Farella Martha, Mary (Orcid ID: 0000-0002-2925-4093)

Breshears David, D. (Orcid ID: 0000-0001-6601-0058)

Gallery Rachel, E (Orcid ID: 0000-0002-9159-8778)

Predicting drivers of collective soil function with woody plant encroachment in complex landscapes

Martha M. Farella[1], David D. Breshears[1, 2], and Rachel E. Gallery[1, 2]

[1]School of Natural Resources and the Environment, Environment and Natural Resources 2, The University of Arizona, 1064 East Lowell Street, Tucson, AZ 85721, USA.

[2]Department of Ecology and Evolutionary Biology, The University of Arizona, 1041 East Lowell Street, P.O. Box 210088, Tucson, AZ 85721, USA.

Corresponding author: Martha Farella (marthag@email.arizona.edu)

Key Points:

- A soil functioning index, which included microbial activity, biomass, soil carbon and nutrients, was evaluated with mesquite encroachment.
- Vegetation influences soil nutrients, microbial biomass, and activity and predicted almost half of the variation in soil functioning.
- Adding topography, soil properties, parent material, and precipitation to vegetation models further improved soil functioning predictions.

Key Words:
Mesquite, biogeochemistry, microbial exoenzyme activity, function, random forest

Index Terms:
Primary- Biogeochemical cycles, processes, and modeling (0412, 0793, 1615, 4805, 4912)
1- Ecosystems, structure and dynamics (4815)
2- Microbiology: ecology, physiology and genomics (4840)
3- Land cover change (1632)
4- Nutrients and nutrient cycling (4845, 4850)

Abstract

Dryland (arid and semiarid) ecosystems are extensive, home to a third of the human population, and a major contributor to terrestrial Net Primary Productivity and associated biogeochemical cycles. Many dryland systems are undergoing woody plant encroachment, which can substantially alter landscape-scale soil nutrient dynamics via long-recognized "islands of fertility" mechanisms. To effectively constrain soil biogeochemistry responses to woody plant encroachment, predictions are needed for microbial biomass and especially microbial activity in addition to existing predictions for soil nutrients—referred to collectively hereafter as "collective soil functioning". Here we evaluated whether collective soil functioning could be predicted from a suite of metrics including plant cover, precipitation, soil physiochemical characteristics, and topographic variables across complex landscapes undergoing woody plant encroachment by mesquite (*Prosopis velutina*). Plant cover alone predicted nearly half of the variability (R^2 = 48.5%) in collective soil functioning and had a significant effect on each component of this index (soil nutrients, microbial biomass and microbial activity). Prediction strength for collective soil functioning increased to 55.4% and the error term decreased by 13.4% when precipitation, soil physiochemical characteristics and topographic metrics were also included in models (plant and environment model). Besides the expected effects of plant cover, other significant predictors of collective soil functioning included state factors such as topography, precipitation, and parent material along with soil age and bulk density. These results illustrate that mesquites influence many components of soil functionality but the strength of this effect depends on which component is analyzed and which environmental variables are considered.

Plain Language Summary

Arid systems contain a significant amount of plant species and are home to one-third of the human population. Woody plant species are invading these historically grass-dominated areas, causing dramatic changes to soil health that are not fully understood. When mesquite trees invade, for example, soil nutrients become concentrated in the areas surrounding them, which changes the distribution of water and nutrients across the landscape. We evaluated the factors that were important in predicting soil health in an area undergoing this conversion. The soil health index we created included carbon, nutrients, the microscopic organisms that live within the soil and the ecosystem services they provide. Plants influenced soil health the most, followed by soil age, the amount of soil pore space, and soil pH. Our results show that the biomass and activity of soil microorganisms changes with mesquite invasion. This is important because if we can recognize changes in soil microorganisms along with other widely used soil variables, we could predict future changes in soil resources that may occur. This could allow us to develop early management plans that would mitigate the impacts of invasive plants in these sensitive and important regions.

1 Introduction

Dryland (arid and semiarid) ecosystems cover 41% of terrestrial land area and are home to more than a third of the human population (MEA, 2005). They account for 30-35 % of terrestrial Net Primary Productivity (NPP) and play large roles in global carbon (C), water, and nitrogen (N) cycles (Campbell & Stafford Smith, 2000; Field, 1998). Given their coverage and importance, the numerous threats these systems face present crucial research, management, and policy challenges (J. F. Reynolds et al., 2007). Over the past 150 years, for example, many xeric systems have undergone a dramatic shift in vegetation from grasslands and savannahs to shrublands due to woody plant encroachment (Archer, 1995; Van Auken,

2000). In the western U.S. alone, over 330 million hectares have undergone this conversion (Mitchel P. McClaran, 2003; Pacala et al., 2001).

Woody plant encroachment affects key ecosystem services such as shifts in plant diversity (Ratajczak et al., 2012), carbon storage (González-Roglich et al., 2014), ecohydrology (Huxman et al., 2005), and the distribution of soil resources (Eldridge et al., 2011). Soils are responsible for providing many essential ecosystem services including: water storage and filtration, nutrient cycling, litter decomposition, carbon sequestration, and climate regulation (Bardgett & Van Der Putten, 2014; Fierer et al., 2013; Zhou et al., 2011). These processes often require the production of microbial exoenzymes, hereafter referred to as "microbial activity". Both biotic (e.g., plant communities) and abiotic (e.g., climate and soil texture) properties can influence soil functioning. For example, plants can directly influence their associated microbial communities through root exudates and the quantity and nutrient stoichiometry of plant litter inputs to soil (Cotrufo et al., 2013; Sinsabaugh et al., 1991). Across larger scales, land use and factors such as climate and geomorphic properties can alter soil microbial community diversity and activity (Cao et al., 2016; Drenovsky et al., 2010; Evans & Wallenstein, 2012). Woody plant encroachment has particularly important implications for soil biogeochemistry. For example, woody plant encroachment by velvet mesquites (*Prosopis velutina*) create "islands of fertility" where they alter nutrient dynamics and the soil resource pools in these islands become elevated compared to non-woody plant patches (Archer et al., 2000; Charley & West, 1975; W H Schlesinger et al., 1990). Shrub size, which can correlate with shrub age, contributes to these patterns with larger mesquites having more pronounced impacts on soil nutrients and microbial biomass than smaller ones (Cable et al., 2009; Throop & Archer, 2008). Although the local effects of mesquites on soil nutrients have been repeatedly quantified and are well recognized, soil biogeochemistry is constrained and regulated in many key ways by soil microbial biomass and activity. The

effect of mesquite encroachment on these more dynamic biological processes is less well known.

Despite the recognized importance of soil microbial communities, there are large uncertainties regarding how spatial variation in microbial communities affects ecosystem processes at larger scales (Berg, 2012). The distribution of soil microorganisms and the relative importance of drivers that influence their dynamics are not fully resolved and, as a result, soil microbial communities are often excluded or oversimplified in ecosystem models. Improvements to models of soil biogeochemical dynamics require the inclusion of microbial biomass and activity in addition to soil nutrients—hereafter referred to as "collective soil functioning". More specifically, understanding the responses of collective soil functioning that accompany woody plant encroachment can inform how changes in vegetation affect landscape scale nutrient dynamics, help identify important factors that facilitate or result from this shift in vegetation, and better inform management efforts aimed at mitigating these changes (Browning et al., 2014).

To evaluate trends in collective soil functioning with woody plant encroachment we developed a "collective soil functioning index". Soil indices are commonly used as measures of soil quality or health and can help to determine the effects of disturbance on soil functioning. Generally, soil indices are calculated from many simultaneously measured soil properties that provide information on the overall health or functioning of the soil ecosystem (Karlen et al., 1997). Soil indices allow for the combination of multiple related variables into a single measure that is straightforward and interpretable (Delgado-Baquerizo et al., 2016; Maestre et al., 2012). Traditionally, soil indices included information pertaining to the physical and chemical properties of a soil (Bone et al., 2014). More recently, soil biological properties, which often respond more rapidly to disturbance and control many of the functions that occur in soils including decomposition, nutrient cycling, and N-fixation, have

been included into soil index calculations (Bradford et al., 2014; Delgado-Baquerizo et al., 2016; Paz-Ferreiro & Fu, 2016). The collective soil functioning index we developed includes measures of soil carbon, nutrients, and biological properties such as microbial biomass and activity.

Here we evaluated the impact of mesquite encroachment, soil physiochemical variation, water dynamics, and topographic features on collective soil functioning. The Santa Rita Experimental Range (SRER) in Arizona, USA is the study site for this experiment. In addition to woody plant encroachment by mesquite, SRER also hosts a range of soil physiochemical variation, climatic gradients, and topographic factors making it an ideal location for the present study. Building upon past research cited above, we predict that while state factors such as parent material and soil age will broadly explain changes in collective soil functionality over climatic gradients, the biotic influence of mesquite cover will overwhelm the abiotic landscape influences on this index. Furthermore, within this index, the microbial activity parameters will respond to mesquite encroachment, contributing to the positive feedbacks with microbial biomass and nutrient availability that are predicted to increase with mesquite encroachment. Disentangling the relative importance of these factors on soil functionality can help us improve model predictions and management policies in an effort to ensure sustainable land management in woody plant encroachment systems worldwide.

2 Materials and Methods

2.1 Site Description

Samples were collected at SRER on August 23 – 29, 2017. SRER is a 52,000 acre long-term semiarid rangeland research facility located in southeastern Arizona (Fig. 1). This historically grass-dominated ecosystem has been undergoing woody plant encroachment over

the past century, transforming the area into a mesquite-dominated savannah (Browning & Archer, 2011).

We sampled sites across SRER capturing gradients in plant communities, precipitation regimes, topographic factors, and soil physiochemical properties (Fig. 1). Soil age, landform type, and parent material delineations for each site across the study area were defined based on geomorphic surface definitions outlined in Batchily et al. (2003). We chose to use geomorphic surface delineations across SRER to guide our sampling strategy because they allowed us to capture a high degree of variation in the aforementioned variables. We sampled a total of 15 sites, which consisted of two replicates for each of the seven geomorphic surfaces (A-G) with the exception of geomorphic surface E, which ended up having three replicates.

Annual precipitation at SRER is bimodal with peaks during the monsoon (late July - August) and the winter rainy season (December - January) (Browning et al., 2008). Mean annual precipitation varies across SRER ranging from 275 mm at the lower elevations (≈ 900 m; sites in geomorphic surfaces B, D, & E) up to 575 mm at the higher elevations (≈ 1,400 m; sites in geomorphic surfaces A & C; Supp. Table 1). There are a variety of soil types across SRER from Pleistocene and Holocene origins derived from both igneous and sedimentary rocks (Batchily et al., 2003).

