

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/324090248>

Desenvolvimento de correlação para estimativa da taxa de geração per capita de resíduos sólidos urbanos no estado de São Paulo: influências da população, renda per capita e consumo...

Article in *Engenharia Sanitaria e Ambiental* · March 2018

DOI: 10.1590/s1413-41522018167380

CITATIONS

18

READS

116

3 authors, including:



Reinaldo Pisani Júnior

Universidade de Ribeirão Preto

19 PUBLICATIONS 143 CITATIONS

SEE PROFILE

Desenvolvimento de correlação para estimativa da taxa de geração *per capita* de resíduos sólidos urbanos no estado de São Paulo: influências da população, renda *per capita* e consumo de energia elétrica

Development of a correlation to estimate per capita municipal solid waste generation rates in São Paulo state, Brazil: population, per capita income and electricity consumption influences

Reinaldo Pisani Junior¹, Marcus Cesar Avezum Alves de Castro², Antonio Álvarez da Costa³

RESUMO

A previsão da geração dos resíduos sólidos urbanos (RSU) é fundamental para escolha e dimensionamento das operações e dos processos envolvidos na cadeia de gerenciamento em âmbito municipal. O desafio da sustentabilidade financeira desse gerenciamento passa pela criação de indicadores que permitam a cobrança individualizada e proporcional da geração *per capita* de cada município. Buscou-se, então, desenvolver equações para estimar a taxa de geração *per capita* (TGP) de RSU nos municípios do estado de São Paulo. Adotou-se a premissa de que as variáveis população (P), renda *per capita* diária (RPC) e consumo *per capita* diário de energia elétrica (Ce) nos municípios eram intervenientes na TGP. A série de dados envolveu 238 municípios que rotineiramente pesaram seus resíduos no ano-base de 2013. Diversas funções foram testadas com o intuito de ajustar as TGPs observadas às calculadas pelo método dos mínimos quadrados. A qualidade dos ajustes foi avaliada por meio da comparação dos valores previstos e observados com a reta bisetriz do primeiro quadrante, dos coeficientes de correlação de Pearson (r) e de determinação (R²), das raízes quadradas dos erros médios ao quadrado (RMSEs) e dos erros percentuais médios (E_p). A equação que melhor representou o conjunto de dados forneceu r de 0,49, R² de 0,24, RMSE de 0,217 kg.hab⁻¹.d⁻¹ e E_p de -14,1%. Apesar de R² relativamente baixo, foi demonstrado, pelo teste de t de Student, que a equação proposta foi capaz de representar valores médios e resultar em mesma variância com probabilidade acima de 99%, fato que permitiu considerar a equação adequada para prever as TGPs nos municípios paulistas.

Palavras-chave: gerenciamento de resíduos sólidos; resíduos sólidos urbanos; taxa de geração *per capita*; análise de regressão.

ABSTRACT

Forecasting municipal solid waste (MSW) generation is fundamental in choosing and scaling the operations and processes involved in municipal management. The challenge for the financial sustainability of this management is to create indicators that enable municipal solid waste fees to be charged in proportion to the amount of MSW generated by each resident. To this end, we sought to develop equations to estimate the per capita waste generation rate (TGP) in the municipalities of the state of São Paulo, Brazil. It was assumed that the variables of population (P), daily per capita income (RPC) and daily per capita consumption of electric energy (Ce) in the municipalities were determining factors of the TGP. The dataset involved 238 municipalities that routinely weighed their solid wastes in the baseline year of 2013. Several functions were tested in order to adjust the observed TGP to those calculated by the least squares method. The quality of the adjustments was evaluated by comparing the expected and observed values with the bisector of the first quadrant, as well as by the Pearson (r) and determination (R²) coefficients, root-mean-square error (RMSE), and mean percentage error (E_p) values. The equation that best represented the dataset had an r of 0,49, R² of 0,24, RMSE of 0,217 kg.person⁻¹.d⁻¹ and E_p of -14,1%. Despite the relatively low R², it was demonstrated by Student's t-test that the proposed equation was able to represent mean values and result in the same variance with probability higher than 99%. This fact allows to consider the equation adequate for predicting the TGP in the municipalities of the state of São Paulo.

Keywords: waste management; municipal solid waste; per capita generation rate; regression analysis.

¹Mestre e Doutor em Engenharia Química, Universidade Federal de São Carlos. Professor e pesquisador do Programa de Pós-graduação em Tecnologia Ambiental, Universidade de Ribeirão Preto - Ribeirão Preto (SP), Brasil.

²Doutor em Hidráulica e Saneamento, Universidade de São Paulo. Professor assistente do curso de Engenharia Ambiental, Universidade Estadual Paulista "Júlio de Mesquita Filho" - Rio Claro (SP), Brasil.

³Matemático pela Faculdade de Filosofia, Ciências e Letras de Passos. Mestrando em Tecnologia Ambiental, Universidade de Ribeirão Preto - Ribeirão Preto (SP), Brasil.

Endereço para correspondência: Reinaldo Pisani Júnior - Rua José Bonifácio, 787 - 13560-610 - São Carlos (SP), Brasil - E-mail: pisanijr@terra.com.br

Recebido: 01/08/16 - **Aceito:** 02/03/17 - **Reg. ABES:** 167380

INTRODUÇÃO

Os processos de produção e o consumo humano geram grande diversidade e quantidade de resíduos que requerem destinação ambientalmente adequada. Os resíduos sólidos urbanos (RSU) representam parcela significativa dos resíduos sólidos que, devido aos riscos potenciais ao meio ambiente e à saúde, necessitam de gestão e gerenciamento integrados em concordância com a manutenção da qualidade ambiental.

A quantificação do montante de resíduos é fundamental para a escolha e o dimensionamento das operações envolvidas na cadeia de gerenciamento. Além disso, a estimativa das quantidades geradas é necessária para totalização das massas ou dos volumes, uma vez definida a vida útil dos sistemas utilizados no gerenciamento. A taxa de geração *per capita* (TGP) é um indicador bastante utilizado para representar a geração de RSU, pois relaciona a massa diária de resíduo produzido por habitante em uma dada localidade.

Segundo o art. 7º da Política Nacional de Resíduo Sólidos, Lei nº 12.305 (BRASIL, 2010b, grifo dos autores), um de seus objetivos é a

