

CONSÓRCIO MOGIANO

A Tabela acima apresenta valores usualmente adotados na avaliação geral de um modelo de regressão.

O RMSE (Root Mean Squared Error) é uma média que representa a raiz quadrada da média dos quadrados dos erros entre os valores previstos pelo modelo e os valores reais da variável dependente. O RMSE é útil para avaliar a precisão das previsões do modelo. Ele indica quanto os valores previstos estão desviando dos valores reais. Quanto menor o valor do RMSE, melhor é o desempenho do modelo.

O MAE (Mean Absolute Error) é uma medida que representa a média dos valores absolutos dos erros entre os valores previstos pelo modelo e os valores reais da variável dependente. O MAE é outra métrica de precisão do modelo. Ele é menos sensível a outliers em comparação com o RMSE. Assim como o RMSE, quanto menor o valor do MAE, melhor o desempenho do modelo.

Já o R^2 (Coeficiente de Determinação) é uma medida estatística que indica a proporção da variabilidade da variável dependente que é explicada pelo modelo de regressão. Ele varia de 0% a 100%, sendo 100% o **melhor resultado possível**, indicando que o modelo explica toda variabilidade da variável dependente. Perante as variáveis consideradas, o coeficiente R^2 , aponta para um ajuste em torno de **71%**, o que caracteriza a tendência de correlação entre os valores observados e estimados.

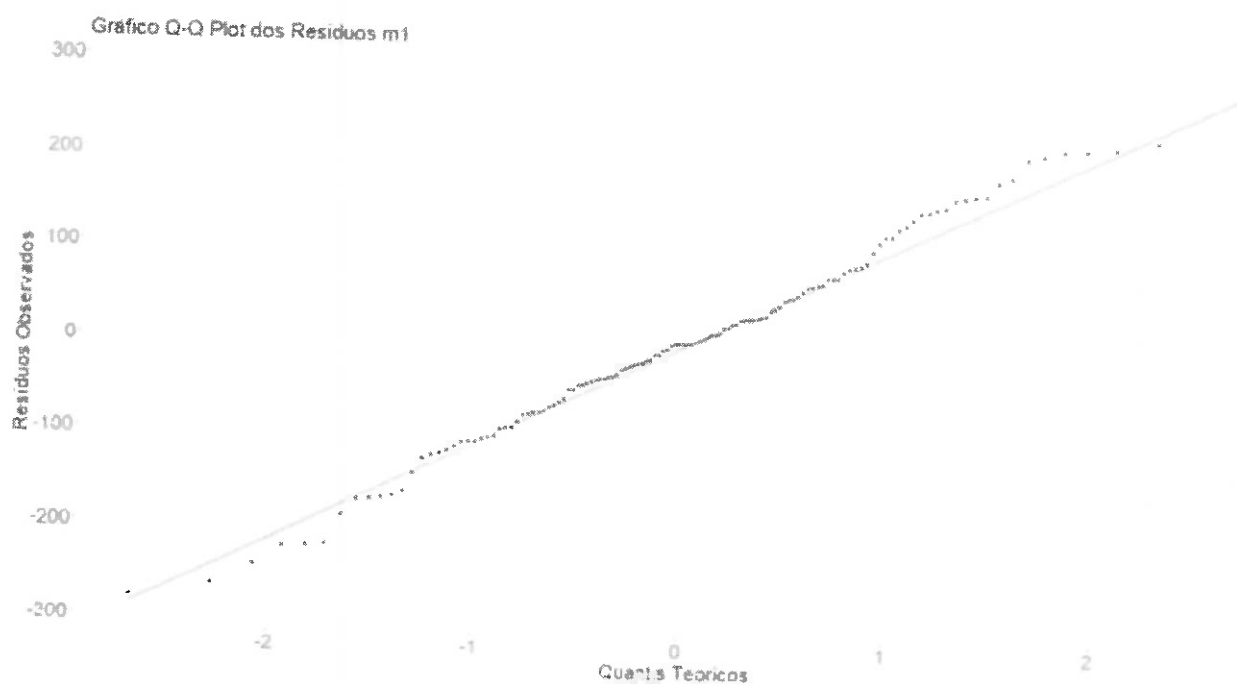
Tabela 1- Coeficientes de regressão analisados.

