



Uso de Inteligência Artificial (IA) na Definição de Limites de Continuidade do Fornecimento de Energia Elétrica

Tema: Qualidade de Energia

Autores: Alejandro Caamaño Antelo Pose - apose@siglasul.com.br, Rafael Ferraz Catramby - rca-tramby@siglasul.com.br, Carlos Frederico Meschini Almeida - cfmalmeida@usp.br, Ivo Ordonha Cyrillo - ivocyrillo@usp.br, Lindemberg Reis - Ireis@abradee.org.br, Sérgio Kinya Fugimoto - projeto.continuidade@abradee.org.br e Rodolfo Ribeiro de Oliveira - rribeiro@cpfl.com.br

Co-Autores: -

Empresa: Siglasul Consultoria Ltda

Resumo

Este artigo visa apresentar uns dos vários resultados obtidos a partir da execução do projeto de Pesquisa e Desenvolvimento (P&D) PD-00063-3078/2022, com gestão da CPFL Energia e financiado por recursos do P&D da Agência Nacional de Energia Elétrica (ANEEL), contando com a participação de 17 distribuidoras cooperadas e 3 entidades executoras (universidade, consultoria e pesquisa). O projeto abordou diretamente o método de definição de limites dos indicadores de continuidade coletivos relativos à duração e à frequência equivalente de interrupção por unidade consumidora, DEC e FEC, respectivamente. Os fundamentos para elaboração de futura proposta metodológica de limites de continuidade do serviço têm como base em extensa revisão bibliográfica, um seminário internacional, um *position* paper, pesquisa direta com os consumidores e simulação de metodologias. Este artigo apresenta uma metodologia alternativa para o cálculo dos alvos regulatórios de DEC e FEC dos conjuntos de unidades consumidoras utilizando Redes Neurais Artificiais (RNA), uma técnica de Inteligência Artificial (IA). Nesta abordagem, aborda-se a técnica de aprendizado de máquina (do inglês, *machine learning*) não-paramétrica capaz de capturar interrelações complexas entre variáveis de entrada (atributos explicativos) e de saída (indicadores de qualidade).

1. Introdução

Em fevereiro de 2022, a SIGLASUL Consultoria LTDA e a Escola Politécnica da Universidade de São Paulo (USP) – deram início, sob a coordenação geral do Instituto ABRADÉE de Energia (iABRADÉE) e da CPFL Energia, como entidades executoras do projeto de Pesquisa e Desenvolvimento (P&D) de código PA3078 e intitulado “Revisão e Aprimoramento das Metodologias de Definição dos Limites Regulatórios para Duração Equivalente de Interrupção por Unidade Consumidora (DEC) e Frequência Equivalente de Interrupção por Unidade Consumidora (FEC)”.

O projeto de P&D, com duração total de 30 meses, objetivou a confecção de novas metodologias com vistas ao estabelecimento de limites coletivos de indicadores de continuidade do serviço. Realizou-se propostas

de novas metodologias que resultassem na definição de limites regulatórios condizentes com as distintas realidades observadas pelas distribuidoras em suas respectivas áreas de concessão.

Nesse artigo, aborda-se principalmente a construção, o treinamento e a aplicação de uma Rede Neural Artificial (RNA) para o cálculo dos alvos regulatórios (limites objetivos) de DEC e FEC dos conjuntos de unidades consumidoras.

2. Desenvolvimento

METODOLOGIA VIGENTE

A ANEEL publicou sua primeira versão da metodologia de definição dos limites coletivos para os indicadores da qualidade do serviço na Resolução Normativa nº 024/2000 (REN 024/00). Ela se baseia em modelos de comparação de desempenho usando técnicas estatísticas de formação de agrupamentos (clusters) dos conjuntos de unidades consumidoras. A premissa básica da metodologia, que ainda é adotada, é a de que conjuntos considerados semelhantes a partir da comparação de seus atributos descritores (variáveis explicativas), devem apresentar desempenhos equivalentes.

Em 2011, a ANEEL atualizou a metodologia, incorporando a ela importantes avanços ainda em uso. Entre os quais, elencam-se: (i) alteração do critério de formação de conjuntos, de geográfico para geoeletrico; (ii) revisão dos atributos utilizados no cálculo; (iii) mudança do método de *clusterização* de *k-means* para o dinâmico; e (iv) separação dos modelos de cálculo dos limites dos indicadores DEC e FEC, que antes eram unificados.

