

# Análise Espacial Exploratória das Perdas Não Técnicas em Sistemas de Distribuição de Energia Elétrica

Gustavo Estevo Felix, Nicole Costa da Silva, Lucas Teles Faria

Universidade Estadual Paulista “Júlio de Mesquita Filho” Câmpus Experimental de Rosana  
Rosana, São Paulo, Brasil

(E-mail: [gustavo.felix@unesp.br](mailto:gustavo.felix@unesp.br), [nicole.costa@unesp.br](mailto:nicole.costa@unesp.br), [lucas.teles@unesp.br](mailto:lucas.teles@unesp.br))

**Abstract:** Non-technical losses generate financial losses to electricity distributors and others damages to society, quality of the energy (increased interruptions, voltage fluctuation). Numerous researches have concentrated on detection these losses. In this context, a new approach is adopted to the problem of losses with a focus on incorporating geographic space. Implementation of spatial data analysis by area, contributes to produce maps that indicate the subareas of the city with higher probability of non-technical losses. The study of the geographic characteristics has been assessed attributes socioeconomic and technical: average income, consumer units rented, and load density. The study of the characteristics of the place where losses occur can provide relevant information for a better understanding of the problem. In this work, it was found that non-technical losses are associated with subareas that have a higher number of consumer units rented, lower average income and lower load density.

**Resumo:** As perdas não técnicas ou perdas comerciais geram incontáveis prejuízos financeiros às distribuidoras de energia elétrica, bem como outros prejuízos à sociedade, à qualidade da energia (aumento de interrupções no sistema, flutuação de tensão). Inúmeros trabalhos concentram-se na localização ou detecção pontual dessas perdas. Neste trabalho adota-se a incorporação do estudo do espaço geográfico ao problema das perdas não técnicas. Este estudo é pertinente por trazer informações relevantes para melhor compreensão do fenômeno. As técnicas de exploração de análise espacial de dados geográficos por áreas contribuem para a produção de mapas temáticos onde estão destacadas as subáreas “quentes” da cidade, isto é, aquelas com maior taxa de incidência de perdas não técnicas. Para o estudo das características do lugar onde ocorrem as perdas, foram avaliados aspectos socioeconômicos e técnicos das subáreas da cidade tais como: renda média, unidades consumidoras alugadas e densidade de carga. Esses atributos auxiliam na explicação da disposição das perdas no espaço geográfico do município em estudo. Neste trabalho constatou-se que as perdas não técnicas estão associadas, em sua maioria, a subáreas que possuem uma maior quantidade de unidades consumidoras alugadas, menor renda média e menor densidade de carga.

**Keywords:** Commercial Losses; Electric Power Distribution System; Non-Technical Losses; Spatial Data Analysis.

**Palavras-chaves:** Análise Espacial de Dados Geográficos; Distribuição de Energia Elétrica; Perdas Comerciais; Perdas Não Técnicas (PNTs).

## 1. INTRODUÇÃO

A energia mensurada e faturada pelas distribuidoras de energia elétrica junto às Unidades Consumidoras (UCs) sempre será inferior à energia comprada e distribuída pela mesma. Essa diferença é devido às perdas no Sistema de Distribuição de Energia Elétrica (SDEE) – perdas técnicas e Perdas Não Técnicas (PNTs).

As PNTs no sistema elétrico brasileiro ocorrem principalmente por questões sociais, comportamentais, educativas e socioeconômicas. Essas perdas podem ser classificadas conforme a sua origem em perdas técnicas ou em perdas não técnicas.

As perdas técnicas são inerentes ao transporte de energia elétrica na rede de distribuição, relacionadas à transformação de energia elétrica em energia térmica nos condutores, perdas nos núcleos dos transformadores, perdas dielétricas, entre outras (ANEEL, 2019).

As perdas não técnicas (PNTs) ou perdas comerciais são provocadas prioritariamente por fraudes, desvios e irregularidades, tais como “gatos”, “gambiaras” e não

dependem unicamente da gestão interna das distribuidoras de energia (MPF, 2019).

Segundo ANEEL (2019), em 2018, o prejuízo das companhias de energia elétrica devido às PNTs superou o montante de 3 bilhões de reais. Observa-se que esse valor é repassado ao consumidor regular por meio da elevação da tarifa de energia elétrica.

As PNTs no Brasil estão associadas a questões de cunho social, aspectos comportamentais, educativos e socioeconômicos (Penin, 2008). Neste cenário, torna-se importante a obtenção de formas de maximizar a energia disponível para a comercialização, utilizando técnicas para redução expressiva dessas perdas.

As pesquisas em PNTs, em sua maioria têm enfoque principal na detecção das UCs com perdas ou UCs irregulares (Angelos *et al.*, 2011; Cabral *et al.*, 2004; Faria, 2012; Jiang, 2002; Nagi *et al.*, 2010; Nagi *et al.*, 2011; Nizar; Dong; Wang, 2008; Nizar; Dong; Zhang, 2008; Ramos *et al.*, 2011; Ramos *et al.*, 2012; Ramos, 2014; Trevizan *et al.*, 2015).

