

Contribuição do GESEL para Consulta Pública 018/2019



Setembro 2019

As contribuições aqui apresentadas são baseadas em estudos desenvolvidos para o projeto de P&D da Aneel “Avaliação da Metodologia de Definição de Metas para Perdas Não Técnicas e Proposição de Aperfeiçoamentos Regulatórios”, projeto executado pelo GESEL para o Grupo CPFL, exceção feita à pergunta 12 onde os argumentos são derivados do projeto “Aspectos Regulatórios Relacionados a Perdas Não Técnicas em Áreas com Severas Restrições Operativas” realizado pelo GESEL para a Light.

1. Respostas ao questionário proposto pela ANEEL

Perdas

Apêndice “A” – Perguntas direcionadas para a Consulta Pública

- 1) O modelo regulatório utilizado a partir de 2008, no qual se inicia o estabelecimento de trajetórias decrescentes de perdas, contribuiu para a redução da média de perdas praticadas pelas concessionárias?

No âmbito do arcabouço da regulação por incentivos que incide sobre as concessionárias de distribuição de energia elétrica no Brasil, o estabelecimento de uma trajetória de perdas reconhecidas regulatoriamente busca incitar as distribuidoras a reduzirem suas perdas ao longo do tempo. Explica-se: caso as perdas sejam superiores às perdas regulatórias, isso impacta negativamente os resultados econômico-financeiros das distribuidoras. Caso as perdas reais sejam inferiores às perdas reconhecidas regulatórias, isso representa ganhos para as distribuidoras ao longo do período regulatório.

Neste contexto, é compreensível a contribuição do modelo regulatório à redução das perdas médias praticadas pelas concessionárias ao longo dos últimos anos. Entretanto, em anos recentes nota-se que um considerável número de concessionárias não está conseguindo atingir os níveis de perdas não técnicas reconhecidas regulatoriamente. Essa constatação foi feita no projeto de exame do tratamento de perdas não técnicas realizado pelo GESEL para o Grupo CPFL.

Basicamente, construiu-se um indicador que mede o percentual de distribuidoras com trajetória de redução que conseguiram reduzir os índices de perdas. O resultado obtido indica uma redução na efetividade da regulação ao longo dos ciclos, vide que a cada ciclo é reduzida a quantidade de distribuidoras com trajetória que conseguem reduzir seus índices de perdas.

Abaixo pode-se verificar o resultado mencionado:

Tabela 1 - Distribuidoras que reduziram perdas entre diferentes CRT (% do total)

Indicador	Ciclo Tarifário	Resultado
Indicador de Reação das empresas ao incentivo	2ºCRT	0,6157895
	3ºCRT	0,4149378
	4ºCRT	0,3720930

Como consequência, nota-se que as perdas vêm impactando cada vez de forma mais acentuada o resultado operacional das empresas. Dada que a regulação é por incentivo, não parece fazer sentido a hipótese que estes resultados são derivados de ineficiência das empresas, sobretudo, considerando-se o número de empresas que não estão conseguindo performance condizente com o estabelecido no âmbito regulatório.

Desta forma, é preciso se discutir a factibilidade das metas, pois o modelo pode estar estabelecendo metas incompatíveis com a capacidade de redução das empresas, especialmente no caso de empresas com níveis de perdas já próximos ao limiar de saturação do combate às perdas. De maneira mais ampla, é imperativa a discussão acerca das bases metodológicas do modelo adotado para o exame de sua factibilidade.

Cabe destacar, que a pouca efetividade da regulação em anos recentes pode ser derivada de problemas metodológicos do modelo vigente, uma vez que sua formulação atual não captura possível deterioração da situação socioeconômica do país nos ciclos. De todo modo, estas mudanças exógenas reforçam a necessidade de discussão sobre as diretrizes regulatórias de perdas não técnicas com vistas a uma possível evolução do modelo.

2) O início da crise econômica em 2015 contribuiu para o aumento das perdas

praticadas entre 2015 e 2018? Se sim, qual variável ou ferramenta estatística pode capturar esse efeito no modelo econométrico?

