

RELAÇÃO DA AMOSTRAGEM DO SOLO COM A VARIABILIDADE DAS ESTIMATIVAS DE ESTOQUE DE CARBONO

Alfredo José Barreto Luiz¹, Cristiano Alberto de Andrade¹, Ruan
Carnier¹

alfredo.luiz@embrapa.br

¹ Embrapa Meio Ambiente – Jaguariúna, SP, Brasil.

1 Introdução

No âmbito do projeto Carbono+, uma cooperação entre a Embrapa e a Bayer Cropscience para a pesquisa científica, tecnológica, desenvolvimento e inovação com a finalidade de propor métricas, ferramentas e protocolos para a contabilidade do carbono na produção nacional de grãos, envolvendo soja e milho no Brasil, um dos objetivos é estabelecer protocolos para a amostragem de solos agrícolas. O foco é na determinação precisa do estoque de carbono (EC) presente no solo, importante para o estabelecimento do *baseline* nos projetos de carbono, bem como para auditorias sobre o possível sequestro de C. Há uma questão central nessa amostragem: qual o número de amostras necessárias para uma estimativa com precisão aceitável?

No planejamento de um levantamento amostral, sempre nos deparamos com a questão de decidir qual o tamanho da amostra (n). E essa decisão é realmente importante, pois uma amostra muito grande pode significar desperdício de recursos, enquanto uma amostra muito pequena pode reduzir a utilidade dos resultados ou mesmo torna-los inúteis para o objetivo pretendido. Além disso, nem sempre é possível tomar uma decisão inteiramente satisfatória, isso porque frequentemente não temos informação suficiente para estarmos certos que nossa escolha do tamanho amostral seja a melhor. Para contornar esses problemas, precisamos lançar mão da teoria da amostragem, que nos fornece uma estrutura dentro da qual podemos pensar inteligentemente sobre o

problema. Segundo Cochran (1977), entre as principais etapas envolvidas na escolha do tamanho amostral está o estabelecimento de requisitos sobre o que é esperado de uma amostra. Esses requisitos devem ser expressos em termos dos limites de erros desejados, ou em termos de algumas decisões ou ações que serão adotadas quando os resultados da amostra forem conhecidos. Embora a responsabilidade de estabelecer os requisitos seja dos demandantes do levantamento, eles frequentemente precisam de orientação dos analistas de métodos quantitativos para colocar seus desejos em termos numéricos.

Outro aspecto fundamental a ser considerado no planejamento amostral é que quase sempre existe um maior interesse por certas subdivisões principais da população, para as quais aqueles limites de erro desejados são definidos. Nesses casos, um cálculo separado pode ser feito para o n em cada subdivisão e o n total será encontrado por adição.

Por fim, Cochran (1977) nos lembra que em quase todo levantamento pode ser considerado uma amostragem multivariada, ou seja, são obtidas medidas de mais de uma variável, representativas de vários itens ou características. Ele diz ainda que esse número algumas vezes é grande e que nem todas estas variáveis são bem comportadas ou apresentam variância e distribuição semelhantes. Se calcularmos o tamanho de amostra mais adequado para uma, isso não garante que seja o mesmo tamanho ideal para outra variável. O que vai nortear a escolha da variável para determinar o tamanho da amostra é o objetivo principal do levantamento e grau de precisão que se exigirá para a estimativa dessa variável. A forma tradicional de amostragem de solos agrícolas tem como objetivo principal permitir a recomendação técnica da correção da acidez e da fertilização do solo, visando como resultado o ótimo econômico entre gasto com insumos e renda do produto. Para esse fim específico, existem recomendações técnicas bem detalhadas (QUAGGIO e CHIBA, 2022). No entanto, é importante entender a justificativa para os bons resultados no uso dessa amostragem recomendada, na qual uma única amostra de solo é analisada, mas ela foi composta, antes da análise, por 20 amostras simples coletadas em uma área considerada homogênea. Essa amostragem funciona adequadamente ao que se propõe porque o produtor agrícola, especialmente aquele com extensas áreas e uso de máquinas agrícolas para o manejo da sua cultura, precisa colher toda a sua área

disponível para plantio, ou seja, mesmo que existam regiões mais férteis e outras menos, ele não irá deixar de cultivar aqui e ali. Assim, o que se pretende ao fazer a amostragem para fins de manejo da fertilidade do solo é conhecer o teor médio dos elementos de interesse em uma grande área de terreno. O cálculo do corretivo ou do fertilizante que será usado é feito para toda aquela área (aqui é importante dizer que o avanço na agricultura de precisão pode mudar essa realidade, mas ainda não é o caso para a maioria dos agricultores), pois toda ela será cultivada, adubada, corrigida, manejada e colhida da mesma forma. Certamente, algumas partes do terreno receberão mais insumo que o necessário, e outras, menos, mas na média, o objetivo de obter a produção próxima ao ótimo econômico será atingido.

