

Predição de Indicadores de Qualidade de Energia Utilizando Técnicas de Processamento de Dados e Redes Neurais

Bruno Mattedi* Klaus Fabian Côco* Patrick Marques Ciarelli*

* Departamento de Engenharia Elétrica, Universidade Federal do Espírito Santo - UFES, Avenida Fernando Ferrari 514, Vitória, Espírito Santo, Brasil

Abstract: In Brazil, one way to evaluate the service performance of electric power distribution utility is through the monitoring of electricity indices of Power Quality: SAIDI (System Average Interruption Duration Index) and SAIFI (System Average Interruption Frequency Index). They are regulated by the Brazilian National Electric Energy Agency (ANEEL), which is responsible for the application of financial penalties in case of non-compliance with regulatory limits. The estimation of these indicators provides insight into the future outlook of the company, enabling the identification of areas that have a tendency to worsen quality over time. In this work, we present a methodology for forecasting both indicators using data preprocessing, recurrent neural networks, and LSTM networks. This article uses real data from an electric utility and aims to estimate the daily SAIDI and SAIFI indicators. The results indicated that there was an improvement in the forecasting of the SAIFI, but there was no apparent benefit for the SAIDI. However, improvement in prediction allows a more appropriate allocation of maintenance teams.

Resumo: Uma das formas de avaliar o desempenho da distribuição de energia elétrica é através do monitoramento dos indicadores coletivos de qualidade de energia: Duração Equivalente de Interrupção por Unidade Consumidora (DEC) e Frequência Equivalente de Interrupção por Unidade Consumidora (FEC). Eles são regulados pela Agência Nacional de Energia Elétrica (ANEEL), o qual cabe a aplicação de penalidades financeiras em caso de não cumprimento dos limites regulatórios. A estimativa destes indicadores fornecem uma visão sobre as perspectivas futuras da empresa, permitindo a identificação de áreas que tendem a piorar a qualidade com o tempo. Neste trabalho é apresentada uma metodologia para prever ambos os indicadores usando pré-processamento de dados, redes neurais recorrentes e redes LSTM. Este artigo usa dados reais de uma concessionária de energia elétrica e tem como objetivo estimar os indicadores diários de DEC e FEC. Os resultados indicaram que houve uma melhora na predição do FEC, mas não houve benefício aparente para o DEC. Todavia, a melhora na predição permite uma alocação mais adequada das equipes de manutenção.

Keywords: Electrical Energy, RNN, LSTM, Forecasting, Distribution of Energy.

Palavras-chaves: Energia Elétrica, RNN, LSTM, Predição, Distribuição de Energia

1. INTRODUÇÃO

Nos últimos anos, o setor elétrico brasileiro tem sofrido um processo de grandes mudanças em sua estrutura. A primeira grande mudança aconteceu em 1996, com a criação da Agência Nacional de Energia Elétrica (ANEEL), que estabeleceu um conjunto de exigências, obrigações, procedimentos, indicadores de natureza técnica e comercial, além de penalidades, com o objetivo de assegurar a qualidade dos serviços prestados por empresas distribuidoras de energia elétrica aos seus consumidores (Dias, 2002), de maneira semelhante ao método de penalidade e recompensa apresentado em Simab e Haghifam (2012).

Para medir o desempenho e a confiabilidade de sistemas de distribuição de energia a partir do aspecto da qualidade do produto, existem indicadores de qualidade estabelecidos pela ANEEL. Dentre eles, destacam-se a Duração Equiva-

lente de Interrupção por Unidade Consumidora (DEC) e Frequência Equivalente de Interrupção por Unidade Consumidora (FEC), estabelecidos no Módulo 8 dos Procedimentos de Distribuição – PRODIST (ANEEL, 2017b).

Em caso de violação das metas estabelecidas no período de apuração dos indicadores, as concessionárias sofrem penalidades (Sperandio et al., 2004) que vão desde a elaboração de um plano de melhorias, na qual a concessionária deve elaborar um plano para atingir as metas de seus conjuntos elétricos e apresentá-lo ao órgão fiscalizador, bem como a realização de compensações aos consumidores na forma de descontos. Por outro lado, o bom desempenho em tais indicadores é recompensado pela ANEEL através de ajuste tarifário anual como incentivo à busca pela melhoria contínua da qualidade do serviço. Conforme dados da ANEEL (2017a), no Brasil as concessionárias realizaram uma quantidade de 83,48 milhões de compensações em

2018, 80,81 milhões em 2019 e 79,49 milhões em 2020, que resultaram num montante de R\$ 483,70 milhões, R\$ 633,89 milhões e R\$ 634,64 milhões, respectivamente. As multas e as compensações onerosas para as distribuidoras afetam não só o Brasil, mas também vários outros países. Por exemplo, estima-se que os custos da falta de energia nos Estados Unidos foram de aproximadamente US\$ 79 bilhões para a economia em 2006 (LaCommare e Eto, 2006).

