Statistical analysis of microbial soil phyla ecology

University College London

8th May 2018

Abstract

Metagenomics has emerged as a new area of research with the drastic decrease in cost of sequencing due to increase of computational power and new generation sequencing methods (NGS) such as Illumina. Therefore, both the development of bioinformatics tools with high throughput DNA sequencing has allowed the emergence of metagenomics allowing to study complete microbial communities in their natural environment such as the microbiome in the gut, marine water samples or in our case soil samples. A small-scale soil metagenomics analysis was carried out by collecting 30 samples from Central London. Following DNA sequencing, Cartesius, a high performance computing clusters was used together with QIIME scripts in order to elucidate a potential correlation between the abundance of bacterial phylum and its ecological environment. Using a statistical model by correlating various metadata variables it was possible to investigate a relationship between the biological and chemical characteristics of a microbial environment and the abundance of bacterial phylum.

Introduction

Metagenomics has progressively replaced the traditional tedious culturing methods since it appears that most microorganisms cannot grow in vitro. An interest for microbial diversity present in our environment has been raised with 1580 publications in 2015 compare to 2002 with only 3 articles. **[1] [2]**

With culture-independent approach, it has been possible to discover unknown species and analyse how microbes are organised in a community soil samples together with their metabolic outcome. Metagenomics relies on gene sequences which encodes functional RNAs. **[3]** To date, similarly to the human genome project, the earth microbiome project (EMP) which started in 2010 provides a catalogue of multiple studies performed on various microbial community samples to elucidate relationships between them based on gradients of environment, biological and chemical factors. **[4]** 16s rRNA is a subunit component of prokaryotic ribosomes sequence, this property enables its use for phylogenetic studies as it is highly conserved across individual group of microorganisms. By amplifying the genomic DNA from hypervariable region of 16s rRNA and tagging the amplified products with unique barcodes, it is possible to sequence the whole DNA of a given soil sample and then compare the sequences to known microbial database.

The research explained in this paper which could be included in the EMP focuses on assessing the impact of specific ecological meta-data; pH, sample depth, nitrogen, phosphorus and potassium on the relative abundance of a specific phylum. Using, a multivariable linear regression (OLS), it was possible to observe some correlations, which need to be carefully handle due to the low and clustered number of samples collected. In addition, the insufficient accuracy of the kits used to analyse the ecological data generated statistics which we cannot rely on being mostly non-significant.

Material and Methods

After the collection of thirty samples from different locations with various vegetation types within Central London; footfall, moisture and sample depth were noted and using the *HI3895 Soil testing kit* metadata was measured by eye to identify pH, phosphorus, nitrogen and potassium ions concentrations. Following collection, various subsequent experiments to isolate the 16S rRNA of

prokaryotes from soil samples prior sequencing were carried out using manufacturer guided kits.
Firstly, genomic DNA was extracted from soil sample removing any **non-organic DNA**, cell debris and proteins using the *DNeasy PowerSoil Kit*. Then, PCR primers were designed to amplify the v4 hypervariable region with Golay barcodes to enable a high level of multiplicity. The PCR reactions to amplify the DNA sequence from the v4 region were performed with the *Bioline BioMix PCR kit.* The solution from the PCR was further purified to remove any dNTPs, primers, proteins with a one hundred nucleotides cut-off using *QIAquick PCR purification kit.* Following purification, double stranded DNA concentration was quantified the using a fluorescent dye which binds to major groove of dsDNA provided by *SpectraMax Quant AccuClear Nano ds-DNA assay kit.* The fluorescence was further measured with *Molecular Devices Spectramax i3* to determine the concentration.
Two gel electrophoresis were carried at crucial stages to verify the integrity of our product, first after isolation of genomic DNA from soil to check the quality of gDNA and post PCR to check the quality of amplicons. Due to the observation of bubble products on the second gel electrophoresis, PCR had to be optimized and carried out again with less cycles.
Using the measured DNA concentration, all samples were diluted to the same concentration and pooled together prior Illumina sequencing on *Illumina's MiSeq sequencer*. Illumina sequencing is a process allowing to sequence simultaneously large number of samples on high throughput instrument, a method called multiplexing. With multiplex libraries, unique barcodes sequences are added to each DNA fragment during library preparation so that each read can be identified and sorted before final data analysis.
The returned DNA sequences were analysed for further investigation. 8.6M read pairs were generated with high quality (91% of bases > Q30). The Q scores is attributed to each base pair read as the probability of an error in base calling, with Q30 corresponding to 1 error in 1000 reads. QIIME (Quantitative Insight Into Microbial Ecology) is a bioinformatics pipeline which provides various python scripts to perform high-throughput analysis on raw DNA data. High computational power was provided with Cartesius high performance computer clusters.
The first task was to join forward and reverse reads from the fasta files of the sequencing data followed by demultiplexing which match the barcodes with the read-files, the output folder contains all the sequences corresponding to a specific sample.
Following this stage, OTU picking, groups similar sequence reads into an operational taxonomic unit (OTU) referring to clusters of organisms. Various reference databases provide the OTU sequences, the one used in the research is SILVA, a more updated database than the default QIIME one, Greengenes. **[5][6]** 97% similarity was chosen as the cut-off to consider it as an individual OTU. A precise type of picking OTU python script was used, open reference OTU, where remaining reads which did not match with known sequences are clustered de novo, thus it does not discard reads that may represent novel OTUs. GitHub repository provides all the python scripts used during this analysis to calculate several diversity metrics such as alpha diversity which gives the diversity of OTUs within a sample and beta diversity observed with PCoA plots displaying the diversity between samples or in other words, how closely related the samples are. **[7]**
In an attempt to find a relationship between metadata and the phylum type, the second level for taxonomy classification, a statistical model was created called a multivariable linear regression where each qualitative metadata was transformed into numerical values to fit into the model. It predicts the linear association between various variables corresponding to the metadata and the percentage of abundance of a specific phylum.

