

Relatório geoestatística

Naimara Vieira do Prado

Análise geoestatística para os dados de Resistência do solo à penetração (RSP)

Resumo: A resistência do solo à penetração é uma medida que detecta esta compactação, contudo ela é fortemente influenciada pela densidade, pelo teor de água no solo. O objetivo deste trabalho foi avaliar a variabilidade espacial da resistência do solo à penetração em um Latossolo Vermelho Distrófico. As amostras de solo foram coletadas em uma área agrícola comercial na cidade de Cascavel, na camada de 0,00-0,20 m de profundidade. Na análise geoestatística foram utilizados os métodos de Mínimos Quadrados Ordinários (OLS), Mínimos Quadrados Ponderados (WLS1), Máxima Verossimilhança (ML) e Máxima Verossimilhança Restrita (RML) para ajuste dos modelos espaciais (exponencial, esférico, gaussiano e família Matérn). Os critérios de validação cruzada foram utilizados para a escolha do melhor modelo ajustado e a partir disso, foram construídos mapas temáticos da RSP utilizando a krigagem ordinária.

A resistência do solo à penetração tem sido frequentemente utilizada como indicador da compactação do solo em sistemas de manejo, por ser um atributo diretamente relacionado ao crescimento das plantas e de fácil e rápida determinação (MERCANTE et al., 2003). Segundo Balastreire (1987), a resistência do solo à penetração (RSP) é uma propriedade dinâmica do solo identificada como a reação do solo à ação de uma força aplicada por meio de uma ponteira cônica. A resistência do solo à penetração pode ser influenciada pela densidade, pelo teor de água e pelo tipo de solo.

A tabela 1 apresenta as estatísticas descritiva dos dados de resistência do solo à penetração (RPS) na camada de 0,10 a 0,20 m, os valores são expressos em MPa.

Tabela 1: Estatísticas descritivas

Estatísticas	Resultados
Média (\bar{x})	2,427
Desvio padrão (S)	0,510
Variância (S^2)	0,261
Coef. variação (CV)	21,04%
Mínimo	0,875
1 Quartil (Q_1)	2,131
Mediana	2,329
3 Quartil (Q_3)	2,771
Máximo	3,709
Assimetria (As)	0,107
Curtose (K)	0,280
n	90

O valor médio de RSP para a área em estudo foi de 2,427 (MPa) na camada de 0,1 a 0,2 m, segundo Canarache (1994) valores de RSP entre 1,10 e 2,59 MPa apresentam pouca limitação para o desenvolvimento radicular, porém, com o intuito de um manejo localizado para os atributos físicos do solo, faz-se necessário análises geoestatísticas para analisar os valores de RSP de acordo com sua localização na área em estudo. Pode-se observar pela mediana que 50% das medidas de RSP são menores ou iguais a 2,329 MPa. Os dados de RSP apresentam assimetria positiva ($As = 0,107 > 0$) e distribuição leptocúrtica ($K = 0,280 < 3$). Na tabela acima, percebe-se que os dados de RSP apresentam coeficiente de variação superior a 20%, o que indica que os dados não são homogêneos (Pimentel Gomes, 2000).

A figura 1 apresenta o histograma dos dados de RSP construído com os 90 elementos amostrais. Verifica-se que, mesmo com o coeficiente de assimetria indicando que os dados possuem assimetria positiva, pelo histograma percebe-se que os dados seguem uma distribuição muito similar à distribuição normal.

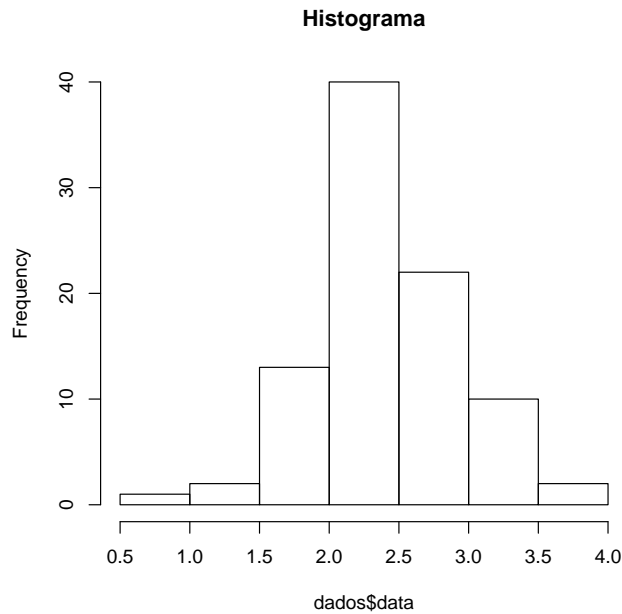


Figura 1: Histograma dos dados de RPS.