A terceira e atual versão da metodologia consta na Nota Técnica nº 102/2014-SRD/ANEEL (NT 102/14), a partir da publicação da Resolução Normativa nº 641/2014 (REN 641/14), que a incorporou ao Módulo 8 das Regras e Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST). As alterações trazidas são em três pontos: (i) atributos selecionados para *clusterização* dos conjuntos; (ii) trajetória dos limites; e (iii) tratamento a situações peculiares.

Em síntese, elencam-se as principais etapas da metodologia de definição dos limites coletivos desde 2014:

a) Seleção dos atributos dos conjuntos a serem utilizados como descritores de seus desempenhos, por meio de análise de correlações e do algoritmo *Stepwise*. Etapa realizada uma única vez, quando ANEEL verifica sua necessidade. A partir daí, apenas há atualizações esporádicas dos valores dos atributos escolhidos, que atualmente são 6 para cada indicador.

b) Padronização dos atributos dos conjuntos pelo método *Z-score*, de modo que, em cada atributo, os dados possuam média igual a 0 e desvio padrão igual a 1. Busca-se que todos tenham o mesmo peso na comparação, independente da ordem de grandeza de suas unidades de medida.

c) **Aplicação do método dinâmico para determinar os conjuntos semelhantes.** Para cada conjunto em análise, forma-se um agrupamento dos mais comparáveis a ele, a depender de 3 critérios: (i) grau de heterogeneidade percentual permitida (20%), número desejável (100) e número mínimo (50) de conjuntos que devem ser selecionados como semelhantes.

d) Cálculo do limite objetivo, a partir do percentil dos desempenhos médios de 3 anos dos conjuntos semelhantes, com ordenação crescente. Para conjunto com rede aérea pertencente ao SIN o percentil é 20, já para aquele com rede aérea fora do SIN (isolados) ou com mais de 50% de rede subterrânea, adota-se o percentil 50. O valor do limite objetivo é um número inteiro e com arredondado para cima.

- e) Definição da trajetória de limites, a partir do valor atual até o objetivo. Caso o limite objetivo seja superior ao atual, mantém-se este último constante todos os anos. Caso contrário, aplica-se uma trajetória linear de redução. Ressalta-se que todos os limites intermediários são números inteiros e com arredondamento simples.
- f) **Tratamento *ex-post*** ao método a fim de considerar situações particulares dos conjuntos, tais como a **heterogeneidade elevada no agrupamento** (método Score ANI) e as trajetórias de redução muito intensas (acima de 8 horas para o DEC e 5 interrupções para o FEC).
- g) Seleção dos primeiros anos da trajetória considerando apenas o período da revisão tarifária de cada distribuidora, variando entre 3 e 5 anos.
- h) Divulgação da proposta ANEEL da trajetória de limites de cada conjunto e do global da distribuidora em análise, por meio de consulta à concessionária de energia elétrica e à sociedade como um todo (consumidores individualizados, conselho de consumidores, empresas, reguladores estaduais, entre outros).
- i) Recolhimento das contribuições dirigidas tanto pela distribuidora em análise quanto pela sociedade.
- j) Análise e alteração (caso necessário) das trajetórias de limites calculadas pela ANEEL com base nas especificidades dos conjuntos e nas justificativas contidas nas contribuições.
- k) Divulgação dos resultados finais.

METODOLOGIA ESTUDADA

No projeto de P&D PA3078, as executoras estudaram e desenvolveram metodologias alternativas para várias das etapas descritas anteriormente. Entre elas, destaca-se uma abordagem mais disruptiva em relação ao método atual, por ser empregada uma técnica de Inteligência Artificial (IA) para a definição dos limites coletivos da qualidade do serviço (DEC e FEC), foco deste artigo.

Assim, o **item “c”** da listagem – indicado em **negrito** – é aquele mais impactado por este estudo, uma vez que se propõe a alteração do método de clusterização dinâmica pelo de Redes Neurais Artificiais (RNA), cuja explicação geral encontra-se na sequência. Ademais, desconsidera-se tanto os critérios de aplicação (grau de heterogeneidade, número desejável e mínimo de conjuntos), quanto o tratamento *ex-post* do item “c” relacionado à elevada heterogeneidade no agrupamento, por se tratar de algo específico apenas para a aplicação da clusterização dinâmica.