Results

The OTU picking step, revealed more than 5 million unique OTUs using the open-reference OTU picking. Alpha diversity which communicates the number of operational taxonomic units per sample was measured and revealed a similar diversity for most of the samples with two outliers *515rcbc20* and *515rcbc12.* The presence of outliers can be explained due to their distant locations. *515rcbc12* was collected at Regent's park differing from the 29 others in Gordon square, as observed on *Figure 1*.

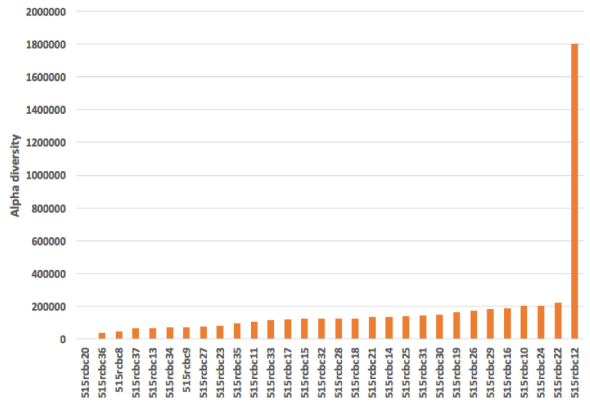

Figure 1: Alpha diversity of the 30 samples from Central London

The similarity observed in terms of species richness from the various soil samples (*Figure 1)* took our analysis to another point of view. The data collected from the taxonomy has revealed more diversity and interesting differences across samples highlighted on the heatmap *Figure 2*. For the purpose of this experiment the taxonomy analysis was limited to the level of classification 2: phylum. Differences observed suggested that an in-depth analysis could reveal some relationships between the ecological metadata of samples and the abundance of a specific phylum. The part above the heatmap framed in red shows some phylum which are much less abundant in some samples contrasted by some bright squares.

In this paper, the 10 most abundant phylum were selected for further analysis. *Figure 3* depicts the relative abundance of each of these phylum which represents 97.6 % of all phylum in the soil samples collected.

A multivariable ordinary least squares model (OLS) was created to analyse the relationship between metadata and the relative abundance of phylum. All metadata variables were assumed to vary linearly. The model coded in a software called Stata **[8]** follows the equation depicted on *Figure 4* were all variables except the pH and Sample Depth are dummy variables taking the value 1 if observed in the sample or 0 otherwise. The $Abundance_i$ corresponds to the abundance of a phylum. The β are the coefficients for which each variables vary relatively and u_i is the error term.

Therefore, by analysing coefficients, it was possible to identify the relative effect of a variable to the other ones on the abundance of a phylum. The results shown in Appendix were translated into more comprehensive tables.

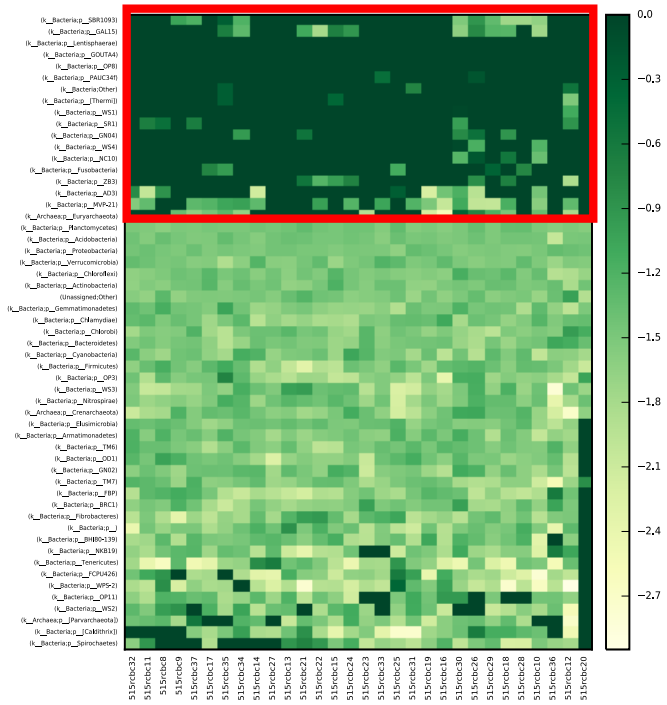

Figure 2: Phylum taxonomy abundance in the 30 samples presented as a heat map where the brightest colours represent the highest abundance. We can observe disparity between samples in terms of phylum abundance.