Soil texture consisted mostly of sand, loamy sand, and sandy loam. Clay % was highest at the alluvial fan and basin floor sites of middle to early Pleistocene soil ages (sites located in geomorphic surfaces B, E, & G; Supp. Table 1) and lowest at the alluvial fan site of Holocene to Late Pleistocene soil ages (sites F1 and F2). Sand % was highest in igneous rock derived soils of site F and lowest at alluvium soils of sites located in geomorphic surfaces B and C and limestone sedimentary rock derived soil of sites E1-E3. Average soil pH ranged from 6.06 – 8.83 across sites with soils derived from limestone sedimentary rock

having more alkaline pH values (sites E1-E3; Supp. Table 1) and soils derived from igneous rock having more acidic pH values (sites in geomorphic surfaces A, F, & G; Supp. Table 1). High elevation sites in the floodplain, on hills and mountains, and in the alluvial fan supported the highest aboveground biomass (sites in geomorphic surfaces A, C, F, & G; Supp. Table 1) while lower elevation sites in the alluvial fan had the lowest aboveground biomass (sites in geomorphic surfaces B, D, & E;).

2.2 Field Sampling and Mesquite Biomass Estimates

At each site, three 5 m x 5 m plots were established for a total of 45 plots. Each of the three plots had a dominant plant cover of either mesquite, grass, or no plants (henceforth referred to as the "bare" plots). Soil samples were collected to 10 cm depth. In mesquite plots, soil samples were collected from opposing sides of the bole, dripline, and a mid-point halfway between the bole and dripline. In the grass plots, soil samples were collected from opposing sides of the grass rhizosphere. In the bare plots, two samples were taken from random locations within the plot with no plant growth. Duplicate samples were homogenized to represent a single composite soil sample for each plot. Samples were then stored on ice and transported to the lab for analysis. Litter depth was measured at the location soil was sampled. Tree height was measured in the field for the trees in the mesquite plots. Mesquite biomass was calculated with the allometric equation (Eqn. 1) presented in McClaran et al. (2013) using the measured tree height values.

$$\text{Eqn. 1: } \ln(Y) = -3.20 + 3.00(\ln(X)) * 1.09$$

Where X = tree height (m) and Y = foliar biomass (kg).

2.3 Laboratory Analysis

Soil samples were sieved using 2-mm mesh and then analyzed for physical, chemical, and biological properties. Percent organic matter (OM %) of samples was determined by the

loss on ignition method of heating samples at 500 °C for 4 hrs (Nelson & Sommers, 1965). Soil pH was measured in deionized water with a 1:2 soil-to-solution ratio using sympHony VWR pH electrode probe (Nicol et al., 2008). Soil texture analysis (sand, silt, and clay %) was determined according to the hydrometer method following dispersal with sodium hexmetaphosphate (Ashworth et al., 2001; ASTM, 2007; Bouyoucos, 1962; Day, 1965). Bulk density was determined with Soil Water Characteristics Hydraulic Properties Calculator software from USDA-ARS (version 6.02.74) using sand, silt, clay, and OM % values. This Pedotransfer Function (PFT), which is widely used in the literature, is based on > 6,000 samples from the USDA/NRCS National Soil Characterization database (Saxton & Rawls, 2006). Furthermore, when recently tested on soils from arid and semiarid regions it produced high accuracy and low biased results, supporting the use of it in the present analysis (Sevastas et al., 2018).

Microbial biomass C and N (MBC and MBN respectively) in the soil was measured using the chloroform fumigation extraction method (Beck et al., 1997). Paired samples that were either fumigated with ethanol-free chloroform or non-fumigated were extracted with 25 mL 0.5 M K_2SO_4. Samples were shaken for 1 hour, filtered, and stored at -20 °C until processing using a non-purgeable-organic-C protocol on a Shimadzu total organic carbon analyzer (TOC 5000) equipped with a total dissolved nitrogen module (Shimadzu Scientific Instruments, Inc., Columbia, MD, U.S.A.). Efficiency factors for MBC (kEC = 0.45; Beck et al., 1997) and MBN (kEN = 0.54; Brookes et al., 1985) were used to calculate microbial biomass C and N concentrations as the difference between fumigated and non-fumigated samples. Values measured in the non-fumigated samples represented soil dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) amounts.

Soil ammonium (NH_4^+), nitrate (NO_3^-), and rates of net N-mineralization were determined using K_2SO_4 extraction (Robertson et al., 1999). Air-dried soil (5.0 g) was

brought up to 60 % water holding capacity and incubated for a week. K_2SO_4 extractions were performed on samples before and after the incubation with 25 mL of 0.5 M K_2SO_4. Samples were shaken for 1 hour, filtered, and stored at -20 °C until processing. All K_2SO_4 extracts were analyzed colorimetrically with 2-phenylphenol for NH_4^+-N (Rhine et al., 1998) and the vanadium method of Doane & Horwáth (2003) for NO_3^--N using Synergy™ 4 Multi-Mode microplate reader. Net N-mineralization was calculated as the difference between the sum of NH_4^+-N and NO_3^--N before and after the incubation. Pre-incubation NH_4^+ and NO_3^- values were used in analysis.

Potential extracellular enzyme activity (EEA) was measured using a fluorometric deep-well microplate technique following Gebhardt et al. (2017) modified from Wallenstein et al. (2012). Prior to the assays, soil pH was used to determine the appropriate buffer solution. Soil slurries were prepared with 2.75 g of soil that was stored at 4 ˚C and 91 mL of either tris or sodium acetate buffer, which was titrated to a pH that closely resembled the soils analyzed. We measured potential activity of seven hydrolytic exoenzymes. Four of the exoenzymes analyzed hydrolyze carbon-rich substrates: α-Glucosidase (AG), β-Glucosidase (BG), β-D-cellubiosidase (CB), and β-Xylosidase (XYL). Two of the exoenzymes analyzed hydrolyze nitrogen rich substrates: N-acetyl-β-Glucosaminidase (NAG) and leucine aminopeptidase (LAP). Phosphatase (PHOS), the final exoenzyme assayed, hydrolyzes phosphorus rich substrates. For each exoenzyme assayed, 100 µL of 200 µM fluorometric substrate was added to 900 µL of soil slurry. Standards for standard curves and assays were incubated at 25 °C for 1 hour or 40 minutes respectively. Fluorescence was measured on Synergy™ 4 Multi-Mode microplate reader with an excitation wavelength of 365 nm and an emission wavelength of 450 nm. Incubation time or standard curve dilutions were adjusted for samples that had activity higher than the detection limit.

DNA was extracted from soils using the DNeasy PowerSoil HTP 96 Kit (Qiagen, Hilden, Germany) and concentration was measured on Synergy™ 4 Multi-Mode microplate reader. DNA concentration has been used as a proxy of soil biomass in other studies and it had significant ($p < 0.001$) correlations (R^2 values ranged from 0.400 – 0.470) with all other metrics of microbial biomass we measured (MBC, MBN, and OM %), supporting its use in the present analysis (S. L. Johnson et al., 2012; Kuske et al., 2002).

2.4 Remote Sensing Data Products

Precipitation was interpolated in ArcGIS (version 2.3.2) from monthly rain gauge measurements on gauges dispersed across SRER. Monthly rain gauge measurements were provided by the Santa Rita Experimental Range Digital Database. Funding for the digitization of these data was provided by USDA Forest Service Rocky Mountain Research Station and the University of Arizona (Mitchel P. McClaran et al., 2002). Three precipitation values were used in analysis: one value that represented the total precipitation that occurred 12 months prior to sampling (1yr precip), one value that represented the total precipitation that occurred 8 months prior to sampling including both winter and summer monsoon amounts (8mo precip), and a final value that represented the total precipitation that occurred 3 months prior to sampling including just the summer monsoon amount (3mo precip).

Topographic factors such as elevation, slope, and aspect were calculated as the average value for each plot from The National Ecological Observatory Network's (NEON) Airborne Observation Platform (AOP) L3 LiDAR data products (Supp. Table 2). NEON is a program sponsored by the National Science Foundation and operated under cooperative agreement by Battelle Memorial Institute. This material is based in part upon work supported by the National Science Foundation through the NEON Program. Aboveground biomass was calculated as the sum of aboveground biomass for all pixels that fell within the plot boundary from NEON's L2 spectrometer data product (Supp. Table 2).

2.5 Indices

To calculate the collective soil functioning index, we first log normalized (and added a constant value to avoid negative values, when needed) and standardized each soil variable measured using the Z-score transformation. Standardized variables were then averaged to obtain the collective soil functionality index. The 16 variables used to calculate this index included: OM %, DOC, TDN, NH_4^+, NO_3^-, all seven potential exoenzyme activities (AG, BG, CB, XYL, NAG, LAP, and PHOS), net N-mineralization, MBC, MBN, and DNA concentration.

We recognize that decreases in one variable can be compensated for by increases in another using the averaging approach applied here (Gamfeldt et al., 2008). Correlation coefficients between most of the soil variables were positive, except for N- mineralization and the nitrogen degrading exoenzymes, where the strongest negative correlation coefficient was only -0.21. Furthermore, none of the negative correlations coefficients were significant (Supp. Table 6). Each individual variable used in the collective soil functioning index calculation was also analyzed independently (Figs. 4 & 5), which allowed us to evaluate similarities and differences that emerged in predictor variable importance between these two approaches.

2.6 Data Analysis

All statistical analysis was performed in R version 3.4.4 (R Core Team, 2018). In most cases, data transformations did not result in normally distributed data. As such, data were left untransformed, unless specified otherwise. Analysis of variance (ANOVA) was performed to determine differences among sites and plant types in measured variables. Although ANOVA generally tolerates violations to the assumption of normality, nonparametric Kruskal-Wallis tests were also ran on the data (Blanca et al., 2017; Glass et

al., 1972; Harwell et al., 1992; Lix et al., 1996). In all cases, nonparametric analysis exhibited the same trends as parametric analysis, so parametric results are presented here since they provide more efficient inferences than nonparametric procedures. Tukey's post-hoc comparison tests were run on significant ($p < 0.05$) differences. Regression analysis was used to analyze relationships between mesquite size, soil physiochemical properties, water dynamics, topographic factors, plant variables, and the variables included in the collective soil functioning index.

