central de segunda ordem, isto é, a covariância. Assim, demonstra-se que a medida de correlação (covariância em casos de média zero) é considerada na função da correntropia [47].

$$V(X, Y) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma^{2n} n!} E[\|x - y\|^{2n}] \quad (25)$$

$$\begin{aligned} V(X, Y) &= \frac{1}{\sqrt{2\pi}\sigma} \times \frac{-1}{2\sigma^2} E[\|x - y\|^2] \\ &= \frac{-1}{2\sqrt{2\pi}\sigma^3} (E[\|x\|^2] + E[\|y\|^2] - 2E[\langle x, y \rangle]) \end{aligned} \quad (26)$$

Da comparação entre as equações do MSE e da correntropia para amostras de saída previstas e reais, expostas em (8) e (27) respectivamente, percebe-se a diferença de comportamento das funções em relação às amostras de erro $e_k = y_k^{previsto} - y_k^{real}$. No MSE, todos os erros têm o mesmo peso de contribuição no cálculo do valor total, ao passo que a correntropia pondera a contribuição das amostras através do kernel gaussiano. Por exemplo, no caso de *outliers*, o termo do somatório em (27) tende a zero, fazendo com que a correntropia praticamente não sinta o efeito do erro mais grosseiro [48].

$$V(Y^{previsto}, Y^{real}) = \frac{1}{N} \sum_{k=1}^N G(y_k^{previsto} - y_k^{real}, \sigma) \quad (27)$$

O erro quadrático médio também pode ser escrito como em (28). Desta expressão pode-se verificar que o termo quadrático para outliers pode amplificar a contribuição de amostras cujas probabilidades estão distantes da média da distribuição de erros, sinalizando que a aplicação do MSE é adequada para a análise de distribuições de erro gaussianas, mas não para o caso de distribuições não simétricas, com média diferente de zero e com *outliers* [48].

$$MSE(X, Y) = E[(X - Y)^2] = \iint (x - y)^2 f_{XY}(x, y) dx dy = \int e^2 f_E(e) de \quad (28)$$

Outra forma de comparação entre a correntropia e o MSE é feita através da análise de uma propriedade da primeira. A chamada métrica induzida de correntropia (em inglês, *correntropy induced metric* ou CIM), cuja equação, vista em (29), denota o significado geométrico da função de correlação generalizada no espaço amostral. De acordo com [48],

quando o erro entre $Y^{previsto}$ e Y^{real} tende a zero, a CIM equivale a norma L2, ou norma euclidiana (distância vetorial em relação à origem). Com o aumento do erro, a métrica se torna a norma L1, como a soma das diferenças absolutas. Quando o erro atinge valores muito elevados, a CIM é equivalente a norma L0 (número de elementos diferentes de zero em um vetor), o que evidencia a baixa sensibilidade para erros muito grandes. Vale ressaltar que as escalas das regiões de normas L2, L1 e L0 dependem da definição da largura da janela do kernel gaussiano. Por exemplo, um valor muito pequeno de janela reduz a zona euclidiana e aumenta a área L0, ao passo que uma janela muito grande proporciona o aumento da zona euclidiana, região na qual a métrica se comporta como o MSE. Ou seja, mesmo o uso de janelas mais largas (menor precisão) não fará com que a correntropia tenha um desempenho pior que o erro médio quadrático.

$$CIM(Y^{previsto}, Y^{real}) = \sqrt{(G_\sigma(0) - V(Y^{previsto}, Y^{real}))} \quad (29)$$

Dessa forma, a função de correlação generalizada pode ser utilizada como uma função objetivo no treinamento de modelos de ML, inclusive para os casos de regressão. Com o propósito de aproximar saídas previstas e reais o máximo possível, é desejável a maximização da similaridade, ou seja, a correntropia entre elas. A partir da maximização da equação (27), define-se o critério de correntropia máxima (em inglês, *maximum correntropy criterion* ou MCC), como apresentado em (30), com $g(\cdot)$ sendo o modelo estruturado com parâmetros ajustáveis w e amostras de entrada x_k que, em conjunto, produzem amostras calculadas $y_k^{previsto}$. Vale ressaltar que, em relação ao MEE mostrado em (23), o cálculo do MCC é mais fácil por precisar apenas do somatório de contribuições gaussianas dos erros (o MEE demanda o cálculo de todas as combinações de diferenças entre saídas previstas e reais), o que reduz o tempo de processamento necessário nas etapas de treino e validação dos modelos [6].

$$MCC = \max \left(\frac{1}{N} \sum_{k=1}^N G(g(w, x_k) - y_k^{real}, \sigma) \right), \quad \text{sendo } g(w, x_k) = y_k^{previsto} \quad (30)$$

Capítulo 4 - Revisão da Literatura

O perfil estocástico da hidrologia sugere a aplicação de modelos suficientemente capazes de analisar dados do escopo hidrológico. No estudo de previsões de aflúncias naturais, modelos empíricos e conceituais podem ser aplicados. Os modelos empíricos se baseiam nas relações matemáticas para expressar o comportamento dos objetos de análise, enquanto os modelos conceituais aplicam informações de processos físicos no estudo realizado [49].

A adoção de modelos estocásticos é uma tendência na geração de vazões afluentes previstas, com históricos de aflúncias tratados como séries temporais. Nessa linha, determinados modelos podem conseguir melhor extração de informação das séries, como características de sazonalidade, periodicidade e tendência. Dentre os modelos empíricos, existem os modelos autorregressivos (como aqueles usados no PREVIVAZ, explicado na seção 2.3.2), destacando-se métodos que consideram efeito de média móvel (ARMA e ARIMA), sazonalidade (SAR e SARIMA) e periodicidade (PAR e PARMA) [50]. Existem também modelos autorregressivos conceituais, isto é, que consideram a adoção de dados de processos físicos correlatos à vazão, como, por exemplo, o PARX, que conta com a adoção de variáveis exógenas [51]. Uma limitação dessas estruturas é referente ao tratamento de meios não lineares, uma vez que esses modelos são baseados em cálculos de combinações lineares.

Modelos de *Machine Learning* são alternativas com capacidade de análise não linear através do processo iterativo de aprendizagem. Mais especificamente, as redes neurais artificiais apresentam resultados satisfatórios, com técnicas aplicadas na análise de previsão de variáveis de múltiplas áreas (mercado acionário, carga elétrica, hidrologia, entre outras) [52].

A modelagem de RNAs abrange o ajuste fino de diversas configurações de forma que o desempenho da estrutura seja otimizado. Especificamente na linguagem de programação Python, a API Keras aparece como ótima opção, com *framework* completo para modelagem, treino e uso de uma rede neural a partir de sintaxe simples [53].

Multilayer Perceptron é a estrutura de rede neural mais difundida, com a possibilidade de variação do número de neurônios e de camadas ocultas. Entretanto,

segundo teorema desenvolvido por Andrei Kolmogorov em 1957, qualquer função contínua e diferenciável, mesmo que não linear, pode ser aproximada para uma RNA do tipo MLP com uma camada escondida. Assim, a adoção de apenas uma camada oculta é suficiente para a otimização da performance da rede [54].

Em relação aos estudos com aplicação de RNAs, [55] desenvolveu uma rede neural com estrutura do tipo *long short-term memory* (arquitetura recorrente) que fez a previsão de tendências de preços de ações através do histórico de cotações e indicadores de análise técnica. Os resultados encontrados foram considerados positivos na comparação com outras metodologias de previsão. Um ponto favorável em relação à rede modelada foi o alto retorno médio por operação financeira a partir das previsões geradas.

O próprio ONS já utilizou RNAs para previsão de aflúncias no horizonte de uma semana, especificamente para a hidrelétrica de Três Marias na bacia do rio São Francisco. Para treinamento do modelo, foram utilizados registros de vazões históricas e precipitação. Dentre os resultados, foi verificado o melhor desempenho da rede a partir da incorporação de dados de chuva [52].

O trabalho proposto em [56] apresentou uma rede neural para a previsão de aflúncias mensais médias com a inclusão de inferência bayesiana e teoria do caos para determinação do espaço amostral de entrada. A previsão final para horizonte de mais de um mês era produzida através de processo recursivo, considerando oito UHEs brasileiras. A estrutura modelada foi comparada a modelos tradicionais (RNA convencional, Fourier, ARMA e SARIMA), apresentando menor MAPE de previsão. Os autores indicaram o estudo futuro por métodos de detecção automática de padrões de sazonalidade em dados de entrada nos casos em que a informação de calendário não é disponível.

Passando da base mensal para a semanal, [57] modelou um previsor vazões semanais que adotou, como dado de entrada, o histórico de aflúncias de hidrelétricas de diferentes bacias hidrográficas. A metodologia consistiu na combinação de técnicas de redes neurais artificiais e lógica *fuzzy*. Na comparação com o modelo PREVIVAZ, a rede neural apresentou melhores resultados, com base em valores de erros relativos percentuais médios.

Reduzindo ainda mais a granularidade temporal, o projeto desenvolvido em [58] consistiu na modelagem de um previsor de vazões afluentes diárias na usina de Três Marias com técnicas de redes neurais. Além do histórico de aflúncias, foram utilizadas informações pluviométricas e fluviométricas oriundas de postos hidrométricos a montante

da região de estudo. O trabalho atingiu seu objetivo, uma vez que o previsor apresentou melhor desempenho que modelos do ONS, como o PREVIVAZH e o Neuro3M (RNA descrita acima). Foi verificado que a aplicação da rede neural trouxe alguns benefícios no processo preditivo, como a redução significativa dos erros médios e da ocorrência de *outliers*, além da previsão sem viés positivo ou negativo. Dentre as recomendações futuras, estão análises para aplicação de ponderação nos dados de chuva de postos pluviométricos e testagem de outras arquiteturas de RNAs, como o modelo recorrente.

O modelo PREVIVAZH, utilizado pelo ONS, também foi adotado como referência comparativa em [37], no qual projetou-se uma RNA para aflúências previstas, em base diária, da UHE de Tucuruí no horizonte de até 12 dias à frente. Dos resultados, foi verificado o melhor desempenho da rede neural, com MAPE, em média, 48% inferior aos erros referentes ao programa do ONS.

Um modelo chuva-vazão foi proposto em [59] para a previsão de vazões diárias com horizonte de previsão de sete dias à frente. No projeto, foi utilizada inferência bayesiana em MLP com seis estratégias diferentes de previsão em relação à janela de vazões e precipitação em atraso. Vale ressaltar que os dados de entrada também contaram com a representação binária de períodos seco, úmido e de transição. Os melhores resultados foram atingidos nos casos de uso dos dados de chuva e uma janela maior de *inputs* em atraso. De acordo com os autores, uma aproximação foi feita ao se adotar precipitação observada como informação prevista, então a aplicação de previsão de chuva está entre as melhorias futuras.

No trabalho feito em [60], foi elaborado um modelo de previsão de vazões diárias no horizonte de dez dias para a hidrelétrica de Furnas, localizada na bacia do Rio Grande. Foram escolhidas a arquitetura de rede neural do tipo MLP e algoritmo de retropropagação Levenberg-Marquardt. Em relação aos *inputs*, foram consideradas alternativas com e sem o uso de precipitação observada e prevista corrigida matematicamente ou não (correção feita através de metodologia usada pelo ONS que considera MLPs de treino bayesiano). Adicionalmente, a modelagem das redes passou por validação cruzada, com a divisão de amostras para etapas de treino, validação e teste. Dentre os diferentes cenários testados, os resultados mostraram que o desempenho da rede com precipitação corrigida foi equivalente ao de modelos conceituais, com a vantagem de levar menos tempo para calibração de parâmetros. Como o estudo foi focado em uma UHE específica, trabalhos futuros incluem a expansão da análise para outras usinas.

A combinação de redes neurais artificiais e teorema de Bayes foi explorada em [61] com modelo testado que aplicou informação de precipitação oriunda do modelo climatológico ETA. Teoria do caos foi adotada para a determinação de dados de entrada em atraso. Além de vazão e chuva, variáveis binárias foram usadas para representação dos meses do ano. Na análise de erros absolutos médios, os resultados indicaram que a RNA com previsão feita de processo recursivo obteve melhor desempenho para a maioria das bacias estudadas, com ganhos de acurácia na previsão que chegaram a 22%.

No trabalho em [62], um MLP foi construído com a adoção de dados de vazão e precipitação. O modelo foi treinado com duas abordagens diferentes: retropropagação de Levenberg Marquardt e algoritmo bayesiano. Os resultados mostraram que a segunda alternativa, obteve menores erros de previsão. Além disso, a adoção de informação de chuva ajudou a melhorar o desempenho do previsor. O estudo foi centrado especificamente em uma pequena central hidrelétrica localizada na bacia do rio São Francisco, sendo necessário, assim, estudo futuro em relação a aplicação do modelo para outras usinas e regiões.

Em [63], foi desenvolvido um previsor neural com horizonte de previsão de 24 horas. O modelo foi arquitetado como um *Multilayer Perceptron* e foi treinado com dados do rio Petro, localizado na bacia do rio São Francisco. As amostras de entrada foram normalizadas e divididas para etapas de treino, validação e teste. O número de neurônios na camada oculta foi escolhido de forma heurística no processo de validação cruzada. Os resultados indicaram que uso de chuva à montante e à jusante da usina melhorou o desempenho do previsor em cerca de 31% na comparação ao modelo que apenas utilizou dados históricos de afluências como *inputs*. Como a rede neural foi modelada para o teste de apenas uma usina, os autores sugeriram estudos futuros sobre a aplicação em outras UHEs e com horizonte de previsão maior.

A partir de pesquisa literária, é possível verificar o grande número de trabalhos com técnicas de redes neurais artificiais aplicadas no processo de previsão de afluências, evidenciando a boa relação entre a estrutura de ML e o escopo hidrológico. Adicionalmente, implementações presentes nos trabalhos descritos serviram de base para a modelagem de rede proposta neste trabalho, como a adoção de dados de entrada hidrológicos, meteorológicos e variáveis binárias temporais, normalização e validação cruzada aplicada aos *inputs*, processo recursivo para preenchimento do horizonte de previsão e adoção de arquitetura MLP.

Como explicitado no capítulo anterior, o desempenho de modelos de *Machine Learning* que adotam funções de erro que se propõem a minimizar a variância do erro não é ótimo quando a análise é feita para distribuições não gaussianas. Nessa linha, *Information Theoretic Learning* aparece como *framework* para análise estatística não linear, a partir da combinação dos conceitos de entropia e estimação de PDF das amostras disponíveis. Alguns trabalhos com propostas de ML com ITL podem ser destacados a partir de mapeamento de produções bibliográficas.

No trabalho apresentado em [64], a previsão de demanda elétrica é gerada a partir do modelo proposto, com base na combinação do método de regressão por vetores de suporte e práticas de ITL. No projeto, foi utilizada uma variação do MCC, com a mistura de duas funções gaussianas como kernels, dando nome à função de erro adotada: critério de máxima correntropia misturada. Os resultados comprovaram que o modelo proposto superou a opção pela regressão por vetores de suporte com aplicação do erro quadrático médio.

No trabalho feito em [6], uma rede neural foi modelada para a previsão de geração eólica com a utilização de conceitos de ITL aplicados no treino da estrutura. Foram testados MLPs com diferentes funções de erro: MCC, MEE, além de variações derivadas das duas primeiras funções. Os resultados mostram que, com a inserção de noções de ITL, a rede obteve melhor desempenho na comparação com a previsão feita com o MSE. A previsão de geração eólica feita com RNA sob conceitos de ITL também foi proposta em [65], com a apresentação de resultados melhores do que a condição de uso do MSE. Além disso, uma vantagem em termos computacionais foi verificada do MCC em relação ao MEE.

No trabalho apresentado em [66], foi proposto o desenvolvimento de uma rede neural do tipo *long short-term memory*, com a adoção do MCC como critério de aprendizado, para a previsão de geração eólica de curto prazo. No projeto, foi considerada a decomposição do histórico de geração, com posterior recomposição utilizando-se o método de *Sample Entropy* (algoritmo para estimação de entropia de uma série temporal). A partir dos resultados, concluiu-se que a rede recorrente proposta teve um desempenho melhor que os métodos convencionais.

Na revisão bibliográfica realizada, não foram encontrados trabalhos em que a maximização da correntropia fosse utilizada em modelos de previsão de vazões, mostrando que a relação entre os conceitos de ITL e previsão hidrológica ainda não foi devidamente

explorada. Pelos exemplos ilustrados nos últimos três parágrafos em que a aplicação de *Information Theoretic Learning* superou os métodos preditivos convencionais de carga elétrica e geração eólica, tem-se boa perspectiva para a adoção do *framework* para ML no escopo de hidrologia. A partir da vantagem de tempo de processamento do MCC em relação ao MEE, como citado no capítulo anterior e no trabalho em [6] descrito, foi feita a escolha pela aplicação do Critério de Máxima Correntropia neste trabalho.

A relação entre o processo de previsão de vazões e a formação de preço *spot* foi tema do trabalho feito em [67], no qual foi feita a comparação entre três modelos previsores de afluências, dentre eles, a rede neural artificial do tipo MLP. Em relação ao PLD, foi considerado o horizonte de curto prazo, a partir da execução do modelo DECOMP. A comparação entre os preços gerados pelos modelos propostos e as referências do ONS e da CCEE foi satisfatória, sobretudo no caso de utilização de modelo climatológico que fornece dados de precipitação.

Levando em consideração que a CCEE adotou o DESSEM recentemente como modelo final da cadeia de otimização do despacho hidrotérmico, a relação entre a previsão de vazões e o PLD em resolução horária – tema deste trabalho – ainda não foi diretamente estudada.

Capítulo 5 - Materiais e Método

Com o objetivo de complementar os métodos utilizados pelo ONS, este trabalho propõe o desenvolvimento de um previsor de vazões afluentes, em base diária, aplicado às usinas hidrelétricas contempladas no processo de otimização do despacho hidrotérmico de curto prazo. O modelo conta com a combinação de técnicas de RNAs e ITL. Adicionalmente, a análise de impacto da previsão de afluências na formação do preço *spot* horário do mercado de energia elétrica é avaliada, com a incorporação das vazões geradas ao deck de entrada do DESSEM. Como forma de avaliação da rede neural criada, as previsões de vazões e PLDs oficiais são utilizadas, além de saídas produzidas de outros dois modelos que serão descritos nesse capítulo. Por fim, é feito o estudo comparativo em relação aos preços oriundos do caso em que as vazões realizadas são adotadas como previsões do arquivo DADVAZ.

Com exceção ao cálculo dos valores de preço, feitos através da execução do programa DESSEM, todo o projeto foi desenvolvido em linguagem de programação Python, com destaque para a biblioteca Keras [53] utilizada para modelagem da rede neural (arquitetura, etapa de aprendizagem e previsão final).

5.1 Descrição dos Dados de Entrada e Saída

Os dados de entrada utilizados na modelagem da rede neural podem ser classificados como hidrológicos, meteorológicos e temporais.

A informação de característica hidrológica é referente ao histórico de vazões afluentes médias diárias de usinas hidrelétricas do SIN. Os valores foram coletados do relatório divulgado diariamente pelo ONS aos agentes com as condições hidrológicas dos reservatórios hidrelétricos. O arquivo de um dia específico contém afluências realizadas de diferentes tipos (natural total, incremental, artificial, entre outras) nos últimos 30 dias. Para este trabalho, foi escolhida a vazão afluente incremental, uma vez que esta é utilizada pelo ONS e a CCEE para as execuções oficiais do DESSEM. O histórico de vazões realizadas

direcionado para a treino e validação da RNA corresponde ao período entre 01 de janeiro de 2018 e 02 de julho de 2021.

Os *inputs* do escopo meteorológico são relativos aos dados de precipitação diária observada em território brasileiro, obtidos da análise de mapas de precipitação, como o exemplar mostrado na Figura 16, publicados pelo Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), pertencente ao Instituto Nacional de Pesquisas Espaciais (INPE). Segundo [68], o mapa é produzido a partir de dados do Instituto Nacional de Meteorologia (INMET), de plataformas de coleta de informações do próprio INPE e de centros estaduais de meteorologia. Em cada mapa diário, foi calculada a precipitação média, em mm, para doze diferentes regiões de bacias hidrográficas no Brasil (Amazonas, Atlântico Leste, Atlântico Nordeste Ocidental, Atlântico Nordeste Oriental, Atlântico Sudeste, Atlântico Sul, Paraguai, Paraná, Parnaíba, São Francisco, Tocantins e Uruguai), baseada nas cores dos pixels distribuídos em cada região, levando em consideração a escala gráfica de cores presente na parte de baixo do mapa. O algoritmo de processamento dos dados de precipitação pode ser visto na Figura 17. O horizonte do histórico de chuva foi o mesmo utilizado na obtenção dos registros de vazões realizadas.

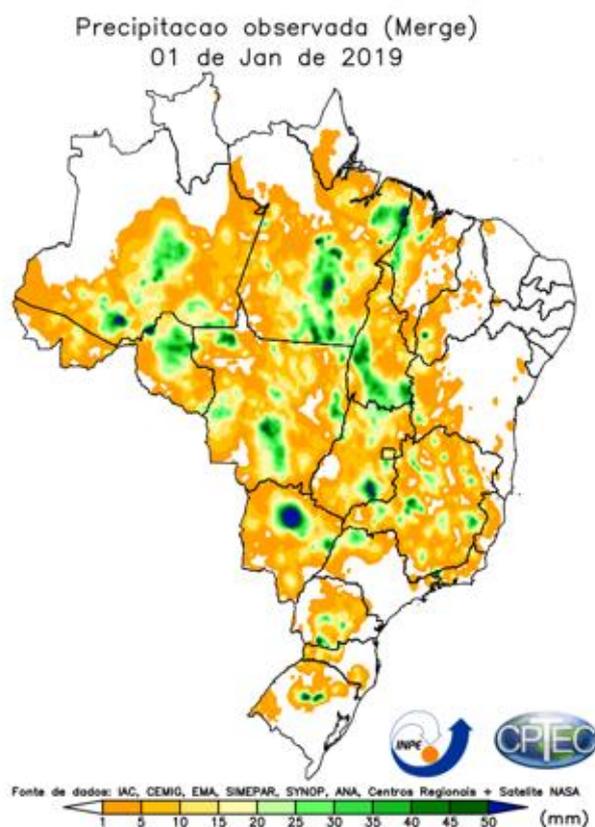


Figura 16 - Mapa de precipitação observada diária. Fonte: [69].

