A 7 part series about soil sampling for turfgrass

Micah Woods

9 February 2020

Part 1

After I read the article by Lawrence et al. on Guiding soil sampling strategies using classical and spatial statistics: A review, I made some calculations, sent a few email inquiries, looked up some standard recommendations for turfgrass testing, reread another article on this topic, sent out a Twitter poll, and started drafting a blog post.

I soon realized that it might be a long blog post. Too long, probably.[1] And the topic of soil sampling and the number of cores (or subsamples) to collect to then put together as a single composite sample (Figure 1) representative of the tested area—that's a particularly dry subject, isn't it? So I decided to break that post I'd drafted into a number of short ones.

In these posts, I'll be considering soil nutrient analyses. I'll often use individual putting greens as an example.

A standard recommendation seems to be what I found in the Soil Sampling Instructions for Turfgrass from Rutgers.

For testing of a single putting green, they recommend a composite sample be made up of 12 to 15 subsamples.

The instructions read,

> "Plan to collect multiple subsamples randomly within each defined area to obtain a representative sample. For a uniform sports field or fairway, 20 to 30 subsample locations may be needed, whereas 12 to 15 subsamples may be adequate on a golf green or tee ... placing the soil subsamples in a clean plastic bucket to obtain a composite sample."

Part 2

The instructions from Rutgers are to take from 12 to 15 samples from a single putting green. I like to check Penn State recommendations too—their websites are a wealth

[1] This document is the 7 blog posts from www.asianturfgrass.com, put together in order.

Figure 1: How many cores should be taken from a single putting green, or sports field, or fairway, or lawn, to put together as a composite sample for soil nutrient analysis?

of reference materials. A good starting point is the Professional Turfgrass Management page with its long list of documents. My favorite from that page is *Turfgrass Fertilization: A Basic Guide for Professional Turfgrass Managers*.

On the soil sampling instructions page, the instructions for turf soils are:

> "Using a soil sampling tube, auger or trowel, and a clean pail, obtain thin slices or borings of soil from 12 or more locations."

I'm not sure where the **12** comes from, but Penn State recommends a minimum of 12 subsamples be put together as a composite sample for an area, and so does Rutgers. This seems to be the standard recommendation.

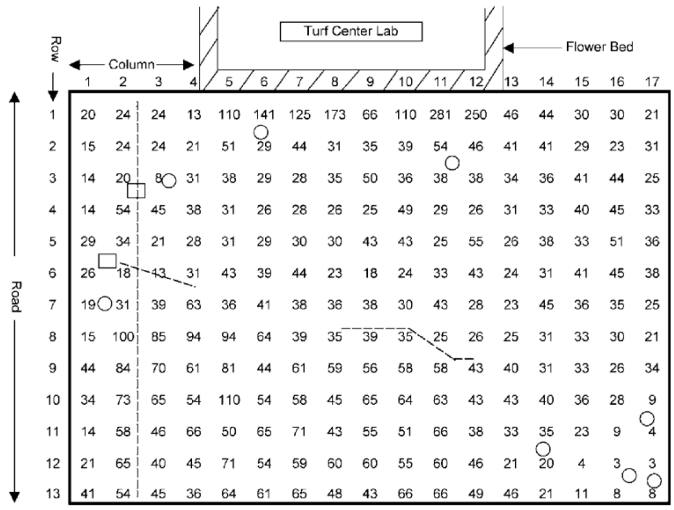
Part 3

There's an article by Donohue, Evaluation of soil nutrient variability for development of turfgrass soil test sampling methods, that describes results from sampling a 2,000 m^2 (21,520 ft^2) lawn at the Virginia Tech Turfgrass Center Laboratory.

221 possible sampling locations were identified, on a 3 m (9.8 foot) grid pattern across this area (Figure 3). I'll now give some quotes directly from the article:

> "To determine the optimum number of subsamples required for a representative composite sample for the turf area, the overall landscape was

Figure 2: Penn State's websites contain a wealth of reference materials about turfgrass management.



Row	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	20	24	24	13	110	141	125	173	66	110	281	250	46	44	30	30	21
2	15	24	24	21	51	29	44	31	35	39	54	46	41	41	29	23	31
3	14	20	8	31	38	29	28	35	50	36	38	38	34	36	41	44	25
4	14	54	45	38	31	26	28	26	25	49	29	26	31	33	40	45	33
5	29	34	21	28	31	29	30	30	43	43	25	55	26	38	33	51	36
6	26	18	43	31	43	39	44	23	18	24	33	43	24	31	41	45	38
7	19	31	39	63	36	41	38	36	38	30	43	28	23	45	36	35	25
8	15	100	85	94	94	64	39	35	39	35	25	26	25	31	33	30	21
9	44	84	70	61	81	44	61	59	56	58	58	43	40	31	33	26	34
10	34	73	65	54	110	54	58	45	65	64	63	43	43	40	36	28	9
11	14	58	46	66	50	65	71	43	55	51	66	38	33	35	23	9	4
12	21	65	40	45	71	54	59	60	60	55	60	46	21	20	4	3	3
13	41	54	45	36	64	61	65	48	43	66	66	49	46	21	11	8	8

Figure 3. Variation in soil P in the sampled area. Circles (○) indicate location of trees, boxes (□) are storm drains and dashes (– –) give location of drainage strips.

