



## **Estratégia de Estimação para Faturamento em Medições Inteligentes**

**Juliana Christina. Carvalho de Araújo**

Departamento de Informática, PUC-Rio  
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900  
juliana@inf.puc-rio.br

**Rodrigo Flora Calili**

Programa de Pós-graduação em Metrologia, PUC-Rio  
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900  
calili@puc-rio.br

**Reinaldo Castro Souza**

Departamento de Engenharia Elétrica, PUC-Rio  
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900  
reinaldo@ele.puc-rio.br

**Weules Fernandes Correia**

Programa de Pós-graduação em Metrologia, PUC-Rio  
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900  
weules.correia@enel.com

### **RESUMO**

O entendimento do faturamento de energia dos medidores inteligentes é de grande importância para as distribuidoras de energia, contudo a legislação brasileira atual não leva em consideração a grande massa de dados gerados com o uso de medidores eletrônicos. Este trabalho consistiu em combinar métodos de imputação de dados e métodos estatísticos de previsão para realização de estimativas para faturamento na ausência do valor faturado causado por erros de comunicação (transmissão dos dados). A base de dados utilizada nesta pesquisa é composta pelos dados de consumo de energia acumulado dos clientes de baixa tensão que participaram do projeto Búzios Cidade Inteligente patrocinado pela Enel Distribuição Rio. A metodologia aqui apresentada mostrou-se eficiente e os resultados trouxeram pequenos erros de medições para a maioria dos clientes que compunham a amostra de estudo.

**PALAVRAS CHAVE. Imputação de dados, Previsão, Medição inteligente, Faturamento de energia.**

### **ABSTRACT**

The understanding of smart meters' billing is of great importance for power distribution utilities, however, current Brazilian legislation does not take into account the large mass of data generated through the use of electronic meters. This work consisted in combining methods of data imputation and forecasting statistical methods to make estimates for billing in the absence of billed value. The database used in this research is the accumulated energy consumption data of the low voltage customers that participated in the Búzios Smart City project sponsored by Enel Distributor Rio. The methodology presented here was robust and the results bring small measurement errors to most of the clients that composed the studied sample.

**KEYWORDS. Data imputation. Forecasting, Smart meter, Energy billing.**



## 1. Introdução

O Sistema de Medição Inteligente (SEM – *Smart Electricity Meter*) é um componente essencial das Redes Elétricas Inteligentes. Medidores inteligentes, além de melhorarem as funções de cobrança primárias, servem como uma rede de sensores sem fio (*Wireless Sensor Networking* - WSN) que podem, inclusive, melhorar a confiabilidade do sistema do elétrico.

A transmissão e processamento do SEM deve ser realizado de forma a evitar perdas que possam influenciar na confiabilidade e estruturação dessas informações. Essa estruturação dos dados é denominada Inteligência do Processo. Os medidores inteligentes são geralmente mais precisos que os medidores convencionais. Esses últimos medem a eletricidade apenas uma vez por período de faturamento (Zufferey, et al. 2016). No entanto, no processo de aquisição e transmissão de grande volume de informações para o processamento centralizado das medições inteligentes, os dados podem ser corrompidos ou perdidos por completo. As leituras não disponíveis ocorrem por vários motivos, por exemplo: falha na comunicação, falta de equipamentos, perda de dados, dentre outros fatores (Chen, et al. 2010). Esses dados faltantes devem ser corrigidos antes que a fatura do cliente seja gerada.

Uma infraestrutura de medição avançada (do inglês *Advanced Metering Infrastructure* - AMI) é também necessária para monitorar e analisar os padrões de demanda para cada usuário. Segundo Yang *et al.* (Yang, et al. 2014), o Sistema de medição inteligente é um dos mais críticos elementos de uma *Smart Grid*, pois este sistema tem de ser altamente confiável e seguro. No entanto as tecnologias hoje existentes, e inclusive testadas no *Smart Grid* de Búzios, não trazem como resposta uma efetividade de transmissão de dados 100%. Além disso, há paradigmas regulatórios sobre este tema e é possível constatar que as regras de faturamento atuais não são indicadas ao faturamento remoto possibilitado pela medição inteligente.

