



**IBAPE NACIONAL**  
Instituto Brasileiro de Avaliações  
e Perícias de Engenharia



**IBAPE BAHIA**  
Instituto Brasileiro de  
Avaliações e Perícias de  
Engenharia da Bahia

# **Avaliação pelo Método Comparativo - Modelo por Krigagem Simples**

**Luiz Carlos dos Santos Borges**



O conteúdo dos trabalhos técnicos apresentados no COBREAP é de inteira responsabilidade de seus autores.

## **AVALIAÇÃO PELO MÉTODO COMPARATIVO**

### ***“MODELO POR KRIGAGEM SIMPLES”***

#### **RESUMO**

Este trabalho se propõe a apresentar uma abordagem sobre o método comparativo direto de dados no mercado imobiliário por análise de regressão, onde o procedimento da estimativa se baseia no conhecimento da função de covariância influenciada pela distância entre as amostras em seu domínio espacial. A escolha da melhor função de ajustamento é pelo método dos mínimos quadrados ordinários, que expressa um valor médio para a localidade e parâmetros de correção que são mensurados em função da posição no espaço geodésico, de aspectos mercadológicos, da infraestrutura urbana e dos atributos do próprio imóvel. O plano dos valores ajustados no espaço bidimensional tem sua confiabilidade decorrente do grau de incerteza aferido para o erro aleatório ou medida residual entre o valor observado e o inferido, e, sendo este influenciado pela disposição espacial da amostra, a equação de regressão incorpora a correção do valor em função da distância que separa cada par de amostras e, com esse procedimento, o modelo adotado é denominado de Modelo por Krigagem Simples.

***Palavras-chave: Avaliação; Método Comparativo; Método dos Mínimos Quadrados Ordinários; Krigagem Simples.***

## **1. INTRODUÇÃO**

A avaliação de imóveis se fundamenta no desenvolvimento de modelos que permite prever o valor em localidades não amostradas a partir de amostras com localidade e intensidade de valor conhecidos, que pode servir ao lançamento de tributos sobre a propriedade imobiliária, ao cálculo da indenização em projetos de desapropriação ou instituição de servidão, ou ainda, por diversos interesses públicos ou privados.

Estes modelos se fundamentam em metodologia preconizada pela Norma Brasileira de Avaliação de Bens (NBR 14653), a qual recomenda a aplicação do método comparativo direto sempre que possível nos trabalhos de avaliação de imóveis, onde o valor de mercado é identificado por meio do tratamento técnico de atributos dos elementos comparáveis e constituintes de uma amostra de dados.

O tratamento técnico mais recomendado pela já referida norma é o da análise de regressão, onde as evidências dos dados coletados no mercado imobiliário, normalmente distribuídas de forma irregular no espaço amostral, podem ser comprovadas estatisticamente quanto às diversas hipóteses de validação do modelo de regressão, e, por fim, verificar sob que aspectos a estimativa de valor no espaço geodésico pode ser generalizada.

O procedimento de ajustamento da equação de regressão espacial é feito com base no método dos mínimos quadrados ordinários e a precisão da estimativa se baseia no conhecimento da função de covariância entre as variáveis aleatórias no domínio da amostra, que incorpora sua localização geográfica no plano UTM, o que permite observar a tendência de comportamento direcional do valor no espaço geodésico da amostra.

A confiabilidade da estimativa, num processo iterativo em que se procura encontrar os parâmetros que minimizam a distancia entre o valor inferido no modelo ajustado e o observado em cada ponto do espaço amostral, é dependente do grau de incerteza aferido para o erro aleatório ou medida residual, e, sendo assim, o vetor que separa cada par de amostras influencia a função de covariância e a análise de regressão é denominada de modelo por Krigagem Simples.

## **2. DESENVOLVIMENTO**

Na engenharia de avaliação, o modelo para estimativa do valor (Y) de uma localização não amostrada (cor vermelha) ou de valor desconhecido (Figura\_1), a ser aceita como verdadeira, é geralmente desenvolvido por análise de regressão com base na aplicação do método dos mínimos quadrados ordinários sobre os valores pesquisados de oferta/transação ou de valores conhecidos (cor verde).