R^2	RMSE	MAE
71%	112	87

As próximas figuras ilustram uma visão geral sobre os resíduos observados no ajuste do modelo. O Gráfico Q-Q Plot (

Figura 2) compara os quantis dos resíduos com os quantis de uma distribuição normal. Devido ao fato de os resíduos seguirem uma distribuição normal, os pontos no gráfico Q-Q plot se aproximam da linha diagonal.

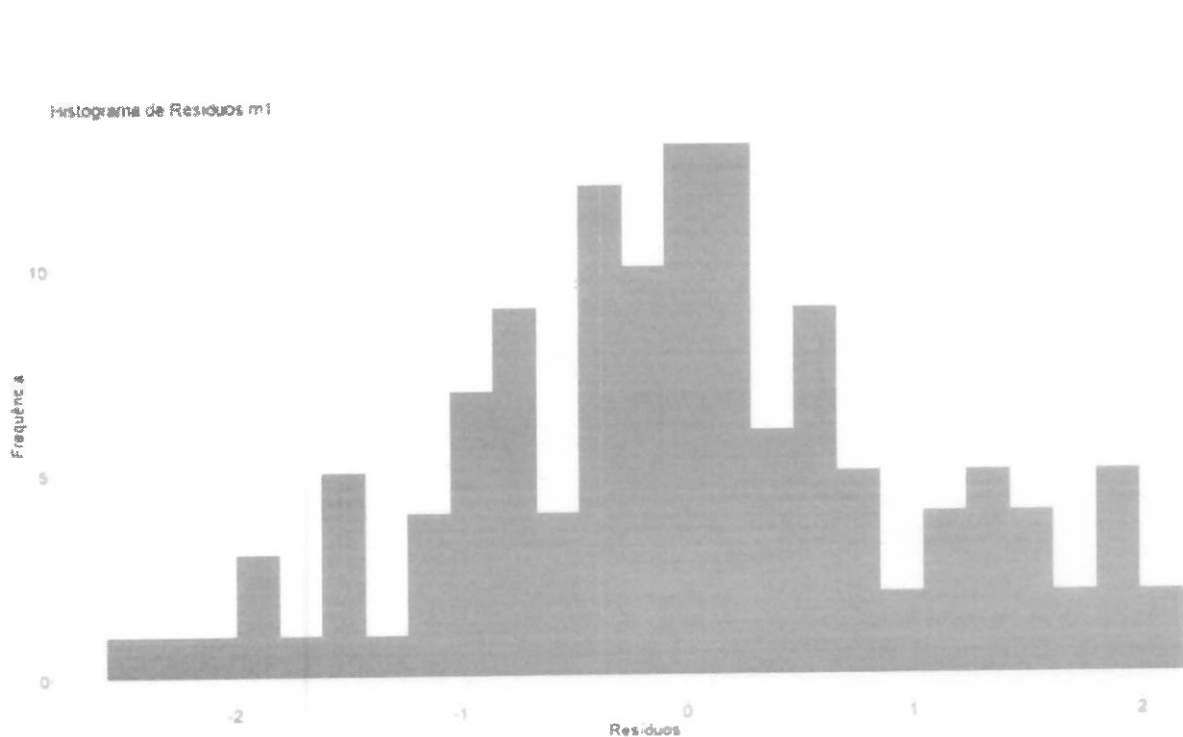
Figura 2 - Gráfico Q-Q Plot dos Resíduos.



Já com o Histograma de Resíduos (

Figura 3) é possível observar que os resíduos estão normalmente distribuídos em torno de zero. Ambos os gráficos indicam a inexistência de altas frequências de discrepância de ajuste.