A figura 2 apresenta o gráfico *Box-plot* para os dados de RSP, construído com os 90 elementos amostrais. O gráfico *Box-plot* é utilizado usualmente para se ter uma idéia do comportamento dos dados, com esse gráfico pode-se notar a assimetria positiva dos dados. Além disso, os dados de RPS na camada de 0,10 a 0,20 m, apresenta um ponto discrepante.

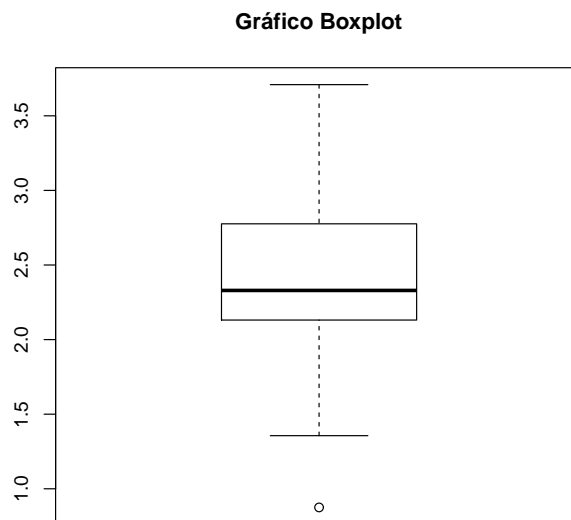


Figura 2: Gráfico *Box-plot* dos dados de RPS.

Na figura 3 podemos observar a localização do ponto discrepante, este ponto, corresponde à parcela de número 1, cujo valor de resistência do solo é 0,875 MPa, sendo um

valor de mínimo.

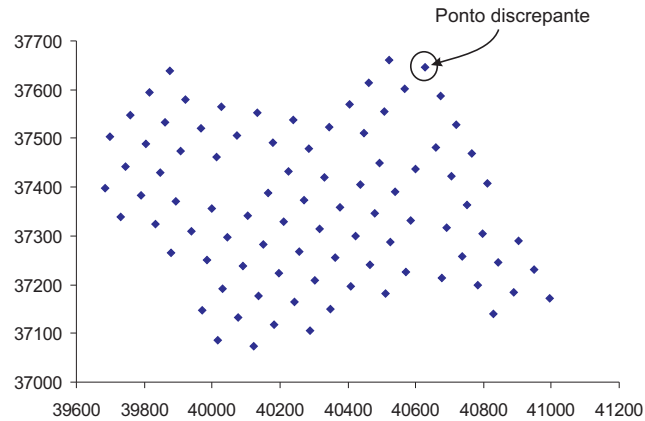


Figura 3: Localização dos pontos na área experimental.

A figura 4 apresenta o gráfico Post-Plot para os dados de RPS na camada de 0,10 a 0,20 m. Podemos observar por sua localização, que o ponto discrepante faz vizinhança com pontos de maior valor, isso indica que possivelmente, esse ponto pode ser um ponto influente para o conjunto de dados.

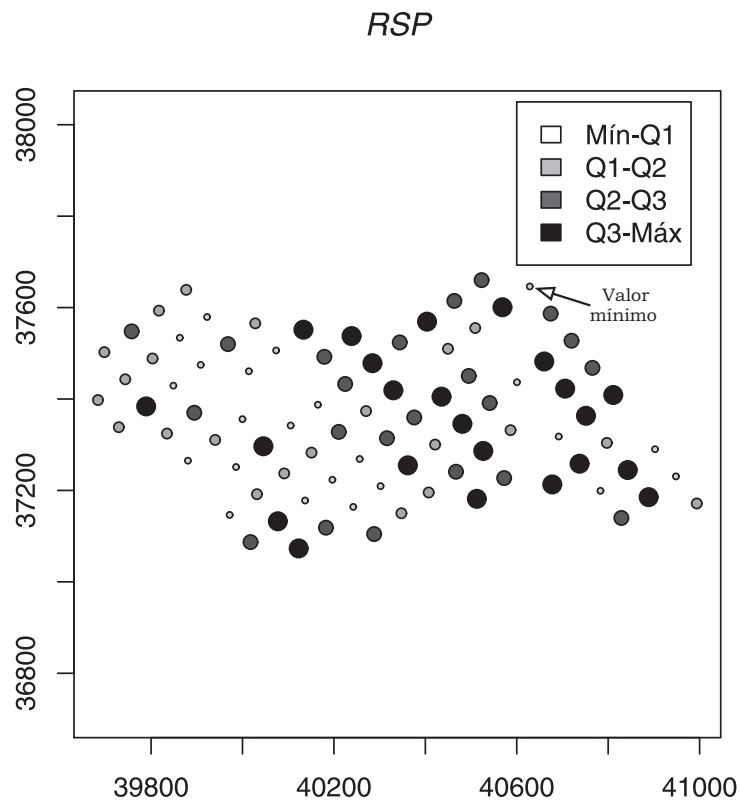


Figura 4: Post-plot dos dados de RPS.