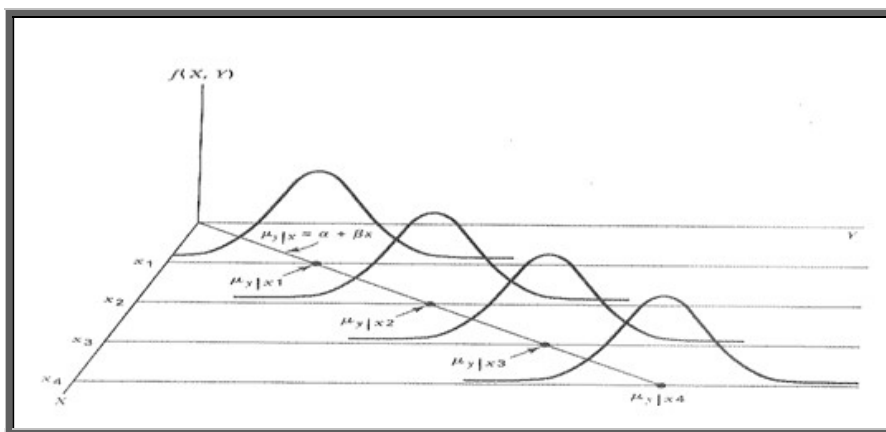
Figura\_1 – Estimativa de Valor

As questões principais, que devem ser observadas no processo estocástico de estimativa do valor, envolvem a interpretação e descrição de sua estrutura espacial e das relações causais que podem ser estabelecidas em função das variáveis explicativas selecionadas para um modelo que é ajustado aos dados pelo método dos mínimos quadrados ordinários (Anexo A – NBR 14653-2).

O método dos mínimos quadrados ordinários admite que a função de comportamento do valor (Y) é aproximada pela distribuição de probabilidade normal (Figura\_2), onde a média (m) da função (f) que explica o comportamento do valor (Equação\_1) é um valor constante e válido para qualquer ponto do espaço geodésico (E, N) da amostra, e não apenas para os pontos amostrais.

$$f[Y_{(E, N)}] = m$$

Equação\_1



Figura\_2 – Distribuição de Probabilidade do Valor

No processo estocástico onde a média (m) é constante para qualquer localização (E, N) do espaço geodésico da amostra, ou seja, para localizações diferentes dos pontos amostrais (Equação\_2), a função de covariância (C) é influenciada pelo vetor (h) que define a distancia entre pares de pontos amostrais (Equação\_3), e, sendo assim, a análise de regressão múltipla é abordada num contexto espacial e denominada de modelo por Krigagem Simples (Wackernagel, 1995).

$$f[Y_{(E, N)}] = f[Y_{(E, N)} + h] = m$$

Equação\_2

$$\text{cov} [Y_{(E, N)} + h, Y_{(E, N)}] = C(h)$$

**Equação\_3**

No modelo por Krigagem Simples (KS), o valor estimado ( $Y^*$ ) é definido a partir da média ( $m_0$ ) do modelo (Equação\_4) com a combinação linear dos parâmetros de vizinhança ( $b_1, b_2$ ) e sua posição geográfica ( $E, N$ ), e dos parâmetros ( $b_i$ ) de cada variável explicativa ( $X_i$ ).

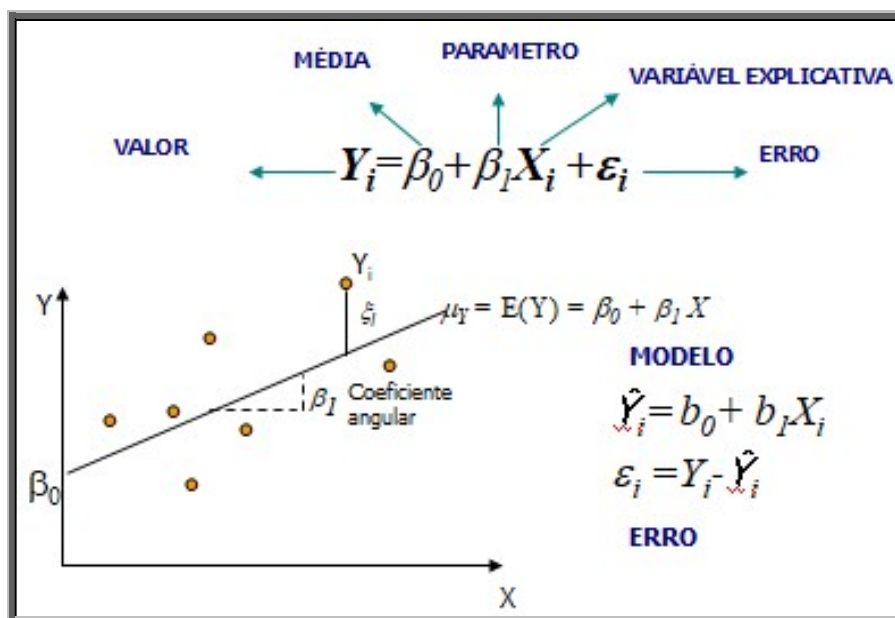
$$Y^* = m_0 + b_1E + b_2N + \sum_{i=1}^n b_i X_i$$

**Equação\_4**

A confiabilidade da estimativa no modelo por Krigagem Simples (KS), em analogia com a análise de regressão múltipla (Figura\_3), se baseia no grau de incerteza aferido para o erro aleatório ( $\varepsilon$ ) ou medida residual (Equação\_5), que corresponde à diferença entre o valor estimado ( $Y^*$ ) e o correspondente valor verdadeiro ( $Y$ ) num ponto ( $E_i, N_i$ ) do espaço amostral (HAIR et al, 2005).

