



## GRUPO DE ESTUDO DE PLANEJAMENTO DE SISTEMAS ELÉTRICOS - GPL

### MODELO DE PREVISÃO DE CONSUMO DE ENERGIA ELÉTRICA EM CENÁRIO PANDÊMICO COM USO DE FERRAMENTAS *MACHINE LEARNING*

LEANDRO VARGAS BARBOSA (1); GLADIS BORDIN (2)  
COMPANHIA ESTADUAL DE GERACAO E TRANSMISSAO DE ENERGIA - CEEE-GT (1);  
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL – UFRGS, DEPARTAMENTO DE ENGENHARIA ELÉTRICA (2)

#### RESUMO

O trabalho propõe um novo modelo estocástico de projeção de consumo de energia elétrica, utilizando técnicas de otimização e *machine learning* aplicadas em uma rede neural recorrente do tipo *Long Short Term Memory*. Ressalta o efeito das mortes por COVID-19 no consumo de energia, e contempla o efeito aleatório e de atenuação de um transitório social, o que costumeiramente é simplificado em modelos de projeção de consumo. Os resultados da aplicação do modelo, através dos estudos de caso no Rio Grande do Sul e Brasil, mostram que as técnicas utilizadas são adequadas à representação do problema sob análise.

PALAVRAS-CHAVE: Previsão de Consumo, Energia Elétrica, COVID-19, *Machine Learning*.

#### 1.0 INTRODUÇÃO

Um problema relevante e complexo relacionado ao planejamento da gestão empresarial de uma concessionária de energia elétrica diz respeito à projeção do consumo de energia elétrica. Sabe-se que as projeções, em geral, não constituem um fim, mas um instrumento que fornece informações para uma consequente tomada de decisão com objetivos específicos. Assim, a projeção de consumo visa atender às necessidades dos consumidores com um nível de qualidade satisfatório.

A base de todo o processo de planejamento da expansão e operação de um sistema elétrico está ancorada em suas projeções de demanda por consumo de energia elétrica, tanto no que se refere aos aspectos técnicos quanto aos aspectos financeiros, como: o planejamento da geração, a transmissão e distribuição, o planejamento da operação e o planejamento financeiro.

A demanda de energia elétrica se reflete diretamente na atividade econômica, e, consequentemente, no desenvolvimento de uma região ou de um país. Assim sendo, a previsão do consumo de energia é uma informação basilar na decisão de investimento de qualquer ramo empresarial.

A projeção usual de consumo de energia elétrica baseia-se principalmente em técnicas econométricas, através de correlações entre a série histórica do consumo de energia elétrica e indicadores sócio-econômicos. No entanto, com o evento mundial da pandemia do COVID-19 e seus efeitos sobre o consumo, uma vez que dados medidos de demanda atendida após o início da pandemia mostram a forte correlação entre o evento e a evolução do consumo, se faz necessário repensar os modelos de projeção de consumo em uso e adequá-los a esse novo contexto, buscando tornar as projeções mais realísticas para tomada de decisão de investimento.

Neste contexto, o presente trabalho propõe um modelo estocástico de projeção de consumo de energia elétrica, com uso de ferramentas e técnicas de *machine learning* [1] de código aberto implementadas em linguagem de programação python para modelar o efeito do evento aleatório COVID-19 sob as projeções de consumo de médio e longo prazo. O modelo proposto faz uso de indicadores econômicos tradicionais, entre outros, como IPCA, indicadores de mercado como IBOV e indicadores sociais e de saúde como registro de óbitos (TABNET SUS), contrastando e correlacionando os indicadores citados com os dados históricos de consumo disponibilizados pelo anuário Empresa de Pesquisa Energética - EPE [2]. Os cenários de consumo são projetados através de uma rede neural recorrente do tipo LSTM – *Long Short Term Memory*.

O presente trabalho se diferencia de estudos anteriores tanto na técnica quanto no enfoque de modelagem. Neste sentido, utiliza uma abordagem de vanguarda no emprego de inteligência artificial e explora variáveis não tradicionais, ou seja, o efeito da pandemia de COVID-19 no consumo de energia elétrica no Brasil. Ainda, o modelo desenvolvido contempla o que costumeiramente é simplificado em modelos de projeção de consumo de energia elétrica: o efeito aleatório e de atenuação de um transitório social.

## 2.0 DESENVOLVIMENTO DO MODELO PROPOSTO

O modelo foi realizado privilegiando-se, ao mesmo tempo, restringindo-se à utilização de dados de entrada de fácil acesso e sem restrição de uso para fins acadêmicos. Igualmente, optou-se por dados de institutos governamentais, de domínio público. De forma similar foram utilizadas ferramentas de código aberto ou sem restrições de uso, em especial a linguagem de programação Python [2] e bibliotecas associadas, principalmente: Tensorflow [3], Statsmodels [4] e Pandas [5].

Os dados elétricos utilizados foram coletados do anuário de consumo de energia elétrica fornecido pela EPE e para as análises mais atuais foram utilizados os dados preliminares disponibilizados através da resenha mensal EPE [6]. Os dados climatológicos utilizados foram coletados no portal do Instituto Nacional de Pesquisas Espaciais - INPE [7], das estações suas meteorológicas automáticas. Os dados econômicos foram coletados por meio do Índice Nacional Preços ao Consumidor Amplo (IPCA) e Índice de Atividade Econômica do Banco Central (IBC-Br) [8]. Da Bolsa de Valores B3 [9] utilizou-se o índice Bovespa (IBOV). Ainda, para estimar a intensidade da pandemia foram utilizados os dados do total de mortes registradas no Sistema Único de Saúde - SUS chamado DataSUS/Tabnet [10].