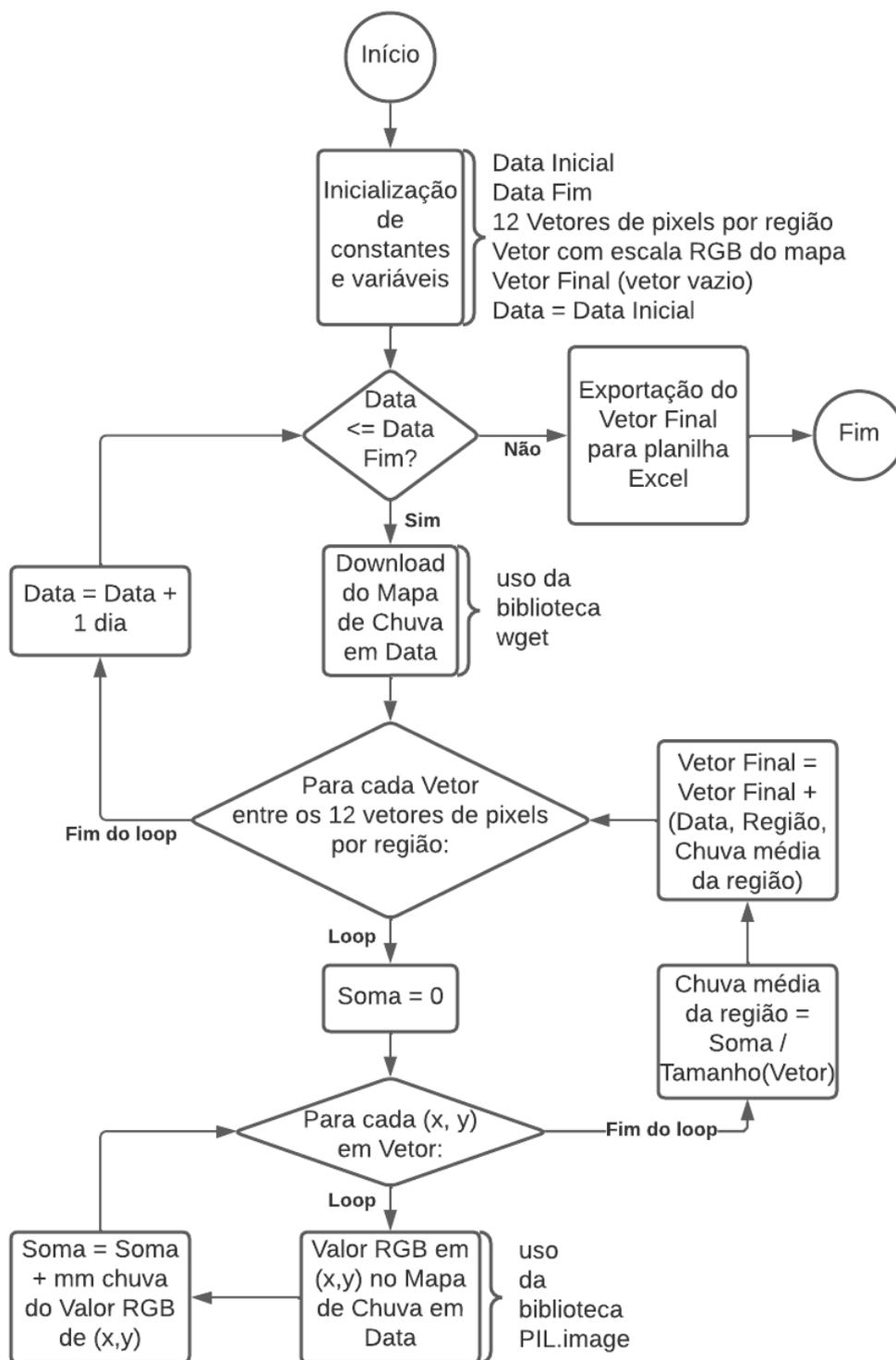


Figura 17 - Algoritmo de processamento dos dados de precipitação

Os dados de entrada temporais são compostos por três variáveis binárias que, combinadas, representam os três tipos de períodos do ano definidos pela CCEE para atribuição de níveis de patamares de carga elétrica: verão, inverno e transição. Em termos

das três variáveis de entrada criadas, esses períodos são, respectivamente, 100; 010 e 001. Os detalhes dos *inputs* temporais podem ser analisados na Tabela 2.

Tabela 2 - Informações dos dados de entrada temporais

Período	Meses	Representação
Verão	novembro a março	[1, 0, 0]
Inverno	maio a agosto	[0, 1, 0]
Transição	abril, setembro e outubro	[0, 0, 1]

5.2 Modelagem do Previsor

A estrutura escolhida para o previsor foi uma rede neural artificial do tipo *Multilayer Perceptron*, com apenas uma camada oculta entre as camadas de entrada e saída. Foram utilizados 10 neurônios na camada de entrada, o que coincide, obviamente, com o número de variáveis de entrada do modelo: as três variáveis binárias combinada que representam o tipo de mês de uma determinada data t (verão, inverno e transição), vazões incrementais e precipitação referentes a uma usina hidrelétrica específica na janela de três dias passados e a chuva em mm do dia t cuja afluência é prevista. Merece menção o fato que, em termos práticos, a precipitação realizada do dia a ser previsto não é conhecida, ou seja, foi feita uma aproximação na associação do dado realizado como previsto. Ainda assim, como a informação de precipitação é resultante de análise qualitativa de mapas com escalas de cores, é esperado que, em média, os mapas de chuva realizado e previsto sejam significativamente semelhantes, fazendo com que a aproximação feita seja considerada aceitável.

O esforço feito para incorporar a chuva e vazão em dias passados aos dados de entrada do modelo é justificado pelo estudo de correlação feito entre as afluições em estágios (dias) anteriores e no dia corrente, como pode ser visto na Figura 18. No gráfico, estão dispostos dez histogramas dos valores de correlações considerando o histórico de vazões de 144 UHEs, referentes às análises do primeiro ao décimo dia passado. Com base na análise gráfica, foi escolhida a janela de três dias passados, uma vez que as distribuições de $t - 1$, $t - 2$ e $t - 3$ são aquelas com maior frequência de correlações mais próximas do valor unitário.

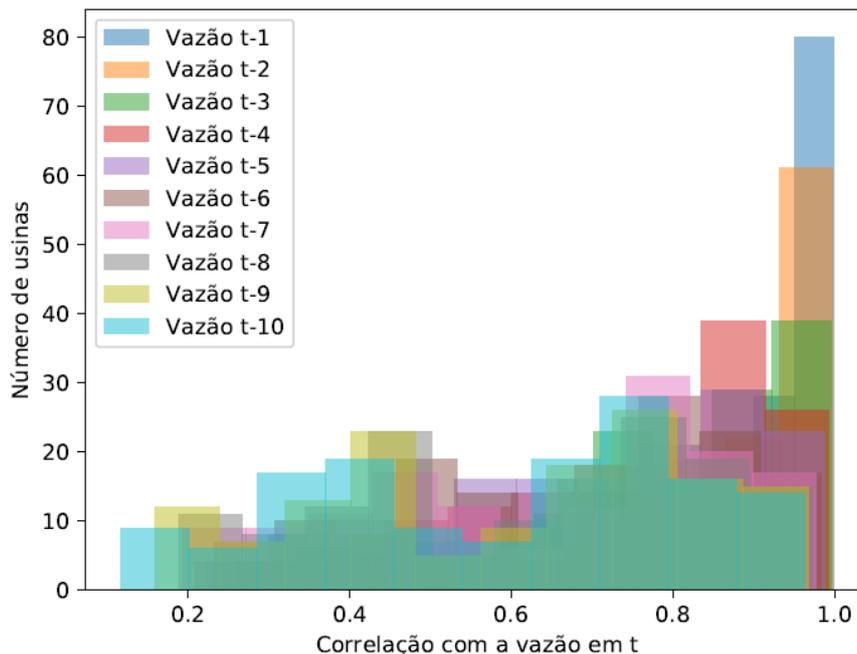


Figura 18 - Histogramas de correlações entre vazões em t-n e t, com t = 1, 2, ..., 10.

Antes de serem utilizados para treino e validação do predictor, os dados de entrada passaram pelo processo de normalização, na tentativa de evitar a ocorrência de cálculos com números muito grandes no interior do modelo, o que tornaria a rede saturada. Cada uma das variáveis de entrada foi normalizada, entre 0 e 1, segundo a equação em (31), na qual x_i^{norm} é a amostra normalizada, l_{sup} e l_{inf} os limites superior e inferior da normalização e x_{max} e x_{min} os valores máximo e mínimo dentre todas as amostras da variável de entrada em transformação. Vale ressaltar que as amostras de saída esperadas também passaram pelo mesmo procedimento.

$$x_i^{norm} = \frac{(l_{sup} - l_{inf})(x_i - x_{min})}{x_{max} - x_{min}} + l_{inf} \quad (31)$$

Em relação à camada oculta da rede, foram testadas diferentes quantidades de neurônios, variando entre 10, 20, 30, 40 e 50 unidades. A função de ativação aplicada em cada um desses neurônios foi do tipo sigmoide, conforme mostrada em (5), o que reforça a necessidade da normalização entre 0 e 1 descrita anteriormente. A terceira e última camada da RNA teve estrutura mais simples, com apenas um neurônio cuja função de ativação foi do tipo linear. A saída desse neurônio e, conseqüentemente, a saída rede foi a vazão afluente incremental em um dia específico t .

A prática de validação cruzada foi implementada no previsor, a partir da aplicação das etapas de treino, validação e teste. Nessa linha, o histórico dos dados de entrada e saída esperada foram divididos da seguinte forma, utilizando, como exemplo hipotético, a previsão de vazão do dia 01 de janeiro de 2021: os primeiros 70% do período entre 01 de janeiro de 2018 e 31 de dezembro de 2020 foram alocados para a etapa de treino, enquanto os outros 30% foram direcionados para validação. Neste trabalho, como a etapa de teste foi utilizada para previsão final, as amostras do horizonte de 01 de janeiro de 2021 para frente foram utilizadas nesse último passo.

Para aprendizagem do modelo, foi adotada uma extensão do algoritmo de retropropagação apresentado na equação (7), chamado de otimizador Adam. De acordo com [70], este método consiste na aplicação da técnica de descida do gradiente, porém com a redução da taxa de aprendizado ao longo das épocas da etapa de treino, feita a partir da normalização da taxa. Com o propósito de evitar que ela caia em ritmo exagerado, é utilizado um termo, chamado de momento, responsável por suavizar o decaimento imposto à taxa de aprendizado. A taxa de aprendizado inicial foi definida como 10^{-3} .

Os conceitos de *Information Theoretic Learning* se fazem presentes na função de erro adotada no treino da RNA: o critério de entropia máxima, apresentado na equação (30). Uma vez que a função de erro deve ser minimizada, foi aplicado o valor negativo do MCC. Na função, a largura da janela do kernel gaussiano foi definida como 10^{-1} , considerando testes preliminares de aderência. Com a aplicação do critério de ITL no treinamento do modelo proposto, este foi nomeado RNA-MCC para efeito de melhor entendimento das estruturas previsoras abordadas neste trabalho.

O número de épocas adotado foi igual a 100. A escolha tentou conciliar a opção por número que não fosse tão elevado (evitando tempo de execução muito longo por conta do número de hidrelétricas e dias previstos considerados) e a condição da não ocorrência de *overfitting* durante o aprendizado do previsor. A Figura 19 apresenta a evolução dos valores da função de erro – o MCC negativo – nas etapas de treino e validação de algumas UHEs, mostrando que as curvas decaem de forma conjunta ao longo das 100 épocas.

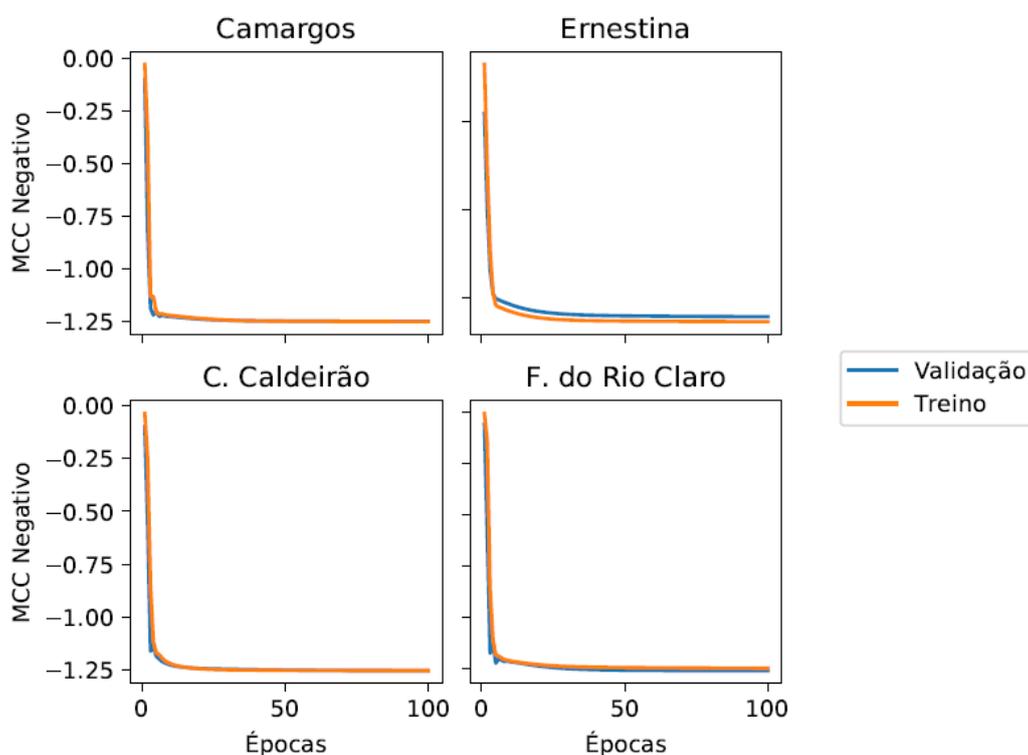


Figura 19 - Evolução dos erros de treino e validação de algumas UHEs.

Além da análise gráfica mostrada na Figura 19, as amostras de validação foram utilizadas para a escolha da melhor estrutura da rede (em relação ao número de neurônios na camada oculta) a partir dos valores da função de erro, isto é, a arquitetura que gerasse o menor MCC negativo de validação seria escolhida para o restante do processo. Essa escolha foi realizada para cada primeiro dia previsto e para cada usina considerada.

Com a escolha da melhor configuração, esta foi adotada para a previsão final, que foi feita para um passo, isto é, uma data. Nessa etapa, é necessária a inversão da normalização dos valores produzidos de acordo com a escala usada em (31). Dois pontos importantes devem ser levantados: a possibilidade que o arquivo DADVAZ tem de apresentar afluências previstas para a janela de até sete dias à frente e o fato de que, em exemplo prático, para que seja prevista a vazão de amanhã, deve-se fazer a previsão do dia de hoje, pois a vazão deste ainda não é realizada. Nessa lógica, de forma operacional, a previsão feita em uma sexta-feira para o dia de sábado, que tem o arquivo DADVAZ com horizonte de sete dias à frente, na verdade foi feita com janela total de 8 dias previstos e a vazão do primeiro dia previsto presente no DADVAZ, na verdade, é o valor referente ao segundo dia previsto na janela de previsão total. A partir dessa lógica, o modelo criado gerou previsões para dias à frente através de passos recursivos da etapa final, ou seja, a

previsão de um dia é executada, o valor gerado é incorporado ao histórico de vazões como se tivesse sido realizado, é gerada outra previsão de afluência e assim por diante, até que todo o horizonte de dias seja preenchido. Para efeito de melhor entendimento, a Figura 20 ilustra o fluxograma do processo realizado para a previsão de cada usina hidrelétrica, enquanto a Tabela 3 resume as configurações gerais da RNA-MCC.

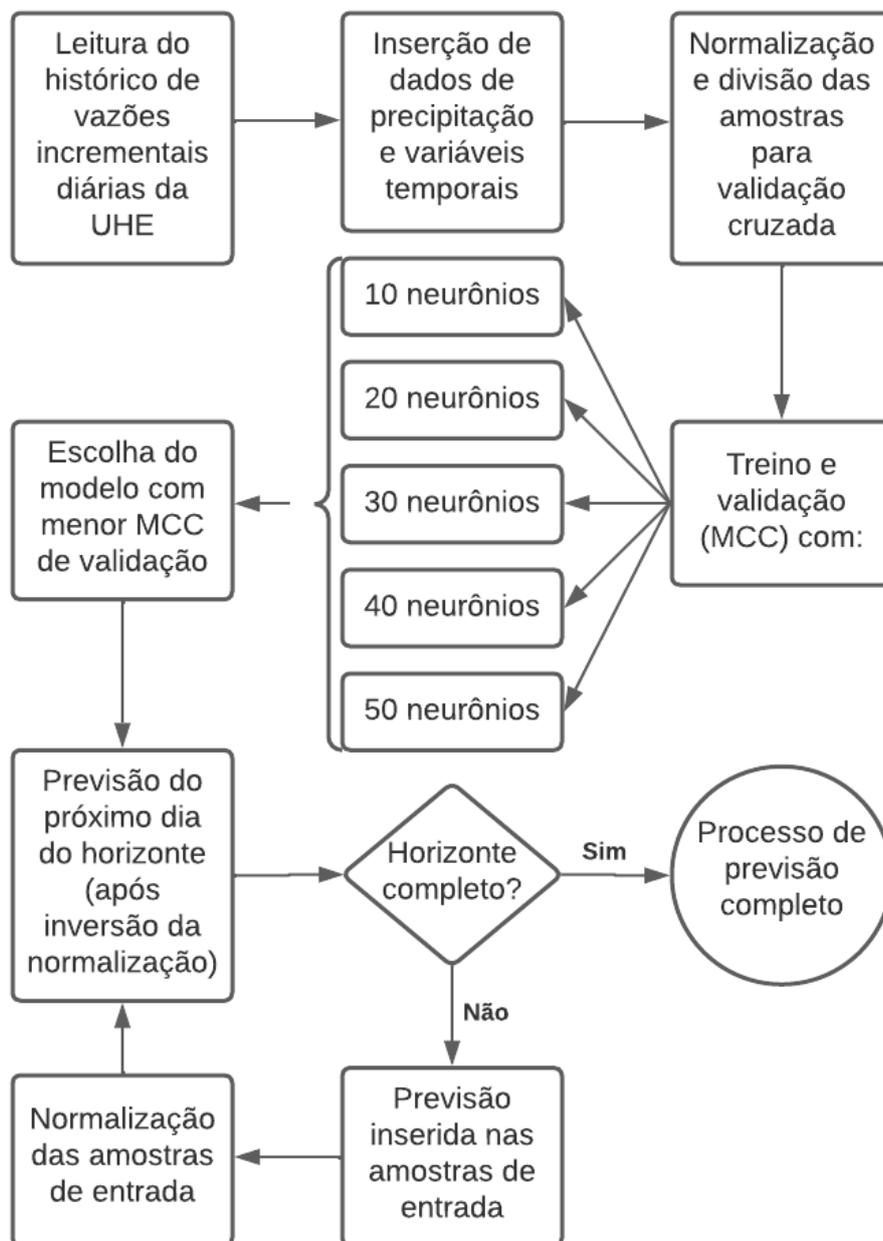


Figura 20 - Fluxograma do preditor de vazões proposto.

Tabela 3 - Configurações do modelo RNA-MCC

Configurações da RNA-MCC	
Arquitetura RNA	MLP com 3 camadas (1 oculta)
Nº Dados de Entrada	10 (vazão e chuva t-1, t-2, t-3; chuva t; 3 variáveis binárias)
Nº Neurônios Ocultos	10, 20, 30, 40 ou 50
Nº Neurônios Saída	1
Função de Ativação	sigmoide (camada oculta) e linear (camada de saída)
Algoritmo de Treino	retropropagação com otimizador Adam
Função de Erro	MCC negativo
Janela do Kernel	0,1
Treino e Validação	70% do histórico para treino, 30% para validação
Nº Épocas	100

5.3 Metodologia

Dentre as 162 usinas hidrelétricas contempladas no arquivo DADVAZ, 144 foram consideradas neste trabalho, através de previsões de vazões feitas para todos os dias dos meses de março, abril e junho do ano de 2021. Merece menção o fato de que as UHEs que não foram utilizadas no estudo possuem vazão natural afluyente diretamente relacionada à defluência das usinas à montante, ou seja, a vazão afluyente incremental é nula. Já a escolha dos meses foi devida ao fato de que esses fazem parte de diferentes períodos definidos pela CCEE para atribuição de níveis de carga, conforme indicado na Tabela 2.

Terminado o passo de previsão hidrológica, as afluências geradas foram incorporadas aos decks de arquivos de entrada do modelo DESSEM, através da modificação dos respectivos arquivos DADVAZ. Nessa etapa, foram utilizados os decks referentes às sextas-feiras dentro dos meses citados acima (totalizando 13 dias adotados), sendo a escolha por esse dia da semana decorrente do menor tempo necessário para execução dos arquivos de sexta-feira. Após a customização feita nos 13 decks do DESSEM, estes foram executados, com a posterior coleta dos valores gerados de PLD horário dos submercados Sudeste/Centro-Oeste, Sul, Nordeste e Norte. Em relação ao passo de formação do preço *spot*, três pontos merecem menção. Primeiro, o DADVAZ foi

o único arquivo de entrada alterado do modelo de otimização do despacho hidrotérmico de curtíssimo prazo, fazendo com que as diferenças dos PLDs produzidos, em comparação com os números oficiais divulgados pela CCEE, sejam provocadas unicamente pela variação das previsões de afluências. Segundo, como o propósito final foi a obtenção de valores de PLD e não CMO, foi necessária coleta dos decks de DESSEM oriundos da CCEE e não do ONS. Por fim, não foram aplicados os valores de piso e teto de PLD aos preços coletados e aos preços oficiais da CCEE.

