

MODELOS DE REGRESSÃO LINEAR MÚLTIPLA APLICADOS À AVALIAÇÃO DE TERRENOS URBANOS - CASO DO MUNICÍPIO DE ITAQUI-RS

Multiple Linear Regression Models Applied to Urban Terrain Assessment - Municipality of Itaquí - RS Case

Jaqueline Prunzel ¹

Marcos Toebe ²

Alexandre Bernardino Lopes²

Virnei Silva Moreira²

Programa de pós-graduação em Ciências Geodésicas, Departamento de Geomática, Universidade Federal do Paraná – UFPR, Curitiba, PR, Brasil.¹

Universidade Federal do Pampa, Campus Itaquí, Departamento de Engenharia, Itaquí, RS, Brasil. ²

Email: jaquelineprunzel@gmail.com, m.toebe@gmail.com, ablopesrp@yahoo.com.br,
virneimoreira@gmail.com

Resumo:

Este trabalho teve como objetivo identificar variáveis explicativas para a modelagem e predição do valor real (ITBI) e venal de lotes urbanos no município de Itaquí-RS utilizando o Método Comparativo Direto de Dados do Mercado, desenvolvido por regressão linear múltipla. Em uma amostra de 67 lotes foram avaliadas variáveis quantitativas: área, testada, valor do m², valor venal e valor do ITBI e variáveis qualitativas: topografia, situação e pavimentação, sendo que para cada variável foram determinados parâmetros estatísticos incluindo a distribuição de dados por assimetria, curtose, teste de normalidade de Kolmogorov-Smirnov, bem como a correlação linear de Pearson e análise de trilha, que mostra as relações de causa e efeito das variáveis explicativas sobre o valor venal e ITBI. A importância de cada variável explicativa sobre o modelo de valor venal e ITBI foi determinada pelo método *stepwise*. O valor venal apresentou correlação positiva, significativa ($p \leq 0,05$) e crescente com a pavimentação, testada, área e valor do m², nessa ordem. Já o valor de ITBI apresentou correlação positiva, significativa ($p \leq 0,05$) e crescente com valor do m² e pavimentação, nessa ordem. A análise de trilha indicou que a área e o valor do m² possuem relação de causa e efeito sobre o valor venal e o valor do m² e a pavimentação apresenta relação de causa e efeito sobre o valor de ITBI. Os modelos: $Y = 16,90 \times \text{área} + 594,69 \times \text{valor do m}^2 - 9526,54 \times \text{topografia}$ ($r^2 = 0,91$) e, $Y = 868,36 \times \text{valor do m}^2 - 23642,19 \times \text{topografia} + 33633,46 \times \text{pavimentação}$ ($r^2 = 0,70$), são adequados para a estimação dos valores venal e de ITBI, respectivamente.

Palavras-chave: Análise de Trilha; Associação linear; Terrenos; Prefeitura municipal de Itaquí-RS.

Abstract:

This work had as objective to identify explanatory variables for modeling and prediction of actual value (ITBI) and venal value of urban plot in the city of Itaqui-RS using the comparative method of Direct market data, developed by multiple linear regression. In a sample of 67 lots, quantitative variables were evaluated: area, front of urban plot, value of m², market value and the value ITBI and qualitative variables: topography, location and flooring, and for each variable were determined statistical parameters including the data distribution for asymmetry, kurtosis, normality test of Kolmogorov-Smirnov, as well as the Pearson linear correlation and path analysis, which shows the relationships of cause and effect of the explanatory variables on the market and ITBI values, respectively. The importance of each explanatory variable on the model of market and ITBI values was determined by the stepwise method. The market value showed a positive, significant correlation ($p \leq 0.05$) and increased with flooring, front of urban plot, area and value of m², in that order. Already ITBI value was positively correlated, significant ($p \leq 0.05$) and with increasing value of m² and flooring, in that order. Path analysis indicated that the area and the value of m² have cause and effect on the market value and, the value of m² and flooring shows cause and effect on the value of ITBI. The models: $Y = 16.90 \times \text{area} + 594.69 \times \text{value of m}^2 - 9526.54 \times \text{topography}$ ($r^2 = 0.91$) and, $Y = 868.36 \times \text{value of m}^2 - 23642.19 \times \text{topography} + 33633.46 \times \text{flooring}$ ($r^2 = 0.70$), are suitable for the estimation of market and ITBI values, respectively.

Keywords: Path analysis; Linear association; Terrain; City Hall of Itaqui – RS

1. Introdução

A engenharia de avaliações e as informações oriundas do Cadastro Técnico Multifinalitário (CTM) são essenciais para a justa cobrança de impostos municipais. Apesar da importância dos referidos mecanismos, sua implantação é um desafio para os municípios de pequeno e médio porte, devido à falta de mão de obra especializada e de recursos financeiros e/ou tecnológicos. No caso do município de Itaqui-RS a Planta Genérica de Valores (PGV) não é atualizada desde 1976 e os métodos que foram utilizados não estão à disposição do setor de cadastro do município.