$$\varepsilon = Y^*_{(E_i, N_i)} - Y_{(E_i, N_i)}$$

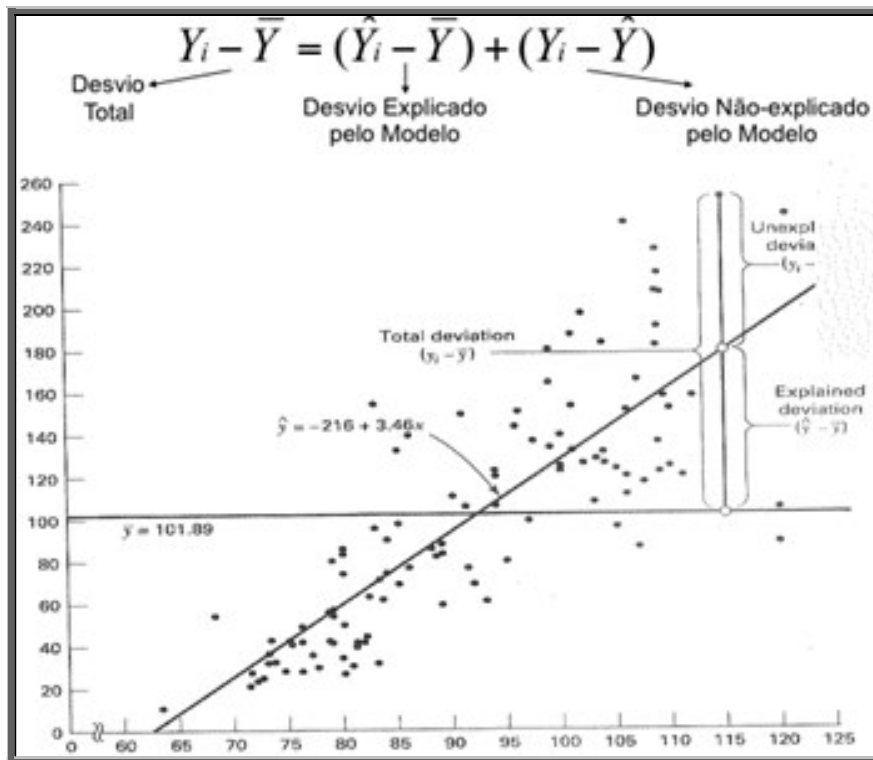
**Equação\_5**



Figura\_4 – Confiabilidade na Regressão Múltipla

A análise de variância entre os valores observados da amostra e os valores preditos no modelo por Krigagem Simples, em analogia com a análise de regressão linear múltipla (Figura\_5), permite definir seu poder de explicação (Equação\_6) através do coeficiente de determinação ( $R^2$ ), desde que observados os aspectos de linearidade, normalidade, homocedasticidade, colinearidade, correlação e demais pressupostos da abordagem clássica para validação da equação de regressão (Anexo C – NBR 14653-2).

$$R^2 = [Y^*_{(E_i, N_i)} - m] / [Y^*_{(E_i, N_i)} - m] + [Y_{(E_i, N_i)} - Y^*_{(E_i, N_i)}] \quad \text{Equação_6}$$



Figura\_5 – Análise de Variância na Regressão Múltipla

O problema de interesse, admitida a hipótese do estimador não viesado, ou seja, o erro é nulo na média (Equação\_7), está em estimar o valor em um ponto não amostrado ( $E_0, N_0$ ) a partir da variância do erro ( $\sigma^2_E$ ) nos pontos amostrados ( $E_i, N_i$ ) em seu domínio espacial (Equação\_8), onde o desenvolvimento dessa expressão permite concluir que a covariância espacial (Equação\_9) é dependente somente do vetor de distancia que separa os pontos amostrados.

$$E[Y^*(E_0, N_0) - Y(E_0, N_0)] = m + \sum_{i=1}^n b_i [E(Y^*(E_i, N_i)) - m] - E[Y(E_0, N_0)] \quad \text{Equação_7}$$

$$\sigma^2_E = \text{var} [Y^*(E_0, N_0) - Y(E_0, N_0)] = E \{ [Y^*(E_0, N_0) - Y(E_0, N_0)]^2 \} \quad \text{Equação_8}$$

$$\text{cov} [Y^*(E_i, N_i), Y^*(E_j, N_j)] = C [(E_i, N_i) - (E_j, N_j)] \quad \text{Equação_9}$$