Sim. A partir de 2015, o Brasil passou a apresentar retração do produto interno bruto (PIB), aumento do desemprego e queda na produção industrial. Tais fatos acabaram por aumentar o endividamento e reduzir o poder de compra das famílias em todo o Brasil, sejam nas camadas sociais mais baixas ou classe média. Conseqüentemente, os índices de perdas não técnicas foram afetados.

De fato, comprova-se uma correlação significativa entre a dificuldade no atendimento às metas regulatórias de perdas e momentos de queda no PIB. Nesse sentido, uma alternativa econométrica poderia ser a inclusão de uma dimensão puramente econômica, considerando indicadores como o PIB e Desemprego. No entanto, esse é um problema multidimensional e as relações não são diretas com as perdas. Com isso, não é simples estudar essas variáveis macroeconômicas e sua transmissão para outras variáveis socioeconômicas. Por conseguinte, a estratégia de tentar representar estes efeitos por meio do modelo econométrico exige bastante cautela. Essa inclusão de dimensionalidade deve ser avaliada levando-se em conta o conceito econométrico da parcimônia, pois do contrário, existe o risco de se incorrer em um problema de heterocedasticidade.

Dada esta dificuldade, entende-se ser possível delinear duas estratégias, que vão além do modelo econométrico, para lidar com anomalias nas expectativas socioeconômicas. Estas estratégias podem ser divididas em *ex-post* e *ex-ante*. Em outras palavras, a estratégia *ex-post* considera que dada uma deterioração observada no ciclo anterior, esta seria compensada no ciclo futuro mediante novas metas compensatórias a serem praticadas no novo período tarifário. Neste caso, é importante considerar que as expectativas com base no histórico de performance das empresas precisam se ajustar para a nova realidade introduzida pela deterioração socioeconômica do período anterior, ou seja, os melhores valores obtidos no passado podem não ser mais condizentes e factíveis frente ao novo contexto de atuação. Para o outro caso, de estratégias *ex-ante*, são considerados mecanismos que ofereçam uma resposta rápida a mudanças abruptas da conjuntura econômica durante o ciclo. Dito isso, são fixadas regras prévias que esclarecem formas de flexibilização da regulação em períodos críticos. Nesse sentido, julga-se que a melhor forma de lidar com possíveis reversões macroeconômicas seria a utilização de um mecanismo de “gatilho” que ofereça ajuste automático à trajetória de redução perdas reconhecidas. O ponto de “gatilho” deve ser determinado como um valor que reflita a

anormalidade do valor frente à série histórica e os seus impactos para os índices de perdas. A variável de PIB é uma boa indicação para cumprir este papel, visto os testes estatísticos realizados que demonstraram sua correlação com as perdas e a sua periodicidade de registro.

3) Diante da atualização dos modelos com base nos dados mais recentes, previsto no Submódulo 2.6 do PRORET, concorda-se em apenas realizar a atualização proposta na Tabela 4, excluindo-se a variável relacionada à violência?

Para examinar a razoabilidade de meramente realizar a atualização proposta na Tabela 4, é preciso antes fazer ponderações sobre alguns pontos críticos da metodologia vigente. Primeiramente, sugere-se que não se remova variáveis sem antes considerar a contribuição de outras variáveis candidatas a entrarem no modelo que já foram utilizadas previamente pela Aneel. Neste sentido, cabe ressaltar que foi utilizado o Teste de Hausman para examinar a pertinência do uso do modelo com efeitos aleatórios. Dada a constatação de correlação entre os efeitos específicos e os regressores do modelo, existem indícios que devem ser adotados modelos de efeitos fixos.

