O INTERPOLADOR IMPORTA: COMO A ESCOLHA DO MÉTODO PODE MUDAR MAPAS E DECISÕES NA AGRICULTURA DE PRECISÃO

Laura Delgado Bejarano[1]; Lucas Rios do Amaral[2]

[1] Doutoranda, Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas – FEAGRI/UNICAMP, Campinas, SP – laura.bejarano@feagri.unicamp.br

[2] Professor associado, Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas – FEAGRI/UNICAMP, Campinas, SP – lucas.amaral@feagri.unicamp.br

Resumo

A interpolação é uma etapa importante na geração de mapas utilizados na Agricultura de Precisão, pois permite estimar valores em locais não amostrados a partir de observações pontuais georreferenciadas. Entretanto, diferentes métodos podem gerar resultados distintos, e o desempenho deles depende das características dos dados disponíveis. A densidade amostral, a distribuição dos pontos, a estrutura espacial da variável e a relação com covariáveis disponíveis são alguns dos fatores a considerar. A validação permite comparar os métodos e verificar qual deles apresenta melhor desempenho no conjunto de dados analisado. Resultados de pesquisa foram organizados em um framework para auxiliar na priorização dos métodos avaliados. Para levar essa abordagem à prática e facilitar essas etapas, desenvolvemos o *Best Fit Interpolator*, um plugin livre e de código aberto para QGIS que incorpora o framework, permite avaliar as características dos dados, comparar diferentes interpoladores por meio da validação cruzada e gerar os mapas de variabilidade. A partir desses aspectos, são apresentadas recomendações práticas para auxiliar na escolha do método e na elaboração de mapas mais consistentes.

Por que a escolha do interpolador merece atenção?

Um dos princípios da Agricultura de Precisão (AP) é reconhecer que uma área agrícola não apresenta condições uniformes ao longo de sua extensão. Propriedades do solo, produtividade, relevo, vigor das plantas e outros fatores relacionados ao sistema produtivo podem apresentar variabilidade espacial dentro de uma mesma área. Compreender essa variabilidade é fundamental para representar as diferenças existentes no campo e auxiliar na tomada de decisão de forma localizada.

Para isso, são realizadas medições ou amostragens georreferenciadas em diferentes pontos da área. No caso da fertilidade do solo, é comum o levantamento por meio de amostragens onde os valores de fósforo, potássio, pH e outros atributos são conhecidos apenas nos locais onde as amostras foram coletadas. Entretanto, a interpretação espacial desses atributos e sua utilização em aplicações de fertilizantes em doses variadas normalmente exigem uma superfície contínua que represente toda a área, tornando a interpolação uma etapa fundamental na construção desses mapas.

A interpolação utiliza os valores conhecidos nos pontos amostrados para estimar a variável nos locais onde não foram realizadas medições e parte do pressuposto que amostras próximas são mais parecidas do que amostras distantes entre si. Para isso,

podem ser utilizados diferentes métodos, como o inverso da distância ponderada (IDW), o spline de placa fina (SPF), a krigagem ordinária (KO) entre outros. Esses métodos possuem características e critérios distintos para estimar valores para os locais não amostrados, o que pode resultar em superfícies e padrões espaciais distintos a partir do mesmo conjunto de dados.

Na prática, isso significa que os mapas de fertilidade do solo e de prescrição de fertilizantes podem variar conforme o método utilizado (Figura 1). Essa diferença tende a ser ainda mais importante quando a quantidade de amostras é reduzida ou quando a variável apresenta uma estrutura espacial pouco definida. Por isso, um método que funciona bem em uma área, para determinado atributo do solo ou em uma condição de amostragem, pode não apresentar o mesmo desempenho em outra.