O conhecimento a priori dos valores dos indicadores de continuidade DEC e FEC pode auxiliar na definição de prioridades para distribuir de forma mais eficaz os investimentos e recursos para a redução de problemas no sistema elétrico, trazendo como consequência direta a melhoria dos índices de confiabilidade (Tronchoni et al., 2008). A atual forma de estimativa dos indicadores utilizada pela concessionária de energia é a média histórica dos meses, obtendo assim uma estimativa imprecisa dos valores. Nesse aspecto, esse trabalho busca utilizar a análise e o processamento dos dados históricos, além do uso de redes neurais, para desenvolver uma forma que permita prever, com o menor erro possível, esses indicadores de continuidade.

Sendo assim, esse trabalho propõe o uso de redes neurais com técnicas de processamento de sinais para estimar os indicadores DEC e FEC. Nos experimentos realizados foi utilizado um tipo de rede rasa e uma rede de aprendizado profundo com camada *Long Short-Term Memory (LSTM)*.

Para o tratamento dos dados foi utilizada uma metodologia que decompõe o sinal em subséries por meio de transformada Wavelet (Percival e Walden, 2006) e redução do *offset* do sinal por meio de oscilação em torno do valor zero. Os resultados obtidos sobre uma base de dados reais mostraram que a abordagem proposta de processamento dos dados apresentaram resultados promissores frente ao treinamento sem esta etapa.

Este trabalho está organizado como segue. Na Seção 2 são descritas as redes neurais utilizadas nos experimentos. Na Seção 3 é descrito o método proposto neste trabalho. Os experimentos, resultados e discussões são apresentados na Seção 4. Por fim, é realizada a conclusão na Seção 5.

2. REDES NEURAIS ARTIFICIAIS

Tendo em vista a capacidade do cérebro humano de aprender e generalizar o conhecimento, foram criados os modelos de Redes Neurais Artificiais (*Artificial Neural Networks - ANN*) com a intenção de reproduzir parte de suas habilidades. Estas redes artificiais são compostas por camadas que, por sua vez, têm a formação a partir de simples unidades conhecidas como neurônios (Haykin et al., 2009). Os neurônios são a menor estrutura de uma ANN e possuem como principais componentes os sinais de entrada, pesos sinápticos, viés e uma função de ativação. A aprendizagem é realizada ajustando os pesos sinápticos a fim de minimizar o erro, geralmente usando um algoritmo de descida de gradiente ou suas variantes.

Um tipo comum de ANN é a *feedforward*, amplamente utilizada em tarefas de sistemas de energia, como controle e previsão, conforme Ahn et al. (2017) e Reynolds et al. (2019). Outro tipo de ANN, utilizado neste trabalho, é a Rede Neural Recorrente (*Recurrent Neural Network - RNN*), que possui uma realimentação, ligando as saídas às

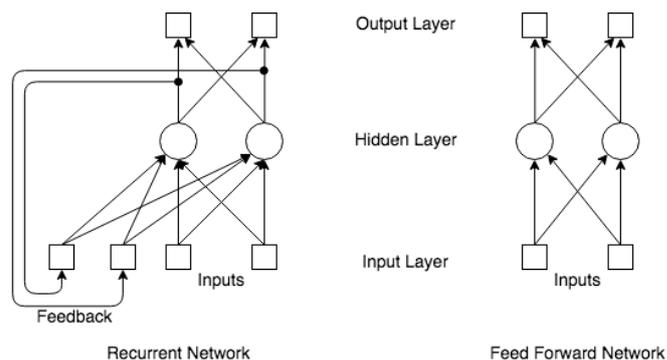


Figura 1. Arquiteturas básicas de rede recorrente e *feedforward*

entradas de todos os neurônios em uma camada (Chandra e Zhang, 2012), formando assim uma forma de memória em relação aos padrões de entrada (Schmidhuber, 2015). A Figura 1 mostra as arquiteturas de ambas as ANNs.