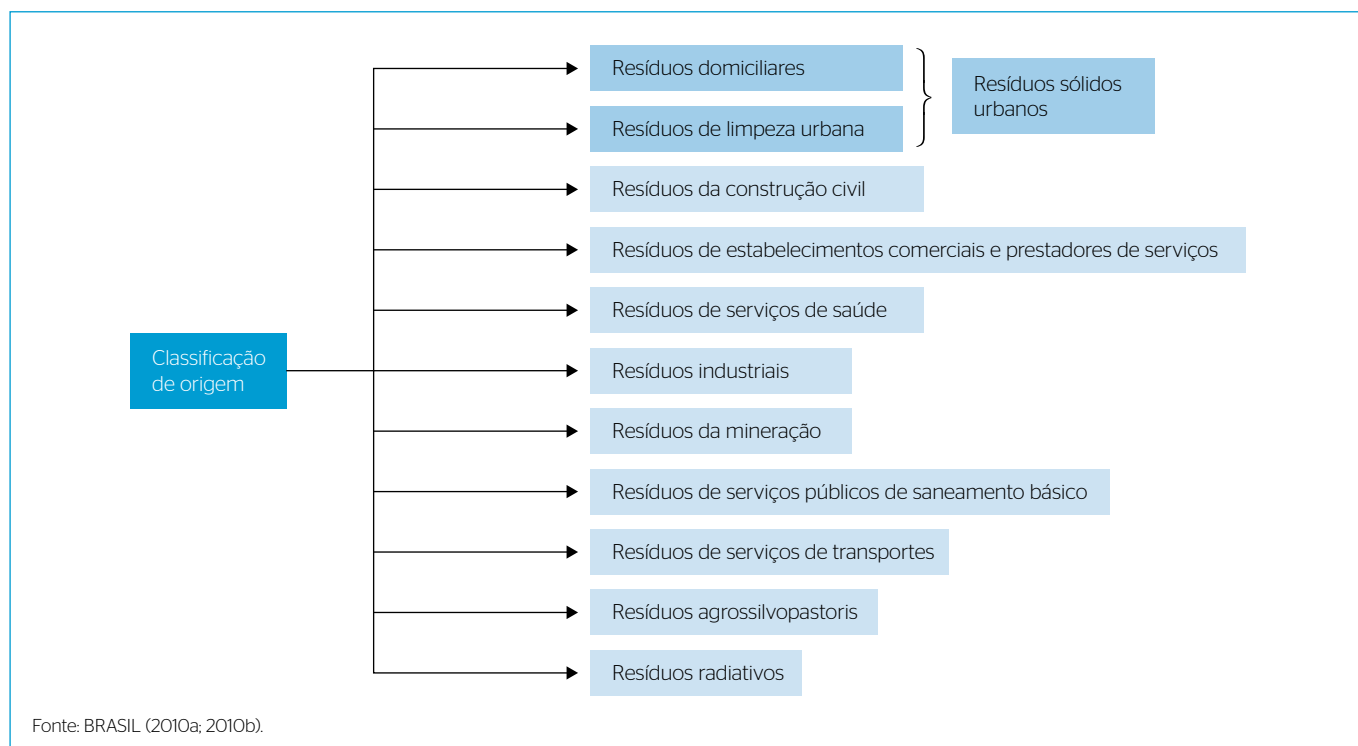
regularidade, continuidade, funcionalidade e universalização da prestação dos serviços públicos de limpeza urbana e de manejo de resíduos sólidos, com adoção de mecanismos gerenciais e econômicos que assegurem a recuperação dos custos dos serviços prestados, como forma de garantir sua sustentabilidade operacional e financeira.

Nessa linha, o estabelecimento de correlações entre a TGP de resíduos com o consumo de energia e renda dos municípios, variáveis com mensuração temporal já estabelecida, permite a cobrança dos serviços individualizada e proporcional à geração de cada município.

Os RSU são formados pelos resíduos domiciliares, aqueles gerados nas moradias, e pelos resíduos de limpeza urbana, decorrentes da varrição, da limpeza de logradouros e vias públicas e de outros serviços de limpeza urbana (Figura 1).

O estado de São Paulo era formado por 655 municípios no ano de 2013; dentre eles, 533 participaram do diagnóstico do manejo de resíduos sólidos de 2013, dos quais 238 afirmaram pesar rotineiramente os resíduos coletados (SECRETARIA NACIONAL DE SANEAMENTO AMBIENTAL, 2015). Esse conjunto de 238 municípios representou aproximadamente 82% da população total em 2013 (42.304.694 habitantes). São Paulo é o estado mais urbanizado do Brasil, com taxa registrada, em 2010, de 95,9% (FUNDAÇÃO SEADE, 2013).

O estado abriga o maior parque industrial e, conseqüentemente, a maior produção econômica do país. Apresenta também o maior Índice de Desenvolvimento Humano (0,783 em uma escala de 0 a 1). Em 2013, gerou cerca de 59.291 t.d⁻¹ de RSU, dos quais 74,6% foram dispostos em aterros sanitários, 15,1%, em aterros controlados, e 8,5%, em lixões (ABRELPE, 2014). Panorama diferente foi descrito pela Companhia Ambiental do Estado de São Paulo (CETESB, 2015), que relatou que 97,8% (39.009 t.d⁻¹) do montante total (39.865 t.d⁻¹) dos RSU foram



Fonte: BRASIL (2010a; 2010b).

Figura 1 - Classificação de origem dos resíduos sólidos, segundo a Lei nº 12.305/2010.

adequadamente dispostos e apenas 2,2% (856 t.d⁻¹) foram depositados em locais inadequados em 2013.

Hockett, Lober e Pilgrim (1995) correlacionaram a taxa de geração de RSU no estado da Carolina do Norte (Estados Unidos) por meio de variáveis socioeconômicas, porém, sem envolver a população de cada município, e obtiveram coeficiente de determinação (R^2) no intervalo de 0,37 a 0,50.

Já Keser, Duzgunb e Aksoy (2012), em diferentes localidades da Turquia, levaram em consideração fatores socioeconômicos e climáticos na determinação da taxa de geração de RSU. Nessa pesquisa, foram utilizados modelos de autorregressão espacial simultânea, regressão não espacial e regressão geograficamente ponderada para as análises de dados espaciais. Obtiveram coeficientes de Pearson (r) no intervalo de 0,47 a 0,59, porém, com valores das raízes quadradas dos erros médios ao quadrado (RMSEs) da ordem de 50% dos valores previstos.

Xu *et al.* (2013) desenvolveram um modelo híbrido que combinava o modelo sazonal autorregressivo de média móvel para a previsão da geração de RSU em uma cidade específica (na China) para múltiplas escalas de tempo, sem a necessidade de considerar outras variáveis. Os resultados alcançados demonstraram que o modelo foi bastante preciso, porém, restrito à cidade avaliada.

Benítez *et al.* (2008) propuseram, para a cidade do México, modelos capazes de relacionar a geração *per capita* de resíduos domiciliares (variável dependente) para diferentes variáveis (independentes), tais como educação, renda por domicílio e número de moradores. A fim de analisar o modelo que melhor explicava a previsão, o conjunto de dados foi analisado a partir dos testes de normalidade, Kolmogorov-Smirnov, multicolinearidade e heterocedasticidade. O valor máximo de R^2 foi de 0,51.

Já no estudo de Navarro, Diamadopoulos e Ginestar (2002), foram demonstrados dois modelos. Tendo como objeto três cidades gregas e espanholas, propuseram técnicas distintas para prever a geração de RSU. A primeira era baseada em dinâmica não linear, na qual foi proposto um modelo comparativo de desempenho, com uma média sazonal autorregressiva em movimento denominado (sARIMA). A segunda técnica assumiu que a geração de RSU era um sistema dinâmico discreto de uma única variável escalar. Apesar dos desvios resultantes de até 5%, não disponibilizaram equações que permitissem utilização.

A relação entre fatores demográficos e socioeconômicos foi alvo de estudo no Brasil realizado por Pinto, Pereira e Freitas (2012) e Dias *et al.* (2012), que correlacionaram a taxa de geração de RSU em Belo Horizonte, Minas Gerais, com a população, o índice de qualidade de vida urbana (IQVU) e a renda *per capita*. Demonstraram que o IQVU e a renda, conjuntamente com o número de moradores, eram variáveis que influenciam a geração de resíduos sólidos em bairros do município. No entanto, as equações descritas são aplicáveis apenas ao município avaliado. Nesse mesmo município foram empregados diferentes métodos

que permitiram chegar à conclusão de que os fatores socioeconômicos, como educação e renda, e demográficos, como estrutura etária e domiciliar, eram importantes na geração de RSU em sub-regiões do município (SILVA; BARBIERI; MÓR, 2012). Sendo assim, este estudo teve como objetivo desenvolver equações que permitissem estimar a TGP de RSU nos municípios do estado de São Paulo, a partir das variáveis população total (P), renda *per capita* e consumo de energia elétrica.