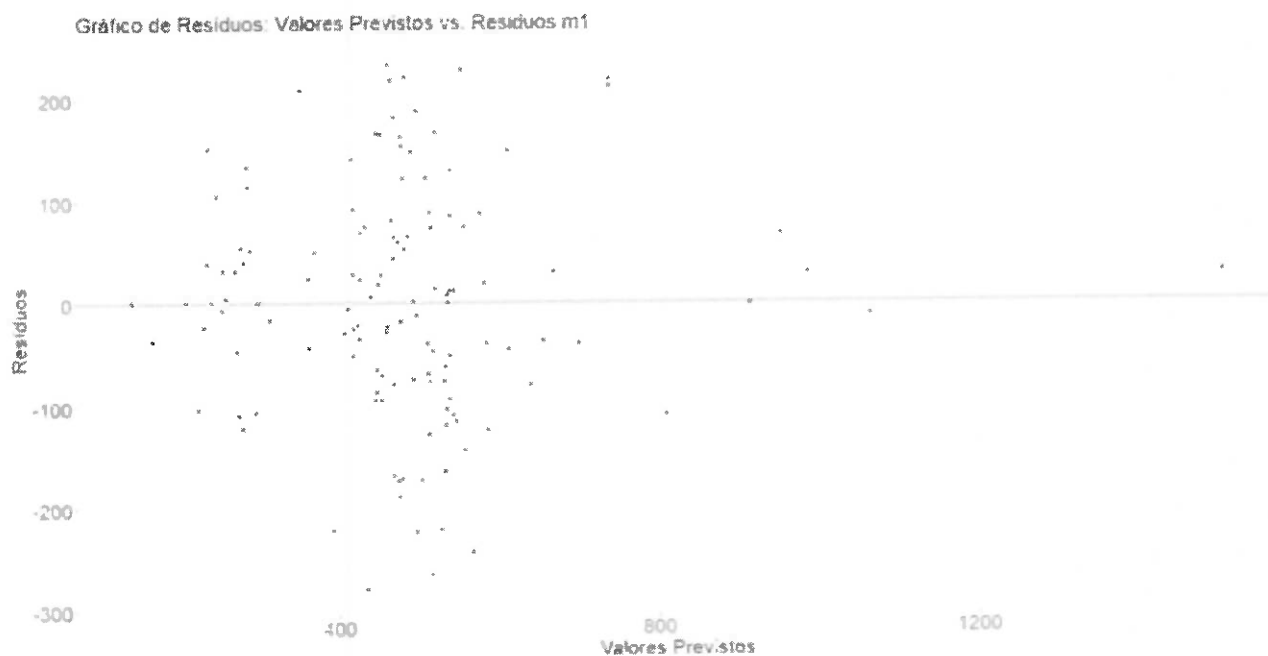
Figura 3 - Histograma de Resíduos.



Com o gráfico “Valores previstos vs. Resíduos” (

Figura 4) é possível validar que os resíduos do modelo (diferença entre os valores reais e previstos) estão distribuídos aleatoriamente em torno de zero o que indica pouca tendência nos resíduos em função do valor retornado pelo modelo.

Figura 4 - Gráfico de Resíduos: Valores Previstos vs. Resíduos.