A estimativa da variância é mínima quando a primeira derivada é zero, sendo a variância ótima ( $\sigma^2_{SKE}$ ) para cada local ou variância do modelo por Krigagem Simples (Equação\_10) obtida pela relação entre a posição da estimativa ( $E_0, N_0$ ) e a da posição amostrada ( $E_i, N_i$ ).

$$\sigma^2_{SKE} = C(E_0, N_0) - \sum_{i=1}^n b_i C [(E_i, N_i) - (E_0, N_0)] \quad \text{Equação_10}$$



### 3. ESTUDO DE CASO

O estudo de caso objetiva analisar o comportamento do mercado de terrenos em uma dada região no município do Recife, incorporando a posição geográfica de cada amostra, obtidas por coordenadas UTM, datum SIRGAS2000, ao conjunto das variáveis explicativas ( $X_i$ ), no intuito de identificar a correlação espacial do valor pela proximidade com a vizinhança.

A abordagem espacial proposta, permite verificar se qualquer mudança de direção no espaço geodésico causa uma variação uniforme do valor (isotropia) ou se a mudança numa dada direção (este/oeste) tem uma influencia diferente em outra direção (norte/sul), e, portanto, com um comportamento do valor não uniforme (anisotropia).

A utilização do posicionamento geodésico dos pontos amostrais, como alternativa para inferir a variação do valor por proximidade através do método dos mínimos quadrados ordinários, não pressupõe que a amostra possua um arranjo espacial predeterminado para os seus elementos constituintes.

A composição da amostra observou o disposto na norma brasileira para avaliação de bens (NBR 14653) para definir a quantidade mínima de dados, que é determinado em função do número de variáveis explicativas quantitativas e dos níveis e dicotomias das variáveis explicativas qualitativas, e, sendo assim, a amostra com 26 (vinte e seis) dados e 19 (dezenove) graus de liberdade, permite o enquadramento no grau de precisão II para as variáveis selecionadas (Figura 6).



Figura 6 – Disposição Espacial das Amostras

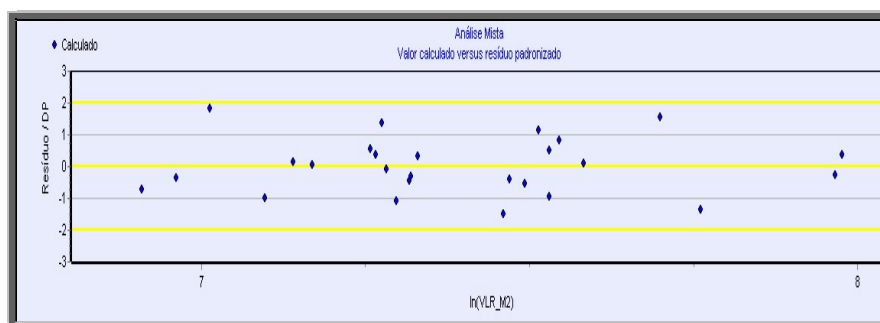
A função de regressão (Equação\_6) que permite inferir a estimativa do valor unitário de terrenos (VAL\_UNIT) é constituída por uma componente constante ou média ( $\beta_0$ ), uma componente que corrige a média por variação das coordenadas UTM de cada amostra ( $\beta_1, \beta_2$ ), uma componente que corrige a média por variação da dimensão de área ou testada do terreno ( $\beta_3, \beta_4$ ), uma componente que corrige a média por variação da renda na localidade ( $\beta_5$ ), uma componente que corrige a média por variação na hierarquia da via ( $\beta_6$ ), e uma componente que expressa o erro aleatório ou medida residual ( $\varepsilon$ ) da estimativa.

$$Y = \beta_0 + \beta_1 \cdot \text{COORD\_E} + \beta_2 \cdot \text{COORD\_N} + \beta_3 \cdot \text{AREA\_TER} + \beta_4 \cdot \text{FRENTE\_TER} + \beta_5 \cdot \text{REND\_SETOR} + \beta_6 \cdot \text{CLASSE\_VIA} + \varepsilon \quad (\text{Equação\_6})$$

A equação de regressão escolhida (Equação\_7) para estimativa pontual do valor unitário de terrenos (Y), explica que há uma correlação linear de média a forte no patamar de 87,91% entre a variável explicada e as variáveis independentes, contudo, o coeficiente de determinação ajustado indica que 70,12% dos valores estimados têm suas variações explicadas pelos parâmetros das variáveis selecionadas. O teste F-Snedecor ( $F=10,78 / p=0,01$ ) comprova que os valores são explicados pelas características selecionadas e não devido ao acaso para a amostra coletada.

$$\begin{aligned} \text{Ln (VLR\_UNIT)} = & 195,47241 - 2,2847659e+12 \cdot 1/\text{COORD\_E}^2 - \\ & 1,9502874e-12 \cdot \text{COORD\_N}^2 + 25,599688 \cdot 1/\text{AREA\_TER}^{1/2} - \\ & 7,7292558 \cdot 1/\text{FRENTE} - 52283,218 \cdot 1/\text{REND\_SETOR}^2 + \\ & 0,11143606 \cdot \text{CLASSE\_VIA} \end{aligned} \quad \text{Equação\_7}$$

A condição de **homocedasticidade** (item A.2.2.3 da NBR 14.653-2/2004) pretendida para o modelo é o termo adotado para testar a hipótese nula de que as variâncias dos erros são iguais nas diversas observações, onde admitimos como confirmada, caso a disposição aleatória e sem padrão definido seja observada na análise gráfica da medida residual versus os valores ajustados pela equação de regressão (Figura\_7).