Como suporte na avaliação do modelo RNA-MCC, mais dois previsores paralelos foram criados. O primeiro deles, nomeado como RNA-MSE, consistiu na modelagem de outra rede neural artificial com as mesmas configurações do modelo principal, com exceção à função de erro das etapas de treino e validação, definida pelo convencional MSE. O segundo previsor paralelo, nomeado como ingênuo, tratava as previsões de vazões como a média histórica de afluências incrementais, para uma determinada hidrelétrica, do período de 01 de janeiro de 2018 a 02 de julho de 2021, sendo o valor aplicado de forma constante para todo o horizonte de previsão. A comparação entre as estruturas RNA-MCC e RNA-MSE teve como principal objetivo a avaliação do efeito do uso dos conceitos de ITL na qualidade das vazões previstas, em termos de exatidão frente às afluências realizadas. Já a adoção do modelo ingênuo serviu, principalmente, para análise do impacto da informação hidrológica na formação do preço *spot*, uma vez que é esperado que a média histórica de vazões tomada como previsão não seja tão exata e, conseqüentemente, os valores de PLD sejam mais descolados dos preços oficiais da CCEE. Assim como o previsor principal, os modelos RNA-MSE passaram pelos passos de geração de vazões previstas e execução de decks modificados de DESSEM, como o processo descrito no início desta seção.

A avaliação de desempenho dos modelos foi dividida em duas partes: hidrológica e de formação do PLD. Na etapa hidrológica, as previsões de vazões dos modelos RNA-MCC, RNA-MSE e ingênuo, assim como as previsões feitas pelo ONS, foram comparadas às afluências realizadas nos mesmos dias, com a utilização do erro absoluto percentual na avaliação de cada usina hidrelétrica. Após o cálculo dos erros, os modelos foram comparados entre si, dois a dois, através da aplicação de testes de hipóteses do tipo T-*pareado unilateral à direita* com nível de significância de 5%, discretizado por UHE e número do dia previsto (variando de 1 a 7) [71]. Em uma rodada desse tipo de teste, os erros dos dois modelos analisados são ordenadamente pareados, sendo possível inferir,

com determinado nível de significância, se a média dos erros de um dos casos é menor do que a média do outro.

Todos os testes de hipóteses realizados neste trabalho foram definidos conforme a seguinte metodologia: tomando duas populações de erros com n amostras cada como exemplos, $X = x_1, x_2, \dots, x_n$ e $Y = y_1, y_2, \dots, y_n$, duas hipóteses são elaboradas. A primeira, chamada de alternativa, consiste na suposição que a média do modelo X é menor que a média do modelo Y , enquanto a segunda hipótese, denominada nula, se baseia no cenário complementar, no qual a média de X é maior ou igual a média de Y . Os primeiros cálculos feitos são a soma das diferenças dos erros pareados, a soma das diferenças ao quadrado e a média das diferenças, apresentadas em (32), (33) e (34) respectivamente, assumindo a distribuição normal para ambas as populações. A variável de definição da escolha da hipótese t_{obs} é definida segundo (35), na qual o desvio padrão amostral s_d é calculado conforme (36).

$$d = \sum_{i=1}^n y_i - x_i \quad (32)$$

$$d_2 = \sum_{i=1}^n (y_i - x_i)^2 \quad (33)$$

$$\bar{d} = \frac{d}{n} \quad (34)$$

$$t_{obs} = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} \quad (35)$$

$$s_d = \sqrt{\frac{d_2 - \frac{d^2}{n}}{n - 1}} \quad (36)$$

Após o cálculo de t_{obs} , o valor é comparado ao parâmetro de valor crítico, definido através de consulta na tabela de distribuição t de Student, a partir do número de amostras e o nível de significância utilizados. Nessa linha, a hipótese alternativa é aceita caso o valor

de t_{obs} for maior que o valor crítico verificado, ao passo que a hipótese nula é escolhida no cenário complementar. Na análise entre dois quaisquer modelos deste trabalho, os testes foram realizados de forma dupla, isto é, cada um dos modelos assumiu o papel da população X do exemplo descrito anteriormente. Nos casos em que não foi possível definir que um dos modelos obteve média de erros menor que o outro, assumiu-se que os respectivos desempenhos foram equivalentes. A Figura 21 apresenta o fluxograma dos testes de hipóteses aplicados, considerando as aflúncias de uma usina específica.

A etapa de avaliação de desempenho dos previsores em termos da formação do preço *spot* foi feita em duas partes. Na primeira parte, os valores de PLD dos modelos RNA-MCC, RNA-MSE e ingênuo foram comparados com preços oficiais da CCEE através do cálculo de erro absoluto, sendo tais erros resumidos em gráficos do tipo *boxplot*, discretizados por submercado. Ademais, foi feita a análise gráfica dos próprios PLDs, a partir de gráficos de linha discretizados por submercado e dia. Como suporte aos gráficos de linha produzidos, uma tabela com valores de erro médio absoluto normalizado (NMAE) – formulado em (37) – em relação à curva de referência da CCEE é apresentada. Na segunda parte da avaliação, foram gerados decks de entrada do DESSEM para as mesmas 13 sextas-feiras contempladas no passo anterior, considerando a própria vazão realizada incorporada ao arquivo DADVAZ de cada deck. Nessa linha, com a execução dos decks, foram produzidos PLDs referentes ao cenário de vazões reais. Tais números foram confrontados com os valores de preço oficiais da CCEE e do modelo RNA-MCC, através de gráficos de linha. Nessa etapa, novamente é feita uma tabela com informação de NMAE, dessa vez, tomando o preço do caso de aflúncias reais como preço de referência. Para efeito de melhor entendimento, a Figura 22 ilustra o fluxograma de atividades realizadas neste trabalho.

$$NMAE(PLD, PLD^{ref}) = \frac{\frac{1}{N} \sum_{k=1}^N |PLD - PLD^{ref}|}{\frac{1}{N} \sum_{k=1}^N |PLD^{ref}|} \quad (37)$$

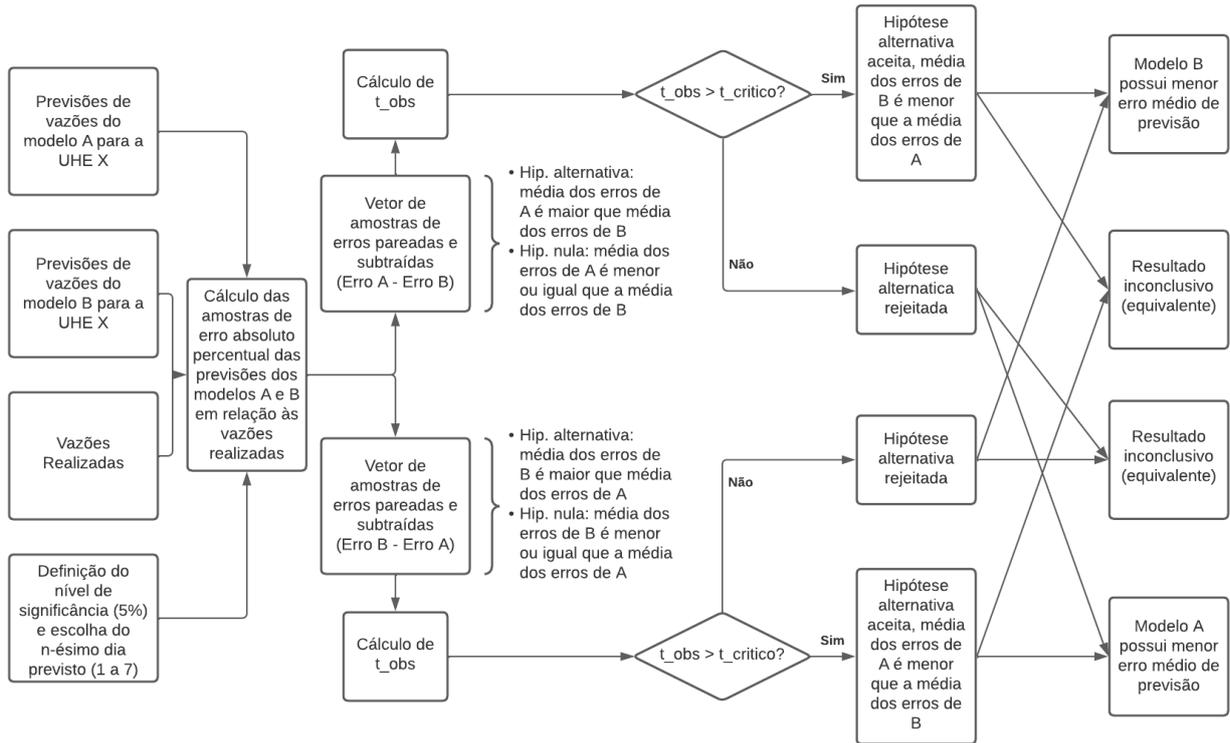


Figura 21 - Fluxograma dos testes de hipóteses do tipo T-pareado unilateral à direita

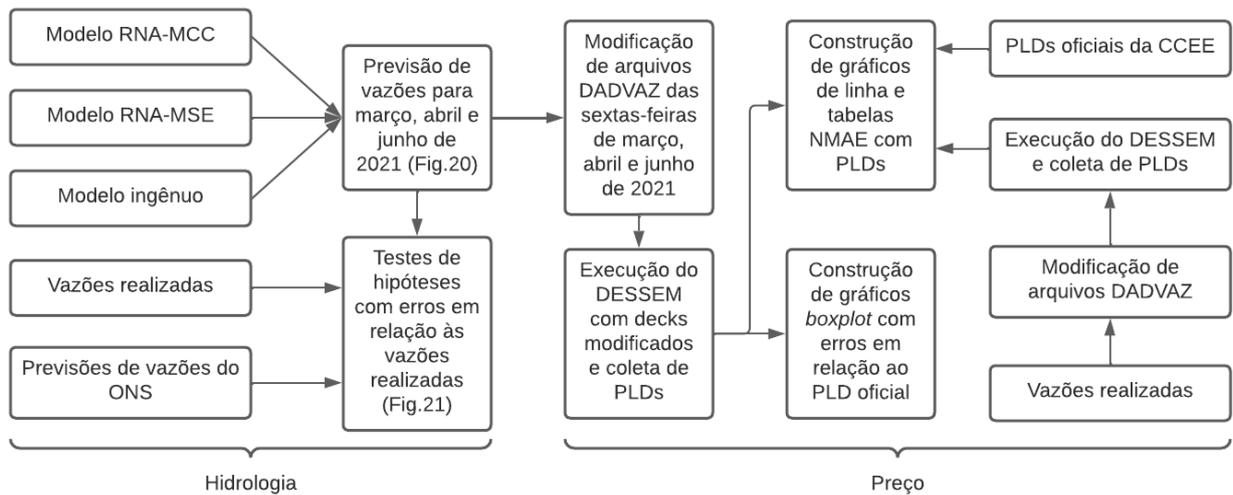


Figura 22 - Fluxograma da metodologia aplicada.

Capítulo 6 - Resultados e Discussão

Os resultados obtidos neste trabalho foram divididos em duas classificações, descritas nas seções deste capítulo: conclusões referentes aos modelos preditivos e suas respectivas afliências previstas e resultados acerca dos PLDs produzidos.

6.1 Resultados quanto à Previsão de Afliências

O primeiro resultado obtido é referente à escolha do tipo de arquitetura do modelo RNA-MCC, em termos no número de neurônios adotados na camada oculta. Considerando 91 dias previstos e 144 usinas hidrelétricas contempladas, obteve-se um cenário de 13.104 previsões. A quantidade de vezes em que cada configuração foi aplicada pode ser vista na Tabela 4. Desses dados, é possível perceber a tendência pela maior quantidade de neurônios, indicando que uma arquitetura de rede neural mais complexa é capaz de reduzir os valores de erro e, conseqüentemente, produzir melhores saídas.

Tabela 4 - Avaliação do número de neurônios ocultos

Nº de Neurônios	Previsões com melhor desempenho	Percentual do Total
10	1.709	13,0 %
20	2.327	17,8 %
30	2.610	19,9 %
40	2.898	22,1 %
50	3.560	27,2 %

Os resultados relativos aos testes de hipóteses feitos com os erros de previsão de afliências estão expostos na Tabela 5 até a Tabela 8. A Tabela 5 mostra a comparação entre o conjunto de previsores do ONS e o modelo RNA-MCC, com a quantidade de usinas hidrelétricas alocadas para cada opção, segundo a média dos erros absolutos percentuais. Verifica-se que, para o primeiro dia previsto, a rede neural proposta foi melhor em 45 UHEs (31,3%), o ONS foi superior em 52 postos (36,1%) e os modelos tiveram

desempenho equivalente em 47 hidrelétricas (32,6%). Desses dados, é possível perceber que o modelo RNA-MCC tem performance equivalente ou melhor que os modelos do ONS em mais da metade dos casos (92 de 144 UHEs, ou 63,9%), indicando uma significativa porção do total. Com o aumento do horizonte de previsão, percebe-se que a performance da RNA-MCC piora frente ao ONS, o que pode ser explicado pelo processo recursivo de previsão da rede neural modelada, que propaga o erro entre previsões e valores reais.

A Tabela 6 quantifica os valores da comparação entre ONS e o modelo RNA-MSE, indicando conclusões semelhantes às aquelas tiradas da tabela anterior, isto é, pico de desempenho obtido no curto prazo e respectiva queda com o aumento do horizonte de previsão. No entanto, há que se destacar que o percentual de equivalência ou superação da previsão do primeiro dia caiu para 54,8% (79 de 144 UHEs).

Tabela 5 - Resultados dos testes de hipóteses: ONS e RNA-MCC

Nº de UHEs	ONS	Equivalente	RNA-MCC
1º dia previsto	52	47	45
2º dia previsto	75	36	33
3º dia previsto	88	33	23
4º dia previsto	97	25	22
5º dia previsto	97	31	16
6º dia previsto	91	40	13
7º dia previsto	80	57	7

Tabela 6 - Resultados dos testes de hipóteses: ONS e RNA-MSE

Nº de UHEs	ONS	Equivalente	RNA-MSE
1º dia previsto	65	38	41
2º dia previsto	86	32	26
3º dia previsto	92	33	19
4º dia previsto	101	26	17
5º dia previsto	107	23	14
6º dia previsto	102	30	12
7º dia previsto	82	54	8

Na comparação entre as duas redes neurais modeladas, exposta na Tabela 7, observa-se que a média dos erros de previsão foi menor ou equivalente no modelo RNA-MCC para quase a totalidade das hidrelétricas analisadas, indicando que a aplicação do critério de correntropia máxima foi benéfica para a melhora do processo de previsão de vazões via RNAs. Na comparação com a revisão literária feita, tal resultado vai ao encontro das conclusões tiradas em [6], [64], [65] e [66], nas quais o MCC superou o desempenho relativo ao MSE.

Por fim, a Tabela 8 apresenta os resultados dos testes de hipóteses envolvendo os previsores RNA-MCC e ingênuo. Dos dados tabelados, nota-se que a RNA-MCC foi melhor ou equivalente em todas as usinas, mostrando que a adoção da média histórica de afluições não foi suficiente para a geração de previsões de vazões mais exatas.

Tabela 7 - Resultados dos testes de hipóteses: RNA-MCC e RNA-MSE

Nº de UHEs	RNA-MCC	Equivalente	RNA-MSE
1º dia previsto	80	62	2
2º dia previsto	78	65	1
3º dia previsto	68	75	1
4º dia previsto	59	83	2
5º dia previsto	58	83	3
6º dia previsto	44	97	3
7º dia previsto	22	121	1

Tabela 8 - Resultados dos testes de hipóteses: RNA-MCC e modelo ingênuo

Nº de UHEs	RNA-MCC	Equivalente	Ingênuo
1º dia previsto	144	0	0
2º dia previsto	142	2	0
3º dia previsto	139	5	0
4º dia previsto	138	6	0
5º dia previsto	134	10	0
6º dia previsto	131	13	0
7º dia previsto	110	34	0

6.2 Resultados quanto à Formação do PLD

Os gráficos do tipo *boxplot* com a análise do erro absoluto entre os PLDs dos modelos RNA-MCC, RNA-MSE e ingênuo e da CCEE, para cada submercado, estão dispostos nas quatro figuras a seguir. Figura 23 e Figura 24 apresentam os resultados para o Sudeste/Centro-Oeste e Sul, respectivamente. É possível perceber que os dados entre as duas figuras são praticamente iguais, o que é justificado por uma questão estrutural, isto é, a forma como as interconexões são realizadas faz com que o Custo Marginal de Operação e, conseqüentemente, o PLD sejam praticamente os mesmos nos dois submercados. Dessa forma, as conclusões podem ser tiradas para os dois gráficos: as distribuições de erros das redes neurais possuem desvios padrões menores e médias mais próximas de zero do que os valores do modelo ingênuo. Além disso, o modelo ingênuo apresenta *outliers* mais espaçados, com maior erro chegando aos R\$ 100,00/MWh, enquanto o maior erro das RNAs ficou no patamar de R\$ 40,00/MWh.

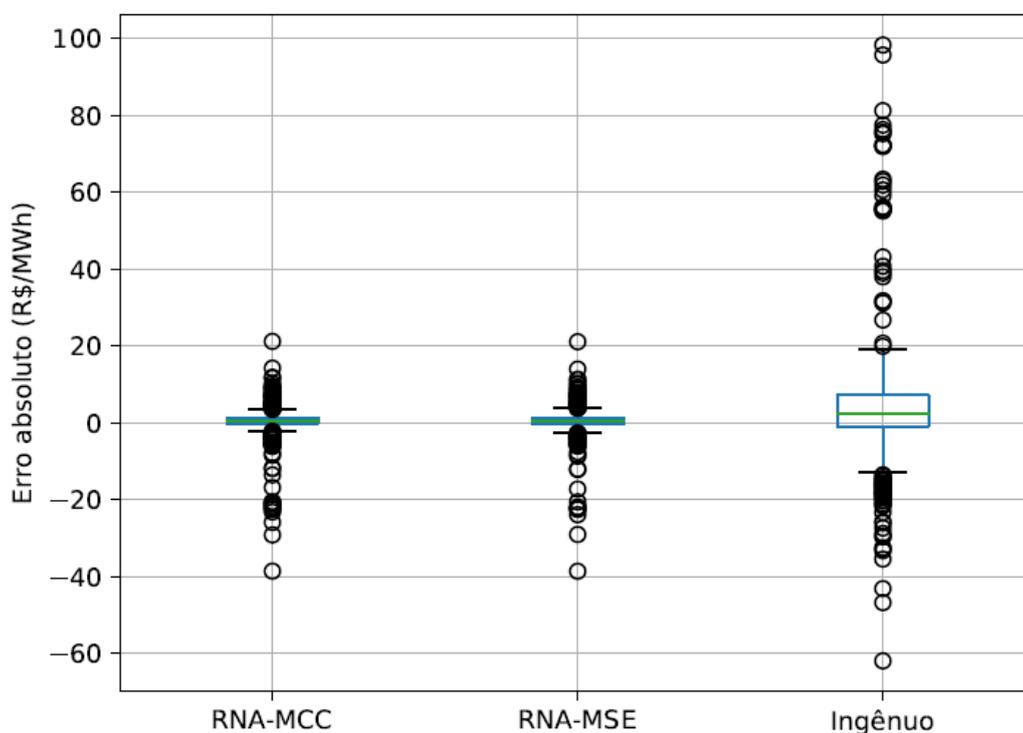


Figura 23 - *Boxplot* com erro de previsão do PLD do submercado Sudeste/Centro-Oeste.

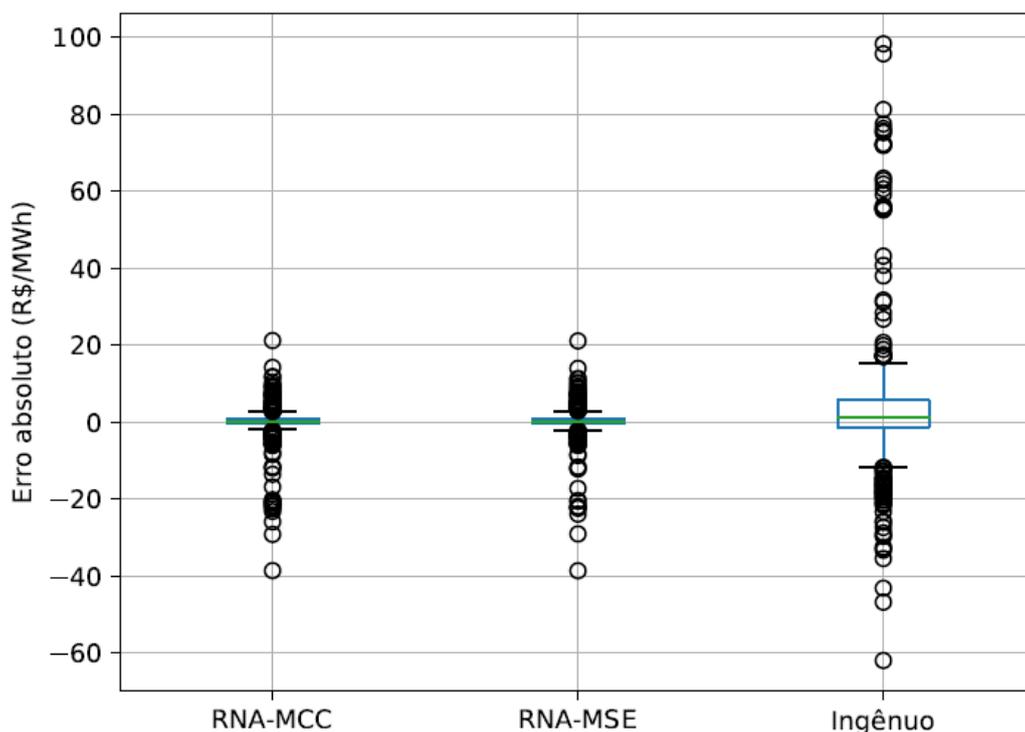


Figura 24 - Boxplot com erro de previsão do PLD do submercado Sul.