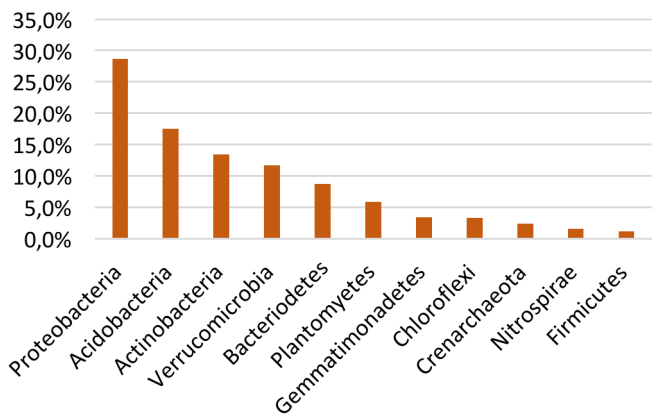

Figure 3: Relative abundance (%) of the 10 most present phylum in the 30 samples

$$
\begin{aligned}
Abundance_i = \ & \beta_0 + \beta_1 pH_i + \beta_2 SampleDepth_i + \beta_3 Dry_i + \beta_4 Moist_i + \beta_5 mediumk_i \\
& + \beta_6 mediumhighk_i + \beta_7 lowk_i + \beta_8 tracek_i + \beta_9 tracelowk_i \\
& + \beta_{10} highnitrogen_i + \beta_{11} mediumnitrogen_i \\
& + \beta_{12} lowmediumnitrogen_i + \beta_{13} lownitrogen_i \\
& + \beta_{14} tracenitrogen_i + \beta_{15} tracelownitrogen_i \\
& + \beta_{16} highphosphorus_i + \beta_{17} mediumphosphorus_i \\
& + \beta_{18} lowmediumphosphorus_i + \beta_{19} lowphosphorus_i \\
& + \beta_{20} tracelowphosphorus_i + \beta_{21} tracephosphorus_i + u_i
\end{aligned}
$$

Figure 4: Equation the multiple variables linear regression coded in Stata.

As shown on *Figure 5*, regarding the pH and sample depth, a percentage variation of phylum abundance was calculated when increasing the variables by one corresponding unit; highlighting the effect of either increasing the pH or the sample depth on the abundance. *Verrucomicrobia*, *Gemmatimonadetes*, and *Acidobacteria* were the most reactive with -1;35%, +0.38%, +0.45% variations in abundance respectively relative to the pH. *Bacteroidetes*, *Firmicutes* and *Acidobacteria* regarding the Sample Depth with -0.81%, +1.12% and +0.43% respectively. *Figure 6,* similarly depicts how variations in moisture, potassium, nitrogen and phosphorus impact the abundance of phylum. The analysis was performed more qualitatively. The highest β coefficients (*Figure 4*) reveal a more sensitive increase of the phylum abundance suggesting a potential preferred environment. Thus, the preferred qualitative ions concentrations were attributed to individual phylum. However, when the difference between the coefficients of the dummy variables was observed to be below 0.1%, any qualitative decision was rejected and a grey squares was added.

These tables did not reveal any trends of preferred phylum chemical and biological environment supporting a more random ecological classification. Nevertheless, the model is mostly not statistically significant; the observed high P-value reinforce the uncertainty of the results with most of them near 0.5. The null hypothesis of no relationships between metadata and phyla abundance could thus not be rejected with high confidence. Further research is required to analyse any correlations between metadata and the abundance of a microbial community as explained in the Discussion part.

pH

Phylum	*Estimated effect of 1 increase in pH on abundance*	**P value**
Crenarchaeota	-0,23%	0.239
Acidobacteria	0,45%	0.589
Actinobacteria	0,28%	0.778
Bacteroidetes	-0,26%	0.741
Chloroflexi	0,13%	0.552
Firmicutes	0,07%	0.416
Gemmatimonadetes	0,38%	0.14
Nitrospirae	0,21%	0.174
Proteobacteria	0,15%	0.858
Verrucomicrobia	-1,35%	0.042

Sample Depth

Phylum	*Estimated effect of 1cm increase depth on abundance*	**P value**
Crenarchaeota	0,17%	0.099
Acidobacteria	0,43%	0.304
Actinobacteria	0,12%	0.820
Bacteroidetes	-0,81%	0.063
Chloroflexi	0,08%	0.499
Firmicutes	1,12%	0.015
Gemmatimonadetes	0,34%	0.017
Nitrospirae	0,07%	0.376
Proteobacteria	-0,40%	0.358
Verrucomicrobia	0,02%	0.937

Figure 5*:* Estimated effects of increasing either pH or Sample depth on microbial abundance

Moisture

Phylum	Dry	Moist	P-value
Crenarchaeota			0,607
Acidobacteria			0,718
Actinobacteria		✓	0,619
Bacteroidetes			0,103
Chloroflexi	✓		0,578
Firmicutes		✓	0,614
Gemmatimonadetes		✓	0,414
Nitrospirae	✓		0,389
Proteobacteria	✓		0,856
Verrucomicrobia	✓		0,645

Potassium

Phylum	High potassium	low potassium	P-value
Crenarchaeota		✓	0,343
Acidobacteria	✓		0,551
Actinobacteria	✓		0,498
Bacteroidetes		✓	0,733
Chloroflexi	✓		0,694
Firmicutes	✓		0,186
Gemmatimonadetes	✓		0,499
Nitrospirae	✓		0,773
Proteobacteria		✓	0,558
Verrucomicrobia		✓	0,583