Figure 3: Figure 3 from Donohue (2002) showing Mehlich 1 phosphorus (P) test results; the units are mg P/dm^3 (mg/L).

first evaluated to identify unusual spots that one should avoid when sampling. This evaluation was made after the grid was marked off, but before samples were actually collected."

Donohue then excluded some areas near a road that may be affected by salt application, areas near a flower bed that may have had fertilizer overthrow, and steeply sloped areas and low areas.

> "This left a total of 163 spots for the subsampling study."

> "Soil test results from the 163 locations within the uniform area were entered into the Virginia Tech mainframe computer for evaluation. For determining the optimum number of subsamples for a representative composite sample, 1, 5 10, 15 and 20 subsamples/composite sample were chosen for evaluation. For each subsample number, five separate composite samples were 'collected' by the computer and compared to determine variation in test results."

In short, random samples were drawn from the 163 soil tests (for n = 1, 5, 10, 15, and 20), then these random samples were averaged as if they were a composite sample, and the means of those random samples and their variation were compared to the known mean of all 163 soil tests.

And here's the final result.

> "Within the uniform area selected in the study, the number of subsamples required for a representative composite sample varied from five subsamples for soil pH to 20 or more for P. Based on the results of this study, 20 subsamples per composite is suggested for a representative sample from a 1,000—2,000 m^2 turfgrass area."

Part 4

I've recommended taking at least 5 subsamples per green, for as long as I can remember.

That's less than the 12 to 15 recommended by Rutgers, less than the 12 or more in the Penn State instructions, and way less than the 20 subsamples per composite recommended by Donohue.

The main reason I've been comfortable with 5 subsamples per green is the consistency of the rootzone material (Figure 4) on most putting greens. And there is also the thought that I'm collecting soil samples to do soil nutrient analyses, with those results to be used to make a fertilizer recommendation. The fertilizer recommendation will be for the *course*; I rarely make fertilizer recommendations for individual greens. So I've been comfortable having smaller numbers of subsamples per green, realizing that

Figure 4: Spreading a uniform rootzone mix when building a putting green at Daitakarazuka Golf Club.

doing so lets me see some of the variability that exists across a property, and at the same time knowing that the total number of subsamples from the *course* is probably going to be 30 or more, because I recommend testing a minimum of 6 greens.

The Donohue article concludes that 20 subsamples are recommended from a 1,000 to 2,000 m^2 area. That's 1 to 2 subsamples per 100 m^2, which works out to a minimum of 5 subsamples per 500 m^2 (5,380 ft^2). That's about the size of a typical putting green.

Part 5

I'd read the soil sampling instructions from Penn State and Rutgers, suggesting a minimum of 12 subsamples be combined as a composite sample for each area. When I give instructions for sampling, I suggest a minimum of 5 subsamples be collected per green (or per area). And the Donohue article I've written about suggests 20 samples be collected from an area about the size of 4 typical putting greens.

I wondered what other people do, so I sent out this poll on Twitter, asking "When collecting samples for soil nutrient analysis on golf course putting greens, how many subsamples do you collect to make a single composite sample for one putting green?"

The results in Table 1, excluding those who answered the survey just to see results, came to a total of 182 respondents.

It seems a lot of turfgrass managers are following the standard recommendations of at least 12 subsamples per area, althought more than half the respondents are collecting

Figure 5: The standard recommendation for turfgrass seems to be a minimum of 12 samples for an area the size of a golf course putting green.

Table 1: Percent of poll respondents collecting this number of subsamples per green to make a composite sample.

Subsamples	Percent
1 to 5	12%
6 to 10	45%
11 or more	43%

fewer than the recommendation of 12. I was surprised there were so few respondents collecting 5 or less subsamples—I would have thought more people would do this the fast and easy way.

Part 6

Now this series gets interesting. I've reviewed what I do, what a Twitter survey says people are doing, what Rutgers and Penn State recommend, and what an intensively sampled lawn in Virginia suggests would be an appropriate number of subsamples to combine in composite samples for turfgrass.