Os trabalhos supracitados são estruturados a partir de modelos que empregam técnicas de sistemas inteligentes; como redes neurais, lógica *fuzzy*, algoritmos de aprendizado de máquinas, mineração de dados e de texto. Porém, não consideram a localização espacial das PNTs.

Segundo Messinis e Hatziargyriou (2018), há a necessidade de desenvolvimento de novos métodos para detecção das PNTs levando em conta o lugar onde essas perdas ocorrem.

Nesse contexto, a análise espacial permite a incorporação do espaço, considerando as características das subáreas onde existem as PNTs. Por meio das técnicas de análise espacial e de modelagem dinâmica, é possível determinar quais subáreas da cidade são mais vulneráveis às PNTs (Faria, 2016; Faria; Melo; Padilha-Feltrin, 2016).

Neste contexto, este trabalho incorpora o espaço geográfico ao estudo das PNTs. As técnicas de análise espacial no espaço geográfico permitem descrever a distribuição das variáveis de estudo, a identificação das possíveis subáreas com maior propensão às perdas, correlacionando as variáveis socioeconômicas e da rede elétrica.

A partir da aplicação da análise exploratória nos dados espacializados de perdas produzem-se mapas temáticos. Esses mapas são de fácil interpretação e fornecem uma informação gráfica das subáreas da cidade mais propensas às perdas. Assim, a utilização dessa ferramenta é adequada ao ambiente dinâmico de uma distribuidora de energia elétrica. Portanto, resulta em uma melhoria significativa do desempenho técnico e financeiro das distribuidoras.

Este estudo segue na Seção 2 com a apresentação da análise espacial de dados geográficos (seção 2.1), análise espacial com dados agregados por áreas (seção 2.1.1) e análise exploratória ponderada geograficamente ou *Geographically Weighted* (GW) (seção 2.2). Na seção 3 é apresentada a execução da análise exploratória em um município do interior do Estado de São Paulo com a descrição da base de dados de perdas (seção 3.1), produção de mapas temáticos (seção 3.2.1), coeficientes de correlação global (seção 3.2.2) e coeficientes de correção local GW (seção 3.2.3). Em último, são apresentadas as conclusões deste trabalho na Seção 4.

## 2. ANÁLISE ESPACIAL EM PERDAS NÃO TÉCNICAS

### 2.1 Análise Espacial de Dados Geográficos

A análise espacial permite incorporar o espaço e levantar as características das subáreas onde ocorrem as PNTs. É composta por um conjunto de ferramentas para explorar e modelar processos que se expressam através de uma distribuição no espaço denominados fenômenos geográficos. Na análise espacial, mensuram-se as propriedades e os relacionamentos, de modo a considerar a localização espacial do fenômeno em estudo de maneira explícita (Druck *et al.*, 2004).

Segundo Faria (2016) este processo é comumente utilizado em várias áreas do conhecimento como: epidemiologia, criminologia, geologia, agronomia, ecologia etc. Conforme Druck *et al.* (2004), os problemas em análise espacial são caracterizados respectivamente como: padrões pontuais, áreas com taxas agregadas e superfícies contínuas.

- ❖ *Padrões pontuais*: Fenômenos expressos por ocorrências identificadas como pontos localizados no espaço. Exemplos: localização de crimes, ocorrências de doenças, localização de espécies vegetais, etc.
- ❖ *Áreas com taxas agregadas*: Dados associados a levantamentos populacionais como censos demográficos que se referem a indivíduos localizados em pontos específicos no espaço. No entanto, por razões de confidencialidade, esses dados são agregados em unidades de análises, delimitadas por polígonos fechados como é o caso dos setores censitários municipais, regionais e federais.
- ❖ *Superfícies contínuas*: Estimadas a partir de um conjunto de amostras regular ou irregularmente distribuídas em campo. Exemplo: levantamento das características geomorfológicas do solo. Nessa categoria estão incluídos os mapas geológicos, topográficos, ecológicos, fitogeográficos e pedológicos.

#### 2.1.1 Análise Espacial com Dados Agregados por Áreas

A análise espacial por áreas é uma subdivisão da análise espacial de dados geográficos. Essas técnicas são úteis quando se lidam com eventos agregados por municípios, bairros ou setores censitários (Druck *et al.*, 2004).

As características ou variáveis espaciais envolvem atributos socioeconômicos, dados técnicos da rede de distribuição de energia elétrica e dados de inspeções em campo advindos do censo demográfico (IBGE, 2010).

Em Smith (2004), relacionam-se as PNTs com as características socioeconômicas do lugar onde elas ocorrem. Portanto, o estudo das características das subáreas onde ocorrem as perdas pode auxiliar no entendimento mais completo das mesmas. Adicionalmente, aproveitam-se as informações públicas disponíveis a partir do censo demográfico (IBGE, 2010). Os dados disponibilizados pelo IBGE estão agregados por áreas; assim o aproveitamento das informações do censo demográfico contribui para o estudo das UCs irregulares.