Concomitantemente, é preciso ressaltar o fato dos modelos utilizados não contemplarem variáveis que considerem o esforço e a eficiência das empresas. Como o que se busca averiguar é o impacto de variáveis socioeconômicas sobre PNTs, é importante que previamente se estabeleça um modelo geral de perdas. De maneira geral, entende-se que a participação das PNT sobre o mercado esteja associada aos seguintes fatores: (i) esforço de combate às perdas; (ii) eficiência no combate às PNT; (iii) características socioeconômicas da área de concessão; (iv) fatores associados à composição e nível do consumo de energia; e (v) configuração da rede. Atualmente, o modelo da Aneel considera apenas variáveis que são representadas nos itens iii, iv e v. O viés de variável omitida surge da interação entre variáveis representadas pelos itens iii e iv, incluídas no modelo, e pelas dos itens i e ii e v contidas no erro. Caso haja interação entre essas variáveis, as estimativas do modelo e, portanto, o índice de complexidade estarão viesadas. Entende-se que existe interação entre essas variáveis no modelo e no erro, portanto, a solução mais efetiva para esse problema seria a inclusão de variáveis que representem o esforço e a eficiência da empresa no combate às perdas. Mesmo que problemas como endogeneidade possam aparecer, existem metodologias para contorná-los, como o uso de variáveis instrumentais. Sendo assim, a maior dificuldade é o acesso às medidas de tais variáveis para sua inclusão no modelo.

Essas inconsistências, apresentadas acima, afetam o resultado do modelo e devem ser tratadas caso a Agência opte por seguir com a estimação econométrica de PNTs. Portanto, a mera atualização do modelo seria insuficiente.

4) Neste sentido, caso se decida por manter os modelos apenas atualizando as variáveis no tempo, é razoável reduzir a quantidade de modelos?

Inicialmente, é preciso enfatizar que a ANEEL busca utilizar vários modelos para lidar com o problema da colinearidade entre as variáveis explicativas. Logo, a redução do número de modelos passa pela adoção de uma metodologia alternativa. Uma forma de contornar esse problema passa pela construção de uma base de dados contendo variáveis relacionadas às diversas dimensões que afetam as perdas não técnicas, relacionadas a questões socioeconômicas, assim como as características individuais das distribuidoras, de suas áreas de concessão e de seus consumidores. Com uma base robusta, a ANEEL poderia utilizar métodos de seleção de variáveis que auxiliem na melhor formatação do modelo através da seleção das variáveis mais relevantes.

Para lidar com essa base de dados, recomenda-se o uso do método de Regressão LASSO (*Least Absolute Shrinkage and Selection Operator*). Esse modelo consiste em um método de seleção de variáveis que visa minimizar o erro quadrático através da penalização dos coeficientes com alto grau de correlação entre si. Seu uso nesse estudo se faz importante especialmente pela necessidade de redução do número de variáveis explicativas, o que colabora para a interpretação dos resultados por meio da seleção de variáveis que mais contribuem para a explicação das PNTs nas diferentes áreas de concessão.

O número de variáveis incluídas no modelo aumenta sua complexidade. Quanto maior o número de variáveis incluídas, menor tende a ser seu viés e maior sua variância, levando a problemas como o *overfitting*¹. Dois métodos poderão ser utilizados para mitigação desse problema. O primeiro é a redução da complexidade do modelo, o segundo é a regularização do mesmo. Na regularização, os coeficientes são reduzidos, podendo chegar até a zero no caso dos modelos LASSO. Esse método é denominado *feature selection* e permite a remoção de algumas variáveis explicativas. A fórmula matemática por trás da regressão LASSO é a descrita pela Equação 1:

¹ É um termo usado em estatística para descrever quando um modelo estatístico se ajusta muito bem ao conjunto de dados anteriormente observado, mas se mostra ineficaz para prever novos resultados.

$$\hat{\beta}_{lasso} = \operatorname{argmin} \frac{1}{2n} \|y - X\beta\|_2^2 + \lambda|\beta|_1 \quad [1]$$

Onde:

- n representa o número de observações;
- y representa o valor observado da variável dependente;
- X representa as variáveis independentes;
- β representa o valor estimado dos coeficientes;
- λ representa o coeficiente de penalização.