O gráfico Post-plot é utilizado para se ter uma ideia sobre a continuidade e verificar se há regimes com concentração de valores altos ou baixos, assim, apresenta a distribuição espacial dos dados. O gráfico apresentado na figura 4 está dividido segundo os quartis da amostra de dados. Com o Post-plot é possível observar a tendência direcional dos dados,

ou seja, podemos visualizar que a parte oeste (direita) do gráfico apresenta valores mais altos que a parte leste (esquerda).

Ao construir os semivariogramas nas direções 0, 45, 90 e 135 tem-se a possibilidade de investigar se os semivariogramas são idênticos para todas as direções, sugerindo o mesmo padrão de continuidade espacial. Com isso, é possível classificar os dados em: anisotrópicos (variações diferentes para cada direção de estudo) ou isotrópicos, e assim analisá-los adequadamente.

Para realizar o estudo variográfico é necessário conhecer a distância máxima entre os pontos amostrais, com intuito de definir o ponto de corte. Neste conjunto de dados a distância máxima entre os pontos é de 1337,626 metros. Deste modo, os semivariogramas em cada uma das direções são construídos utilizando *cutoff* 50% da distância máxima entre os pontos, conforme sugere Clark (1979), correspondendo a 668,5 metros. Para a construção do semivariograma, é necessário escolher as distâncias (h) nas quais as semi-variâncias experimentais serão estimadas, neste caso, como trata-se de uma grade regular de 75×75 , foram divididos em 10 *lags*, visando atingir a distância mínima de 75 metros entre pontos.

A figura 5 apresenta os semivariogramas para cada uma das direções indicadas, podemos observar que os semivariogramas direcionais apresentam comportamentos diferentes para cada uma das direções, porém, como a escala do eixo y (semivariância) é pequena, podemos afirmar que dos dados de RSP na camada de 0,1 a 0,2 m são isotrópicos.

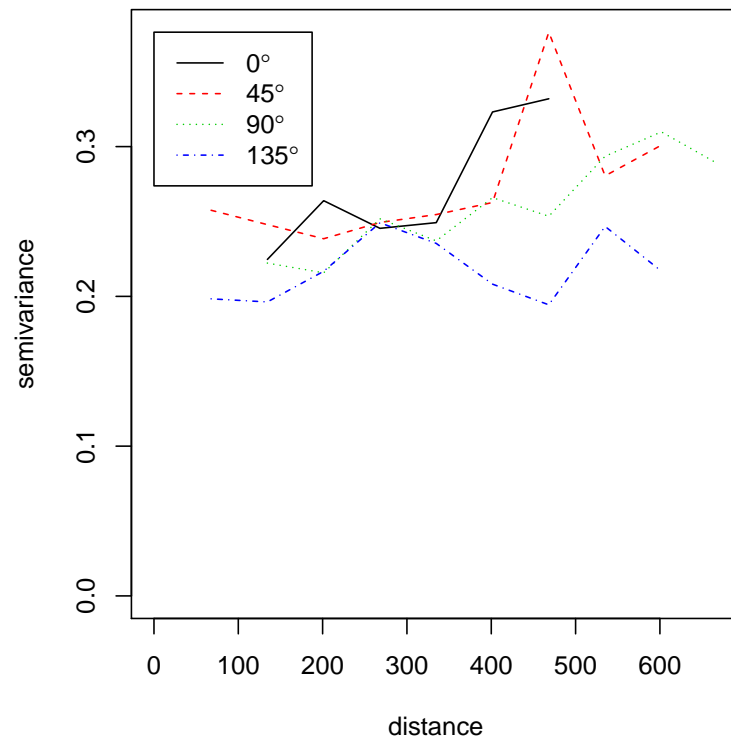


Figura 5: Semivariogramas direcionais.

Considerando que há isotropia nos dados pode-se construir um semivariograma omnidirecional, indicando que os semivariogramas apresentam o mesmo comportamento para todas as direções analisadas.

A figura 6 apresenta o semivariograma omnidirecional. O semivariograma é construído considerando o mínimo de 30 pares para o cálculo das semivariâncias (Guerra, 1988).