A correta avaliação dos imóveis urbanos é um requisito básico para a implantação de uma planta de valores, influenciando diretamente na arrecadação tributária do município. A avaliação deve ser feita em função das diferentes características do imóvel, considerando forma, dimensão, topografia, valor de mercado e localização geográfica, seguindo as normas técnicas.

Conforme o Ministério das Cidades, na aplicação do Método Comparativo Direto (MCD), a Análise de Regressão Múltipla prevista na NBR 14653-2, é aplicada para a identificação dos fatores que influenciam a determinação de preços, bem como o valor de mercado.

No Brasil, a utilização da regressão linear multivariada aplicada à avaliação de terrenos urbanos não é uma realidade nos municípios de pequeno e médio porte, embora sua aplicação seja discutida em trabalhos acadêmicos como por (Baptistella, 2005), que usou dados cadastrais na comparação entre os métodos de redes neurais e regressão linear múltipla na avaliação de

imóveis. Estudo realizado por (Alves, 2005) conclui que a análise multivariada de imóveis apresenta um bom ajuste de dados e uma boa capacidade preditiva. Entretanto ajustes e análises devem ser aplicadas de modo a evitar problemas de multicolinearidade. Em trabalho similar, (Nguyen & Cripps, 2001) aponta as causas das diferenças dos resultados preditivos com a utilização de regressão linear multivariada e de redes neurais.

Para o estudo das associações existentes entre variáveis, pode-se utilizar inicialmente o coeficiente de correlação linear de Pearson (r), que mede o sentido e a intensidade da relação linear entre duas variáveis aleatórias (Ferreira, 2009). O sentido da correlação pode ser positivo ou negativo, no intervalo de $-1 \leq r \leq 1$, sendo que a intensidade da correlação linear será maior, quanto mais próximo a $|1|$ for o r . Quando um conjunto de variáveis é estudado simultaneamente, recomenda-se a utilização de técnicas complementares às correlações lineares (Cruz & Carneiro, 2006). Embora a regressão múltipla e redes neurais já tenham sido utilizadas na predição dos valores venais de terrenos, recomenda-se a utilização da análise de trilha como técnica complementar aos coeficientes de correlação linear, por permitir o desdobramento dos coeficientes de correlação em efeitos diretos e indiretos sobre a variável principal (Cruz & Carneiro 2006), (Wright, 1921). A análise de trilha permite estabelecer qual estratégia será mais eficiente na seleção, para incrementar o modelo (Souza, 2013). Variáveis com alta associação medida por meio do coeficiente de correlação linear e, com efeitos diretos de mesma intensidade e direção, são consideradas variáveis de causa e efeito, sendo indicadas para a predição da variável principal (Cruz & Carneiro, 2006), (Wright, 1921).

A avaliação de métodos que visem homogeneizar bem como diminuir o carácter subjetivo das avaliações é importante para municípios de pequeno e médio porte, já que estes estão sujeitos a especulações imobiliárias em função de fatores econômicos pontuais, não apresentando assim qualquer procedimento científico nas avaliações de mercado (Steiner et al, 2008). No caso do município objeto do estudo, a implantação de uma Universidade pública causou um aumento significativo de preços imobiliários em bairros específicos, sem critério definido. As análises dos critérios que influenciam no valor dos imóveis podem ser indicativas de locais de investimento imobiliário.

Em consonância com os apontamentos acima, o trabalho tem como objetivo analisar as variabilidades do valor venal e de ITBI dos imóveis (terrenos) do município de Itaquí em função de variáveis quantitativas (área, testada e valor do m^2) e qualitativas (topografia, situação, pedologia e pavimentação). Os objetivos específicos são: 1) aplicar o MCDDM desenvolvido a partir de regressão linear múltipla; 2) interpretar e avaliar os resultados considerando a relação existente entre as variáveis que podem influenciar no valor venal considerado pela prefeitura afim de tributos e no valor de compra e venda que está sendo aplicado no município; 3) analisar a aplicabilidade da análise de trilha e da regressão linear múltipla no contexto do trabalho.

2. Materiais e métodos

2.1 Classificação das variáveis

O MCDDM desenvolvido a partir da regressão linear múltipla, no qual visa obter uma função objetiva, explica a relação entre as variáveis dependentes (resposta) e independentes (características do terreno) que podem ser do tipo quantitativo: 1) área do lote e 2) índice de localização (valor oriundo da planta de valores) e qualitativa: 3) Topografia (irregular, declive, aclave ou plana); 4) Pedologia (alagado, inundável, combinação ou seco); 5) Situação (quantidade de frentes, posição do imóvel na quadra); 6) Testada (indica a metragem da frente do lote) e 7) Pavimentação (terra, calçamento irregular, calçamento regular, asfalto), como descrito por (Heil, 2010).

