

# *Previsão de carga horária como parâmetro do modelo de precificação e despacho hidrotérmico de curto prazo (DESSEM)*

Marina da Silva, Priscila dos Santos, Leonardo Forero, Harold de Mello Junior

Departamento de Engenharia Elétrica  
Universidade do Estado do Rio de Janeiro (UERJ)  
Rio de Janeiro, Brasil

[marinaenguerj@outlook.com](mailto:marinaenguerj@outlook.com), [priscila\\_cs1995@hotmail.com](mailto:priscila_cs1995@hotmail.com)  
[leofome@hotmail.com](mailto:leofome@hotmail.com), [harold.dias@gmail.com](mailto:harold.dias@gmail.com)

Allan Gurwicz

Departamento de Engenharia Elétrica  
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)  
Rio de Janeiro, Brasil

[agurwicz@ele.puc-rio.br](mailto:agurwicz@ele.puc-rio.br)

**Abstract**— A previsão de carga é utilizada na controlabilidade do Sistema Integrado Nacional (SIN) e para a garantia da segurança elétrica e energética do atendimento na ponta. Atualmente, é essencial em termos de tomada de decisão em tempo real para o Operador Nacional do Sistema (ONS) e para o planejamento da coordenação hidrotérmica de médio e longo prazo. A previsão precisa da carga resulta em redução de custos operacionais e na maior confiabilidade do sistema elétrico, auxiliando na definição do número de unidades geradoras que precisam ser construídas nos próximos anos e na quantidade de energia que deve ser gerada a cada hora ao longo dos dias. Estima-se, para 2020, o início da operação do sistema DESSEM, que calculará o custo de energia no SIN baseado em dados de previsão de carga, e otimizará a expedição hidrotérmica a cada hora. Dada a relevância desta tarefa, este artigo apresenta três técnicas para a previsão de carga com granularidade horária: um modelo autorregressivo integrado de médias móveis (ARIMA) e dois modelos não-lineares, baseados em redes neurais artificiais recorrentes (GRU e LSTM), que têm obtido resultados superiores quando comparado a outras técnicas de previsão. Verifica-se que as redes LSTM e GRU apresentam resultados com menor erro percentual absoluto que o modelo ARIMA. Este tipo de modelo foi usado em um caso real por uma comercializadora de energia na cidade do Rio de Janeiro a qual conseguiu melhorar em 7% suas previsões, reduzindo custos de operação.

**Palavras-chave**— planejamento, setor elétrico, previsão de carga, preço horário, DESSEM, ARIMA, LSTM

## I. INTRODUÇÃO

O Sistema Elétrico Brasileiro (SEB) vem passando por diversas reestruturações com a criação de novas empresas com atribuições específicas relacionadas à operação e ao planejamento deste setor, tornando-se cada vez mais complexo. Recentemente foram discutidas diretrizes para permitir o máximo de acoplamento entre a operação e a formação de preços. Uma das propostas de aprimoramento do SEB, da

Consulta Pública nº 33/2017 do Ministério de Minas e Energia (MME), foi a implementação até 2020 de preços horários no Mercado de Curto Prazo (MCP), independente da opção de geração. A metodologia para determinação do preço da energia no Brasil é operacionalizada pelo Modelo de Planejamento da Operação de Sistemas Hidrotérmicos Interligados de Longo e Médio Prazo (NEWAVE) e pelo Modelo de Planejamento da Operação de Sistemas Hidrotérmicos Interligados de Curto Prazo (DECOMP). Atualmente, está sendo validado pelo Operador Nacional do Sistema (ONS) e pela Câmara de Comercialização de Energia Elétrica (CCEE), o Modelo de Despacho Hidrotérmico de Curto Prazo (DESSEM) [1], que auxiliará no planejamento e programação da operação dos sistemas hidrotérmicos no curtíssimo prazo, com granularidade horária [2]. Concomitantemente com a evolução do mercado e a complexidade do sistema brasileiro, faz-se necessário identificar a demanda de energia com a maior precisão possível, uma vez que as empresas concessionárias devem fornecer aos consumidores energia elétrica com qualidade e continuidade asseguradas, com especial atenção aos aspectos de segurança. E, por conseguinte, reduzir os custos de operação.