The article Guiding soil sampling strategies using classical and spatial statistics: A review by Lawrence et al. recommends doing things differently. In short, they recommend that "soil samples should not be composited." I recommend reading the full article if you are interested. The article is about testing agricultural fields. It is quite relevant to turfgrass sites too. Below are some quotes from the article, with notes from me interspersed. Any italicized or bold text is mine.

They get right to the purpose of soil nutrient analysis in the first sentence saying that it is "a key practice to increase the efficiency of nutrient management." As an aside, I'll take the opportunity here to point out that that is how soil testing is supposed to work, but in the turfgrass industry, soil testing seems to actually *decrease the efficiency* of nutrient management because of misinterpreted soil tests. But it is easy to fix this with a method such as MLSN.

In their review, Lawrence et al. looked at results from other articles, and calculated "an estimate for the range of soil sampling densities that would be required to achieve a margin of error of 10% of each study's mean at a 5% precision level." For K, the median sampling density is less than 5 samples per hectare; for P the median sampling density is 8.4 samples per hectare. I think it is interesting to consider these sampling densities in comparison to what is standard in the turfgrass industry today. To achieve that aforementioned level of precision with P, for example, one would need on average a single sample for every 1,190 m^2 (12,809 ft^2). With K, fewer samples would be required.

> "Compositing soil samples before analysis is a common method for reducing soil analysis costs ... Compositing, which effectively calculates the mean of a number of soil cores, is unlikely to represent the true population parameter of interest and will typically present a higher estimate of nutrient concentration than is accurate ... the distribution of a nutrient would be positively skewed ... For fertilizer management, this may cause the farmer [or the turfgrass manager] to apply **an inadequate amount of fertilizer**"

Then in a section on the "Current State of Extension Soil Sampling Recommendations," this:

> "54% of extension sources suggested 'zig-zag,' 'Z', or 'W' sampling, with 44% mentioning the need to take 'representative' samples. This suggestion is at odds with the requirements of a design-based approach in which random selection of locations is paramount. As a result, **it is highly likely that these methods often result in biased results, especially if the samples are composited before analysis and the soil property is log-normally distributed.**"

In the general suggestions and conclusion to the article, more things to think about:

> "If no prior information on the soil property is available, log-normality should be assumed, **and soil samples should not be composited.** If the collected cores display log-normality, the geometric mean or median should be used instead of the arithmetic mean. **Only in instances in which normality has previously been established should soil samples be composited.**"

> "... it is not surprising that extension recommendations often suggest practices such as compositing and Z-sampling, which do not have a strong foundation in peer-reviewed literature. **Fortunately, there are theoretically simple ways to correct these recommendations, such as suggesting soil cores to be individually analyzed.**"

Part 7

The summary so far is this:

- standard recommendations for turfgrass sampling are to take 12 or more cores from each area and composite them into one sample
- I've been taking 5 cores and compositing them
- Donohue's research in Virginia led him to recommend taking 20 subsamples and compositing them
- The article by Lawrence et al. suggests that cores should not be composited

I've been wondering if my usual practice of 5 subsamples per green were enough (or too many). Last year I took 30 cores from a 1,092 m^2 double green in Bangkok (Figure 6) and tested them individually. That works out to 1 sample every 36 m^2.

Figure 6: 30 cores were collected from this double green in Bangkok and the cores were tested individually at Brookside Laboratories.

