Commenter J

Nitrogen Management Project Protocol
Minimum Data Standard Public Comment Period

Reserve Spatial Sampling Memo Intro:

I'm not sure the right question is being asked. Experimentally, to have statistical confidence in a treatment effect (alternative management practice), it's not primarily a question of how many chambers per field but rather of how many fields (or replicate blocks) per treatment. If each of a paired field has 100 chambers there remain zero degrees of freedom for testing the treatment (management) effect. If one has a limited number of chambers to deploy (say 20), it is much more powerful to put one chamber in each of 20 fields each randomly assigned to 2 management practices (e.g. tillage) than any other sort of experimental arrangement. Putting more chambers in each field will not affect the power to detect differences, it will simply increase one's confidence in the mean value from that particular field (which is valuable but does not affect treatment confidence).

If the Reserve accepts experimental evidence for a management effect based on 1 treatment replicate (for example, 2 fields, each treated differently), regardless of the number of analytical or chamber replicates, the evidence will be immediately suspect. In the statistical literature this is known as pseudoreplication and conclusions are fraught with all sorts of problems ranging from legacy to positional effects. It sounds from the description above that this is a risk, but perhaps I'm misreading.

Reserve Spatial Sampling Memo Question 1:

(a) For long-term sampling we recommend one chamber per treatment replicate, and a minimum of 4 treatment replicates. A treatment replicate is, in this context, a "field." So, no fewer than 4 fields, one plot per field, one chamber per plot.

For short-term experiments, replication varies by the strength of the expected treatment effect. For a recent small plot experiment, Ruan and Robertson (2013, Global Change Biology 19, 2478–2489) used 2 replicated NT and CT plots in each of 3 replicate fields (3 fields x 2 treatments x 2 replicate plots) for a total of 16 plots with 2 chambers per plot for a total of 32 chambers to detect a single treatment effect (CT vs. NT).
(b) One chamber per plot in the long-term experiment, 2 chambers per plot in the short-term.
(c) 50 x 50 cm
(d) There are no detectable differences with only one set of paired fields. It depends on the time frame and the degree of differences expected.
(e) Chambers are placed to ensure between-row and within-row coverage.
(f) For a short-term experiment, I would expect a minimum number of 3 fields per treatment to be adequate, with 1-2 plots per field and 1-2 chambers per plot. Whether the number is sufficient or overkill will be borne out by the statistics – the more fields, the more power for detecting significant management effects. I would not be too prescriptive here – rather be very strict about the level of statistical confidence acceptable: $p<0.05$ with at least 2 field seasons to cover climate variability is generally considered acceptable scientific practice.
(g) See context statement above: Midwest field crops.

(h) No – so long as statistical significance is demanded over at least two field seasons, and the basic experimental design is statistically sound (no pseudoreplication, for example), the answer should not change.

Reserve Spatial Sampling Memo Question 2:

One is looking for a placement effect, and this placement effect should be sufficiently robust to express itself across fields. The correct test for this – as it is for any management effect – is to establish plots of functional chambers in each of n replicate fields. Then to compare percentages of fluxes – not fluxes – within functional groups across fields.

For example, in the situation above: Plot1, Plot2, and Plot3 should be in different fields. In field 1, the furrows, berms, and fertilizer bands will individually contribute 20%, 40%, and 40%. In field 2, the percentages will be 18, 45, and 47. In field 3, the percentages will be 22, 41, and 37. The average percentages can then be tested for significance and as appropriate applied in an area-weighted manner to the fluxes from each field to provide a weighted flux estimate. Should resources allow, doing multiple plots per field would also be welcome, and the percentages within fields averaged by field. But the final comparison would still involve only 3 fields (2 degrees of freedom) to test the placement effect.

Reserve Spatial Sampling Memo Question 3:

I don't think it's possible to be too prescriptive. Rather, allow respondents to submit their data and analysis to peer review, which would include both gas flux experts as well as a statistician, and with the expectation that sampling designs meet acceptable standards with respect to:
(a) Robust experimental design – appropriately replicated treatment plots blocked in multiple fields, with chambers placed in such a way to capture management-imposed variability (e.g. within row vs. between row).
(b) A regular and event-based sample frequency appropriate to the season for at least 2 years.
(c) Three or more samples taken per chamber closure period (flux period).
(d) Fluxes weighted by management-imposed variability as appropriate.
(e) Treatment effects significant at the p=0.05 level.