#### 2.2 Análise Exploratória Ponderada Geograficamente

O resumo estatístico ponderado geograficamente é uma etapa fundamental da análise exploratória que antecede a aplicação de modelos ponderados geograficamente ou *Geographically Weighted Models* (GWMs) (Gollini *et al.*, 2015).

As equações (1), (2), (4) e (5) correspondem a métricas para a análise estatística ponderada geograficamente ou *Geographically Weighted* (GW). São elas: a média GW (1) e o desvio padrão GW (2). Onde  $w_{ij}$  são os elementos da matriz de ponderação espacial entre as subáreas  $i$  e  $j$ ;  $z_i$  e  $z_j$  são respectivamente as variáveis espaciais aleatórias referentes às subáreas  $i$  e  $j$  e  $n$  é o número de pontos localizados nos centroides das subáreas ou setores censitários.

Os elementos da matriz de ponderação  $w_{ij}$  são obtidos a partir da função de *kernel* biquadrado em (3) onde  $d_{ij}$  é a distância Euclidiana entre os centroides das subáreas  $i$  e  $j$  e  $b$

é a largura de banda ou *bandwidth* que regula o nível de suavização ou decaimento da função de *kernel*.

$$m(z_i) = \frac{\sum_{j=1}^n w_{ij} z_j}{\sum_{j=1}^n w_{ij}} \quad (1)$$

$$s(z_i) = \frac{\sqrt{\sum_{j=1}^n w_{ij} (z_j - m(z_i))^2}}{\sum_{j=1}^n w_{ij}} \quad (2)$$

$$w_{ij} = \begin{cases} 1 - \left(\frac{d_{ij}}{b}\right)^4, & \text{se } |d_{ij}| < b \\ 0, & \text{caso contrário} \end{cases} \quad (3)$$

$$\rho(z_i, y_i) = \frac{c(z_i, y_i)}{s(z_i)s(y_i)} \quad (4)$$

$$c(z_i, y_i) = \frac{\sum_{j=1}^n w_{ij} [(z_j - m(z_i))(y_j - m(y_i))]}{\sum_{j=1}^n w_{ij}} \quad (5)$$

$$r_s = 1 - \frac{6 \sum d^2}{n(n+1)(n-1)} \quad (6)$$

Em (4) tem-se o coeficiente de correlação de Pearson ponderado geograficamente (GW) e; em último, em (5) tem-se a covariância GW.

Observa-se que o desvio padrão em (2) evidencia as subáreas com alta taxa de variabilidade para uma dada variável. Essas áreas podem ser melhor exploradas através de um modelo geograficamente ponderado (GW), por exemplo.

Em (4) apresenta-se a correlação GW, a qual permite realizar uma avaliação preliminar da relação não estacionária entre a variável dependente e uma variável independente de uma regressão GW, por exemplo. O ajuste da ponderação local  $w_{ij}$  é feito de modo que os pontos mais próximos exerçam maior influência. Esse ajuste é feito por meio da função de *kernel* biquadrado em (3).

Nesse estudo, além de uma análise ponderada a nível local (1)–(5) é feita também uma análise global. Nesse contexto em (6) apresenta-se o coeficiente global de correlação de Spearman. Onde  $n$  é o número de unidades de observação;  $d^2$  é o quadrado da diferença entre a posição ordinal de duas variáveis em uma mesma unidade geográfica.

Em contraposição às métricas locais, às métricas globais não levam em consideração a influência de subáreas vizinhas representadas pelos elementos da matriz de ponderação  $w_{ij}$ .

O coeficiente de Spearman indica o grau de dependência entre duas variáveis e é uma importante ferramenta para o auxílio à formulação de hipóteses sobre a dependência espacial no fenômeno em estudo e o estabelecimento de relações de causa e efeito (Ferreira, 2014).

### 3. RESULTADOS E DISCUSSÕES

#### 3.1 Descrição da Base de Dados

Neste trabalho as PNTs são representadas pela *Taxa de Perdas*, que é obtida pela razão entre o número de UCs irregulares e o número de UCs inspecionados por setor

cenitário. Buscam-se por fatores locais (variáveis explicativas ou variáveis independentes) que estejam associadas às PNTs com o objetivo de detectar quais subáreas possuem maior vulnerabilidade às mesmas. Destacam-se como possíveis atributos explicativos ao fenômeno das PNTs as variáveis socioeconômicas: renda média e o percentual de UCs alugadas por subárea.

A percepção visual das variáveis explicativas é melhor compreendida através de mapas temáticos, podendo ser avaliada mais objetivamente com o auxílio das métricas citadas na seção 2.2.

Neste contexto, os atributos avaliados para o estudo das PNTs são: *Densidade de Carga* (KVA/km<sup>2</sup>), *Renda Média* (R\$) e *UCs alugadas* (%). A *Densidade de Carga* refere-se a uma característica da rede de distribuição; as demais, são variáveis socioeconômicas oriundas do censo demográfico (IBGE, 2010). Na Tabela 1 são apresentadas as descrições dos dados de entrada empregados para a análise exploratória espacial das PNTs.

