OPTIMIZING THE SCALE OF SPATIAL AGGREGATION IN THE ANALYSIS OF MINIRHIZOTRON DATA

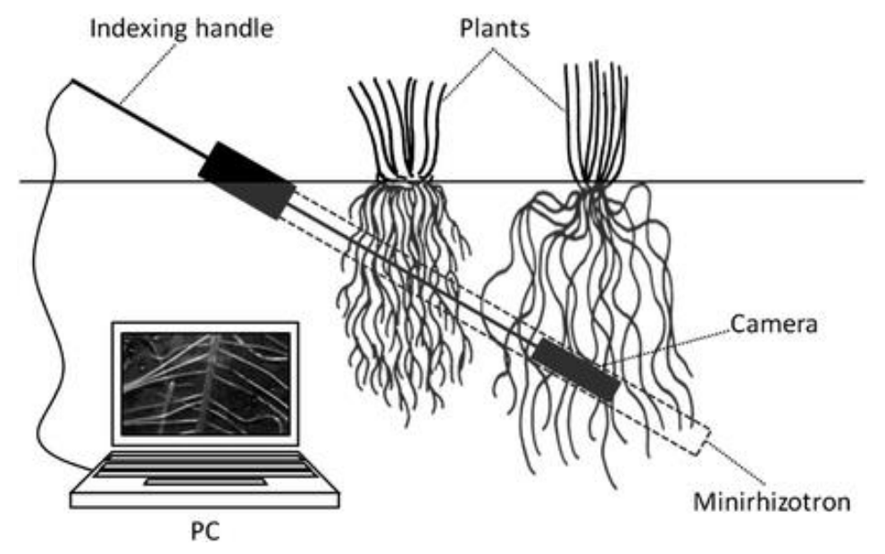
INSTITUTE OF FOOD AND AGRICULTURAL SCIENCE
STATISTICAL CONSULTING UNIT

Simon Riley[1,2], James Colee[1], Edzard van Santen[1,2]
[1]Statistical Consulting Unit, Institute of Food and Agricultural Science, University of Florida, Gainesville, FL, USA
[2]Agronomy Department, Institute of Food and Agricultural Science, University of Florida, Gainesville, FL, USA

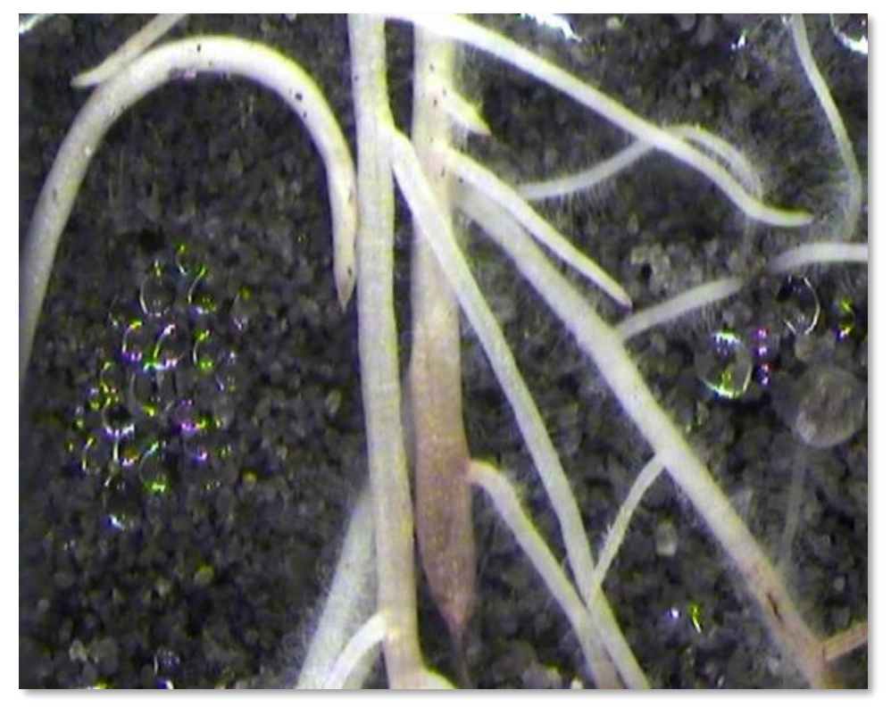
Introduction

Minirhizotron systems consist of clear plastic tubes, which are typically inserted into the soil prior to planting in line with the crop row and at an angle to the soil surface (Fig 1A), along with a boom-mounted camera with which the growing root system is imaged (Fig 1B). Minirhizotron studies thus measure the distribution of crop roots across the soil profile and its evolution through time, under field conditions and with minimal distortion to the root system. This makes them unique among platforms and methods for studying roots.