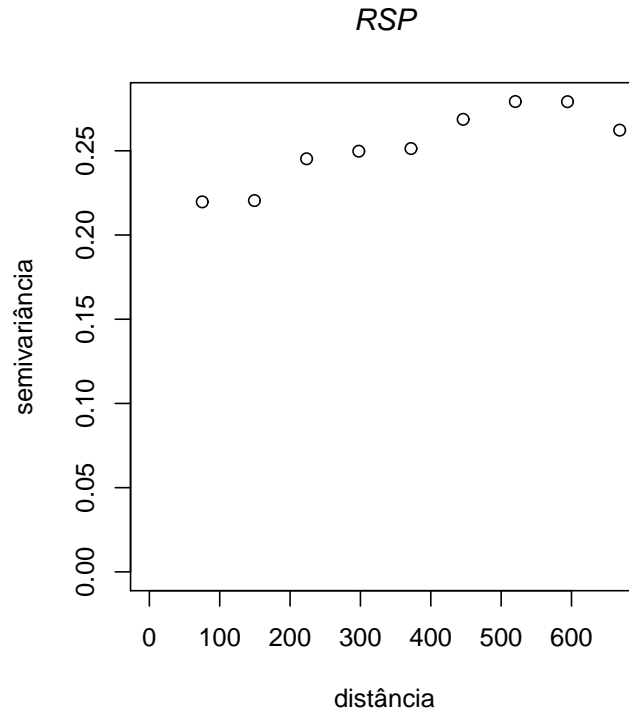


Figura 6: Semivariograma omnidirecional.

Para o cálculo das semivariâncias foi utilizado o estimador de Matheron (1963), considerando *cutoff* de 668,5 metros, 50% da distância máxima e 10 *lags*.

A tabela 2 apresenta os parâmetros do semivariograma experimental juntamente com o valor obtido da função de autocorrelação espacial.

Tabela 2: Parâmetros do semivariograma omnidirecional

Distância (h)	$\gamma(h)$	ρh	N de pares
75,166	0,219	0,302	262
149,333	0,220	0,357	325
223,500	0,245	0,400	371
297,666	0,249	0,397	502
371,833	0,251	0,380	488
446,000	0,268	0,389	450
520,166	0,279	0,411	329
594,333	0,279	0,367	274
668,500	0,262	0,366	270

Observa-se que para todas as variâncias calculadas, a quantidade de pares ultrapassa o limite de 30 pares. Analisando os valores da função de autocorrelação espacial, podemos notar que a maior correlação ocorre na distância 520,166 m e a menor correlação ocorre na distância de 75,166 m, muito próximo da menor distância entre os dados.

Uma forma de investigar se há dependência espacial nos dados, é com a construção de envelopes simulados, apresentado na figura 7.

A figura 7 apresenta o gráfico de envelopes, este sugere se há continuidade espacial nos dados, ou seja, se há pontos fora da linha de confiança, pode-se afirmar que há continuidade espacial nos dados.

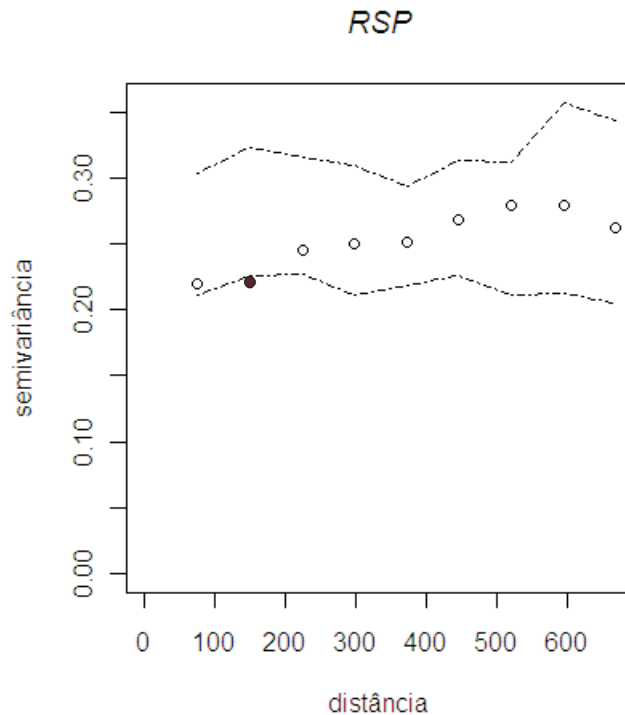


Figura 7: Gráfico de envelopes.

