**Broad-scale controls of root carbon to soil carbon**

**Downloaded from Google Docs on 2-26-20 by JM**

List of contributors (to be confirmed; order to be refined)

Avni Malhotra, Samantha Weintraub, Jessica Moore, Katerina Georgiou, Stuart Grandy, Asmeret Asefaw-Berhe, Marie-Anne de Graaff, Emily Kyker-Snowman, Fraterrigo, Jennifer M, Sharon Billings, Courtney Meier, Kate Lajtha,Will Wieder

**Introduction**

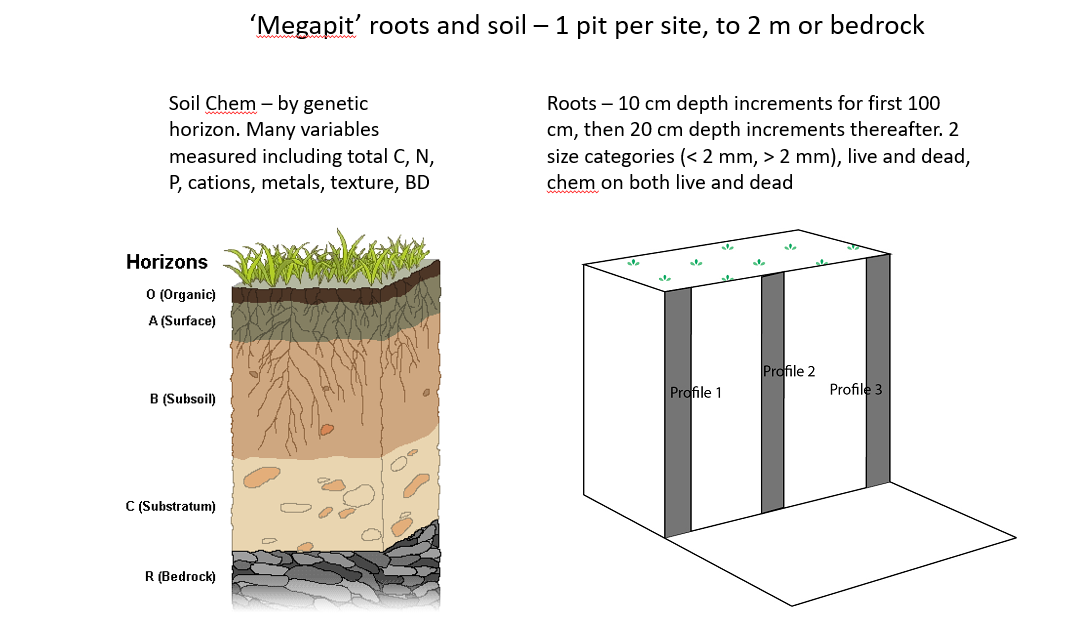
* Soils store vast stocks of carbon that are vulnerable to global change
* Plants also respond to global change and a key unknown is the belowground response
* We know that SOC is generally related to vegetation (Jobaggy and Jackson 2000). More recently we also know that mycorrhizae and SOC are spatially correlated (<https://rdcu.be/bWJ5T>). Mechanisms of these linkages are poorly understood.
* It is postulated that root-derived carbon is more likely to be stabilized as long-term soil carbon (Sokol papers, Jackson et al. 2017) but only demonstrated in lab or ecosystem specific settings.
  + This relationship likely depends on the ecoregion because decomp can be limited by moisture, temperature, etc.
  + It also likely depends on depth - given the dominance of roots at shallower depths.
* However, this notion remains untested across broad scales and beyond grassland ecosystems.
* Here, we leverage paired measurements of root and soil properties across a large climate, vegetation and soil gradient.
* When does the 1:1 soil-root relationship fall apart?
  + Arid systems
  + Permafrost
  + microbes?!