As redes recorrentes têm apresentado bons desempenhos para previsão de séries temporais caóticas (Bone et al., 2000). A Rede Neural Não Linear Autoregressiva (*Nonlinear Autoregressive Neural Network - NAR*) é uma rede rasa deste tipo, que usa somente como entrada dados passados da variável que se quer prever.

Nas redes com retropropagação no tempo, como a mencionada anteriormente, o erro é retroalimentado e tende a explodir ou desaparecer. Isso ocorre pois, conforme demonstrado por Hochreiter e Schmidhuber (1997), a forma como essas redes comumente aprendem faz com que o erro seja escalado de forma a crescer ou decrescer exponencialmente. No primeiro caso, gera uma oscilação nos pesos, fazendo com que a rede tenha um aprendizado instável. Já para o decréscimo, o erro rapidamente desaparece, assim a rede não aprende em um intervalo de tempo aceitável.

Dessa forma, por apresentar dificuldades em utilizar informações passadas no instante necessário, em (Hochreiter e Schmidhuber, 1997) foi apresentada a proposta da camada *Long Short-Term Memory (LSTM)* de modo a proporcionar a retropropagação do erro sem os problemas anteriores. Neste modelo, apresentado na Figura 2, foi introduzido o conceito de célula de memória, permitindo às unidades conterem informações fora do fluxo. Assim, dados mais antigos podem ser utilizados, mesmo vários passos depois, ou esquecidos, caso não sejam mais relevantes.

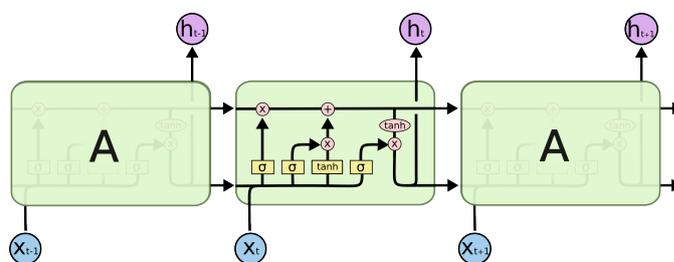


Figura 2. Camada LSTM. Fonte: Cui et al. (2018)

Existem muitas aplicações que utilizam redes profundas em sistemas de energia (Wang et al., 2019; Fan et al., 2019). Afinal, de acordo com a profundidade da rede, informações mais complexas podem ser aferidas, pois as

instruções mais a frente podem utilizar os dados extraídos pelas camadas anteriores (Goodfellow et al., 2016).

Nesse caso, o uso de redes profundas torna-se uma boa motivação por não apresentar esse problema na retropropagação do erro. Assim, foi usada a rede neural LSTM de modo a proporcionar a passagem do erro sem os problemas anteriores (Gers, 2001). As redes LSTM vem sendo utilizadas para aplicações envolvendo séries temporais com longas dependências temporais (Ma et al., 2015). Desta forma, neste trabalho, busca-se avaliar o desempenho das redes NAR e LSTM na tarefa de predição de DEC e FEC.

3. MÉTODO PROPOSTO

3.1 Pré-processamento

Para o desenvolvimento deste estudo, foram utilizados os dados diários de DEC e FEC oriundos de um Conjunto Elétrico que compõe a área da concessionária estudada, referentes ao período de 01/01/2010 a 29/01/2016, formando assim duas séries temporais (DEC e FEC). A Figura 3 mostra um exemplo da disposição dos dados de DEC e FEC para o primeiro semestre 2014. Pode-se observar que existem vários picos de valores, inclusive alguns bem proeminentes. De maneira complementar, buscando uma melhor análise dos dados, a Tabela 1 apresenta uma breve estatística da base de dados.