2.2 Coleta dos dados

A metodologia empregada está limitada a uma única tipologia, no caso, os terrenos urbanos de alguns loteamentos do município de Itaqui. Os dados coletados foram apenas aqueles que estão disponíveis na base de dados do setor de Cadastro Imobiliário Municipal da Prefeitura. O ponto de partida para o início do trabalho foi à escolha dos lotes, sendo coletados 67 lotes a partir das guias de imposto sobre a transmissão de bens imóveis (ITBI), baseados na compra e venda de imóveis no município referente ao período de janeiro até maio de 2015. Simultaneamente foram apurados os valores venais dos mesmos, calculados atualmente pela prefeitura. É imprescindível destacar que os dados dos proprietários não foram expostos, a fim de garantir a privacidade das informações fiscais dos contribuintes. Após a coleta desses dados junto a Prefeitura Municipal de Itaqui, cuja população é de 38.000 habitantes (IBGE, 2015), foram feitas visitas *in loco* para avaliar as condições físicas dos terrenos escolhidos.

A distribuição dos imóveis comercializados no período e que foram objeto do presente estudo estão destacados pelos respectivos códigos e pelos círculos vermelhos e pretos na Figura 1. A maior concentração de pontos na região sudeste do mapa indica maior comercialização de terrenos nessa região, que fica nas proximidades de um campus da Universidade Federal do Pampa, implantado recentemente na localidade.

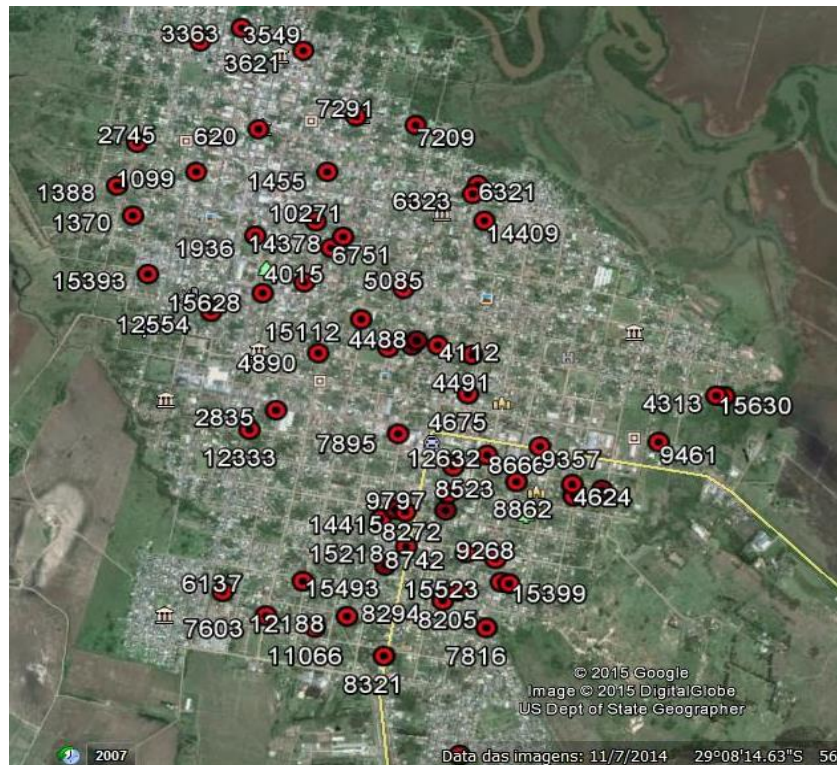


Figura 1: Localização dos imóveis avaliados no presente trabalho.

Fonte: Google Earth, 2014.

2.3 Análise de dados

Para cada variável avaliada (área - em m², testada - em m, valor do m² - em reais, topografia, pavimentação, valor venal - em reais e, valor ITBI - em reais) foram calculados os valores mínimo, máximo, média, mediana, erro padrão, variância, coeficiente de variação e verificada a distribuição dos dados através da assimetria, da curtose e do teste de normalidade de Kolmogorov-Smirnov, com base nas 67 amostras. Também foram construídos histogramas de frequência e gráficos de dispersão entre as sete variáveis, sendo posteriormente calculados os coeficientes de correlação linear de Pearson (r) entre os pares de variáveis e, verificada a significância do r por meio do teste t de *Student*, a 5% de probabilidade. A seguir, para determinar o efeito direto e indireto de cada variável explicativa (área - em m², testada - em m, valor do m² - em reais, topografia, pavimentação) sobre as variáveis principais (valor venal - em reais e, valor ITBI - em reais), em relação de causa e efeito, foram realizadas análises de trilha separadamente, para o valor venal e para o valor de ITBI. Essas análises de trilha foram realizadas em diagrama de única cadeia com cinco variáveis explicativas Figura 2, conforme descrito na literatura (Cruz & Carneiro, 2006), (Wright, 1921), com utilização do sistema de equações normais $X'X\beta = X'Y$. Antes da realização da análise de trilha, foi realizado o diagnóstico de multicolinearidade entre as variáveis explicativas. Para isso, utilizou-se como método de diagnóstico, o número de condição (NC), que é a divisão do maior pelo menor autovalor da matriz de correlação $X'X$ (Gujarati, 2006).

Considerou-se multicolinearidade fraca quando $NC < 100$, multicolinearidade moderada a severa quando $100 \leq NC \leq 1.000$ e, multicolinearidade severa quando $NC > 1.000$ (Montgomery & Peck,

1982). Caso a multicolinearidade foi moderada a severa ou severa, eliminou-se variáveis explicativas altamente correlacionadas, conforme sugerido por (Toebe & Cargnelutti Filho, 2013), (Fávero et al., 2009) e (Hair et al., 2009).