O estudo de faturamento de medidores inteligentes é de grande importância para as distribuidoras de energia, contudo a legislação brasileira atual não leva em consideração a grande massa de dados gerados com o uso de medidores eletrônicos. Atualmente, o faturamento de energia elétrica é estimado de acordo com a legislação imposta pela Resolução 414/2010 da Agência Nacional de Energia Elétrica (ANEEL 2012). A distribuidora deve realizar as leituras em intervalos que variam de um mínimo de 27 dias a um máximo de 33 dias. Quando a leitura não está disponível, ela é calculada como a média aritmética dos últimos 12 meses de faturamento.

Este trabalho consistiu, portanto, em (i) realizar a imputação de dados faltantes na base de dados decorrentes de possíveis erros de transmissão dos medidores; e (ii) realizar a previsão a quantidade de energia elétrica consumida em kWh por cliente, também decorrentes da falha da transmissão dos dados de energia elétrica. Os dados utilizados nesta pesquisa são os dados de consumo de energia acumulados dos clientes de baixa tensão de Búzios e que tiveram os medidores eletromecânicos trocados por medidores inteligentes com transmissão de dados via PLC. Estes clientes fizeram parte do projeto Cidade Inteligente Búzios, cuja distribuidora responsável é a Enel Distribuição Rio.

Neste contexto, as contribuições deste trabalho são:

- I. Mostrar que o uso de uma estratégia que combinada de imputação de dados (Interpolação Linear (LI) e *Spline* (SI)) e métodos estatísticos de previsão (Holt e Box&Jenkins) melhora significativamente a precisão do faturamento.
- II. Realizar um estudo de grande escala, com mais de 5.000 medidores, e avaliarmos os resultados em um estudo de caso real.
- III. Analisar os erros percentuais absolutos (APEs) obtidos a partir da combinação de dois métodos de imputação de dados (Interpolação Linear (LI) e *Spline* (SI)). Também comparamos o desempenho da metodologia proposta com a performance da metodologia ANEEL (ou seja, a média aritmética dos últimos 12 meses).



O programa computacional R foi utilizado para realizar a imputação de dados faltantes, a previsão e para avaliar o desempenho preditivo do modelo adotado, bem como compará-lo com a metodologia hoje adotada pelas distribuidoras para fazer a medição e o faturamento.

O restante deste trabalho é organizado da seguinte forma: A seção 2 apresenta uma visão geral da metodologia proposta e a seção 3 apresenta o método de imputação de dados utilizados. Os métodos de previsão aplicados neste trabalho são apresentados na Seção 4. A Seção 5 descreve um estudo de caso, utilizando dados reais da Enel Distribuição Rio, para avaliar a metodologia proposta. As considerações finais são apresentadas na Seção 6.

## 2. Metodologia

A Figura 1 ilustra as situações relacionadas às falhas de transmissão de dados de medição dos clientes e as tratativas adotadas para se estimar a medição para o faturamento. Imagine que a data de medição deste cliente seja o dia 17, deve-se prever três passos à frente (os dias 15, 16 e 17), objetivando se obter o dia de medição para faturamento, o dia 17. Este valor de medição estimado encontra-se em linha tracejada azul. Todavia, se houver dados faltantes na base de dados, que na figura 2-1 estão representados pela linha tracejada vermelha (nos dias 8 e 9, por exemplo), estes dados devem ser preenchidos para melhorar a acurácia dos modelos de previsão adotados. Assim, no processo metodológico adotado, quando se tem o estipulado para medição do faturamento (que varia conforme a solicitação do cliente), deve-se, primeiramente, preencher os *missing data* (dias 8 e 9) e, em seguida, utilizar algum modelo de previsão que estime o valor do consumo de energia até a data da medição para faturamento (dia 17). Obviamente, os modelos de previsão, irão também prever os dias 15 e 16, mostrados na Figura 1, mas estes dados não são o objeto principal da metodologia proposta, mas sim o dia, 17.

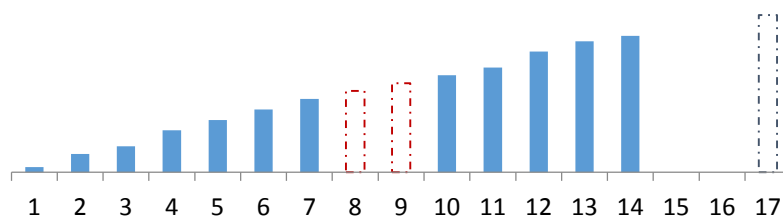


Figura 1: Ilustração da metodologia proposta

As etapas da metodologia proposta neste trabalho são resumidas na Figura 2 e detalhadas nas seções deste capítulo. A metodologia proposta para dados de medição do tipo acumulado de dia pode ser dividida em três fases: (i) limpeza da base de dados de medição (basicamente inconsistências encontradas na base de dados de medição dos clientes); (ii) imputação de dados faltantes (quando há baixa efetividade de medições); e (iii) previsão das medições dos consumos de energia.