Figure 1
(A) Minirhizotron System (Image credit: Rewald & Ephrath, 2013)

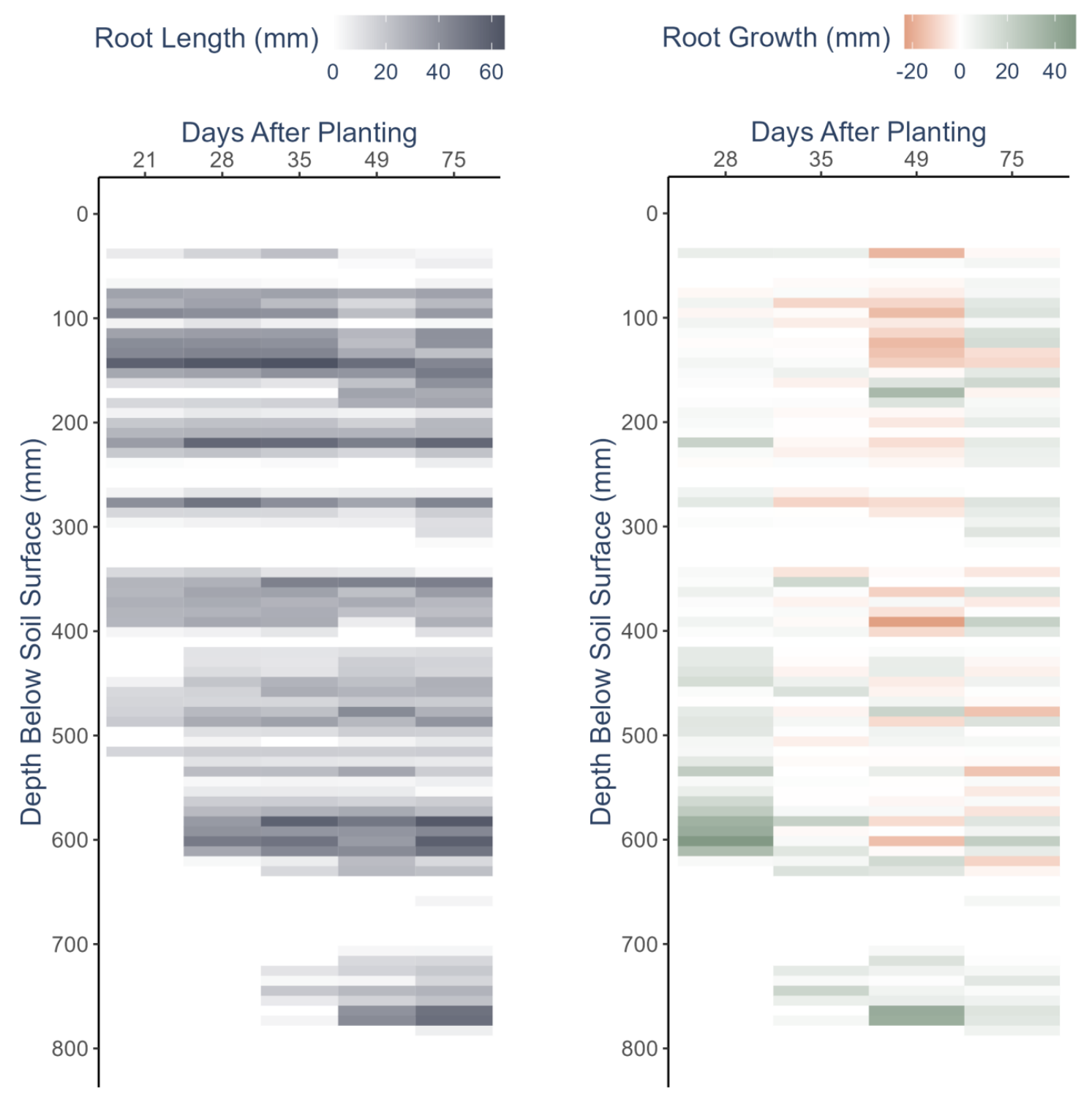


(B) Example minirhizotron photo (Image credit: Roland & Na, 2016)