Considerando a importância dos dados de previsão de carga para a modelagem do preço que será implementada no Brasil a partir de 2020 [3], este artigo consiste em aplicar técnicas computacionais para a previsão de carga horária com o objetivo de auxiliar os agentes de mercado no planejamento diário de suas operações. A fim de alcançar este propósito foram empregados: Modelo autorregressivo integrado de médias móveis, ARIMA [4]; Modelos não-lineares, baseados em Redes Neurais Artificiais Recorrentes (RNN) [5]: GRU [6] e LSTM [7].

A previsão de carga horária de sete dias a frente foi realizada com duas séries temporais: (i) carga de energia do submercado

Sudeste/Centro-Oeste e (ii) temperatura média por hora referente ao submercado Sudeste/Centro-Oeste dos últimos 15 anos (janeiro de 2003 a dezembro de 2017).

## II. PREVISÃO DE CARGA

### A. Definição

Um sistema hidrotérmico, que é o caso do brasileiro, é aquele que combina usinas hidrelétricas com termelétricas a fim de conciliar a melhor confiabilidade provida pelas termelétricas com a vantagem econômica e ambiental inerentes às hidrelétricas. Como o Brasil apresenta predominância de usinas hidrelétricas no seu parque de geração, utilizam-se modelos matemáticos para encontrar a solução ótima de equilíbrio entre o benefício presente do uso da água e o benefício futuro do seu armazenamento, medido pela não utilização dos combustíveis das usinas termelétricas. Com base nas previsões de condições hidrológicas, na demanda de energia, nos preços de combustível, no custo de déficit, na entrada de novos projetos e na disponibilidade de equipamentos de geração e transmissão, o modelo de precificação obtém o despacho (geração) ótimo para o período em estudo, definindo a geração hidráulica e a geração térmica para cada submercado. Como resultados deste processo são obtidos os Custos Marginais de Operação (CMO) para o período estudado, para cada patamar de carga e para cada submercado [8]. Está previsto para 2020 o início de operação de um novo modelo, denominado DESSEM, para o cálculo preço de curtíssimo prazo, com discretização diária [3].

Tendo em vista a importância do comportamento futuro do sistema para planejamento e operação do sistema elétrico, a análise da previsão de carga é fundamental. A previsão que será feita neste projeto é de curtíssimo prazo, com horizonte de uma semana e discretização horária. Ao se trabalhar com curtíssimo prazo, a adaptabilidade da previsão é fundamental devido aos impactos ambientais da mesma.

A carga possui comportamento frequentemente não-linear, influenciado por diversos fatores como condições climáticas, hora do dia e dia da semana. Em consequência disso, modelos de regressão linear tendem a não apresentar bons resultados. Por outro lado, a aplicação de redes neurais artificiais tem aumentado em tarefas de previsão de carga devido à melhor capacidade de modelar dados não-lineares, como é o caso do presente trabalho [5].

### B. Séries Temporais

Série temporal é uma função aleatória de uma variável independente relacionada com o tempo, tratando-se de uma coleção de observações feitas sequencialmente ao longo do tempo. A característica mais importante é que as observações vizinhas são dependentes e o interesse é justamente analisar e modelar tal dependência [9]. Nesse caso, a ordem dos dados é extremamente importante.

## III. MODELAGEM

### A. Modelo Linear

Neste trabalho o modelo linear utilizado foi o ARIMA, modelo autorregressivo integrado de médias móveis [4]. Nele a série temporal é gerada através de um processo estocástico [9]. A notação utilizada para representação desse modelo é  $ARIMA(p, d, q)$ , onde  $p$  representa o número de termos autorregressivos,  $d$  é o número de diferenciações necessárias para a série se tornar estacionária e  $q$  é o número de termos de média móvel [9].

### B. Redes Neurais Artificiais

As redes neurais artificiais são, de fato, redes inspiradas na estrutura do cérebro, tendo como propósito apresentar características semelhantes às do comportamento humano. São compostas por neurônios artificiais que são capazes de calcular determinadas funções matemáticas, sendo muito eficientes no aprendizado de padrões a partir de dados não-lineares [5].

Ao contrário dos modelos lineares, as redes neurais apresentam poucas premissas, o que aumenta sua flexibilidade. Além disso, a capacidade de aproximar qualquer função contínua com precisão arbitrária, dado um número suficiente de neurônios, é outra grande vantagem. Elas conseguem lidar com dados ruidosos e imprecisos, o que as torna uma ferramenta poderosa de previsão de séries.