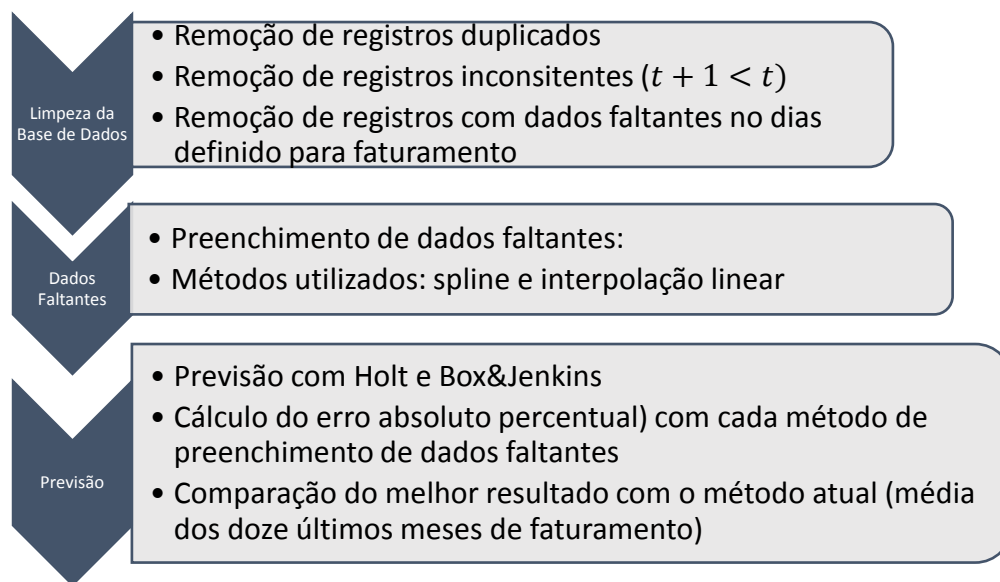


Figura 2: Etapas da metodologia proposta

### 3. Imputação de Dados Faltantes

Dados faltantes ocorrem quando, por algum motivo, os dados de medição não foram transmitidos ou não foram transmitidos de forma correta para o Centro de Medição da Distribuidora, representando, portanto, valores “nulos” na base de dados utilizada para faturamento. Estes dados devem ser preenchidos com alguma metodologia de *missing data*, com vistas a melhorar as previsões, caso nos dias da medição para faturamento, não exista a medição (também causada por algum motivo de erro de transmissão dos dados). Se no dia da medição para faturamento, os dados não chegarem corretamente no Centro de Medição da distribuidora, esta deveria utilizar alguma metodologia de previsão que estime este valor. Hoje em dia, o leitorista deve ir ao cliente e coletar os dados, caso algum problema de transmissão de dados ocorra, o que onera os custos de medição.

Métodos para lidar com dados faltantes são bem estabelecidos em Estatística (Peppanen, et al. 2016). A abordagem mais comum para lidar com as entradas de dados ausentes é ignorá-los completamente. Entretanto, a simples exclusão de dados faltantes reduz a quantidade de dados utilizáveis e pode levar a estimativas tendenciosas (Peppanen, et al. 2016). Para contornar esse problema, surgiram técnicas estatísticas que envolvem a substituição dos dados faltantes por estimativas de valores plausíveis a serem imputados aos dados faltantes. Essas técnicas têm por objetivo “completar” os bancos de dados e possibilitar a análise com todos os dados em estudo. Além disso, os preenchimentos de falhas na base de dados são de especial importância para previsão de séries temporais, uma vez que a ausência de dados completos em uma série de tempo pode comprometer a estimação e a previsão que é baseada no histórico de dados. Os métodos de interpolação linear e interpolação por *spline* cúbica foram utilizados nesse trabalho. A seguir uma breve discussão sobre os métodos citados.

#### 3.1. Interpolação Linear

A interpolação linear (LI) (Hazewinkel 2001) é um método de interpolação que se utiliza de uma função linear, para representar, por aproximação, uma suposta função  $f(x)$  que originalmente representaria as imagens de um intervalo descontínuo (ou degenerado) contido no domínio de  $f(x)$ .