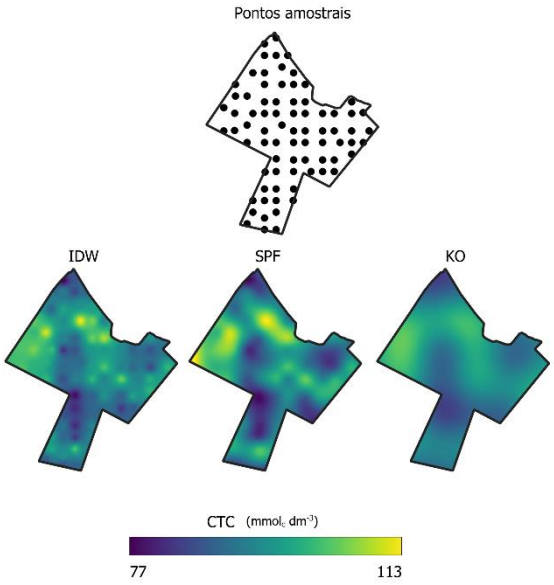


Figura 1. *Comparação entre superfícies interpoladas por IDW, SPF e KO a partir do mesmo conjunto de pontos amostrados da capacidade de troca de cátions (CTC) do solo.*

As diferenças entre os mapas gerados por diferentes métodos de interpolação devem ser avaliadas com cautela, pois o mapa do atributo do solo nem sempre é o produto final. Ele pode ser utilizado para definir doses de insumos, delimitar zonas de manejo ou direcionar novas amostragens. Embora os métodos apresentem padrões espaciais semelhantes, podem ocorrer diferenças nas doses recomendadas, como observado nos mapas de recomendação de calcário (Figura 2). Essas diferenças podem ser pequenas ou localizadas, mas sua relevância prática depende da magnitude das variações e da área afetada. Assim, a escolha do método de interpolação deve considerar não apenas a qualidade das estimativas, mas também seus possíveis efeitos sobre as decisões de manejo.

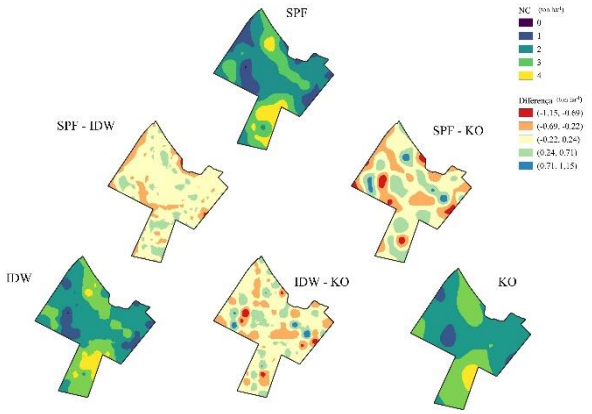


Figura 2. *Mapas de recomendação de calcário (NC) obtidos a partir das superfícies interpoladas pelos métodos IDW, SPF e KO e mapas das diferenças entre cada par de métodos.*

A escolha do método de interpolação

Não existe um único método capaz de apresentar o melhor desempenho em todas as situações. O desempenho dos interpoladores depende das características dos dados disponíveis, principalmente do número de amostras e da estrutura espacial da variável (Delgado Bejarano *et al.*, 2026). A distribuição dos pontos na área também pode influenciar esse desempenho. Nos resultados que obtivemos, o SPF apresentou bom desempenho em condições com poucas observações, enquanto a KO se tornou mais competitiva com o aumento do número de amostras e quando a estrutura espacial estava mais bem definida. Por isso, a escolha do interpolador deve considerar as características do conjunto de dados, e não apenas a preferência por determinado método.

O que determina o desempenho da interpolação?

Na prática, o planejamento da amostragem frequentemente parte da densidade amostral, ou seja, da relação entre o número de amostras e o tamanho da área, normalmente expressa em amostras por hectare. Embora seja um indicador útil da intensidade de amostragem, a densidade não deve ser interpretada isoladamente. Em áreas pequenas, uma densidade elevada pode resultar em pequenas distâncias entre os pontos, mas ainda fornecer um número total de observações insuficiente para representar adequadamente a variabilidade. Por outro lado, em áreas muito grandes, densidades mais baixas podem ser adotadas para reduzir os custos de coleta e análise. Nesses casos, mesmo com um número elevado de amostras, o espaçamento excessivo entre os pontos pode limitar a caracterização da variabilidade espacial.