Random forest modeling was used to identify the most important predictors of collective soil functionality while including both continuous and categorical variables simultaneously in the analysis. Each of our random forest analyses built 5000 trees and the square root of the total number of predictor variables was used to determine the number of variables randomly sampled as candidates at each split. This analysis was done with the randomForest R package. Each tree in the random forest analysis was built on two-thirds of the data and tested on the remaining third. The importance of each predictor variable was determined by the average decrease in prediction accuracy across all trees, or the average increase in mean square error (mse) between the data the tree was built on and the remaining third of the data the tree was tested on. Significance of each predictor variable was determined with the rfPermute R package.

To determine whether the inclusion of additional explanatory variables improved model performance two model iterations were run. The "plant only model" included the three plant variables measured: plant cover type, aboveground biomass, and litter depth. The "plant and environment model" included plant variables, soil physiochemical properties, water dynamics, and topographic variables. Assessment on the inclusion and exclusion of predictor variables on model performance was determined by repeating the random forest analysis over

1,000 permutations. Model accuracy was assessed with R^2 and mse values from the permutations.

Differences among mesquite-associated soils were not always significant but, when they were, the bole samples were always highest. As such, when we wanted to analyze mesquite-specific impacts on soil dynamics, bole samples were used in analysis (e.g. Fig. 2). In the random forest analysis, all three mesquite-associated soil samples were included in the analysis with the "plant cover" explanatory variable accounting for this difference.

3 Results

3.1 Microbial activity patterns echo those of microbial biomass and soil nutrients as a function of plant cover

Across all sites, plant cover type influenced belowground dynamics. Mesquite-associated soils generally had the highest soil nutrient, potential exoenzyme activities, and microbial biomass values while bare soils had the lowest (Fig. 2). Soils associated with mesquite trees had significantly higher ($p < 0.01$) OM %, DOC, TDN, NH_4^+, and NO_3^- than soils associated with grasses or bare ground patches (Fig. 2a-e). Variability in soil nutrient values was also largest in soils associated with mesquite trees. Plant cover type also had a significant impact on exoenzyme activities. Individually, carbon-acquiring AG and BG exoenzyme activities were significantly higher in mesquite soils compared to bare soils but grass and mesquite AG and BG activities were not significantly different from each other (Fig. 2f & g). Potential XYL exoenzyme activity was the highest in the grass soils and there was no significant difference between mesquite and grass or bare soil XYL activity (Fig. 2i). Potential CB, NAG, and PHOS potential exoenzyme activity were highest in mesquite soils but potential activity differences among plant cover type were not significant (Fig. 2h, j, & l). Potential LAP exoenzyme activity displayed a similar trend to XYL with activity in the grass

soils being the highest but, again, this was not a significant difference (Fig. 2 k). Plant cover type also had a significant impact on microbial biomass ($p < 0.05$). Microbial biomass carbon, nitrogen, and DNA concentration were significantly higher in mesquite soils compared to bare patch soils but could not be differentiated among soils associated with mesquites or grasses (Fig. 2m-o).

Mesquite biomass accounted for some of the variability observed in soil nutrients and microbial biomass values (significant R^2 values range from 0.171 – 0.300; Supp. Table 4). Compared to all of the other soil variables analyzed, MBN had the largest amount of variance explained by mesquite biomass ($R^2 = 0.300$; Supp. Table 4). The NEON derived aboveground biomass parameter also had a significant positive correlation with mesquite biomass, which provides support that our estimates of mesquite biomass, derived from allometric equations, accurately reflected actual tree biomass.

3.2 Collective soil functioning and individual metrics for microbial activity, microbial biomass and soil nutrients are predicted by plant cover and improved with environmental variables

Random forest modeling was used to help explain the variation observed in collective soil functioning (Fig. 3). When plant variables were the only explanatory variables included in the models (plant only model), only 48.5 % of the variation in collective soil functioning was explained and all three plant variables were significant predictors (Fig. 3a). When models included plant variables, soil physiochemical properties, water dynamics, and topographic factors (plant and environment model), 55.4 % of the variability in collective soil functioning was explained and the error term decreased from 0.217 to 0.188 (Fig. 3b). Plant variables, soil age, bulk density, and topographic factors were the most important significant predictors of the collective soil functioning index (Fig. 3b). The inclusion and exclusion of

explanatory variables had a significant effect on model performance. Models with only plant variables (plant only model) had significantly lower R^2 values than the plant and environment model (6.84 % reduction in model performance; $p = 2.2 \times 10^{-16}$).

Not surprisingly, individual soil nutrients followed similar trends to the collective soil functioning index (Fig. 4). Bulk density, soil age, topographic factors, and plant variables were significant predictors of OM %, DOC, TDN, and NO_3^- (Fig. 4). Aspect and soil pH were the most important predictor variables of NH_4^+ (Fig. 4d).

Notability, predictive models for seven out of ten individual metrics of microbial activity and biomass had R^2 values exceeding 0.400 (Fig. 5). Considering all individual soil functioning variables, microbial exoenzymes had the highest R^2 values in the predictive models, exceeding 0.600 for NAG and PHOS activities (Fig. 5 e & g). Soil pH was a dominant predictor variable for all exoenzyme activities except CB and XYL (Fig. 5 a-g). All microbial biomass metrics (MBC, MBN, and DNA concentration) had R^2 values greater than 0.400 (Fig. 5 h-j). Plant variables were significant predictors of microbial biomass (Fig. 5 h-j).

Regression analysis showed that bulk density, precipitation, topographic factors, and plant variables each had significant relationships with all individual variables included in the collective soil functioning index (Supp. Table 5). Bulk density had a significant negative relationship with all soil nutrient and microbial biomass variables and many of the soil activity measures. Precipitation regimes, especially the 1-year and 8-month intervals, had significant positive relationships with all soil nutrients and many of the soil exoenzyme activities. Slope and elevation both had significant positive relationships with most of the soil variables included in the collective soil functioning index. Aboveground biomass and litter depth also had significant positive relationships with most soil variables included in the collective soil functioning index (Supp. Table 5).

There were also significant relationships between many of the soil variables included in the collective soil functioning index. All of the soil nutrient variables had significant positive relationships with other soil nutrients and microbial biomass variables, with correlation coefficients ranging from 0.39 to 0.86 (Supp. Table 6). There were also significant positive correlations between soil exoenzyme activities and soil nutrients, especially OM % and NO_3^-. Microbial biomass measures had significant positive correlations with both other microbial biomass measures and all exoenzyme activities except NAG. All exoenzyme activities except LAP, had strong positive correlations with other exoenzyme activities. N-mineralization did not have significant correlations with any of the other measured soil variables (Supp. Table 6).

4 Discussion

4.1 In drylands, plant cover alone has a large impact on soil biogeochemistry

Similar to other studies, our results support the "fertility island" hypothesis that has been proposed to explain mesquite impacts on soil biogeochemistry (Fig. 2; Charley & West, 1975; Reyes-Reyes et al., 2002; Schade & Hobbie, 2005; W H Schlesinger et al., 1990; Wheeler et al., 2007) and extends it to include microbial activities and biomass. Concentrations of soil resources under vegetation patches can be a result of both biotic and abiotic factors, which is why it is important to consider both in terms of soil function. Symbiotic N-fixation and root exudates can cause soil carbon and nutrient accumulation under vegetation and depletion in non-vegetated areas (Aguiar & Sala, 1999; Perakis et al., 2017; William H Schlesinger & Pilmanis, 1998; Wang et al., 2007). The bare interspace regions between shrubs can experience greater degrees of erosion and nutrient runoff when rain or wind events occur (W H Schlesinger et al., 1990), which can redistribute soil, litter, and nutrients to vegetated areas at the expense of neighboring bare areas (Alvarez et al.,

2012; Puigdefabregas et al., 1999). Vegetated areas with more litter can have higher rates of water infiltration and less variation in shallow soil water content, which can enhance nutrient cycling in water-limited systems (Ludwig et al., 2005; Potts et al., 2010). Even though microbial activities and biomass are more dynamic and change faster than nutrient pools, a snapshot sample showed similar, albeit weaker, relationships with plant cover (Fig. 2). This emphasizes the importance of nitrogen fixing woody plant cover as a dominant driver of soil carbon and nutrient cycling in dryland systems.

There is no consensus on the sequence of changes that occur belowground with woody plant encroachment. Maestre et al. (2011) suggest that the increase in soil carbon and nutrients that occurs with encroachment enhances microbial exoenzyme activity. Similarly, microbial biomass may respond most rapidly to encroachment, followed by an increase in soil carbon stocks, and then an increase in some measures of microbial activity such as substrate use efficiency (Cable et al., 2009). It is likely that the functional changes in microbial activity that are associated with woody plant encroachment are also explained by distinct changes in microbial community composition (Hollister et al., 2010; Li et al., 2017; Yannarell et al., 2014).

We suggest that as mesquites establish, they initially have direct impacts on microbial activities. This is especially true for nitrogen-specific activities from symbiotic diazotroph bacteria in mesquite root nodules that fix atmospheric nitrogen (Allen & Allen, 1981; H. B. Johnson & Mayeux, 1990). Over time, this symbiosis will eventually be reflected in changes to the quantity and nutrient stoichiometry of plant inputs to soil communities via litter and root exudates (Liao & Boutton, 2008). The observed shift in soil microbial activities and increased concentrations of DOC, TDN, NH_4^+, and MBN (Supp. Table 4) with woody plant encroachment could represent enhanced microbial biomass and soil carbon and nutrient

cycling. All of these trends are then self-reinforcing: more activity results in greater nutrient availability for microbial and plant growth, which further supports greater rates of activity.

4.2 Environmental variables add modest enhancement to the predictive ability of plant cover for the collective soil functionality index

While plant cover is of paramount importance, the addition of soil physiochemical properties, precipitation, and topography improved the prediction power of the models and identified other factors such as topography, soil age, and bulk density that contributed to the variability observed in collective soil functioning (Fig. 3-5). Consistent with other studies (Chen et al., 2019; Florinsky et al., 2004; Liu et al., 2007), topographic variables such as slope, elevation, and aspect were consistently important drivers of soil functioning and enhanced the predictive capability of models. Although their impact on soil processes is well recognized, soil physiochemical variables such as texture and pH did not emerge as significant predictors of collective soil functioning when measures of soil carbon, nutrients, and dynamics microbial processes were analyzed together in the collective soil functioning index. Binning multiple soil variables into a single value can obscure relationships between explanatory and response variables, however, which is why it is important to have a certain degree of *a priori* understanding about these relationships before analysis (Bradford et al., 2014).