As simulações apresentadas neste estudo foram realizadas no *software R* versão 3.6.1 (R Core Team, 2015). O *R* é um *software* livre para a computação estatística e gráfica. Nele estão implementadas uma série de bibliotecas conhecidas para análise numérica, assim como o pacote *GWmodel*, empregado neste estudo.

**Tabela 1. Descrição das variáveis analisadas por subáreas.**

Variáveis Analisadas	Siglas	Descrições
<i>Renda Média</i>	<i>REN</i>	Rendimento médio mensal (com e sem renda) em reais
<i>Densidade de Carga</i>	<i>DENC</i>	Densidade de Carga por setor censitário em KVA/km <sup>2</sup>
<i>UCs Alugadas (%)</i>	<i>ALUG</i>	Percentual de domicílios particulares alugados
<i>Taxa de Perdas</i>	<i>TXPER</i>	Razão entre o número de UCs irregulares e o número de UCs inspecionados por setor censitário

#### 3.2 Aplicação da Análise Exploratória em um Município do Interior Paulista

Após a apresentação da metodologia e da descrição das variáveis de estudo é apresentada a implementação da análise exploratória nos dados especializados de PNTs em uma cidade do interior paulista.

##### 3.2.1 Mapas Temáticos

As Figs. 1–4 apresentam mapas temáticos com a distribuição espacial das variáveis derivadas da Tabela 1. Associa-se cada cor a um intervalo de classe. A segmentação por intervalos de valores para a construção das legendas dos mapas temáticos é feita com base no critério de intervalos de desvio padrão. Segundo Ferreira (2014), esta técnica é a que melhor considera a distribuição estatística das variáveis a serem mapeadas.

Neste contexto alguns comentários são feitos a partir da análise das Figs. 1-4:

- ❖ Conforme Fig. 1, a variável *Renda Média (REN)* é relevante na região sudoeste; logo, trata-se de uma região de alto poder aquisitivo. Percebe-se também que as regiões da porção leste da cidade são aquelas com menor poder aquisitivo.
- ❖ Conforme Fig. 2, a variável *Taxa de Perdas (TXPER)* apresenta uma distribuição de mediana para alta na região central da cidade.
- ❖ Conforme Fig. 3, observa-se que as *UCs Alugadas (ALU)* há uma predominância na região oeste do município em estudo.
- ❖ Conforme Fig. 4, a variável *Densidade de Carga (DENC)* apresenta uma distribuição mediana para alta em grande parte da região central para sudoeste. Percebe-se também que as regiões da porção leste da cidade são aquelas com menores densidades de carga.

### 3.2.2 Coeficiente de Correlação Global de Perdas

Além de avaliar as variáveis visualmente via mapas temáticos Figs. 1-4, é importante também avaliar a correlação global entre as variáveis em estudo. Segundo Ferreira (2014), o coeficiente de correlação de Spearman em (6) permite estimar se duas variáveis geográficas estão associadas, sem necessariamente que uma delas seja estimada a partir da outra por meio de um modelo linear. Na Tabela 2 agrupam-se os intervalos do coeficiente de correlação a fim de atribuir uma magnitude à correlação global.

Nesse contexto, na Tabela 3 apresenta-se a matriz simétrica dos coeficientes globais de correlação de Spearman  $r_s$  para as variáveis da Tabela 1.

Observa-se que um coeficiente de correlação positivo implica em uma relação diretamente proporcional entre as variáveis; um coeficiente negativo, implica em uma relação inversamente proporcional.

**Tabela 2. Intervalos do coeficiente de correlação e magnitude da correlação.**

Intervalos de $r_s$	Magnitude da Correlação
$0 < r_s \leq 0,3$	Fracamente positiva
$0,3 < r_s \leq 0,6$	Moderadamente positiva
$0,6 < r_s \leq 1$	Altamente positiva
$-0,3 \leq r_s < 0$	Fracamente negativa
$-0,6 \leq r_s < -0,3$	Moderadamente negativa
$-1 \leq r_s < -0,6$	Altamente negativa
$r_s = 0$	Não há correlação

Fonte: Adaptado de Ferreira (2014).

**Tabela 3. Matriz simétrica de coeficientes globais de correlação de Spearman entre pares de variáveis.**

Variáveis	TXPER	REN	DENC	ALUG
TXPER	1	-0,02	0,13	0,14
REN	-0,02	1	0,40	0,39
DENC	0,13	0,40	1	0,58
ALU	0,13	0,39	0,58	1

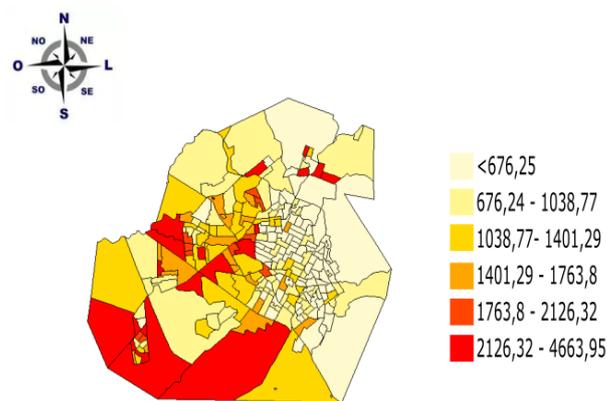


Fig. 1 Distribuição por subárea da *Renda Média* em reais (R\$).