REDES NEURAIS ARTIFICIAIS (RNA)

Uma Rede Neural Artificial é um modelo computacional baseado em Aprendizagem de Máquina (*Machine Learning*) inspirado no funcionamento do cérebro humano. Baseia-se em um algoritmo que emula uma rede densa interconectada de neurônios que recebem, processam e transmitem estímulos elétricos que, *grosso modo*, podem ser entendidos como pacotes informacionais. Uma RNA pode ser utilizada para ampla variedade de tarefas, como reconhecimento de padrões em um agrupamento de dados, classificação de informações, processamento de linguagem natural ou a tomada de decisões baseada em um grupo de estímulos recebidos.

Este método pode ser qualificado como um “*otimizador (ou aproximador) universal*”, pois é capaz de aprender e de se ajustar de forma autônoma (sem a necessidade de se atribuir uma função que correlaciona entradas e saídas), permitindo que ela se adapte a diferentes tipos de problemas (tendo, portanto, aplicação “universal”) e encontre soluções eficientes (“ótimas”) para a minimizar o erro entre a saída desejada e a saída gerada pelo modelo (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Dessa forma, uma RNA é capaz de encontrar soluções para diferentes tipos de problemas, independentemente de sua complexidade, desde que haja um conjunto de dados de treinamento suficientemente grande, representativo e diversificado.

Os modelos de RNA possuem a vantagem de serem mais generalistas, por não dependerem de formas funcionais definidas *a priori*, e, portanto, poderem calcular relações mais complexas entre os atributos de entrada e os atributos de saída.

Ademais, este tipo de modelo possui calibração flexível, admitindo a expansão da base de dados de referência para treinamento ao longo do tempo de modo a ajustar o modelo a mudanças no ambiente.

A Rede Neural Artificial também apresenta desafios a serem considerados. Primeiro, a relação matemática entre entradas e saídas não é facilmente interpretável (comumente referido como “modelo *Black Box*”), sobretudo em redes profundas com muitas camadas ocultas (*deep network*). O objetivo da Rede Neural, ao contrário de modelos econométricos, não é definir relações estatísticas entre as variáveis, mas reproduzir, da melhor maneira possível, os valores de saída. Segundo, a RNA requer um grande conjunto de dados para ser treinada, caso contrário, a otimização alcançada pode ser pouco eficaz e incapaz de aderir a dados novos. Terceiro, os resultados obtidos podem ser sensíveis à arquitetura da rede. Aquelas formadas por poucas camadas ocultas são rapidamente treináveis, mas podem gerar resultados pouco satisfatórios, enquanto as redes profundas podem capturar relações mais complexas, mas exigem maior capacidade computacional.

A arquitetura da RNA é composta por várias camadas de neurônios interconectados. Cada neurônio possui pesos e vieses que são ajustados durante o treinamento da rede neural, permitindo que ela aprenda e melhore sua precisão ao realizar tarefas específicas.

As camadas de uma rede neural podem ser divididas em três tipos:

- i. **camada de entrada**, que recebe os dados de entrada;
- ii. **camadas ocultas**, que são compostas por vários neurônios que processam os dados de entrada e geram saídas intermediárias. A quantidade de camadas ocultas e de neurônios em cada camada pode variar de acordo com a complexidade do problema que a rede neural está tentando resolver; e
- iii. **camada de saída**, responsável por gerar a saída final da rede neural.

Ao receber os dados de entrada, cada neurônio da camada de entrada aplica uma função de ativação, que é responsável por “filtrar” essas entradas e decidir se o neurônio irá transmitir a informação para a camada seguinte. As funções de ativação mais utilizadas são: (i) a função linear, que varia de $-\infty$ a $+\infty$; (ii) a função linear corrigida, ou função ReLU (*Rectified Linear Unit*), que varia de 0 a $+\infty$; (iii) a função sigmóide, que varia de 0 a 1; e (iv) a função tangente hiperbólica (ou tanH), que varia de -1 a +1. A escolha do tipo de função depende do formato dos dados que estão sendo processados por cada neurônio e da quantidade de camadas ocultas necessárias para se obter uma saída satisfatória.

Cada conexão entre camadas possui um peso e, por sua vez, cada camada um viés[1] que irão ponderar as saídas da camada anterior e processá-las como entrada para a seguinte. O objetivo da RNA é encontrar os valores dos pesos e vieses que resultem na saída mais aderente ao problema proposto. Este processo de calibração, denominado treinamento, é repetido várias vezes, ajustando os pesos e vieses para minimizar o erro entre a saída desejada e a saída gerada pelo modelo. Durante o treinamento, a RNA irá bonificar os pesos que mais contribuíram para um bom resultado e penalizar os pesos que mais contribuíram para um resultado ruim.