MATERIAL E MÉTODOS

Estimar e projetar as quantidades de resíduos gerados são ações fundamentais para o gerenciamento integrado de resíduos sólidos; nesse contexto, são descritos métodos e parâmetros estatísticos utilizados para verificar o nível de confiança da amostra disponível, avaliar a dependência entre as variáveis e investigar a qualidade do ajuste do modelo ao conjunto de dados.

No estudo de correlação, procura-se verificar a dependência da TGP de RSU em função da população total, renda *per capita* e consumo *per capita* diário de energia elétrica nos municípios do estado de São Paulo. As populações totais e rendas *per capita* foram utilizadas como variáveis independentes, pois suas influências foram evidenciadas na literatura, já a variável consumo *per capita* de energia elétrica foi escolhida por estar associada indiretamente ao padrão de consumo, à infraestrutura disponível e à utilização de bens e serviços.

Base de dados

Os dados iniciais de geração de resíduos sólidos, para os municípios que rotineiramente pesam seus resíduos, foram obtidos junto ao Sistema Nacional de Informações sobre Saneamento (SNIS), que disponibilizou, em 2015, o diagnóstico do manejo de resíduos sólidos urbanos 2013–2015 (SECRETARIA NACIONAL DE SANEAMENTO AMBIENTAL, 2015). Nesse documento são explicitadas, dentre muitas informações, separadamente as quantidades dos resíduos domiciliares, dos resíduos de limpeza urbana e dos resíduos de serviços de saúde, assim como a utilização de balança para a totalização dos montantes e a cobertura dos serviços de coleta. Portanto, não se admite por hipótese que haja parcelas de outros tipos de resíduos naqueles tipificados na pesquisa. TGPs de RSU que apresentaram inconsistências claras, pela comparação dos valores informados ao SNIS com os constantes nos Planos Municipais de Gestão Integrada de Resíduos Sólidos (PMGIRSS), foram substituídas pelos valores informados nos PMGIRSS.

Já as populações totais (P), as rendas *per capita* anuais e os consumos anuais de eletricidade (CE) de cada município foram obtidos mediante consulta ao site eletrônico da Fundação Estadual de Análise de Dados (FUNDAÇÃO SEADE, 2013). As rendas *per capita* dos municípios paulistas estavam atualizadas com referência ao ano-base de 2010, sendo assim, os seus valores foram convertidos em dólares americanos para

a cotação média nesse ano, de R\$ 1,76 por USD. Além disso, a fim de correlacionar os resultados, as rendas *per capita* anuais de eletricidade foram utilizadas na forma de rendas *per capita* diárias (RPC) (USD. hab⁻¹.d⁻¹) e consumos *per capita* diários de eletricidade (Ce) em cada município (kWh.hab⁻¹.d⁻¹).

Representatividade da amostra

O espaço amostral, uma vez que nem todos os municípios paulistas participaram do Diagnóstico do Manejo de Resíduos Sólidos Urbanos – 2013-2015 – e informaram pesar rotineiramente seus RSU, pôde ser estabelecido na forma de dois critérios principais, com base no número de municípios ou na população abrangida, em relação ao total do estado de São Paulo.

A representatividade da amostra foi avaliada por meio do nível de confiança, associado a uma margem de erro. A margem de erro corresponde a um intervalo no qual deve estar compreendido o valor estimado no universo da pesquisa. O nível de confiança expressa o grau de certeza de que o valor encontrado esteja realmente dentro da margem de erro estipulada. A relação entre essas variáveis foi expressa pela Equação 1 (ISRAEL, 1992):

$$n = \frac{N \cdot Z^2 \cdot p \cdot (1-p)}{(N-1) \cdot e^2 + Z^2 \cdot p \cdot (1-p)} \quad (1)$$

Na qual:

n = o tamanho da amostra;

N = o tamanho do universo amostral (646 municípios ou 42.304.694 habitantes em 2013);

Z = o desvio do valor médio aceitável para o intervalo de confiança adotado (determinado na curva de distribuição normal);

e = a margem de erro estipulada; e

p = a proporção esperada ou heterogeneidade do espaço amostral, normalmente 50% quando não estão disponíveis informações a respeito da variabilidade da amostra.

Sendo assim, foram adotados valores de margem de erro e intervalo de confiança que, a partir do número total de municípios ou da população do estado, resultassem no número de municípios ou de população que participaram da amostra.

Avaliação do modelo

Inicialmente, foram ajustadas funções lineares, com o intuito de representar as TGPs de RSU, em função das variáveis população total, RPC média e consumo médio de energia elétrica *per capita* de cada município do espaço amostral, combinadas duas a duas independentemente ou as três variáveis simultaneamente. Posteriormente, foi feita a combinação de três variáveis independentes, porém, utilizando funções logarítmicas para a P e Ce e dependência linear para a RPC

(Equações 2 a 6). Para isso, foram utilizados os programas computacionais BIOESTAT 5.3, gratuito, voltado para aplicações em Ecologia e desenvolvido pelo Instituto MAMIRAUÁ (AYRES, 2007), e Microsoft Excel, por intermédio da função “SOLVER” para aplicação do método dos mínimos quadrados.

$$TGP = a + b \cdot (P) + c \cdot (RPC) \quad (2)$$

$$TGP = a + b \cdot (P) + c \cdot (Ce) \quad (3)$$

$$TGP = a + b \cdot (RPC) + c \cdot (Ce) \quad (4)$$

$$TGP = a + b \cdot (P) + c \cdot (RPC) + d \cdot (Ce) \quad (5)$$

$$TGP = a + b \cdot \ln(P) + c \cdot (RPC) + d \cdot [\ln(Ce)]^e \quad (6)$$

Nas quais:

a, b, c, d e e = constantes de ajuste;

P = a população total (hab.);

Ce = o consumo *per capita* diário de energia elétrica (kWh.hab⁻¹.d⁻¹); e

RPC = a renda *per capita* média diária (USD.hab⁻¹.d⁻¹) em cada município do estado de São Paulo.

A avaliação estatística das qualidades dos ajustes envolveu análises de correlação linear, de regressão linear e de regressão multivariável e o teste *t* de Student.

Dependência entre variáveis

r, também conhecido como coeficiente de correlação produto-momento ou o coeficiente r de Pearson (Equação 7), tem por objetivo medir o grau da correlação linear entre duas variáveis quantitativas. Tal avaliação apresenta-se como um índice adimensional com valores situados entre -1,0 e 1,0, normalmente representado pela letra r, e reflete a intensidade de uma relação linear entre dois conjuntos de dados. Se r for igual a 1, há uma correlação perfeita positiva entre as duas variáveis; se r for igual a -1, obtém-se uma correlação negativa perfeita entre as duas variáveis, indicando que movem-se em direções opostas. Isto é, se uma aumenta, a outra sempre diminui. Enquanto para duas variáveis perfeitamente correlacionadas positivamente (r=1), elas estariam em perfeita proporção na mesma direção. Entretanto, se o resultado final alcançado for igual a 0, as duas variáveis não dependem linearmente uma da outra, podendo existir uma outra dependência que seja não linear. Assim, o resultado r=0 deve ser investigado por outros meios (LEGATES; MCCABE JR., 1999; MORIASI *et al.*, 2007). Pela Equação 7 é possível definir r.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y}_{medida})}{\sqrt{\left(\sum(x_i - \bar{x})^2\right) \cdot \sqrt{\left(\sum(y_i - \bar{y}_{medida})^2\right)}} \quad (7)$$

Sendo:

x_i = a variável independente observada;

\bar{x} = o valor médio da variável independente x_i ;

y_i = a variável dependente observada; e

\bar{y}_{medida} = o valor médio da variável dependente y_i .