Como pode-se observar na figura 7 há apenas um ponto totalmente fora da linha de confiança, com isso temos indícios de que há continuidade espacial, será realizada a análise para explicitar a estrutura de dependência espacial dos dados de RPS na camada de 0,1 a 0,2m.

A seguir serão ajustados os modelos: exponencial, esférico, gaussiano e família Matérn com $k=0,7$. Para cada um dos métodos de estimação: mínimos quadrados ordinários (OLS), mínimos quadrados ponderados (WLS1), máxima verossimilhança (ML) e máxima verossimilhança restrita (RML), para posterior escolha do melhor modelo ajustado.

A família Matérn é válida para φ_3 , $k > 0$ e corresponde ao modelo exponencial quando $k = 0,5$ e, ao modelo gaussiano quando $k \rightarrow \infty$. Deste modo, foi ajustado o modelo da família Matérn considerando $k = 0,7$.

Para cada modelo ajustado a partir de um chute inicial, com 99 interações realizadas pelo software R, foram obtidos as estimativas da média (β), efeito pepita (φ_1), contribuição (φ_2) e alcance (a), conforme a tabela 4.

Com os parâmetros acima, é possível calcular o valor do efeito pepita relativo (EPR) sugerido por Cambardella (1990), que indica o grau de dependência para cada modelo

(tabela 3). O efeito pepita relativo é obtido por $EPR = \frac{\varphi_1}{\varphi_1 + \varphi_2}$

$EPR < 0,25$	Forte dependência espacial
$0,25 \leq EPR < 0,75$	Média dependência espacial
$EPR > 0,75$	Fraca dependência espacial

A tabela 4 apresenta os modelos ajustados com os respectivos parâmetros obtidos.

Método	Modelos	β	φ_1	φ_2	$\varphi_1 + \varphi_2$	$\varphi_3 = g(\mathbf{a})$	\mathbf{a}	EPR
OLS	Exponencial	2,427	0,145	0,119	0,264	99,999	299,573	0,551
	Esférico	2,427	0,172	0,087	0,259	249,999	249,999	0,665
	Gaussiano	2,427	0,194	0,067	0,261	149,999	259,622	0,742
	Matérn (k=0,7)	2,427	0,168	0,096	0,265	99,999	344,687	0,635
WLS1	Exponencial	2,427	0,187	0,092	0,279	249,999	748,933	0,670
	Esférico	2,427	0,178	0,081	0,260	299,999	299,999	0,686
	Gaussiano	2,427	0,207	0,061	0,268	249,999	432,704	0,772
	Matérn (k=0,7)	2,427	0,185	0,086	0,271	149,999	517,031	0,682
ML	Exponencial	2,426	0,205	0,052	0,257	120,900	362,201	0,798
	Esférico	2,432	0,213	0,044	0,258	358,018	358,018	0,826
	Gaussiano	2,431	0,221	0,036	0,258	189,400	327,884	0,859
	Matérn (k=0,7)	2,426	0,211	0,045	0,257	108,762	374,891	0,823
RML	Exponencial	2,4230	0,2091	0,0570	0,266	169,3096	507,2062	0,786
	Esférico	2,4328	0,2111	0,0537	0,264	381,2423	381,2423	0,797
	Gaussiano	<u>2,4320</u>	<u>0,2214</u>	<u>0,0448</u>	<u>0,266</u>	<u>205,3644</u>	<u>355,4486</u>	<u>0,832</u>
	Matérn (k=0,7)	2,4240	0,2143	0,0516	0,265	143,8373	143,8373	0,801

β : média; φ_1 : efeito pepita; φ_2 : contribuição; $\varphi_3 = g(\mathbf{a})$: função do alcance; \mathbf{a} : alcance; $EPR = \left(\frac{\varphi_1}{\varphi_1 + \varphi_2} \right)$: efeito pepita relativo;

Observa-se na tabela 4 que o maior raio de dependência espacial é proporcionado pelo modelo exponencial com o método de mínimos quadrados ponderados (WLS1), este modelo apresentou moderada dependência espacial ($EPR = 0,670$). O modelo ajustado que apresentou menor valor de efeito pepita relativo ($EPR = 0,551$) foi o exponencial pelo método de mínimos quadrados (OLS). Assim, percebe-se a necessidade de critérios para escolha do melhor modelo ajustado. Uma técnica bastante utilizada é a validação cruzada. Segundo Faraco (2008) o melhor modelo ajustado será aquele que apresentar os menores Erro Médio (EM) e Erro Médio Reduzido (ER), Desvio padrão do erro médio (S_{EM}), menor possível, Desvio padrão do erro reduzido (S_{ER}) mais próximo de 1 e Erro Absoluto (EA)

seja o menor possível. Além disso, para os métodos ML e RML foram calculados os parâmetros de Akaike (AIC) e Akaike corrigido (CAIC) e Bayesian Information Criterion (BIC).

