Statistical Challenges in Analyses of Chamber-based Soil CO₂ and N₂O Emissions Data

Measurements of soil greenhouse gas (GHG) emissions have gained a lot of recent attention in an effort to potentially increase agriculture’s role in mitigating climate effects. However, it may not be well recognized that the nature of chamber-based GHG data is such that analyses require advanced statistical techniques to fully explore experimental treatment effects. Moreover, for soil GHG data, some experimental design approaches can enhance while others can weaken an experiment’s ability to detect treatment differences.

In the January–February 2015 issue of the *Soil Science Society of America Journal*, researchers from the Department of Plant, Soil, and Microbial Sciences and the Kellogg Biological Station of Michigan State University identified and explored the implications of key choices in experimental design and statistical analyses relevant to chamber-based soil GHG studies.

Among the statistics-related frequently asked questions puzzling the agricultural scientists who embark on conducting GHG monitoring experiments are: (1) Should I block my experimental plots? If so, what would be the best criteria for conducting the blocking effectively? (2) I have only a limited amount of resources for my GHG monitoring. Should I focus on monitoring a greater number of experimental plots? Or should I have multiple monitoring stations in just a few experimental plots? (3) What type of statistical data analysis should be used after I collect my data and how do I implement these analyses?

This study addresses some of these questions and provides suggestions for selecting optimal data analyses strategies along with sample SAS codes for implementing them. The actual data for CO₂ and N₂O monitoring from three agricultural experiment sites were used for illustrations. The sites are the Main Cropping System Experiment of the Kellogg Biological Station’s Long-term Ecological Research site and two sites from the Sustainable Corn experiment network.

While many observational studies noted only relatively weak (for CO₂) or non-existing (for N₂O) correlations between fluxes and landscape topography, the topography-centered experimental design of the Sustainable Corn sites enabled the researchers to detect topographical effects on spatial distribution patterns of both gases, especially in drier soil. Thus, blocking by topography can be an important consideration in increasing efficiency of GHG experiments. Increasing the number of treatment replicates, that is, the numbers of plots per treatment, is the best strategy for increasing statistical power of the comparisons among the studied treatments in GHG studies. However, increasing the intensity of sub-sampling at a variety of levels, including increasing the number of gas-measuring chambers per plot and the number of headspace measurements per flux determination, can also bring notable improvements in statistical power. That is, it can raise the experimenter’s chances of detecting statistically significant differences among the treatments.

The nature of GHG data is such that it almost invariably brings about violations of two of the key assumptions of the regular analysis of variance, that is, the assumption of independence between the observations and the assumption of homogeneity of variances. Temporal autocorrelations are likely to be present in repeated flux observations taken from the same gas-chamber again and again during the measuring season, thus violating the independence assumption while occasional high-flux events are likely to produce data of much greater variability than that observed on a regular basis. The solutions are to use the repeated-measures analysis and the analysis that accounts for heterogeneous variances. Unfortunately, these statistical tools remain underutilized by the GHG experimenters. The judicious application of these tools can be crucial for the efficient analysis of GHG data.