É importante destacar que as configurações baseadas no modelo econométrico, devido aos problemas já mencionados na proposta 3, incorrem no problema de entrada e saída de séries no tempo, que podem afetar de forma significativa o ordenamento do *ranking* de complexidade e, com isso, a meta das empresas. Isso gera um risco regulatório associado à inconstância do índice de complexidade atribuídos às distribuidoras entre diferentes períodos de Revisão Tarifária.

5) Caso se opte por rever o modelo e mesmo considerando sua limitação, é razoável manter a premissa de estimação por efeitos aleatórios? Senão, como contornar esta limitação sem gerar os efeitos colaterais citados?

A opção pelo modelo de efeitos aleatórios permitiria o isolamento de efeitos relacionados à área de concessão que não variam no tempo. Entretanto, é importante se averiguar a existência de correlação entre efeitos individuais e os regressores do modelo. Apenas caso não haja correlação o modelo de efeitos aleatórios é indicado. No caso do modelo utilizado atualmente pela ANEEL, a existência dessa correlação fica evidenciada pelo Teste de Hausman, descrito em Hausman (1978), devendo-se optar pelo modelo de efeitos fixos, sob a pena de se obter estimadores inconsistentes. A opção pelo modelo de efeitos fixos levaria à necessidade adicional de se considerarem apenas os parâmetros estimados relacionados a variáveis socioeconômicas na construção dos índices de complexidade. Isso porque os efeitos fixos não são atribuídos inteiramente a aspectos não gerenciáveis por parte das distribuidoras. O mesmo deveria ser aplicado para os coeficientes de variáveis de eficiência e esforço, que não deveriam ser utilizados na formação do índice de complexidade.

Possivelmente, o efeito colateral relacionado à mudança do coeficiente da variável “subnormal” para negativo se dá devido a correlações com variáveis que não estão no modelo ou entre os

próprios regressores, o que afeta a estimação dos parâmetros e pode inverter seu sinal. No caso de impacto pelas variáveis omitidas, a solução deveria passar pela análise da inclusão de novas variáveis explicativas. No caso de problemas de colinearidade, para um número não muito pequeno de amostras, é possível avaliar combinações de coeficientes dentro de dado intervalo de confiança que são indiferentes a qualidade do ajuste dos dados (ex. *Robust Conjecture*).

6) Há sugestões de aprimoramentos para o critério de porte definido no Submódulo 2.6 do Proret?

Recomenda-se o uso de uma metodologia de clusterização associada ao tamanho. Essa metodologia busca identificar grupos de distribuidoras com características semelhantes dentro de determinada dimensão (ex. tamanho do mercado).

Para isso, iria se identificar um conjunto de variáveis de tamanho às quais se acredita possuírem correlação com PNT. Posteriormente, a configuração dos grupos que maximiza as similaridades intragrupo e minimiza as similaridades intergrupo seria utilizada para um número determinado de grupos. Isso contribuiria para uma melhor configuração dos critérios de porte.

7) Para obter o ponto de partida, deve-se manter a média dos últimos quatro anos ou um intervalo menor?

A adoção de um período maior é importante para suavizar os possíveis ajustes intertemporais ou eventos atípicos no decorrer dos anos da apuração. Além disso, o uso da média para o intervalo completo do ciclo anterior garante que o incentivo seja uniforme no tempo. No caso contrário (uso de intervalo menor), poderia haver uma sinalização de desincentivo ao combate às perdas nos anos de apuração. Afinal, as empresas poderiam intensificar o combate as perdas nos anos não cobertos pela apuração para garantir ganhos extraordinários que compensem um relaxamento no combate às perdas nos anos de apuração, visando obter um ponto de partida dado por uma média superior.

8) Qual a melhor forma para determinar o valor de perdas referente às melhores práticas das empresas de grande e pequeno porte? Quartis, decis? Outra métrica?

Em uma análise teórica utilizando o arcabouço microeconômico, o nível eficiente de perdas pode ser exemplificado pelo ponto onde o custo marginal de combate às perdas é igual à receita marginal auferida com a redução de perdas. Este ponto deveria ser reconhecido individualmente para cada empresa, uma vez que se espera que cada empresa tenha uma curva de custo marginal particular. Entretanto, tal abordagem se torna inviável diante da falta de dados específicos sobre os custos de combate às perdas e da assimetria de informação inerente a relação concessionária-regulador.