The resulting images – which can number in the 10's of thousands in even a moderately sized study - are then digitally "traced" to produce spatially- and temporally- referenced measures of root length and diameter, among others (Fig 2A). These data are thus doubly-repeated measures, with very high serial correlations in both dimensions, although differencing can be used to reduce, if not eliminate, temporal autocorrelation (Fig 2B)(Gelman & Hill, 2006). Finally, the data are typically aggregated into depth classes prior to analysis (e.g., Zurweller et al., 2018). This introduces the modifiable areal unit problem (MAUP), in which estimates and inferences are conditional on arbitrary choices regarding the size and location of the areal units (i.e., depth classes)(Jelinski & Wu, 1996). It is also not clear from previous research how the choice of depth class size may interact with spatial- and temporal-autocorrelation to affect type 1 error rates, or how these and the magnitude of the effect at one level may affect statistical power when data are analyzed on a different, larger scale. Given the cost associated with collecting and tracing the images, it is important to ensure the analytical methods are as efficient as possible.

Figure 2



(A) Example measures of root length from a single minirhizotron

(B) Example measures from a minirhizotron after differencing

Objectives

1) To determine whether and to what extent the choice of depth class size affects type I error rates in minirhizotron studies, and to characterize the degree to which this effect is a function of the magnitude of spatial and/or temporal autocorrelation in the measurements.
2) To determine whether and to what extent the choice of depth class affects statistical power in minirhizotron studies, and to characterize the degree to which this effect is a function of both the spatial and/or temporal autocorrelation in the measurements as well as the magnitude of the effect size.
3) To determine, if possible, the optimal scale of spatial aggregation in minirhizotron studies.

Methods

Simulated measures of crop root growth were drawn from a multivariate *t* distribution with one degree of freedom. Specifically, the following formulae were used to simulate the data in R (see Table 1 for explanation and factor levels):

$$Y \sim t_1(\mu, \Sigma)$$
$$\mu = 0 \text{ or } \mu \sim Normal(0, E)$$
$$\Sigma = \rho_s^{|d_s|} \otimes \rho_t^{|d_t|}$$

Where $|d_s|$ is an 80x80 matrix of distances in space and $|d_t|$ is a 5x5 matrix of distances in time. Four replicates were generated for each simulation. The data from each replicate for a given time point was then averaged over one of 6 depth class sizes (Table 1). One thousand simulated data sets were generated for each treatment combination. These simulated data sets were then analyzed in SAS proc mixed, with depth class, time point and their interaction treated as fixed effects. The R side effects were modeled as an anisotropic spatial power correlation structure. Results of the type III tests of fixed effects and covariance parameter estimates were saved and re-exported to R for analysis, with true positives and false negatives both modeled as binomially distributed and estimated via a generalized linear model.

Table 1: Treatment factors and levels employed in the study.

Factor	Symbol	Factor Levels or Dimensions
Temporal Correlation	ρ_t	0, 0.3
Spatial Correlation	ρ_s	0, 0.3, 0.5, 0.7, 0.9
Effect Size Ratio	E	1, 3, 6
Samples per Depth Class		4, 5, 8, 10, 16, 20*

* Four samples per depth class corresponds to 20 depth classes. Twenty samples per depth class corresponds to 4 depth classes.