Figura 25 e Figura 26 ilustram os gráficos relativos aos submercados Nordeste e Norte, cujos resultados indicaram o mesmo panorama anterior: distribuições achatadas e próximas de zero, enquanto o modelo ingênuo obteve maiores erros. Entretanto, nesses dois últimos submercados, RNA-MCC e RNA-MSE apresentaram maiores outliers, chegando a quase R\$ 100,00/MWh no Nordeste e R\$ 150,00/MWh no Norte.

O modelo ingênuo, que já era o predictor com o pior desempenho nos dois gráficos anteriores, teve seu desempenho piorado na análise dos próximos dois submercados. Essa piora pode ser explicada através do perfil sazonal de vazões da região Norte, como mostrado na Figura 1 com a hidrelétrica de Jirau. De forma simplificada, o primeiro semestre possui vazões muito mais elevadas que o segundo semestre. Consequentemente, a geração hidrelétrica é mais alta na primeira metade do ano, permitindo que os preços sejam mais baixos no Norte, podendo chegar ao piso de PLD. Ademais, por conta do intercâmbio, o Nordeste consegue usufruir da produção oriunda da região vizinha, o que leva ao nivelamento de PLDs entre submercados. O mesmo não acontece com o Sudeste/Centro-Oeste e Sul devido aos limites de transmissão existentes. Como o modelo ingênuo fez uso de médias históricas de vazões, suas previsões foram bem mais baixas que as afluições realizadas dos meses de março, abril e junho, o que levou ao cálculo de

preços maiores e bem descolados dos números oficiais da CCEE, tanto para Norte quanto para o Nordeste.

Vale destacar que, em todos os gráficos, as distribuições de erros das duas redes neurais artificiais modeladas ficaram bem próximas, mostrando que o melhor desempenho da RNA-MCC sobre a RNA-MSE na etapa de previsão de vazões não foi transmitido para o passo da formação de preço *spot*. Essa constatação pode ser justificada pelo fato de as vazões previstas pelos dois modelos estarem, na maior parte do tempo analisado, no mesmo patamar de condição hidrológica, mesmo com a redução de erros proporcionada pela RNA-MCC. Nessa perspectiva, o despacho de usinas térmicas calculado pelo método de programação linear do DESSEM não possui muitas variações entre os dois cenários, o que leva a PLDs próximos das duas redes neurais.

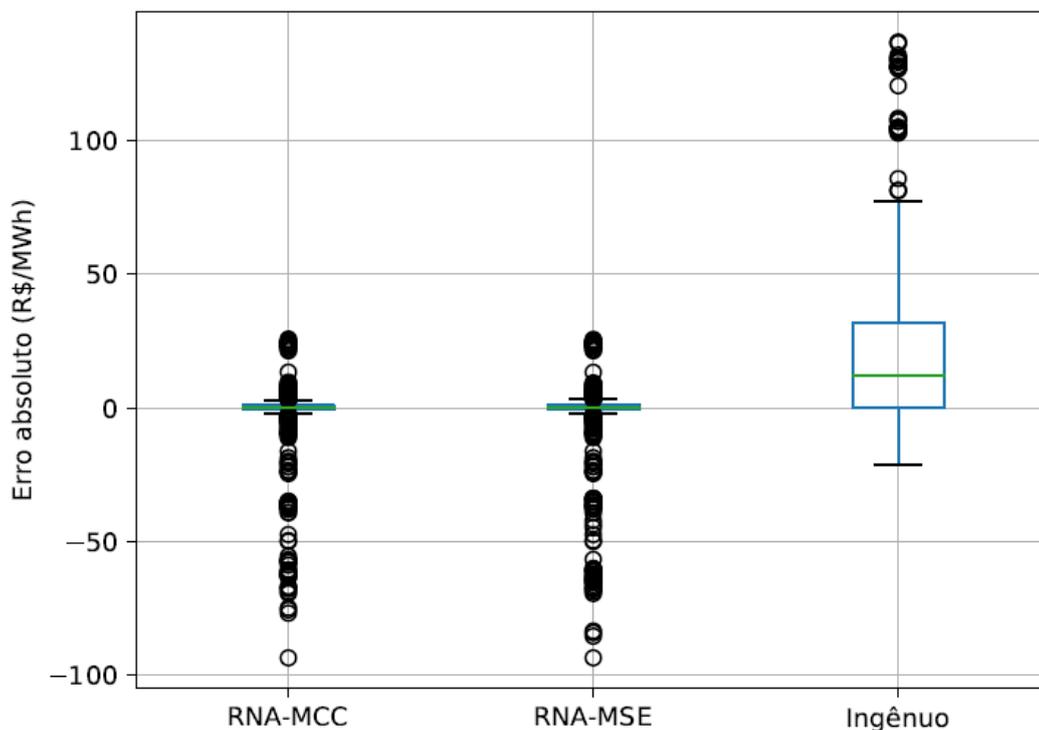


Figura 25 - Boxplot com erro de previsão do PLD do submercado Nordeste.

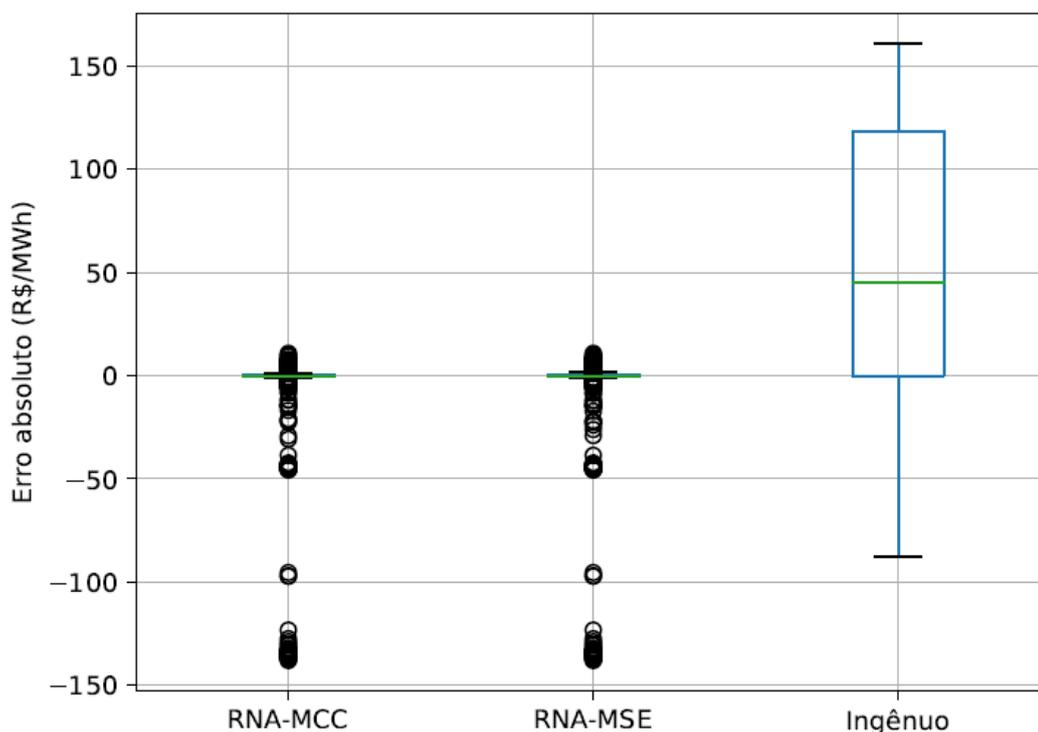


Figura 26 - Boxplot com erro de previsão do PLD do submercado Norte.

Outra maneira de se avaliar os modelos e seus respectivos PLDs é feita através dos gráficos de linhas das 13 sextas-feiras analisadas, expostos da Figura 27 até a Figura 39. Em cada gráfico, são contabilizados 48 estágios de tempo, uma vez que o DESSEM gera valores de PLD em períodos de 30 em 30 minutos para o primeiro dia do horizonte de previsão. As conclusões tomadas dos gráficos de *boxplot* são passadas para os gráficos de linha. Por exemplo, de uma forma geral, as curvas pertencentes aos modelos RNA-MCC e RNA-MSE seguem a referência da CCEE com maior proximidade quando comparadas à curva do predictor ingênuo, inclusive captando movimentos de subida e descida dos preços em momentos de maior volatilidade. Os maiores descolamentos de preços das redes neurais ocorrem nos submercados Nordeste e Norte, sobretudo nos dois primeiros dias estudados. Nesses mesmos submercados, o modelo ingênuo apresenta preços bem maiores que a CCEE, motivados pela não extração da sazonalidade de vazões por parte do predictor, como explicado anteriormente. Adicionalmente, as curvas de preço da RNA-MCC e RNA-MSE seguem o mesmo perfil na maior parte do tempo, o que também foi indicado nos gráficos de *boxplot*. O resumo do desempenho da previsão de PLD dos modelos pode ser visto na Tabela 9 com valores de NMAE em relação à referência da CCEE. Considerando todos os submercados, as redes neurais têm valores de NMAE equivalentes (4,20% da RNA-MCC e 4,21% da RNA-MSE), enquanto o modelo ingênuo possui erro de 17,06%.

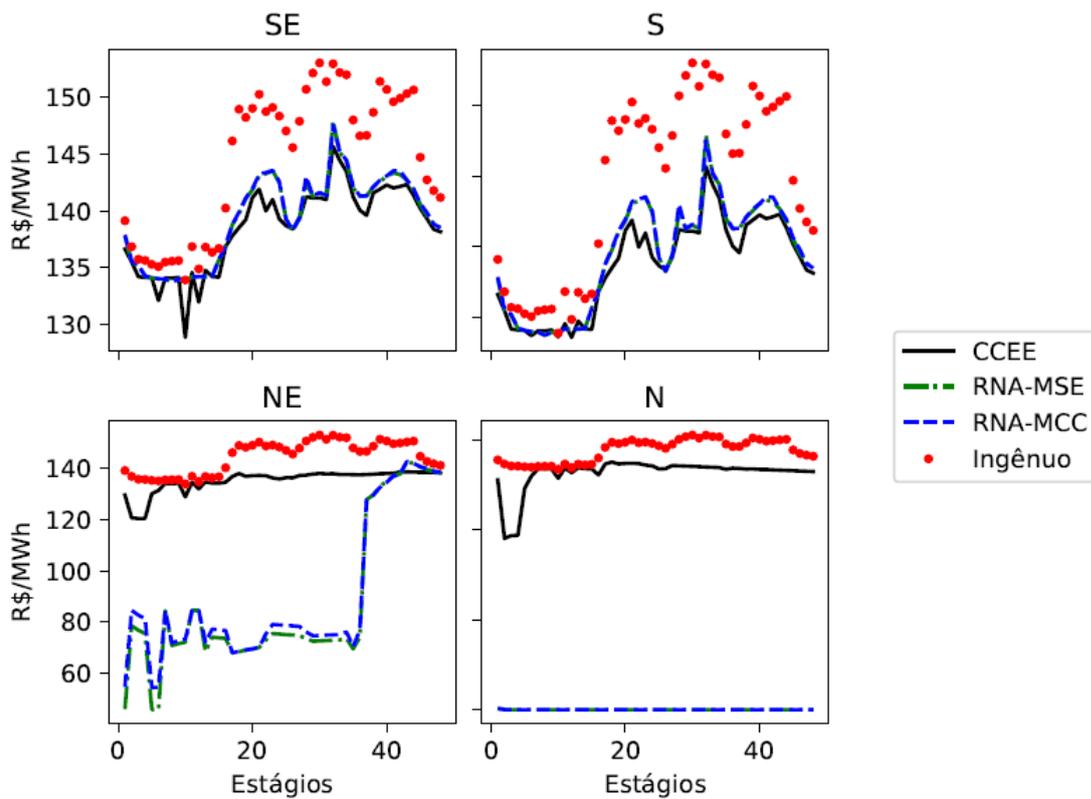


Figura 27 - Curvas de PLD para o dia 05/03/2022.

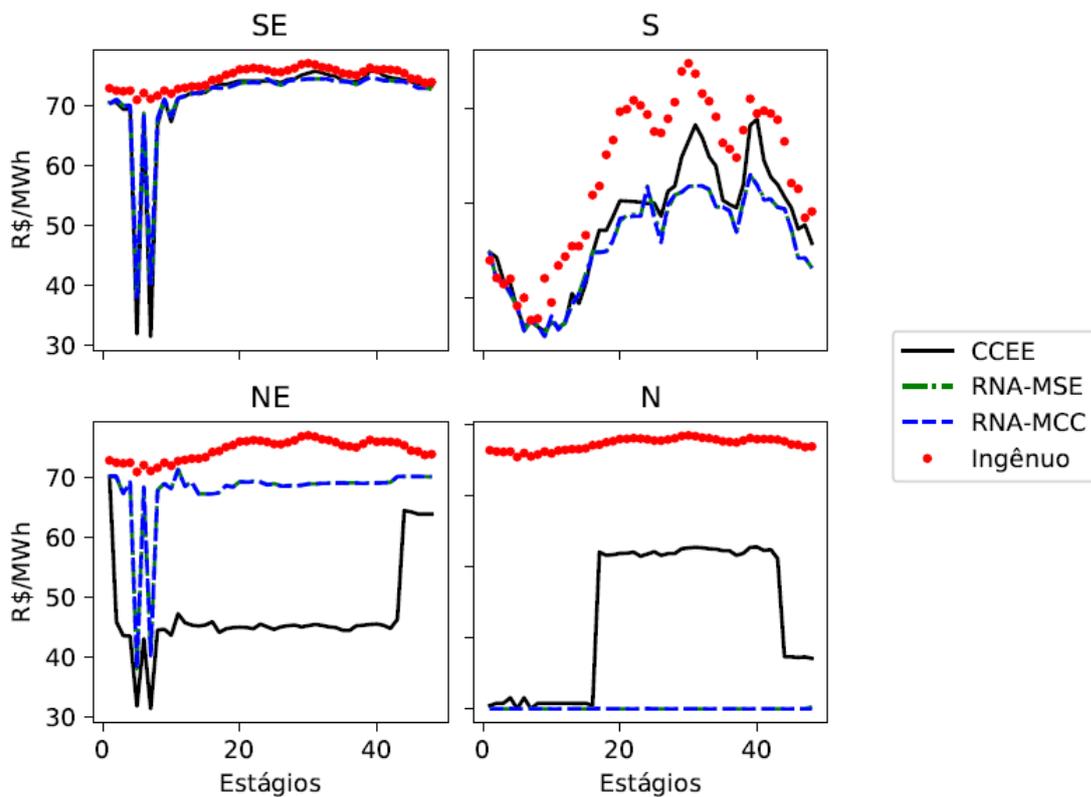


Figura 28 - Curvas de PLD para o dia 12/03/2022.

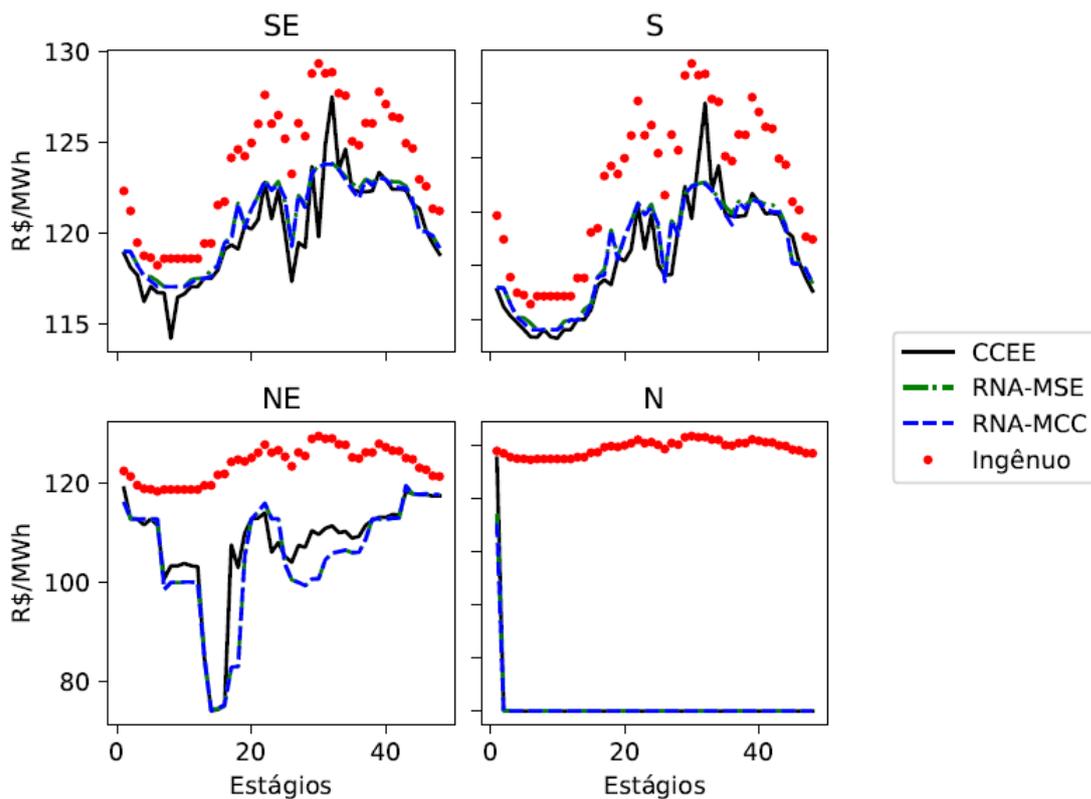


Figura 29 - Curvas de PLD para o dia 19/03/2022.

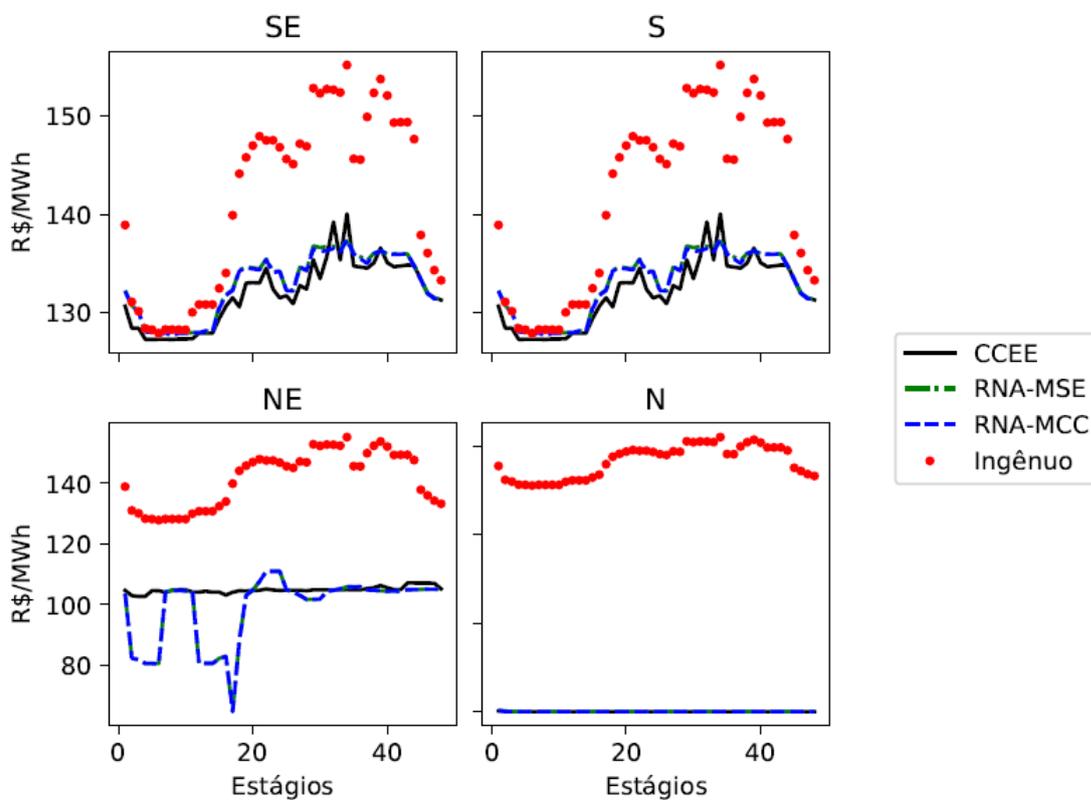


Figura 30 - Curvas de PLD para o dia 26/03/2022.

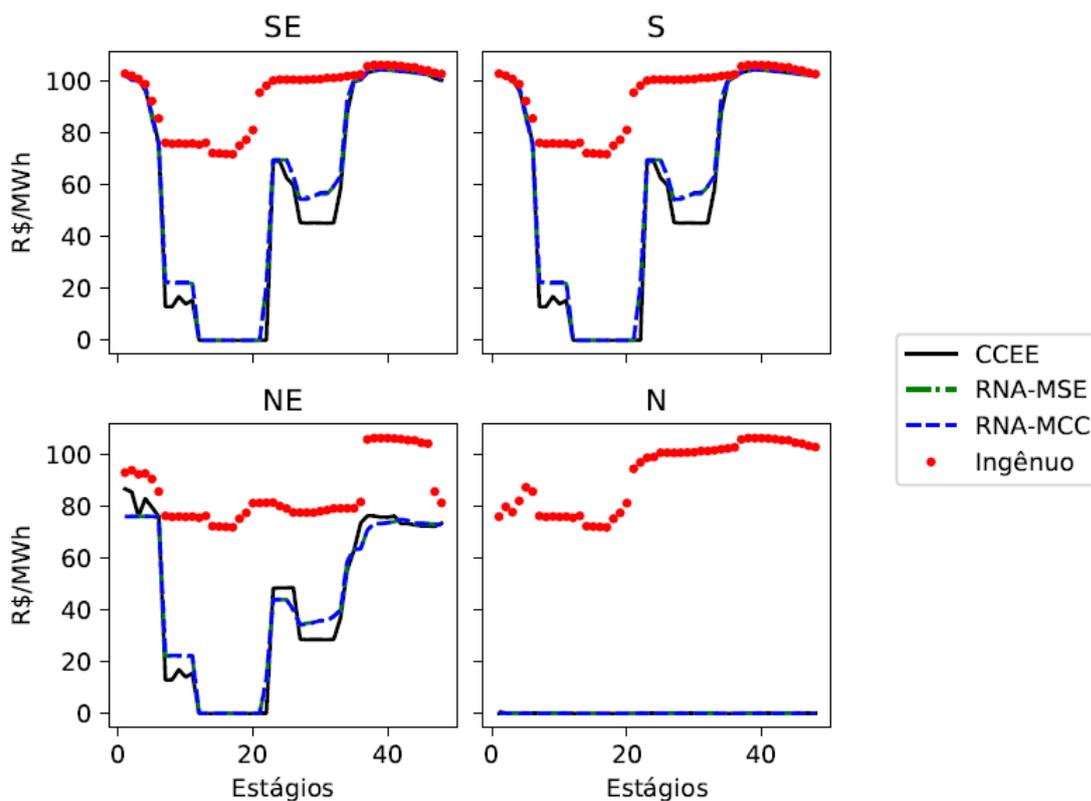


Figura 31 - Curvas de PLD para o dia 02/04/2022.