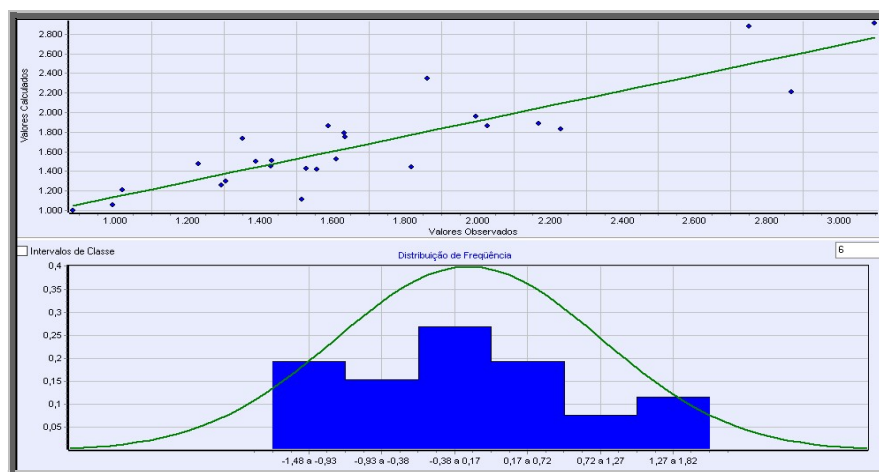


Figura\_7 – Condição de Homocedasticidade

A condição de homocedasticidade (item A.2.2.3 da NBR 14.653-2/2004) também pode ser observada com o teste de Breusch-Pagan (BP), sendo a análise realizada com ajuste da regressão entre a variável resposta definida pelos resíduos padronizados ( $\mu$ ) e a variável explicativa definida pelos valores ajustados ( $Y^*$ ), onde a estatística do teste  $\chi^2_{BP}$  ( $SQE = 18,9687$ ) dessa regressão é maior do que  $\chi^2_{(1)} = 13,12$  (estatística da distribuição qui-quadrado), indicando que aceitamos a hipótese.



A condição de **normalidade** pretendida para o modelo é observada com o exame do histograma da medida residual padronizada, examinando-se a semelhança da forma com a da curva normal, item A.2.2.2 da NBR 14.653-2/2004, bem como, pela comparação da freqüência relativa da medida residual padronizada no intervalo  $[-1, +1]$ , 66% a 74% de probabilidade na distribuição normal e 73% no modelo, no intervalo  $[-1,64, +1,61]$ , com 85% a 95% de probabilidade na distribuição normal e 96% no modelo, e no intervalo de  $[-1,96, +1,96]$ , com 95% a 100% de probabilidade na distribuição normal e 100% no modelo (Figura\_8).



Figura\_8 – Característica de Normalidade

A condição de **normalidade** pretendida para o modelo (item A.2.2.2 da NBR 14.653-2/2004) foi observada também a partir do teste de Jarque\_Bera (JB), onde o exame permitiu verificar que a simetria ( $S = 0,73$ ) em torno da média e a quantidade de dados na cauda da amostra ( $C = 2,86$ ), se aproximam do indicador de simetria ( $S = 0$ ) e de curtose ( $C \leq 3$ ) da distribuição normal.

Os parâmetros estatísticos das variáveis independentes indicam que a AREA\_TER, a FRENTE\_TER, a COORD\_E, a COORD\_N e a RENDA\_SETOR, apresentam um nível de significância menor do que 10% ou nível de confiança maior do que 90% (Figura\_9), enquanto que a variável CLASSE\_VIA tem um nível de confiança maior do que 75%.

| Variável   | Escala      | T. Observado | Significância [%] |
|------------|-------------|--------------|-------------------|
| COORD_E    | $1/x^2$     | -4,25        | 0,04              |
| COORD_N    | $x^2$       | -1,96        | 6,53              |
| AREA_TER   | $1/x^{1/2}$ | 5,40         | 0,01              |
| FRENTE     | $1/x$       | -4,34        | 0,04              |
| SETOR_CEI  | $1/x^2$     | -2,00        | 5,96              |
| CLASSE_VIA | $x$         | 1,23         | 23,21             |
| VLR_M2     | $\ln(y)$    |              |                   |