Soil age, one of the most significant abiotic explanatory variables for the collective soil functionality index, can provide information on soil texture, horizon development, and soil carbon and nutrient pools. For example, older soils at SRER have higher clay content and more horizon development than younger soils with higher sand content (Batchily et al., 2003; Wheeler et al., 2007). Soil age can also inform soil nutrient status; typically, soils shift from N- to P-limitation as they age (Selmants & Hart, 2010; Vitousek & Farrington, 1997). Soil

nutrient status can influence which plant traits dominate, such as mycorrhizae associates and cluster root development to mine N and P in young to old soils respectively (Lambers et al., 2008). Associations between plant traits and soil age at SRER have been observed; older clay-rich soils have greater perennial grass productivity and mesquite biomass than young sandy soils (Browning et al., 2008; Subirge, 1983). Given the strong influence of soil age on soil biogeochemistry, it is no surprise that it consistently emerged as a significant predictor of soil functionality in this analysis.

Bulk density provides information about the level of compaction and porosity in soil; soils with lower bulk densities have less compaction, higher porosity, and greater water holding capacities (Gupta & Larson, 1979). Since water availability is a dominant control over biological activity and nutrient cycling in drylands, it is also no surprise that bulk density consistently emerged as an important predictor variable of soil functioning (Figs. 3-5; Huxman et al., 2004; James F Reynolds et al., 2004). Soil bulk density regulates the amount of nutrient accumulation that occurs during dry periods, with higher bulk density soils having lower soil water availability and nutrient concentrations than lower bulk density soils (Chaudhari et al., 2013; Stutter & Richards, 2012). Under prolonged periods of drought, nutrient accumulation in soil may occur due to reduced microbial growth, limited uptake by plants, and a buildup of microbial necromass during dry periods (Borken & Matzner, 2009; Sardans & Peñuelas, 2007). A further increase in soil nutrients may occur with a rain event due to cell lysis, the release of intracellular solutes, and/or the disruption of soil aggregates releasing previously protected organic matter (Borken & Matzner, 2009; Fierer & Schimel, 2002; Halverson et al., 2000). This helps to explain the strong predictive power of soil bulk density in the index of collective soil functioning.

Many studies have shown the strong correlation between pH and microbial community structure and function across a range of ecosystems (Brockett et al., 2012;

Docherty et al., 2015; Salazar et al., 2011; Sinsabaugh et al., 2008; Štursová & Baldrian, 2011). Directly, soil pH can regulate exoenzyme activities through its effects on the production of exoenzymes, induced conformational changes to the exoenzyme, potential irreversible inactivation of the exoenzyme, and overall availability of substrates and coenzymes (Frankenberger & Johanson, 1982; Tabatabai, 1994). For six of the seven exoenzymes analyzed, pH was a significant predictor variable and in some cases (NAG and PHOS) it was over 3 times more important than the next significant predictor variable (Fig. 5). Across large scales, pH can reflect controls on weathering and plant community composition, which in turn impact nutrient supply and quality, and subsequent microbial activity. For example, soils at site E, derived from limestone sedimentary rock, had the highest pH of all sites and the lowest activity for all exoenzymes except LAP (Supp. Table 3). Furthermore, LAP was the only exoenzyme to have a positive correlation with pH suggesting that activity might be suboptimal under acidic conditions (Supp. Table 5). Together these results suggest that soil pH could persistently influence microbial activity and emphasize the importance of considering this variable when predicting these responses.

Despite the importance of microbial metabolism in all known biogeochemical cycles, there are large uncertainties in earth system models regarding the scaling of soil microbial community processes from the micro- to macroscale (Berg, 2012). This uncertainty arises from the heterogeneous nature of soil properties across scales and the lack of information regarding the relative importance of drivers that influence their dynamics. Here we showed that variables such as plant cover and soil physiochemical properties, which have slower response times than soil microbial communities, can have persistent impacts on nutrient cycling and nutrient pools (Figs. 2-5). As such, even though the microbial parameters we measured represent a snapshot in time, the ability to recognize more stable drivers of these trends can provide insight into microbial processes across landscapes.

5 Conclusions

Mesquites, credited with creating "islands of fertility" in semiarid landscapes, have strong impacts on soil resource pools but their effects on soil microbial process are less established. Microbial communities and their activities often respond more rapidly to disturbance than soil physiochemical parameters, making them good indicators of changes in soil functionality with woody plant encroachment. We developed an index of collective soil functioning in a complex landscape undergoing woody plant encroachment that included not only soil carbon and nutrient pools but also microbial biomass and activity. Plant cover alone accounted for 48.5% of the variability in collective soil functioning, and soils associated with mesquites had elevated levels of nutrients, microbial exoenzyme activity, and microbial biomass compared to other plant cover types. Some of these trends intensified overtime as mesquites aged. Prediction models improved by ca. 7 % and the error term decreased by 13.4 % with the inclusion of soil physiochemical properties, water dynamics, and topographic features. The additional explanatory variables of soil age, bulk density, and pH were significant for collective soil functionality in a woody plant encroached system. For many individual soil nutrient variables, soil age and bulk density consistently emerged as significant predictors, suggesting the roles these variables may have in regulating soil water and nutrient dynamics in drylands. Soil pH was the most important predictor variable for many exoenzymes assayed, illustrating the significant role of soil pH in regulating microbial exoenzyme production and activities. Collectively, our results highlight the relationship between microbial activities and accumulation of microbial biomass and soil nutrients in association with woody plant encroachment—consistent with the "island of fertility" conceptual framework.

Acknowledgments, Samples, and Data

- This work was supported by NSF DDRI 1735693. REG and DB acknowledge the USDA National Institute of Food and Agriculture McIntire Stennis project 1016938.
- We would also like to thank: Willem J.D. van Leeuwen, Jessica Mitchell, Tom Gebhardt, Josh Farella, Noelle Espinosa, Dawson Fairbanks, Clarisa Avalos, Alexis Arizpe, Mallory Barnes, Amy Hudson, Chance Muscarella, Emma Jong, Dong Yan, David Moore, Greg Barron-Gafford, Bill Smith, Russ Scott, Cynthia Norton, Xian Wang, and the UA ART Lab for their expertise, input, analytical assistance, and help with field work and laboratory analyses
- Data supporting conclusions presented in the paper can be found at the University of Arizona Data Repository (Farella et al., 2020); https://arizona.figshare.com/articles/Data_for_Predicting_drivers_of_collective_soil_function_with_woody_plant_encroachment_in_complex_landscapes_/12273191

References

Aguiar, M. R., & Sala, O. E. (1999). Patch structure, dynamics and implications for the functioning of arid ecosystems. *Trends in Ecology & Evolution*, *14*(7), 273–277. https://doi.org/10.1016/S0169-5347(99)01612-2

Allen, O. N., & Allen, E. K. (1981). *The Leguminosae. A source book of characteristics, uses and nodulation.* Madison, Wisconsin, USA: Macmillan Publishers Ltd.

Alvarez, L. J., Epstein, H. E., Li, J., & Okin, G. S. (2012). Aeolian process effects on vegetation communities in an arid grassland ecosystem. *Ecology and Evolution*, *2*(4), 809–821. https://doi.org/10.1002/ece3.205

Archer, S. (1995). Tree-grass dynamics in a Prosopis-thornscrub savanna parkland: Reconstructing the past and predicting. *Écoscience*, *2*(1), 83–99.

Archer, S., Boutton, T. W., & Hibbard, K. A. (2000). Trees in Grasslands: Biogeochemical Consequences of Woody Plant Expansion. In E.-D. Schulze, M. Heimann, S. Harrison, E. Holland, J. Lloyd, I. C. Prentice, & D. S. Schimel (Eds.), *Global Biogeochemical Cycles in the Climate System* (pp. 115–130). San Diego: Academic Press.

Ashworth, J., Keyes, D., Kirk, R., & Lessard, R. (2001). Standard procedure in the hydrometer method for particle size analysis. *Communications in Soil Science and Plant Analysis*, *32*(5–6), 633–642. https://doi.org/10.1081/CSS-100103897

ASTM. (2007). *Standard Test Method for Particle-Size Analysis of Soils*. *The Annual Book of ASTM Standards* (Vol. D422-63). https://doi.org/10.1520/D0422-63R07E02

Van Auken, O. W. (2000). Shrub Invasions of North American Semiarid Grasslands. *Annual Review of Ecology and Systematics*, *31*(2000), 197–215.

Bardgett, R. D., & Van Der Putten, W. H. (2014). Belowground biodiversity and ecosystem functioning. *Nature*, *515*(7528), 505–511. https://doi.org/10.1038/nature13855

Batchily, A. K., Post, D. F., Bryant, R. B., & Breckenfeld, D. J. (2003). Spectral Reflectance and Soil Morphology Characteristics of Santa Rita Experimental Range Soils. In *McClaran, Mitchel P.; Ffolliott, Peter F.; Edminster, Carleton B., tech. coords. Santa Rita Experimental Range: 100 years (1903 to 2003) of accomplishments and contributions; conference proceedings; 2003 October 30-November 1; Tucson, AZ. Proc. RMRS-P-30.* (pp. 175–182).

Beck, T., Joergensen, R. G., Kandeler, E., Makeschin, F., Nuss, E., Oberholzer, H. R., & Scheu, S. (1997). An inter-laboratory comparison of ten different ways of measuring soil microbial biomass C. *Soil Biology and Biochemistry*, *29*(7), 1023–1032. https://doi.org/10.1016/S0038-0717(97)00030-8

Berg, M. P. (2012). Patterns of Biodiversity at Fine and Small Spatial Scales. In D. H. Wall, R. D. Bardgett, V. Behan-Pelletier, & W. H. van der Putten (Eds.), *Soil Ecology and Ecosystem Services* (pp. 136–152). Oxford University Press.