### 3.2 Spline

O objetivo nos splines cúbicos (Hazewinkel 2001) é determinar um polinômio de terceiro grau para cada intervalo entre os pontos amostrados. Ou seja, a aproximação de pontos consecutivos pela Interpolação por Spline (SI) obedece à função:

$$s_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i, i = 1, \dots, n \quad (1)$$

Para dedução dos splines cúbicos, nos baseamos na observação de que, como cada par de pontos amostrados é ligado por um polinômio cúbico, a segunda derivada no interior de cada intervalo é uma reta. Logo, para  $n+1$  pontos dados, existem  $n$  intervalos e, conseqüentemente,  $4n$  constantes indeterminadas. As  $4n$  equações para determinar as  $4n$  constantes são:

- (i) O valor das splines cúbicas tem que ser igual nos pontos interiores;
- (ii) A primeira e a última spline têm que passar nos pontos finais;
- (iii) A primeira derivada nos  $n$  pontos interiores tem de ser igual;
- (iv) A segunda derivada nos pontos interiores tem de ser igual;
- (v) A segunda derivada é nula nos pontos finais (spline natural).

## 4. Métodos Estatísticos de Previsão

Como dito, a necessidade da previsão decorre do fato de não haver a medição, causada por algum problema de comunicação, nas datas consideradas para medição para faturamento de um determinado cliente. Foram utilizados dois métodos de previsão lineares, quais sejam: Holt e Box&Jenkins. Esses métodos são brevemente discutidos a seguir. Estes dois métodos foram escolhidos, pois os dados de medição de energia apresentam um comportamento linear.

### 4.1. Amortecimento Exponencial de Holt

O método de Amortecimento Exponencial de Holt (AEH) é recomendado para séries temporais que apresentam tendência (Holt 2004). Considera-se que uma série temporal  $Z_t$  a ser prevista é composta localmente da soma de nível  $\mu_t$ , tendência  $T_t$  e resíduo  $a_t$  com média zero e variância constante  $\sigma_a^2$ , isto é:

$$Z_t = \mu_t + T_t + a_t \quad (2)$$

Os valores do nível e da tendência da série serão estimados por:

$$\bar{Z}_t = \alpha Z_t + (1 - \alpha)(\bar{Z}_{t-1} + \hat{T}_{t-1}), 0 < \alpha < 1 \text{ e } t = 2, \dots, N, \quad (3)$$

$$\hat{T}_t = \beta(\bar{Z}_t - \bar{Z}_{t-1}) + (1 - \beta)\hat{T}_{t-1}, 0 < \beta < 1 \text{ e } t = 2, \dots, N \quad (4)$$

Onde  $\alpha$  e  $\beta$  são denominados constantes de amortização, tal que  $\alpha \geq \beta$ .

Assim, a previsão do valor  $Z_{t+h}$  com origem em  $t$  é dada por  $\hat{Z}_t(h) = \bar{Z}_t + h\hat{T}_t$  para todo  $h > 0$

### 4.2. Box&Jenkins

Box&Jenkins (B&J) (Box e Jenkins 1976) compreende uma classe específica de modelos estocásticos lineares (Chatfield 1975). A metodologia consiste em ajustar modelos autorregressivos integrados de média móvels a um conjunto de dados. É representado por ARIMA ( $p, d, q$ ), onde  $p$  é o número de defasagens da série,  $d$  é a ordem de integração para tornar a série estacionária e  $q$  o número de defasagens dos erros aleatórios ( $a_t$ ). O modelo ARIMA( $p, d, q$ ) pode ser descrito como:

$$\nabla^d \phi(B)Z_t = \theta(B)a_t \quad (5)$$

Onde,  $B$  é o operador de atraso;  $\nabla^d = (1 - B)^d$ ;  $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  e  $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ .



Os dados que compõem a série devem ter distribuição normal e variância constante. Para construir o modelo seguimos um algoritmo composto por quatro etapas:

(i) Identificação – A estrutura do modelo é identificada analisando-se a função de autocorrelação (FAC<sup>1</sup>) e a função de autocorrelação parcial (FACP<sup>2</sup>); busca-se identificar os valores de p, q, e d;

(ii) Estimação – Após a identificação da estrutura do modelo, os parâmetros  $\varphi_i$ 's e  $\theta_j$ 's são estimados de forma a minimizar a soma dos quadrados dos resíduos;

(iii) Verificação – As estruturas identificadas são validadas através de testes de sobrefixação. Verifica-se, ainda, se o modelo ajustado é adequado aos dados através de uma análise de resíduos (e.g. teste de Ljung-Box (Ljung e Box 1978)). Se os resíduos são autocorrelacionados, então, a dinâmica da série não é completamente explicada pelos coeficientes do modelo ajustado, devendo-se voltar à fase de identificação;

(iv) Previsão – Após a verificação, aplica-se a equação do modelo estimado para prever valores futuros da variável em estudo.

### 4.3. Métrica de Avaliação da Previsão

Para avaliar qualquer modelo de previsão, necessita-se de uma métrica para se verificar se a previsão realizada está próxima do dado real. No caso específico deste trabalho, adotou-se o APE (erro percentual absoluto). A base de dados considerada considera a leitura real do consumo dos clientes do projeto Rede Inteligente Búzios no dia 3 de outubro. No parágrafo a seguir uma descrição da métrica adotada:

$$APE = \left| \frac{Z_t - \hat{Z}_t}{Z_t} \right| \cdot 100 \quad (6)$$

## 5. Estudo de Caso - Cidade Inteligente Búzios

### 5.1. Limpeza de Dados e Preenchimento de Dados Faltantes

A metodologia proposta foi realizada considerando como horizonte de análise o período de primeiro de janeiro a três de outubro de 2016. Definiu-se o dia três de cada mês, como dia de referência para medição e faturamento. O dia três foi escolhido por apresentar baixo número de dados faltantes e proporcionar um horizonte de dados históricos grande para previsão. Especificamente, escolheu-se o dia 3 de outubro de 2016 para se testar a metodologia proposta, uma vez que este dia apresentava menor quantidade de dados faltantes.

Primeiramente, realizou-se uma limpeza na base de dados inicial a fim de eliminar inconsistências como duplicação de clientes e erros de medição, ou seja, a leitura em  $t + 1$  para o cliente  $i$  deve ser igual ou superior à leitura realizada para este mesmo cliente em  $t$ , uma vez que os dados de medição são cumulativos. Ademais, o registro de cada cliente  $i$  em todo o horizonte de tempo deve ser único. No total a base é composta por 5.199 clientes do projeto Cidade Inteligente Búzios.

A base de dados foi ainda dividida em treinamento e teste. Dessa forma, o desempenho da metodologia de previsão no dia do faturamento pode ser avaliado. A previsão cobriu o dia de faturamento (três de outubro) e a base de treinamento, o período de primeiro de janeiro a dois de outubro de 2016.

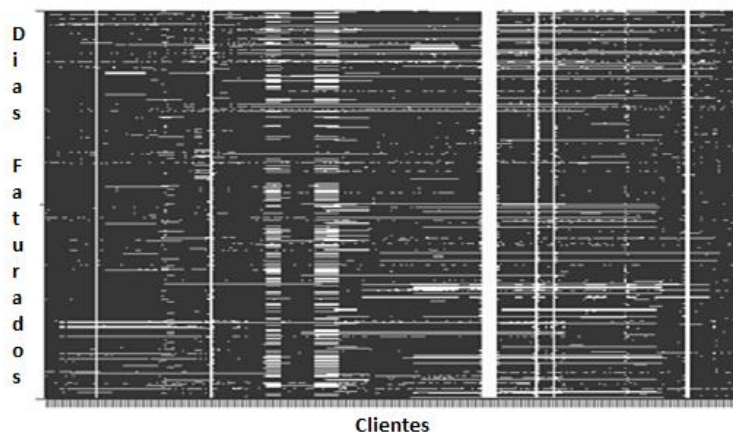
A Figura 3 demonstra o padrão dos dados faltantes na base de utilizada neste trabalho. No eixo das abscissas estão as datas de medições de primeiro de janeiro a três de outubro de 2016

<sup>1</sup> Medida padronizada da dependência linear de lag k:  $\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{Cov[Z_t, Z_{t+k}]}{\sqrt{var(Z_t) \cdot var(Z_{t+k})}}$

<sup>2</sup> Medida de dependência linear ou correlação linear entre  $Z_t$  e  $Z_{t+k}$  eliminando a dependência dos termos intermediários  $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$