Figura\_9 – Nível de Confiança das Variáveis

A matriz de correlações (Figura\_10) entre as variáveis independentes e de cada variável independente com a variável dependente, permite observar uma forte dependência linear entre as variáveis AREA\_TER e FRENTE\_TER, e, nestas circunstâncias, só é possível o uso da equação de regressão com multicolinearidade (ITEM A.2.1.5.4, NBR 14653-2), quando o imóvel avaliando atende aos padrões estruturais do modelo, ou seja, o terreno avaliando deve

ter localização similar ao do terreno amostrado e com as mesmas características de área e frente do terreno.

| Tabela de Correlações (valores em percentual) |                    |         |         |          |        |             |            |        |
|-----------------------------------------------|--------------------|---------|---------|----------|--------|-------------|------------|--------|
| Variável                                      | Forma Linear       | COORD_E | COORD_N | AREA_TER | FRENTE | SETOR_CENSO | CLASSE_VIA | VLR_M2 |
| COORD_E                                       | 1/x <sup>2</sup>   |         | 53      | 46       | 39     | 65          | 48         | 70     |
| COORD_N                                       | x <sup>2</sup>     | -20     |         | 21       | 18     | 37          | 54         | 41     |
| AREA_TER                                      | 1/x <sup>1/2</sup> | 3       | -21     |          | 95     | 21          | 17         | 78     |
| FRENTE                                        | 1/x                | 7       | -17     | 91       |        | 20          | 7          | 71     |
| SETOR_CENSO                                   | 1/x <sup>2</sup>   | -43     | 12      | -24      | -20    |             | 50         | 42     |
| CLASSE_VIA                                    | x                  | 11      | 43      | -34      | -33    | 33          |            | 27     |
| VLR_M2                                        | ln(y)              | -43     | -21     | 45       | 17     | -11         | -21        |        |

Figura\_10 – Matriz de Correlações do Modelo

A função ajustada pelo modelo por Krigagem Simples (KS) permite mensurar o valor em cada ponto do domínio espacial da amostra, definido pelo polígono convexo de área mínima que engloba os pontos amostrados (Figura\_11) e que corresponde ao campo geométrico aplicado a variável regionalizada (MATHERON, 1962).



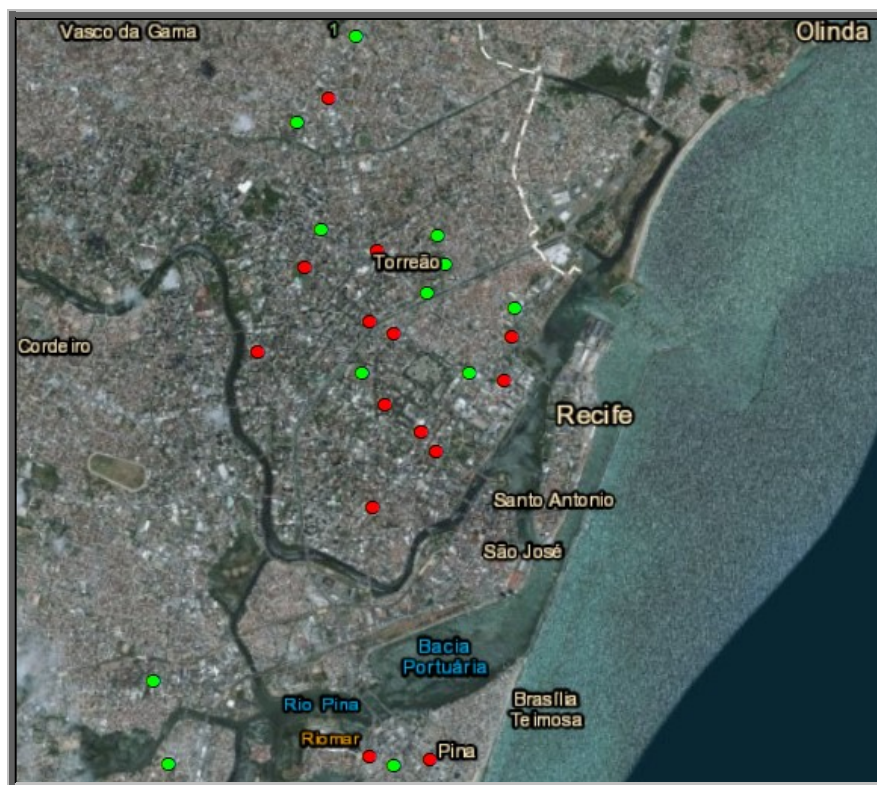
Figura\_11 – Domínio Espacial da Amostra

A disposição espacial do erro relativo (Figura\_12), com base na posição dos pontos amostrados, permite verificar que ocorre uma variação de até 10% (cor amarela) para 15 (quinze) eventos, de 10% a 20% (cor verde) para 5 (cinco) eventos, de 20% até 28% (cor azul claro) para 5 (cinco) eventos, e de 28,67% (cor azul escuro) para um único evento.