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Bone, J., Barraclough, D., Eggleton, P., Head, M., Jones, D. T., & Voulvoulis, N. (2014). Prioritising soil quality assessment through the screening of sites: The use of publicly collected data. *Land Degradation and Development*, *25*(3), 251–266. https://doi.org/10.1002/ldr.2138

Borken, W., & Matzner, E. (2009). Reappraisal of drying and wetting effects on C and N mineralization and fluxes in soils. *Global Change Biology*, *15*(4), 808–824. https://doi.org/10.1111/j.1365-2486.2008.01681.x

Bouyoucos, G. J. (1962). Hydrometer Method Improved for Making Particle Size Analyses of Soils1. *Agronomy Journal*, *54*(5), 464–465. https://doi.org/10.2134/agronj1962.00021962005400050028x

Bradford, M. A., Wood, S. A., Bardgett, R. D., Black, H. I. J., Bonkowski, M., Eggers, T., et al. (2014). Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(40), 14478–14483. https://doi.org/10.1073/pnas.1413707111

Brockett, B. F. T., Prescott, C. E., & Grayston, S. J. (2012). Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. *Soil Biology and Biochemistry*, *44*(1), 9–20. https://doi.org/10.1016/j.soilbio.2011.09.003

Brookes, P., Landman, A., Pruden, G., & Jenkinson, D. S. (1985). Chloroform fumigation and the release of soil nitrogen: A rapid direct extraction method to measure microbial

biomass nitrogen in soil. *Soil Biology and Biochemistry*, *17*(6), 837–842. https://doi.org/10.1016/0038-0717(85)90144-0

Browning, D. M., & Archer, S. R. (2011). Protection from livestock fails to deter shrub proliferation in a desert landscape with a history of heavy grazing. *Ecological Applications*, *21*(5), 1629–1642. https://doi.org/10.1890/10-0542.1

Browning, D. M., Archer, S. R., Asner, G. P., McClaran, M. P., & Wessman, C. A. (2008). Woody plants in grasslands: Post-encroachment stand dynamics. *Ecological Applications*, *18*(4), 928–944. https://doi.org/10.1890/07-1559.1

Browning, D. M., Franklin, J., Archer, S. R., Gillan, J. K., & Guertin, D. P. (2014). Spatial patterns of grassland – shrubland state transitions : a 74-year record on grazed and protected areas. *Ecological Applications*, *24*(6), 1421–1433.

Cable, J. M., Ogle, K., Tyler, A. P., Pavao-Zuckerman, M. A., & Huxman, T. E. (2009). Woody plant encroachment impacts on soil carbon and microbial processes: Results from a hierarchical Bayesian analysis of soil incubation data. *Plant and Soil*, *320*(1–2), 153–167. https://doi.org/10.1007/s11104-008-9880-1

Campbell, B. D., & Stafford Smith, D. M. (2000). A synthesis of recent global change research on pasture and rangeland production: Reduced uncertainties and their management implications. *Agriculture, Ecosystems and Environment*, *82*(1–3), 39–55. https://doi.org/10.1016/S0167-8809(00)00215-2

Cao, H., Chen, R., Wang, L., Jiang, L., Yang, F., Zheng, S., et al. (2016). Soil pH, total phosphorus, climate and distance are the major factors influencing microbial activity at a regional spatial scale. *Scientific Reports*, *6*(April), 1–10. https://doi.org/10.1038/srep25815

Charley, J. L., & West, N. E. (1975). Plant-Induced Soil Chemical Patterns in Some Shrub-Dominated Semi-Desert Ecosystems of Utah. *The Journal of Ecology*, *63*(3), 945–963. https://doi.org/10.2307/2258613

Chaudhari, P. R., Ahire, D. V, Ahire, V. D., Chkravarty, M., & Maity, S. (2013). Soil Bulk Density as related to Soil Texture, Organic Matter Content and available total Nutrients of Coimbatore Soil. *International Journal of Scientific and Research Publications*, *3*(1), 2250–3153. https://doi.org/10.2136/sssaj2015.11.0407

Chen, L., Xiang, W., Wu, H., Ouyang, S., Zhou, B., Zeng, Y., et al. (2019). Tree species identity surpasses richness in affecting soil microbial richness and community composition in subtropical forests. *Soil Biology and Biochemistry*, *130*(November 2018), 113–121. https://doi.org/10.1016/j.soilbio.2018.12.008

Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K., & Paul, E. (2013). The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: Do labile plant inputs form stable soil organic matter? *Global Change Biology*, *19*(4), 988–995. https://doi.org/10.1111/gcb.12113

Day, P. R. (1965). Particle Fractionation and Particle-Size Analysis. In C. A. Black (Ed.), *Methods of soil analysis, Part 1.* (pp. 545–567). Madison, Wisconsin, USA: American Society of Agronomy. https://doi.org/10.2134/agronmonogr9.1.c43

Delgado-Baquerizo, M., Maestre, F. T., Reich, P. B., Jeffries, T. C., Gaitan, J. J., Encinar, D.,

et al. (2016). Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nature Communications*, *7*(1), 10541. https://doi.org/10.1038/ncomms10541

Doane, T. a., & Horwáth, W. R. (2003). Spectrophotometric Determination of Nitrate with a Single Reagent. *Analytical Letters*, *36*(12), 2713–2722. https://doi.org/10.1081/AL-120024647

Docherty, K. M., Borton, H. M., Espinosa, N., Gebhardt, M., Gil-Loaiza, J., Gutknecht, J. L. M., et al. (2015). Key edaphic properties largely explain temporal and geographic variation in soil microbial communities across four biomes. *PLoS ONE*, *10*(11), 1–23. https://doi.org/10.1371/journal.pone.0135352

Drenovsky, R. E., Steenwerth, K. L., Jackson, L. E., & Scow, K. M. (2010). Land use and climatic factors structure regional patterns in soil microbial communities. *Global Ecology and Biogeography*, *19*(1), 27–39. https://doi.org/10.1111/j.1466-8238.2009.00486.x

Eldridge, D. J., Bowker, M. A., Maestre, F. T., Roger, E., Reynolds, J. F., & Whitford, W. G. (2011). Impacts of shrub encroachment on ecosystem structure and functioning: Towards a global synthesis. *Ecology Letters*, *14*(7), 709–722. https://doi.org/10.1111/j.1461-0248.2011.01630.x

Evans, S. E., & Wallenstein, M. D. (2012). Soil microbial community response to drying and rewetting stress: Does historical precipitation regime matter? *Biogeochemistry*, *109*(1–3), 101–116. https://doi.org/10.1007/s10533-011-9638-3

Farella, M. M., Breshears, D. D., & Gallery, R. (2020). Data for "Predicting drivers of collective soil function with woody plant encroachment in complex landscapes." https://doi.org/10.25422/azu.data.12273191

Field, C. B. (1998). Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science*, *281*(5374), 237–240. https://doi.org/10.1126/science.281.5374.237

Fierer, N., & Schimel, J. P. (2002). Effects of drying-rewetting frequency on soil carbon and nitrogen transformations. *Soil Biology and Biochemistry*, *34*(6), 777–787. https://doi.org/10.1016/S0038-0717(02)00007-X

Fierer, N., Ladau, J., Clemente, J. C., Leff, J. W., Owens, S. M., Pollard, K. S., et al. (2013). Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. *Science*, *342*(6158), 621–4. https://doi.org/10.1126/science.1243768

Florinsky, I. V., McMahon, S., & Burton, D. L. (2004). Topographic control of soil microbial activity: A case study of denitrifiers. *Geoderma*, *119*(1–2), 33–53. https://doi.org/10.1016/S0016-7061(03)00224-6

Frankenberger, W. T., & Johanson, J. B. (1982). Effect of pH on enzyme stability in soils. *Soil Biology and Biochemistry*, *14*(5), 433–437. https://doi.org/10.1016/0038-0717(82)90101-8

Gamfeldt, L., Hillebrand, H., & Jonsson, P. R. (2008). Multiple functions increase the importance of biodiversity for overall ecosystem functioning. *Ecology*, *89*(5), 1223–1231. https://doi.org/10.1890/06-2091.1

Gebhardt, M., Fehmi, J. S., Rasmussen, C., & Gallery, R. E. (2017). Soil amendments alter

plant biomass and soil microbial activity in a semi-desert grassland. *Plant and Soil*, *419*(1–2), 53–70. https://doi.org/10.1007/s11104-017-3327-5

Glass, G. V, Peckham, P. D., & Sanders, J. R. (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Educational Research*, *42*(3), 237–288.

González-Roglich, M., Swenson, J. J., Jobbágy, E. G., & Jackson, R. B. (2014). Shifting carbon pools along a plant cover gradient in woody encroached savannas of central Argentina. *Forest Ecology and Management*, *331*, 71–78. https://doi.org/10.1016/j.foreco.2014.07.035

Gupta, S. C., & Larson, W. E. (1979). Estimating soil water retention characteristics from particle size distribution, organic matter percent, and bulk density. *Water Resources Research*, *15*(6), 1633–1635. https://doi.org/10.1029/WR015i006p01633

Halverson, L. J., Jones, T. M., & Firestone, M. K. (2000). Release of intracellular solutes by four soil bacteria exposed to dilution stress. *Soil Science Society of America Journal*, *64*(5), 1630–1637.

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo Results in Methodological Research : The One- and Two-Factor Fixed Effects ANOVA Cases Author (s): Michael R . Harwell , Elaine N . Rubinstein , William S . Hayes and Corley C . Olds Published by : American Educational Research A. *American Educational Research Association and American Statistical Associatoin*, *17*(4), 315–339.

Hollister, E. B., Schadt, C. W., Palumbo, A. V., James Ansley, R., & Boutton, T. W. (2010). Structural and functional diversity of soil bacterial and fungal communities following woody plant encroachment in the southern Great Plains. *Soil Biology and Biochemistry*, *42*(10), 1816–1824. https://doi.org/10.1016/j.soilbio.2010.06.022

Huxman, T. E., Snyder, K. a, Tissue, D., Leffler, a J., Ogle, K., Pockman, W. T., et al. (2004). Precipitation pulses and carbon fluxes in semiarid and arid ecosystems. *Oecologia*, *141*(2), 254–268. https://doi.org/10.1007/s00442-004-1682-4

Huxman, T. E., Wilcox, B. P., Breshears, D. D., Scott, R. L., Snyder, K. A., Small, E. E., et al. (2005). Ecohydrological Implications of Woody Plant Encroachment. *Ecology*, *86*(2), 308–319. https://doi.org/10.1890/03-0583

Johnson, H. B., & Mayeux, H. S. (1990). Prosopis glandulosa and the nitrogen balance of rangelands: extent and occurrence of nodulation. *Oecologia*, *84*(2), 176–185. https://doi.org/10.1007/BF00318269

Johnson, S. L., Kuske, C. R., Carney, T. D., Housman, D. C., Verne, G.-G. La, & Belanp, J. (2012). Increased temperature and altered summer precipitation have differential effects on biological soil crusts in a dryland ecosystem. *Global Change Biology*, *18*(8), 2583–2593. https://doi.org/10.1111/j.1365-2486.2012.02709.x

Karlen, D. L., Mausbach, M. J., Doran, J. W., Cline, R. G., Harris, R. F., & Schuman, G. E. (1997). Soil quality: A concept, definition, and framework for evaluation: (A guest editorial). *Soil Science Society of America Journal*, *61*(1), 4–10. https://doi.org/10.2136/sssaj1997.03615995006100010001x

Kuske, C., Ticknor, L., Miller, M., Dunbar, J., Davis, J., Barns, S., & Belnap, J. (2002).