Além da densidade, a distribuição dos pontos na área também deve ser considerada. Diferentes arranjos podem proporcionar coberturas espaciais maiores ou menores, mesmo com o mesmo número de amostras. O planejamento amostral pode ser otimizado utilizando informações auxiliares, como dados de sensoriamento remoto ou proximal, conhecimento agronômico prévio da área ou modelos espaciais, para orientar a distribuição dos pontos. Essas informações podem ajudar a identificar regiões com maior variabilidade (manchas pequenas) ou com padrões de variabilidade diferentes, onde uma amostragem mais intensa ou com pontos estrategicamente alocados pode ser necessária, e regiões mais homogêneas, onde o número de pontos pode ser reduzido e uma grade regular tende a representar bem a variabilidade presente. Dessa forma, busca-se melhorar a representação da variabilidade espacial e evitar amostragens redundantes, sem necessariamente aumentar os custos de coleta e análise. A Figura 3 ilustra como mudanças na densidade e na distribuição dos pontos podem alterar a cobertura da área (JAMSHIDI *et al.*, 2026; Pusch *et al.*, 2023)

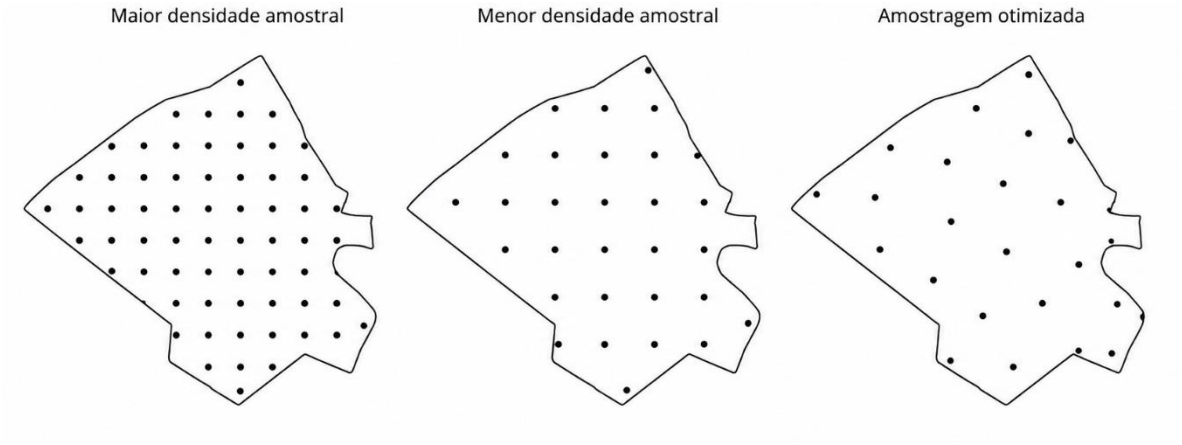


Figura 3. *Exemplos de configurações amostrais em uma mesma área, incluindo maior densidade, menor densidade e posicionamento otimizado das amostras.*

Outro aspecto importante é a estrutura espacial da variável, ou seja, a forma como os valores se organizam no espaço e como essa organização varia em função da distância. Quando valores próximos tendem a ser mais semelhantes entre si do que valores mais distantes, há dependência espacial. O Índice de Moran é uma métrica que permite avaliar a presença de autocorrelação espacial, enquanto o semivariograma descreve como a variabilidade entre pares de observações aumenta com a distância (Anselin, 1995; Oliver; Webster, 2015). A presença de uma estrutura espacial bem definida é fundamental para que a krigagem possa explorar essa relação durante a estimativa (Oliver; Webster, 2014).

Informações auxiliares, também chamadas de covariáveis, podem contribuir para a interpolação quando apresentam relação com o atributo que se deseja mapear. Dados de relevo, sensores proximais e sensoriamento remoto, por exemplo, podem fornecer informações adicionais sobre a variabilidade da área. Entretanto, a inclusão de uma covariável não garante melhor desempenho, pois sua contribuição depende da relação que apresenta com a variável de interesse (Pusch *et al.*, 2021), sendo necessário conhecimento técnico e análises preliminares

Como avaliar o método escolhido

A escolha do interpolador não deve ser baseada apenas na aparência do mapa. Uma forma de avaliar seu desempenho é por meio da validação cruzada, sendo que no caso de poucos dados, que geralmente é o caso da amostragem de solo, deve-se utilizar o método *leave-one-out* ("tire-uma-fora" - LOOCV). Nesse procedimento, uma observação é temporariamente retirada do conjunto de dados e seu valor é estimado com base nas demais amostras. Em seguida, o valor estimado é comparado ao valor observado, e o processo é repetido até que todas as amostras tenham sido utilizadas para validação. Essa estratégia permite avaliar o desempenho da interpolação utilizando todo o conjunto de dados disponível em diferentes etapas do processo (Piikki *et al.*, 2021).