A tabela 5 apresenta os parâmetros calculados pelo método de validação cruzada, AIC, CAIC e BIC, respectivamente.

Tabela 5: Parâmetros para escolha dos modelos ajustados

Método	Modelos	EM	ER	S _{EM}	S _{ER}	EA	AIC	CAIC	BIC
OLS	Exponencial	-0,00054	-0,00056	0,50644	1,04927	35,49292	-	-	-
	Esférico	-0,00043	-0,00044	0,50593	1,04448	35,77361	-	-	-
	Gaussiano	-0,00030	-0,00031	0,50397	1,04456	35,43549	-	-	-
	Matérn (k=0,7)	-0,00044	-0,00046	0,50528	1,04795	35,32732	-	-	-
WLS1	Exponencial	-0,00023	-0,00023	0,50365	1,04500	34,82279	-	-	-
	Esférico	-0,00028	-0,00028	0,50210	1,04382	35,25508	-	-	-
	Gaussiano	4,43e-05	5,76e-05	0,50023	1,04136	35,43549	-	-	-
	Matérn (k=0,7)	-0,00026	-0,00027	0,50348	1,04802	34,95280	-	-	-
ML	Exponencial	-7,27e-05	-7,15e-05	0,50123	1,00889	35,03481	138,6	139,1	148,6
	Esférico	1,82e-05	2,09e-05	0,49915	1,00916	34,89847	138,2	138,6	148,2
	Gaussiano	4,59e-05	4,94e-05	0,49890	1,00909	34,83145	138,2	138,6	148,2
	Matérn (k=0,7)	-6,00e-05	-5,87e-05	0,50103	1,00890	34,83145	138,6	139,1	148,6
RML	Exponencial	-9,68e-05	-9,37e-05	0,50138	1,00688	34,82613	137,1	137,5	147,1
	Esférico	-5,42e-06	-1,69e-06	0,49923	1,00699	34,80718	136,8	137,2	146,8
	Gaussiano	4,37e-05	4,91e-05	0,49885	1,00707	34,69970	136,7	137,6	146,7
	Matérn (k=0,7)	-7,97e-05	-7,64e-05	0,50116	1,00690	34,81259	137,0	137,4	147,0

EM: erro médio; ER: erro reduzido; S_{EM}: desvio padrão do erro médio; S_{ER}: desvio padrão do erro reduzido; EA: erro absoluto; AIC: Método de Akaike; CAIC: Método de Akaike corrigido; BIC: Método de Akaike modificado.

Na tabela 5 pode-se assumir que o modelo que melhor se ajusta ao semivariograma experimental é o modelo gaussiano pelo método de máxima verossimilhança restrita (RML). A figura 8 apresenta o semivariograma experimental com o modelo ajustado.

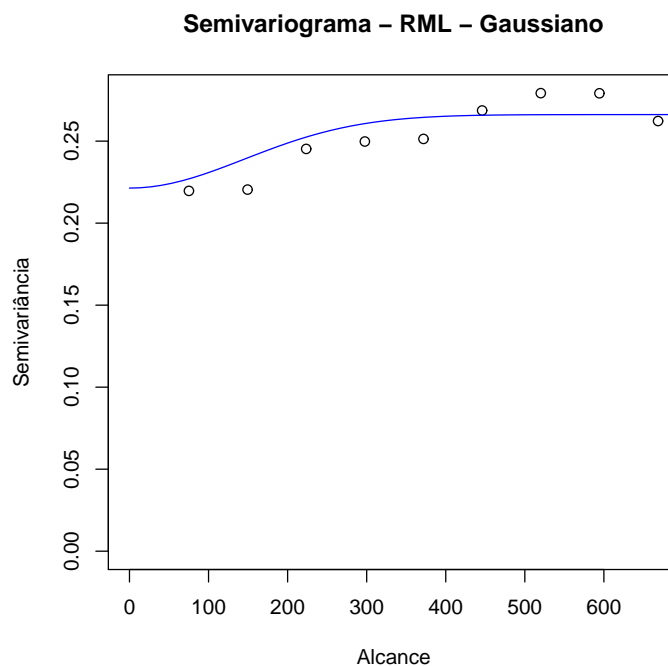


Figura 8: Semivariograma experimental com o modelo ajustado.

Os parâmetros associados ao melhor modelo gaussiano por RML são os seguintes:

Efeito pepita ($\varphi_1 = 0,2214$): Representa as variações locais de pequena escala, como erros de amostragem, é o valor de semivariância para distâncias menores que a mínima distância amostrada.