- LSTM

Uma das redes neurais utilizada neste trabalho foi a rede LSTM (*Long Short Term Memory*) [7]. Trata-se de um tipo de rede neural recorrente (RNN), amplamente bem sucedida em uma gama de diferentes problemas devido à capacidade de distinguir entre ocorrências recentes e distantes. Estas redes atribuem pesos diferentes para cada ocorrência e são capazes de ignorar memórias que possam ser consideradas irrelevantes para prever a próxima saída. Desta forma, elas têm maior capacidade de lidar com sequências longas de dados quando comparadas a outras redes neurais recorrentes [10].

As LSTMs diferem das redes neurais recorrentes tradicionais por possuírem uma estrutura de módulo de repetição mais complexo, na qual os portões (*gates*) são responsáveis por manter o fluxo de erro constante. Desde a sua introdução, nos anos 90, a arquitetura das redes LSTM vem passando por diversas modificações. [11], em 1999, apresentaram o portão de esquecimento (*forget gate*), não existente no modelo proposto inicialmente por [7], e em 2000 introduziram as conexões *peephole*, em que os portões enxergam o estado anterior da célula.

- GRU

Mais recentemente, foram introduzidas as chamadas *Gated Recurrent Unit* (GRU) [6] que combinam os portões de entrada e esquecimento, criando um portão único de atualização. Este tipo de rede pode ser considerado uma versão compacta da rede LSTM na medida, requerendo, portanto, menos parâmetros.

De uma forma geral, a GRU não possui o estado celular e usa o estado oculto para transferir a informação. A GRU também não possui o portão de saída (*output*). Desta forma, por ser mais simples, a rede GRU é também mais rápida para treinar e precisa de menos dados para generalizar [12],[13].

Para resolver o problema da dissipação do gradiente de uma rede neural recorrente padrão, a GRU usa dois portões, *reset* e *update gate*. Basicamente, são dois vetores que decidem quais informações devem ser passadas para a saída. Eles podem ser treinados para manter informações de muito tempo atrás, sem dissipá-las com o tempo ou remover informações irrelevantes para a previsão.

#### IV. METODOLOGIA

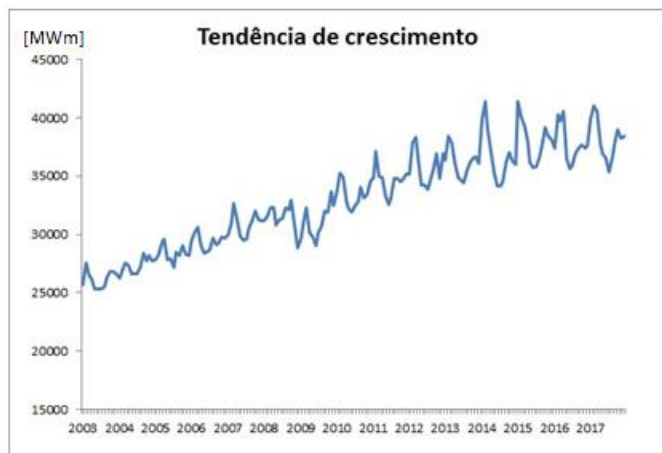
##### A. Análise dos Dados de Entrada

###### 1) Carga

A série temporal de carga foi analisada e verificou-se que os perfis horários dos dias úteis apresentam o mesmo comportamento, sendo que os horários de pico ocorrem no intervalo de 14 horas às 16 horas e os horários de menor consumo de 2 horas às 5 horas. Vale ressaltar que nos últimos anos houve mudanças nos hábitos de consumo de energia da população brasileira, e o horário de pico, que antes era entre 18 horas e 21 horas, foi deslocado para o período da tarde [14].

Analisando a evolução da carga no período em questão, encontrou-se a componente de tendência deste crescimento através da função *trend* do Excel. A curva obtida é mostrada na Figura 1 (referente à média mensal). A tendência observada será importante para realização da previsão com o modelo ARIMA.

Figura 1. Evolução da carga no período de análise



###### 2) Temperatura

Para a série de temperatura foi considerada os graus-dia de resfriamento (*Cooling Degree Days – CDD*), que é uma medida de quão quente está um determinado dia e o consumo de energia decorrente da utilização de equipamentos de ar condicionado. Todas as horas do dia que ultrapassem a temperatura base selecionada contribuem para o CDD, enquanto as horas com temperatura abaixo são desconsideradas. A carga por refrigeração é assumida como proporcional ao número de CDDs, considerando que, para as temperaturas abaixo da base, não há necessidade de refrigeração porque a população não aciona condicionadores de ar.