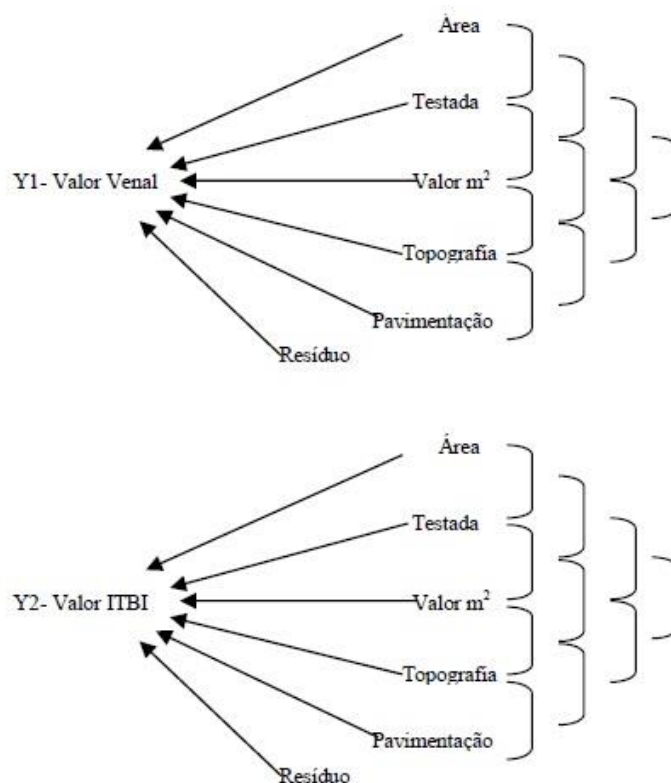


Figura 2: Diagrama causal de única cadeia, com os efeitos diretos das variáveis explicativas e da variável residual (Resíduo) sobre: (Y1) o valor venal, e (Y2) valor com base no ITBI e as inter-relações entre as variáveis explicativas na análise de trilha.

Posteriormente à análise de trilha, foi realizada a regressão linear múltipla pelo método *stepwise* para o valor venal e valor de ITBI, separadamente. Para isso, foi utilizada a seguinte equação:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (1)$$

Onde: Y = Variável dependente: valor venal do terreno ou valor do ITBI; x_n = n-ésima variável independente; β_n = n-ésimo coeficiente ou parâmetro do modelo; ε = efeito da variável residual. Na regressão linear multivariada, a estimação dos parâmetros é feita por meio do método mínimos quadrados. Assim estimam-se os valores dos parâmetros β_i para $i=1, 2, 3, \dots, n$, os quais minimizam a soma dos quadrados das distâncias (SQD), buscando diminuir ao máximo a quantidade da variável aleatória " ε " existente no processo (Heil, 2010). Como foi realizado o procedimento *stepwise*, foram mantidas no modelo de regressão múltipla, apenas variáveis com coeficientes estatisticamente significativos ($p \leq 0,05$). As análises estatísticas foram realizadas com auxílio do programa GENES (Cruz & Genes, 2013), do software Statistica 7.0® (STATSOFT, 2005) e do aplicativo Microsoft Office Excel®.

3. Resultados e discussões

As estatísticas descritivas para as variáveis explicativas: Área do terreno (m²); Testada (m); Valor do metro quadrado (em reais); Topografia, e Pavimentação, assim como as variáveis dependentes consideradas: Y1 – Valor venal obtido pela prefeitura (em reais) e Y2 – Valor com base nas guias de ITBI (em reais) são apresentadas no Quadro 1. As variáveis Pedologia e Situação foram avaliadas nos 67 terrenos, mas excluídas na análise desse trabalho por não apresentarem variabilidade entre os terrenos avaliados. Em média, os terrenos apresentam 575,17m², com testada de 14,32m, valor de 18,21 reais por m², valor venal de 10478,71 reais e valor de ITBI de 33026,87 reais, Quadro 1. Com exceção da topografia, todas as demais variáveis explicativas e dependentes apresentaram coeficiente de variação muito alto ($33,78\% \leq CV \leq 110,50\%$), conforme classificação de (Pimentel, 2009). Esses dados indicam ampla variabilidade entre os 67 terrenos, o que pode ser justificado pela não estratificação da amostra por bairros. Dessa forma, foram contemplados nesse estudo, terrenos de todas as regiões da cidade, com ampla oscilação das características avaliadas. Essa variabilidade é importante na construção de modelos representativos, com aplicação genérica, independentemente das condições de localização de cada terreno dentro do município.

Quadro 1: Estatísticas descritivas com base na avaliação de 67 terrenos do município de Itaquiraçu-RS.