Entretanto, quando se consideram outros objetivos, a construção de uma amostra aleatória não é tarefa trivial e, na verdade, a dificuldade em obtê-la em levantamentos agrícolas vem desafiando os estatísticos envolvidos com esse tema há décadas (LUIZ et al., 2012).

Um aspecto importante a ser destacado é que os componentes do solo, por natureza, apresentam uma variabilidade espacial, tanto no sentido horizontal quanto vertical. A rocha da qual se originou o solo, o relevo, o intemperismo sofrido, a cobertura vegetal original, o manejo agrícola empregado, o tipo de cultura que cresceu ali no passado, tudo isso afeta a maior ou menor variabilidade de cada componente do solo, assim como das suas propriedades. Ou seja, mesmo que fosse possível obter e analisar um número enorme de amostras, ainda permaneceria um certo grau de incerteza sobre as estimativas calculadas usando os dados coletados.

No caso presente, o objetivo é estimar o estoque de carbono em áreas agrícolas. A recomendação presente no Boletim 100, uma importante referência no assunto de amostragem de solo para fins de manejo da fertilidade, traz que cada amostra composta (de 20 sub amostras) deve representar no máximo 10 hectares (QUAGGIO e CHIBA, 2022). Atualmente, entretanto, as áreas cultivadas com as chamadas commodities ocupam muito mais terreno, com talhões alcançando 300 a 500 hectares tornando quase impraticável a coleta com esse nível de detalhamento. Ao considerar o problema de forma tradicional, as unidades básicas de amostragem constituem-se em talhões agrícolas. Isso causa

um sério problema para o estabelecimento de amostras aleatórias, pois, para garantir a aleatoriedade, é necessário assegurar que todos os componentes da população tenham chance não nula e conhecida de pertencer à amostra. Para isso, é necessário que se conheçam todos os elementos de uma população finita ou a distribuição de probabilidade no caso de populações infinitas (LUIZ et al., 2012). Mas qual seria um 'componente' da população? No caso da amostragem para fins de fertilização, são coletadas fatias de solo com 20 cm de profundidade, que são coletadas em 20 pontos espalhados no talhão (de no máximo 10 ha) e depois misturadas para compor uma só amostra para ser analisada no laboratório. No caso específico do estoque de C do solo, calculado com resultados de teor de C e densidade, há que se considerar a impossibilidade de obtenção de amostra composta para a densidade, devido à natureza da variável e o procedimento de coleta.

2 Resultados e Discussão

No primeiro ano do projeto Carbono+ foram coletadas amostras de solo para determinação do teor de C e da densidade do solo, dentre outras variáveis de interesse, para obtenção do estoque de C do solo até 1 m de profundidade (usados aqui resultados para as camadas 0-5, 5-10, 10-20 e 20-30 cm), em 54 propriedades agrícolas (usados aqui banco de dados completos de 38 propriedades), em dois talhões agrícolas e em área de mata de mesmo tipo de solo, adjacente aos talhões agrícolas. Foram retiradas amostras indeformadas de solo para densidade e amostras deformadas de solo a partir de 14 trincheiras abertas em cada um dos talhões agrícolas. A área de mata não foi usada no presente estudo. Após esse primeiro ano do projeto, um dos talhões amostrados continuará a ser manejado pelo produtor conforme seus critérios e o outro passará a ser manejado com práticas sustentáveis recomendadas pela equipe do projeto. Porém, no momento inicial, os dois talhões podem ser considerados iguais com relação ao passado. Assim, dispomos de 28 amostras coletadas em cada fazenda, nas quais o solo foi separado em camadas de diferentes espessuras: de 0 a 5 cm; de 5 a 10 cm; de 10 a 20 cm; e de 20 a 30 cm de profundidade.

Para o cálculo do estoque de carbono até 30 cm (EC30) é necessário saber o teor de carbono em cada camada de solo, assim como a sua densidade. O estoque em cada camada é obtido pela multiplicação do percentual de carbono pela densidade e pela espessura de cada camada. Ao final, os estoques de todas as quatro camadas são somados para a obtenção do estoque total de 0 a 30 cm. Esse procedimento foi reproduzido em cada uma das 38 fazendas participantes do estudo, para cada um dos 28 pontos (trincheiras) de coleta considerados representantes de um mesmo talhão no primeiro ano, antes da diferenciação do manejo.

Foram, portanto, calculadas as médias, os erros padrão (EP) e os coeficientes de variação para o EC30 em cada fazenda, usando para isso os 28 valores amostrados em cada uma. A partir das 38 fazendas foram escolhidas três delas, com valores máximo, médio e mínimo de coeficiente de variação (CV) para o estoque de carbono de 0 a 30 cm (Tabela 1).