Foram utilizados os valores de consumo por capital de cada um desses estados como base de ponderação das temperaturas para, assim, encontrar um valor médio representativo de temperatura para toda a região em estudo. Esses dados de consumo foram baseados no estudo realizado pela EPE no Plano Decenal de Expansão de Energia (PDE) 2027 [3].

##### B. Definição de variáveis

###### 1) Modelo ARIMA

As variáveis foram escolhidas com base na influência que estas apresentam sobre o desenvolvimento da carga. Desta forma, foram definidas as seguintes variáveis:

- Dia útil
- Sábado
- Domingo/feriado nacional
- Meses do ano de janeiro a outubro
- Crescimento econômico (PIB)
- Temperaturas regionais expressas como carga ponderada (CDD)
- Fator de tendência que representa o crescimento da população e dos aparelhos elétricos ao longo do tempo, ou seja, captura o crescimento da carga no longo prazo que é independente dos ciclos econômicos

Um sistema de relações econométricas foi aplicado aos dados de carga horária para o subsistema Sudeste/Centro-Oeste para o período de janeiro de 2003 a dezembro de 2017, o que permitiu 24 relacionamentos para as previsões de carga por hora para o submercado de estudo com base nas variáveis explicativas citadas anteriormente.

Os parâmetros do modelo são definidos pelos gráficos de autocorrelação e autocorrelação parcial e, dessa forma, a previsão da série parametrizou-se o modelo ARIMA(1, 0, 1). A modelagem foi feita no Python e, como resultado das 24 regressões (uma para cada hora do dia), encontrou-se os valores de coeficientes para as 16 variáveis.

### 2) GRU

Foram testados diferentes parâmetros variáveis da rede a fim de identificar o que resulta em menor erro do modelo de previsão. Os parâmetros são:

- *Window*, que representa o tamanho da janela deslizante,
- *Batch size*, que indica o tamanho do lote de dados que é passado para rede por vez,
- *Epoch*, número de épocas de treinamento, ou seja, a quantidade de vezes que os dados de treino são passados inteiramente pela rede, e
- *Units*, número de processadores por camada.

A configuração da rede GRU que resultou nos menores valores de erros está expressa na Tabela 1.

Os dados foram processados e apresentados da mesma forma que foram para o modelo ARIMA: a série foi dividida de acordo com as 24 horas do dia. Deste modo, o conjunto de entrada da rede possui 15 anos a serem analisados x 365 dias = 5.475 dados para cada hora do dia. Desses dados 80% foram separados para treino, 10% para validação e 10% para teste.

Tabela 1 – Melhores parâmetros GRU

<i>Epoch</i>	100
<i>Batch size</i>	128
<i>Window</i>	10
<i>Units</i>	64

### 3) LSTM:

Foram testados diferentes parâmetros variáveis da rede a fim de identificar o que resulta em menor erro do modelo de previsão. Os parâmetros são:

- *Window*, que representa o tamanho da janela deslizante,
- *Batch size*, que indica o tamanho do lote de dados que é passado para rede por vez,
- *Epoch*, número de épocas de treinamento, ou seja, a quantidade de vezes que os dados de treino são passados inteiramente pela rede, e
- *Unit*, número de processadores por camada.

Foram utilizados dois casos diferentes na metodologia LSTM os valores que apresentaram os menores erros estão apresentados nas Tabelas 2 e 3.

#### *LSTM (Caso 1)*

Este primeiro caso considera o mesmo processamento de dados realizado para os modelos ARIMA e GRU: série dividida de acordo com as 24 horas do dia. Deste modo, o conjunto de entrada da rede possui 15 anos a serem analisados x 365 dias =

5.475 dados para cada hora do dia. Desses dados 80% foram separados para treino, 10% para validação e 10% para teste.

Tabela 2 – Melhores parâmetros LSTM (Caso 1)

<i>Epoch</i>	150
<i>Batch size</i>	256
<i>Window</i>	10
<i>Units</i>	128

#### *LSTM (Caso 2)*

No segundo caso, a série foi dividida em dias da semana, sábados e domingos. Desta forma, têm-se três saídas da previsão, diferente do primeiro caso, em que os dados horários de carga não são categorizados em dias da semana, na apresentação para a rede.