Nitrogen

Phylum	high nitrogen	low nitrogen	P-value
Crenarchaeota			0,042
Acidobacteria		✓	0,772
Actinobacteria			0,371
Bacteroidetes	✓		0,810
Chloroflexi		✓	0,810
Firmicutes			0,107
Gemmatimonadetes	✓		0,191
Nitrospirae		✓	0,547
Proteobacteria	✓		0,391
Verrucomicrobia		✓	0,746

Phosphorus

Phylum	High phosphorus	low phosphorus	P-value
Crenarchaeota			0,175
Acidobacteria		✓	0,434
Actinobacteria		✓	0,734
Bacteroidetes	✓		0,333
Chloroflexi		✓	0,473
Firmicutes			0,497
Gemmatimonadetes			0,585
Nitrospirae		✓	0,482
Proteobacteria	✓		0,449
Verrucomicrobia		✓	0,886

Figure 6: Moisture, Potassium, Nitrogen and Phosphorus qualitative analysis for each individual phylum. Each "tick" shows a preference for each phylum to belong to a metadata level.

Discussion

The model described gives an insight on how changes in metadata impact the abundance of a specific phylum. Our data demonstrated a relatively pronounced shift in abundance for individual phylum such as *Verrucomicrobia*, *Acidobacteria* when increasing pH, but only small alterations for most of the other microbial groups in concordance with previously EMP conducted studies. **[9]** Similarly, variations observed in microbial taxonomy abundance across depth was supported by Kathryn G. Eilers et al who pointed out variation in terms of community composition depending on the depth. *Bacteroidetes* abundance was shown to decline with depth; strengthening our -0.81% decrease per centimetre in our study confirming the correlation between microbial community and sample depth. **[10]**

However, regarding the levels of mineral resources for microbial growth, potassium, nitrogen and phosphorus it has not be possible to draw an interesting conclusion as microbial phyla abundance seems to vary randomly in relation to ions concentrations. Noah Fierer et al, similarly tried to divide bacterial phyla into an ecological manner through measurement of carbon mineralization rate essential for growth and cellular respiration, however with higher confidence (P values < 0.001) **[11]**

The reliability of these results is still very low considering the very high P-value and the drawn conclusion should be carefully discussed. Various reasons are associated with the poor quality of our data. First of all, the 30 samples are very clustered in terms of microbial richness as shown on *Figure 1* not considering the outliers, this is due to the clustered geographical locations of our samples. It would be therefore interesting to conduct an experiment with a larger amount of samples from distant locations to increase the variance of the chemical and biological gradients. Furthermore, with the use of replicates the risk of false positive or negatives could be decreased ensuring the integrity of our data to be in agreement with EMP studies where multiple samples on the same location are collected and then merged as replicates.

Moreover, the quality of the kits used to measure the chemical and biological characteristics of each sample could be perform with more appropriate quantitative measurements methods to have a higher accuracy.

Finally, it would be interesting to deepen the taxonomy analysis if we had more time by analysing the ecological environment of individual microbial species at higher level of classification. Thus by expanding the metagenomics studies, this research could potentially contribute to the growing EMP catalogue and span gradients of chemistry and biology for individual phylum.

In conclusion, while the methods used to extract DNA followed by sequencing were performed at very high throughput, the ecological characteristics of the samples were carried out with low quality procedures not providing significant scientific results regarding to this research. Additional metagenomics studies will be needed to predict with higher confidence the microbial environment of a precise sample based on its ecological properties.

References

1. Hugenholtz, P. & Tyson, G.W., 2008. Metagenomics. *Nature*, 455(7212), pp.481–483.

2. Garrido-Cardenas, J.A. & Manzano-Agugliaro, F., 2017. The metagenomics worldwide research. *Current Genetics*, 63(5), pp.819–829.

3. Thompson, L.R. et al., 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, 551(7681), p.457.

4. Anon, EMP - Earth Microbiome Project. Available at: https://qiita.ucsd.edu/emp/ [Accessed May 6, 2018].

5. Quast, C. et al., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1), pp. D590–D596.

6. DeSantis, T.Z. et al., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology*, 72(7), pp.5069–72.

7. Anonymous. QIIME scripts Available at: https://github.com/Biotech97/BIOC3301.git

8. Anon, Title stata.com regress — Linear regression. Available at: https://www.stata.com/manuals/rregress.pdf

9. Bartram, A.K. et al., 2014. Exploring links between pH and bacterial community composition in soils from the Craibstone Experimental Farm. *FEMS Microbiology Ecology*, 87(2), pp.403–415.

10. Eilers, K.G. et al., 2012. Digging deeper to find unique microbial communities: The strong effect of depth on the structure of bacterial and archaeal communities in soil. *Soil Biology and Biochemistry*, 50, pp.58–65.

11. Fierer, N., Bradford, M.A. & Jackson, R.B., 2007. Toward an ecological classification of soil bacteria. *Ecology*, 88(6), pp.1354–1364.