Tabela 1. Resultados médios e medidas de variabilidade para estoque de carbono no solo até 30 cm de profundidade (EC0a30), para camadas estratificadas até 30 cm (0 a 5, 5 a 10, 10 a 20 e 20 a 30 cm, respectivamente EC0a5, EC5 a 10, EC10a20 e EC20a30) e tamanho amostral estimado para obter um CV de 10%.

	Fazenda	EC0a30	EC0a5	EC5a10	EC10a20	EC20a30
Média	Max	99,7	23,1	20,0	32,3	26,9
	Med	83,6	22,3	15,8	25,6	19,8
	Min	79,2	16,9	14,3	25,1	22,9
Variância	Max	873,1	33,8	30,5	134,4	87,0
	Med	187,8	13,1	13,1	47,5	33,3
	Min	22,3	6,2	2,2	2,6	3,1
Erro Padrão EP	Max	5,6	1,1	1,1	2,2	1,8
	Med	2,6	0,7	0,7	1,3	1,1
	Min	0,9	0,5	0,3	0,3	0,3
Coeficiente de Variação CV(%)	Max	29,6	25,2	27,6	35,9	34,7
	Med	16,4	16,2	22,9	26,9	29,1
	Min	6,0	14,7	10,3	6,4	7,7
Tamanho amostral para CV = 10%	Max	8,8	6,4	7,6	12,9	12,0
	Med	2,7	2,6	5,2	7,2	8,5
	Min	0,4	2,2	1,1	0,4	0,6

Nesse breve trabalho, foram utilizados estes dados para demonstrar como a estratificação física das amostras pode alterar a variabilidade das estimativas.

Na sequência, a partir do conjunto de 28 valores de cada fazenda, foram extraídas amostras aleatórias, com reposição, pelo método conhecido como *bootstrap*, variando o tamanho da amostra de 1 até o máximo de 28, repetidamente por 10.000 vezes. Dessa forma, é como se tivéssemos coletado 10 mil vezes uma única amostra de solo da fazenda, ou duas amostras, ou três, até 28 amostras. Em cada uma dessas vezes, foi possível calcular a média, o erro padrão, a variância e o CV da estimativa do estoque de carbono (exceto para o EP e o CV da amostra tamanho 1, que não existem, pois de uma observação não é possível estimar a variabilidade). Depois, calculamos a média das 10 mil médias, dos 10 mil EP, das 10 mil variâncias e dos 10 mil CV para cada tamanho de amostra.

Nas figuras 1 e 2 são apresentados os CV das três fazendas selecionadas. Em cada figura, são comparados os comportamentos do CV em função do tamanho da amostra, em cada fazenda, para o valor do EC30 e para o valor em cada uma das camadas que compõem esse estoque.

Na figura 1, podemos observar as curvas de CV para o EC30 e o EC de 0 a 5 cm de profundidade (EC0a5) nas três fazendas. Na fazenda que apresentou o maior valor original de CV para EC30, o CV foi menor de 0 a 5 centímetros de profundidade do que de 0 a 30, nas duas outras fazendas o CV foi praticamente igual ou menor de 0 a 30 do que de 0 a 5.

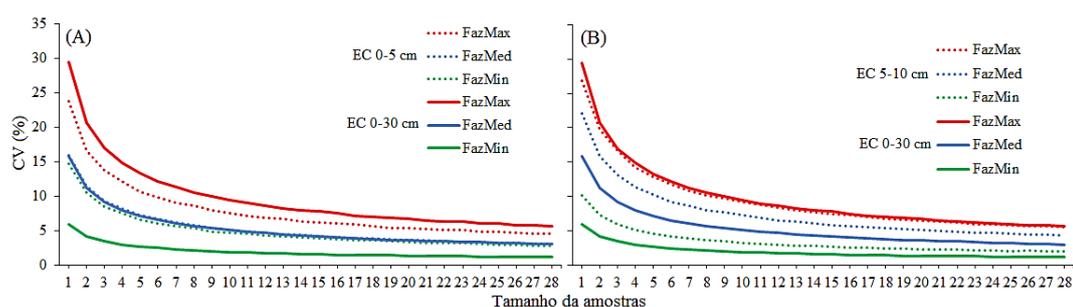


Figura 1. Comparação das curvas de CV, em função do tamanho da amostra simulada por bootstrap, entre o estoque de carbono acumulado de 0 a 30 e o determinado de 0 a 5 cm (A) e de 5 a 10 cm (B), em três fazendas.