Como valores extremos (-1 e 1) dificilmente são encontrados na prática, é importante qualificar intervalos para o coeficiente r . Para Cohen (1988), valores entre 0,10 e 0,29 podem ser considerados pequenos; resultados entre 0,30 e 0,49 podem ser avaliados como médios, e de 0,50 a 1,00, como grandes. Dancy e Reidy (2006) apontaram para uma classificação ligeiramente diferente: r de 0,10 até 0,30 como fraca dependência entre as variáveis; 0,40 a 0,60 como moderada; e de 0,70 a 1,00 como forte correlação. Seja como for, o certo é que quanto mais próximo a 1 (independente do sinal), maior é o grau de dependência estatística linear entre as variáveis. No outro oposto, quanto mais próximo de zero, menor é a força dessa relação.

Qualidade de ajuste

A avaliação da qualidade dos ajustes foi feita a partir do cálculo de alguns parâmetros amplamente utilizados em avaliações estatísticas: o R^2 , RMSE e erro percentual médio (E_p), além da comparação dos valores previstos e observados para TGP com a reta bissetriz no primeiro quadrante, incluindo o intervalo compreendido entre -RMSE e +RMSE. R^2 é o quadrado de r (Equação 8), e apresenta-se como uma das formas de avaliar a qualidade do ajuste do modelo utilizado. Descreve a proporção da variabilidade em uma variável que é explicada pela variabilidade da outra. Basicamente, esse coeficiente indica quanto o modelo foi capaz de explicar os dados coletados. O valor de R^2 pode variar de 0 a 1, sendo pouco comum uma correlação perfeita ($R^2=1$) na prática, uma vez que existem muitos fatores interferentes no fenômeno analisado que podem não estar presentes nos modelos testados, desse modo, valores mais elevados de R^2 estão associados a menores variâncias de erros, ou seja, quanto mais próximo de 1, mais explicativo é o modelo e melhor é a qualidade do ajuste (LEGATES; MCCABE JR., 1999; MORIASI *et al.*, 2007).

$$R^2 = \left\{ \frac{\sum_{i=1}^n [y_{medida_i} - \bar{y}_{medida}] \cdot [y_{calc_i} - \bar{y}_{calc}]}{\left[\sum_{i=1}^n (y_{medida_i} - \bar{y}_{medida})^2 \right]^{0,5} \cdot \left[\sum_{i=1}^n (y_{calc_i} - \bar{y}_{calc})^2 \right]^{0,5}} \right\}^2 \quad (8)$$

Na qual:

y_{medida_i} = a variável dependente observada para cada elemento do conjunto de dados;

\bar{y}_{medida} = o valor médio da variável dependente observada do conjunto de dados;

y_{calc_i} = a variável dependente calculada para cada elemento do conjunto de dados; e

\bar{y}_{calc} = o valor médio da variável dependente calculada para conjunto de dados; e

n = o número de elementos contidos no conjunto de dados de cada amostra.

A barra superior indica a média de todo o período de avaliação, observando-se que R^2 é limitado na medida em que uniformiza as diferenças entre as médias e as variâncias observadas e previstas, uma vez que avalia apenas as relações lineares entre as variáveis (LEGATES; MCCABE JR., 1999).

Diversos são os desvios comumente utilizados na avaliação de modelos, com o intuito de avaliar as imprecisões dos elementos de interesse, o que auxilia na análise e interpretação dos resultados. Dentre eles, as RMSEs (Equação 9) quantificam os erros em termos de unidades da variável, é a estatística não negativa que não tem limite superior, porém, no extremo oposto, valores iguais a 0 indicam um ajuste perfeito (ABDULLAH, 2009; MORIASI *et al.*, 2007; YAFEE, MCGEE, 2000).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_{medida_i} - \bar{y}_{calc_i})^2}{n - v}} \quad (9)$$

Sendo que:

y_{medida_i} = a variável dependente observada para cada elemento do conjunto de dados;

\bar{y}_{calc} = o valor médio da variável dependente calculado para todo o conjunto de dados;

n = o número de elementos contidos no conjunto de dados da amostra; e

v = o número de variáveis independentes envolvidas no modelo.

E_p (Equação 10) expressa a proporção do desvio encontrado entre os valores observados e previstos, que se traduz em melhor previsão quanto menor for E_p :

$$E_p = \frac{100}{n} \sum_{i=0}^n \left[\frac{y_{medida_i} - y_{calc_i}}{y_{medida_i}} \right] \quad (10)$$

Na qual:

y_{medida_i} = a variável dependente observada para cada elemento do conjunto de dados;

y_{calc_i} = a variável dependente calculada para cada elemento do conjunto de dados; e

n = o número de elementos contidos no conjunto de dados de cada amostra.

Teste *t* de Student para correlação

O teste *t* de Student utiliza conceitos estatísticos para rejeitar ou não uma hipótese nula (H_0), quando a estatística de teste (*t*) segue uma distribuição normal. Essa premissa é normalmente usada quando a estatística de teste segue uma distribuição normal, mas a variância da

população é desconhecida. Assim, utiliza-se uma variância amostral, e, com esse ajuste, a estatística de teste passa a seguir uma distribuição *t* de Student. Isso significa que, se forem realizadas amostras de determinado tamanho a partir de uma única população e se forem calculadas as médias de uma variável dessa população que apresenta uma distribuição normal, a distribuição dessas médias seguiria uma distribuição *t* de Student. O formato da distribuição *t* de Student é dependente do número de graus de liberdade, ou seja, quanto maior esse número, mais concentrada é a distribuição. Quanto maiores os graus de liberdade, mais a distribuição *t* de Student se aproxima da distribuição normal.

Ao calcular o valor de *t* pela Equação 11 e aplicá-lo na função “densidade de probabilidade da distribuição *t* de Student”, avalia-se a integral dessa função para valores maiores ou iguais a *t*. Essa área representa a probabilidade de a média das amostras em estudo terem apresentado valores observados ou valores mais extremos. Se a probabilidade desse resultado ocorrer for pequena, pode-se concluir que o resultado observado é estatisticamente relevante. Essa probabilidade também é chamada de valor *p* ou *p*. Consequentemente, o nível de confiança é igual a 1-*p*. Normalmente, utiliza-se um ponto de corte, *p*<0,05 (nível de confiança de 95%) para o valor *p* ou para o nível de confiança para definir se a hipótese nula deve ser rejeitada ou não. Se o valor de *p* for menor que esse ponto de corte, a hipótese nula é rejeitada. Caso contrário, a hipótese nula não é rejeitada (TIBONI, 2010).

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} \tag{11}$$

Na qual:

t = valor para comparação na função densidade de probabilidade da distribuição *t* de Student;

r = coeficiente de correlação de Pearson; e

n = número de pares de valores testados.

RESULTADOS E DISCUSSÃO

Nas Figuras 2 a 4 são mostradas as TGP de RSU para os municípios do estado de São Paulo que informaram pesar rotineiramente seus resíduos (SECRETARIA NACIONAL DE SANEAMENTO AMBIENTAL, 2015), referente ao ano-base de 2013.