Pode-se dizer que uma Rede Neural Artificial que cumpre satisfatoriamente seu papel é aquela que: (i) foi treinada com dados representativos, ou seja, dados que possuam a saída desejada dado um conjunto de entrada; (ii) passou por uma rotina iterativa de treinamento longa o suficiente para a correta calibração dos pesos; e (iii) possua bom desempenho face a um conjunto de dados representativos inteiramente novos, isto é, que não compunham a base de treinamento.

O modelo RNA requer um conjunto de dados que possua tanto as variáveis de entrada quanto as de saída observadas e cuja relação matemática desconhecida se pretende simular com o modelo pós-treino. Portanto, as observações da base de referência ideal devem possuir três características desejáveis: (i) serem

numerosas; (ii) serem representativas do fenômeno a ser reproduzido; e (iii) serem diversificadas, a fim de abranger muitos cenários possíveis.

Ademais, deseja-se que os resultados do modelo possam ser escaláveis a novas situações. Por esse motivo, o modelo RNA precisa passar por duas etapas antes da sua aplicação, a saber, o treinamento e a avaliação de desempenho (comumente referido como “Teste”). Esta divisão é feita por meio de um sorteio aleatório.

A divisão de uma base de dados para a Rede Neural entre treinamento e teste é uma prática comum na construção de modelos de aprendizagem de máquina. A proporção da divisão depende da quantidade de observações disponíveis na base. Caso seu número não seja uma restrição significativa, é habitual utilizar a razão[2] 40:60, isto é, 40% da base de referência para treinamento e 60% para avaliação do modelo.

A cada rodada o modelo passa por um ajuste de pesos e vieses – o treinamento – e por um procedimento de avaliação, em que se compara os resultados obtidos na saída com os apurados na base de dados, a fim de minimizar o erro. A forma de avaliação pode ser feita por meio de uma função de perdas, entre as mais comuns estão a soma do erro quadrático (SSE ou *Sum of Squared Error*), o erro quadrático médio (MSE ou *Mean Squared Error*), a aderência (ou *fitness*, que calcula o erro médio percentual).

O algoritmo de otimização mais utilizado para este fim é o *Gradient Descent* (ou Gradiente Descendente). Este consiste em utilizar o valor da função de perdas para calcular um gradiente, ou seja, a direção da mudança de valores, dos pesos e vieses, a fim de ajustá-los para gradualmente minimizar o erro calculado, num processo chamado *Backpropagation*.

Em síntese, o algoritmo de otimização pode ser descrito pelas seguintes etapas:

1. Inicialmente, são atribuídos valores aleatórios aos pesos e vieses.
2. O modelo é exposto a uma base de dados de treinamento que contém as variáveis de entrada e saída e, a partir do processamento dessas informações em suas camadas, gera uma saída.
3. O erro entre a saída gerada pelo modelo e a saída desejada é calculado por meio de uma função de perda.
4. O gradiente dos pesos e vieses é calculado com base na função de perda. Algebricamente, ele indica a direção e intensidade com que cada peso e viés será corrigido na próxima iteração.
5. Pesos e vieses são ajustados de forma a minimizar o erro, utilizando o gradiente calculado (etapa de *backpropagation*).

6. Repete-se o processo a partir do passo 2, ajustando os pesos e vieses de forma a minimizar o erro entre a saída desejada e a saída gerada pelo modelo.

Conforme exposto, uma RNA treinada pode ser utilizada para processar novos dados. A saída reproduz as condições que se esperaria obter caso as variáveis de entrada inseridas tivessem desempenho equivalente ao observado na base de referência usada para o treinamento da Rede Neural.

APLICANDO RNA NA METODOLOGIA DE DEFINIÇÃO DOS LIMITES COLETIVOS DA QUALIDADE DO SERVIÇO

Para a construção do modelo de RNA voltado à definição dos limites da qualidade do serviço, desenvolveu-se as seguintes etapas: (i) definição dos atributos descritivos da continuidade do fornecimento de energia elétrica; (ii) definição da arquitetura do modelo; (iii) definição da base de referência para treinamento e validação do modelo; e (iv) treinamento do modelo.