Appendix

Source	SS	df	MS
Model	.005823484	17	.000342558
Residual	.001281234	12	.000106769
Total	.007104717	29	.00024499

Number of obs =	30
F(17, 12) =	3.21
Prob > F =	0.0227
R-squared =	0.8197
Adj R-squared =	0.5642
Root MSE =	.01033

Abundance1	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	-.0023113	.0018668	-1.24	0.239	-.0063787	.001756
SampleDepth	.0016842	.0009412	1.79	0.099	-.0003665	.0037348
Dry	.0073469	.0129435	0.57	0.581	-.0208546	.0355484
Moist	.0062804	.0127934	0.49	0.632	-.021594	.0341549
Low	0	(omitted)				
mediumk	-.0133047	.0135434	-0.98	0.345	-.0428132	.0162037
mediumhighk	-.0491743	.0237902	-2.07	0.061	-.1010087	.0026601
lowk	-.0118917	.0119078	-1.00	0.338	-.0378367	.0140532
tracek	-.0064016	.012867	-0.50	0.628	-.0344364	.0216333
tracelowk	0	(omitted)				
mediumnitrogen	-.0527162	.0218799	-2.41	0.033	-.1003885	-.005044
lownitrogen	-.0483057	.020825	-2.32	0.039	-.0936796	-.0029319
tracenitrogen	-.0430145	.0219673	-1.96	0.074	-.0908771	.0048481
tracelownitrogen	0	(omitted)				
highphosphorus	-.0654008	.0176108	-3.71	0.003	-.1037715	-.0270302
mediumphosphorus	-.04994	.0148346	-3.37	0.006	-.0822619	-.0176181
lowphosphorus	-.0462174	.015968	-2.89	0.013	-.0810086	-.0114262
highn	0	(omitted)				
lowmediumn	-.0586859	.0225176	-2.61	0.023	-.1077475	-.0096244
lowmediumpho	-.0031514	.015527	-0.20	0.843	-.0369818	.0306791
tracelowpho	-.0627369	.0206534	-3.04	0.010	-.1077367	-.017737
tracepho	0	(omitted)				
_cons	.1180069	.0292254	4.04	0.002	.0543302	.1816836

Source	SS	df	MS
Model	.024281651	17	.001428332
Residual	.022427776	12	.001868981
Total	.046709427	29	.00161067

Number of obs =	30
F(17, 12) =	0.76
Prob > F =	0.7018
R-squared =	0.5198
Adj R-squared =	-0.1604
Root MSE =	.04323

Abundance4	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	-.0026413	.0078104	-0.34	0.741	-.0196586	.0143761
SampleDepth	-.0080741	.0039378	-2.05	0.063	-.0166538	.0005056
Dry	-.0953012	.0541541	-1.76	0.104	-.2132929	.0226905
Moist	-.0949943	.053526	-1.77	0.101	-.2116175	.0216289
Low	0	(omitted)				
mediumk	-.0248738	.0566638	-0.44	0.668	-.1483337	.0985861
mediumhighk	-.0316659	.0995353	-0.32	0.756	-.2485348	.1852029
lowk	.0213484	.049821	0.43	0.676	-.0872022	.1298989
tracek	-.0116342	.053834	-0.22	0.833	-.1289284	.1056601
tracelowk	0	(omitted)				
mediumnitrogen	.0102096	.091543	0.11	0.913	-.1892454	.2096645
lownitrogen	-.0317891	.0871294	-0.36	0.722	-.2216278	.1580496
tracenitrogen	-.0296554	.0919084	-0.32	0.753	-.2299067	.1705959
tracelownitrogen	0	(omitted)				
highphosphorus	.132362	.0736814	1.80	0.098	-.0281761	.2929
mediumphosphorus	.1174762	.0620663	1.89	0.083	-.0177546	.252707
lowphosphorus	.1239703	.0668081	1.86	0.088	-.0215919	.2695326
highn	0	(omitted)				
lowmediumn	-.037048	.0942107	-0.39	0.701	-.2423154	.1682195
lowmediumpho	.0423815	.0649631	0.65	0.526	-.099161	.183924
tracelowpho	-.0142285	.0864113	-0.16	0.872	-.2025024	.1740454
tracepho	0	(omitted)				
_cons	.1969861	.1222755	1.61	0.133	-.0694294	.4634015

Source	SS	df	MS
Model	.010964844	17	.000644991
Residual	.023280913	12	.001940076
Total	.034245756	29	.001180888

Number of obs =	30
F(17, 12) =	0.33
Prob > F =	0.9810
R-squared =	0.3202
Adj R-squared =	-0.6429
Root MSE =	.04405