A partir da obtenção dos dados foram elaborados os estudos de caso. Inicialmente, o estudo de caso abrange o Estado do Rio Grande do Sul - RS, testando o modelo desenvolvido numa abordagem simplificada e avaliando a aderência da utilização dos dados climatológicos do INPE e a série histórica de consumo disponibilizada pela EPE. A primeira análise utiliza a série histórica de 2013 a 2017 para prever a carga de 2018, buscando obter um cenário pré-pandemia em 2018. O modelo ajustado foi reaplicado na série 2015 a 2019 para prever 2020, sendo que a janela de cinco (05) anos foi definida por ser o formato fornecido pelo anuário da EPE. Posteriormente, o estudo de caso abrange o Brasil (Seções 2.5 e 2.6).

### 2.1 Dados Utilizados

Visualizando os dados de 2018 do RS pelo primeiro nível de agrupamento EPE Setor N1, verifica-se na Figura 1 que os grupos mais significativos são o Industrial, o Residencial, o Comercial e o Rural.

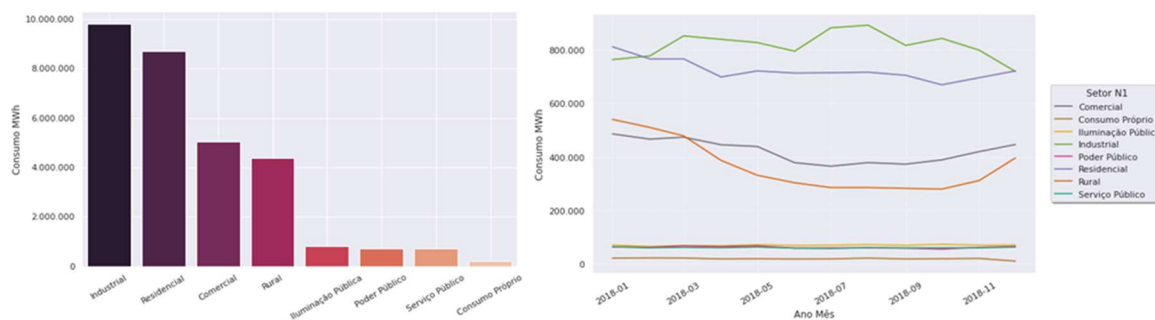


Figura 1 - Consumo por Setor N1 RS em 2018 (à esquerda) e Perfil de Consumo Mensal por Setor N1 RS em 2018 (à direita)

Ao analisar a distribuição deste consumo ao longo dos meses, pode-se notar quais são as cargas que tem uma variação ao longo do ano. A Figura 2 evidencia que o setor Rural tem a maior variação ao longo do ano. Para validar esta informação e verificar que esta oscilação se repete ao longo da série de 2013 à 2018, foi realizada uma análise de dispersão através do desvio padrão.

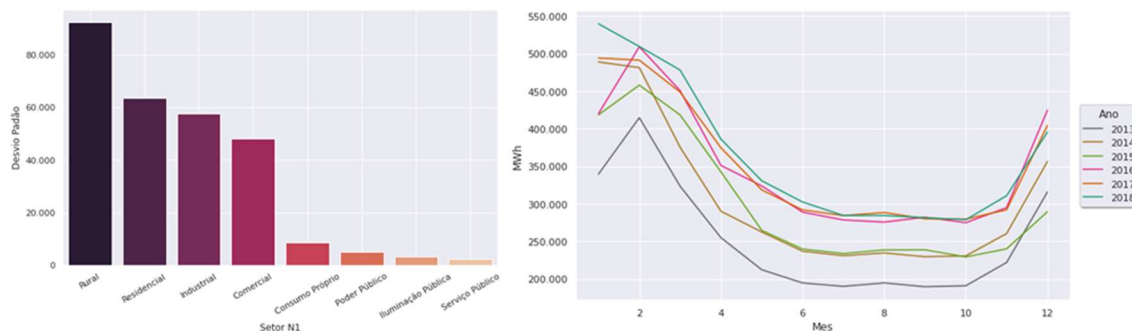


Figura 2 – Desvio Padrão de consumo por Setor N1 RS de 2013 à 2018 (à esquerda) e Perfil de consumo de 2013 a 2018, Detalhe: Setor N1 = Rural e UF = RS (à direita)

Assim, fazendo um novo recorte e ponderando que o mais desafiador é a previsão de uma carga variável ao longo do tempo, do que uma carga constante ou que varie muito suavemente. Considerando somente os dados dos consumidores rurais do RS, visando aprofundar o estudo nesta classe de consumo. Ao decompor o sinal de consumo ao longo do tempo, utilizando o método multiplicativo do Statsmodels [4], na Figura 3 (à esquerda), evidencia-se a sazonalidade da carga que se trata de um sinal periódico no tempo e com uma tendência de alta. A fim de comprovar a sazonalidade e verificar se é possível associar a mesma a sucessão de meses e estações do ano foi aplicado o

teste de autocorrelação da série de consumo Rural no RS sob ela própria defasando-a mês a mês, obtém-se o gráfico (à direita) onde percebe-se uma correlação negativa após uma defasagem de 4 meses, o que enseja uma possível dependência das estações do ano, na qual a característica senoidal reforça essa hipótese e a envoltória decrescente reafirma a tendência de incremento no consumo com o passar dos meses.