**Research Questions**

* **Q1: Is SOC a function of root C over broad spatial scales, and can we predict it based on roots + covariates from the conceptual framework.**
  + **Q1aTotal stocks (summed across the whole profile)**
    - SOC ~ bgb + MAT + MAP + nutrients + %clay + myc type + landcover + (1| layer\_bot)
    - Notes
      * layer\_bot is a random effect because sites differ in their max sampling depth. Alternatively, we could standardize the SOC and bgb values by max depth. Layer\_bot ranges from 46 to 213 cm, with most sites ~200 cm
      * Sites TOOL, PUUM, and KONA were removed because they have all NA bgb measurements; 44 NEON sites were used in this analysis
  + **Q1b: By horizon (Organic vs mineral + more detailed horizons + depth)**
    - Hypothesize that O is more variable than mineral horizon
    - Hypothesize that more shallow depths have significant SOC~bgb than deeper soils
    - SOC ~ bgb + (horizon or depth interval) + MAT + MAP + nutrients + %clay + myc type + landcover
  + **Q1c: Surface soils only ( + nutnet + periodic + cedar creek)**
* **Q2: Does root carbon and soil carbon relate within vertical profiles (vertical distribution determined via beta)?**
  + beta(SOC) ~ beta(bgb) + MAT + MAP + AET + nutrients + %clay + myc type + landcover + nutrient profiles
  + Bonus question: Do CLM beta SOC and beta roots show similar trends?
    - Outputs for both from CLM
* **Which covariates help explain the variability in SOC~f(bgb)**
  + Climate
  + Landcover/veg/biome
  + Nutrient availability
  + Soil type/texture
  + Mycorrhizal type
  + Aboveground biomass

**Methods**

We used NEON megapit data for all questions, and NEON periodic and Nutnet data for surface soil dynamics. [Belowground data list](https://docs.google.com/spreadsheets/d/1HaqsKCO6GUZ4rdRuMHydAgcsBKTUkZETbCFlWrhkp_Q/edit#gid=0)



Statistical model

Linear mixed models explaining SOC:

* 1. Q1: SOC ~ bgb + MAT + MAP + nutrients + %clay + myc type + landcover
  2. Q2:
  3. Where beta is calculated using the following:

Y = 1-(beta)d

Where,

d is the bottom depth of the layer

Y is the cumulative fraction of roots at that depth

**need to fix the following line numbers once we figure out landcover:**

**154, 161, 167, 197, 206, 261**

**also for lines 67 and 68 we were unable to run with the new tarball**

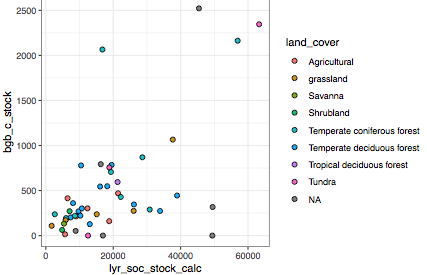
# **Results**

Fig. 1: Relationship between total stocks of root C and stocks of soil C (line 156)

Model

Random effects: layer\_bot max

Fixed effects: bgb, land\_cover, mat, map



**Stats run on Nov 20**

full.mod<-lmer(data=somNEONMegaSoilRoot\_wholeprofilestats, lyr\_soc\_stock\_calc\_sum ~

bgb\_c\_stock\_sum + mat + map + clay + (1|layer\_bot\_max))

Random effects:

Groups Name Variance Std.Dev.

layer\_bot\_max (Intercept) 160240912 12659

Residual 66834581 8175

Number of obs: 47, groups: layer\_bot\_max, 13

Fixed effects:

Estimate Std. Error df t value Pr(>|t|)

(Intercept) 17422.424 5069.817 21.487 3.436 0.00242 \*\*

bgb\_c\_stock\_sum 9.755 3.379 41.988 2.887 0.00612 \*\*

mat -664.784 242.408 41.009 -2.742 0.00900 \*\*

map 2.346 3.030 41.121 0.774 0.44314

clay 347.337 103.798 29.785 3.346 0.00223 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> r2(full.mod)

# R2 for mixed models

Conditional R2: 0.786

Marginal R2: 0.272

> vif(full.mod)

bgb\_c\_stock\_sum mat map clay

1.538840 1.731726 1.429055 1.028896

Reduced model results

Formula: lyr\_soc\_stock\_calc\_sum ~ bgb\_c\_stock\_sum + mat + clay + (1 | layer\_bot\_max)