Abundance2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	.0044545	.0079575	0.56	0.586	-.0128834	.0217925
SampleDepth	.0043074	.004012	1.07	0.304	-.004434	.0130487
Dry	.0206225	.0551745	0.37	0.715	-.0995924	.1408374
Moist	.0199801	.0545346	0.37	0.720	-.0988405	.1388007
Low	0	(omitted)				
mediumk	.0486263	.0577315	0.84	0.416	-.0771598	.1744124
mediumhighk	.0732442	.1014108	0.72	0.484	-.1477109	.2941993
lowk	.0160091	.0507597	0.32	0.758	-.0945868	.126605
tracek	.0341201	.0548484	0.62	0.546	-.0853843	.1536244
tracelowk	0	(omitted)				
mediumnitrogen	-.0052496	.0932678	-0.06	0.956	-.2084627	.1979635
lownitrogen	.0410029	.0887711	0.46	0.652	-.1524127	.2344185
tracenitrogen	.0364771	.0936402	0.39	0.704	-.1675474	.2405015
tracelownitrogen	0	(omitted)				
highphosphorus	-.0817629	.0750698	-1.09	0.297	-.2453258	.0818001
mediumphosphorus	-.0712317	.0632357	-1.13	0.282	-.2090106	.0665471
lowphosphorus	-.0953688	.0680669	-1.40	0.187	-.2436737	.0529362
highn	0	(omitted)				
lowmediumn	.0279122	.0959858	0.29	0.776	-.1812229	.2370474
lowmediumpho	-.0515978	.0661872	-0.78	0.451	-.1958072	.0926116
tracelowpho	.0054897	.0880394	0.06	0.951	-.1863317	.1973111
tracepho	0	(omitted)				
_cons	.0939019	.1245794	0.75	0.466	-.1775334	.3653372

Source	SS	df	MS
Model	.002828827	17	.000166402
Residual	.001701444	12	.000141787
Total	.004530272	29	.000156216

Number of obs =	30
F(17, 12) =	1.17
Prob > F =	0.3964
R-squared =	0.6244
Adj R-squared =	0.0924
Root MSE =	.01191

Abundance5	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	.0013149	.0021512	0.61	0.552	-.0033722	.0060021
SampleDepth	.0007557	.0010846	0.70	0.499	-.0016074	.0031189
Dry	.0091895	.0149158	0.62	0.549	-.0233093	.0416883
Moist	.0078048	.0147428	0.53	0.606	-.024317	.0399267
Low	0	(omitted)				
mediumk	.0102684	.0156071	0.66	0.523	-.0237365	.0442733
mediumhighk	.0062097	.0274153	0.23	0.825	-.0535231	.0659425
lowk	-.0109979	.0137223	-0.80	0.438	-.0408963	.0189005
tracek	-.0001638	.0148277	-0.01	0.991	-.0324705	.0321429
tracelowk	0	(omitted)				
mediumnitrogen	-.0023837	.0252139	-0.09	0.926	-.0573201	.0525528
lownitrogen	.0111126	.0239983	0.46	0.652	-.0411752	.0634004
tracenitrogen	.0053355	.0253146	0.21	0.837	-.0498203	.0604913
tracelownitrogen	0	(omitted)				
highphosphorus	-.0241346	.0202943	-1.19	0.257	-.0683521	.0200829
mediumphosphorus	-.0107418	.0170951	-0.63	0.542	-.0479888	.0265052
lowphosphorus	-.010645	.0184011	-0.58	0.574	-.0507376	.0294476
highn	0	(omitted)				
lowmediumn	-.0058867	.0259487	-0.23	0.824	-.0624241	.0506507
lowmediumpho	.0307679	.017893	1.72	0.111	-.0082175	.0697534
tracelowpho	.0036833	.0238005	0.15	0.880	-.0481736	.0555401
tracepho	0	(omitted)				
_cons	.0146835	.0336787	0.44	0.671	-.0586961	.0880631

Source	SS	df	MS
Model	.013536361	17	.000796257
Residual	.035677608	12	.002973134
Total	.049213969	29	.001697033

Number of obs =	30
F(17, 12) =	0.27
Prob > F =	0.9932
R-squared =	0.2751
Adj R-squared =	-0.7520
Root MSE =	.05453

Abundance3	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	.0028391	.0098509	0.29	0.778	-.0186241	.0243024
SampleDepth	.0011533	.0049666	0.23	0.820	-.0096679	.0119745
Dry	.0292145	.0683025	0.43	0.676	-.1196038	.1780328
Moist	.0403036	.0675103	0.60	0.562	-.1067886	.1873958
Low	0	(omitted)				
mediumk	-.0136946	.0714678	-0.19	0.851	-.1694097	.1420205
mediumhighk	.1278386	.12554	1.02	0.329	-.1456895	.4013666
lowk	-.0661932	.0628372	-1.05	0.313	-.2031038	.0707173
tracek	-.0475069	.0678987	-0.70	0.497	-.1954455	.1004318
tracelowk	0	(omitted)				
mediumnitrogen	.1040391	.1154595	0.90	0.385	-.1475256	.3556037
lownitrogen	.1031641	.1098929	0.94	0.366	-.1362719	.3426001
tracenitrogen	.1233438	.1159205	1.06	0.308	-.1292252	.3759129
tracelownitrogen	0	(omitted)				
highphosphorus	-.0619955	.0929315	-0.67	0.517	-.2644758	.1404848
mediumphosphorus	-.0281519	.0782817	-0.36	0.725	-.1987131	.1424094
lowphosphorus	-.0183944	.0842624	-0.22	0.831	-.2019863	.1651975
highn	0	(omitted)				
lowmediumn	.0978867	.1188242	0.82	0.426	-.161009	.3567824
lowmediumpho	-.0052123	.0819354	-0.06	0.950	-.1837342	.1733096
tracelowpho	.0509059	.1089871	0.47	0.649	-.1865566	.2883684
tracepho	0	(omitted)				
_cons	.0251304	.1542213	0.16	0.873	-.3108889	.3611496

Source	SS	df	MS
Model	.000793881	17	.000046699
Residual	.00022374	12	.000018645
Total	.001017622	29	.00003509