Nota-se, pelas Figuras 2 a 4, dispersão acentuada da TGP de RSU em relação às variáveis independentes. Há municípios com geração superior a 1,0 kg.hab⁻¹.d⁻¹ para populações na faixa de 3 a 400 mil habitantes; por outro lado, cidades com TGP inesperadamente baixas, inferiores a 0,5 kg.hab⁻¹.d⁻¹, para populações no intervalo de 3 a 40 mil habitantes. Percebe-se, pelas Figuras 2 e 3, o mesmo comportamento entre TGP quanto a *P* e *CE*, cuja tendência foi de crescimento em relação ao logaritmo dessas variáveis. Porém, o consumo total de energia

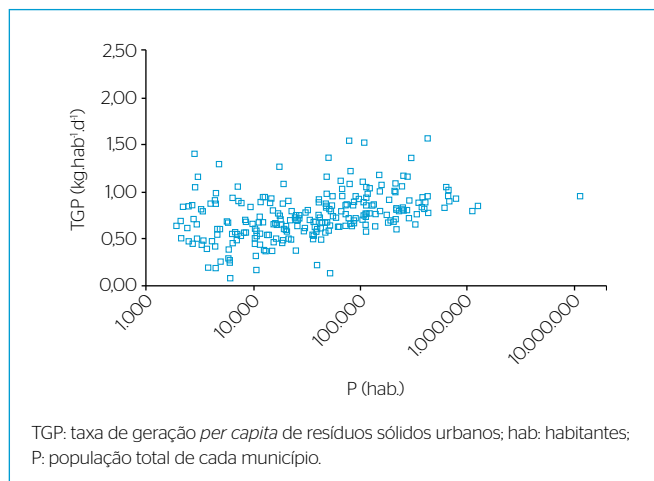


Figura 2 – Taxa de geração *per capita* de resíduos sólidos urbanos em função da população total dos municípios paulistas que rotineiramente pesaram seus resíduos.

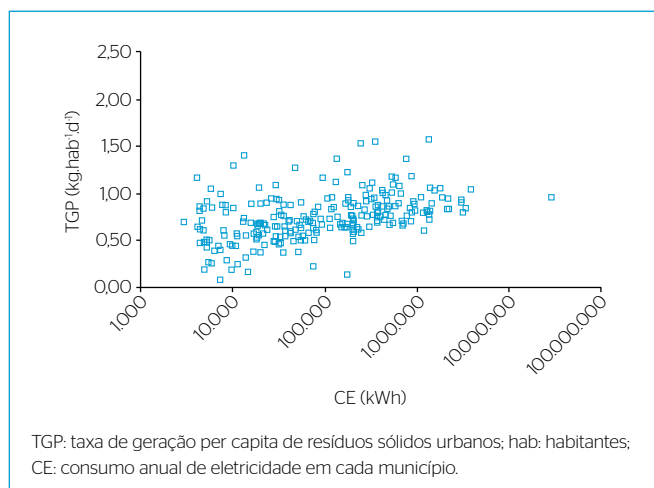


Figura 3 – Taxa de geração *per capita* de resíduos sólidos urbanos em função do consumo anual de energia elétrica nos municípios paulistas que rotineiramente pesaram seus resíduos.

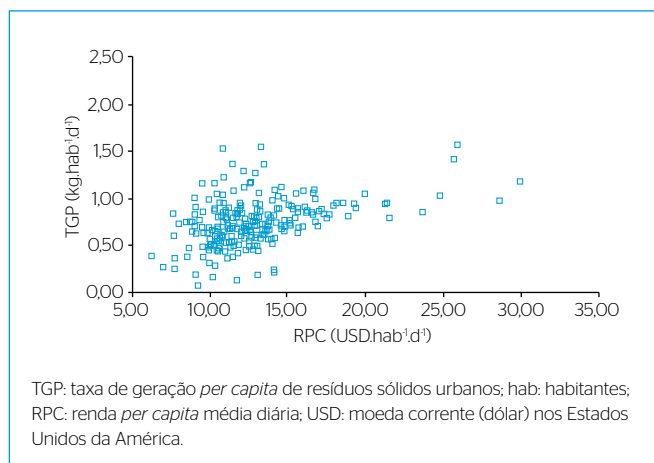


Figura 4 – Taxa de geração *per capita* de resíduos sólidos urbanos em função da renda *per capita* média diária nos municípios paulistas que rotineiramente pesaram seus resíduos.

elétrica é bastante dependente da população de cada município, portanto, optou-se por utilizar o consumo diário *per capita* de eletricidade para tornar as variáveis independentes entre si. Já na Figura 4, apesar da grande dispersão dos pontos, o comportamento foi de crescimento linear de TGP a partir da renda *per capita* diária de cada município.

As TGPs de RSU de nove municípios foram glosadas, por apresentarem valores exageradamente altos; assim, os valores informados ao SNIS foram substituídos pelas gerações encontradas nos PMGIRSs. Na Tabela 1 são mostradas as informações referentes a esses municípios. Logo, apenas três cidades foram excluídas do conjunto de dados, pois seus planos não foram localizados.

O universo amostral foi formado por 238 municípios, cuja população somada foi de 34.551.664 habitantes (82%), diante do número de municípios que responderam à coleta de informações junto ao SNIS (533 cidades) e ao total que compunha o estado de São Paulo, de 645 municípios (42.304.694 habitantes) em 2013. Pelo critério de número de cidades, pôde-se afirmar que 238 municípios representam o total de 645 com uma margem de erro de 5,1% e intervalo de confiança de 94,9%. Já com base na população, 34.551.664 habitantes representam a população total de 42.304.694 habitantes com margem de erro de apenas 0,015% e intervalo de confiança de 99,985%. Portanto, concluiu-se que a amostra foi representativa do estado de São Paulo e que os resultados obtidos são passíveis de extrapolação para os demais municípios, incluindo aqueles que não participaram do diagnóstico do manejo de resíduos sólidos urbanos 2013–2015.

Regressão linear múltipla

A regressão linear múltipla de TGP em função de P e RPC (Equação 1) resultou na Equação 12:

$$TGP = 0,348 + 4,19 \cdot 10^{-9} \cdot P_{(hab)} + 0,0306 \cdot RPC_{(USD.hab^{-1}.d^{-1})} \quad (12)$$

Tabela 1 - Municípios para os quais as informações foram glosadas ou suprimidas.

Município	População (hab.)	TGP _{SNIS} (kg.hab ⁻¹ .d ⁻¹)	TGP _{PMGIRS} (kg.hab ⁻¹ .d ⁻¹)
Cruzália	2.222	1,24	0,84
Lutécia	2.682	1,56	0,86
Lavínia	8.666	1,99	Não informado
Iacanga	10.414	1,57	0,60
Ilhabela	29.837	2,86	Não informado
Paraguaçu Paulista	42.858	3,20	Não informado
Registro	54.107	1,53	0,86
Atibaia	130.606	1,58	0,86
Guarulhos	1.260.840	1,53	0,84

hab: habitantes; TGP_{SNIS}: taxa de geração *per capita* de resíduos sólidos urbanos disponível no Sistema Nacional de Informações sobre Saneamento; TGP_{PMGIRS}: taxa de geração *per capita* de resíduos sólidos urbanos obtida no Plano Municipal de Gestão Integrada de Resíduos Sólidos.

Na qual:

TGP = taxa de geração *per capita* de resíduos sólidos urbano (kg.hab⁻¹.d⁻¹);

P = população total (hab.); e

RPC = renda *per capita* média diária (USD.hab⁻¹.d⁻¹) em cada município do estado de São Paulo.

Essa correlação múltipla apresentou r de 0,43, R² de 0,18, RMSE de 0,225 kg.hab⁻¹.d⁻¹ e E_p de -15,2%, indicando média ou intermediária correlação múltipla.