Figura\_12 – Disposição Espacial do Erro Relativo

Por outro lado, a disposição espacial do erro absoluto (Figura\_13), permite verificar que ocorre a variação com valor negativo do erro para 15 (quinze) eventos, e de 11(once) eventos para uma variação com valor positivo, onde a distribuição aleatória do erro é um indicativo da ausência de autocorrelação espacial nos resíduos para os dados observados.



Figura\_13 – Disposição Espacial do Erro Absoluto



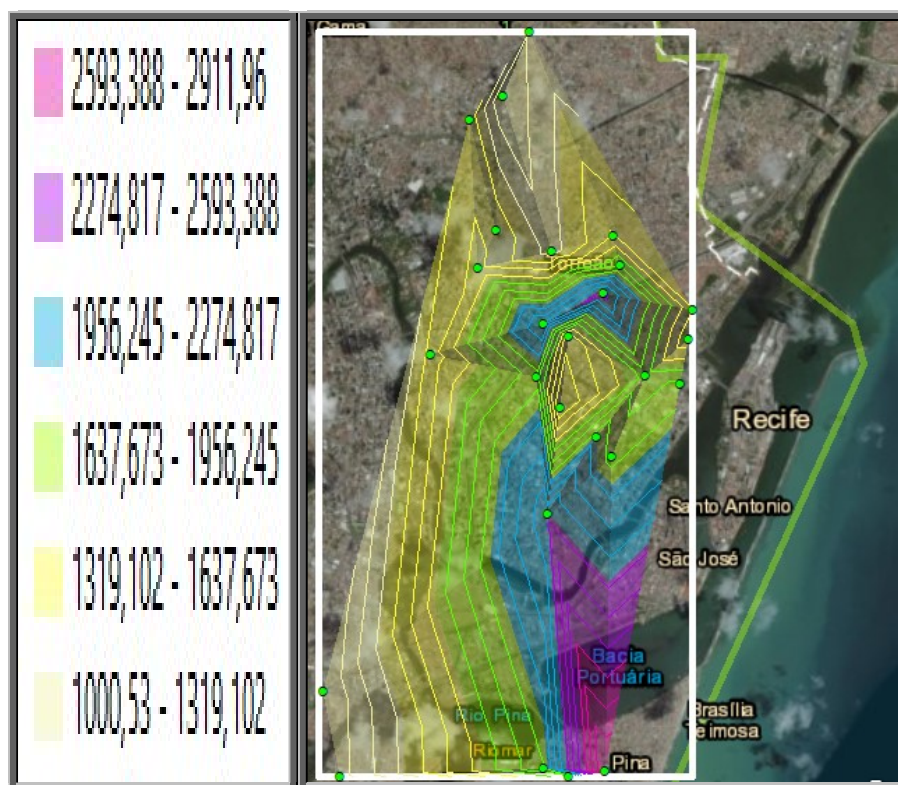
### 3. CONSIDERAÇÕES FINAIS

Os conceitos aqui apresentados se fundamentam na análise de regressão clássica, sendo o modelo por Krigagem Simples (KS) definido pela incorporação do contexto espacial na regressão múltipla, onde a média constante é ponderada por suas variáveis explicativas, dentre estas, a mensuração da correlação espacial entre pares de amostras.

O contexto espacial se fundamenta na identificação da posição espacial das amostras por coordenadas planas UTM e sua incorporação na equação de regressão, e, no modelo probabilístico, comprovamos que os valores unitários de terrenos estão de alguma forma relacionados com os vizinhos e a correlação espacial é estabelecida em função de sua distribuição espacial.

A adoção desse procedimento não apenas descreve o fenômeno do valor nos pontos amostrais, mas para toda a região amostrada, e, sendo assim, representa um passo além da informação contida nos pontos amostrados, permitindo também, estimativas em várias escalas e localizações diferentes das observações da amostra (Wackernagel, 1998).

O espaço amostral para interpolação do valor é o perímetro definido pelo vetor que une os pontos amostrais identificados por coordenadas planas UTM (Figura\_14), no entanto, de acordo com o definido na NBR 14653-2 (item 9.2.1), o domínio amostral para interpolação do valor corresponde ao limite superior e inferior da característica medida por cada variável, e, sendo assim, a linha na cor branca reflete o perímetro definido entre a máxima coordenada UTM e a mínima na direção norte, bem como, da máxima coordenada UTM e a mínima na direção leste.