Comparison of soil bacterial communities in rhizospheres of three plant species and the interspaces in an arid grassland. *Applied and Environmental Microbiology*, *68*(4), 1854–1863. https://doi.org/10.1128/AEM.68.4.1854

Lambers, H., Raven, J. A., Shaver, G. R., & Smith, S. E. (2008). Plant nutrient-acquisition strategies change with soil age. *Trends in Ecology and Evolution*, *23*(2), 95–103. https://doi.org/10.1016/j.tree.2007.10.008

Li, H., Zhang, J., Hu, H., Chen, L., Zhu, Y., Shen, H., et al. (2017). Shift in soil microbial communities with shrub encroachment in Inner Mongolia grasslands, China. *European Journal of Soil Biology*, *79*, 40–47. https://doi.org/10.1016/j.ejsobi.2017.02.004

Liao, J. D., & Boutton, T. W. (2008). Soil microbial biomass response to woody plant invasion of grassland. *Soil Biology and Biochemistry*, *40*(5), 1207–1216. https://doi.org/10.1016/j.soilbio.2007.12.018

Liu, W., Xu, W., Han, Y., Wang, C., & Wan, S. (2007). Responses of microbial biomass and respiration of soil to topography, burning, and nitrogen fertilization in a temperate steppe. *Biology and Fertility of Soils*, *44*(2), 259–268. https://doi.org/10.1007/s00374-007-0198-6

Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research*, *66*(4), 579–619. https://doi.org/10.3102/00346543066004579

Ludwig, J. A., Wilcox, B. P., Breshears, D. D., Tongway, D. J., & Imeson, A. C. (2005). Vegetation patches and runoff-erosion as interacting ecohydrological processes in semiarid landscapes. *Ecology*, *86*(2), 288–297. https://doi.org/10.1890/03-0569

Maestre, F. T., Puche, M. D., Guerrero, C., & Escudero, A. (2011). Shrub encroachment does not reduce the activity of some soil enzymes in Mediterranean semiarid grasslands. *Soil Biology and Biochemistry*, *43*(8), 1746–1749. https://doi.org/10.1016/j.soilbio.2011.04.023

Maestre, F. T., Quero, J. L., Gotelli, N. J., Escudero, A., Ochoa, V., Delgado-Baquerizo, M., et al. (2012). Plant Species Richness and Ecosystem Multifunctionality in Global Drylands. *Science*, *335*(6065), 214–218. https://doi.org/10.1126/science.1215442

McClaran, M. P., McMurtry, C. R., & Archer, S. R. (2013). A tool for estimating impacts of woody encroachment in arid grasslands: Allometric equations for biomass, carbon and nitrogen content in Prosopis velutina. *Journal of Arid Environments*, *88*, 39–42. https://doi.org/10.1016/j.jaridenv.2012.08.015

McClaran, Mitchel P. (2003). A Century of Vegetation Change on the Santa Rita Experimental Range. *USDA Forest Service Proceedings RMRS-P-30*, (520), 16–33.

McClaran, Mitchel P., Angell, D. L., & Wissler, C. (2002). *Santa Rita Experimental Range digital database: user's guide*. *USDA Forest Service - General Technical Report RMRS-GTR*. Ogden, UT. Retrieved from http://www.fs.fed.us/rm/pubs/rmrs_gtr100.pdf

MEA. (2005). *Ecosystems and Human Well-Being: Desertification Synthesis*. Washington, D.C.

Nelson, D. W., & Sommers, L. E. (1965). Total Carbon, Organic Carbon, and Organic Matter. In *Methods of Soil Analysis. Part 2. Chemical and Microbiological Properties*

(pp. 539–579). Madison, Wisconsin: American Society of Agronomy, Inc.

Nicol, G. W., Leininger, S., Schleper, C., & Prosser, J. I. (2008). The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Environmental Microbiology*, *10*(11), 2966–2978. https://doi.org/10.1111/j.1462-2920.2008.01701.x

Pacala, S. W., Hurtt, G. C., Baker, D., Peylin, P., Houghton, R. A., Birdsey, R. A., et al. (2001). Consistent Land- and Atmosphere-Based U.S. Carbon Sink Estimates. *Science*, *292*(5525), 2316–2320. https://doi.org/10.1126/science.1057320

Paz-Ferreiro, J., & Fu, S. (2016). Biological Indices for Soil Quality Evaluation: Perspectives and Limitations. *Land Degradation and Development*, *27*(1), 14–25. https://doi.org/10.1002/ldr.2262

Perakis, S. S., Pett-Ridge, J. C., & Catricala, C. E. (2017). Nutrient feedbacks to soil heterotrophic nitrogen fixation in forests. *Biogeochemistry*, *134*(1–2), 41–55. https://doi.org/10.1007/s10533-017-0341-x

Potts, D. L., Scott, R. L., Bayram, S., & Carbonara, J. (2010). Woody plants modulate the temporal dynamics of soil moisture in a semi-arid mesquite savanna. *Ecohydrology*, *3*(December 2009), 20–27. https://doi.org/10.1002/eco.91

Puigdefabregas, J., Sole, A., Gutierrez, L., del Barrio, G., & Boer, M. (1999). Scales and processes of water and sediment redistribution in drylands: results from the Rambla Honda field site in Southeast Spain. *Earth-Science Reviews*, *48*(1–2), 39–70. https://doi.org/10.1016/S0012-8252(99)00046-X

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Ratajczak, Z., Nippert, J. B., & Collins, S. L. (2012). Woody encroachment decreases diversity across North American grasslands and savannas. *Ecology*, *93*(4), 697–703. https://doi.org/10.1890/11-1199.1

Reyes-Reyes, G., Baron-Ocampo, L., Cuali-Alvarez, I., Frias-Hernandez, J. T., Olalde-Portugal, V., Varela Fregoso, L., & Dendooven, L. (2002). C and N dynamics in soil from the central highlands of Mexico as affected by mesquite (Prosopis spp.) and huizache (Acacia tortuoso): A laboratory investigation. *Applied Soil Ecology*, *19*(1), 27–34. https://doi.org/10.1016/S0929-1393(01)00169-X

Reynolds, J. F., Smith, D. M. S., Lambin, E. F., Turner, B. L., Mortimore, M., Batterbury, S. P. J., et al. (2007). Global Desertification: Building a Science for Dryland Development. *Science*, *316*(5826), 847–851. https://doi.org/10.1126/science.1131634

Reynolds, James F, Kemp, P. R., Ogle, K., & Fernández, R. J. (2004). Modifying the "pulse-reserve" paradigm for deserts of North America: precipitation pulses, soil water, and plant responses. *Oecologia*, *141*(2), 194–210. https://doi.org/10.1007/s00442-004-1524-4

Rhine, E. D., Mulvaney, R. L., Pratt, E. J., & Sims, G. K. (1998). Improving the Berthelot Reaction for Determining Ammonium in Soil Extracts and Water. *Soil Science Society of America Journal*, *62*(2), 473. https://doi.org/10.2136/sssaj1998.03615995006200020026x

Robertson, G., Wedin, D., Groffman, P., Blair, J., Holland, E., Nadelhoffer, K., et al. (1999). Soil carbon and nitrogen availability. Nitrogen mineralization, nitrification and soil respiration potentials. *Standard Soil Methods for Long-Term Ecological Research*, 258–271.

Salazar, S., Sánchez, L. E., Alvarez, J., Valverde, A., Galindo, P., Igual, J. M., et al. (2011). Correlation among soil enzyme activities under different forest system management practices. *Ecological Engineering*, *37*(8), 1123–1131. https://doi.org/10.1016/j.ecoleng.2011.02.007

Sardans, J., & Peñuelas, J. (2007). Drought changes phosphorus and potassium accumulation patterns in an evergreen Mediterranean forest. *Functional Ecology*, *21*(2), 191–201. https://doi.org/10.1111/j.1365-2435.2007.01247.x

Saxton, K. E., & Rawls, W. J. (2006). Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions. *Soil Science Society of America Journal*, *70*(5), 1569–1578. https://doi.org/10.2136/sssaj2005.0117

Schade, J. D., & Hobbie, S. E. (2005). Spatial and temporal variation in islands of fertility in the Sonoran Desert. *Biogeochemistry*, *73*(3), 541–553. https://doi.org/10.1007/s10533-004-1718-1

Schlesinger, W H, Reynolds, J. F., Cunningham, G. L., Huenneke, L. F., Jarrell, W. M., Virginia, R. a, & Whitford, W. G. (1990). Biological Feedbacks in Global Desertification. *Science*, *247*(4946), 1043–1048. https://doi.org/10.1126/science.247.4946.1043

Schlesinger, William H, & Pilmanis, A. M. (1998). Plant-soil Interactions in Deserts. *Biogeochemistry*, *42*(1), 169–187. https://doi.org/10.1023/A:1005939924434

Selmants, P. C., & Hart, S. C. (2010). Phosphorus and soil development: Does the Walker and Syers model apply to semiarid ecosystems? *Ecology*, *91*(2), 474–484. https://doi.org/10.1890/09-0243.1

Sevastas, S., Gasparatos, D., Botsis, D., Siarkos, I., Diamantaras, K. I., & Bilas, G. (2018). Predicting bulk density using pedotransfer functions for soils in the Upper Anthemountas basin, Greece. *Geoderma Regional*, *14*, e00169. https://doi.org/10.1016/j.GEODRS.2018.e00169

Sinsabaugh, R. L., Antibus, R. K., & Linkins, A. E. (1991). An enzymic approach to the analysis of microbial activity during plant litter decomposition. *Agriculture, Ecosystems and Environment*, *34*(1–4), 43–54. https://doi.org/10.1016/0167-8809(91)90092-C

Sinsabaugh, R. L., Lauber, C. L., Weintraub, M. N., Ahmed, B., Allison, S. D., Crenshaw, C., et al. (2008). Stoichiometry of soil enzyme activity at global scale. *Ecology Letters*, *11*(11), 1252–1264. https://doi.org/10.1111/j.1461-0248.2008.01245.x

Štursová, M., & Baldrian, P. (2011). Effects of soil properties and management on the activity of soil organic matter transforming enzymes and the quantification of soil-bound and free activity. *Plant and Soil*, *338*(1), 99–110. https://doi.org/10.1007/s11104-010-0296-3

Stutter, M. I., & Richards, S. (2012). Relationships between soil physicochemical, microbiological properties, and nutrient release in buffer soils compared to field soils. *Journal of Environmental Quality*, *41*(2), 400–409.

https://doi.org/10.2134/jeq2010.0456

Subirge, T. G. (1983). *Soil-plant relations on the Santa Rita Experimental Range*. University of Arizona.