A correlação múltipla entre a variável dependente TGP e as variáveis independentes, população total (P em hab.) e Ce (Equação 13), resultaram em r de apenas 0,24; R² de 0,06; E_p de -17,2%; e RMSE igual 0,241 kg.hab⁻¹.d⁻¹:

$$TGP = 0,675 + 3,80 \cdot 10^{-8} \cdot P_{(hab)} + 0,00745 \cdot Ce_{(kW.hab^{-1}.d^{-1})} \quad (13)$$

Na qual:

TGP = taxa de geração *per capita* de resíduos sólidos urbano (kg.hab⁻¹.d⁻¹);

P = população total (hab.); e

Ce = consumo *per capita* diário de energia elétrica (kWh.hab⁻¹.d⁻¹).

Essa combinação de variáveis piorou a qualidade de ajuste quando comparada com a regressão que empregou as variáveis P e RPC, em razão de valores de r, R² e módulo de E_p menor e RMSE mais elevados.

A regulação de TGP em função linear de RPC e Ce (Equação 14) forneceu o melhor ajuste de duas variáveis até o momento, pois r foi de 0,44; R² de 0,19; E_p de -15,3% e RMSE igual a 0,223 kg.hab⁻¹.d⁻¹:

$$TGP = 0,345 + 0,0284 \cdot RPC_{(USD.hab^{-1}.d^{-1})} + 0,00408 \cdot Ce_{(kW.hab^{-1}.d^{-1})} \quad (14)$$

Na qual:

TGP = taxa de geração *per capita* de resíduos sólidos urbano (kg.hab⁻¹.d⁻¹);

RPC = renda *per capita* média diária (USD.hab⁻¹.d⁻¹); e

Ce = consumo *per capita* diário de energia elétrica (kWh.hab⁻¹.d⁻¹) em cada município do estado de São Paulo.

Já a correlação multilinear empregando as três variáveis independentes simultaneamente (P, Ce e RPC) resultou em r de 0,44; R² de 0,19; E_p de +9,1%; e RMSE de 0,279 kg.hab⁻¹.d⁻¹ (Equação 15):

$$TGP = 0,339 + 6,19 \cdot 10^{-9} \cdot P_{(hab)} + 0,00422 \cdot Ce_{(kW.hab^{-1}.d^{-1})} + 0,0162 \cdot RPC_{(USD.hab^{-1}.d^{-1})} \quad (15)$$

Na qual:

TGP = taxa de geração *per capita* de resíduos sólidos urbano (kg.hab⁻¹.d⁻¹);

P = população total (hab.);

Ce = consumo *per capita* diário de energia elétrica (kWh.hab⁻¹.d⁻¹) em cada município do estado de São Paulo; e

RPC = renda *per capita* média diária (USD.hab⁻¹.d⁻¹).

A Equação 15 não melhorou sensivelmente a qualidade do ajuste, uma vez que o valor de RMSE (0,279 kg.hab⁻¹.d⁻¹) foi maior que os obtidos com as Equações 12 a 14 para R² próximos a 0,19; exceção feita a 0,06 fornecido pela Equação 13. No entanto, foram mostradas, nas Figuras 2 e 3, as dependências de TGP em função do logaritmo neperiano da população e do consumo de energia elétrica. Sendo assim, é possível que a função expressa no logaritmo natural de P e Ce resulte em melhor ajuste.

Correlação para Ln(P) e Ln(Ce)

Na Tabela 2 são mostrados os valores de r para as TGPs de RSU em relação às variáveis independentes P, RPC e Ce.

Nota-se, pela Tabela 2, que correlações envolvendo Ln(P), RPC e Ln(Ce) potencialmente ajustariam melhor as TGPs de RSU. Sendo assim, o método dos mínimos quadrados permitiu determinar os coeficientes presentes na Equação 6, que resultou na Equação 16:

$$TGP = -11,280 + 0,0356 \cdot \ln(P) + 0,0179 \cdot RPC + 11,347 \cdot [\ln(Ce)]^{0,0103} \quad (16)$$

Na qual:

TGP = taxa de geração *per capita* de resíduos sólidos urbano (kg.hab⁻¹.d⁻¹);

P = população total (hab.);

RPC = renda *per capita* média diária (USD.hab⁻¹.d⁻¹); e

Ce = consumo *per capita* diário de energia elétrica (kWh.hab⁻¹.d⁻¹) em cada município do estado de São Paulo.

Essa equação é válida para os seguintes intervalos: 1.923 hab ≤ P ≤ 11.446.275 hab; 6,23 USD.hab⁻¹.d⁻¹ ≤ RPC ≤ 29,89 USD.hab⁻¹.d⁻¹; e 2,51 kWh.hab⁻¹.d⁻¹ ≤ Ce ≤ 87,04 kWh.hab⁻¹.d⁻¹.

Os parâmetros de avaliação da qualidade do ajuste são mostrados na Tabela 3. Apesar do baixo valor de R² (0,24), o desvio percentual associado à previsão foi aceitável, em torno de -14% – o sinal negativo indicou uma tendência de estimar valores acima dos observados. O valor de RMSE foi da ordem de 29% (0,217 kg.hab⁻¹.d⁻¹) do valor médio observado (0,746 kg.hab⁻¹.d⁻¹). Portanto, a Equação 16 forneceu o melhor ajuste dentre as equações testadas. Nas Figuras 5 a 7 são mostradas as TGPs de RSU observadas e previstas pela Equação 16 em função das variáveis independentes.

Tabela 2 - Valores de r entre as taxas de geração *per capita* de resíduos sólidos urbanos observadas para os municípios do estado de São Paulo em relação às variáveis independentes.

	Variáveis independentes					
	P (hab)	Ln(P)	RPC (USD.hab ⁻¹ .d ⁻¹)	Ln(RPC)	Ce (kWh.hab ⁻¹ .d ⁻¹)	Ln(Ce)
r	0,12	0,41	0,43	0,42	0,21	0,32

r: Coeficiente de Pearson; P: população total de cada município; Ln(P): logaritmo natural da população total de cada município; RPC: renda *per capita* média diária; Ln(RPC): logaritmo natural da renda *per capita* média diária; Ce: consumo *per capita* diário de energia elétrica em cada município; Ln(Ce): logaritmo natural do consumo *per capita* diário de energia elétrica em cada município.

Foi possível perceber, nas Figuras 5 a 7, que melhor tendência de comportamento dos resultados era obtida ao se utilizar escala logarítmica nos eixos das variáveis P e Ce. Fato corroborado por maiores coeficientes de Pearson calculados para Ln(P), Ln(Ce) e RPC, quando comparados com os obtidos para P e Ce (Tabela 2). Além disso, percebe-se grande dispersão das TGPs, em relação aos valores observados e previstos, o que implicou em R² relativamente baixo. Há municípios com TGPs excessivamente baixas e inesperadamente elevadas. Por exemplo, existem

Tabela 3 - Parâmetros de avaliação da qualidade dos ajustes para Equação 16.

	Parâmetros de avaliação da qualidade do ajuste			
	r (-)	R ² (-)	RMSE (kg.hab ⁻¹ .d ⁻¹)	Ep (%)
Equação 16	0,490	0,240	0,217	-14,100

r: coeficiente de correlação de Pearson; R²: coeficiente de determinação; RMSE: raiz quadrada do erro quadrático médio; hab: habitantes; Ep: erro percentual médio.