ETAPA 1: DEFINIÇÃO DOS ATRIBUTOS DESCRITIVOS DA QUALIDADE

A seleção e definição dos atributos que irão compor a camada de entrada (*input*) é uma etapa fundamental. Conforme já abordado, ao longo do projeto de P&D PA3078, foram estudados distintos temas, inclusive este, que se relaciona ao **item “a”** da listagem presente no capítulo de explicação da METODOLOGIA VIGENTE.

Para a confecção das simulações presentes neste artigo, optou-se pela utilização dos resultados obtidos no referido projeto de P&D,

ETAPA 2: DEFINIÇÃO DA ARQUITETURA DO MODELO

Na sequência, deve-se definir o formato da camada de saída. Duas alternativas podem ser seguidas para o cálculo dos limites. A primeira se baseia no cálculo dos limites dos indicadores de DEC e FEC separadamente, ou seja, no treinamento de dois modelos RNA. A segunda é o cálculo de ambos os indicadores por apenas um modelo RNA que possua dois neurônios de saída, um para DEC e outro para FEC. Uma vantagem desta última alternativa é reduzir a complexidade de cálculo e associar em um único modelo as estimativas dos limites dos indicadores. Contudo, esta última alternativa possui dois desafios para ambos os indicadores, a saber: (i) requer que as observações da base de referência sejam representativas, o que reduz consideravelmente o tamanho da base de treinamento; e (ii) exigir que os atributos de entrada sejam relevantes.

Diante dessas limitações, entende-se que a melhor opção é o treinamento de dois modelos RNA separadamente, um para o DEC e outro para o FEC.

A Rede Neural também requer a construção das camadas ocultas ou intermediárias, responsáveis pelas operações matemáticas que conectam *inputs* e *outputs*. Um modelo simplificado conterá poucas camadas intermediárias, que tem como benefício o tempo necessário para a treinamento dos pesos, embora quando aplicados a problemas complexos possam gerar saídas pouco aderentes aos valores reais. Para problemas de regressão complexos e com características pouco determinísticas, recomenda-se o uso de modelos de Redes Neurais Artificiais Profundas. Estes modelos incluem muitas camadas intermediárias, com muitos neurônios em cada uma. Esta é a opção eleita para o estudo em tela. O número ótimo de camadas não é determinado *a priori*, sendo necessário a realização de testes.

A rede pode ser dividida em 3 grupos de neurônios: entrada (em azul e à esquerda), intermediários ou ocultos (em rosa ao centro) e saída (em verde à direita). Para fins de calibração da Rede Neural, deve-se possuir uma base de dados de treinamento que contenha atributos da camada de entrada e resultados dos DEC e FEC esperados na camada de saída.

Neste projeto de P&D, empregou-se a seguinte arquitetura:

- **Camada de Entrada** com $x+2$ neurônios, sendo x a quantidade de atributos utilizados acrescidos de duas variáveis *dummy* (0 ou 1) que identificam, respectivamente, os conjuntos aéreos ou subterrâneos e os pertencentes ao SIN ou isolados;
- **Camada Intermediária** com 128 neurônios ativada por ReLU;
- **Camada Intermediária** com 64 neurônios ativada por ReLU;
- **Camada Intermediária** com 32 neurônios ativada por ReLU; e
- **Camada de Saída** com 1 neurônio ativado por ReLU.

ETAPA 3: DEFINIÇÃO DA BASE DE REFERÊNCIA PARA TREINAMENTO E VALIDAÇÃO DO MODELO

Os maiores desafios na implementação deste tipo de método são a construção da base de referência para o treinamento e a avaliação de desempenho de uma RNA são consideradas.

Assim, torna-se necessário definir um critério para classificar as observações entre eficientes e não eficientes, com o objetivo de calibrar a RNA de modo que a saída seja equiparável a uma situação de eficiência na gestão de indicadores de qualidade, reduzindo o viés dos conjuntos pouco eficientes. Uma alternativa para essa classificação é a utilização de algum método de *benchmarking* menos disruptivo para a classificação da base de dados, como o estimador de Mínimos Quadrados Ordinários (MQO) ou a Fronteira Estocástica Eficiente (do inglês, *Stochastic Frontier Analysis*, SFA).