Tabela 3 – Melhores parâmetros LSTM (Caso 2)

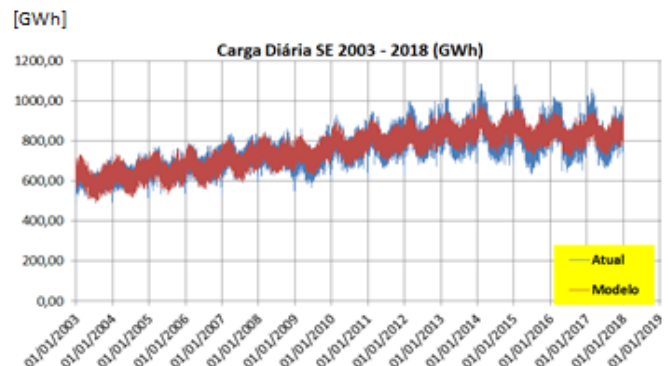
<i>Epoch</i>	150
<i>Batch size</i>	1029
<i>Window</i>	20
<i>Units</i>	128

## V. RESULTADOS

### A. ARIMA

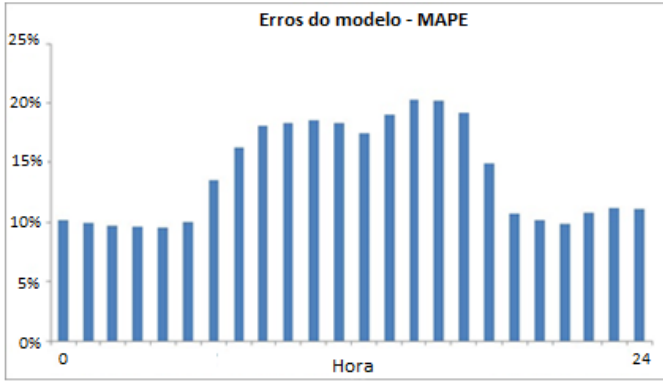
Na Figura 2 é apresentada a comparação da previsão obtida com o modelo ARIMA com os valores reais de carga. Vale ressaltar que os valores dos anos de 2018 e 2019 são previsões de dados inexistentes inicialmente na série original.

Figura 2. Comparação da previsão do modelo ARIMA com os dados originais da série



A acurácia da previsão do modelo ARIMA pode ser verificada com o MAPE obtido disposto na Fig. 3. Observa-se que o erro está na faixa de 10 a 20% e com valores acima de 15% entre 8 e 18 horas.

Figura 3. Erros do modelo ARIMA



### B. GRU

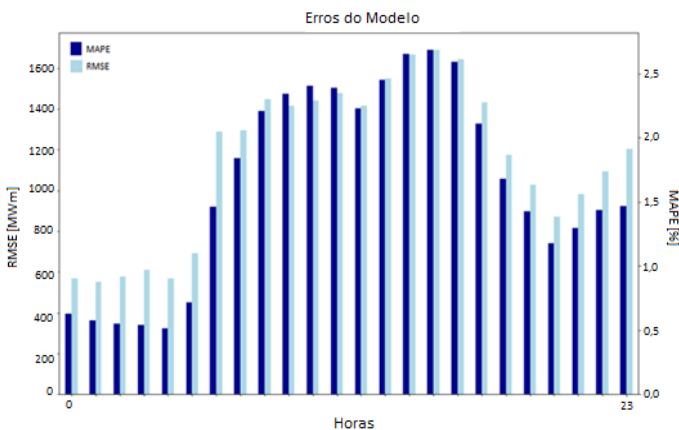
O MAPE neste modelo foi 1,84%, bem próximo ao obtido com a rede LSTM. A vantagem do modelo GRU é a capacidade de realizar a previsão com menor esforço computacional que os demais modelos.

### C. LSTM (Caso 1)

Na fase de teste deste modelo foram gerados os 24 gráficos comparativos, um para cada hora do dia, entre os valores da série temporal de entrada do modelo e os valores de saída.