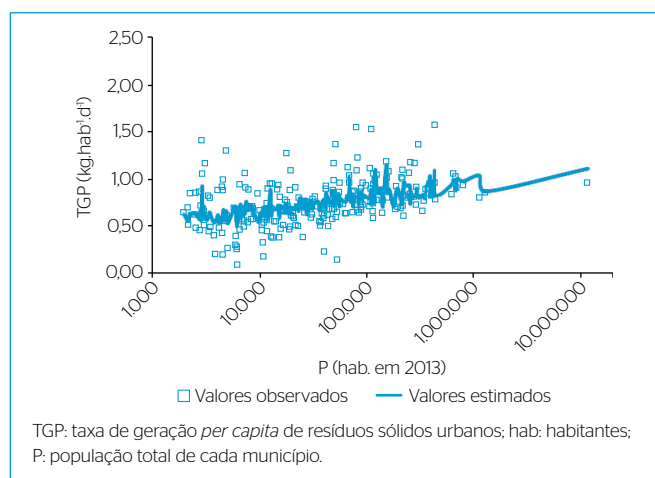


Figura 5 - Taxas de geração *per capita* de resíduos sólidos urbanos em função da população total dos municípios do estado de São Paulo que alegaram pesar rotineiramente seus resíduos.

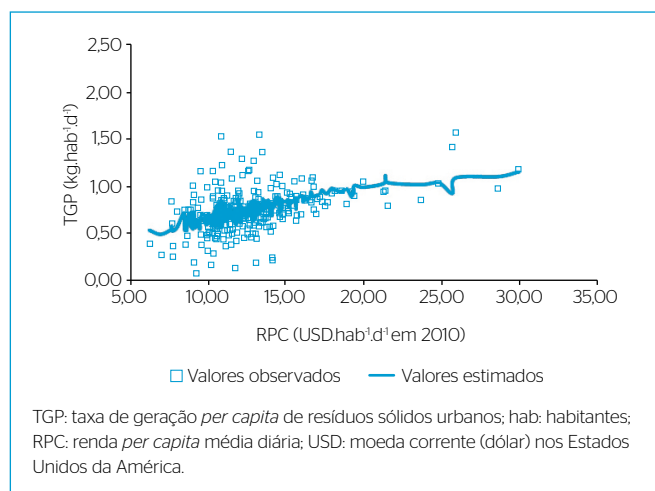


Figura 6 - Taxas de geração *per capita* de resíduos sólidos urbanos em função da renda *per capita* diária nos municípios do estado de São Paulo que alegaram pesar rotineiramente seus resíduos.

municípios com populações inferiores a 10 mil habitantes com TGP acima de $1,00 \text{ kg.hab}^{-1}.\text{d}^{-1}$, que podem ser decorrentes da disposição final inadequada de resíduos da construção civil em conjunto com RSU. Ao mesmo tempo em que foram encontradas cidades com população no intervalo de 6 a 60 mil com TGP da ordem de $0,20 \text{ kg.hab}^{-1}.\text{d}^{-1}$, que talvez sejam decorrentes de inconsistência nos dados informados pelos municípios ao SNIS. Comprovou-se também que as TGPs foram crescentes em relação às variáveis independentes, P , renda *per capita* e C_e nos municípios.

Na Figura 8 foram plotadas as TGPs observadas e previstas pela Equação 16 para se avaliar a qualidade do ajuste; quanto mais próximos os pontos forem da reta diagonal, melhor é a previsão. Nota-se, portanto, que grande parte dos pontos obedece a essa condição. Porém, percebe-se que há condições que correspondem claramente a subestimativas e superestimativas de TGPs, mostradas pelas regiões tracejadas na figura. No entanto, a maioria dos pontos encontra-se dentro do intervalo de boa previsão, ou seja, entre o valor médio e uma RMSE para mais e para menos ($0,217 \text{ kg.hab}^{-1}.\text{d}^{-1}$). Em acréscimo, percebe-se pela

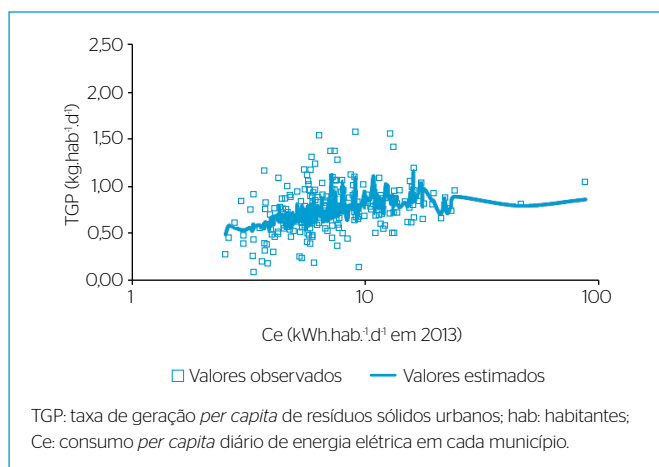


Figura 7 - Taxas de geração *per capita* de resíduos sólidos urbanos em função do consumo *per capita* de energia elétrica total nos municípios do estado de São Paulo que alegaram pesar rotineiramente seus resíduos.

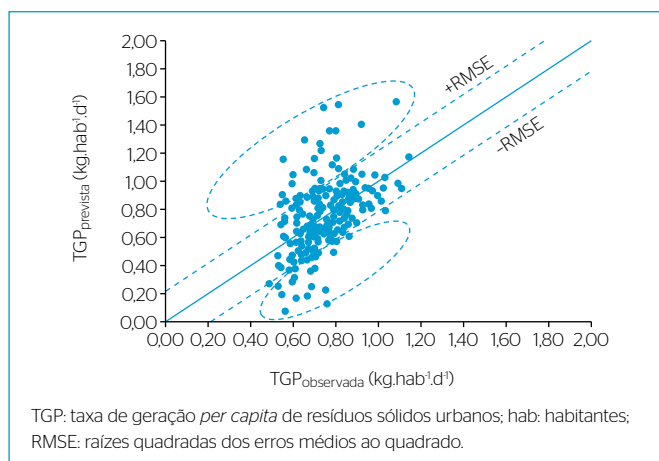


Figura 8 - Taxa de geração *per capita* de resíduos sólidos urbanos observadas e previstas pela Equação 16 em relação à diagonal.

Figura 9 a aleatoriedade da distribuição dos erros diante das TGPs previstas pela Equação 16 e que não ocorreu erro sistemático na previsão.

Apesar de R^2 igual a 0,24, o teste de t de Student mostrou que a Equação 16 foi capaz de representar valores médios de TGP com probabilidade acima de 99%; o valor de t crítico bicaudal foi de 2,60; e o valor de t do conjunto de dados foi de 1,00, fato que comprovou que as médias foram praticamente iguais. O mesmo pôde ser afirmado quanto a presumir variâncias equivalentes entre os dois conjuntos de resultados (observados e previstos para TGP), uma vez que t crítico bicaudal foi de 2,59; enquanto o valor de t do conjunto de dados foi de 1,00. Portanto, a equação pôde ser considerada adequada para prever as TGPs nos municípios do estado de São Paulo.