Figura\_14 – Domínio Espacial da Amostra (NBR 14653-2)

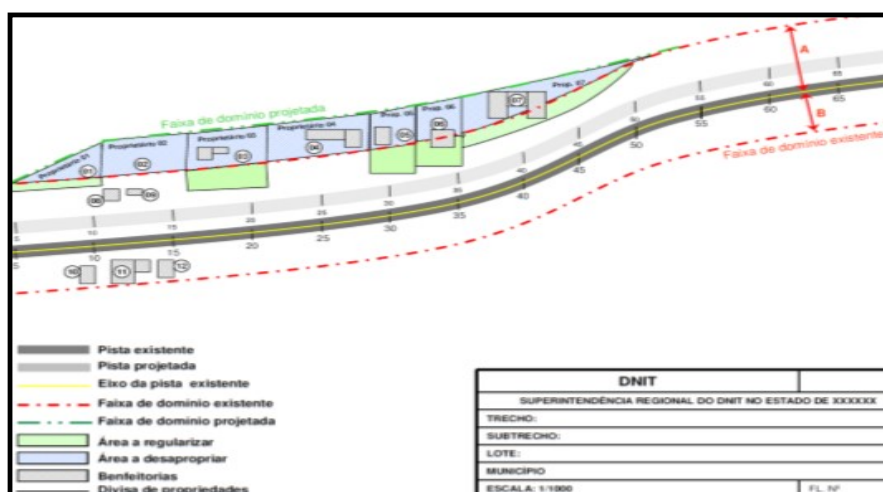
A correlação espacial do valor no espaço geodésico permite observar também que a tendência direcional de comportamento do valor unitário de terreno, tem uma variação mais acentuada na direção de leste para oeste do que a observada na direção de norte para o sul, e, neste caso, a conclusão é de que a tendência direcional de comportamento do valor ocorre de forma anisotrópica.

A operação do modelo por Krigagem Simples pode ser repetida em intervalos do domínio espacial, como por exemplo, para o ponto médio da face de quadra adotada no Cadastro Imobiliário das prefeituras, o que permite sua adaptação na implantação da Planta de Valores Genéricos (Figura\_15), como base de cálculo para apuração do valor do terreno no lançamento do ITBI – Imposto sobre a Transmissão de Bens Imóveis e/ou do IPTU – Imposto sobre a Propriedade Predial e Territorial Urbana.



Figura\_15 – Planta de Valores Genéricos

A adaptação do modelo pode ainda ser efetivada no cálculo de estimativas do valor em Projetos de Indenização (Figura\_16), seja por desapropriação (Item 11.1 – NBR 14653-2) ou por instituição de servidão (Item 11.2 – NBR 14653-2), onde o valor do terreno é obtido a partir das coordenadas UTM de seu centróide, de suas características próprias e das características vinculadas ao espaço urbano onde está inserido.



Figura\_16 – Projetos de Indenização (Fonte: DNIT)

#### 4. REFERENCIA BIBLIOGRÁFICA

- **IRIZARRY, Rafael A.** Introduction to Data Science. Data Analysis and Prediction Algorithms with R, 2019.
- **ABUNAHMAN**, Sérgio Antônio: Curso Básico de Engenharia legal e de Avaliações, Ed. PINI, SP, 2008.
- **J.F. HAIR JR., R.E. ANDERSON, R.L. TATHAM E W.C. BLACK;** trad. Adonai Schlup Sant' Anna e Anselmo Chaves Neto: Análise Multivariada de Dados, 5ª ed., Porto Alegre, Bookman, 2005.
- **ABNT:** NBR 14653-2 – Avaliação de Bens – Parte 2: Imóveis Urbanos, 2004.
- **ABNT:** NBR 14653-1– Avaliação de Bens – Parte 1: Procedimentos Gerais, 2001.
- **DANTAS, R.:** Engenharia de Avaliações. Uma Introdução à Metodologia Científica, Ed. PINI, SP, 1998.
- **WACKERNAGEL, H.** Multivariate Geostatistics. An Introduction with Applications, Springer, 1998.
- **KAHMEN, H. FAIG, W:** Surveying, Walter de Gruyter, Berlim, New York, 1988.
- **ABNT:** NBR 14.166 – Rede de Referência Cadastral Municipal – Procedimento, 1997.
- **ROBINSON, A. H., MORRISON, J. L., MUEHRCKE, P. C., KIMBERLING, J., GUPTILL, S. C.** Elements of Cartography, John Wiley & Sons, INC, 1995.
- **ABNT:** NBR 13.133 – Norma Brasileira para Execução de Levantamento Topográfico, 1994.
- **GEMAEL, C.** Introdução ao Ajustamento de Observações, Aplicações Geodésicas, Curitiba, UFPR, 1994.
- **BURROUGHS, PA.** Principles of Geographical Information Systems for Land Resources Assessment, Oxford University, 1986.
- **Matheron, G..** Traité de Géostatistique Appliquée – Tome I. Paris, Editions Technip. 1962.