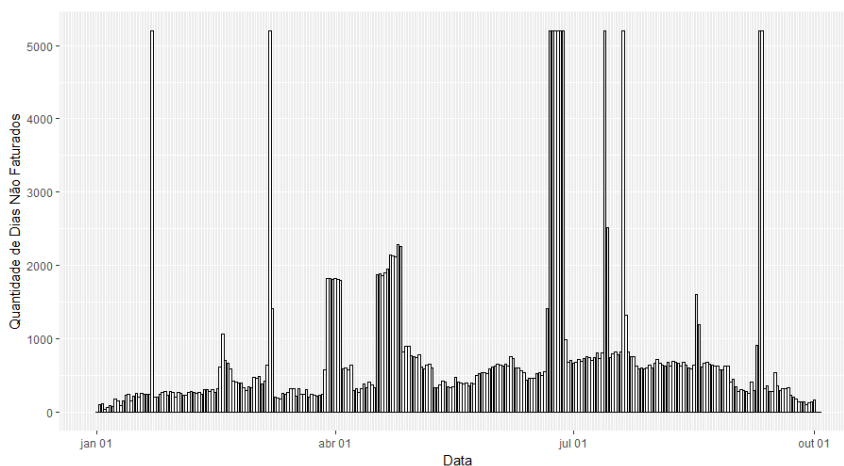


e no eixo das ordenadas os registros da amostra utilizada (5.199 clientes). Os dados registrados aparecem preenchidos em preto e os dados faltantes em cor branca ou acinzentada.



**Figura 3:** Padrões dos dados faltantes na base de dados

Dos 277 dias considerados, 12 não apresentam medição para nenhum dos 5.199 clientes, tendo ocorrido problemas de transmissão de dados, são eles: 12 de janeiro, 7 de março, 22 a 27 de junho, 13 e 20 de julho, 10 e 11 de setembro. A Figura 4 **Dados de consumo não coletado por dia faturado** ilustra a distribuição dos dados de consumo não coletados por dia faturado, considerando os 5.199 clientes presentes na base.



**Figura 4:** Dados de consumo não coletado por dia faturado.

O número máximo de dados faltantes (dias sem medição) por cliente presentes na base foi de 243 dias. Já o número máximo de dados faltantes consecutivos para um mesmo cliente foi de até 194 dias.

O preenchimento de falhas na base de dados é de especial importância para previsão de séries temporais, uma vez que a ausência de dados completos em uma série de tempo pode comprometer a estimação e a previsão que é baseada no histórico de dados.

## 5.2. Resultados da Previsão do Dia de Faturamento

Na Tabela 1, é apresentado o resumo dos resultados com as estratégias testadas 1 passo à frente, ou seja, previu-se apenas o consumo na data de medição para faturamento. Foram testados os seguintes métodos: Holt com Interpolação Linear (Holt + LI); Holt com Spline (Holt + SI); Box&Jenkins com Interpolação Linear (B&J + LI); Box&Jenkins com Spline (B&J + SI). Os

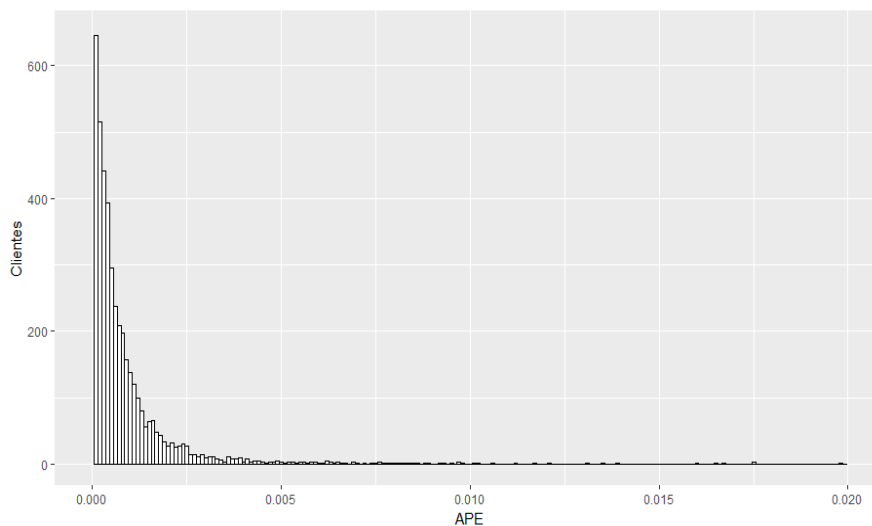


resultados com Holt e Interpolação Linear se mostram mais acurados do que a metodologia atual da ANEEL (que considera a média dos últimos 12 meses de medição para realizar o faturamento atual). Apesar de comportar-se de modo semelhante ao Holt, o método de Box&Jenkins é mais complexo e necessita de um profissional para a correta estimação dos modelos. Além disso, a estratégia de combinar Holt com Imputação Linear é simples, computacionalmente rápida e pode ser mais facilmente extrapolada para grandes conjuntos de clientes de uma distribuidora de energia.