Number of obs =	30
F(17, 12) =	2.50
Prob > F =	0.0555
R-squared =	0.7801
Adj R-squared =	0.4687
Root MSE =	.00432

Abundance6	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	.0006572	.0007801	0.84	0.416	-.0010425	.0023568
SampleDepth	.0011209	.0003933	2.85	0.015	.000264	.0019778
Dry	.0014648	.0054089	0.27	0.791	-.0103202	.0132498
Moist	.0043065	.0053462	0.81	0.436	-.0073419	.0159548
Low	0	(omitted)				
mediumk	-.0091968	.0056596	-1.62	0.130	-.021528	.0031344
mediumhighk	-.0067725	.0099416	-0.68	0.509	-.0284334	.0148884
lowk	-.012479	.0049761	-2.51	0.028	-.023321	-.0016369
tracek	-.0104921	.0053769	-1.95	0.075	-.0222075	.0012232
tracelowk	0	(omitted)				
mediumnitrogen	.0216922	.0091433	2.37	0.035	.0017706	.0416138
lownitrogen	.0114571	.0087025	1.32	0.213	-.007504	.0304182
tracenitrogen	.0198347	.0091798	2.16	0.052	-.0001664	.0398359
tracelownitrogen	0	(omitted)				
highphosphorus	.0019373	.0073593	0.26	0.797	-.0140973	.0179718
mediumphosphorus	-.0100915	.0061992	-1.63	0.130	-.0235983	.0034154
lowphosphorus	-.0082121	.0066728	-1.23	0.242	-.0227509	.0063266
highn	0	(omitted)				
lowmediumn	.0153738	.0094098	1.63	0.128	-.0051283	.035876
lowmediumpho	-.0026636	.0064885	-0.41	0.689	-.0168009	.0114736
tracelowpho	.0042875	.0086308	0.50	0.628	-.0145173	.0230923
tracepho	0	(omitted)				
_cons	-.0084469	.0122129	-0.69	0.502	-.0350565	.0181627

Source	SS	df	MS
Model	.00670422	17	.000394366
Residual	.002174283	12	.00018119
Total	.008878502	29	.000306155

Number of obs =	30
F(17, 12) =	2.18
Prob > F =	0.0875
R-squared =	0.7551
Adj R-squared =	0.4082
Root MSE =	.01346

Abundance7	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	.0038465	.0024318	1.58	0.140	-.001452	.0091451
SampleDepth	.0034096	.0012261	2.78	0.017	.0007382	.006081
Dry	.0107384	.0168615	0.64	0.536	-.0259997	.0474764
Moist	.0184231	.0166659	1.11	0.291	-.0178889	.0547351
Low	0	(omitted)				
mediumk	-.0073077	.0176429	-0.41	0.686	-.0457484	.0311329
mediumhighk	.0069916	.0309915	0.23	0.825	-.060533	.0745162
lowk	-.0242955	.0155123	-1.57	0.143	-.058094	.009503
tracek	-.0165356	.0167618	-0.99	0.343	-.0530565	.0199853
tracelowk	0	(omitted)				
mediumnitrogen	.0514943	.028503	1.81	0.096	-.0106083	.1135969
lownitrogen	.0325953	.0271287	1.20	0.253	-.0265131	.0917038
tracenitrogen	.0384486	.0286168	1.34	0.204	-.0239019	.1007992
tracelownitrogen	0	(omitted)				
highphosphorus	.0257069	.0229416	1.12	0.284	-.0242784	.0756923
mediumphosphorus	-.0061709	.0193251	-0.32	0.755	-.0482766	.0359347
lowphosphorus	-.0007923	.0208015	-0.04	0.970	-.0461148	.0445302
highn	0	(omitted)				
lowmediumn	.0386578	.0293336	1.32	0.212	-.0252546	.1025702
lowmediumpho	.0223953	.020227	1.11	0.290	-.0216756	.0664662
tracelowpho	.0135226	.0269051	0.50	0.624	-.0450987	.0721438
tracepho	0	(omitted)				
_cons	-.0659755	.0380719	-1.73	0.109	-.148927	.016976

Source	SS	df	MS
Model	.037065376	17	.002180316
Residual	.025605669	12	.002133806
Total	.062671045	29	.002161071

Number of obs =	30
F(17, 12) =	1.02
Prob > F =	0.4965
R-squared =	0.5914
Adj R-squared =	0.0126
Root MSE =	.04619