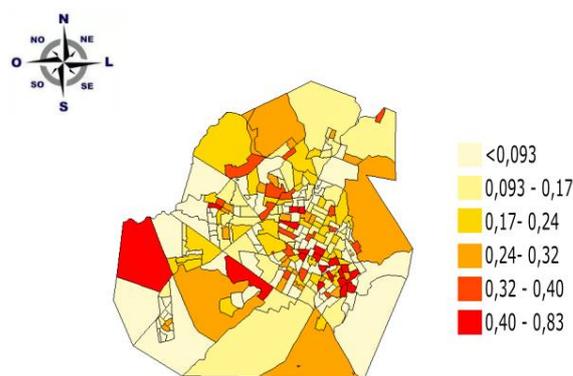


Fig. 2 Distribuição percentual da *Taxa de Perdas* por subáreas.

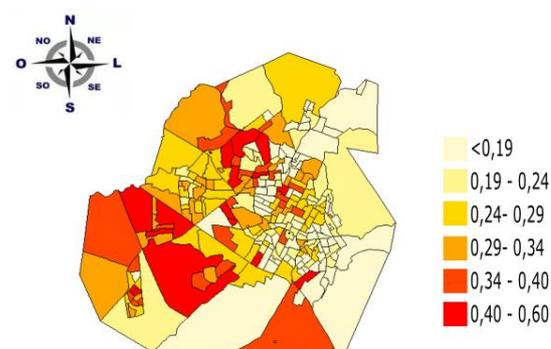


Fig. 3 Distribuição percentual das *UCs Alugadas (%)* por subáreas.

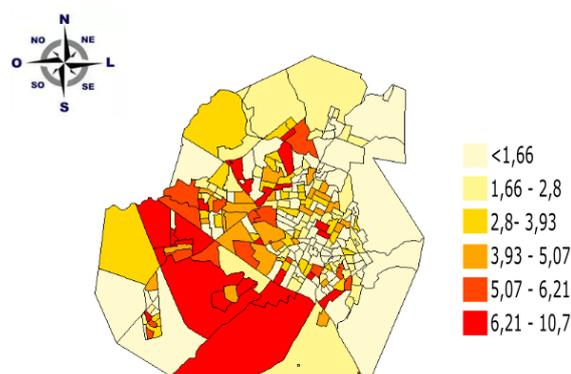


Fig. 4 Distribuição da *Densidade de Carga* MVA/km<sup>2</sup> por subáreas.

Seguem algumas análises a partir dos dados da Tabela 3:

- ❖ As variáveis *TXPER* e *ALUG* possuem correlação global fracamente positiva de 0,14. Portanto, existe uma fraca tendência de maior concentração das PNTs em subáreas com maior número de UCs alugadas.
- ❖ As variáveis *TXPER* e *DENC* possuem correlação global fracamente positiva 0,13. De fato, espera-se que a distribuição de *DENC* se relacione de forma proporcional com as PNTs. Subáreas com maior *DENC* tendem a possuir maior número de UCs; portanto, maior possibilidade de acréscimo das PNTs.
- ❖ Observa-se que a correlação entre a *TXPER* e *REN* é fracamente negativa de -0,02. De fato, espera-se que a *REN* seja inversamente proporcional às PNTs. A baixa renda implica que os indivíduos não dispõem de um poder aquisitivo suficiente para os gastos essenciais como: moradia, alimentação, saúde, água e eletricidade.

Em último, observa-se que a ausência de correlação a nível global não implica em ausência de correlação em nível local (Druck *et al.*, 2004; Gollini *et al.*, 2015).

Nesse contexto, para uma análise mais detalhada, na próxima seção é feita a análise exploratória ponderada geograficamente ou *Geographically Weighted* (GW), levando em conta a influência das subáreas vizinhas.

### 3.2.3 Correlação Local Ponderada Geograficamente

Além da apresentação em mapas temáticos da disposição espacial das variáveis em estudo (Figs.1–4) emprega-se o coeficiente de correlação GW de Pearson em (4) para avaliar a correlação local GW entre pares de variáveis conforme Figs. 5–7.

Depreende-se da Tabela 3 que as variáveis independentes (*REN*, *DENC* e *ALU*) estão fracamente correlacionadas a nível global com a variável em estudo *TXPER*. No entanto, a partir da análise das Figs. 5–7 conclui-se que a correlação local GW está disposta de forma heterogênea na cidade. Algumas subáreas exibem correlação local moderada positiva; outras, correlação GW moderada negativa.

Da Tabela 3 a correlação global entre a *TXPER* e *REN* é fracamente negativa (-0,02). No entanto, na Fig. 5 tem-se uma correlação local GW heterogênea de *TXPER* com *REN*. A correlação é moderada positiva no extremo oeste e fracamente negativa em algumas subáreas da região central.