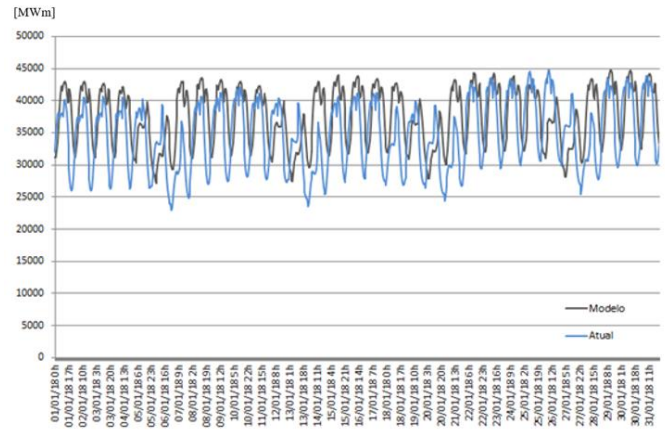
Foram calculados os erros da fase de teste (RMSE e MAPE), como apresentado na Figura 4. Conforme pode ser observado, o erro MAPE máximo obtido foi de 2,5%, resultado bastante satisfatório.

Figura 4. Erros de teste da LSTM (Caso 1)



Em seguida, foi realizada a previsão dos dados da série que não foram utilizados no treinamento da rede para calcular os erros de validação. A Fig. 5 ilustra o resultado obtido. O deslocamento observado em alguns dias refere-se ao tamanho da janela escolhido.

Figura 5. Comparação da previsão do modelo LSTM (Caso 1) com os dados originais da série (em MWm)

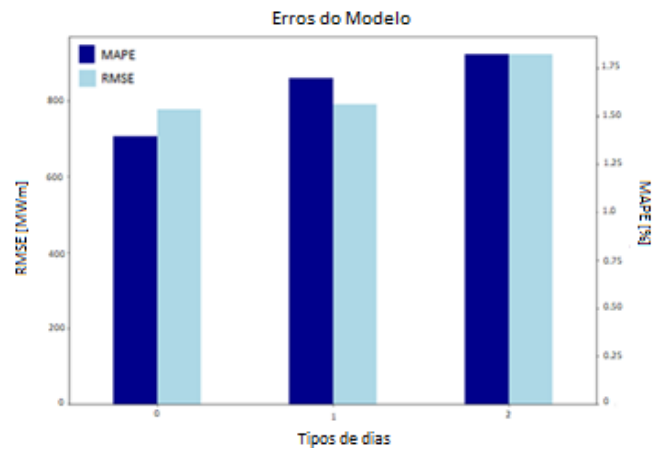


### D. LSTM (Caso 2)

Na fase de teste da rede neural foram gerados os três gráficos comparativos, representando os tipos de dias, entre os valores da série temporal de entrada do modelo e os valores de saída.

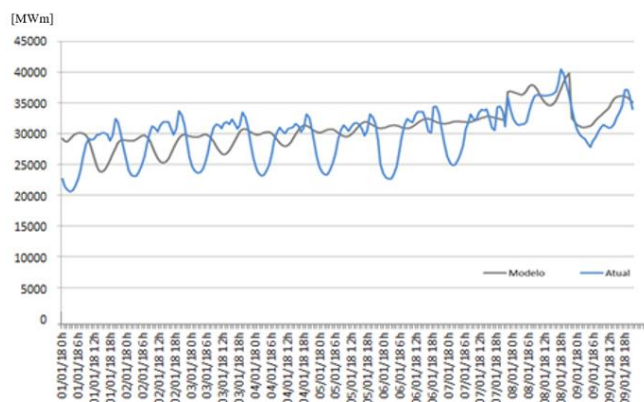
Foram calculados os erros da fase de teste (RMSE e MAPE), como apresentado na Figura 6. Conforme pode ser observado, o erro MAPE máximo obtido foi de 1,75%, apresentando ótimo desempenho da rede.

Figura 6. Erros de teste da LSTM (Caso 2)



Em seguida, foi realizada a previsão dos dados não apresentados no treinamento da rede. Na Fig. 7 tem-se o resultado obtido.

Figura 7. Comparação da previsão do modelo LSTM (Caso 2) com os dados originais da série (em MWm)



## VI. CONCLUSÃO

O presente projeto teve como objetivo a previsão de carga horária dias a frente do subsistema Sudeste/Centro-Oeste, considerando a importância desses dados para a modelagem do preço horário que será implementada no Brasil a partir de 2020. Para isso, foram utilizadas duas metodologias de previsão, o modelo linear autorregressivo ARIMA e o modelo de redes neurais artificiais LSTM e GRU.