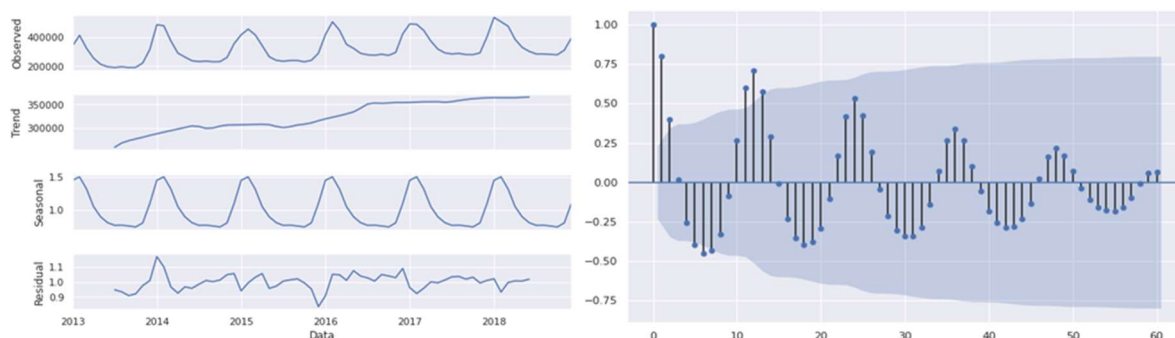


Figura 3 – Decomposição multiplicativa (à esquerda) e Autocorrelação (à direita),  
Detalhe: Setor N1 = Rural e UF = RS

Ainda foi realizado uma breve análise sobre os outros setores N1 onde se verifica, seguindo os mesmos critérios, uma periodicidade bem definida que se pode correlacionar a sucessão das estações do ano, nos setores N1: o Residencial Comercial, o Rural e o Poder Público. E um comportamento não diretamente correlacionado a sucessão dos meses do ano, nos setores N1: Industrial, Iluminação Pública, Serviço Público, Consumo Próprio, devido à característica sazonal verificada nos dados elétricos e da atividade rural depender essencialmente do clima e das estações, buscou-se os dados provenientes das estações automáticas do INMET. Nestas estações estão disponíveis os seguintes dados agrupados por mês, e sumarizados pelo valor médio: Número de dias com precipitação; Precipitação total, mensal [mm]; Pressão Atmosférica, média mensal [mb]; Temperatura média, mensal [°C]; Vento, velocidade máxima mensal [m/s] Vento, velocidade média mensal [m/s], e estas variáveis foram correlacionadas ao consumo onde foram verificadas que as correlações mais significativas foram a sucessão dos anos e a temperatura. Para verificar essa dependência entre as séries climáticas e o consumo de energia foi analisada a correlação pelo método de Pearson onde os resultados estão graficados na Figura 4 (à esquerda) e a variável com maior correlação (Temperatura) esta detalhada no seu perfil (à direita).

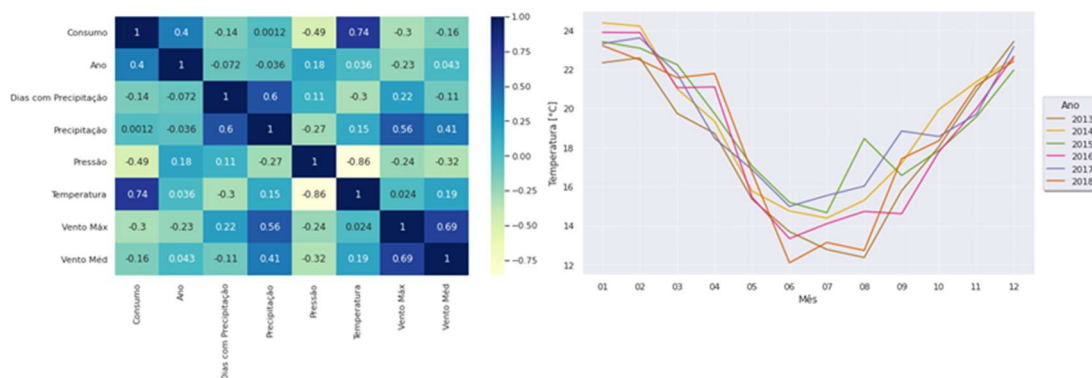


Figura 4. Correlação entre as variáveis (à esquerda) e Perfil Médio de Temperatura no RS (à direita)

Considerando a hipótese de que com a evento da pandemia de COVID-19 houve impactos diretos na economia e, por conseguinte, no consumo de energia elétrica. No trabalho publicado pelo Banco Central do Brasil [11], as mortes por COVID-19 foram correlacionadas a economia e ao indicador de atividade econômica IBC-BR (Figura 5 à direita), no entanto, o estudo sobre as mortes por COVID EPICOID [12] revela que há no Brasil uma sub notificação de casos e óbitos por COVID-19. Assim, uma vez que não há estatística previa, e considerando que outras doenças deixaram de ser tratadas no momento correto devido à pandemia, assumiu-se que o mais confiável seria utilizar o número de óbitos totais disponibilizados pela plataforma tabnet/DATASUS [10], apesar de não ser o número total de óbitos que estaria disponível no portal dos cartórios de registro civil [13], os dados fornecidos pelos mesmos não contemplam todo o período deste estudo, e assim, entende-se que os dados do SUS representam uma parcela significativa do total (Figura 5 à esquerda).

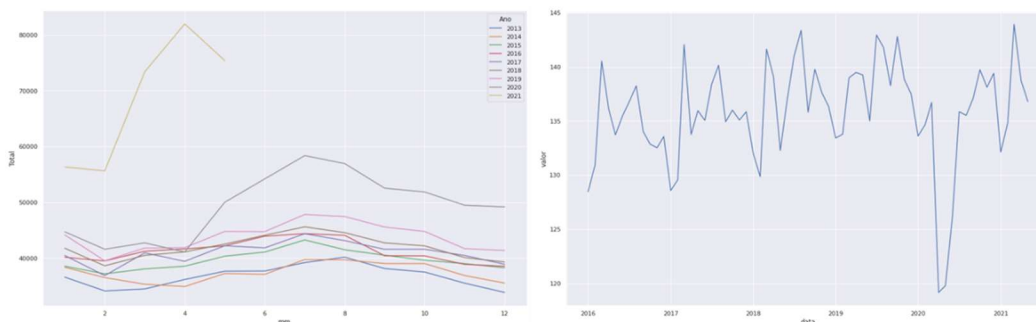


Figura 5 – Óbitos no SUS (à esquerda) e IBC-Br (à direita)

Para melhor definir o cenário econômico também foram agregadas como variáveis de entrada o IPCA e o IBOV, e estes foram escolhidos pela possibilidade de obter-se estudos de projeção oficiais futuras pelo próprio Banco Central do Brasil e pela Bovespa (Figura 6).