Como alternativa, primeiramente, é importante salientar que a determinação de limites que simbolizem as melhores práticas das distribuidoras deve ponderar a distância em que as distribuidoras estão desses valores e o cenário em que estão inseridas. A métrica utilizada na metodologia atual considerou que uma vez que algumas distribuidoras em momentos distintos conseguiram atingir níveis baixos de perdas, esses valores também poderiam ser alcançados pelas demais distribuidoras. No entanto, devido região em que estão inseridas, algumas distribuidoras terão mais facilidade em manter esses resultados e não podem ser utilizadas como referências.

Alternativas para determinação desses limites deverão ser analisadas quando da formação dos novos clusters e poderão ser estruturadas de uma combinação entre a adoção de percentis e a análise da distância das distribuidoras do grupo desse valor.

9) O ano de referência para a obtenção da meta deveria ser alterada para um média?

Essa média deveria ser a mesma do ponto de partida?

De acordo com a regra definida no Submódulo 2.6 do PRORET, a meta calculada a partir de cada ranking é obtida por meio de uma ponderação que considera as perdas praticadas pela concessionária e pelo seu *benchmark*, além da incerteza na comparação entre as duas empresas quanto às posições no ranking.

Tendo em vista as inconsistências da metodologia e os erros que estão sendo repassados à matriz de probabilidade e ao *ranking*, se limitar a rever apenas o período de referência não garantirá maior efetividade do modelo.

Sendo assim, sabendo que as distribuidoras apresentam capacidades de redução diferenciadas entre si e que algumas estão próximas do limite de saturação, a imposição de

metas baseadas no comportamento histórico de um grupo semelhante pode ser mais efetiva do que quando realizado com base no resultado de uma única empresa.

Para isso, será apresentada na Questão **Erro! Fonte de referência não encontrada.** uma proposta que reduz a instabilidade da metodologia, por meio de uma comparação menos vulnerável às incertezas dos dados de entrada, e que tende a definir metas mais aderentes a capacidade de cada empresa, conforme a similaridade de seus pares.

10) Há sugestões de aperfeiçoamentos para os limites de redução, ou seja, os parâmetros descritos acima estão adequados?

Sugere-se que os limites de redução sejam calculados com base nos estudos de variações históricas em perdas praticadas pelas distribuidoras. Uma possibilidade seria estudar a distribuição de probabilidade dessas variações com o objetivo de se determinar a factibilidade do atingimento de determinada trajetória de redução. Trajetórias baseadas em reduções anuais raras poderiam ser descartadas, buscando essa factibilidade.

Outra possibilidade seria estimar qual seria a curva de saturação de perdas. Essa abordagem demandaria uma segmentação de empresas que favorecesse sua comparabilidade. A partir dessa segmentação, seria determinada uma curva que traçasse a trajetória de perdas (variação anual de perdas vs. nível de perdas). Variações anuais para um dado nível de perdas que se distanciassem até determinado ponto dessa trajetória média poderiam ser consideradas limites de redução.

11) Essa transformação das perdas não técnicas medidas pela faturada deve estar em conformidade com o período definido para o ponto de partida e a meta?

Para transformação do resultado da base medida para faturada, a ANEEL deverá considerar o mesmo período utilizado para determinação do ponto de partida, uma vez que ele é a principal referência do modelo. Além disso, a utilização de 4 anos garantirá com que variações atípicas não sejam capturadas.

12) É razoável reabrir a discussão das variáveis socioeconômicas e do modelo econométrico antes da publicação do Censo de 2020 e incluir variáveis que melhor

retratam cada área de concessão?

A opção pelo uso do modelo econométrico para uso nas definições das metas entre períodos tarifários (que ocorrem entre 3 e 5 anos) não deveria se basear em séries com baixa periodicidade como o Censo. Idealmente, as séries utilizadas deveriam ter periodicidade menor que os ciclos tarifários.