CONCLUSÕES

A partir da realização do presente estudo foi possível concluir que:

- Os 238 municípios que formaram o espaço amostral da pesquisa foram capazes de representar o estado de São Paulo (645 municípios em 2013) com uma margem de erro máxima de 5,1% e intervalo de confiança mínimo de 94,9%;
- As variáveis independentes P , RPC e C_e se mostraram interferentes na TGP de RSU, sobretudo ao se considerar $\ln(P)$ (r de 0,41), RPC (r de 0,43) e $\ln(C_e)$ (r de 0,32), coeficientes que são compatíveis com considerável grau de correlação entre as variáveis dependentes e independentes;
- A dependência de TGP em função das variáveis independentes (P , C_e e RPC) mostrou possuir apenas um domínio, não sendo necessário subdividir as equações testadas em diferentes intervalos de validade;
- Foi constatada grande dispersão dos resultados observados para as TGPs para os municípios de informaram pesar rotineiramente os seus resíduos no estado de São Paulo, cujo valor máximo foi de $1,57 \text{ kg.hab}^{-1}.\text{d}^{-1}$ e mínimo de $0,080 \text{ kg.hab}^{-1}.\text{d}^{-1}$, que se refletiu em baixos coeficientes de determinação para as equações testadas (R^2 no intervalo de 0,06 a 0,24);

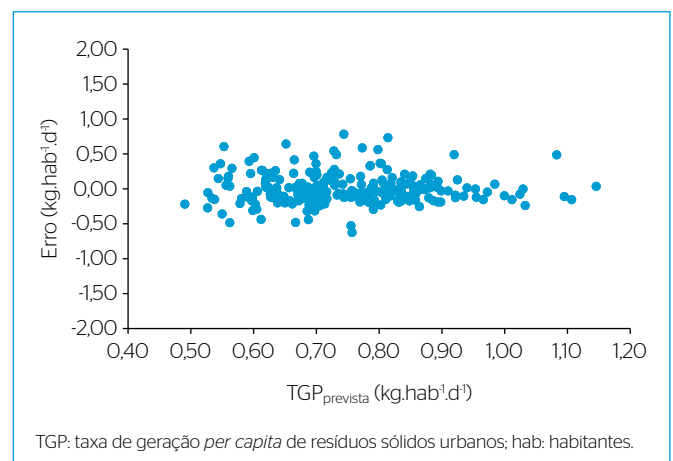


Figura 9 - Desvios entre as taxas de geração *per capita* de resíduos sólidos urbanos previstas e estimadas em função dos valores previstos.

- Correlações lineares envolvendo as variáveis P, RPC e Ce não foram satisfatórias para prever as TGP's observadas;
- Porém, ao se considerar a função envolvendo variáveis na forma de Ln (P), Ln (Ce) e RPC, o E_p foi de -14,1% e os valores observados de TGP's estiveram predominantemente no intervalo compreendido entre os valores médios e as RMSE para mais e para menos, o que confere segurança à previsão realizada para a maioria dos casos;
- Os testes *t* de Student realizados para verificar a hipótese de médias iguais pareadas e variâncias equivalentes entre os valores previstos e observados de TGP's comprovaram a validade da equação proposta (Equação 16) com intervalo de confiança acima de 99%.

REFERÊNCIAS

- ABDULLAH, W.S. (2009) Exponential model for predicting the sorption isotherms-evaluation, statistical appraisal and comparison with existing models. *Jordan Journal of Civil Engineering*, v. 3, n. 4, p. 344-355.
- ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE LIMPEZA PÚBLICA E RESÍDUOS ESPECIAIS (ABRELPE). (2014) *Panorama dos resíduos sólidos no Brasil 2013*. São Paulo: ABRELPE. 112 p.
- AYRES, M. (2007) *BioEstat: aplicações estatísticas nas áreas das ciências biomédicas*. 5. ed. Belém: Universidade Federal do Pará. 364p.
- BENÍTEZ, S.O.; LOZANO-OLVERA, G.; MORELOS, R.A.; VEGA, C.A. (2008) Mathematical modeling to predict residential solid waste generation. *Waste Management*, v. 28, p. S7-S13. <https://doi.org/10.1016/j.wasman.2008.03.020>
- BRASIL. (2010a) *Decreto nº 7.404*. Brasília: Presidência da República.
- _____. (2010b) *Lei de nº 12.305: Política Nacional de Resíduos Sólidos (PNRS)*. Brasília: Presidência da República.
- COHEN, J. (1988) *Statistical power analysis for the behavioral sciences*. 2. ed., Hillsdale: Lawrence Earlbaum Associates.
- COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO (CETESB). (2015) *Inventário Estadual de Resíduos Sólidos Domiciliares 2014*. São Paulo: Cetesb. 124p.
- DANCEY, C.; REIDY, J. (2006) *Estatística sem matemática para psicologia: usando SPSS para Windows*. Porto Alegre: Artmed. 608p.
- DIAS, D.M.; MARTINEZ, C.B.; BARROS, P.T.V.; LIBÂNIO, M. (2012) Modelo para estimativa da geração de resíduos sólidos domiciliares em centros urbanos a partir de variáveis socioeconômicas conjunturais. *Engenharia Sanitária Ambiental*, v. 17, n. 3, p. 325-332. <http://dx.doi.org/10.1590/S1413-41522012000300009>
- FUNDAÇÃO SISTEMA ESTADUAL DE ANÁLISE DE DADOS (FUNDAÇÃO SEADE). (2013) *Portal*. Disponível em: <<http://www.seade.gov.br/>>. Acesso em: 20 maio 2016.
- HOCKETT, D.; LOBER, D.J.; PILGRIM, K. (1995) Determinants of *per capita* municipal solid waste generation in the Southeastern United States. *Journal of Environmental Management*, v. 45, p. 205-217. <https://doi.org/10.1006/jema.1995.0069>
- ISRAEL, G.D. (1992) *Determining sample size*. Fact sheet PEOD-6. Flórida: Florida Cooperative Extension Service, University of Florida. p. 1-5.
- KESER, S.; DUZGUNB, S.; AKSOY, A. (2012) Application of spatial and non-spatial data analysis in determination of the factors that impact municipal solid waste generation rates in turkey. *Waste Management*, v. 32, n. 3, p. 359-371. DOI: 10.1016/j.wasman.2011.10.017
- LEGATES, D.R.; MCCABE JR., G.J. (1999) Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, v. 35, n. 1, p. 233-241. DOI: 10.1029/1998WR900018
- MORIASI, D.N.; ARNOLD, J.G.; LIEW, M.W.V.; BINGNER, R.L.; HARMEL, R.D.; VEITH, T.L. (2007) Model evaluation guidelines for systematic Quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers*, v. 50, n. 3, p. 885-890. DOI: 10.13031/2013.23153
- NAVARRO, E.J.; DIAMADOPOULOS, E.; GINESTAR, D. (2002) Time series analysis and forecasting techniques for municipal solid waste management. *Resources, Conservation and Recycling*, v. 35, p. 201-214. DOI: 10.1016/S0921-3449(02)00002-2
- PINTO, M.R.; PEREIRA, D.R.M.; FREITAS, R.C. (2012) Fatores sociais, econômicos e demográficos associados à geração de lixo domiciliar na cidade de Belo Horizonte. *Reuna*, Belo Horizonte, v. 17, n. 2, p. 27-44.
- SECRETARIA NACIONAL DE SANEAMENTO AMBIENTAL. (2015) *Sistema Nacional de Informações sobre Saneamento: diagnóstico do manejo de resíduos sólidos urbanos 2013*. Ministério das Cidades. 154p.
- SILVA, H.; BARBIERI, A.F.; MÓR, R.L.M. (2012) Fatores demográficos e geração de resíduos sólidos domiciliares no município de Belo Horizonte. *Revista Brasileira de Estudos de População*, v. 29, n. 2, p. 421-449.
- TIBONI, C.G.R. (2010) *Estatística Básica para cursos de administração, ciências contábeis, tecnológicos e de gestão*. São Paulo: Atlas. 344p.
- XU, L.; GAO, P.; CUI, S.; LIU, C. (2013) A hybrid procedure for MSW generation forecasting at multiple time scales in Xiamen City China. *Waste Management*, v. 33, p. 1324-1331. DOI: 10.1016/j.wasman.2013.02.012
- YAFEE, R.; MCGEE, M. (2000) *Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS*. Orlando: Academic Press. 509p.