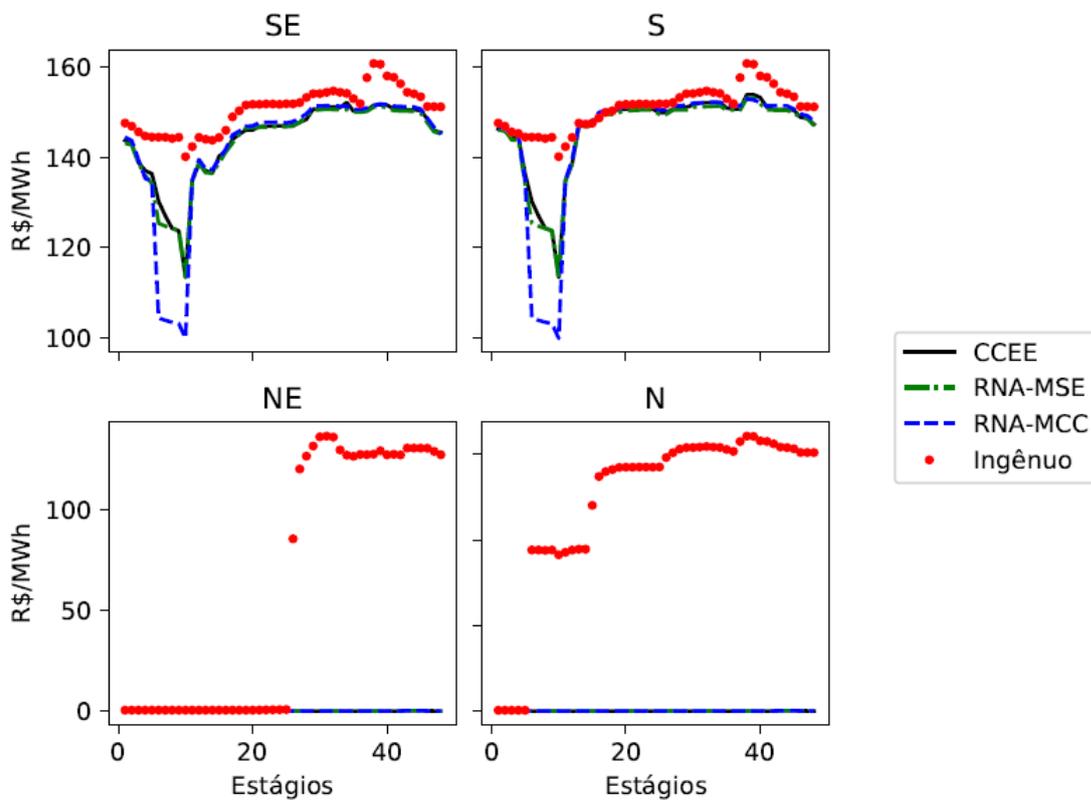


Figura 32 - Curvas de PLD para o dia 09/04/2022.

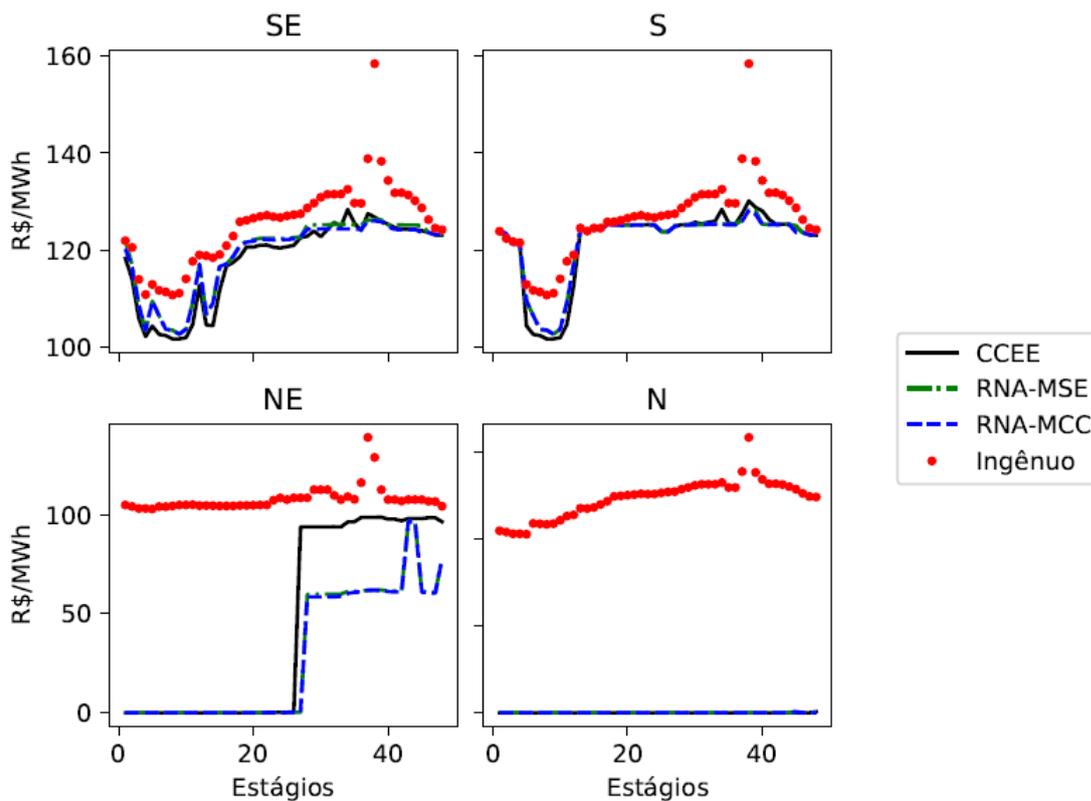


Figura 33 - Curvas de PLD para o dia 16/04/2022.

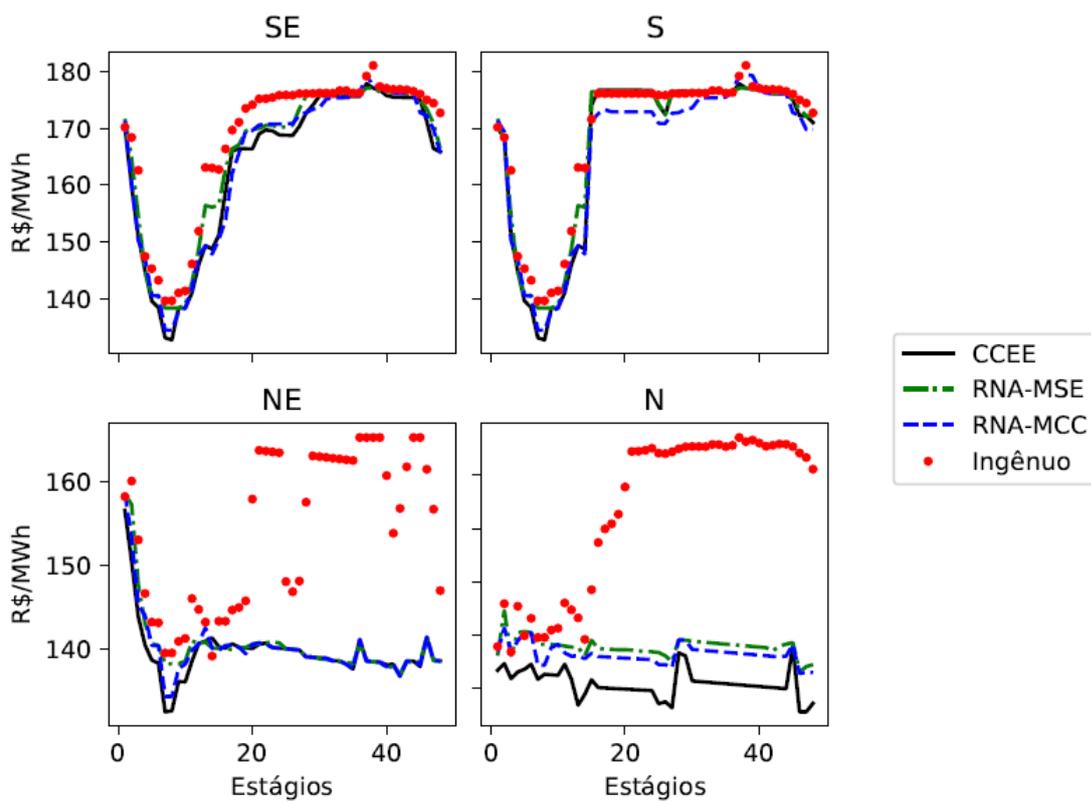


Figura 34 - Curvas de PLD para o dia 23/04/2022.

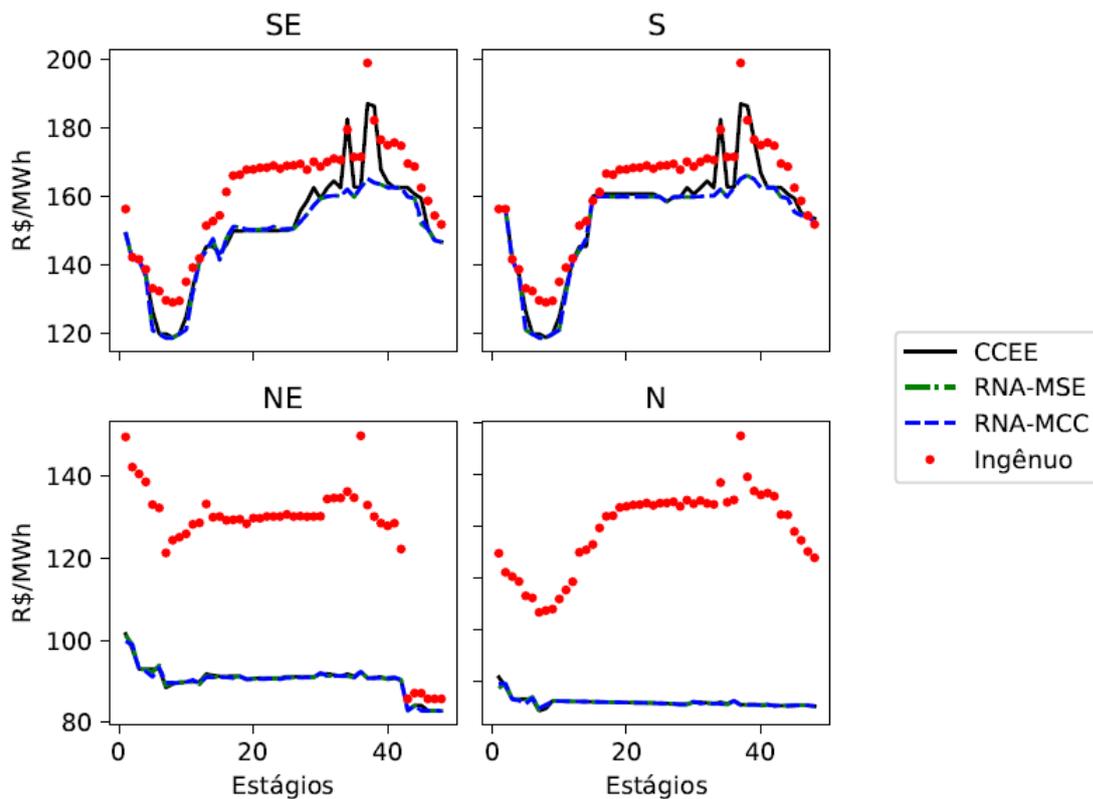


Figura 35 - Curvas de PLD para o dia 30/04/2022.

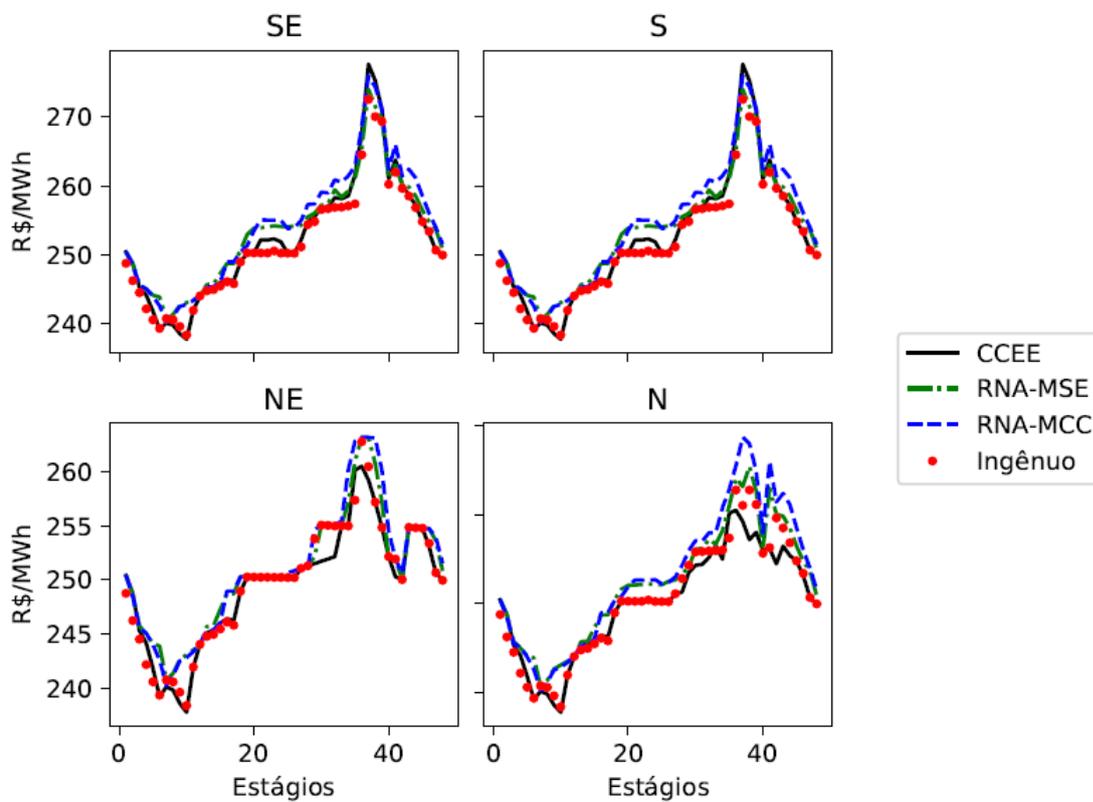


Figura 36 - Curvas de PLD para o dia 04/06/2022.

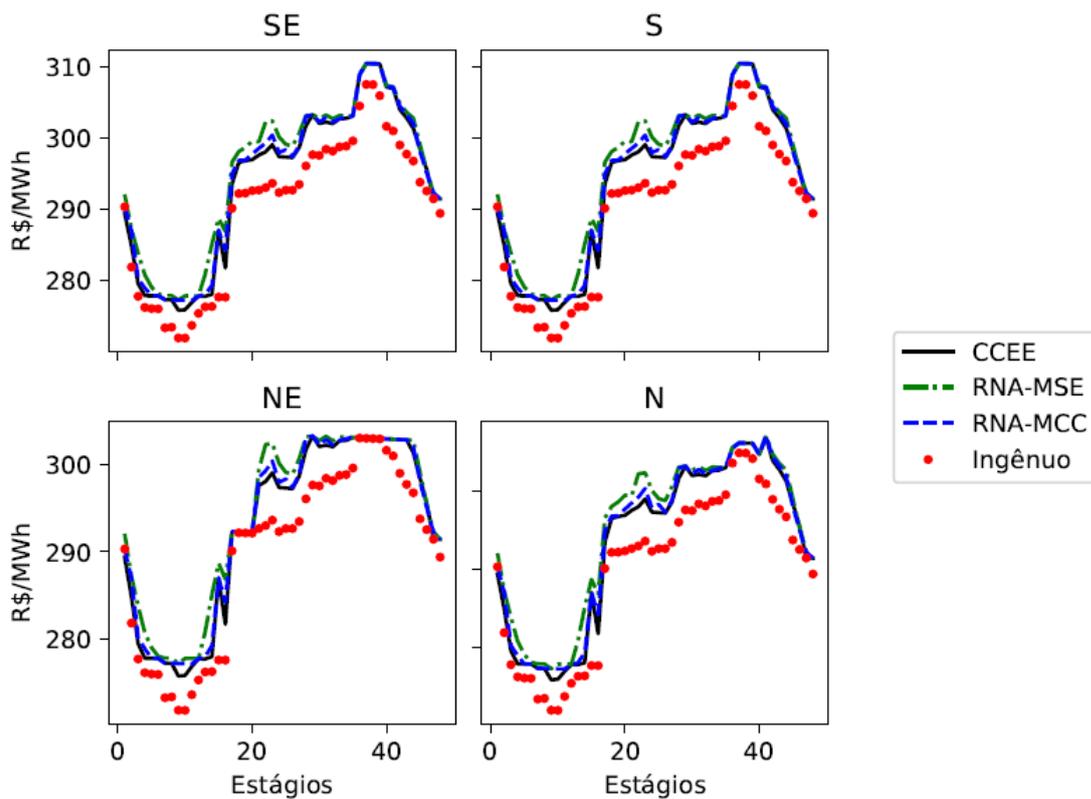


Figura 37 - Curvas de PLD para o dia 11/06/2022.

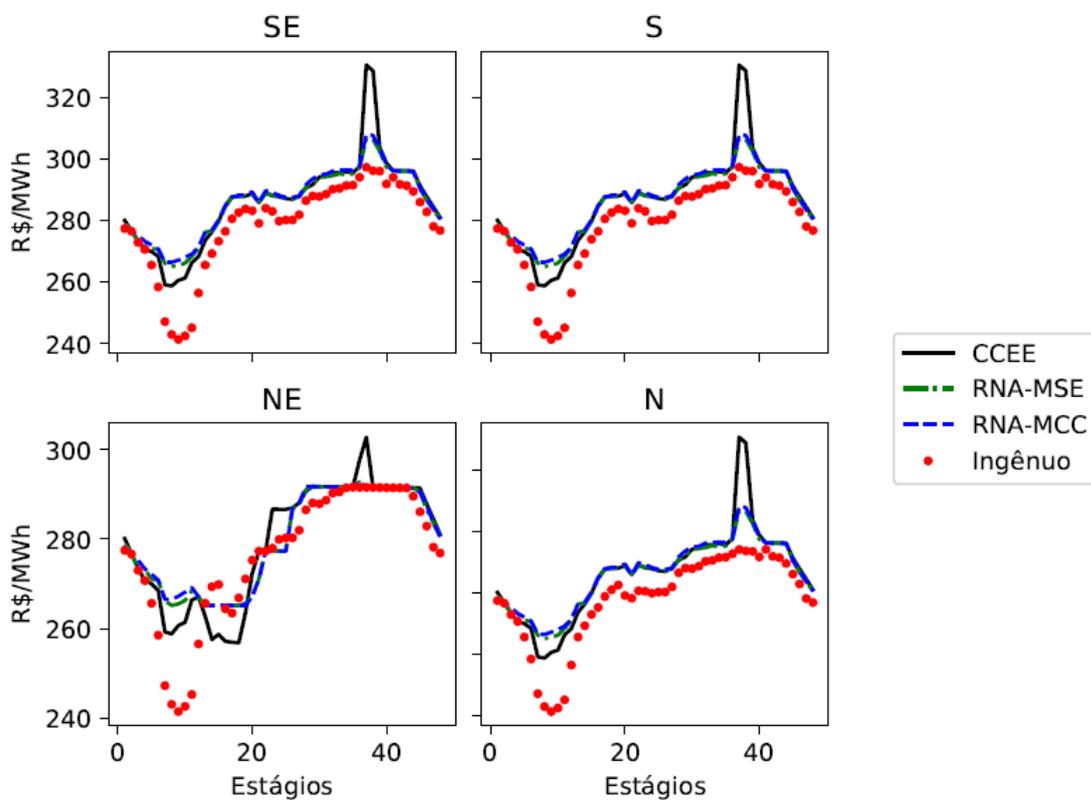


Figura 38 - Curvas de PLD para o dia 18/06/2022.