Além disso, como enfatizado pelo ente regulador, a grande dificuldade do estabelecimento dos níveis adequados de reconhecimento regulatório de perdas não técnicas consiste em evitar que custos derivados da ineficiência das distribuidoras no combate às perdas não técnicas sejam repassados para as tarifas e, ao mesmo tempo, evitar que sejam imputados às distribuidoras custos originados por questões não gerenciáveis pelas distribuidoras.

Neste contexto, destaca-se a problemática das áreas com severas restrições à operação (ASRO). Tratam-se de áreas dominadas pelo poder paralelo onde a força do estado para garantia da ordem inexistente. Logo, a prestação de serviços nessas áreas acaba por ficar comprometida. No caso específico do setor elétrico, nota-se um elevado nível de furto de região nessas áreas. Portanto, é perceptível o caráter não gerenciável destas perdas não técnicas por parte das distribuidoras. Adicionalmente, embora exista relação entre ASRO e áreas subnormais, é preciso ressaltar que são fenômenos distintos. Ou seja, apesar de existir áreas de interseção, também se verifica áreas subnormais que não são ASRO e ASRO que não são áreas subnormais.

Esta é uma problemática crítica para distribuidoras do estado do Rio Janeiro, mas que vem começando a se fazer presente em concessionárias de distribuição de outros estados. Como ilustração, pode ser o mencionado o caso da CPFL Piratininga que apresenta algumas regiões na Baixa Santista com características bastante semelhantes às verificadas nas ASRO das concessionárias de distribuição fluminense.

Observa-se assim a necessidade do tratamento regulatório de perdas não técnicas no Brasil considerar essa dimensão em sua metodologia. Considerando as normas vigentes, a solução mais imediata é inserir a variável ASRO no modelo de complexidade. Neste sentido, deve ser mencionado que exercícios realizados pelo GESEL com a inserção da variável ASRO no modelo de complexidade.

Os resultados obtidos indicam a pouca efetividade desta variável quando inserida como uma variável adicional no modelo. Este resultado está associado ao fato que, apesar de ser um

problema de grande dimensão em algumas distribuidoras, o número de distribuidoras que possuem ASRO é limitado no montante total de distribuidoras. Alternativamente, buscou-se somar a variável ASRO na variável subnormal (descontando-se as áreas de interseção). Neste caso, apesar do resultado ter significância, é preciso ressaltar que a ANEEL na AP nº 55/2018 indicou que tais variáveis não devem ser somadas por terem dimensionalidades distintas.

Concomitantemente, é natural levantar-se a hipótese da utilização de outras variáveis que permitam refletir melhor a presença de ASRO em algumas concessionárias de distribuição. Porém, esta hipótese acaba por ser refutada diante à constatação que indicadores oficiais de violência não são capazes de refletirem a efetiva violência e ausência de Estado nessas localidades. Explica-se: trata-se de um estado de violência latente, a qual só se manifesta quando provocada. Logo, indicadores oficiais de violência não são capazes de embasarem o argumento que existem imensas dificuldades na prestação de serviços públicos nessas regiões.

Esta dificuldade não está restrita ao escopo do modelo de complexidade. Deve ser ressaltado que qualquer abordagem estatística/econométrica que requeira construção de base de dados oficiais terá o mesmo problema. Desta forma, sugere-se que o tratamento de ASRO seja feito de uma forma separada. Em síntese, a meta e velocidade de redução das perdas não técnicas seriam calculadas a partir da metodologia convencional e/ou a ser estabelecida para áreas onde as distribuidoras possuem capacidade gerencial, sendo reconhecidas as perdas não técnicas nas ASRO. Este reconhecimento separado é feito na Jamaica onde 80% das perdas não técnicas em áreas de risco são reconhecidas e no Panamá onde são reconhecidos 60% das perdas não técnicas de áreas onde a distribuidora não consegue atuar.

