A common soil handling technique can generate incorrect estimates of soil biota effects on plants

Introduction

An active area of research seeks to understand how soil biota effects on plants vary across experimental factors (i.e. regions, treatments). The study biotas are obtained by gathering soil sample(s) from randomly selected location(s) within each experimental unit, with an experimental unit being a site within a study region or a field plot of a manipulative experiment. Then, plant growth is measured in glasshouse containers housing soil and biota from the various sites or plots. Results of these glasshouse bioassays are sensitive to a common soil handling decision. In particular, it is common to either: (1) fill each container with soil from one experimental unit (e.g. Callaway et al., 2004; Hood et al., 2004; Kardol et al., 2006; Wardle et al., 2012), or (2) fill each container with a mixture of soils from multiple experimental units (i.e. all sites within a region, all plots that received the same treatment) (e.g. Van der Putten et al., 1993; Nijjer et al., 2007; Felker-Quinn et al., 2011; Pendergast et al., 2013; Rodríguez-Echeverría et al., 2013; Yang et al., 2013; Gundale et al., 2014; Pizano et al., 2014; Hilbig & Allen, 2015; Larios & Suding, 2015) (Fig. 1). We define samples generated from these two approaches as ‘individual soil samples’ (ISS) and ‘mixed soil samples’ (MSS). The term ‘individual soil sample’ is slightly misleading, as ISS are often formed by mixing multiple samples gathered from the same experimental unit (i.e. pooling subsamples). Combining subsamples from individual experimental units is a perfectly acceptable approach. Conversely, the express purpose of this paper is to illustrate that, without exception, the MSS approach of mixing together soils from multiple experimental units (Fig. 1) is fatally flawed. Hypotheses regarding differences among regions or treatments cannot be legitimately tested by the MSS approach of mixing together soils from multiple sites within regions or multiple plots receiving the same treatment. The importance of this point is clearly underappreciated: we estimate, 52% of published studies use MSS in place of the correct ISS methodology (of 76 evaluated studies using ISS or MSS, 40 used MSS) (K. O. Reinhart & M. J. Rinella, unpublished, 2015).

Estimating treatment (e.g. region, nutrient) differences entails computing residual variance. Residual variance describes variation not explained by treatments, and it is needed to compute relevant statistics (i.e. P-values, confidence intervals). In experiments considered here, there are two contributors to residual variance in plant growth: (1) spatial variation in soil biotas (i.e. site-to-site variation not explained by region, plot-to-plot variation not explained by treatment) and (2) glasshouse variation owing to environmental gradients (e.g. temperature) and plant genetics. With the ISS approach (Fig. 1), having two contributors to residual variance poses no unique analytic challenges, and standard regression and analysis of variance (ANOVA) approaches give correct inferences. With the MSS approach, all information regarding residual variation in soil mutualists and pathogens is lost, and if this variation is nonzero, MSS and ISS are guaranteed to give different inferences, with the MSS inferences being incorrect. More specifically, if residual variation in mutualists and/or pathogens is nonzero, statistical estimates from MSS will be falsely precise and evidence for differences among treatments (e.g. regions, nutrients) will be weaker than reported. Assuming only factors being studied cause soil mutualists/pathogens to vary spatially is highly unrealistic, particularly given that plant disease expression (e.g. Martin & Loper, 1999) and soil microbe compositions (e.g. Ettema & Wardle, 2002; Ritz et al., 2004) are known to vary widely across even small spatial scales (i.e. < 1.0 m).

Exploring effects of mixing soils

While inferences from MSS are a priori expected to be falsely precise (i.e. confidence intervals incorrectly narrow, P-values incorrectly small), the magnitude of the inaccuracy is data-dependent. Unfortunately, the magnitude of the inaccuracy arising from MSS cannot typically be quantified, because the necessary information (i.e. soil biota abundance data for each field unit and data describing relationship between soil biota abundance and plant performance) is not collected in most studies. Data from Reinhart & Clay (2009) and Reinhart et al. (2010a) provide a rare opportunity to evaluate the magnitude of inferential inaccuracies arising from MSS approaches. Their data describe site-to-site variation in soil pathogen (Pythium spp.) densities (propagules/g of soil) and the relationship between these densities and biomass of the host tree Prunus serotina in a growth chamber. Pythium densities were measured at 10 eastern US sites separated by a maximum distance of 700 km. While P. serotina was subjected to only one Pythium species in the growth chamber (i.e. P. sylvaticum), we believe P. sylvaticum is a suitable proxy for what may have been multiple Pythium species sampled from the field (Reinhart et al., 2010a). Pythium sylvaticum is a widely distributed pathogen with effects on P. serotina comparable to other Pythium spp. (Reinhart et al., 2010b).
Parameterizing the simulation