A correlação global entre a *TXPER* e *DENC* é fracamente positiva de 0,13 – Tabela 3. No entanto, na Fig. 6 é analisada a distribuição heterogênea da correlação GW entre *TXPER* com *DENC*. A correlação local GW é moderada positiva em uma parte da região central, extremo leste e extremo sul e é fracamente negativa na porção oeste.

Em último, conforme Tabela 3, a correlação global entre a *TXPER* com *ALUG* é fracamente positiva de 0,14. Todavia, na Fig. 7 a correlação local GW entre a *TXPER* com *ALUG* encontra-se distribuída de forma heterogênea. A correlação GW é moderada positiva na região central, oeste e sul e é moderada negativa em algumas subáreas da região nordeste.

A cada subárea da cidade está associada um valor de coeficiente de correlação local GW. A Tabela 5 exhibe esses coeficientes locais de correlação GW apresentados nas Figs.

5–7 e contabilizados nos intervalos descritos na Tabela 1. Depreende-se que a maioria das subáreas possuem uma correlação fracamente negativa entre *TXPER* e *REN* (64,2%); fracamente negativa entre *TXPER* e *DENC* (67,8%) e fracamente positiva entre *TXPER* e *ALUG* (57,5%).

A correlação GW fracamente negativa entre *TXPER* e *REN* na maioria das subáreas (64,2%) está de acordo com o Smith (2004) que relaciona às PNTs com a precarização socioeconômica.

A correlação GW fracamente positiva entre *TXPER* e *ALUG* na maioria das subáreas (57,5%) é compreensível. O vínculo de um indivíduo em uma residência alugada é mais frágil do que o vínculo de alguém que é proprietário da mesma. Em caso de descoberta de uma fraude o indivíduo que aluga a casa (locatário) pode ocultar o ilícito mais facilmente do que se fosse proprietário da mesma.

Destaca-se que a correlação global e local podem apresentar resultados distintos, visto que o coeficiente da correlação a nível global representa todas as subáreas por um único valor numérico (Rodrigues, 2012).

Por meio das funções de *kernel* e da seleção de uma largura de banda, os coeficientes de correlação local GW afastam-se de seus valores globais, isto é, áreas distantes exercem uma menor influência do que áreas mais próximas. Ademais, o parâmetro da largura de banda *b* controla a taxa de decaimento das funções de *kernel*, podendo ser especificado respectivamente como número fixo (para amostras razoavelmente regulares) ou incorporando um número fixo de subárea – recomendado para amostras altamente irregulares, como é o caso da distribuição espacial das PNTs.

Dessa forma, suportado por Gollini *et al.* (2015) o parâmetro de largura de banda adotado neste estudo representa um número fixo de 48 subáreas; que representa 16% do total de setores censitários ou de subáreas.

Em último, segundo Gollini *et al.* (2015), a ausência da correlação entre variáveis em escala global não implica que essa mesma ausência esteja presente em nível local – Tabelas 3 e 5 e Figs. 5–7. Dessa forma, para a identificação das áreas mais propensas à ocorrência das PNTs a análise exploratória local ponderada geograficamente é a que fornece maior nível de detalhamento e um sistema de apoio à decisão para a busca das PNTs.

**Tabela 5. Coeficientes de correlação local GW contabilizados percentualmente por intervalos de classes.**

<i>Intervalos de Coeficientes Locais de Correlação GW</i>	<i>TXPER e REN</i>	<i>TXPER e DENC</i>	<i>TXPER e ALUG</i>
0 a 0,3	15,6%	25,2%	<b>57,5%</b>
0,3 a 0,6	2,3%	5,6%	9,3%
0,6 a 1,0	0,0%	0,4%	0,4%
-0,3 a 0	<b>64,2%</b>	<b>67,8%</b>	29,5%
-0,6 a -0,3	17,9%	1,0%	3,3%
-1,0 a -0,6	0,0%	0,0%	0,0%

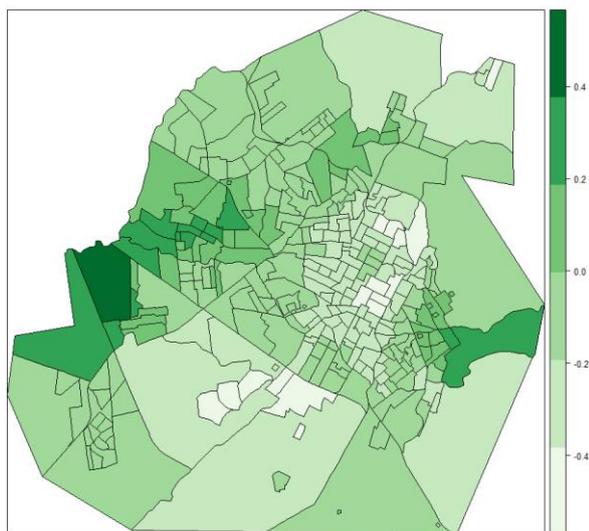


Fig. 5 Correlação ponderada geograficamente (GW) entre a Taxa de Perdas (TXPER) com a Renda Média (REN).