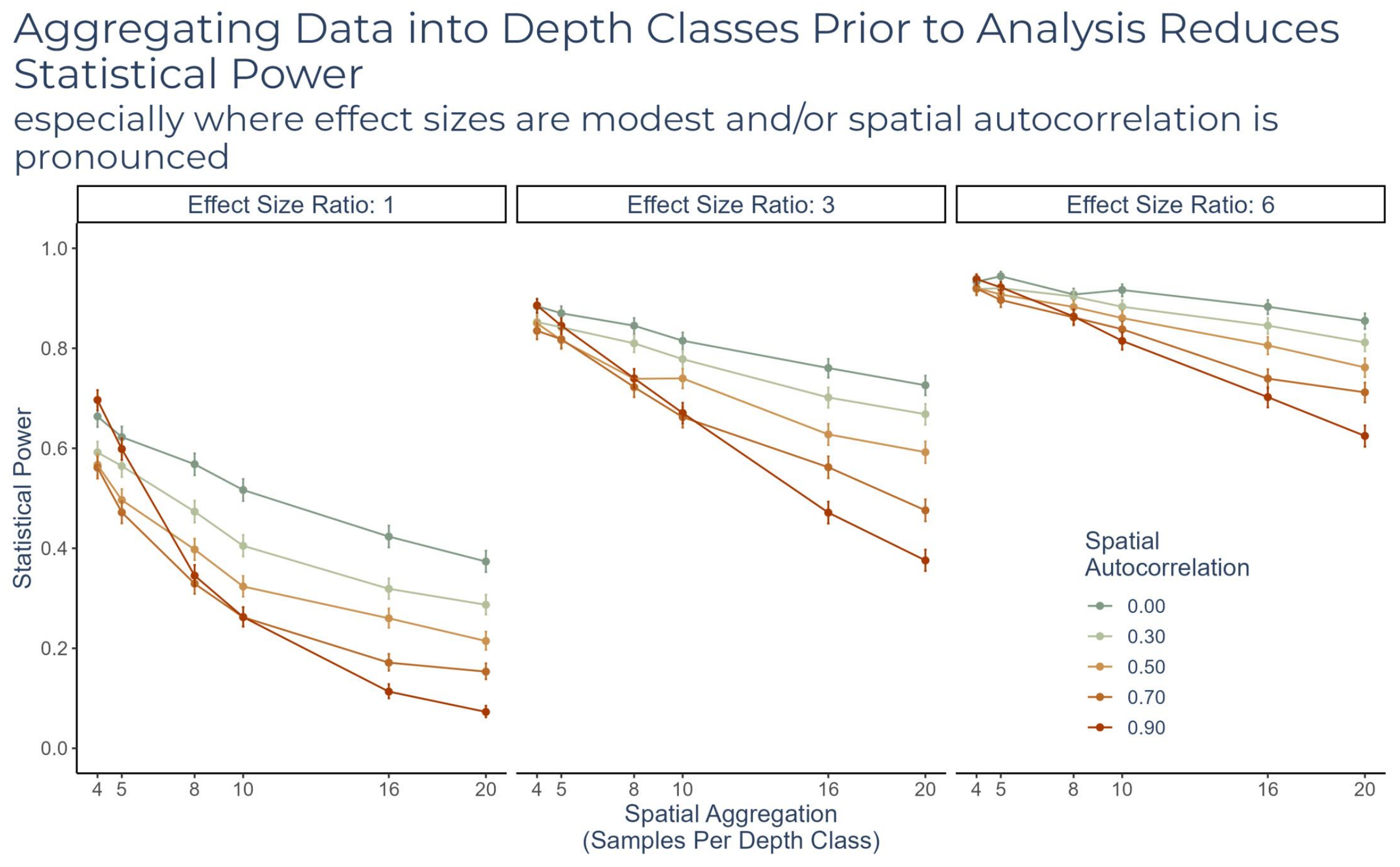
Results

Although the effect of depth class was statistically significant with regard to the depth x time point interaction, it did not have any practical impact on type I error rates, regardless of the degree of spatial or temporal autocorrelation: in all instances, type I error rates were kept below the nominal 0.05 rate (results not shown).

Power, however, was heavily impacted by the magnitude of the effect size, the degree of spatial autocorrelation, and the choice of depth class size (Figure 3). Specifically, there is a pronounced loss of power as the size of the depth classes is increased, and this is exacerbated when spatial autocorrelation is high and the true effect size modest.

These findings suggest that analyzing data on a scale larger than that at which the effects and spatial correlation are manifest can have deleterious impact on statistical power.

Figure 3: Statistical power to detect differences among depth classes as a function of effect size, spatial autocorrelation and spatial aggregation



References

Gelman, A., & Hill, J. (2006). Concepts and methods from basic probability and statistics. In *Data Analysis Using Regression and Multilevel/Hierarchical Models* (pp. 13–27). Cambridge University Press.

Jelinski, D. E., & Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, *11*(3), 129–140. https://doi.org/10.1007/BF02447512

Rewald, B., & Ephrath, J. E. (2013). Minirhizotron Techniques. In A. Eshel & T. Beeckman (Eds.), *Plant Roots: The Hidden Half* (Fourth). Chapman and Hall/CRC. http://ebookcentral.proquest.com/lib/ufl/detail.action?docID=1172599

Rowland, D., & Na, C. (2016, January 8). UF/IFAS Researchers Studying Corn Root Development in Florida's Sandy Soils. *UF/IFAS Extension - Panhandle Agriculture*. https://nwdistrict.ifas.ufl.edu/phag/2016/01/08/ufifas-researchers-studying-corn-root-development-in-floridas-sandy-soils/

Zurweller, B. A., Rowland, D. L., Tillman, B. L., Payton, P., Migliaccio, K., Wright, D., & Erickson, J. (2018). Assessing above- and below-ground traits of disparate peanut genotypes for determining adaptability to soil hydrologic conditions. *Field Crops Research*, *219*, 98–105. https://doi.org/10.1016/j.fcr.2018.01.020