A partir dos pares de valores observados e estimados, diferentes métricas podem ser utilizadas para avaliar o erro e a concordância. Uma medida de erro amplamente utilizada é a raiz do erro quadrático médio (RMSE), enquanto que uma medida interessante para comparar a concordância/correlação é o coeficiente de correlação de concordância de Lin (LCCC). Este índice verifica o quanto os valores estimados se aproximam dos valores observados em relação à linha 1:1, ou seja, correlação absoluta perfeita. Assim, além de indicar se existe associação entre os valores, o LCCC permite avaliar se as predições estão efetivamente próximas dos valores reais, sendo que valores mais próximos de 1 indicam maior concordância.

No entanto, embora essas métricas permitam avaliar quantitativamente o desempenho dos métodos, nem sempre é possível relacionar os erros observados com suas consequências agronômicas ou econômicas. Nesse contexto, a avaliação visual dos mapas por produtores e consultores também pode contribuir para identificar diferenças com possível relevância prática. Como apresentado na Figura 2, os métodos produziram mapas com padrões espaciais semelhantes, embora existam diferenças locais nas doses de calcário recomendadas. Determinar se essas diferenças representam alterações relevantes nos custos ou nos resultados agronômicos ainda é um desafio, sendo uma questão importante para pesquisas futuras.

Durante a pesquisa que publicamos na revista *Precision Agriculture* (Delgado Bejarano *et al.*, 2026), organizamos os resultados obtidos sob diferentes condições de amostragem e de estrutura espacial em um *framework* que sintetiza o desempenho dos métodos avaliados. Sua principal utilidade é orientar a priorização de métodos, reduzindo o número de alternativas e o tempo de processamento. Quando há disponibilidade computacional para avaliar diferentes interpoladores, o *framework* pode servir de referência, mas a escolha deve considerar o desempenho obtido no conjunto de dados analisado de acordo com o resultado da validação cruzada. Como sua construção se baseou em métodos e cenários específicos, suas indicações devem ser interpretadas com cautela.

Para levar essa lógica à prática, desenvolvemos o *Best Fit Interpolator*, um plugin livre e de código aberto para QGIS, no qual o *framework* foi integrado ao fluxo de análise. A ferramenta permite avaliar automaticamente as características dos dados, realizar a validação comparativa dos métodos considerados, indicar o interpolador com melhor desempenho e gerar a superfície interpolada. Além desse fluxo, o plugin disponibiliza ferramentas independentes para a interpolação por IDW, SPF, KO, Random Forest

(RF), *Support Vector Machine* (SVM) e Krigagem dos Resíduos da Regressão (*Regression Kriging*), permitindo avaliar a correlação com variáveis auxiliares, modelar semivariogramas, configurar e avaliar cada método separadamente, comparar métricas de desempenho e criar os mapas dos atributos analisados.

Em conjuntos de dados mais densos, como os de produtividade agrícola, diferentes métodos de interpolação tendem a apresentar desempenhos semelhantes, e a escolha pode depender também da capacidade computacional e do tempo de processamento. Nesse contexto, o *Best Fit Interpolator* permite operacionalizar a comparação entre métodos, facilitando a avaliação de seu desempenho e a escolha da alternativa mais adequada de forma prática e acessível.

Recomendações práticas

Antes de gerar um mapa interpolado, alguns cuidados simples podem tornar a escolha do método mais consistente. Mais do que utilizar sempre o mesmo interpolador, é importante observar as características do conjunto de dados e verificar se o método escolhido representa adequadamente a variabilidade da área.