Data: somNEONMegaSoilRoot\_wholeprofilestats

Random effects:

Groups Name Variance Std.Dev.

layer\_bot\_max (Intercept) 185871351 13633

Residual 63195802 7950

Number of obs: 47, groups: layer\_bot\_max, 13

Fixed effects:

Estimate Std. Error df t value Pr(>|t|)

(Intercept) 18028.908 5143.299 20.977 3.505 0.002108 \*\*

bgb\_c\_stock\_sum 10.999 2.997 40.253 3.670 0.000703 \*\*\*

mat -569.980 204.637 39.007 -2.785 0.008211 \*\*

clay 350.219 101.112 31.741 3.464 0.001547 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: lyr\_soc\_stock\_calc\_sum

Chisq Df Pr(>Chisq)

bgb\_c\_stock\_sum 13.471 1 0.0002422 \*\*\*

mat 7.758 1 0.0053475 \*\*

clay 11.997 1 0.0005328 \*\*\*

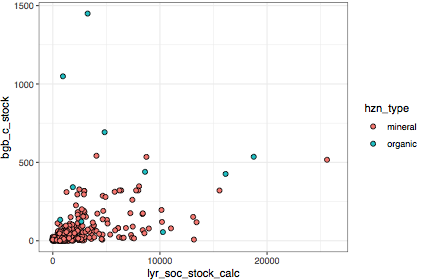
Conditional R2: 0.812

Marginal R2: 0.260

**Stats run by JM on 2-26-20**

|  |  |  |
| --- | --- | --- |
| Model | AIC | Marginal Psuedo-R2 |
| SOC ~ |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Fig 2: Horizon-dependent relationships



Data were analyzed separately for organic and mineral horizons.

Organic horizon results:

Formula: lyr\_soc\_stock\_calc\_sum ~ bgb\_c\_stock\_sum + (1 | layer\_bot\_max)

Data: somNEON\_organic\_wholeprofile

Random effects:

Groups Name Variance Std.Dev.

layer\_bot\_max (Intercept) 53249972 7297

Residual 954576 977

Number of obs: 19, groups: layer\_bot\_max, 12

Fixed effects:

Estimate Std. Error df t value Pr(>|t|)

(Intercept) 2388.576 2584.207 9.921 0.924 0.37727

bgb\_c\_stock\_sum 8.891 2.770 9.976 3.209 0.00937 \*\*

Conditional R2: 0.989

Marginal R2: 0.402

Mineral horizon results:

Formula: lyr\_soc\_stock\_calc\_sum ~ bgb\_c\_stock\_sum + mat + clay + (1 | layer\_bot\_max)

Data: somNEON\_mineral\_wholeprofile

Random effects:

Groups Name Variance Std.Dev.

layer\_bot\_max (Intercept) 115543380 10749

Residual 64121254 8008

Number of obs: 47, groups: layer\_bot\_max, 15

Fixed effects:

Estimate Std. Error df t value Pr(>|t|)

(Intercept) 14383.210 4239.862 27.283 3.392 0.00213 \*\*

bgb\_c\_stock\_sum 15.908 5.685 37.170 2.798 0.00810 \*\*

mat -406.685 190.393 41.516 -2.136 0.03862 \*

clay 288.438 102.320 31.752 2.819 0.00823 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Analysis of Deviance Table (Type II Wald chisquare tests)

Response: lyr\_soc\_stock\_calc\_sum

Chisq Df Pr(>Chisq)

bgb\_c\_stock\_sum 7.8294 1 0.005140 \*\*

mat 4.5626 1 0.032677 \*

clay 7.9467 1 0.004818 \*\*

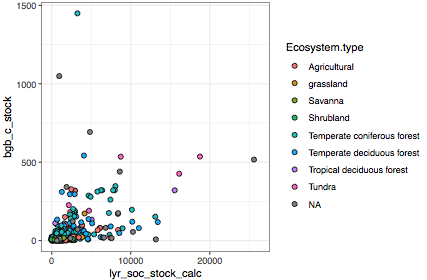
---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