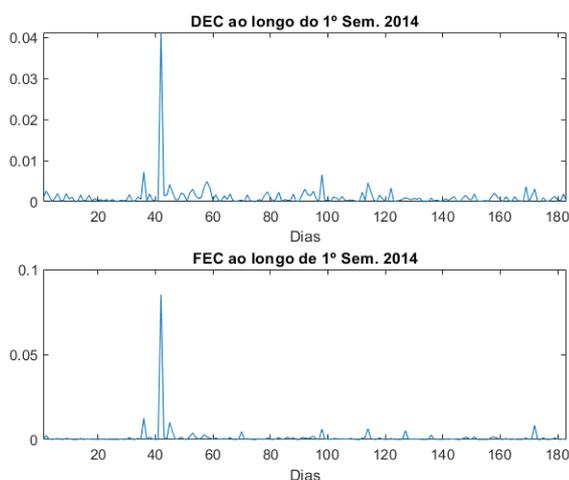


Figura 3. Exemplo de série do DEC para o ano de 2014

Tabela 1. Análise estatística dos indicadores coletivos de qualidade da rede elétrica

Indicador	Qnt	$\bar{x}(10^3)$	Min(10^3)	Max (10^3)	$\sigma(10^3)$
DEC	2220	0,4547	1,0483e-03	41,1690	2,1
FEC	2220	0,2144	7,9328e-04	85,1780	2,7

Cada uma das séries recebeu tratamento de *outliers*, onde foram mantidos inalterados os dados no intervalo de valor de $[0, \sigma]$, sendo σ o desvio padrão calculado a partir de toda a série temporal até $t-1$, sendo t o instante a ser estimado. Os valores acima desse intervalo foram retirados da série e preenchidos com um valor por meio do algoritmo *Piecewise Cubic Hermite Data Interpolation* (PCHIP) (Fritsch e Carlson, 1980; Fritsch, 1985).

Em seguida, os dados foram interpolados, aumentando o número de amostras entre cada ponto com 3 amostras, a fim de suavizar as variações abruptas das séries. Para isto, foi usado o spline cúbico por partes de forma a preservar a forma após a interpolação, conforme ilustrado na Figura 4.

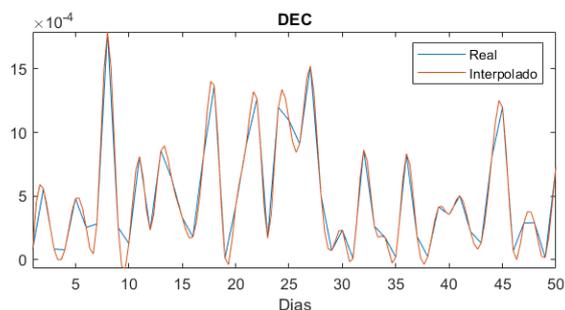


Figura 4. DEC antes e depois da interpolação

Como as 2 séries possuem apenas valores positivos, foram geradas 2 novas séries a partir da raiz quadrada dos valores das séries tratadas para reduzir a amplitude dos dados. Os valores das novas séries foram alternados entre positivo e negativo ao multiplicar por 1 e -1 , respectivamente, a cada $n = 5$ pontos buscando promover uma excursão oscilatória das amostras com objetivo de melhorar a aprendizagem neural com a despolarização dos dados. A Figura 5 ilustra antes e depois de realizar tal procedimento.

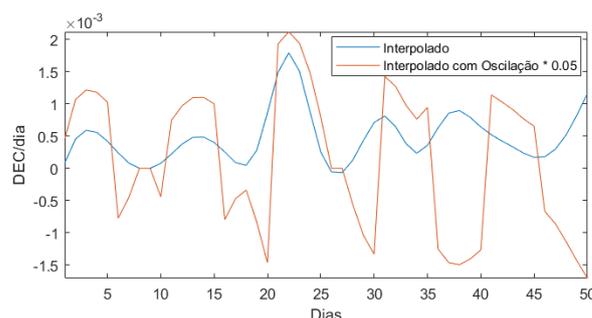


Figura 5. Comparação do DEC antes (azul) e depois (laranja) da raiz seguida da oscilação. Ambos os sinais estão interpolados

Cada uma das 2 séries foi decomposta em 3 níveis de detalhe (d1, d2 e d3) e um de aproximação (a3) utilizando uma Transformada Wavelet chamada *Maximal Overlapped Discrete Wavelet Transform* (MODWT) (Percival e Walden, 2006), usando a base db10. Após isto, cada série foi recomposta em 3 subséries com intervalo de frequência baixa (d3 e a3), média (d2) e alta (d1), resultando, no total, em 3 subséries para análise do DEC e 3 para o FEC.