This section describes technical details of parameterizing the simulation and may be skipped by readers seeking a general understanding. Pythium spp. were found at only three of 10 sites, and at these sites, the mean and standard deviation of pathogen density was 3.9 and 1.8 log(propagules/g of soil), respectively. Correspondingly, our system for simulating pathogen densities for regions (Step 1) was to generate 10 Pythium presence/absence values with presence probability 0.3. (Our results were similar in an additional simulation that assumed Pythium occurred at all sites.) For sites simulated to have Pythium, Pythium density was simulated as Exp[N(3.9, 1.8)] where N(μ, σ) denotes the normal distribution with, mean μ, standard deviation σ. The growth chamber biomass data from 22 pots of Reinhart et al. (2010a) were well-approximated by:

\[ \log y_i \sim N(-1.8 - 0.2 \times \log(x_i + 1), 0.13), \]

Eqn 1

where \( x_i \) is \( P. \text{sylvaticum} \) density (propagules/g of soil) and \( y_i \) is mean \( P. \text{serotina} \) shoot biomass (g/plant) per pot in pot \( i = 1, \ldots, 22 \), respectively.

Computing the confidence intervals (Step 4) involved fitting the following simple linear regression model to each simulated dataset:

\[ \log z_i \sim N(B_0 + B_1 r_i, \sigma), \]

Eqn 2

where \( z_i \) is simulated \( P. \text{serotina} \) shoot biomass (g/plant) for pot \( i = 1, \ldots, 20 \), and \( r_i \) is an indicator equaling 0.0 for Region 1 and 1.0 for Region 2. It is easy to show that \( B_1 \) is the following log response ratio:

\[ \log \frac{\text{region 1 mean biomass/plant}}{\text{region 2 mean biomass/plant}}. \]

Eqn 3

The analytical summaries of interest in Step 4 were 95% confidence intervals on this log ratio. Negative log response ratios indicate plant biomass was lower with soil originating from Region 1 than Region 2 and vice versa for positive values. Mathematica 9 code (Wolfram Research Inc., Champaign, IL, USA) for conducting the simulation is provided (Supporting Information Notes S1).

Simulation results

It is critical to recall that the simulation was set up such that there was no pathogenicity difference between the two regions. Thus, for any given simulated experiment, a proper 95% confidence interval on the soil origin effect has a 95% chance of overlapping the zero line (Fig. 2). It follows that if proper methods are used, there should be a tendency to detect no difference in pathogenicity between regions, and 95 of our 100 simulated confidence intervals should overlap zero. (Because of sampling variability, the number need not be exactly 95.) A confidence interval that fails to overlap zero represents a false rejection of the null hypothesis of no difference between regions (i.e. type I error). With the ISS approach, 97 of 100
confidence intervals overlapped zero (Fig. 2a), thereby indicating the reliability of this approach. Conversely, the MSS approach performed very poorly, leading to a false rejection of the null hypothesis in 50 of 100 experiments (Fig. 2b). This poor performance occurred because mixing soil samples eliminated site-to-site variation in soil biota, thereby leading to falsely precise estimates. This can be seen by comparing the narrow MSS confidence intervals with the broader ISS confidence intervals (Fig. 2).

Conclusions

Soil handling decisions may dramatically affect conclusions of experiments seeking to quantify effects of soil biota on plants. Our results lead us to question results of replicated field studies that use MSS approaches to assess effects of soil biota on plant growth (e.g., Van der Putten et al., 1993; Nijj et al., 2007; Felker-Quinn et al., 2011; Pendergast et al., 2013; Rodriguez-Echeverría et al., 2013; Yang et al., 2013; Gundale et al., 2014; Pizano et al., 2014; Hilbig & Allen, 2015; Larios & Suding, 2015). In the types of studies we describe in this paper, a one-to-one correspondence should be maintained between field units (i.e. plots, sites) and greenhouse containers. Finally, for the reasons described in this paper, inferences from all studies (e.g. soil chemistry, soil microbial community composition), not just soil biota effects studies, are nearly certain to be invalid if they are derived by performing tests on mixtures of soils from multiple experimental units.

Author contributions

K.O.R. and M.T.R. identified the methodological problem, M.T.R. designed the simulation and performed the analysis, K.O.R. provided data to parameterize the simulation, and M.T.R. and K.O.R. wrote the manuscript.

Kurt O. Reinhart* and Matthew J. Rinella

United States Department of Agriculture – Agricultural Research Service, Fort Keogh Livestock & Range Research Laboratory, 243 Fort Keogh Road, Miles City, MT 59301-4016, USA (*Author for correspondence: tel +1 406 874 8211; email kurt.reinhart@ars.usda.gov)

References


### Supporting Information

Additional supporting information may be found in the online version of this article.

**Notes S1** Mathematica code for simulation.

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

**Key words:** microbial ecology, plant–soil biota experiment, plant–soil feedback, soil handling technique, type I error.