Figura 6 – IPCA(à esquerda) e IBOV (à direita)

Nos gráficos da Figura 5 – Óbitos no SUS (à esquerda) e IBC-Br (à direita) Figura 5 e Figura 6 estão caracterizados os comportamentos do mercado, do IBOV se espera medir o otimismo do mercado que se consolidará ou não nas medidas de IPCA e IBC-br, enquanto no quantitativo de óbitos verificamos a consolidação dos resultados da pandemia.

## 2.2 Rede neural

Uma rede neural artificial é um arranjo de células matemáticas computacionais inspiradas no cérebro humano e são utilizados como classificadores ou regressores matemáticos que, a partir de um conjunto de dados, é ajustada para classificar ou estimar um valor com base em dados inéditos.

Um Neurônio Perceptron ou um neurônio computacional possui um determinado número de entradas distintas, estas, por sua vez, são multiplicadas por uma variável de peso individualizado. O produto de peso pela entrada é somado a uma constante chamada “bias”, este somatório é o argumento de uma função de ativação, arbitrária a qual retorna um valor de saída. (Figura 7 à esquerda), este modelo é chamado de “perceptron” o qual teve as primeiras publicações e implementações por Rosenblatt em 1957 [14].

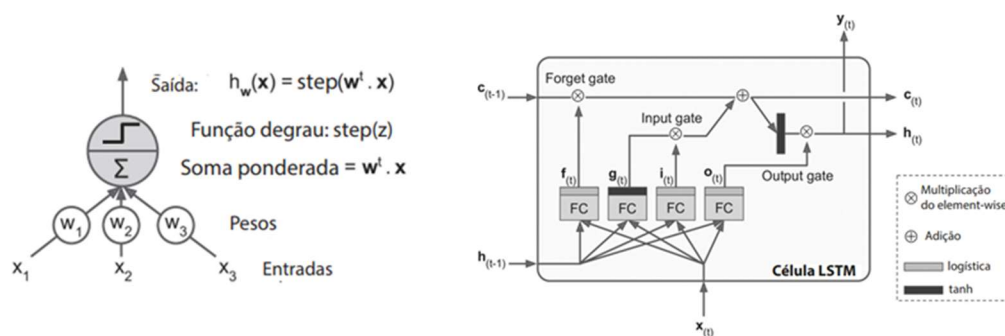


Figura 7 - Neurônio Perceptron (à esquerda) e Célula LSTM (à direita)

Fonte: [15]

Uma Célula LSTM é uma variação das redes neurais recorrentes (RNN), com uma Longa memória de curto prazo (LSTM), é eficiente na predição de séries temporais longas, essa arquitetura se difere de outras por possuir em laço interno de realimentação (*feedback*) e uma célula de memória, que retém informações do passado por mais tempo. A Célula LSTM foi proposta em 1997 por Sepp Hochreiter e Jurgen Schmidhuber e aperfeiçoada em 2014 por Alex Graves, Haşim Sak e em 2015 por Wojciech Zaremba e sua estrutura básica está descrita na Figura 7 (à

direita), a qual tem-se um comportamento básico, similar ao neurônio Perceptron, com uma modificação que é comparada a uma memória onde o estado da própria célula em uma iteração anterior representada pelo índice  $(t-1)$  influencia em um estado atual de índice  $(t)$  onde os sinais  $h_t$  funcionam como um estado de curto prazo e  $c_t$  como um estado de longo prazo.

A ideia-chave da utilização de células LSTM é que ao arranjá-las em rede a mesma possa aprender o que armazenar no estado de longo prazo. Como o estado de longo prazo passa primeiro por um “forget gate”, a memória pode ser apagada ou renovada a cada iteração, ao passo que um ajuste dessa memória pode ser feito através do “input gate” através da operação de adição. Já a memória de curto prazo é a mesma saída da própria célula onde a memória de longo prazo é filtrada pela função tanh e multiplicada pela “output-gate”. Uma célula LSTM pode aprender a reconhecer uma entrada importante, armazená-la no estado de longo prazo, aprender a preservá-la pelo tempo necessário e aprender a extrai-la sempre que for preciso. [15]

Obtém-se uma rede neural ao agrupar-se estes neurônios ou células e interconectá-los em níveis chamados camadas podendo-se variar o formato do arranjo, o número de camadas, o tipo construção do neurônio. Para cada variável alterada obtém-se arquiteturas diferentes que são especializados para resolver diferentes tipos de problemas, na qual a primeira camada deste arranjo é denominada camada de entrada e não processa nenhum cálculo, somente serve de entrada para os parâmetros avaliados possuindo a mesma dimensão dos parâmetros de entrada. Além disso, a última camada é a chamada camada de saída e entrega o valor final a qual, neste caso, é formada por apenas um neurônio. Todas as camadas intermediárias entre as camadas de entrada e saída são chamadas camadas ocultas. A Figura 8 ilustra um exemplo desta arquitetura de rede.