A decomposição usando a MODWT tem a finalidade de decompor as séries originais em subséries aditivas e somáveis com menor complexidade. Aditivas, pois não apresentam efeito de borda como na Transformada Wavelet Discreta, sendo possível efetuar a adição de uma amostra ao final de cada subsérie antes da transformada sem a necessidade de modificar as demais amostras que compõem as respectivas subséries. Somáveis, pois a série original é recomposta pela soma das subséries da decomposição MODWT.

Figura 6 mostra um exemplo do resultado de operações de pré-processamento gerando 3 subséries de uma série de

DEC, onde também é possível observar que a inserção da oscilação vem a gerar uma espécie de onda portadora, fazendo com que as frequências tenham um comportamento semelhante. A Figura 7 mostra as etapas do processo de pré-processamento dos dados.

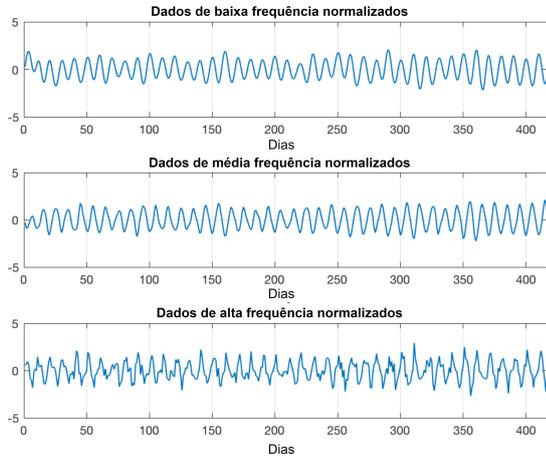


Figura 6. Dados após o pré-processamento e decomposição

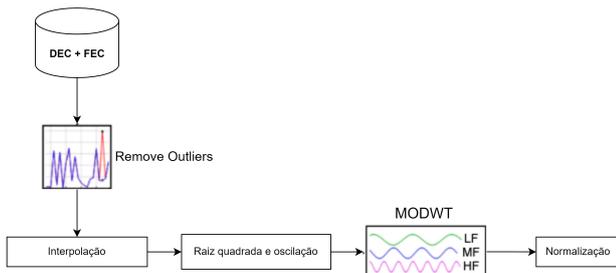


Figura 7. Diagrama do tratamento de dados realizado

3.2 Redes Neurais Artificiais

Foram utilizadas duas configurações de redes, NAR e LSTM, para fazer a predição de cada uma das 3 subséries (baixa, média e alta frequência). Em função da interpolação, cada rede irá estimar um horizonte de 3 valores, sendo dois deles interpolados e o terceiro o valor real. Dessa forma, a quantidade de valores estimados em um instante é igual ao horizonte de predição (HP), no caso $HP = 1$ dia, multiplicado pela interpolação escolhida ($numero_{elementos} = HP * Interpol$ e $\hat{y} = \hat{y}_{last}$). Assim, para cada grupo de saída é considerado apenas o valor do último para comparação.

Como a MODWT é somável, a predição de cada subsérie pode ser somada para formar a predição da série interpolada normalizada diária. Para desfazer a normalização de cada subsérie, o valor predito é obtido ao somar a média \bar{x} e multiplicar pelo desvio padrão σ . Os valores de \bar{x} e σ foram recalculados para cada dia dos dados de teste, acrescentando ao vetor de treino os de teste até $d - 1$. Em seguida, as subséries de frequência são somadas e seus valores são elevados ao quadrado para desfazer a raiz com oscilação, conforme Equação (1).

$$\hat{y} = \left(\sum_{n=baixa}^{alta} prev(n) * \sigma(n) + \mu(n) \right)^2 \quad (1)$$

Na etapa de treinamento, para as redes NAR, foi utilizada a faixa de parâmetros mostrada na Tabela 2, onde o parâmetro *delay* define a quantidade de dados passados que a rede utilizará para realizar a predição, ou seja, para um valor de x dias, os dados para treinamento corresponderão às entradas $(y_n, y_{n-1}, \dots, y_{n-x+1})$, obtendo a saída y_{n+1} .

Além disso, foi configurada com uma camada de entrada, uma oculta e uma de saída, com o algoritmo de treinamento Levenberg-Marquardt Lourakis et al. (2005), função tangente hiperbólica sigmoidal (tansig) como função de ativação e a função *Mean Squared Error* (MSE) para avaliar o desempenho da rede.