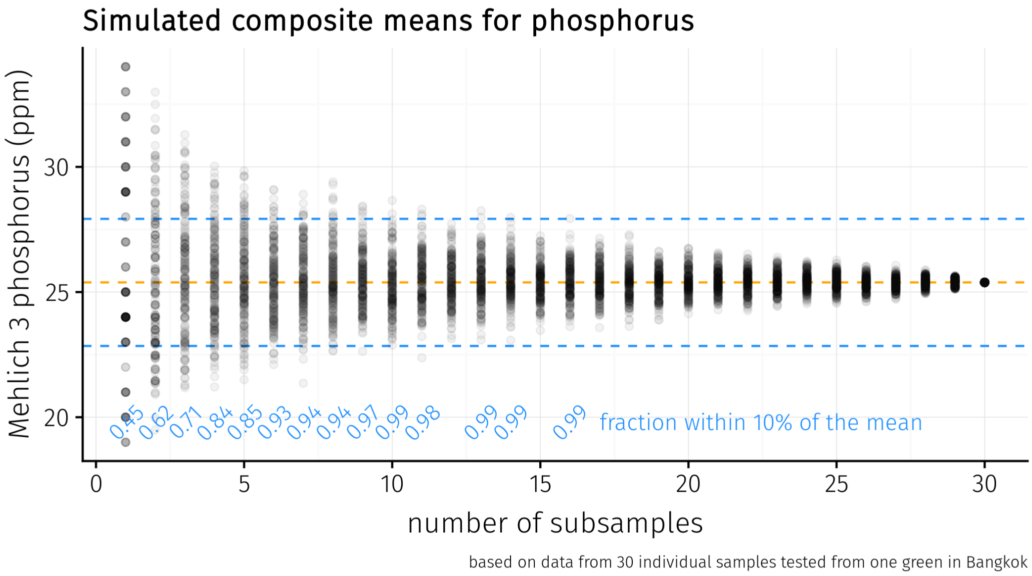


Figure 7: 200 simulated means for 1 to 30 cores calculated from the soil test results from a 1,092 m^2 green. The lines show a 10% margin above the known mean (calculated from all 30 cores) and a 10% margin below the mean.

I made some calculations about the number of samples required. I'll use the Mehlich 3 phosphorus results as an example. These are the soil test phosphorus data for those 30 cores in units of mg/kg.

`31 23 29 25 25 29 24 27 26 20 19 30 29 23 24 28 21 24 32 32 22 20 33 24 23 34 24 25 24 21`

I took random draws from those results, selecting 1 core, or 2, or 3, all the way up to 30 cores. And from those cores that were randomly drawn, I calculated the geometric mean, as recommended by Lawrence et al. I did that 200 times, and then plotted the means that were obtained from taking 1 up to 30 cores (Figure 7). This is simulating composite sampling from 1 core up to 30 cores on the same green.

Looking at the fraction of the simulations that returned a mean within 10% of the known mean, the fraction goes above 90% (marked with a horizontal dashed line in Figure 8 below) once there were 6 or more subsamples.

For phosphorus, once 6 cores were composited from this area, the mean soil test P was within 10% of the known mean more than 90% of the time.

I made the same calculation for other elements, and it was 6 subsamples for K, 5 for Ca, 4 for Mg, and 6 for S.

This green was 1,092 m^2 so 6 cores from this area would be 1 core from every 182 m^2. Another way to express this is 0.55 cores required per 100 m^2 or 0.51 samples required per 1,000 ft^2. That's how many samples would be required to get a test result

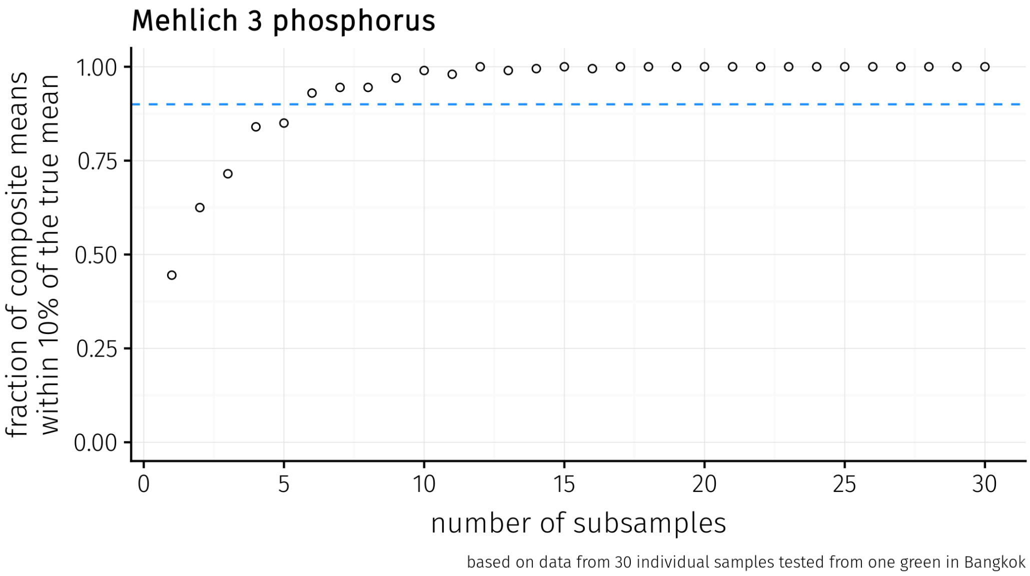


Figure 8: The fraction of simulated means that were within 10% of the known mean, calculated from 200 random draws from the test data.

within 10% of the true mean of the area, 90% of the time.

The 5 subsamples I've been compositing per green is about twice that sampling density; if nutrient variability within the green I tested in Bangkok is similar to other greens, then I am probably getting a mean value within 10% of the true mean almost 100% of the time.

The implications of all this include:

- a few subsamples per green are probably enough to get a result close to the mean value of the green
- current recommendations for 12 or more subsamples seem to be excessive, but I need to check variability on more greens
- testing individual cores, rather than composited samples, is probably a better way to do soil testing