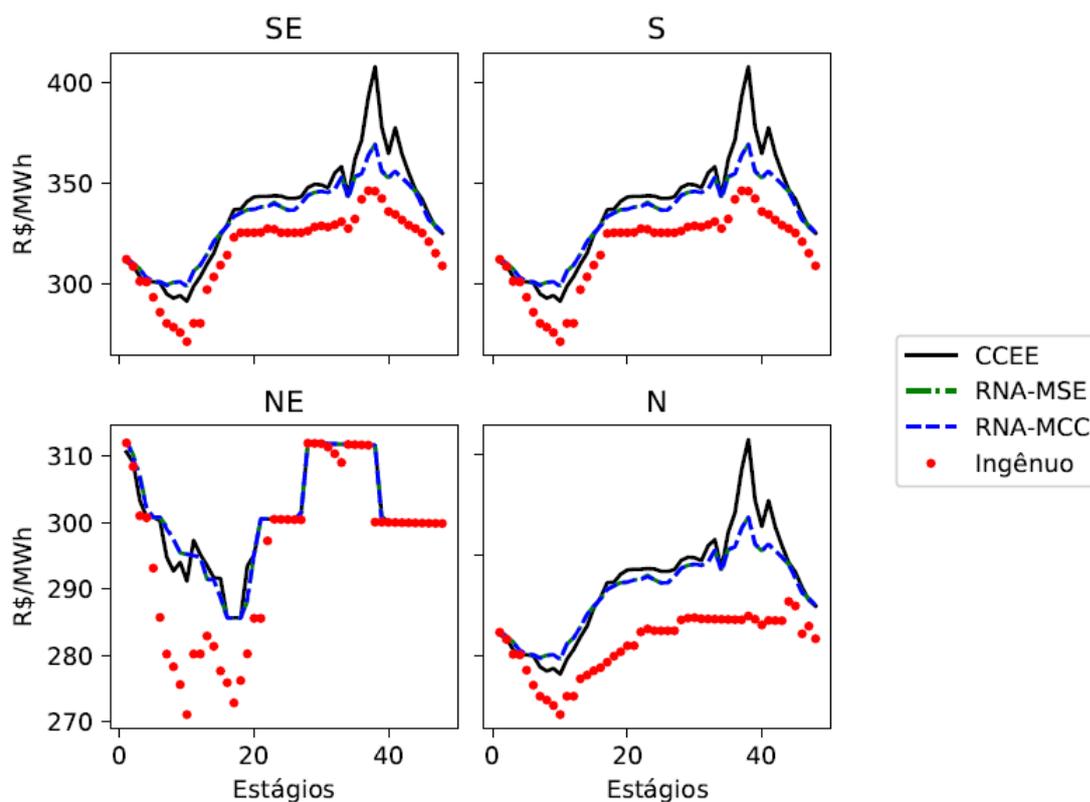


Figura 39 - Curvas de PLD para o dia 25/06/2022.

Tabela 9 - NMAE dos PLDs dos modelos em relação aos valores oficiais da CCEE

Submercados	RNA-MCC	RNA-MSE	Ingênuo
Sudeste/Centro-Oeste	1,18 %	1,14 %	5,30 %
Sul	1,12 %	1,01 %	4,67 %
Nordeste	5,64 %	5,78 %	17,55 %
Norte	11,44 %	11,53 %	51,90 %
Todos	4,20 %	4,21 %	17,06 %

A última análise feita é ilustrada da Figura 40 a Figura 52, com gráficos de linha para as mesmas 13 sextas-feiras dos gráficos anteriores, mas com a inclusão das curvas de preços referentes ao caso em que a vazão realizada foi incorporada como afluência prevista nos arquivos DADVAZ. Dos gráficos a seguir, um resultado interessante foi observado: em média e para os quatro submercados, as curvas de preços do modelo RNA-MCC seguem mais de perto os valores do caso de vazões realizadas do que as referências oficiais da CCEE. Levando em consideração que as previsões de vazões do ONS tiveram menores médias de erros em mais hidrelétricas do que as previsões da RNA-MCC, pode-se supor

que a rede neural obteve melhor desempenho na previsão de aflúncias das UHEs mais influentes à formação de PLD.

A Tabela 10 apresenta de forma quantitativa os resultados descritos no parágrafo anterior, através dos valores de NMAE dos casos RNA-MCC e CCEE em relação aos PLDs do cenário de hidrologia real. Os preços do modelo proposto possuem menores erros em todos os submercados, com desempenho geral bem superior ao da CCEE: 0,76% de NMAE para a RNA-MCC e 4,36% para a CCEE.

Dos resultados obtidos na Figura 40 a Figura 52 e na Tabela 10, dois pontos de vista podem ser tomados: aquele relativo ao conjunto CCEE/ONS e outro referente ao comercializador de energia elétrica. No primeiro caso, a CCEE tem como objetivo acertar a curva de preço do cenário de vazões real, de modo que o PLD oficial seja cada vez mais aderente ao preço referente às condições reais do sistema, fazendo com que seja cada vez menos necessário a adoção de encargos decorrentes de geração despachada após a execução do DESSEM que causam aumentos nos custos globais de operação. Para isso, a previsão de vazões oficial deve ser cada vez mais assertiva. Nessa linha, os resultados mostram margem de melhora no processo preditivo de aflúncias feito pelo ONS, visando a aproximação da curva de PLD oficial em relação aos preços do cenário de vazão realizada. Já na perspectiva de uma empresa comercializadora, o alvo não é a curva de PLD das aflúncias performadas, mas sim os preços *spot* oficiais da CCEE. Assim, é interessante que o processo de previsão de vazões considere a minimização das amostras de erro em relação às aflúncias previstas do ONS, informação que, de fato, vai ser utilizada como dado de entrada no modelo de otimização do despacho hidrotérmico de curtíssimo prazo.

Como mencionado anteriormente, a revisão literária feita neste trabalho indicou que a relação entre a aflúncias previstas e a formação do PLD em resolução horária ainda não foi diretamente estudada devido à recente implantação do DESSEM oficialmente. Em comparação qualitativa, a previsão de preço *spot* feita em [67] apresentou desempenho satisfatório assim como neste trabalho, porém tratada em base semanal e com formação de preço feita através do uso direto de RNAs, não sendo a forma oficial de geração de PLD semanal (que seria feita através do uso do modelo DECOMP). Nessa linha, os resultados apresentados neste trabalho têm caráter precursor, mostrando que o escopo tratado pode ser explorado em trabalhos futuros.

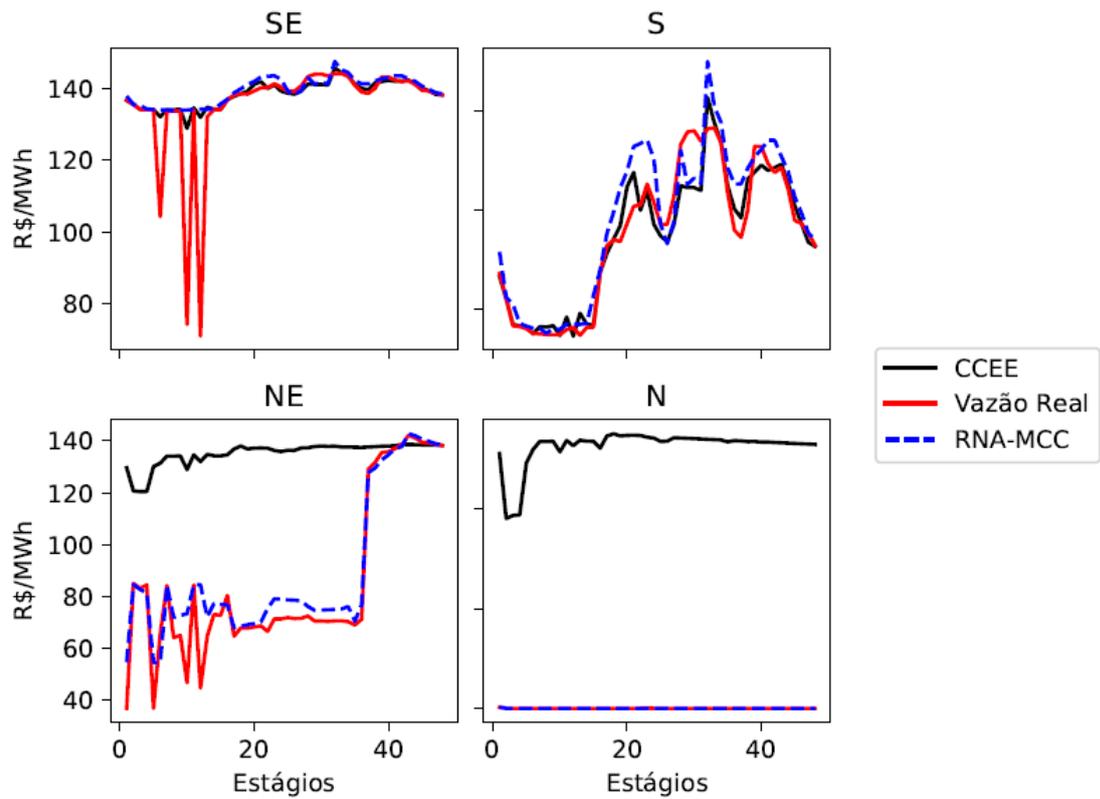


Figura 40 - Curvas de PLD para o dia 05/03/2022, considerando preços do caso de vazão realizada.

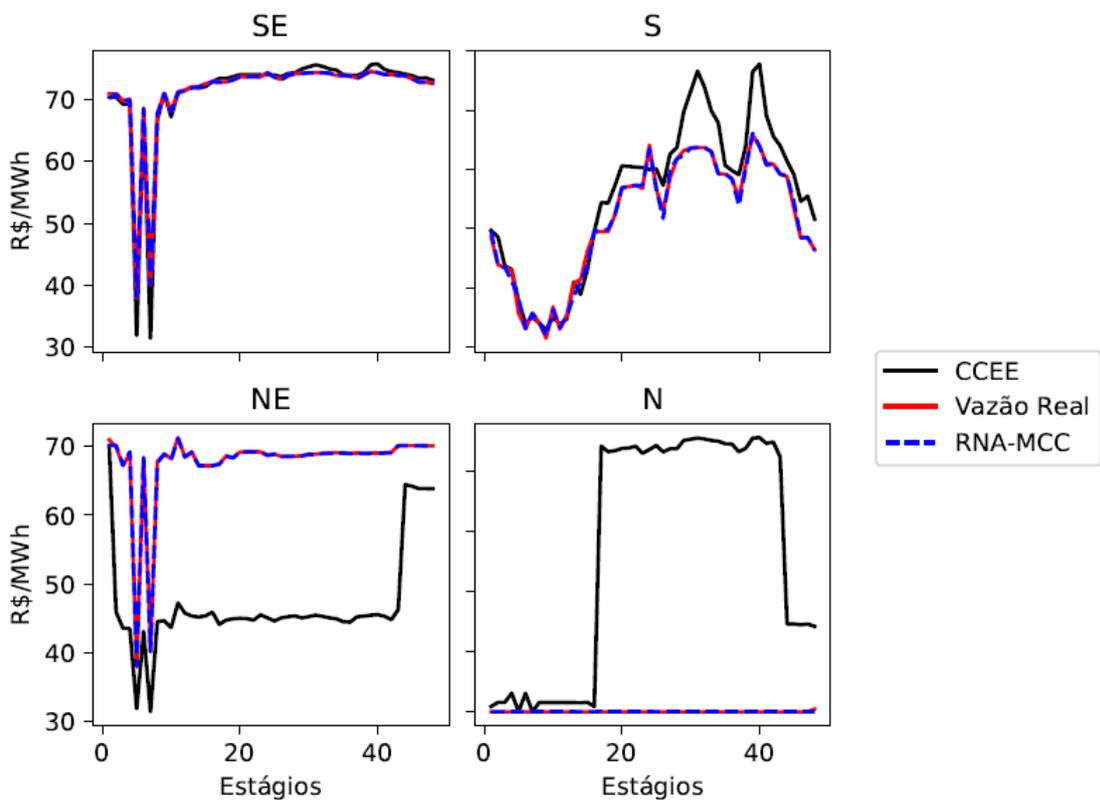


Figura 41 - Curvas de PLD para o dia 12/03/2022, considerando preços do caso de vazão realizada.

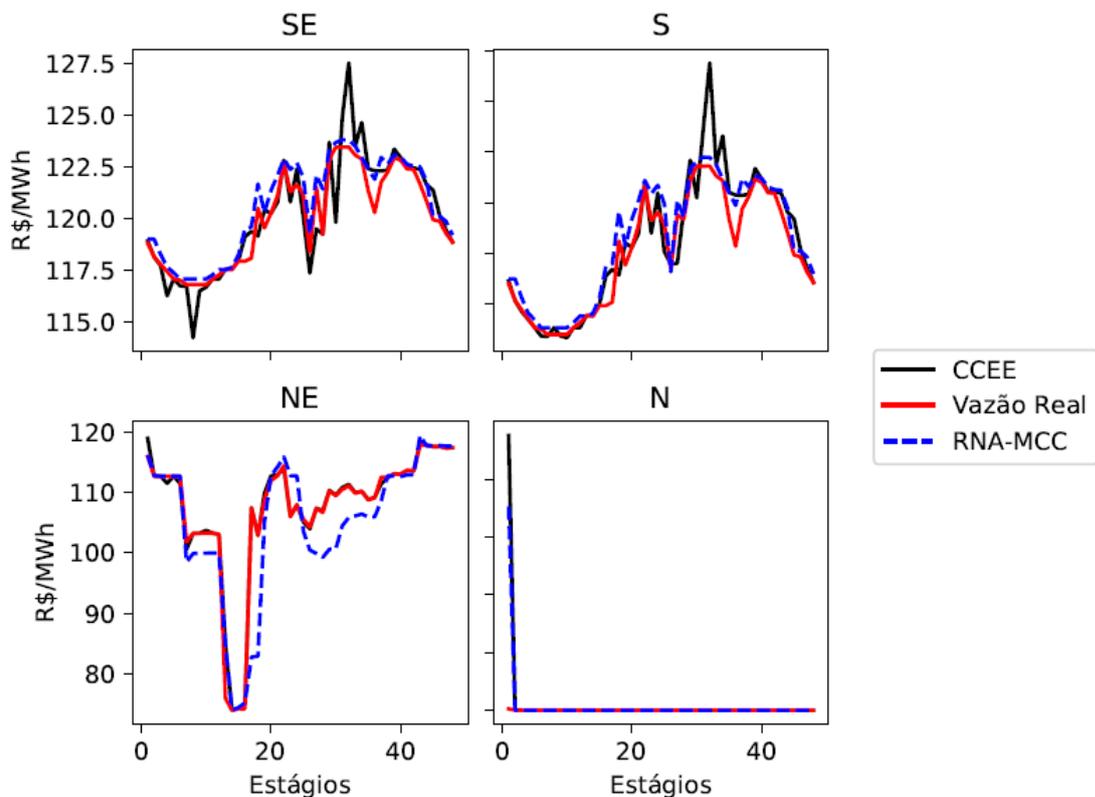


Figura 42 - Curvas de PLD para o dia 19/03/2022, considerando preços do caso de vazão realizada.

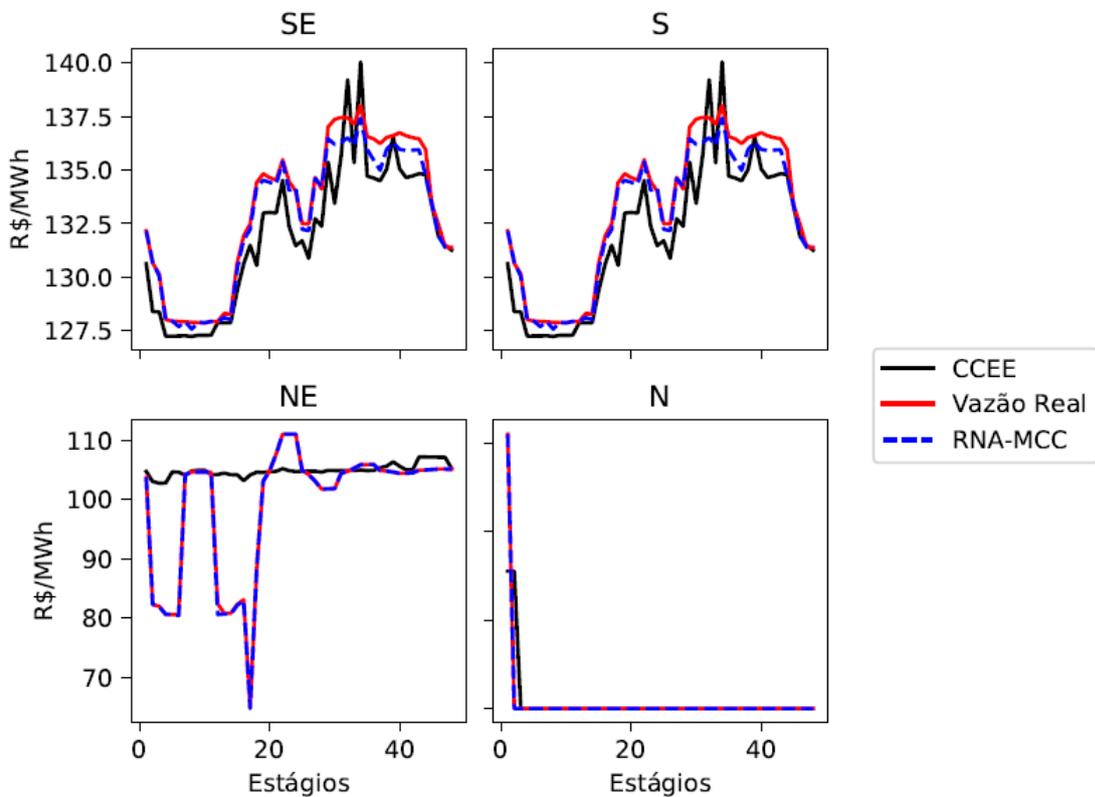


Figura 43 - Curvas de PLD para o dia 26/03/2022, considerando preços do caso de vazão realizada.

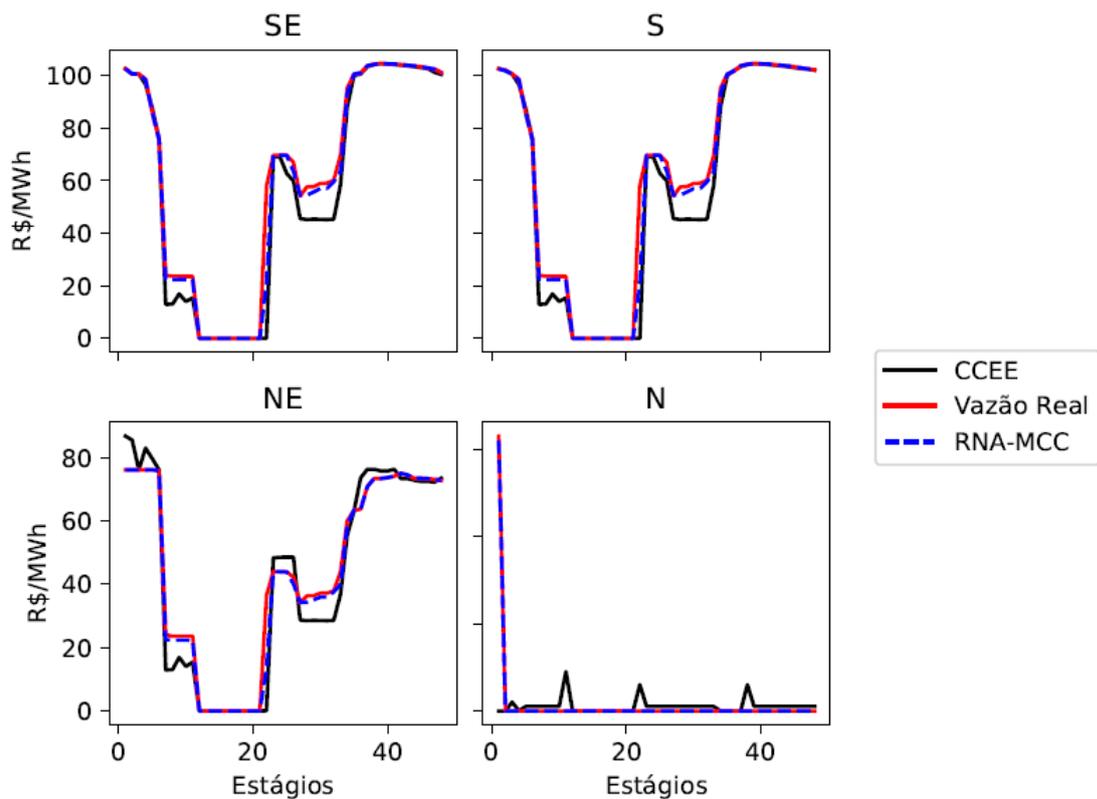


Figura 44 - Curvas de PLD para o dia 02/04/2022, considerando preços do caso de vazão realizada.

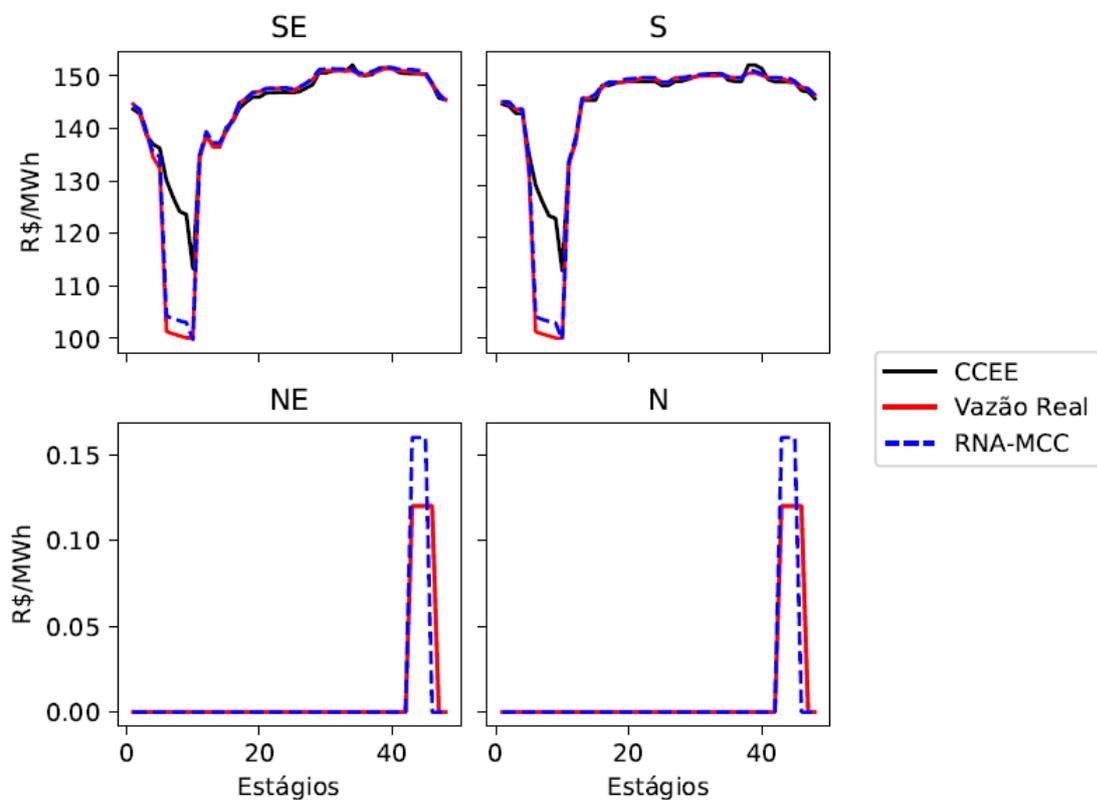


Figura 45 - Curvas de PLD para o dia 09/04/2022, considerando preços do caso de vazão realizada.