Tabela 2. Intervalo de variação dos parâmetros

	Neurons		Delay	
	Min	Max	Min	Max
NAR	5	80	3	30

Para a predição de indicadores por meio de técnicas de aprendizado profundo, foi utilizada uma arquitetura de seis camadas, mostrada na Figura 8, sendo quatro ocultas, uma de entrada e uma de saída.

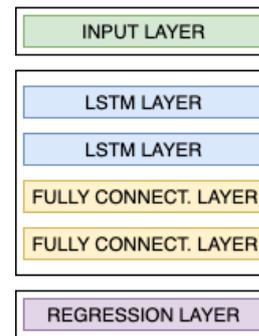


Figura 8. Arquitetura da Rede Profunda com duas camadas LSTM em Alto Nível

Em relação às camadas ocultas, as duas primeiras são do tipo LSTM, com a quantidade de células na primeira camada variando na faixa de 400 a 600, onde a segunda camada possui 2/3 do número de células da primeira. A primeira camada densa possui o mesmo número de neurônios da camada anterior de LSTM, seguida pela segunda camada densa que possui somente um neurônio, pois o horizonte da predição é de um dia. Assim, para cada dia a ser estimado são realizadas n predições, sendo n o número de interpolações.

4. EXPERIMENTO E RESULTADOS

Esta seção compara os resultados obtidos nas propostas para predição dos indicadores de continuidade DEC e FEC por meio de uma rede NAR e uma rede profunda com camadas LSTM para o mês de janeiro nos dias 01/01/2016 à 29/01/2016, cuja ocorrência dos erros absolutos em porcentual de predição estão na Tabela 3. Estes erros são calculados pela Equação (2), sendo que *RealSerie* corresponde aos valores reais do intervalo estimado, no

caso, o mês de janeiro. Além disso, a Tabela 3 destaca em negrito a rede que obteve o melhor desempenho para cada indicador em cada dia.

$$e_{abs} = \frac{Real - Prev}{\max(RealSerie) - \min(RealSerie)} \times 100 \quad (2)$$

Tabela 3. Erro percentual absoluto dos indicadores de continuidade para janeiro de 2016 e média dos módulos dos erros

Dia	DEC		FEC	
	NAR	LSTM	NAR	LSTM
1	29,17	30,8	59,14	69,85
2	34,81	47,61	20,66	19
3	35,19	49,83	-7,79	-3,99
4	18,94	38,45	-5,06	9,98
5	28,13	-7,12	-8,12	11,86
6	-7,28	-16,91	-16,73	-35,75
7	-7,79	-1,14	-9,2	-36
8	7,58	-3,89	-8,44	1,19
9	42,55	45,95	-1,47	-10,79
10	-47,61	-24,61	-3,26	-3,32
11	-29,57	-31,49	-0,05	-6,96
12	36,82	33,06	16,34	15,77
13	7,23	-16,08	-0,61	3,83
14	-33,68	-3,41	-11,55	-10,65
15	52,86	64,77	29,85	23,44
16	-37,35	-32,78	-22,6	-9,02
17	-34,12	2,03	-10,74	31,71
18	-38,89	-31,68	52,75	48,15
19	58,06	34,82	30,85	75,32
20	5,99	23,93	16,81	6,38
21	19,13	37,96	12,05	18,48
22	-14,88	8,14	1,8	2,7
23	-33,32	0,33	6,8	42,2
24	-17,21	-4,11	33,27	10,48
25	22,58	19,15	-9,57	-6,94
26	3,91	5,84	8,43	-22,07
27	-6,86	-25,09	7,08	-11,83
28	1,62	-3,76	-16,55	-1,74
29	-34,11	-15,31	-2,02	-42,35
Média	25,77	22,76	14,81	20,41

A Tabela 3 permite observar que ambas as redes apresentaram um desempenho semelhante para o DEC e para o FEC, sendo a análise complementada pela Figura 9 que, através do gráfico de erro acumulado, indica que os desempenhos das redes foram semelhantes para predição do DEC, mas para o FEC a rede NAR obteve um menor erro acumulado. Isto pode também ser observado nos valores de RMSE da Tabela 4.

Tabela 4. RMSE das redes propostas para os indicadores DEC e FEC, multiplicados por 10^3

RMSE	DEC		FEC	
	NAR	LSTM	NAR	LSTM
	0.5148	0.4913	0.2584	0.3519

A Figura 10 apresenta a comparação de desempenho entre o pré-processamento dos dados com apenas normalização e tratamento de outliers e o método proposto.