Contribuição ($\varphi_2 = 0,0448$): Representa as diferenças espaciais entre os valores de uma variável tomada em dois pontos separados por distâncias cada vez maiores.

Patamar ($\varphi_1 + \varphi_2 = 0,266$): Para os dados de RSP da área em estudo, os valores do semivariograma experimental, estabilizam-se a partir desse valor (0,266). O patamar é uma estimativa a variância dos dados, se compararmos esse valor com a variância apresentada na tabela 1, onde, $S^2 = 0,261$, percebe-se que os valores são muito próximos.

Alcance ($a = 355,4486$): Indica que a partir dessa distância (355,4 metros), um ponto amostral não apresenta mais influência sobre o ponto vizinho.

Efeito pepita relativo ($EPR = 0,832$): Indica que os dados de RSP na camada de 0,1 a 0,2 m, na área em estudo apresenta fraca dependência espacial, segundo a classificação de Cambardella (1994).

$\beta = 2,4320$: É uma estimativa para a médias dos dados de RSP, se compararmos esse valor com a mdia apresentada na tabela 1, onde, $\bar{x} = 2,427$, percebe-se que os valores são muito próximos.

A figura 9 apresenta o mapa temático para os valores de RPS na camada de 0,1 a 0,2 m, dividido em 5 classes de tamanhos iguais.

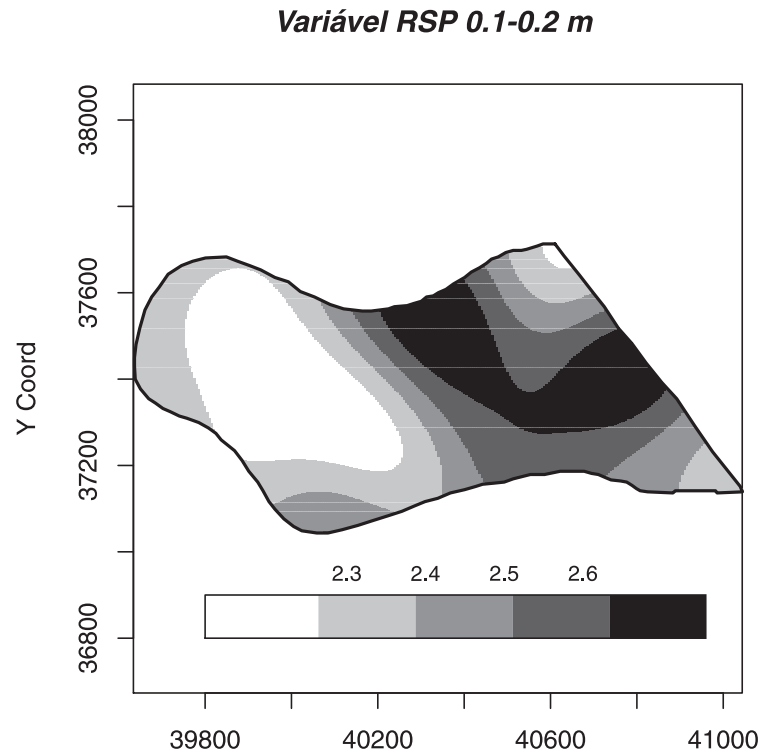


Figura 9: Mapa temático RSP com 5 classes iguais.

Pode-se observar no mapa temático (figura 9) que os menores valores de RSP encontram-se na parte leste (esquerda) do mapa, conforme foi apresentado no gráfico Post-plot (figura 4).

A tabela 6 apresenta a porcentagem de área referente a cada classe. Considerando que a área em estudo apresenta uma área total de 57,16 ha.

Tabela 6: Porcentagem de área em cada classe do mapa temático

Classes (MPa)	n de pixels	área (ha)	% da área total
2,202 – 2,301	29320	14,66	25,65
2,301 – 2,399	26579	13,29	23,25
2,399 – 2,498	17161	8,57	14,99
2,498 – 2,597	21181	10,59	18,53
2,597 – 2,695	20101	10,05	17,58
Total	114342	57,16	100

Pode-se observar pela tabela 6 que a menor classe (2,202; 2,301) corresponde a 25,65% da área em estudo. Além disso, se agruparmos as classes (2,202; 2,301) e (2,301; 2,399), sendo ambas inferiores ao valor da média ($\bar{x} = 2,427$ MPa) juntas correspondem a 48,9% da área total. A classe que contém os máximos valores (2,597; 2,695) de RSP ocupa 17,58% da área.