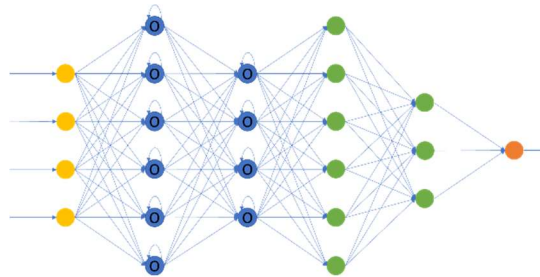


Figura 8 - Arquitetura LSTM+MLP

No presente trabalho também é avaliada uma variação de RNN, chamada Bidirecional RNN (BRNN), essa estrutura proposta por Mike Schuster e Kuldip K. Paliwal [16], onde a célula utiliza tanto dados de um estado futuro ( $t + 1$ ), quanto de um estado passado ( $t - 1$ ) para determinar o estado ( $t$ ).

### 2.3 Otimização da Rede por Algoritmo genético

O Algoritmo genético é uma técnica de otimização, baseada nos princípios de seleção natural, nas quais a partir de uma população inicial e uma avaliação de aptidão, os indivíduos mais aptos são selecionados e geram descendentes, propagando seus genes e os menos aptos morrem sem deixar descendentes e os descendentes podem ou não sofrer algum tipo de mutação aleatorizada. Neste caso, cada indivíduo representa uma arquitetura para a Rede Neural, variando profundidade (número de camadas ocultas), número de neurônios por camada, número de interconexões e complexidade das primeiras camadas (LSTM ou BiLSTM). A aptidão é a medida do menor MAPE (do Inglês Mean Absolute Percentual Error). Seleção é o processo pelo qual uma certa proporção de indivíduos é selecionada para acasalar entre si e gerar novos descendentes. Assim como na seleção natural da vida real, os indivíduos mais aptos têm maiores chances de sobreviver e, portanto, de passar seus genes para a próxima geração. Embora existam versões com mais indivíduos, geralmente o processo de seleção combina dois indivíduos, criando pares de indivíduos. Existem quatro estratégias principais: (1) Emparelhamento: consiste em emparelhar os cromossomos mais aptos, dois a dois e realizar o cruzamento de genes seguindo uma regra pré estabelecida ou de maneira aleatória. (2) Aleatório: esta estratégia consiste em selecionar indivíduos aleatoriamente do pool de acasalamento. (3) Roleta: esta estratégia também segue um princípio aleatório, mas os indivíduos mais aptos têm maiores probabilidades de serem selecionados. (4) Torneio: com essa estratégia o algoritmo primeiro seleciona alguns indivíduos como candidatos e, em seguida, seleciona o indivíduo mais apto. Esta opção tem a vantagem de não exigir que os indivíduos sejam classificados primeiro por aptidão. Para este experimento será utilizada a estratégia “roleta”.

### 2.4 Estudo de Caso RS

Seguindo o desenvolvimento da rede LSTM, foi realizado o teste de conceito com o consumo rural do RS utilizando a otimização pelo algoritmo genético discutido anteriormente para definir o arranjo dessa estrutura. Foram testadas variações de até seis camadas ocultas de profundidade com três variações de “dropout” que é quando se desliga (aleatoriamente) algumas interconexões, portanto nas conexões entra as camadas poderia ser de três formas: totalmente interconectada; 90% interconectada (dropout = 0,1); ou 80% interconectada (dropout = 0,2). Cada uma das camadas poderia ter até 64 neurônios (onde 0 (zero) descartaria a camada), onde as duas primeiras poderiam assumir duas formas LSTM (unidirecional) ou BLSTM (bidirecional LSTM), as quatro últimas camadas poderiam

assumir apenas o formato MLP, após os testes descritos, onde foi possível aproximar a série de 2018 com um erro (MAPE) de 0,0327.

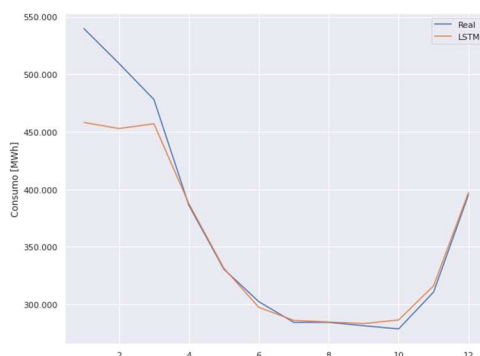


Figura 9 - Teste Valor Real vs Predito LSTM

## 2.5 Estudo de Caso BR

Inicialmente, este estudo foi segmentado em classes de consumo, para viabilizar uma análise mais detalhada, considerando somente o consumidor rural no estado do Rio Grande do Sul. Em seguida, foram realizadas generalizações, para qualquer classe de consumo, para qualquer unidade da federação e para o Brasil.

Na generalização por classe de consumo há duas possibilidades de teste. No primeiro teste utilizou-se uma rede com  $n$  saídas, sendo uma para cada  $n$  Setor N1, e o segundo teste foi realizado para  $n$  redes com uma única saída, sendo uma rede para cada Setor N1. Nesta segunda possibilidade obteve-se a melhor performance tanto em tempo de processamento quanto em precisão. Assim, esta segunda possibilidade foi a utilizada para o estudo de caso BR.

Na generalização por Unidade da Federação - UF ou Brasil – BR, foi observado que a previsão de um estado é similar a qualquer outro estado e que o algoritmo genético proposto seria capaz de ajustar a rede da mesma forma que o fez para o estado do RS, segmentado por Setor N1 ou não. O modelo proposto teve suas entradas ajustadas para considerar todas as estações meteorológicas disponíveis no INMET e os dados econômicos e de órbitas foram aplicados apenas ao estudo de caso BR, uma vez que estes dados coletados para o estudo são divulgados para todo o território nacional. Os dados climáticos foram reavaliados como no estudo de caso anterior e os resultados foram equivalentes.