Já a Figura 11 apresenta a média histórica, forma de estimativa empregado pela concessionária, em comparação com os resultados obtidos através das redes neurais discursadas ao longo deste artigo.

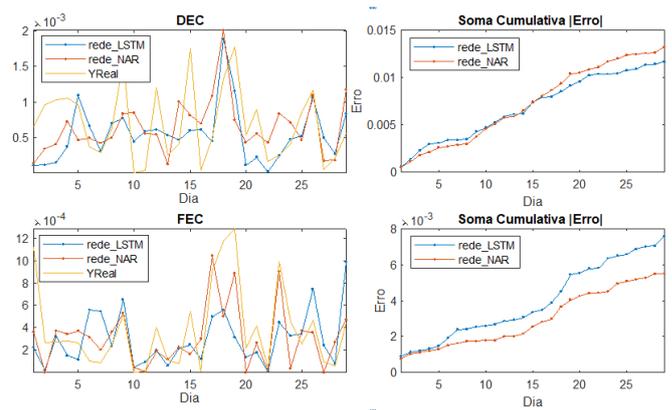


Figura 9. Comparação entre valores preditos e reais

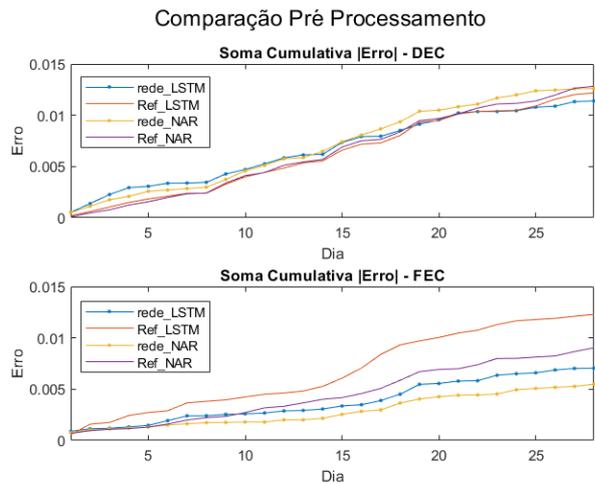


Figura 10. Comparação entre o desempenho das redes com (azul e amarelo) e sem (laranja e roxo) o pré-processamento proposto

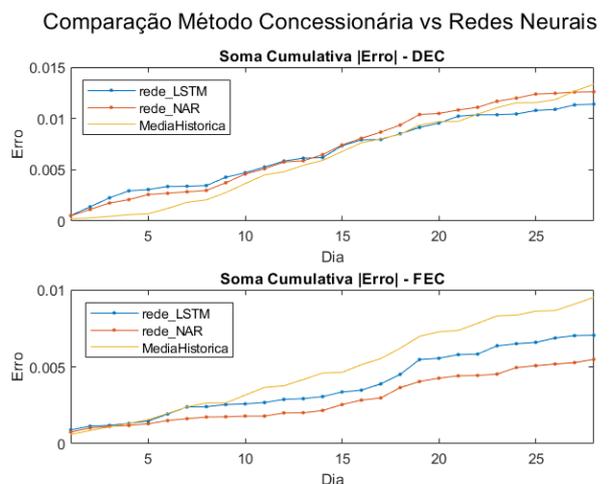


Figura 11. Comparação entre o método atual de predição e o proposto

Assim, por meio das Figuras 10 e 11, nota-se que houve um desempenho superior das redes na predição do índice FEC em relação ao método utilizado pela concessionária, e

também superior se comparado com a aplicação das redes neurais sem a utilização do pré-processamento proposto.

5. CONCLUSÃO

O maior objetivo deste trabalho foi a aplicação de técnicas de processamento de sinais e de redes neurais para a predição dos indicadores de continuidade coletivos de uma distribuidora de energia elétrica e, a partir dos resultados obtidos, melhorar a alocação de recursos da empresa.

Foi realizada a predição do DEC e FEC por meio de uma rede NAR e de uma rede profunda com camadas LSTM, e observou-se que para o DEC não houve ganho significativo com os métodos aplicados. Todavia, para o indicador FEC foi possível observar que há um ganho de desempenho se comparado tanto com a aplicação de redes neurais sem a utilização dos métodos decorridos de pré-processamento como frente ao utilizado pela concessionária atualmente.