No modelo ARIMA, foram utilizadas 24 regressões de acordo com as horas do dia, apresentando um erro MAPE na faixa de 10 a 20%, com valores acima de 15% entre 8 e 18 horas. Como proposta para trabalhos futuros, as previsões com este tipo de modelo podem tentar ser melhoradas com o uso do modelo autorregressivo integrado de médias móveis sazonal – SARIMA. Neste modelo não é necessário a diferenciação da série para torná-la estacionária já que este possui uma parte não sazonal e outra sazonal, podendo trazer melhores resultados.

Com as redes LSTM, foram feitos dois casos: um com o mesmo perfil usado no modelo ARIMA (considerando as 24 horas do dia), e o outro diferenciando os dias úteis dos sábados e dos domingos/feriados. O primeiro caso apresentou um erro MAPE máximo de 2,5%, enquanto que para o segundo o erro foi de 1,75%. Os melhores resultados obtidos com as redes LSTM são decorrentes da capacidade de tais modelos em ajustar séries de comportamento não-linear. No entanto, apesar do segundo caso ter apresentado erros menores, ao se analisar as Figuras 5 e 6, é possível observar que o primeiro caso apresentou uma previsão com perfil mais próximo ao original. Adicionalmente, apresentou-se para rede GRU os dados da série processados do mesmo modo que no caso 1, considerando as 24 horas do dia, obtendo-se um MAPE de 1,84%.

Desta forma, conclui-se que, dentre os modelos de previsão utilizados neste projeto, as redes neurais LSTM, no caso em que

foram consideradas as 24 horas, e o modelo baseado em GRU apresentaram desempenho superior ao do segundo caso, com dados organizados por tipo de dias, e ao do modelo ARIMA.

## AGRADECIMENTOS

Os autores agradecem o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), da Coordenação de Aperfeiçoamento do Ensino Superior Educação Pessoal (CAPES) e da Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

## REFERÊNCIAS

- [1] Mapeamento Organizacional do Setor Elétrico Brasileiro. Disponível em: <http://www.abradee.com.br/setor-eletrico/visao-geral-do-setor/>. Acesso em: Janeiro 2018.
- [2] Gonçalves, D.N., Rodrigues, M.A.P., Montezuma, L.R., Neri, T.F., "Gestão de Interligações do SIN por Indicadores de Desempenho", X STPC - Seminário Técnico de Proteção e Controle, 2010.
- [3] EPE - Plano Decenal de Expansão de Energia 2027. Disponível em: [http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/Documents/PDE%202027\\_aprovado\\_OFICIAL.pdf](http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/Documents/PDE%202027_aprovado_OFICIAL.pdf). Acesso em: 05 fev. 2019.
- [4] G. Box and G. Jenkins, Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day, 1970.
- [5] Haykin, S. Redes Neurais - Princípios e prática 2a Edição. Porto Alegre: Bookman, 2001.
- [6] Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; and Bengio, Y. 2014b. Learning phrase representations using rnn encoder-decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP, 1724–1734*.
- [7] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Computation*, 9(8):1735–1780, 1997.
- [8] CCEE - Visão Geral das Operações na CCEE. 2018. Disponível em: [https://www.grancursospresencial.com.br/novo/upload/VISAO\\_GERAL\\_OPERACAO\\_14\\_05\\_2010\\_20100514154651.pdf](https://www.grancursospresencial.com.br/novo/upload/VISAO_GERAL_OPERACAO_14_05_2010_20100514154651.pdf). Acesso em: 02 dez. 2018.
- [9] Morettin, P.A.; Tolo, C. M. C. Análise de séries Temporais. 3. ed. São Paulo: Blucher, 2018.
- [10] T. Ergen and S. S. Kozat, "Online training of LSTM networks in distributed systems for variable length data sequences," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 5159–5165, Oct. 2018.
- [11] Felix A. Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. *Neural Computation*, 12:2451–2471, 1999
- [12] L. Jing et al., "Gated orthogonal recurrent units: On learning to forget," 2017, arXiv:1706.02761. [Online]. Available: <https://arxiv.org/abs/1706.02761>
- [13] M. Bai, B. Zhang, and J. Gao, "Tensorial recurrent neural networks for longitudinal data analysis," 2017, arXiv:1708.00185. [Online]. Available: <https://arxiv.org/abs/1708.00185>
- [14] Pires, Camilla. Metodologia para previsão de carga e geração no horizonte de curtíssimo prazo. 98 p. Dissertação (Pós-graduação em Engenharia Elétrica) - Universidade Federal de Santa Maria, 2016.