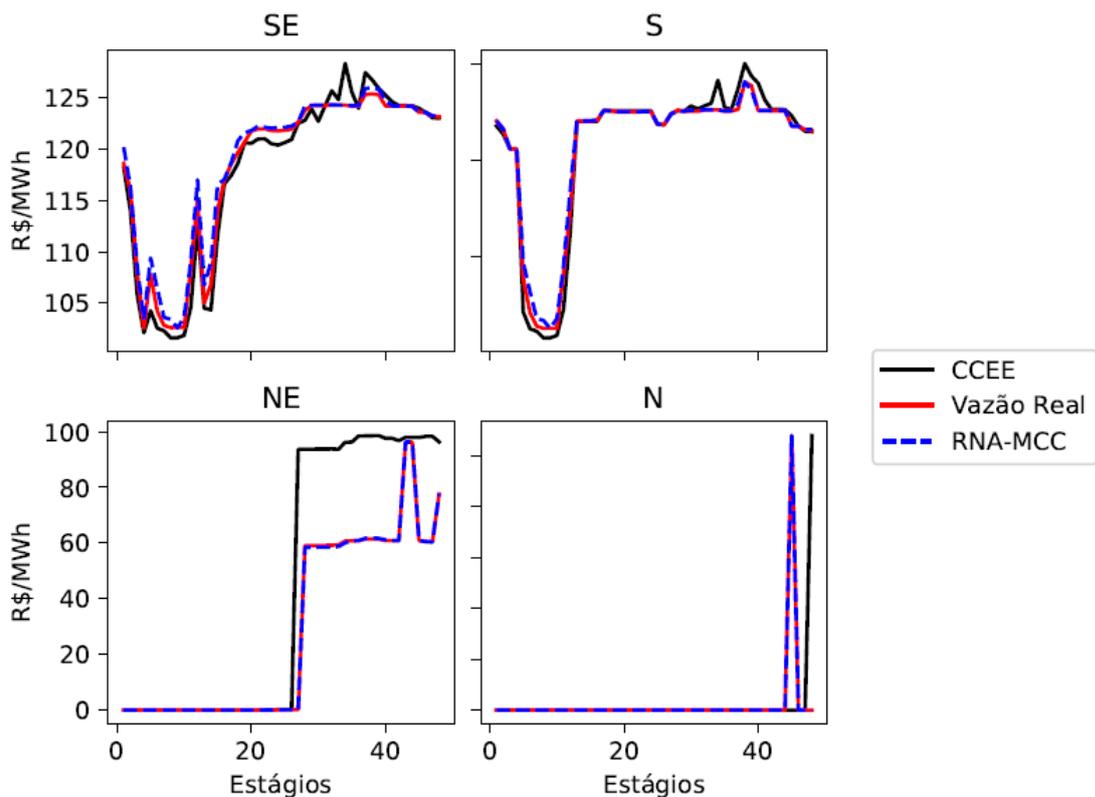


Figura 46 - Curvas de PLD para o dia 16/04/2022, considerando preços do caso de vazão realizada.

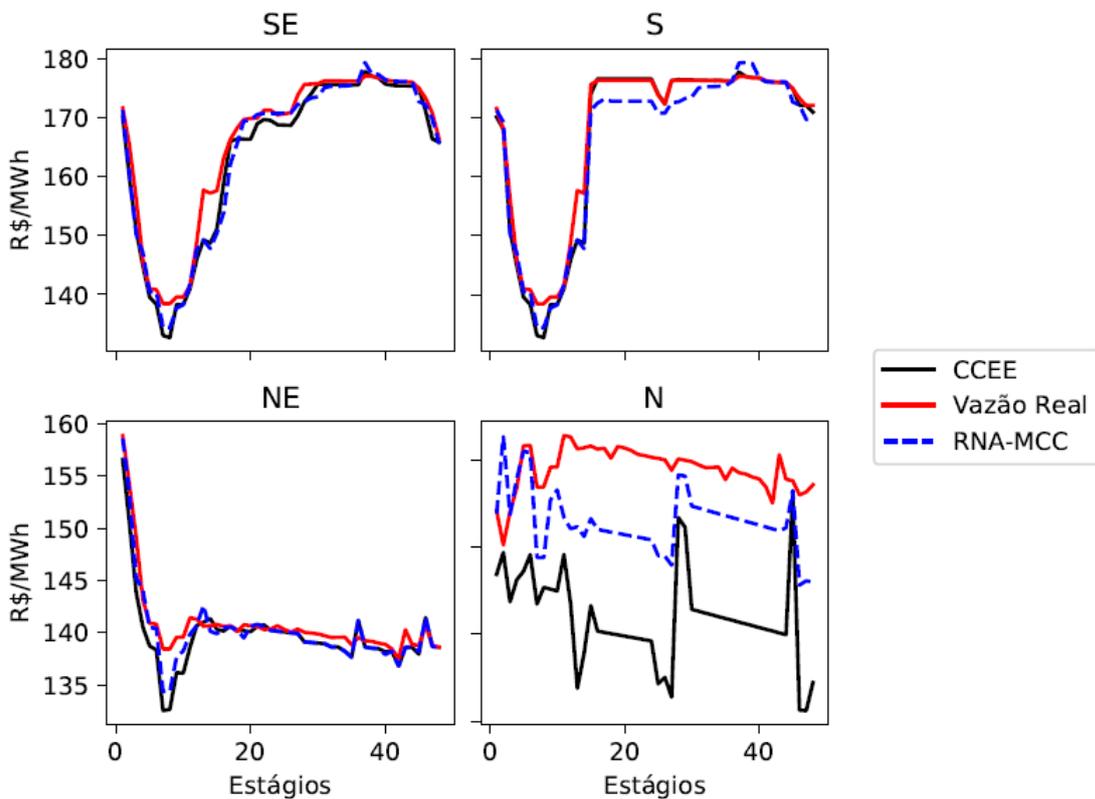


Figura 47 - Curvas de PLD para o dia 23/04/2022, considerando preços do caso de vazão realizada.

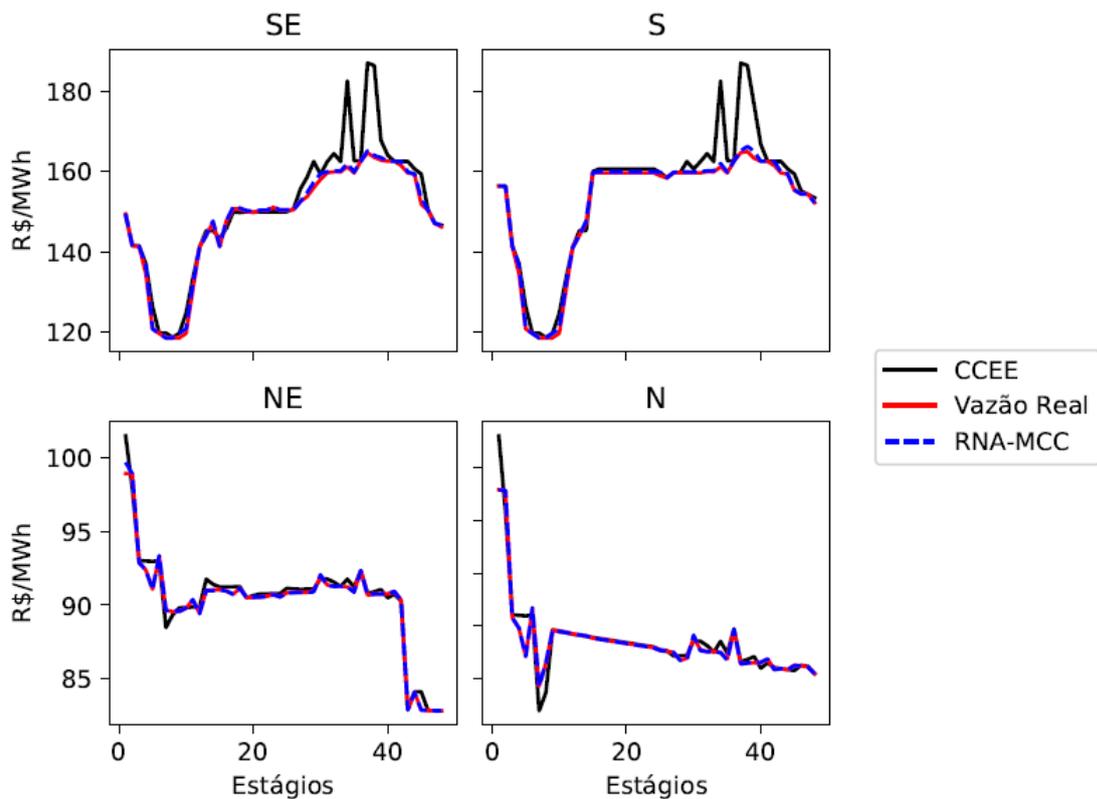


Figura 48 - Curvas de PLD para o dia 30/04/2022, considerando preços do caso de vazão realizada.

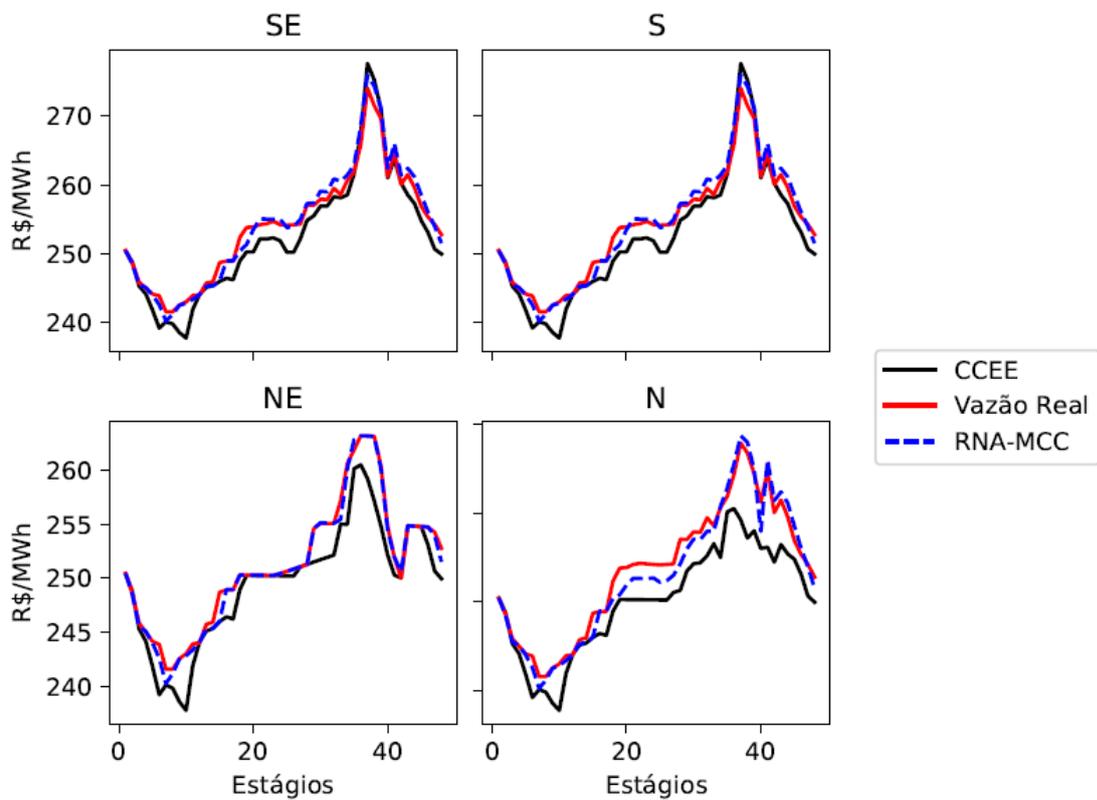


Figura 49 - Curvas de PLD para o dia 04/06/2022, considerando preços do caso de vazão realizada.

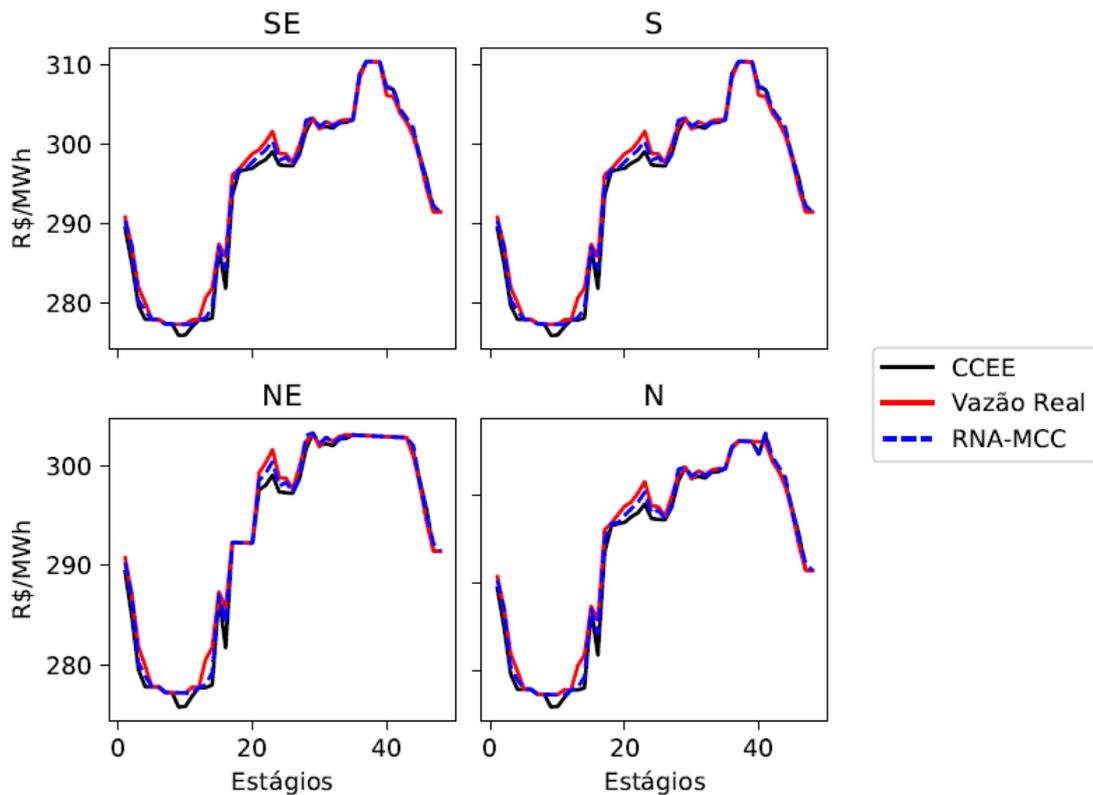


Figura 50 - Curvas de PLD para o dia 11/06/2022, considerando preços do caso de vazão realizada.

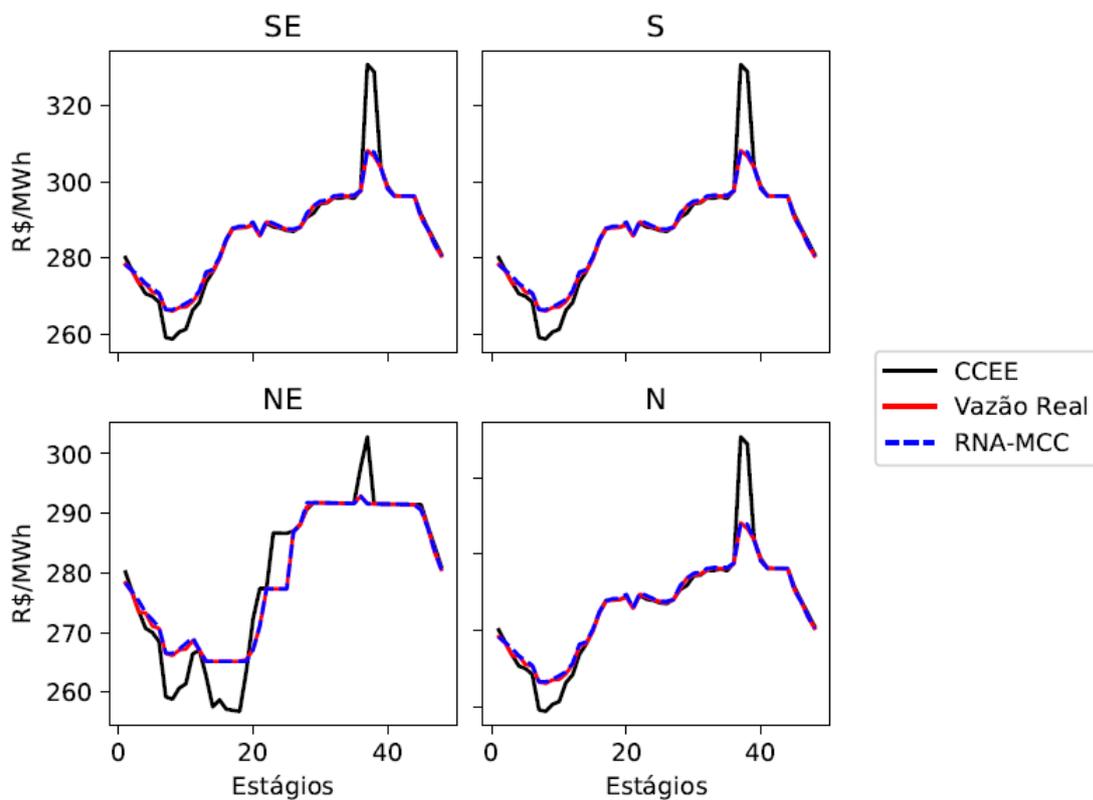


Figura 51 - Curvas de PLD para o dia 18/06/2022, considerando preços do caso de vazão realizada.

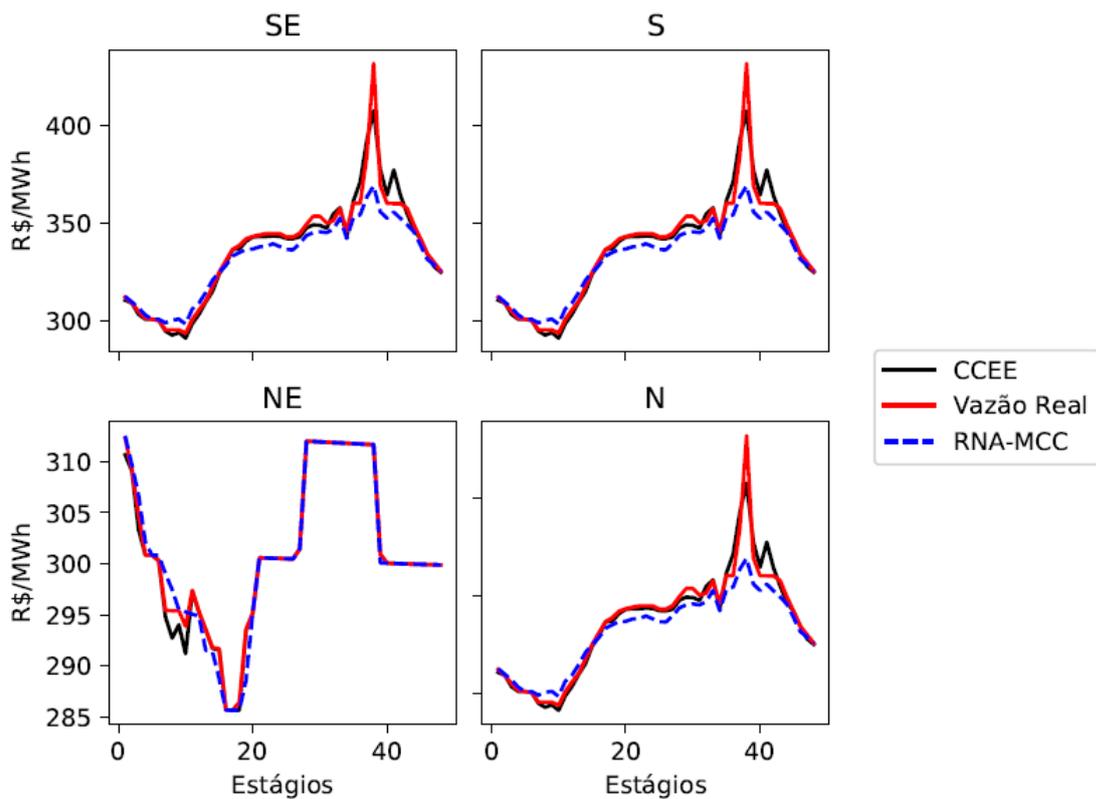


Figura 52 - Curvas de PLD para o dia 25/06/2022, considerando preços do caso de vazão realizada.

Tabela 10 - NMAE dos PLDs da RNA-MCC e da CCEE em relação aos preços do cenário de vazões realizadas

Submercados	RNA-MCC	CCEE
Sudeste/Centro-Oeste	0,84 %	1,26 %
Sul	0,66 %	1,03 %
Nordeste	0,65 %	5,93 %
Norte	0,92 %	12,79 %
Todos	0,76 %	4,36 %

Capítulo 7 - Conclusão

A previsão de vazões afluentes aos reservatórios de usinas hidrelétricas do SIN é um dado de entrada do processo de otimização do despacho hidrotérmico, atividade feita através da utilização de modelos computacionais executados pelo ONS e CCEE com diferentes propósitos. Nessa perspectiva, é desejável que as afluições previstas tenham cada vez menores erros, de forma que as saídas do processo – entre elas, o preço *spot* do mercado de energia elétrica – sejam mais consistentes em relação às reais condições do sistema.

O uso de redes neurais artificiais modeladas com a função de regressão de séries temporais é uma opção fora do conjunto de previsores adotados pelo ONS para a geração de vazões previstas. Dentre as vantagens das RNAs, podem ser citadas a capacidade de adaptabilidade frente às mudanças do meio estudado e o potencial de extração de informações dos dados, como traços de sazonalidade, periodicidade e tendência.