Na Figura 10 pode-se verificar que, no ano de 2018, o consumo do Brasil não teve uma variabilidade tão grande quando o estudo de caso (RS), mas, ainda assim, o método obteve um bom ajuste dos valores preditos em comparação aos valores reais verificados. O que é comprovado na mesma figura, a qual traz os valores de erro dentro dos limites estipulados como meta para este estudo de  $MAPE \leq 5\%$ . Ressalta-se que, ajustando-se os dados disponíveis até a data deste estudo, não estavam publicados os dados do anuário EPE. Ou seja, a publicação de 2021 com os dados de 2020. Neste caso, como solução utilizou-se os dados da "resenha mensal da EPE" [6], que são os dados preliminares que irão compor o anuário com uma pequena diferença. Nesta resenha os dados são publicados em quatro setores: Comercial, Industrial, Residencial e Outros. A partir destes dados foi criado o "Consumo" para representar o consumo total.

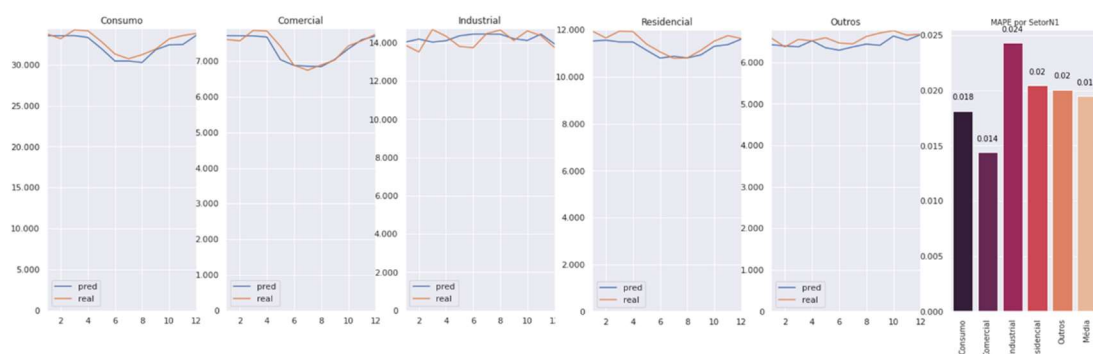


Figura 10 - Teste Valor Real vs Predito LSTM BR

## 2.6 Estudo de Caso BR em Ambiente Pandêmico

Aplicando a mesma metodologia ao período de 2015 a 2019 para então prever o consumo de 2020 e valendo-se exclusivamente da sazonalidade do clima e da inércia intrínseca do consumo de energia, não foi possível prever o consumo de 2020 com a mesma performance. No caso a previsão passa de um erro médio de 0,019 para o 0,047 Figura 11.

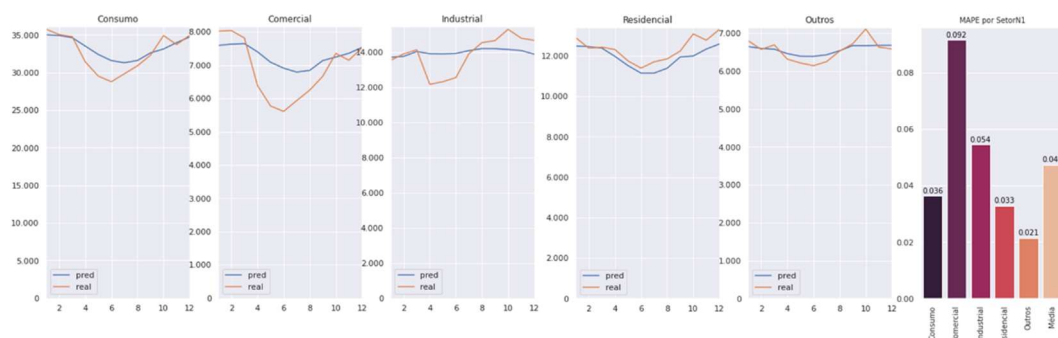


Figura 11 - Teste Valor Real vs Predito 2020 e MAPE (BR Resenha)

Foi definida uma população inicial de 16 genomas e 32 épocas de iteração para cada classe de consumo a fim de garantir a repetibilidade e a comparação com os mesmos parâmetros objetivando a definição da rede via algoritmo genético. Além disso, para assegurar a correta escolha de variáveis, foram realizados 16 cenários de testes de forma a testar todos os arranjos possíveis para as variáveis escolhidas e os resultados destes testes são mostrados no Quadro 1. O melhor arranjo testado (além das variáveis climáticas usadas em todos os cenários de teste), foi obtido pelas entradas IBCBR, IPCA e Óbitos, nas quais é possível observar que estas variáveis estão presentes (cada uma) em quatro dos cinco melhores resultados evidenciando sua relevância.

Quadro 1 - Cenários de Teste

Entradas					MAPE						#
Clima	IBOV	IBCBR	IPCA	Óbitos	Consumo	Comercial	Industrial	Residencial	Outros	Média	
1		1	1	1	0,032	0,079	0,038	0,013	0,017	0,036	1
1	1	1	1	1	0,031	0,079	0,042	0,021	0,019	0,038	2
1	1	1		1	0,029	0,084	0,041	0,019	0,023	0,039	3
1		1	1		0,029	0,087	0,037	0,034	0,018	0,041	4
1	1		1	1	0,026	0,086	0,056	0,019	0,022	0,042	5
1				1	0,036	0,085	0,053	0,014	0,021	0,042	6
1	1	1			0,032	0,084	0,044	0,027	0,023	0,042	7
1		1		1	0,037	0,087	0,046	0,019	0,023	0,042	8
1	1	1	1		0,031	0,087	0,044	0,031	0,021	0,043	9
1			1	1	0,038	0,086	0,055	0,015	0,022	0,043	10
1	1				0,034	0,09	0,055	0,033	0,022	0,047	11
1					0,036	0,092	0,054	0,033	0,021	0,047	12
1		1			0,039	0,082	0,044	0,052	0,023	0,048	13
1			1		0,035	0,09	0,057	0,038	0,022	0,048	14
1	1		1		0,036	0,094	0,055	0,039	0,023	0,049	15
1	1			1	0,042	0,099	0,06	0,045	0,026	0,054	16

Pode-se evidenciar na Figura 12, a qual ilustra o melhor arranjo testado quando comparada com a **Fonte de referência não encontrada**, que houve uma melhoria na previsão de todas as classes e que a configuração considerando apenas o clima seria a 12ª melhor previsão dentre as 16 testadas. Ainda é possível verificar no Quadro 1 que a inclusão da variável "óbitos" melhora a previsão da 12ª para 6ª posição.