Neste projeto de P&D PA3078, optou-se por utilizar o método da SFA, sendo um método de estimação de fronteira eficiente de produção. Este modelo, considera que os desvios em relação à fronteira estimada possuem, além do termo de ineficiências na gestão da empresa (denominado de ineficiência técnica), um componente estocástico, isto é, aleatório, não-gerenciável (COELLI et al., 2005). No caso específico da fronteira de qualidade, os fatores não-gerenciáveis (aleatórios ou exógenos) incluem aqueles que impactam os indicadores de qualidade, mas são difíceis ou impossíveis de serem gerenciados pela empresa, por exemplo, características geográficas e climáticas da área de concessão tais como relevo, vegetação, descargas elétricas, temperatura e precipitação.

Enquanto o estimador de MQO possui a forma funcional " $y_i = f(x_i, \beta) + \mu_i$ " em que μ_i indica o erro da regressão, ou seja, o desvio dos valores estimados pela função " $f(x_i, \beta)$ " em relação ao valor apurado da variável explicada " y_i ", o SFA substitui ao termo de erro " μ_i " pelos componentes " u_i ", que captura a ineficiência técnica (gerenciável) e " v_i " que captura o termo estocástico (não gerenciável). Para isto, atribui-se ao componente técnico uma distribuição de probabilidades estritamente positiva, usualmente a distribuição seminormal, normal-truncada, gama ou exponencial. Esta distribuição de probabilidades tem como objetivo garantir que não exista eficiência gerenciável negativa. Já o termo aleatório " v_i " é representado pode assumir qualquer valor real, e costuma ser modelado como um ruído branco, isto é, assume a forma de uma distribuição de probabilidade Normal com média zero. Este segundo componente, portanto, permite desvios tanto negativos (que reduzem a eficiência em relação à fronteira) quanto positivos (que aumentam a eficiência).

Ademais, no referido projeto de P&D PA3078, considera-se para a formação da base de referência apenas os conjuntos cujo termo de eficiência esteja acima do percentil 60. Neste contexto, resulta-se em 1.204 conjuntos de referência para DEC e 1.209 conjuntos de referência para FEC, o que representa, aproximadamente, 40% dos 3.061 conjuntos de unidades consumidoras avaliados.

ETAPA 4: TREINAMENTO DO MODELO

Destaca-se que, no âmbito do projeto de P&D PA3078, o treinamento da rede foi conduzido ao longo de 9.999 gerações, representando um valor significativo, uma vez que os modelos se tornam estáveis a partir de 2.999 gerações.

[1]Introduz um ruído em cada camada para a evitar que o algoritmo de otimização "trave" em uma situação de mínimo local ao invés de continuar a busca por um mínimo global (problema do *lock-in*).

[2]Também pode-se utilizar outras razões, como 30:70.

RESULTADOS OBTIDOS

Os limites objetivos (alvos regulatórios) resultantes da simulação da abordagem estudada neste artigo sobre interligência artificial por meio do uso de Redes Neurais Artificiais (RNA) foram comparados com aqueles simulados com a aplicação da atual metodologia da ANEEL. Apenas cabe resaltar que os atributos utilizados em um e outro caso são os novos obtidos pelo desenvolvimento do projeto de P&D PA3078 .

Na sequência, a partir da execução dessas duas simulações, discorre-se sobre os principais resultados numéricos e suas conclusões, tanto para o global do país quanto por empresa.

Na simulação da RNA, os limites objetivos calculados para o país como um todo registraram um valor de **9,20 horas de DEC**, representando aumento marginal de 21% em comparação à metodologia estabelecida

pela ANEEL, no qual se obteve **7,61 horas**. Já para os limites objetivos de FEC, notou-se uma elevação de 4% do valor de **4,40 interrupções** obtido pela metodologia de Redes Neurais em comparação às **4,32 interrupções** calculadas com a metodologia vigente da ANEEL.

Assim, nota-se o patamar superior dos valores obtidos com os alvos regulatórios calculados com a RNA, tanto para o DEC quanto para o FEC.

Ao se ampliar essa análise nível Brasil para cada distribuidora incorporada no estudo, mantém-se a mesma conclusão. Ressalta-se que na simulação da RNA, além dos resultados, acrescentam-se os percentuais de elevação ou de redução em relação a metodologia vigente, bem como ícones de seta para cima ou para baixo, a depender do caso.

Conforme se observa, a abordagem com as Redes Neurais Artificiais (RNA) apresenta valores superiores em quase todas as empresas no DEC e no FEC.