Convencionalmente, o MSE é aplicado na etapa de aprendizagem de redes neurais voltadas para a atividade de regressão, o que é uma boa abordagem considerando a ocorrência de distribuições de erros de previsão com perfil gaussiano. Como ilustrado na introdução deste trabalho, as distribuições referentes às amostras de erro de previsões de vazões feitas pelo ONS para algumas hidrelétricas apresentam valores de curtose e assimetria que descaracterizam as curvas do perfil normal simétrico, indicando que, para esses casos, o treino de RNAs baseado na minimização do erro quadrático médio não é uma solução ótima. Nesse cenário, os conceitos de *Information Theoretic Learning, framework* derivado do campo da Teoria da Informação, proporcionam análise estatística mais apurada, com possibilidade de aplicação em estudos que envolvam distribuições não gaussianas e com *outliers* bastante distantes da média.

Este trabalho apresentou o desenvolvimento de um previsor, nomeado como RNA-MCC, modelado a partir da combinação de técnicas de RNAs e ITL, sendo uma alternativa complementar ao conjunto de modelos adotados pelo ONS para a previsão de vazões afluentes diárias. Após a geração de afluições previstas, os valores foram incorporados em um dos arquivos de entrada do modelo de otimização do despacho hidrotérmico de curtíssimo prazo, o DESSEM, proporcionando a produção de PLDs horários após a

execução do programa. Duas análises foram feitas a partir dos dados obtidos: uma do ponto de vista hidrológico, com o estudo de exatidão das previsões de vazões criadas através de testes de hipóteses aplicados às médias dos erros de previsão por UHE e dia previsto, e outra no escopo da formação de preço, avaliando-se o impacto das afliências previstas nos valores de PLD com base em análise gráfica (*boxplots* dos desvios de preço e gráficos de linha com as curvas de PLD). Com o objetivo de validar o desempenho da RNA-MCC, mais dois previsores foram criados, sendo uma rede neural, nomeada RNA-MSE, cuja etapa de treino foi desenvolvida com o MSE como função de erro e um modelo ingênuo que utilizou as médias históricas de vazões das hidrelétricas como previsões.

Dos testes de hipóteses realizados, os resultados mostraram que a RNA-MCC, em comparação ao ONS, obteve médias de erros de previsão menores ou equivalentes em mais da metade das hidrelétricas contempladas no estudo, considerando o horizonte do 1º dia previsto. Tal desempenho não se mantém com o aumento do horizonte de previsão, o que pode ser justificado pela propagação do erro presente no processo recursivo de previsão da rede neural. Em relação a RNA-MSE, a RNA-MCC apresentou desempenho de previsão melhor ou equivalente em quase todas as usinas, indicando que a adoção de técnicas de *Information Theoretic Learning* foi benéfica para o passo de aprendizagem do modelo, o que vai ao encontro da perspectiva inicial de melhora da análise estatística de distribuições de erros não gaussianas. Quanto ao modelo ingênuo, como esperado, as médias históricas, tomadas como previsões, fizeram com que o predictor não fosse melhor em nenhuma das UHEs, na comparação com a RNA-MCC.

A avaliação hidrológica ruim do modelo ingênuo foi transmitida para as curvas de PLD horário e *boxplots* de erros gerados, resultados que evidenciam o fato de que previsões de vazões de má qualidade podem comprometer os valores de preço *spot* previstos. Essa conclusão ficou ainda mais clara no enfoque dos submercados Nordeste e Norte, uma vez que a média histórica de vazões não captou o viés de alta das afliências de hidrelétricas do Norte no primeiro semestre do ano, fazendo com que as curvas de preços do predictor ficassem bem mais altas do que os números oficiais da CCEE.

Entre os modelos RNA-MCC e RNA-MSE, os PLDs gerados apresentaram poucas diferenças, o que leva à conclusão de que, apesar da redução de erros de previsão de vazões propiciada pela rede neural com traços de ITL, as previsões de afliências dos dois modelos proporcionaram cenários de condições hidrológicas equivalentes na maior parte do tempo e, conseqüentemente, levaram a cenários de geração térmica parecidos, conforme

otimização pelo DESSEM. Já na comparação com a referência da CCEE, os preços das redes neurais mostraram desempenho satisfatório, acompanhando as curvas oficiais com proximidade (maiores descolamentos foram percebidos apenas nos dois primeiros dias estudados, especificamente nos submercados Nordeste e Norte) e seguindo movimentos de descida e subida de preços com maior volatilidade.

Um resultado interessante foi verificado na inclusão do cenário de preços considerando a vazão realizada: em média, as curvas de PLD da RNA-MCC ficaram mais próximas dos valores do caso de aflúncias reais do que as linhas de referência da CCEE. Como as médias de erros de previsões de vazões do ONS foram menores para mais hidrelétricas do que para a rede neural, pode-se concluir que as previsões do modelo proposto se mostraram mais exatas que as dos modelos oficiais nas usinas com maior influência na formação de preço. Vale ressaltar que o objetivo da modelagem da RNA-MCC não é estimular a substituição dos previsores utilizados pelo ONS, mas sim indicar o potencial de melhora no processo de previsão de vazões no escopo da programação diária da operação do SIN.

Na perspectiva de uma empresa comercializadora de energia elétrica, que tenta antecipar o valor do preço do mercado de curto prazo com base na previsão de dados do setor (vazões, carga, entre outros), a redução à zero do erro da previsão de aflúncias é irrelevante se esta não levar à minimização do erro de previsão do preço spot. Tomando como referência os gráficos produzidos neste trabalho, o preço alvo (do ponto de vista do agente comercializador) deve ser a curva da CCEE e não a curva da vazão realizada. Assim, pode-se inferir que o mais importante não é acertar os valores realizados de vazões, mas sim as previsões de aflúncias feitas pelo ONS, informação que, de fato, é inserida nos decks oficiais da CCEE e, conseqüentemente, impacta a formação do PLD oficial.

Em suma, o trabalho apresentou resultados satisfatórios, tanto na parte hidrológica quanto na etapa de formação de preço. Os resultados no passo de previsão de vazões indicaram que a rede neural proposta superou os modelos oficiais em quantidade considerável das hidrelétricas. O desempenho do modelo também foi superior ao da RNA com função de erro convencional, evidenciando o bom proveito dos conceitos de ITL no tratamento de erros de previsão de aflúncias. Na etapa de preços, as curvas de PLD horário referentes à RNA-MCC tiveram melhor desempenho que os valores do modelo ingênuo. Este último, por sinal, apresentou resultados esperados, com grandes desvios nas etapas de hidrologia e preço. O resultado mais interessante foi encontrado no caso de

inclusão das curvas de PLD considerando as afluições realizadas, com os preços oriundos do modelo proposto mais aderentes do que os valores oficiais da CCEE. Diante dos dados obtidos e conclusões feitas, pode-se dizer que este trabalho possui contribuições ao estado da arte, tais como a combinação de técnicas de Redes Neurais Artificiais e *Information Theoretic Learning* na atividade de previsão de vazões e a análise de impacto no cálculo do preço *spot* horário do mercado brasileiro de energia elétrica de curto prazo.

7.1 Trabalhos Futuros

As sugestões de melhorias para trabalhos futuros apresentadas a seguir podem ser divididas em aplicações de avanço do estudo e alargamento da experiência. Nas aplicações de avanço do estudo, destaca-se a possibilidade do uso de outras importantes variáveis de entrada do DESSEM (por exemplo, carga elétrica e a geração de fontes intermitentes) como objetos de estudo da análise de impacto na formação do preço *spot*. Outra sugestão seria a adoção de PLDs previstos via modelagem proposta nas estratégias de compra e venda de energia de uma empresa comercializadora fictícia de energia elétrica, visando a análise de impacto financeiro nas operações de *trading*, o que, no final, acaba sendo o principal objeto de interesse de um agente comercializador. Os parágrafos seguintes abordam sugestões do escopo de alargamento da experiência.

A partir do que foi concluído previamente neste capítulo, um ponto de melhoria para o estudo é a substituição das afluições realizadas utilizadas como dados de entrada do modelo, passando-se a adotar as previsões de vazões publicadas pelo ONS. Outro *input* a ser incorporado ao estudo é a previsão de precipitação diária, além de representação mais apurada da informação de chuva (por exemplo, dados descritos em coordenadas geográficas).

Por conta da grande quantidade de hidrelétricas contempladas neste trabalho, o número massivo de execuções comprometeu o desenvolvimento de análises mais profundas, tal como o impacto da mudança da quantidade de neurônios na camada oculta no erro final de previsão. Além disso, algumas configurações da modelagem de rede neural foram mantidas constantes, como o número de épocas e a janela de precipitação e vazões passadas. Nessa linha, desenvolvimentos futuros devem compreender as análises mais profundas mencionadas e a flexibilidade na escolha das configurações para cada usina, de forma a permitir a otimização individualizada do desempenho de previsão.

Quanto ao processo de validação cruzada, um ajuste futuro é a divisão das amostras de treino, validação e teste de forma aleatória, e não sequencial (como foi adotado neste trabalho). Uma alternativa seria a aplicação da técnica de reamostragem aleatória com substituição, conhecida como método de *bootstrap*. Por fim, como a adoção de noções de ITL foi considerada positiva neste trabalho, estudos futuros podem considerar, além do MCC, outras funções de erro baseadas nos conceitos de entropia, como o MEE.

Bibliografia

- [1] ONS, “Histórico da Operação - Geração de Energia,” 2022. [Online]. Available: www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/geracao_energia.aspx. [Acesso em 6 Janeiro 2022].
- [2] ONS, “Previsão de Vazões Diárias ao Reservatório de Três Marias Usando a Técnica de Redes Neurais,” Rio de Janeiro, 2009.
- [3] ONS, “Histórico da Operação - Dados Hidrológicos/Vazões,” 2022. [Online]. Available: http://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/dados_hidrologicos_vazoes.aspx. [Acesso em 4 Maio 2022].
- [4] ONS, “Submódulo 4.5 - Programação Diária da Operação - Procedimental,” Rio de Janeiro, 2020.
- [5] CCEE, “Conceitos de Preço,” [Online]. Available: <https://www.ccee.org.br/web/guest/precos/conceitos-precos>. [Acesso em 2021 Dezembro 30].
- [6] J. Principe, R. Bessa, V. Miranda, A. Botterud e J. Wang, “Information Theoretic Learning applied to Wind Power Modelling,” *The 2010 International Joint Conference on Neural Networks*, pp. 1-8, 2010.
- [7] MME, “Portaria Nº 203,” Brasília, 2014.
- [8] ONS, “PMO Maio/2022 - Principais Informações da Modelagem de Curto Prazo (DECOMP),” RIO de Janeiro, 2022.
- [9] CCEE, “Regras de Comercialização - Mecanismo de Realocação de Energia,” São Paulo, 2022.
- [10] CCEE, “Regras de Comercialização - Preço de Liquidação das Diferenças,” São Paulo, 2022.
- [11] M. P. Maceira, L. A. Terry, F. S. Costa, J. M. Damázio e A. G. Melo, “Chain of optimization models for setting the energy dispatch and spot price in the brazilian system,” *Proceedings of the power system computation conference - PSCC*, vol. 2, pp. 24-28, 2002.
- [12] PSR, “SDDP - User Manual,” 2018.
- [13] FGV, “Notas dos Encontros dos Comercializadores de Energia do RIO de Janeiro,” RIO de Janeiro, 2017.
- [14] M. Pereira, N. Campodónico e R. Kelman, “Long-term hydro scheduling based on stochastic models,” em *EPSOM*, Zurique, 1998.
- [15] CCEE e ONS, “Treinamento NEWAVE-DECOMP. O SIN e os Modelos para Planejamento da Operação Energética,” São Paulo, 2015.
- [16] MegaWhat, “Energia Natural Afluyente (ENA),” [Online]. Available: <https://www.megawhat.energy/verbetes/318/energia-natural-afluyente-ena>. [Acesso em 4 Abril 2022].
- [17] CEPEL, “Manual de Referência do Modelo NEWAVE,” Rio de Janeiro, 2006.
- [18] ONS, “Submódulo 4.3 - Programação Mensal da Operação Energética - Responsabilidades,” Rio de Janeiro, 2020.
- [19] ONS, “Submódulo 4.3 - Programação Mensal da Operação Energética - Procedimental,” Rio de Janeiro, 2020.

- [20] CEPEL, “Modelo DECOMP - Determinação da Coordenação da Operação de Curto Prazo - Manual do Usuário,” Rio de Janeiro, 2013.
- [21] R. B. d. S. Santos, “Estudo da Representação da Resposta da Demanda na Programação Diária da Operação e seus Impactos na Otimização do Despacho e Custos de Operação,” Niterói, 2021.
- [22] D. L. D. D. Jardim, M. E. P. Maceira e D. M. Falcão, “Stochastic streamflow model for hydroelectric systems using clustering techniques,” *2001 IEEE Porto Power Tech Proceedings*, vol. 3, 2001.
- [23] Senado Federal, “Projeto de Lei do Senado nº 232, de 2016,” [Online]. Available: <https://www25.senado.leg.br/web/atividade/materias/-/materia/126049>. [Acesso em 5 Janeiro 2022].
- [24] MME, “Consulta Pública Nº 033/2017,” Brasília.
- [25] ONS, “Submódulo 18.2 - Relação dos Sistemas e Modelos Computacionais,” Rio de Janeiro, 2020.
- [26] CCEE, “Painel de Preços,” [Online]. Available: <https://www.ccee.org.br/web/guest/precos/painel-precos>. [Acesso em 11 Janeiro 2022].
- [27] ONS, “Submódulo 4.5 - Programação Diária da Operação - Responsabilidades,” Rio de Janeiro, 2020.
- [28] CEPEL, “Modelo DESSEM - Manual do Usuário,” Rio de Janeiro, 2019.
- [29] S. Zions, *Linear and Integer Programming*, Englewood Cliffs: Prentice Hall, 1974.
- [30] ONS, “Submódulo 4.6 - Análise e tratamento dos dados hidroenergéticos e previsão e geração de cenários de vazões - Procedimental,” Rio de Janeiro, 2020.
- [31] ONS, “Submódulo 4.6 - Análise e tratamento dos dados hidroenergéticos e previsão e geração de cenários de vazões - Responsabilidades,” Rio de Janeiro, 2020.
- [32] ONS, “Submódulo 2.4 - Critérios para estudos energéticos e hidrológicos,” Rio de Janeiro, 2020.
- [33] CCEE, “Acervo CCEE - Deck de Preços DESSEM,” [Online]. Available: <https://www.ccee.org.br/acervo-ccee?periodo=365&fc=39789%2C39792>. [Acesso em 12 Maio 2022].
- [34] ONS, “Aplicação do Modelo SMAP/ONS para Previsão de Vazões no Âmbito do SIN,” Rio de Janeiro, 2021.
- [35] ONS, “Relatório de Validação do Modelo CPINS (Cálculo e Previsão de Incrementais e Naturais a Sobradinho),” Rio de Janeiro, 2006.
- [36] S. Haykin, *Neural Networks and Learning Machines*, Upper Saddle River: Pearson Prentice Hall, 2008.
- [37] C. Ferreira, “Previsão de Vazões Naturais Diárias Afluentes ao Reservatório da UHE Tucuruí Utilizando a Técnica de Redes Neurais Artificiais,” Goiânia, 2012.
- [38] C. Moreira, “Neurônio,” *Ciência Elementar*, vol. 1, 2013.
- [39] D. G. Silva, D. G. Fantinato, J. C. Canuto, L. T. Duarte, A. O. Neves, R. Suyama, J. Montalvão e R. Attux, “An Introduction to Information Theoretic Learning, Part I: Foundations,” *Journal of Communications and Information Systems*, vol. 31, nº 1, pp. 69-80, 2016.
- [40] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford: Claredon Press, 1995.

- [41] J. C. Principe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*, Springer, 2010.
- [42] J. C. Principe, D. Xu e J. W. Fisher III, "Information-Theoretic Learning," em *Advances in Unsupervised Adaptive Filtering*, 2000.
- [43] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379-423, 623-653, 1948.
- [44] A. Rényi, "On Measures of Entropy and Information," *Fourth Berkley Symposium on Mathematical Statistics and Probability*, pp. 547-561, 1961.
- [45] E. Parzen, "On the Estimation of a Probability Density and the Mode," *Annals Math Statistics*, vol. 33, 1962.
- [46] D. Erdgomus e J. C. Principe, "An Error-Entropy Minimization Algorithm for Supervised Training of Non-Linear Adaptive Systems," *IEEE Transactions on Signal Processing*, vol. 50, 2022.
- [47] J. C. Principe, I. Santamaría e P. P. Pokharel, "Generalized Correlation Function: Definition, Properties and Application to Blind Equalization," *IEEE Transactions on Signal Processing*, vol. 54, nº 6, pp. 2187-2197, 2006.
- [48] J. C. Principe, W. Liu e P. P. Pokharel, "Correntropy: Properties and Applications in Non-Gaussian Signal Processing," *IEEE Transactions on Signal Processing*, 2007.
- [49] W. Collischonn, B. C. Da Silva e C. E. M. Tucci, "Previsão de Vazão com Modelos Hidroclimáticos," *Revista Brasileira de Recursos Hídricos*, vol. 11, nº 3, pp. 15-29, 2006.
- [50] P. B. Xavier, "Análise e Comparação de Modelos de Previsão de Vazões para o Planejamento Energético Utilizando Séries Temporais," Goiânia, 2009.
- [51] P. Sangy, "Inclusão da Precipitação na Geração de Cenários de Vazão," Rio de Janeiro, 2016.
- [52] ONS, "Previsão de Vazões Diárias ao Reservatório de Três Marias Usando a Técnica de Redes Neurais," Rio de Janeiro, 2009.
- [53] Keras, "Keras API Reference," [Online]. Available: <https://keras.io/api/>. [Acesso em 24 Maio 2022].
- [54] Z. L. Kovács, *Redes Neurais Artificiais. Fundamentos e Aplicações, um texto básico*, São Paulo: Livraria da Física, 2006.
- [55] D. M. Q. Nelson, "Uso de Redes Neurais Recorrentes para Previsão de Séries Temporais Financeiras," Belo Horizonte, 2017.
- [56] C. M. Leocádio e V. H. Ferreira, "Autonomous Neural Models Applied to Medium-Term Water Inflow Forecasting," *16th International Conference on Intelligent System Applications to Power Systems*, pp. 1-6, 2011.
- [57] R. Ballini, F. Guilhon, S. Soares e F. Gomide, "Previsão de Vazões Semanais Utilizando Redes Neurais Nebulosas," 2016.
- [58] N. D. S. Mendes, "Previsão das Vazões Afluentes Diárias por Rede Neural para a Operação do Sistema Interligado Nacional," Rio de Janeiro, 2019.
- [59] C. M. Leocádio e V. H. Ferreira, "Inferência Bayesiana no Desenvolvimento de Previsores Neurais de Vazão Diária Utilizando Informações de Precipitação," *Learning and Nonlinear Models - Journal of the Brazilian Neural Network Society*, vol. 10, nº 3, pp. 157-165, 2021.
- [60] T. L. Dias, M. Cataldi e V. H. Ferreira, "Aplicação de Técnicas de Redes Neurais e

- Modelagem Atmosférica para Elaboração de Previsões de Vazão na Bacia do Rio Grande (MG),” *Engenharia Sanitária e Ambiental*, vol. 22, nº 1, pp. 169-178, 2017.
- [61] V. H. Ferreira e A. P. A. Da Silva, “A Non-Parametric Approach for River Flow Forecasting based on Autonomous Neural Network Models,” em *Advances in Electric Power and Energy Systems: Load and Price Forecasting*, 2017, pp. 285-296.
- [62] V. H. Ferreira, G. Flores e M. Z. Fortes, “Modelo Chuva-Vazão para Previsão de Vazão Afluente Diária Utilizando Redes Neurais Artificiais,” *XIII Simpósio Brasileiro de Automação Inteligente*, 2017.
- [63] V. H. Ferreira e G. Flores, “A Rain-Streamflow Model for Prediction of Limnimetric Behavior of Reservoirs Using Artificial Neural Networks,” *Simpósio Brasileiro de Sistemas Elétricos*, 2018.
- [64] J. Duan, X. Tian, W. Ma, X. Qiu, P. Wang e L. An, “Electricity Consumption Forecasting using Support Vector Regression with the Mixture Maximum Correntropy Criterion,” *Entropy*, vol. 21, nº 7, 2019.
- [65] R. J. Bessa, V. Miranda e J. Gama, “Entropy and Correntropy Against Minimum Square Error in Offline and Online Three-Day Ahead Wind Power Forecasting,” *IEEE TRANSACTIONS ON POWER SYSTEMS*, vol. 24, nº 4, pp. 1657-1666, 2009.
- [66] J. Duan, P. Wang, W. Ma, X. Tian, S. Fang, Y. Cheng, Y. Chang e H. Liu, “Short-Term Wind Power Forecasting Using the Hybrid Model of Improved Variational Mode Decomposition and Correntropy Long Short-Term Memory Neural Network,” *Energy*, 2021.
- [67] C. A. C. Kuki, F. L. R. Torres, V. Faria, R. Queiroz, L. M. M. Lima e J. W. M. Lima, “Estratégias de Previsão de Vazão para Formação do Preço de Energia Elétrica no Curto Prazo,” *LII Simpósio Brasileiro de Pesquisa Operacional*, 2020.
- [68] CPTEC/INPE, “Informações sobre Produtos de Monitoramento Climático da Página do CPTEC/INPE,” São Paulo, 2021.
- [69] CPTEC/INPE, “Clima Monitoramento Brasil - CPTEC/INPE,” CPTEC/INPE, [Online]. Available: <http://clima1.cptec.inpe.br/monitoramentobrasil/pt>. [Acesso em Agosto 2021].
- [70] S. Ruder, “An Overview of Gradient Descent Optimization Algorithms,” Galway, 2016.
- [71] E. L. Lehmann e J. P. Romano, *Testing Statistical Hypothesis*, vol. III, Springer.