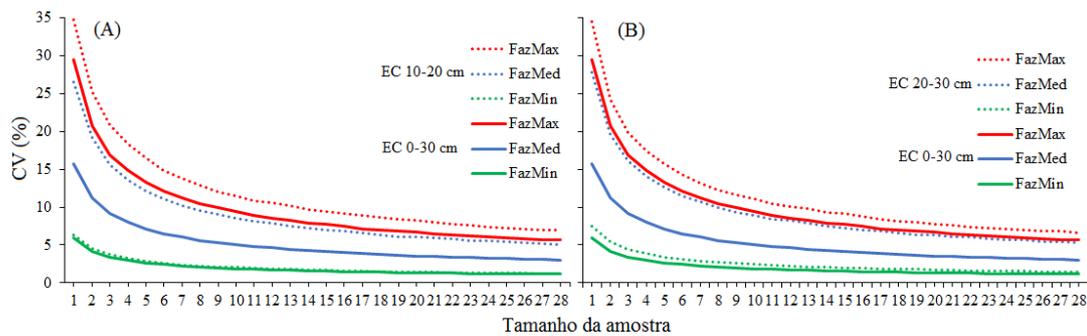


Figura 2. Comparação das curvas de CV, em função do tamanho da amostra simulada por *bootstrap*, entre o estoque de carbono acumulado de 0 a 30 e o determinado de 10 a 20 cm (A) e de 20 a 30 cm (B), em três fazendas.

Em todas as comparações, para as fazendas com CV originalmente médio ou baixo, as estimativas feitas para o estoque de carbono acumulado de 0 a 30 cm sempre resultaram em valores de CV mais baixo, com qualquer tamanho de amostra utilizado. Apenas em dois casos, para a fazenda com CV originalmente máximo, entre 0 e 5 cm e entre 5 e 10 cm de profundidade, o CV da estimativa do estoque de carbono acumulado de 0 até os 30 cm foi ligeiramente superior ao das camadas individuais. O mesmo pode ser observado na tabela 1, na qual eventualmente o tamanho de amostra necessário para produzir um CV de 10% pode ser menor em uma ou outra camada do que o estimado para as camadas somadas de 0 a 30 cm. Como em um levantamento amostral com mais de uma variável de interesse, o tamanho amostral deve satisfazer o critério de erro para o pior caso, entretanto, sempre o tamanho amostral estimado para EC30 foi menor que o maior entre as camadas individuais.

Isso quer dizer que o fato de empregar amostras fracionadas, ou estratificadas, coletando e analisando no laboratório as camadas de diferentes espessuras e profundidades, para depois calcular o estoque de carbono de 0 a 30 cm e só então calcular a variabilidade entre os pontos amostrais, melhora a precisão da estimativa em comparação a análise da variabilidade do estoque de cada camada separadamente.

Se o planejamento amostral fosse feito com base na variabilidade das camadas, seria necessário um número maior de pontos amostrais para obter a mesma variabilidade, quando comparado ao número de amostras necessárias para a estimativa do estoque de todas as camadas somadas.

3 Conclusão

Como afirmado por Cochran (1977), se formos calcular o melhor tamanho de amostra, segundo algum requisito do demandante da pesquisa, a partir de cada uma das variáveis observadas, teremos de escolher, para um mesmo erro amostral da estimativa, aquele tamanho que atende aos requisitos no pior caso. Assim, para o presente objetivo de estimar o estoque de carbono até os 30 cm, embora as amostras sejam estratificadas fisicamente, na prática, tanto para a coleta como para a análise dos seus componentes, é mais vantajoso calcular o tamanho de amostra a partir da variabilidade encontrada para o estoque de 0 a 30 e não sobre cada camada separadamente. O tamanho amostral igual a 9 (inteiro maior que 8,8) se mostrou o número de pontos coletados suficiente para, no pior caso, garantir um CV de 10% na estimativa da média.

Agradecimentos

Os autores agradecem à Embrapa e à Bayer pelo financiamento do projeto que permitiu a realização deste trabalho.

Referências bibliográficas

- COCHRAN, W.G. **Sampling techniques**. 3ed. New York: John Wiley & Sons, 1977. 428p.
- LUIZ, A. J. B.; FORMAGGIO, A. R.; EPIPHANIO, J. C. N.; ARENAS-TOLEDO, J. M.; GOLTZ, E.; BRANDÃO, D. Estimativa amostral objetiva de área plantada regional, apoiada em imagens de sensoriamento remoto. **Pesquisa Agropecuária Brasileira**, v. 47, n. 9, p. 1279-1287, 2012.
- QUAGGIO, J. A.; CHIBA, M. K. Amostragem do solo para fins de fertilidade. In: CANTARELLA, H.; QUAGGIO, J. A.; MATTOS JR, D.; BOARETTO, R. M.; RAIJ, B. van. (Eds.). **Recomendações de adubação e calagem para o Estado de São Paulo**. Campinas: Instituto Agrônomo de Campinas, 2022. p. 17-27. (IAC. Boletim Técnico, 100).