Estatística	Variável e unidade de mensuração						
	Área	Testada	Valor do m ²	Topografia	Pavimentação	Valor Venal	Valor ITBI
	m ²	m	R\$	Índice	Índice	R\$	R\$
Mínimo	217,00	7,00	4,07	0,70	0,50	1338,89	1000,00
Máximo	1254,00	33,00	41,52	1,00	2,00	43845,12	200000,00
Média	575,17	14,32	18,21	0,97	0,90	10478,71	33026,87
Mediana	464,00	13,00	8,23	1,00	1,00	7036,65	23300,00
Erro padrão	35,59	0,59	1,74	0,01	0,06	1285,60	4458,53
Variância	84854,82	23,40	201,80	0,01	0,27	110734771	1331860479
CV (%)	50,65	33,78	78,01	8,42	57,22	100,42	110,50
Curtose+3 ⁽¹⁾	2,02 ^{ns}	7,66*	1,63*	9,85*	3,43 ^{ns}	5,44*	13,71*
Assimetria ⁽²⁾	0,67*	1,77*	0,58*	-2,83*	1,25*	1,72*	2,99*
P-valor do teste de normalidade	0,03	0,04	0,00	0,00	0,00	0,01	0,00

⁽¹⁾*Curtose difere de três, pelo teste t, em nível de 5% de probabilidade. ^{ns} = Não-significativo. ⁽²⁾*Assimetria difere de zero, pelo teste t, em nível de 5% de probabilidade. ^{ns} = Não-significativo.

Em relação à curtose foram verificados desvios significativos ($p \leq 0,05$) e positivos (leptocúrtica) para testada, topografia, valor venal e valor de ITBI, indicando elevada concentração de valores na região central do banco de dados. Já para o valor do m² foi verificada distribuição platicúrtica com maior concentração de valores extremos. Com exceção da topografia, todas as demais variáveis apresentaram assimetria positiva, o que indica alguns terrenos muito grandes, com elevada testada, alto valor do m², alto valor venal e valor ITBI, contribuindo para elevação do valor médio em relação à mediana. Os desvios de assimetria e

curtose Figura 3, associados ao elevado número de observações ($n = 67$ terrenos) contribuíram para o não ajuste a distribuição normal dos resíduos ($p \leq 0,05$) de todas as variáveis explicativas avaliadas. Embora a não-normalidade dificulte a aplicação de métodos paramétricos de análise de dados, destaca-se que de acordo com o Teorema do Limite Central, mesmo que os dados não se ajustem a distribuição normal, a distribuição amostral dos estimadores terá distribuição aproximadamente normal quando o tamanho de amostra for grande ($n \geq 30$ observações).

Com base nos gráficos de dispersão Figura 3 e nos coeficientes de correlação Quadro 2 pode-se verificar que as variáveis área, testada, valor do m^2 e pavimentação apresentam coeficientes significativos e positivos com o valor venal. Já o valor de ITBI apresentou correlação positiva, significativa ($p \leq 0,05$) e crescente com valor do m^2 e pavimentação, nessa ordem. Entre as variáveis explicativas foi constatada correlação positiva e significativa entre área e testada e entre valor do m^2 e pavimentação. Ainda, foi constatada correlação negativa entre testada e topografia. As correlações entre os demais pares de variáveis explicativas foram não significativas, indicando baixa colinearidade entre as mesmas (Gujarati, 2006), (Hair et al., 2009) e (Fávero et al., 2009). A correlação entre valor venal e valor do ITBI ($r = 0,562^*$), embora significativa, indica uma correlação moderada entre os dois cálculos de valores dos terrenos. Dessa forma, é provável que a utilização das variáveis explicativas e os seus coeficientes sejam dependentes do tipo de valor que é considerado na tributação (venal ou de ITBI).