Vale ressaltar que as equipes que trabalham para restaurar a energia tem sua demanda em função da frequência de interrupções, FEC, e prioridade de atendimento de acordo com o DEC gerado por cada uma delas. Assim, através da melhora da predição do FEC é possível melhorar a estratégia de alocação das equipes, trazendo menos penalidades para a empresa.

Para trabalhos futuros, é sugerido um modelo que implemente uma rede mais complexa ou variações no processamento dos dados, de modo a permitir melhorar o desempenho de predição dos indicadores DEC e FEC.

AGRADECIMENTOS

Agradecemos o apoio da NVIDIA Corporation com a doação da GPU Titan V usada para esta pesquisa.

REFERÊNCIAS

- Ahn, J., Cho, S., e Chung, D.H. (2017). Analysis of energy and control efficiencies of fuzzy logic and artificial neural network technologies in the heating energy supply system responding to the changes of user demands. *Applied Energy*, 190, 222 – 231.
- ANEEL (2017a). Compensação pela transgressão dos limites de continuidade. <http://www.ANEEL.gov.br/indicadores-de-compensacao-de-continuidade>. Accessed: 2018-04-19.
- ANEEL (2017b). Módulo 8. <http://www.aneel.gov.br/modulo-8>. Accessed: 2018-04-19.
- Bone, R., Crucianu, M., e de Beauville, J.P.A. (2000). Two constructive algorithms for improved time series processing with recurrent neural networks. In *Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No. 00TH8501)*, volume 1, 55–64. IEEE.
- Chandra, R. e Zhang, M. (2012). Cooperative coevolution of elman recurrent neural networks for chaotic time series prediction. *Neurocomputing*, 86, 116–123.
- Cui, Z., Ke, R., e Wang, Y. (2018). Deep bidirectional and unidirectional lstm recurrent neural network for network-wide traffic speed prediction. *arXiv preprint arXiv:1801.02143*.
- Dias, E.B. (2002). *Avaliação de indicadores de continuidade e seu impacto no planejamento de sistemas de distribuição*. Ph.D. thesis, Universidade de São Paulo.
- Fan, C., Wang, J., Gang, W., e Li, S. (2019). Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. *Applied Energy*, 236, 700 – 710.
- Fritsch, F. (1985). Pchip, piecewise cubic hermite data interpolation.
- Fritsch, F.N. e Carlson, R.E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2), 238–246.
- Gers, F. (2001). *Long short-term memory in recurrent neural networks*. Ph.D. thesis, Verlag nicht ermittelbar.
- Goodfellow, I., Bengio, Y., e Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- Haykin, S.S., Haykin, S.S., Haykin, S.S., e Haykin, S.S. (2009). *Neural networks and learning machines*, volume 3. Pearson Upper Saddle River.
- Hochreiter, S. e Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- LaCommare, K.H. e Eto, J.H. (2006). Cost of power interruptions to electricity consumers in the united states (us). *Energy*, 31(12), 1845–1855.
- Lourakis, M.I. et al. (2005). A brief description of the levenberg-marquardt algorithm implemented by levmar. *Foundation of Research and Technology*, 4(1), 1–6.
- Ma, X., Tao, Z., Wang, Y., Yu, H., e Wang, Y. (2015). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, 54, 187–197.
- Percival, D.B. e Walden, A.T. (2006). *Wavelet methods for time series analysis*, volume 4. Cambridge university press.
- Reynolds, J., Ahmad, M.W., Rezgui, Y., e Hippolyte, J.L. (2019). Operational supply and demand optimisation of a multi-vector district energy system using artificial neural networks and a genetic algorithm. *Applied Energy*, 235, 699 – 713.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85–117.
- Simab, M. e Haghifam, M.R. (2012). Quality performance based regulation through designing reward and penalty scheme for electric distribution companies. *International Journal of Electrical Power & Energy Systems*, 43(1), 539–545.
- Sperandio, M. et al. (2004). Classificação de conjuntos consumidores de energia elétrica via mapas auto-organizáveis e estatística multivariada.
- Tronconi, A.B. et al. (2008). Identificação de causas de desligamentos não programados em redes de distribuição.
- Wang, W., Hong, T., Xu, X., Chen, J., Liu, Z., e Xu, N. (2019). Forecasting district-scale energy dynamics through integrating building network and long short-term memory learning algorithm. *Applied Energy*, 248, 217 – 230.