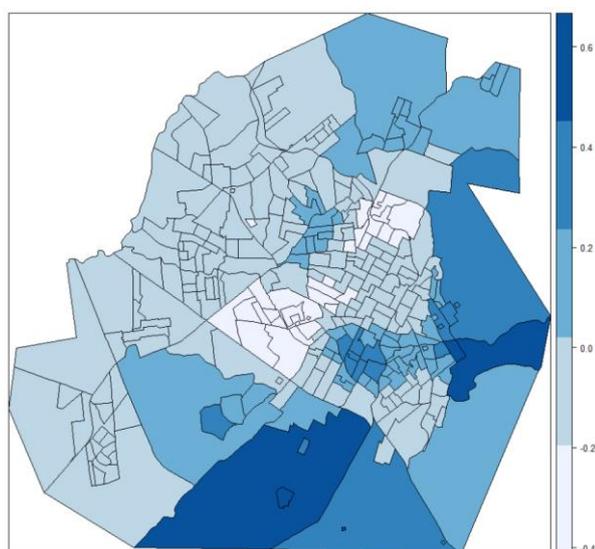


Fig. 6 Correlação ponderada geograficamente (GW) da Taxa de Perdas (TXPER) com a Densidade de Carga (DENC).

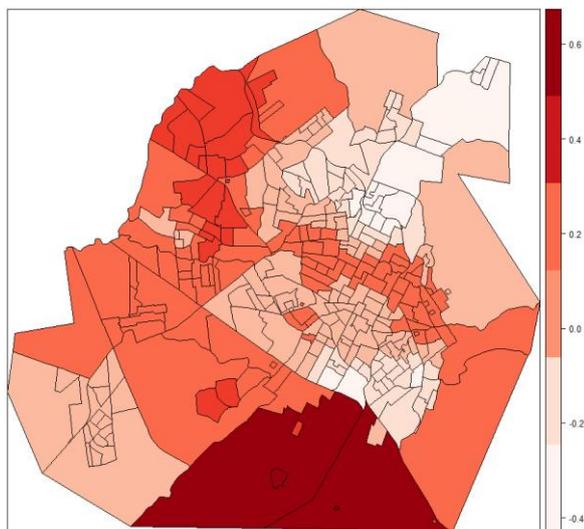


Fig. 7 Correlação ponderada geograficamente (GW) da Taxa de Perdas (TXPER) com UCs Alugadas (ALUG).

#### 4. CONCLUSÕES

Neste trabalho foi realizada a incorporação do espaço geográfico ao problema das PNTs com o auxílio das técnicas de análise exploratória espacial a fim de descrever a distribuição das variáveis de estudo. As observações das variáveis com respeito à distribuição e aos vizinhos e a busca por padrões na distribuição espacial dos dados foram representados através da produção de mapas temáticos.

A maioria dos trabalhos avaliados em perdas são focados na detecção pontual das mesmas, logo este trabalho surge como um complemento às linhas de pesquisa convencionais relacionadas à detecção das PNTs nas redes de distribuição de energia elétrica.

Observa-se que as análises deste estudo preliminar foram importantes para a caracterização das subáreas mais vulneráveis da cidade. A análise da correlação a nível local é eficiente para identificar o perfil real do cliente fraudador, levando em consideração a variável dependente TXPER com as variáveis independentes ou explicativas: DENC, ALUG e REN.

A maioria das subáreas apresentam correlação local GW fracamente negativa entre TXPER –REN (64,2%) e TXPER–DENC (67,8%) e fracamente positiva entre TXPER –ALUG (57,5%).

Observa-se em último que os coeficientes de correção são divididos em três níveis (valor absoluto): 0 – 0,3 (nível 1); 0,3 – 0,6 (nível 2) e 0,6 – 1,0 (nível 3). A correlação local GW entre TXPER e REN foi a única que apresentou número relevante de subáreas no nível 2 de moderada negativa (17,9%, Tabela 5). Portanto, atenção especial deve ser dada às PNTs em subáreas vulneráveis socioeconomicamente – zona leste da cidade, Fig. 1.

Em trabalhos futuros, após a análise exploratória realizada neste estudo, é sugerida a continuidade deste trabalho com a execução de modelos espaciais mais sofisticados para compreensão da relação encontrada entre as variáveis em muitas subáreas da cidade alvo desta análise.

#### AGRADECIMENTOS

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão da bolsa de Iniciação Científica (Processo: 19/04417-0).

#### REFERÊNCIAS

- Agência Nacional de Energia Elétrica (ANEEL) (2019). Relatório sobre perdas de energia. Disponível em: <<https://www.aneel.gov.br/>>.
- Angelos, E.W.S., Saavedra, O.R., Cortés, O.A.C., Souza, A.N. (2011). Detection and Identification of Abnormalities in Customer Consumptions in Power Distribution Systems. *IEEE Trans. on Power Delivery*, v. 26.
- Cabral, J.E., Pinto, J.O., Gontijo, E.M., Filho, J.R. (2004). Fraud detection in electrical energy consumers using rough sets. *IEEE International Conference on Systems, Man and Cybernetics*.
- Druck, S. Carvalho.M.S, Câmara,G, Monteiro, A.V (2004). *Análise Espacial de Dados Geográficos*, Embrapa, Brasília, DF.