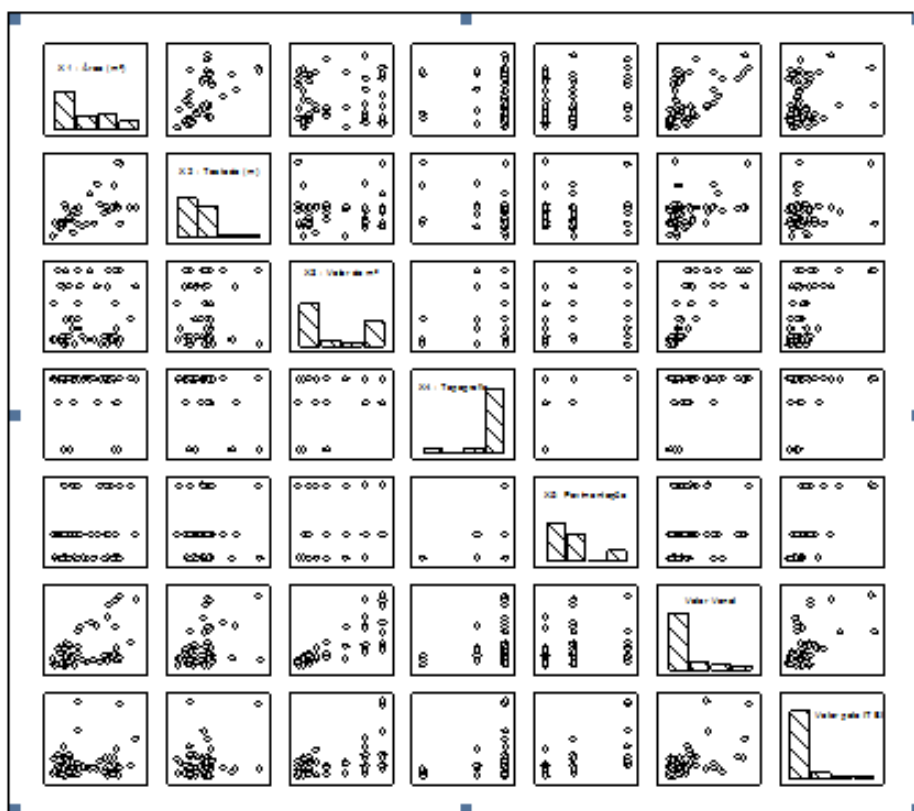


Figura 3: Matriz com o histograma de frequência (na diagonal) e gráficos de dispersão entre as cinco variáveis explicativas e as duas variáveis dependentes avaliadas em 67 terrenos do município de Itaquí - RS.

Quadro 2: Estimativas dos coeficientes de correlação linear de Pearson (r) entre as cinco variáveis explicativas e as duas variáveis dependentes avaliadas em 67 terrenos do município de Itaquí - RS.

Variável	Área	Testada	Valor do m ²	Topografia	Pavimentação	Valor Venal	Valor ITBI
Área	1,000	0,638*	-0,017	-0,084	0,151	0,470*	0,192
Testada	0,638*	1,000	0,128	-0,278*	0,108	0,440*	0,211
Valor do m ²	-0,017	0,128	1,000	0,199	0,372*	0,778*	0,512*
Topografia	-0,084	-0,278*	0,199	1,000	0,236	0,143	0,161
Pavimentação	0,151	0,108	0,372*	0,236	1,000	0,333*	0,614*
Valor Venal	0,470*	0,440*	0,778*	0,143	0,333*	1,000	0,562*
Valor ITBI	0,192	0,211	0,512*	0,161	0,614*	0,562*	1,000

(1)* Coeficiente de correlação linear de Pearson (r) significativo a 5% de probabilidade, pelo teste t, com 65 graus de liberdade.

Foram realizadas duas análises de trilha (*path analysis*) por meio do sistema de equações normais $X'X\beta = X'Y$, conforme proposto por (Cruz & Carneiro, 2006) e (Wright, 1921), sendo que na primeira foi considerada como variável dependente, o valor venal e na segunda, o valor com base no ITBI. Estas análises têm por objetivo fracionar as correlações em efeitos diretos e indiretos, como se podem observar a partir dos diagramas causais as variáveis explicativas sobre a variável principal Figura 2. Com base no número de condição do modelo, de $NC = 5,9337$, verificou-se baixo grau de multicolinearidade entre as variáveis explicativas, não sendo necessária a exclusão de nenhuma variável explicativa para a realização da análise de trilha.

Nota-se que para o valor venal, as maiores correlações foram com valor do m², área, testada e pavimentação, respectivamente Quadro 2. A testada apresentou pouco efeito direto sobre o valor venal e efeitos indiretos via área e valor do m² Quadro 3. Já a pavimentação apresentou efeito indireto via valor do m² e, a área e o valor do m² apresentaram efeitos diretos de magnitudes similares aos coeficientes de correlação com o valor venal. Logo, pode-se afirmar que existe relação de causa e efeito entre valor do m² e valor venal, o que era esperado devido ao método utilizado pela Prefeitura para o cálculo do valor venal. E em menor magnitude, entre área e valor venal. A pavimentação e a testada atuam indiretamente via valor do m² e via área do terreno. Dessa forma, com base na análise de trilha, para a obtenção do valor venal devem ser avaliadas a área e o valor do m².

Quadro 3: Efeitos diretos e indiretos de variáveis explicativas (X1 - Área do terreno, em m²; X2 - Testada, em m; X3 - Valor do m², em reais; X4 - Topografia; X5 - Pavimentação) sobre o valor venal e o valor com base no ITBI, e coeficiente de determinação da análise de trilha (R²) com base em dados de 67 terrenos do município de Itaquí - RS.

Análise de trilha com efeitos diretos e indiretos sobre valor venal					
	Área	Testada	Valor do m ²	Topografia	Pavimentação
Efeito direto	0,447	0,076	0,782	0,058	-0,047
Efeito Indireto via Área	-	0,285	-0,008	-0,037	0,067
Efeito Indireto via Testada	0,049	-	0,010	-0,021	0,008
Efeito Indireto via Valor do m ²	-0,013	0,100	-	0,155	0,291
Efeito Indireto via Topografia	-0,005	-0,016	0,011	-	0,014
Efeito Indireto via Pavimentação	-0,007	-0,005	-0,018	-0,011	-
Correlação	0,470*	0,440*	0,778*	0,143	0,333*
R ² da Análise de Trilha	0,845				
Análise de trilha com efeitos diretos e indiretos sobre valor com base no ITBI					
	Área	Testada	Valor do m ²	Topografia	Pavimentação
Efeito direto	0,085	0,067	0,328	0,011	0,470
Efeito Indireto via Área	-	0,054	-0,001	-0,007	0,013
Efeito Indireto via Testada	0,043	-	0,009	-0,019	0,007
Efeito Indireto via Valor do m ²	-0,006	0,042	-	0,065	0,122
Efeito Indireto via Topografia	-0,001	-0,003	0,002	-	0,003
Efeito Indireto via Pavimentação	0,071	0,051	0,175	0,111	-
Correlação	0,192	0,211	0,512*	0,161	0,614*
R ² da Análise de Trilha	0,489				