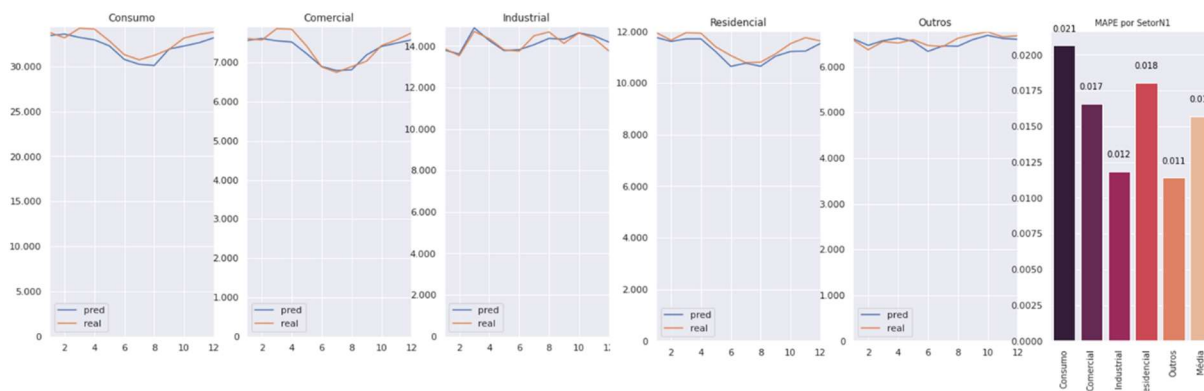


Figura 12 - Teste Valor Real vs Predito 2020 e MAPE (BR Resenha IBCBR, IPCA, Óbitos)

Reaplicando esta mesma estrutura para a previsão de 2018 a fim de verificar se as novas variáveis interfeririam positivamente também num cenário sem pandemia. Pode-se verificar que o consumo na Figura 13 teve um desempenho levemente superior quando comparados os erros (MAPE) da **Erro! Fonte de referência não encontrada.**, demonstrando que o uso destas novas variáveis é tanto relevante em período de pandemia quanto fora dele.

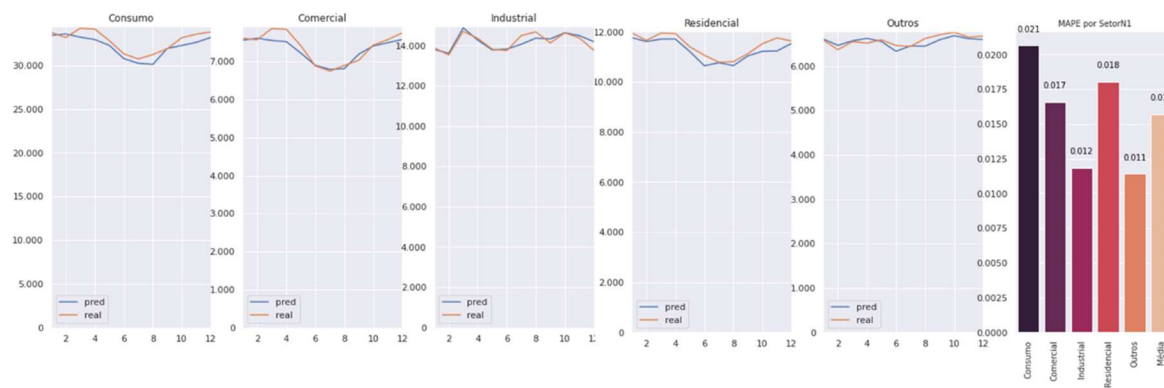


Figura 13 - Teste Valor Real vs Predito 2018 e MAPE (BR Resenha IBCBR, IPCA, Óbitos)

Assim, aplicando-se a variável “óbitos” ao modelo de previsão pode-se prever o impacto social de uma pandemia ou calamidade de saúde no consumo de energia elétrica. Ainda, o modelo pode ser utilizado para projetar o consumo durante a atenuação de eventos pandêmicos.

## CONCLUSÕES

A análise das aplicações do modelo aqui proposto comprovou que a técnica escolhida, regressor LSTM/MLP, associada a um dimensionamento da rede neural através de algoritmo genético, foi eficiente na previsão do consumo de energia elétrica. A inclusão da variável do número de óbitos, associada as tradicionais variáveis econômicas e sociais, mesmo em um contexto adverso como a pandemia de COVID-19, resultou em melhoria na previsão de consumo de energia elétrica.

Atendendo a regulamentação da Agência Nacional de Energia Elétrica - ANEEL quanto ao erro de projeção, o modelo permite obter uma previsão com erro menor do que 5%, resultando em um erro médio final de 3,6%, a exceção foi a classe comercial que apresentou um erro de 7,9% em ambiente pandêmico, devido à falta de uma variável sócio-econômica que se correlacionasse com esta classe. No entanto, em um contexto fora do período de pandemia, obteve-se um erro médio de 1,6%, e um melhor caso de erro de 1,1% para a classe “Outros” e um pior caso de 2,1% para a classe “Consumo Total”, ambos os resultados são considerados muito bons.