Abundance9	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	.0015296	.0083454	0.18	0.858	-.0166534	.0197126
SampleDepth	-.0040191	.0042075	-0.96	0.358	-.0131864	.0051483
Dry	-.0037836	.0578637	-0.07	0.949	-.1298579	.1222907
Moist	-.0176596	.0571926	-0.31	0.763	-.1422716	.1069524
Low	0	(omitted)				
mediumk	.0271783	.0605454	0.45	0.662	-.1047388	.1590953
mediumhighk	-.0758712	.1063536	-0.71	0.489	-.3075958	.1558534
lowk	.0436292	.0532338	0.82	0.428	-.0723572	.1596156
tracek	.0264379	.0575217	0.46	0.654	-.0988912	.1517669
tracelowk	0	(omitted)				
mediumnitrogen	-.0518032	.0978137	-0.53	0.606	-.2649211	.1613146
lownitrogen	-.0887315	.0930979	-0.95	0.359	-.2915744	.1141113
tracenitrogen	-.1156115	.0982043	-1.18	0.262	-.3295802	.0983572
tracelownitrogen	0	(omitted)				
highphosphorus	.1350056	.0787287	1.71	0.112	-.0365295	.3065407
mediumphosphorus	.0996383	.0663179	1.50	0.159	-.044856	.2441326
lowphosphorus	.0837284	.0713845	1.17	0.264	-.071805	.2392618
highn	0	(omitted)				
lowmediumn	-.1007124	.1006642	-1.00	0.337	-.3200409	.1186161
lowmediumpho	-.0236205	.0694132	-0.34	0.740	-.1748588	.1276178
tracelowpho	-.0032876	.0923305	-0.04	0.972	-.2044585	.1978833
tracepho	0	(omitted)				
_cons	.3051704	.1306515	2.34	0.038	.0205052	.5898356

Source	SS	df	MS
Model	.000941659	17	.000055392
Residual	.000808588	12	.000067382
Total	.001750247	29	.000060353

Number of obs =	30
F(17, 12) =	0.82
Prob > F =	0.6531
R-squared =	0.5380
Adj R-squared =	-0.1165
Root MSE =	.00821

Abundance8	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	.0021444	.001483	1.45	0.174	-.0010867	.0053756
SampleDepth	.000688	.0007477	0.92	0.376	-.0009411	.002317
Dry	.0071442	.0102826	0.69	0.500	-.0152596	.029548
Moist	.0115686	.0101633	1.14	0.277	-.0105754	.0337126
Low	0	(omitted)				
mediumk	.0031414	.0107591	0.29	0.775	-.0203007	.0265835
mediumhighk	.0048835	.0188994	0.26	0.800	-.0362947	.0460618
lowk	-.0055717	.0094598	-0.59	0.567	-.0261829	.0150395
tracek	-.0006854	.0102218	-0.07	0.948	-.0229568	.021586
tracelowk	0	(omitted)				
mediumnitrogen	-.0152583	.0173818	-0.88	0.397	-.05313	.0226135
lownitrogen	-.0077079	.0165438	-0.47	0.650	-.0437537	.028338
tracenitrogen	-.0092441	.0174512	-0.53	0.606	-.047267	.0287789
tracelownitrogen	0	(omitted)				
highphosphorus	-.0188505	.0139904	-1.35	0.203	-.0493329	.0116319
mediumphosphorus	-.0114387	.0117849	-0.97	0.351	-.0371158	.0142384
lowphosphorus	-.0135258	.0126853	-1.07	0.307	-.0411646	.014113
highn	0	(omitted)				
lowmediumn	-.0114898	.0178884	-0.64	0.533	-.0504652	.0274856
lowmediumpho	-.0092293	.012335	-0.75	0.469	-.0361049	.0176462
tracelowpho	-.0073956	.0164074	-0.45	0.660	-.0431443	.0283532
tracepho	0	(omitted)				
_cons	.0079626	.0232172	0.34	0.738	-.0426234	.0585486

Source	SS	df	MS
Model	.031339196	17	.001843482
Residual	.01287706	12	.001073088
Total	.044216257	29	.001524699

Number of obs =	30
F(17, 12) =	1.72
Prob > F =	0.1721
R-squared =	0.7088
Adj R-squared =	0.2962
Root MSE =	.03276

Abundance10	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pH	-.0134738	.0059182	-2.28	0.042	-.0263684	-.0005793
SampleDepth	.0002426	.0029838	0.08	0.937	-.0062585	.0067437
Dry	.0266173	.0410343	0.65	0.529	-.0627887	.1160233
Moist	.0127005	.0405583	0.31	0.760	-.0756685	.1010696
Low	0	(omitted)				
mediumk	-.0528661	.042936	-1.23	0.242	-.1464155	.0406833
mediumhighk	-.0523926	.075421	-0.69	0.500	-.2167209	.1119357
lowk	.0165189	.0377509	0.44	0.669	-.0657333	.098771
tracek	.0041906	.0407917	0.10	0.920	-.084687	.0930681
tracelowk	0	(omitted)				
mediumnitrogen	-.0102124	.069365	-0.15	0.885	-.1613457	.1409208
lownitrogen	-.0026531	.0660207	-0.04	0.969	-.1464998	.1411936
tracenitrogen	.012214	.0696419	0.18	0.864	-.1395227	.1639507
tracelownitrogen	0	(omitted)				
highphosphorus	-.0033416	.0558307	-0.06	0.953	-.1249863	.1183031
mediumphosphorus	-.0099697	.0470296	-0.21	0.836	-.1124383	.092499
lowphosphorus	-.0014833	.0506226	-0.03	0.977	-.1117804	.1088138
highn	0	(omitted)				
lowmediumn	.0830226	.0713864	1.16	0.267	-.072515	.2385602
lowmediumpho	.0083726	.0492246	0.17	0.868	-.0988786	.1156238
tracelowpho	.0171712	.0654765	0.26	0.798	-.1254898	.1598323
tracepho	0	(omitted)				
_cons	.1904458	.092652	2.06	0.062	-.0114255	.3923172