- Faria, L.T. (2012). Sistema Inteligente Híbrido Intercomunicativo para Detecção de Perdas Comerciais. Dissertação (Mestrado em Engenharia Elétrica) – Faculdade de Engenharia, Universidade Estadual Paulista.
- Faria, L.T. (2016). Estimaco Espaco-Temporal das Perdas No Tcnicas no Sistema de Distribuico de Energia Eltrica. Tese (Doutorado em Engenharia Eltrica) – Faculdade de Engenharia, Universidade Estadual Paulista.
- Faria, L.T.; Melo, J.D.; Padilha-Feltrin, A. (2016). Spatial-temporal estimation for non-technical losses. *IEEE Trans. Power Delivery*, Piscataway, v. 31, n. 1, p. 362-369.
- Ferreira, M.C. (2014). *Iniciao  anlise geoespacial: teoria, tcnicas e exemplos para geoprocessamento*. So Paulo: Ed. Unesp.
- Gollini, I. Binbim L., Charlton, M., Brunson, C. Harris, P. (2015). GWmodel: an R package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, v. 63, p. 1-50.
- Instituto Brasileiro de Geografia e Estatística – IBGE (2010). Censo demogrfico 2010: Agregado por setores censitrios. 2ª Ed. Braslia, DF. Disponvel em: <<http://www.ibge.gov.br>>.
- Jiang, R.J. (2002). Wavelet based feature extraction and multiple classifiers for electricity fraud detection. *IEEE/PES Transmission and Distribution Conference and Exhibition, Anais, Asia Pacific*.
- Messinis, G.M., Hatziargyriou, N.D. (2018). Review of non-technical loss detection methods, *Electrical Power Systems Research*, v. 158, p. 250-266.
- Ministrio Pblico Federal (MPF). 3ª Cmara de Coordenao e Reviso (Org.). Perdas no tcnicas de energia eltrica, 2019.
- Nagi, J., Yap, K.S., Tiong, S.K., Ahmed, S.K, Mohamad, M. (2010). Nontechnical Loss Detection for Metered Customers in Power Utility Using Support Vector Machines. *IEEE Trans. Power Delivery*, v. 25, p. 1162-1171.
- Nagi, J., Yap, K.S., Tiong, S.K. Ahmed, S.K., Nagi, F. (2011). Improving SVM-Based Nontechnical Loss Detection In Power Utility Using the Fuzzy Inference System. *IEEE Trans. Power Delivery*, v. 26, p. 1284-1285
- Nizar, A.H., Dong, Z.Y., Wang, Y. (2008). Power Utility Nontechnical Loss Analysis with Extreme Learning Machine Method. *IEEE Trans. Power Systems*, v. 23, p. 946-955.
- Nizar, A.H., Dong, Z.Y., Zhang, P. Nizar, A.H., Dong, Z.Y., Zhang, P. (2008). Detection Rules for Non-Technical Losses Analysis in Power Utilities. *IEEE Power and Energy Society General Meeting*, Pittsburgh, PA, USA.
- Penin, C.A.S. (2008). Combate, preveno e otimizao das perdas comerciais de energia eltrica. Tese (Doutorado em Engenharia Eltrica) – Escola Politcnica, Universidade de So Paulo.
- R Core Team (2015). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponvel em: <<http://www.R-project.org/>>.
- Ramos, C.C.O., Sousa, A.N., Papa, J.P., Falco, A.X. (2011). A New Approach for Nontechnical Losses Detection Based on Optimum-Path Forest. *IEEE Trans. Power Systems*, v. 26, p. 181-189.
- Ramos, C.C.O., Souza, A. N., Falco, A.X., Papa, J. P. (2012). New insights on nontechnical losses characterization through evolutionary-based feature selection. *IEEE Trans. on Power Delivery*, v. 27, p. 140-146.
- Ramos, C.C.O. (2014). Caracterizao das Perdas Comerciais em Sistemas de Energia atravs de Tcnicas Inteligentes. Tese (Doutorado em Cincias) – Escola Politcnica, Universidade de So Paulo.
- Rodrigues, T.C.V. (2012). Regresso binomial negativa geograficamente ponderada: modelando super disperso espacial. Dissertao (Mestrado em Estatística) – Instituto de Cincias Exatas, Universidade de Braslia.
- Smith, T. B. (2004). Electricity theft: a comparative analysis. *Energy Policy*. v. 32, p. 2067-2076.
- Trevizzan, R.D., Rossoni, A., Bretas, A.S., Gazzana, D.S., Martin, R.P., Bretas, N.G., Bettiol, A.L; Carniato, A.; Passos, L. F.N. (2015). Non-technical losses identification using optimum-path forest and state estimation. In: *PowerTech*, Eindhoven.