No DEC, 43 (86%) das 50 Distribuidoras obtiveram limites objetivos mais relaxados que os valores calculados com a metodologia da ANEEL. Outras 5 (10%) empresas obtiveram limites mais apertados, enquanto 2 (4%) concessionárias permaneceram com o mesmo valor em ambas as simulações.

No FEC, 31 (62%) das 50 Distribuidoras obtiveram limites objetivos mais relaxados que os valores calculados com a metodologia da ANEEL. Outras 15 (30%) empresas obtiveram limites mais apertados, enquanto 4 concessionárias permaneceram com o mesmo valor em ambas as simulações.

Por meio dos histogramas de DEC, nota-se que os Limites Objetivos iguais ou inferiores a 5 horas na metodologia vigente foram relaxados com a RNA em direção a faixas superiores e assim por diante. Observa-se, no RNA, uma maior concentração das faixas de 6 a 8 horas. Isto explica os efeitos observados nas visões Distribuidoras e Brasil, uma vez que estas apenas são consequência dos resultados por conjunto.

Uma explicação possível a este resultado, pode estar relacionada ao fato da Rede Neural – com suas relações entre neurônios, pesos e vieses – estar conseguindo capturar os efeitos das dificuldades (ou, até mesmo, das restrições) que muitos conjuntos com a rede prioritariamente aérea possuem para performarem abaixo do patamar de 5 horas, dado suas especificidades intrínsecas.

Já nos histogramas de FEC, verifica-se que os Limites Objetivos iguais ou inferiores a 2 interrupções na metodologia vigente foram relaxados com a RNA em direção a faixas superiores. No entanto, a cauda mais longa na vigente torna-se mais curta na RNA, mantendo a concentração dos dados em patamares de 3 a 6 interrupções. Como o alvo regulatório de um conjunto é sempre o valor médio performado por outro, esta concentração pode evidenciar uma maior homogeneidade entre os conjuntos como um todo na apuração desse indicador.

3. Conclusão

Os resultados da aplicação de Rede Neurais Artificiais (RNA), ao se cotejar com os padrões obtidos de simulação com a metodologia vigente da ANEEL, estabelece limites mais elevados em ambos os indicadores. Na simulação da RNA para o DEC, observa-se maior concentração de alvos regulatórios nas faixas de 6 a 8 horas. Já para o FEC, nota-se maior concentração de Limites Objetivos nas faixas de 3 a 6 interrupções, além de terem uma cauda menor em relação a metodologia vigente, indicando performance mais homogênea.

Conclui-se que o emprego da RNA para a definição dos limites coletivos da qualidade do serviço representa um avanço importante na modernização regulatória do setor elétrico brasileiro. Há indícios de que esta metodologia, baseada em técnica de Inteligência Artificial (IA), consiga capturar dadas especificidades que afetam a continuidade do fornecimento de energia dos conjuntos de unidades consumidoras e que não

estejam capturadas pelos atributos, permitindo maior precisão, flexibilidade e isonomia no trato regulatório entre a ANEEL e as Distribuidoras.

Por fim, entende-se que o projeto de P&D PA3078 intitulado “*Revisão e Aprimoramento das Metodologias de Definição dos Limites Regulatórios para Duração Equivalente de Interrupção por Unidade Consumidora (DEC) e Frequência Equivalente de Interrupção por Unidade Consumidora (FEC)*”, trouxe este e outros avanços de métodos e de simulações, visando o repensar de vários aspectos metodológicos da qualidade do serviço no âmbito nacional.

4. Referências bibliográficas

ANEEL. (2000). Resolução Normativa nº 024. Brasília.

ANEEL. (2014). Nota Técnica nº 102/2014. Brasília.

ANEEL. (2014). Resolução Normativa nº 641. Brasília.

ANEEL. (2021). PRODIST módulo 8. Brasília.

Dados Abertos - ANEEL. (2023). aneel.gov. Fonte: <https://dadosabertos.aneel.gov.br/dataset/indicadores-coletivos-de-continuidade-dec-e-fec>

ANEEL. (2021). Tomada de Subsídio nº 022, Processo nº 48500.000448/2021-51. Brasília.

KONONENKO, I.; KUKAR, M. Machine learning and data mining. 1a Edição ed. [s.l.] Elsevier Science & Technology, 2007.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. Elements of Statistical Learning. 2a Edição ed. [s.l.] Springer, 2009. v. 27.

COELLI, T. J. et al. An introduction to efficiency and productivity analysis. New York: Springer, 2005.