É notório que essa solução passa pela existência de metodologia robusta capaz de possibilitar a ANEEL delimitar com exatidão o que é efetivamente ASRO. Para isso, resalta-se o caráter fundamental de se implementar medição de fronteira capaz de delimitar o perímetro das áreas onde efetivamente nota-se um comportamento anômalo das perdas não técnicas. Ao mesmo tempo, dada a necessidade de se caracterizar a impossibilidade de combater perdas não técnicas nessas áreas, devem ser feitos estudos baseados em indicadores alternativos que permitam a caracterização da problemática social e a violência presente nessas regiões. Tais estudos devem ser validados juntos a instituições que fazem estudos sobre a dinâmica da violência e, quando possíveis, por órgãos de segurança pública ao serem consultados pela ANEEL.

- 13) É desejável que se mude para outras ferramentas para definir os níveis regulatórios de perdas não técnicas? Senão, em que medidas tais ferramentas podem complementar o modelo regulatório atual?

Sim, a necessidade de alteração da ferramenta para definição dos níveis regulatórios de perdas foi explanada ao longo do documento, porém ainda existe oportunidade para adicionar mecanismos que garantirão validade de aplicação da metodologia em diversos cenários.

O uso de uma metodologia de criação de *clusters* associados às variáveis socioeconômicas, poderia contribuir para que a ANEEL não dependesse de um modelo completo de perdas não técnicas sem deixar de considerar as especificidades de cada distribuidora. Para isso, o processo de definição das metas e das trajetórias deveria se dar de maneira distinta. O processo de clusterização busca identificar grupos de distribuidoras com características semelhantes dentro de determinada dimensão, como a complexidade socioeconômica ou o tamanho do mercado.

Para isso, iria se identificar um conjunto de variáveis de tamanho e socioeconômicas às quais se acredita possuem correlação com as PNT. Posteriormente, seria selecionado um grupo de variáveis não correlacionadas para formulação dos clusters. Por fim, as variáveis que menos contribuem para segmentação dos grupos seriam removidas da análise e a nova segmentação dos grupos seria calculada.

Para definição das metas e da trajetória, analisar-se-ia a evolução das PNT das empresas de acordo com cada grupo de complexidade-tamanho, capturando informações anuais de variação média de PNT e de volatilidade dessa variação. Utilizando essas informações, seria definida uma velocidade de redução para cada grupo, baseada no percentil específico das variações anuais de PNT de cada distribuidora, estipulados de acordo com uma análise por parte do regulador, que buscaria identificar os objetivos de redução de perdas para além de um período tarifário específico e a factibilidade da trajetória definida para cada grupo de distribuidoras. Por fim, a partir do ponto de partida de cada empresa, aplicar-se-ia a redução sobre o ano anterior, até que o final do ciclo da empresa, criando assim as trajetórias de redução que chegam a uma meta para o fim do período.

Essa abordagem possui o benefício de depender de um menor conjunto de observações e séries históricas e tende a refletir melhor empresas que são mais parecidas. Além disso, integraria

também de maneira mais clara qual é o objetivo de longo prazo do regulador para as perdas não técnicas.

14) Há sugestões de aperfeiçoamentos para a obtenção dos dados da energia injetada e das perdas totais das distribuidoras?

As distribuidoras no âmbito de Audiências e Consultas Públicas tem identificado inconsistências nos valores disponibilizados nas bases de dados da Agência, tanto nos valores homologados quanto nas informações reais. Esses problemas quando identificados a tempo podem ser corrigidos e refletidos no momento do processo, porém existem casos em que o mesmo não ocorre e as distribuidoras acabam por ter seus limites alterados posteriormente, o que impõe imprevisibilidade ao resultado. A formação de uma base de dados única é de fundamental importância para garantir que as análises e estudos realizados pela Agência, e até mesmo pelas distribuidoras estejam corretos.

Além disso, as informações divulgadas no site da Agência também precisam estar em conformidade para que possam ser utilizadas pelos agentes e demais interessados. Nesse sentido, é importante essa adequabilidade se dê em termos de periodicidade, uniformidade, consistência e solidez dos dados.