**Tabela 1:** Resumo dos Resultados com os Métodos Propostos

Método	APE			
	Q1	Média	Mediana	Q3
Holt + LI	0.01%	0.04%	0.08%	0.09%
Holt + SI	0.01%	0.04%	0.10%	0.09%
B&J + LI	0.01%	0.04%	0.08%	0.09%
B&J + SI	0.01%	0.04%	0.10%	0.09%
ANEEL	10.68%	31.44%	13582.09%	79.41%

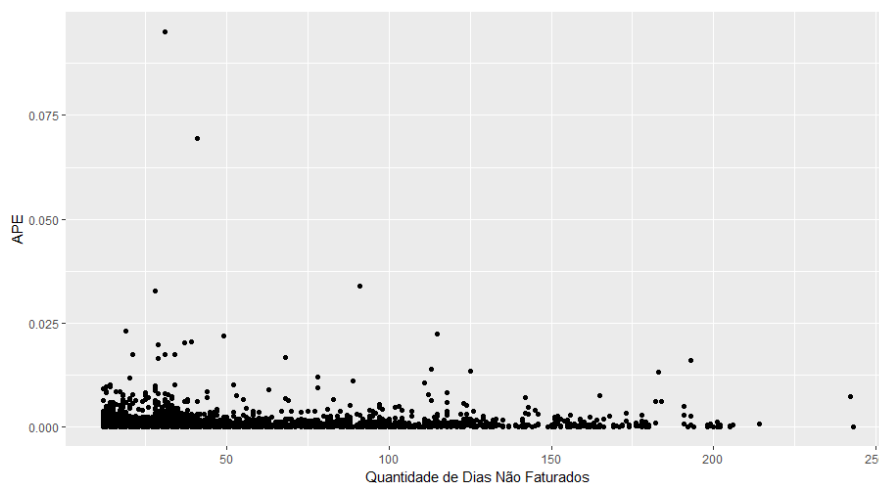
Os erros de previsão obtidos com Holt não ultrapassam os 0,05% para a grande maioria dos clientes da amostra. Esse resultado pode ser observado no histograma da Figura 5.



**Figura 5:** Histograma do Erro de Previsão no dia do Faturamento (03/out).

Não foi detectada uma relação proporcional de aumento entre o erro de previsão e a porcentagem de dados faltantes por cliente. Na Figura 6 tem-se a dispersão dos erros de previsão em relação ao percentual de dados faltantes. Pode-se constatar que a maioria dos erros de previsão se concentra em torno de zero mesmo com o aumento do percentual de dados faltantes. Como o consumo é altamente linear, a metodologia de interpolação linear demonstrou-se ser suficiente robusta.





**Figura 6:** Dispersão dos Erros de Previsão no dia do Faturamento (03/out).em relação aos Dados Faltantes.

O método aqui proposto também foi comparado com a metodologia atualmente utilizada (Resolução Normativa 414). Segundo a Resolução 414 quando não há a medição na data de medição/faturamento e nos três dias antecedentes e posteriores a esta data, deve-se considerar a média do faturamento dos últimos doze meses.

O modelo proposto mostrou-se mais preciso que a metodologia atualmente empregada (utilizando medidor eletromecânico), como pode ser visto na tabela Tabela 2, que compara o APE da metodologia atual (Resolução 414) e a metodologia proposta (Holt + Interpolação Linear 1 passo à frente).

Verificou-se que mais de 75% dos casos da amostra estudada, o método considerado neste estudo é mais ou pelo menos igualmente consistente do que o utilizado pela Resolução 414 (ANEEL, 2010), que considera a média dos últimos 12 meses para efeito de faturamento, em caso de impedimento de acesso para leitura. Sendo que para horizontes mais curtos (até seis dias), a metodologia aqui proposta se mostra superior ou pelo menos igual em mais de 94% dos casos. Mesmo no caso de 29 passos à frente de previsão (H29), quando não se possui os 29 pontos de mediação para faturamento, o modelo proposto é mais robusto que o atualmente considerado pela Resolução 414 em 77% dos clientes estudados. Os resultados são resumidos na Tabela 2 a seguir.