O modelo fornece projeções de consumo de energia elétrica relevantes para o planejamento de sistemas elétricos e, como consequência, o melhor uso dos recursos energéticos, com o uso de variáveis e informações públicas de fácil acesso, no que se refere a dados históricos e projeções.

As contribuições do trabalho são a comprovação de que o número de óbitos influencia o consumo de energia e que um enfoque socio-econômico-ambiental permite uma melhor previsão de consumo. Ainda, a estrutura do modelo algorítmico ajusta via algoritmo de otimização genética uma rede neural a partir de quaisquer entradas numéricas e assim pode estimar um valor de saída com base em dados históricos e projeções correlatas, desde que sejam testadas e pesquisadas entradas adequadas para cada estudo de caso.

O trabalho aqui desenvolvido revelou aperfeiçoamentos no modelo proposto, como, por exemplo, o estudo do número de internações que não resultaram em óbitos e de um indicador econômico que melhor caracterize a atividade comercial. A comparação deste método com métodos usualmente empregados nos estudos de projeção de consumo de energia elétrica também é desejável, visando melhoramentos.



## BIBLIOGRAFIA

- [1] M. Harrison, *Machine Learning – Guia de Referência Rápida: Trabalhando com dados estruturados em Python*, 1ed ed., São Paulo, SP: Novatec Editora, 2020.
- [2] Empresa de Pesquisa Energética, “EPE,” EPE, 2019. [Online]. Available: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/dados-abertos/dados-do-anuario-estatistico-de-energia-eletrica>. [Acesso em Março 2021].
- [3] G. v. Rossum, “Python,” [Online]. Available: <https://www.python.org/>. [Acesso em 2021].
- [4] A. M. Abadi, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015.
- [5] S. Seabold e J. Perktold, “statsmodels: Econometric and statistical modeling with python,” *9th Python in Science Conference*, 2010.
- [6] “Pandas,” 2021. [Online]. Available: [pandas.pydata.org](https://pandas.pydata.org).
- [7] Empresa de Pesquisa Energética, “EPE - Resenha Mensal,” [Online]. Available: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/resenha-mensal-do-mercado-de-energia-eletrica>. [Acesso em julho 2021].
- [8] Banco Central do Brasil, “Dados Abertos,” [Online]. Available: <https://dadosabertos.bcb.gov.br/dataset/>. [Acesso em ago 2021].
- [9] B3 - Brasil Bolsa Balcão, “Market Data e Índices,” [Online]. Available: [http://www.b3.com.br/pt\\_br/market-data-e-indices/indices/indices-amplos/indice-ibovespa-ibovespa-estatisticas-historicas.htm](http://www.b3.com.br/pt_br/market-data-e-indices/indices/indices-amplos/indice-ibovespa-ibovespa-estatisticas-historicas.htm). [Acesso em jun 2021].
- [10] Ministério da Saúde, “DataSus,” [Online]. Available: <https://datasus.saude.gov.br/>. [Acesso em mar 2021].
- [11] B. C. d. B. -. BCB, “Intensidade da pandemia e atividade econômica,” Junho 2021. [Online]. Available: <https://www.bcb.gov.br/content/ri/relatorioinflacao/202106/ri202106b1p.pdf>. [Acesso em 22 Julho 2021].
- [12] H. P. M. A. B. A. H. B. B. F. Mesenburg MA, *Chronic non-communicable diseases and COVID-19: EPICOVID-19 Brazil results*, vol. 55, 2021, p. 38.
- [13] Transparência, “Registro Civil,” [Online]. Available: <https://transparencia.registrocivil.org.br/registros>. [Acesso em abril 2021].
- [14] F. ROSENBLATT, “THE PERCEPTRON: A PROBABILISTIC MODEL FOR,” *Psychological Review*, vol. 65, nº 6, pp. 386-408, 23 abril 1958.
- [15] A. Géron, *Mãos à Obra Aprendizado de Máquina com Scikit-Learn & TensorFlow: Conceitos, Ferramentas e Técnicas Para a Construção de Sistemas Inteligentes*, Rio de Janeiro: Alta Books, 2019.
- [16] M. Schuster e K. K. Paliwal, “Bidirectional Recurrent Neural Networks,” *IEEE Transactions on Signal Processing*, vol. 45, pp. 2673-2681, 1997.

## DADOS BIOGRÁFICOS



LEANDRO VARGAS BARBOSA, nasceu em Caxias do Sul – RS, em 1986. Formou-se como Técnico em Eletrotécnica (2005) pela Escola Técnica Estadual 25 de Julho. Graduiu-se em Matemática Licenciatura pelo Centro Universitário Leonardo Da Vinci – UNIASSELVI (2014). Atualmente é formando em Engenharia Elétrica pela Universidade Federal do Rio Grande do Sul – UFRGS (2021). Desde 2006 atua como Técnico em Eletrotécnica na Companhia Estadual de Energia Elétrica - CEEE-T, atualmente lotado na área de Operação da Transmissão.



GLADIS BORDIN, possui mestrado em Engenharia Elétrica pela Universidade Federal de Santa Maria – UFSM (1988) e doutorado em Engenharia Elétrica pela Universidade Federal de Santa Catarina – UFSC (2000). Professora da Universidade Federal do Rio Grande do Sul - UFRGS, junto ao Departamento de Engenharia Elétrica. Pesquisadora na área de Sistemas Elétricos de Potência, com ênfase em Planejamento da Expansão de Sistemas Elétricos, atuando principalmente nos seguintes temas: projeção de demanda; modelagem e simulação de mercados, empresas e consumidores de energia elétrica.