Para o valor com base no ITBI, as correlações de maior intensidade foram verificadas entre pavimentação e ITBI, entre valor do m² e ITBI Quadro 2. Com base na análise de trilha, constatou-se que o valor do m² apresenta efeito direto sobre o valor de ITBI e indireto sobre a pavimentação Quadro 3. A pavimentação também atua diretamente sobre o valor de ITBI e indiretamente sobre o valor do m². Logo, para o cálculo do valor de ITBI, pode-se inferir que a pavimentação e o valor do m² apresentam relação de causa e efeito e devem ser mantidos nos modelos de predição.

Para a equação de regressão linear múltipla foi utilizado o método *stepwise*. Este tem por objetivo selecionar as variáveis que mais influenciam a variável dependente, desta forma, pode-se diminuir o número de variáveis a compor a equação de regressão. O procedimento de seleção de variáveis é realizado de forma iterativa, adicionando e removendo variáveis, a partir de um critério de seleção. Um dos critérios de seleção mais usados é o teste F, o qual foi adotado neste trabalho com 5% de significância (Alves, 2013).

As correlações lineares e a análise de trilha indicaram que a área do terreno e o valor de m² tem relação de causa e efeito sobre o valor venal Quadros 2 e 3. A partir dos modelos de predição encontrados pelo método *stepwise*, pode-se observar que na equação do valor venal essas duas

variáveis permanecem no modelo Quadro 4. Conforme já supracitadas, a testada e a pavimentação atuam indiretamente e, portanto, não foram significativas no modelo de regressão múltipla para o valor venal. Já a topografia foi considerada significativa no modelo de predição, embora apresente reduzida correlação com o valor venal. Como essa variável apresenta reduzida colinearidade com as demais variáveis explicativas é provável que a significância no modelo esteja associada a reduzida variância conjunta e a explicação de porção adicional em relação à resultante da utilização de apenas área e valor do m². Assim, sugerem-se estudos adicionais para a verificação da necessidade ou não de avaliação da topografia. Com base nos dados desse estudo, o modelo de predição do valor venal deve ser:

$$Y=16,90 \times \text{área} + 594,69 \times \text{valor do m}^2 - 9526,54 \times \text{topografia}, \text{ com } r^2=0,91 \quad (2)$$

Para o valor de ITBI, a pavimentação e o valor do m² tem maior correlação com o valor de ITBI Quadro 2 e ambos apresentam efeitos diretos e indiretos sobre o valor de ITBI Quadro 3, indicando relação de causa e efeito. Essas duas variáveis explicativas foram significativas também no modelo de regressão múltipla e devem ser consideradas na avaliação do valor de ITBI. Também foi constatado que a variável topografia foi significativa nesse modelo, podendo a significância ser atribuída às mesmas justificativas acima descritas no modelo de valor venal. No entanto, o elevado erro padrão desse coeficiente, associados as reduzidas correlações e efeitos diretos, indicam que a utilização da topografia nesses estudos deve ser vista com cautela e demanda novos estudos de validação. Com base nos dados do presente trabalho, para a modelagem do valor de ITBI, recomenda-se o modelo:

$$Y=868,36 \times \text{valor do m}^2 - 23642,19 \times \text{topografia} + 33633,46 \times \text{pavimentação}, \text{ com } r^2=0,70 \quad (3)$$

Ressalta-se, contudo, que para uma melhor análise conclusiva dos resultados, os modelos necessitam de validação (Souza, 2013) em nível local e em outras regiões para utilização ampla desses modelos.

Quadro 4: Modelos de predição do valor venal (Y1) e do valor com base no ITBI (Y2) gerados com base nas cinco variáveis explicativas (X1 - Área do terreno, em m²; X2 - Testada, em m; X3 - Valor do m², em reais; X4 - Topografia; X5 - Pavimentação) mensuradas em 667 terrenos do município de Itaqui - RS.

Regressão Múltipla para estimação do Valor Venal						
Variáveis explicativas	X1	X2	X3	X4	X5	Correlação
Coefficiente	16,90	-8,35	594,69	-9526,54	-685,45	0,96
Erro padrão	2,52	141,66	42,96	1848,18	1183,20	R ² Ajustado
P-valor	0,00	0,95	0,00	0,00	0,56	0,91
Predição do valor venal: $Y1 = 16,90X1 + 594,69X3 - 9526,54X4$						
Regressão Múltipla para estimação do Valor com base no ITBI						
Variáveis explicativas	X1	X2	X3	X4	X5	Correlação
Coefficiente	11,95	213,43	868,36	-23641,19	33633,46	0,85
Erro padrão	15,17	854,57	259,15	11149,60	7137,95	R ² Ajustado
P-valor	0,43	0,80	0,00	0,04	0,00	0,70
Predição do valor de ITBI: $Y2 = 868,36X3 - 23641,19X4 + 33633,46X5$						

4. Conclusões

O valor venal apresentou correlação positiva, significativa ($p \leq 0,05$) e crescente com a pavimentação, testada, área e valor do m², nessa ordem. Já o valor de ITBI apresentou correlação positiva, significativa ($p \leq 0,05$) e crescente com valor do m² e pavimentação, nessa ordem.