Tabatabai, M. A. (1994). Soil Enzymes. In R. W. Weaver, S. Angle, P. Bottomley, D. Bezdicek, S. Smith, A. Tabatabai, & A. Wollom (Eds.), *Methods of Soil Analysis. Part 2. Microbiological and Biochemical Properties* (pp. 775–833). Madison, Wisconsin: Soil Science Society of America.

Throop, H. L., & Archer, S. R. (2008). Shrub (Prosopis velutina) encroachment in a semidesert grassland: Spatial-temporal changes in soil organic carbon and nitrogen pools. *Global Change Biology*, *14*(10), 2420–2431. https://doi.org/10.1111/j.1365-2486.2008.01650.x

Vitousek, P. M., & Farrington, H. (1997). Nutrient Limitation and Soil Development: Experimental Test of a Biogeochemical Theory. *Biogeochemistry*, *37*(1), 63–75. https://doi.org/https://doi.org/10.1023/A:1005757218475

Wallenstein, M. D., Haddix, M. L., Lee, D. D., Conant, R. T., & Paul, E. a. (2012). A litter-slurry technique elucidates the key role of enzyme production and microbial dynamics in temperature sensitivity of organic matter decomposition. *Soil Biology and Biochemistry*, *47*, 18–26. https://doi.org/10.1016/j.soilbio.2011.12.009

Wang, Y. P., Houlton, B. Z., & Field, C. B. (2007). A model of biogeochemical cycles of carbon, nitrogen, and phosphorus including symbiotic nitrogen fixation and phosphatase production. *Global Biogeochemical Cycles*, *21*(1), 1–15. https://doi.org/10.1029/2006GB002797

Wheeler, C. W., Archer, S. R., Asner, G. P., & Mcmurtry, C. R. (2007). Climatic/Edaphic Controls on Soil Carbon/Nitrogen Response To Shrub Encroachment in Desert Grassland. *Ecological Applications*, *17*(7), 1911–1928. https://doi.org/10.1890/06-1580.1

Yannarell, A. C., Menning, S. E., & Beck, A. M. (2014). Influence of Shrub Encroachment on the Soil Microbial Community Composition of Remnant Hill Prairies. *Microbial Ecology*, *67*(4), 897–906. https://doi.org/10.1007/s00248-014-0369-6

Zhou, J., Xue, K., Xie, J., Deng, Y., Wu, L., Cheng, X., et al. (2011). Microbial mediation of carbon-cycle feedbacks to climate warming. *Nature Climate Change*, *2*(2), 106–110. https://doi.org/10.1038/nclimate1331

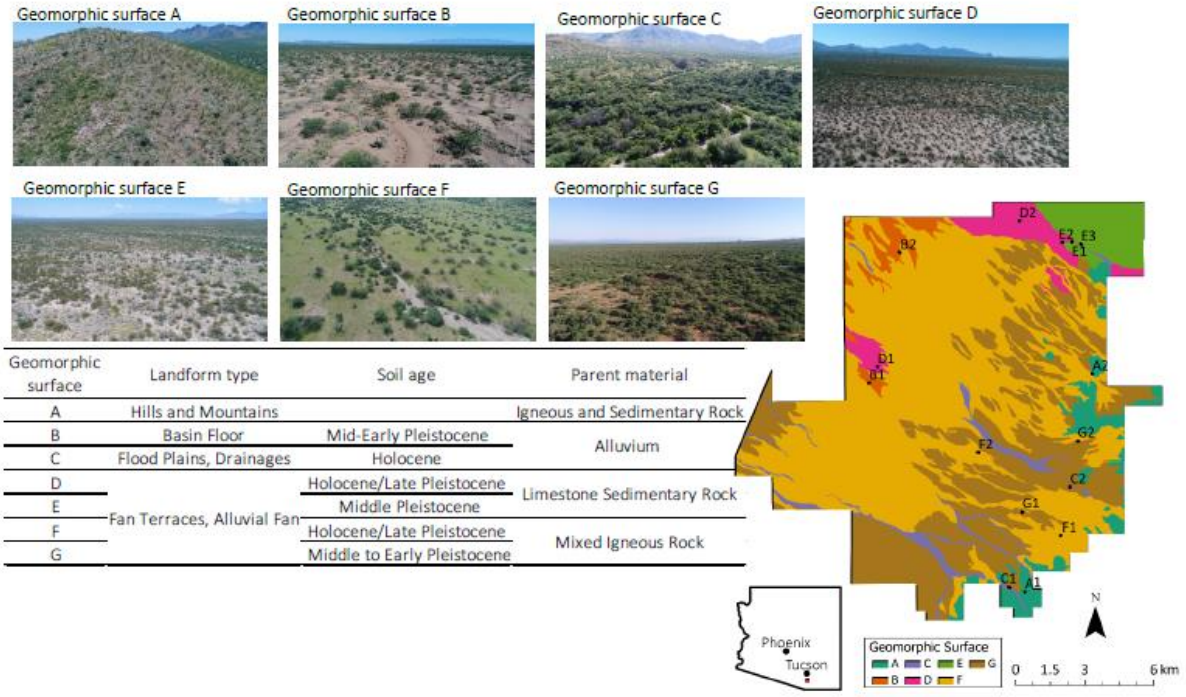

Fig. 1: Geomorphic surfaces and location of sites across the study area. Photographs provide an overview of the diversity of the geomorphic surfaces. Map shows location of study area (red) within Arizona, geomorphic surface distribution across the study area, and location of the sites included in the analysis (labeled A-G corresponding to each geomorphic surface and numbers indicate the site replicate). Geomorphic surfaces (defined in Batchily et al. 2003) captured variation in landform type, soil age, and parent material. The table inset defined each of these classifications for the geomorphic surfaces across the study area. Drone imagery courtesy of Willem J.D. van Leeuwen.

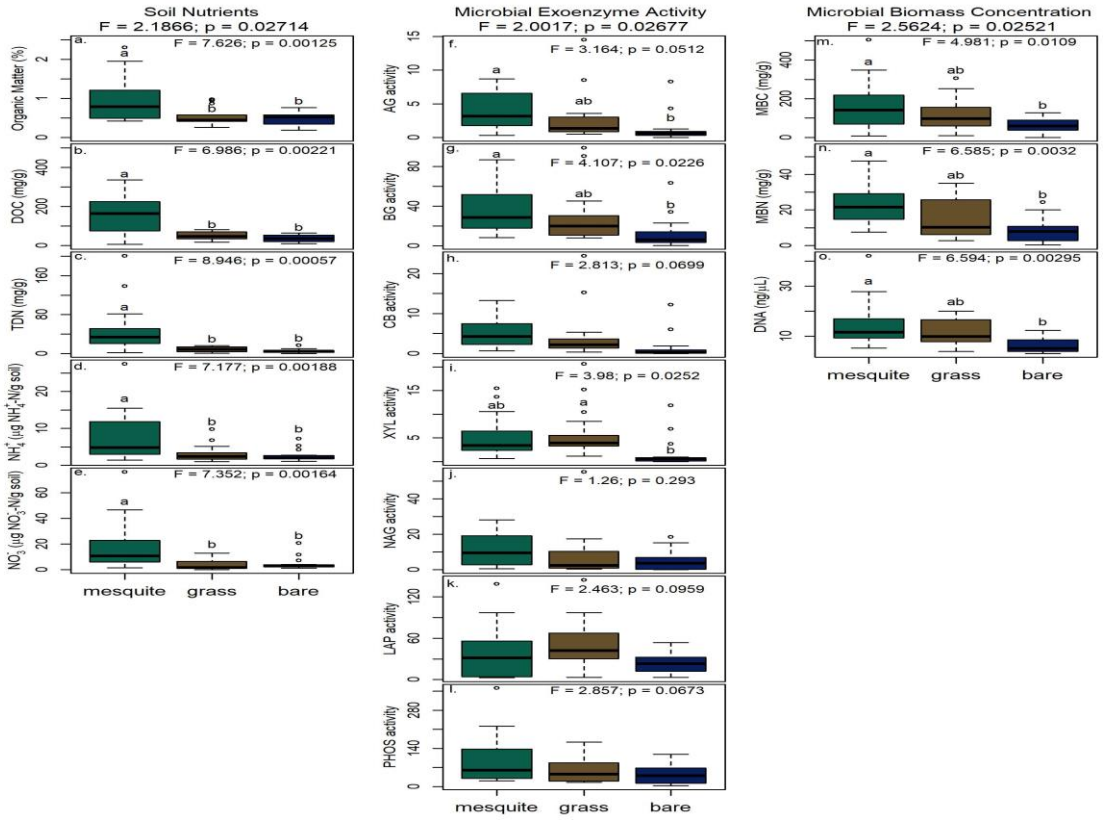

Fig. 2: Differences in soil nutrients (a-e), exoenzyme activity (f-l), and microbial biomass concentration (m-o) in soils collected from mesquite bole (N = 18), grass (N = 17), and bare ground (N = 15) samples. Boxes in the boxplots represent the 75th (upper portion) and 25th (bottom portion) percentiles of soil organic matter % (a), dissolved organic carbon (DOC; b), total dissolved nitrogen (TDN; c), ammonium concentration (NH4+; d), nitrate concentration (NO3-; e), α-Glucosidase exoenzyme activity (AG; f), β-Glucosidase exoenzyme activity (BG; g), β-D-cellubiosidase exoenzyme activity (CB; h), β-Xylosidase exoenzyme activity (XYL; i), N-acetyl-β-Glucosaminidase exoenzyme activity (NAG;j), leucine aminopeptidase exoenzyme activity (LAP; k); Phosphatase exoenzyme activity (PHOS; l) ; microbial biomass carbon (MBC; m), microbial biomass nitrogen (MBN; n), and DNA concentration (o). The band in the middle of the box represents the median value. Outliers are symbolized as open points and whiskers extend to lower and upper quartiles times 1.5 the interquartile range. Values at the top of each column represent MANOVA results for all variables in the column. Letters, which represent Tukey HSD significance, are presented when $p < 0.05$. All exoenzyme activities presented in (nmol activity/g dry soil/hr).