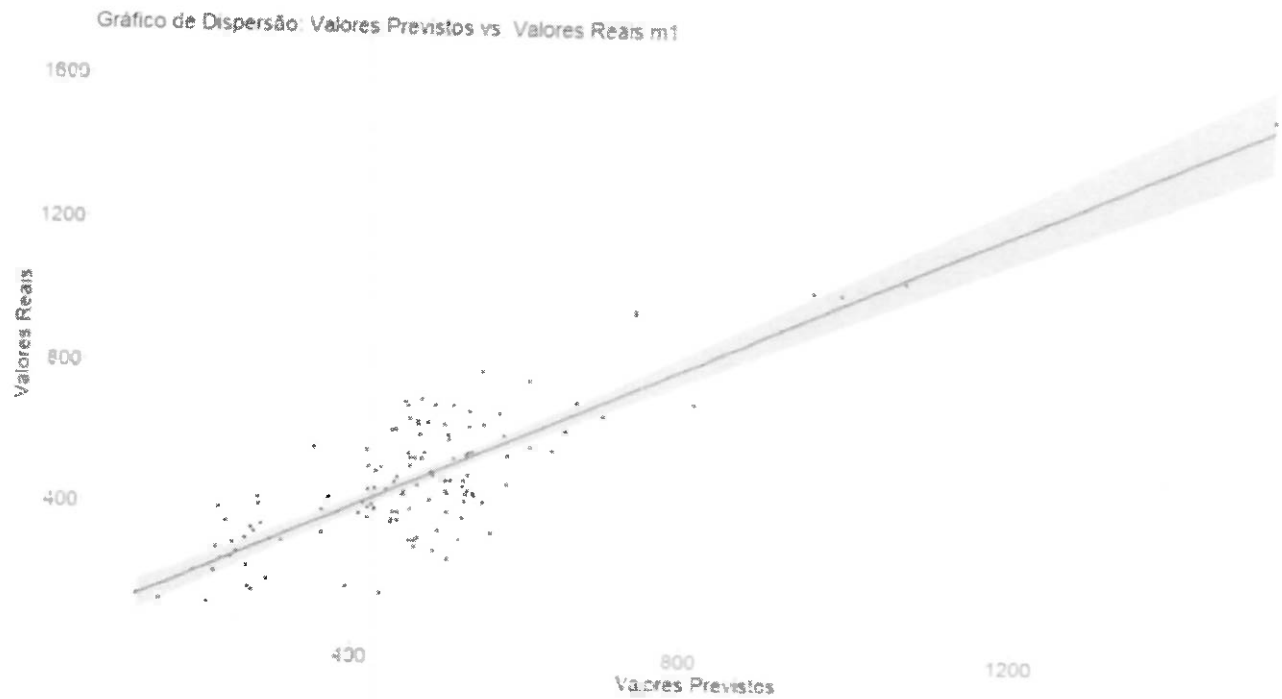


Com intuito de exibir a tendência de predição do modelo ajustado, o Gráfico de Dispersão (

Figura 5) mostra a relação linear entre valores observados e estimados, corroborando assim os coeficientes de ajuste já discutidos (Tabela 1).

CONSÓRCIO MOGIANO

Figura 5 - Gráfico de Dispersão: Valores Previstos vs. Valores Reais.



4. DISCUSSÃO SOBRE FATORES QUE INFLUENCIAM NA VALORAÇÃO.

O modelo de regressão obtido, conforme expresso na Equação 1, determina o valor da unidade do terreno (Vm^2T) através das variáveis preditoras:

- $area_terreno, gleba$
- $dist_alta_tensao$
- $zona_homogenea_1$
- $zona_homogenea_3$
- $zona_homogenea_6$
- $zona_homogenea_8$
- $zona_homogenea_10$
- $zona_homogenea_12$
- $zona_homogenea_14$
- $dens_imobiliaria$
- $desn_comercial$
- $zona_homogenea_2$
- $zona_homogenea_4$
- $zona_homogenea_7$
- $zona_homogenea_9$
- $zona_homogenea_11$
- $zona_homogenea_13$
- $zona_homogenea_15$

O processo de escolha destas variáveis é pautado pela aplicação de técnicas e testes estatísticos conhecidos para determinação dos fatores que influem na valoração da unidade de área.

$$\begin{aligned}Vm^2T = & 394.0 + (area_terreno * 0.023) + (gleba * -381.62) \\& + (dens_comercial * 59.10) + (dens_imobilia * -1.69) \\& + (dist_alta_tensao * -0.03) + (zona_homogenea_1 * -82.48) \\& + (zona_homogenea_2 * 501.95) + (zona_homogenea_3 * -1.842.99) \\& + (zona_homogenea_4 * -84.16) + (zona_homogenea_6 * 364.47)\end{aligned}$$