A análise de trilha indicou que a área e o valor do m² possuem relação de causa e efeito sobre o valor venal e o valor do m² e a pavimentação apresenta relação de causa e efeito sobre o valor de ITBI.

Os modelos: $Y=16,90 \times \text{área} + 594,69 \times \text{valor do m}^2 - 9526,54 \times \text{topografia}$ ($r^2=0,91$) e, $Y=868,36 \times \text{valor do m}^2 - 23642,19 \times \text{topografia} + 33633,46 \times \text{pavimentação}$ ($r^2=0,70$), são adequados para a estimação dos valores venal e de ITBI, respectivamente.

AGRADECIMENTOS

Os autores agradecem à Prefeitura Municipal de Itaqui, à equipe do projeto GeoItaqui por auxiliar com recursos financeiros e na coleta de dados e a Universidade Federal do Pampa – UNIPAMPA.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alves, V. “Avaliação de imóveis urbanos baseada em métodos estatísticos multivariados”. *Diss.*, Universidade federal do Paraná, 2005.
- Alves, V. Marleide F. “Seleção de variáveis *stepwise* aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas.” *XI Simpósio brasileiro de automação inteligente*. Fortaleza, Anais, 2013.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. “*NBR 14653-1*”. Avaliação de bens – Parte 1: Procedimentos gerais. Rio de Janeiro. 2001.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. “*NBR 14653-2*”. Avaliação de bens – Parte 2: imóveis urbanos. Rio de Janeiro, 2011.
- Baptistella, M. “*O uso de redes neurais e regressão linear múltipla na engenharia de avaliações: Determinação dos valores venais de imóveis urbanos*”. *Diss.*, Universidade Federal do Paraná, 2005.
- Cruz, C. D., Carneiro, P. C. S. *Modelos biométricos aplicados ao melhoramento genético*, 585p. Viçosa: UFV, 2006.
- Cruz, C., D. Genes. “A software package for analysis in experimental statistics and quantitative genetics.” *Acta Scientiarum Agronomy* 35 (2013): 271-276. Acessado maio 11, 2015, doi: 10.4025/actasciagron.v35i3.21251.
- Fávero, L.P., Belfiore, P., Silva, F.L. DA., Chan, B.L. *Análise de dados: modelagem multivariada para tomada de decisões*. Elsevier, Rio de Janeiro, 2009.
- Ferreira, D.F. *Estatística básica*. UFLA, Lavras, 2009.
- Gujarati, D.N. *Econometria Básica*. Elsevier, Rio de Janeiro, 2006.
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L. *Análise multivariada de dados*. Bookman, Porto Alegre, 2009.
- Heil, J. “*Emprego da estatística multivariada como proposta para o cálculo do valor venal e tributação imobiliária*”. *Diss.*, Universidade Federal do Paraná, 2010.
- IBGE - Instituto Brasileiro de Geografia e Estatística. Censo Demográfico, 2015. Disponível em: <http://www.cidades.ibge.gov.br>.
- Ministério das Cidades. Programa Nacional de capacitação das cidades. Avaliação em massa de imóveis para fins fiscais: Discussão, análise e identificação de soluções para problemas e casos práticos. Brasília, 2012, 114p.
- Montgomery, Douglas C., Peck, Elizabeth A. *Introduction to linear regression analysis*. New York: John Wiley e Sons, 1982.
- Nguyen N., Cripps A. “Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks”. *Journal of Real Estate Research*: 22 (2001): 313-336.
- Pimentel, Gomes F. *Curso de estatística experimental*. FEALQ, Piracicaba, 2009.
- Souza, T. V. “*Aspectos estatísticos da análise de trilha (path analysis) aplicada em experimentos agrícolas*”. *Diss.*, Universidade Federal de Lavras, 2013.
- STATSOFT. *Statistica 7.0 Software*. Tucksa: USA, 2005.
- Steiner, M. T. A. Chaves N. A. Braulio S. N. Alves V. “Métodos estatísticos multivariados

aplicados à engenharia de avaliações”. *Revista Gestão & Produção*: 15 (2008).

Toebe, M., Cargnelutti Filho, A. “Multicollinearity in path analysis of maize (*Zea mays* L.)”. *J. Cereal Sci.*: 57 (2013): 453–462. Acessado Agosto 12, 2015, doi:10.1016/j.jcs.2013.01.014.

Wright, S. “Correlation and causation”. *J. Agric. Res.*: 20 (1921): 557–585.

Recebido em dezembro de 2015.

Aceito em março de 2016.