**Tabela 2:** Percentual de Ganho da Metodologia Proposta em Relação à Metodologia Atual na Data definida para Faturamento

<b>Metodologia Proposta</b>	<b>Último Dado Histórico</b>	<b>% Ganho do modelo Holt em Relação a Resolução 414</b>
H1	02/out	99%
H2	01/out	97%
H3	30/set	97%
H4	29/set	96%
H5	28/set	96%
H6	27/set	95%
H7	26/set	93%
H8	25/set	91%
H9	24/set	92%
H10	23/set	92%
H11	22/set	91%
H12	21/set	89%
H13	20/set	88%



<b>Metodologia Proposta</b>	<b>Último Dado Histórico</b>	<b>% Ganho do modelo Holt em Relação a Resolução 414</b>
H15	18/set	86%
H16	17/set	86%
H17	16/set	85%
H18	15/set	84%
H19	14/set	84%
H20	13/set	83%
H21	12/set	81%
H22	11/set	82%
H23	10/set	82%
H24	09/set	80%
H25	08/set	80%
H26	07/set	80%
H27	06/set	78%
H28	05/set	78%
H29	04/set	77%

## 6. Considerações finais

A metodologia aqui apresentada se mostrou robusta e os resultados trazem pequenos erros de medições para maioria dos clientes que compunham a amostra estudada. Além disso, esta metodologia proposta pode ser usada pela ANEEL em alternativa à utilizada hoje em dia (Resolução Normativa 414), que é mais adequada quando da utilização de medidores eletromecânicos. Neste caso, a proposta atual, utilizada para medição inteligente, considera o preenchimento dos dados faltantes com Interpolação Linear e utilizar a previsão com modelos Holt para preenchimento dos dados até a data de medição para faturamento dos clientes.

## 8. Agradecimentos

Este estudo é parte do projeto Búzio Cidade inteligente conduzido pela distribuidora de energia Enel em parceria com a Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio. Os autores agradecerem a Enel por sua contribuição na versão original deste artigo.

## Bibliografia

- ANEEL. “Normative Resolution N°414/2010: Updated until REN 499/2012 [In: Portuguese].” Brasília, 2012.
- Box, G.E.P., e G.M. Jenkins. Time Series Analysis: Forecasting and Control. San Francisco, California: Holden-Day Inc., 1976.
- Chatfield, C. The analysis of time series: Theory and practice. London: Chapman and Hall, 1975.
- Chen, Jiyi, Wenyan Li, Adriel Lau, Jiguo Cao, e Ke Wang. “Automated Load Curve Data Cleansing in Power Systems.” IEEE Transactions on Smart Grid, July de 2010: 213 - 221.
- Hazewinkel, Michiel. “Linear interpolation.” Em Encyclopedia of Mathematics. Springer, 2001.



- Hazewinkel, Michiel, ed. "Spline interpolation." Em *Encyclopedia of Mathematics*. Springer, 2001.
- Holt, C. C. "Forecasting Trend Seasonal by Exponentially Wheighted Averages." *International Journal of Forecasting*, 2004: 5-13.
- Ljung, G. M., e G. E. Box. "On a measure of lack of fit in time series models." *Biometrika* 65, n° 2 (1978): 297-303.
- Makridakis, S., e M. Hibon. "The M3-Competition: Results, conclusions and implications." *International Journal of Forecasting*, 2000: 451-476.
- Peppanen, Jouni, Xiaochen Zhang, Santiago Grijalva, e Matthew J. Reno. "Handling Bad or Missing Smart Meter Data through." *Innovative Smart Grid Technologies Conference (ISGT)*, IEEE Power & Energy Society. 2016.
- Yang, Z., Y.X. Chen, Y.F. Li, E. Zio, e R. Kang. "Smart electricity meter reliability prediction based on accelerated degradation testing and modeling." *International Journal of Electrical Power and Energy Systems* 56, 2014: 209-219.
- Zufferey, T., A. Ulbig, S. Koch, e G. Hug. "Forecasting of Smart Meter Time Series Based on Neural Networks." *Data Analytics for Renewable Energy Integration (DARE)*, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD). Riva del Guarda, 2016. 19-23.