Na prática, recomendamos:

- Tenha cautela ao interpolar conjuntos de dados muito pequenos. Como referência, com menos de 50 amostras, a quantidade e a distribuição dos pontos podem ser insuficientes para caracterizar adequadamente a dependência espacial por meio de semivariogramas, dificultando a aplicação de métodos geoestatísticos, como a krigagem (Kerry; Oliver, 2007). Nesses casos, outras estratégias, como a delimitação de zonas de manejo, podem ser mais adequadas. Quando a interpolação for necessária, os resultados devem ser interpretados com maior cautela, sendo que as diferenças entre os métodos de interpolação podem ser mais marcantes;
- Garanta uma amostragem representativa da área. Evite grandes regiões sem observações e, quando houver indícios de maior heterogeneidade, considere aumentar a densidade amostral. Informações já disponíveis da área, como imagens históricas de satélite, mapas de produtividade, histórico de manejo e conhecimento prévio do campo, podem ajudar a identificar essas características antes do planejamento da amostragem;
- Evite densidades muito baixas. Distâncias excessivas entre os pontos podem dificultar a representação da variabilidade e a caracterização da estrutura espacial, resultando em redução do retorno da aplicação de fertilizantes em doses variadas. Como referência, quando o alcance da dependência espacial é conhecido, recomenda-se utilizar espaçamentos entre pontos inferiores à metade desse alcance (Oliver; Webster, 2014);
- Considere o objetivo do mapa. Superfícies mais contínuas podem facilitar determinadas decisões de manejo, mas a escolha deve considerar a finalidade do mapeamento e o desempenho da interpolação. SPF e KO tendem a entregar superfícies mais suavizadas/contínuas;
- Teste diferentes interpoladores sempre que possível. Utilize o mesmo procedimento de validação, como o LOOCV, para comparar o desempenho dos métodos. Considere também a avaliação visual dos mapas e a experiência de produtores e consultores, pois diferenças nas métricas estatísticas nem sempre se traduzem em mudanças relevantes nas decisões de manejo;
- Avalie as covariáveis antes de utilizá-las. Verifique sua relação com o atributo a ser mapeado, por exemplo por meio da correlação. Quanto maior essa relação, maior tende a ser o potencial da covariável para contribuir com a amostragem dirigida e predição por métodos multivariados de interpolação.

Referências

ANSELIN, L. Local Indicators of Spatial Association—LISA. Geographical Analysis. v. 27, 2, p. 93–115. 1995.

DELGADO BEJARANO, L.; LOUREIRO GONÇALVES OLIVEIRA, A.; FIOLO POZZUTO, J. V.; CASTAÑEDA SÁNCHEZ, D.; RIOS DO AMARAL, L. Performance of interpolation methods in digital soil

mapping: the influence of data characteristics. Precision Agriculture. v. 27, 1. 2026.

JAMSHIDI, S.; CHERKAUER, K.; ARMSTRONG, S.; SMUKLER, S. M.; PAUL, S. S. Sample optimization for spatial prediction of soil properties using divergence-based representativeness metrics. Pedosphere. p. 1–23. 2026.

KERRY, R.; OLIVER, M. A. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. Geoderma. v. 140, 4, p. 383–396. 2007.

OLIVER, M. A.; WEBSTER, R. Basic Steps in Geostatistics: The Variogram and Kriging. Cham: Springer International Publishing. 2015.

OLIVER, M. A.; WEBSTER, R. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. Catena. v. 113, p. 56–69. 2014.

PIIKKI, K.; WETTERLIND, J.; SÖDERSTRÖM, M.; STENBERG, B. Perspectives on validation in digital soil mapping of continuous attributes—A review. Soil Use and Management. v. 37, 1, p. 7–21. 2021.

PUSCH, M.; OLIVEIRA, A. L. G.; FONTENELLI, J. V.; AMARAL, L. R. D. Soil Properties Mapping Using Proximal and Remote Sensing As Covariate. Engenharia Agricola. v. 41, 6, p. 634–642. 2021.

PUSCH, M.; SAMUEL-ROSA, A.; SERGIO GRAZIANO MAGALHÃES, P.; RIOS DO AMARAL, L. Covariates in sample planning optimization for digital soil fertility mapping in agricultural areas. Geoderma. v. 429, October 2022. 2023.