CANARACHE (1994) propôs um modelo para mostrar que os valores de RSP variam sob diferentes condições, tais como as combinações de texturas, densidades e teores de umidade, e a partir deste modelo concluiu de maneira geral que valores da RSP no intervalo [1,10;2,59] MPa apresentam poucas limitações ao crescimento das raízes, enquanto que valores no intervalo [2,6;5,0] MPa indicam que existem algumas limitações. Segundo esta informação foi possível reclassificar o mapa obtido anteriormente nestas duas classes e determinar a porcentagem de cada classe.

A figura 10 apresenta o mapa temático para os valores de RPS na camada de 0,1 a 0,2 m, dividido em 2 classes, conforme foi proposto por Canarache (1994). Pode-se observar no Post-plot (figura 4) que o ponto discrepante (valor mínimo), localizado na parte superior direita do mapa, faz vizinhança com pontos de alto valor, porém, nota-se que o ponto discrepante influenciou a região em que está localizado no mapa, pois toda a região superior esquerda do mapa, apresentou valores baixo de RSP de 1,10 a 2,59 MPa. A tabela 7 apresenta a porcentagem de área referente a cada classe.

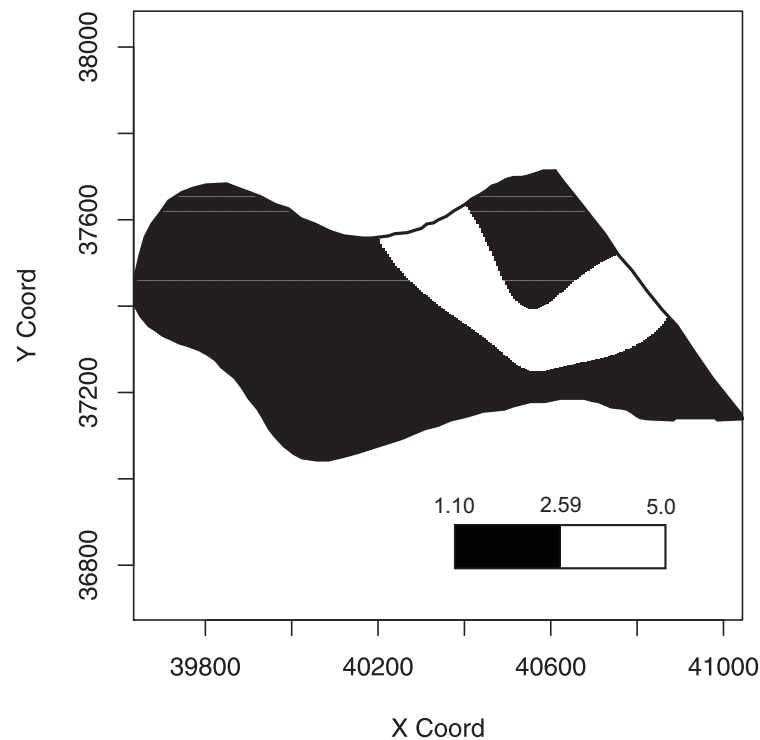


Figura 10: Mapa temático RSP com 2 classes.

Tabela 7: Porcentagem de área em cada classe do mapa temático

Classes (MPa)	nde pixels	área (ha)	% da área total
1,10 – 2,59	91695	45,84	80,20
2,59 – 5,00	22647	11,32	19,80
Total	114342	57,16	100

Pode-se observar pela tabela 6 que 80,20% da área em estudo apresenta poucas limitações ao crescimento radicular, pois, nessa área os valores de RSP está no intervalo

1,10; 2,59. Por outro lado 19,20% da área em estudo apresenta limitação ao crescimento radicular (2,6;5,0).

Com auxílio das análises geoestatísticas foi possível detectar que os valores de RSP apresentam fraca dependência espacial na camada de 0,1 - 0,2 m, deste modo, foi possível obter que o alcance de dependência espacial nessas camadas foram de 355,4 m, ou seja, os valores de RSP são influenciados pela distância entre os pontos de amostragem.

Com o estudo da RSP do solo foi possível perceber que na área em estudo há indícios de limitações ao desenvolvimento radicular na camada de 0,1 a 0,2 m, cerca de 80% da área experimental. Os valores de RSP devem ser monitorados, considerando também, os valores de densidade, porosidade e umidade do solo, para evitar que ocorra compactação do solo.

A compactação do solo deve ser evitada pois pode propiciar diminuição na produtividade, se os valores de RPS do solo são superiores a 6,5 MPa é indicado o uso de implementos adequados ou a implantação de culturas com sistema radicular com capacidade de romper a camada compactada, porém, na área em estudo os valores de RSP não ultrapassam de 3,709 MPa.