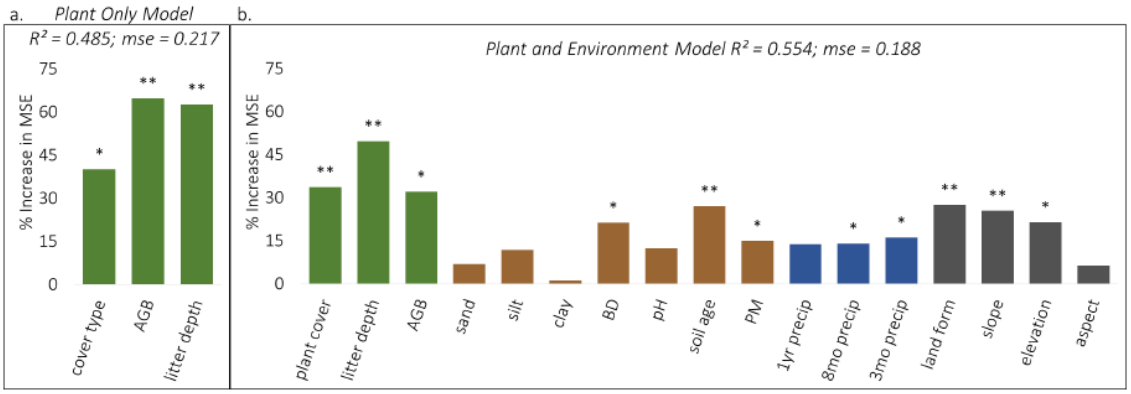

Fig. 3: Predictor variable importance values for the collective soil functionality index. Graph shows the random forest variable importance (% increase in MSE; mean square error) of plant variables only (a) and plant and environment variables (b) for the collective soil functionality index. Plant variables (green bars), soil physiochemical properties (brown bars), precipitation values (blue bars), and topographic factors (gray bars) are symbolized according to variable type. The "*" and "**" indicate significant levels ($p < 0.05$ and $p < 0.01$, respectively) of predictor variables. AGB, aboveground biomass; BD, bulk density; PM, soil parent material.

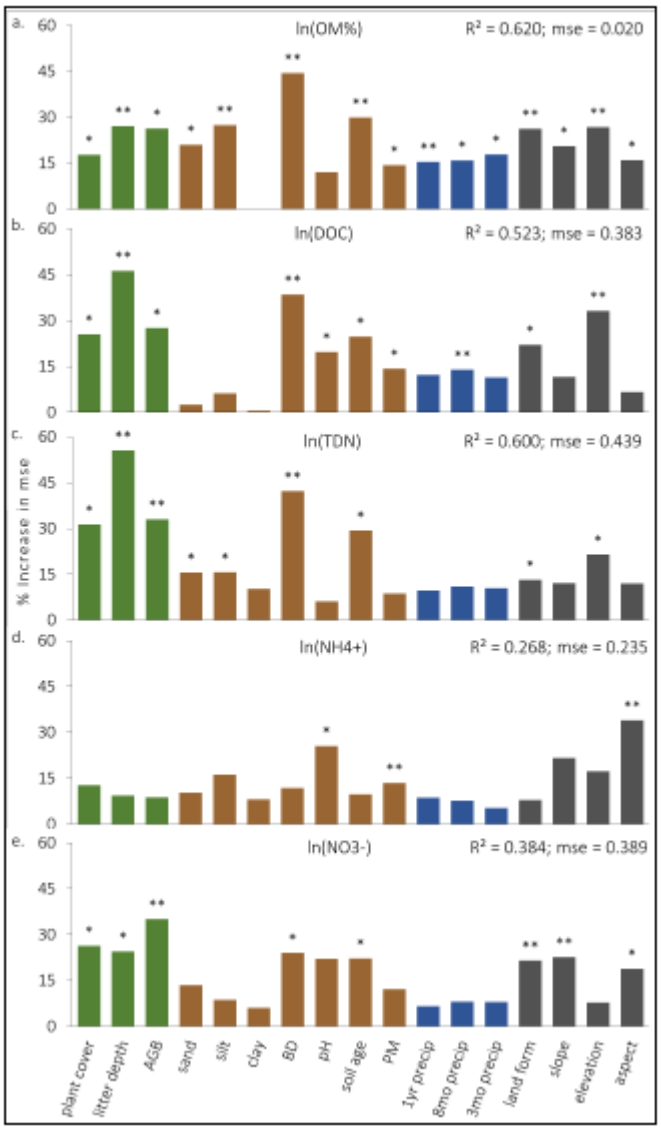

Fig. 4: Predictor variable importance values for individual soil nutrients. Graph shows the random forest variable importance (% increase in MSE; mean square error) of plant and environment variables for percent soil organic matter (OM %; a), dissolved organic carbon (DOC; b), total dissolved nitrogen (TDN; c), ammonium concentration (NH4+; d), and nitrate concentration (NO3-; e). Plant variables (green bars), soil physiochemical properties (brown bars), precipitation values (blue bars), and topographic factors (gray bars) are symbolized according to variable type. The "*" and "**" indicate significant levels ($p < 0.05$ and $p < 0.01$, respectively) of predictor variables. AGB, aboveground biomass; BD, bulk density; PM, soil parent material.

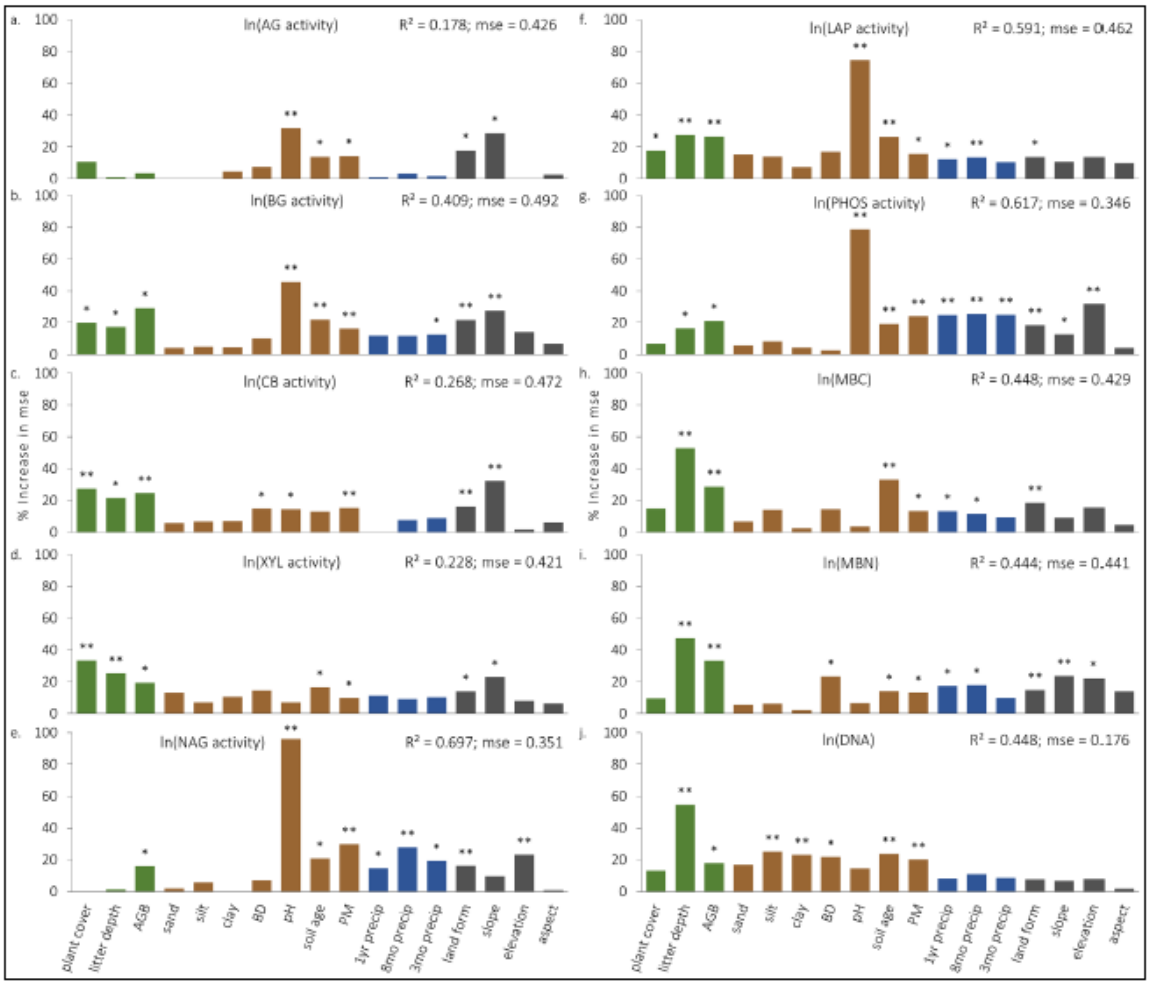

Fig. 5: Predictor variable importance values for individual microbial activity and biomass metrics. Graph shows the random forest variable importance (% increase in MSE; mean square error) of plant and environment variables for potential exoenzyme activities (a-g) and microbial biomass concentrations (h-j). Plant variables (green bars), soil physiochemical properties (brown bars), precipitation values (blue bars), and topographic factors (gray bars) are symbolized according to variable type. The "*" and "**" indicate significant levels ($p < 0.05$ and $p < 0.01$, respectively) of predictor variables. AGB, aboveground biomass; BD, bulk density; PM, soil parent material. AG, α-Glucosidase exoenzyme activity; BG, β-Glucosidase exoenzyme activity; CB, β-D-cellubiosidase exoenzyme activity; XYL, β-Xylosidase exoenzyme activity; NAG, N-acetyl-β-Glucosaminidase exoenzyme activity; LAP, leucine aminopeptidase exoenzyme activity; PHOS, Phosphatase exoenzyme activity; MBC, microbial biomass carbon; MBN, microbial biomass nitrogen; DNA, DNA concentration.