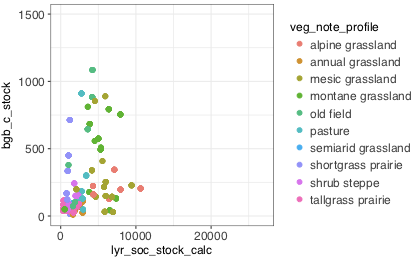
Conditional R2: 0.702

Marginal R2: 0.166

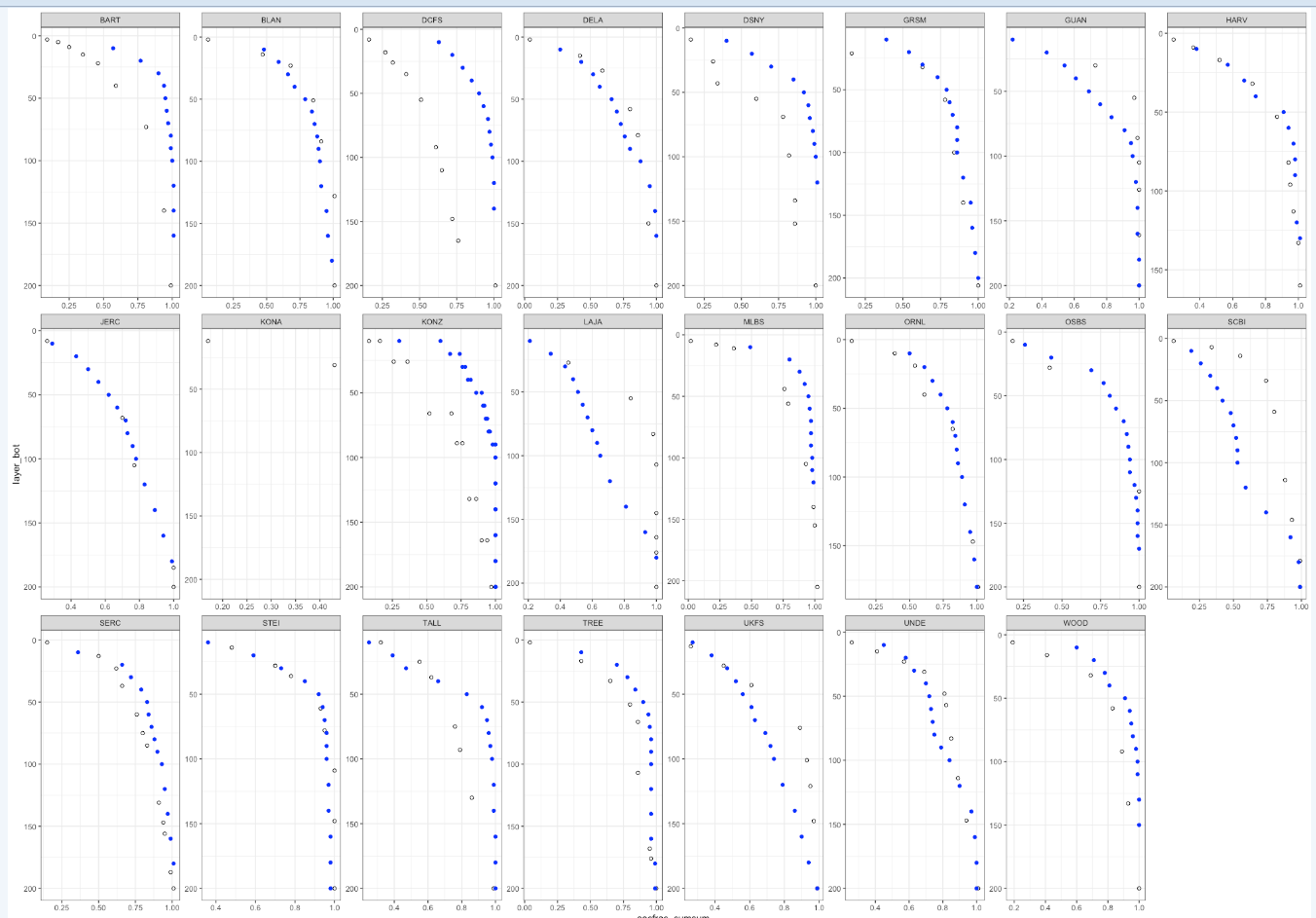
NEON layer specific data



NutNet total stock comparison



Root and soil carbon relationships by depth



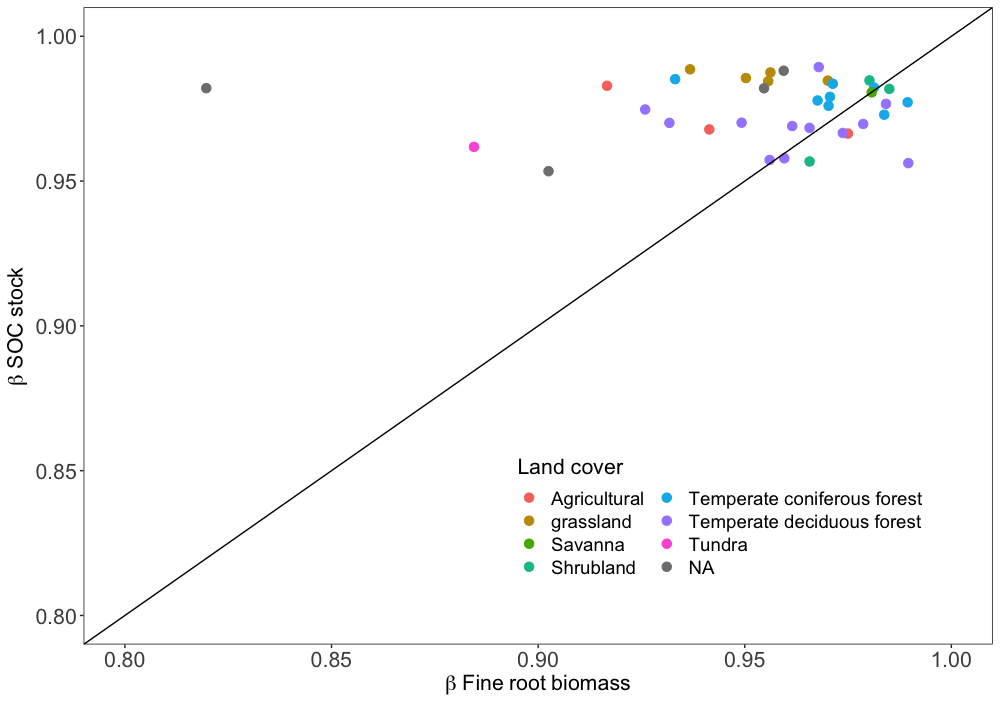


Figure X: The correlation between beta coefficients for fine root biomass (< 4 mm diam) and total (bulk) soil C stocks with depth. Beta was calculated for root biomass as fraction of total root biomass in the profile that accumulates in 10 cm depth increments. These beta curves

Interpretation: Points above the 1:1 line

Total SOC is related to bgb, but focused on depth profiles the relationship seems disconnected. Maybe because of transport, where the SOC is produced it doesn’t stay there. Maybe SOC~bgb isn’t related in o horizon or surface soils because it’s decomposing and being respired.

Could we have factors that describe macropores, earthworms, etc.

How does beta SOC (distribution of SOC with depth in the profile) vary with ecoregion, texture, nutrients, litter inputs. This has implications for global change. What is the depth where you capture 80-90% of SOC? What factors capture that?

Does ELM have a soc/root lag?

Could have a model output for root and soil beta map and scatter plot (will).

Vulnerability could be discussed using the beta curves. Shallower the carbon perhaps more vulnerable?

**These are results from the betas:**

Call:

lm(formula = beta\_soc ~ beta\_roots + land\_cover, data = beta.all)

Residuals:

Min 1Q Median 3Q Max

-0.0186864 -0.0017398 -0.0000225 0.0025458 0.0210514

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.056562 0.082350 12.830 1.69e-12 \*\*\*

beta\_roots -0.089140 0.087061 -1.024 0.3157

land\_covergrassland 0.014632 0.005981 2.446 0.0218 \*

land\_coverSavanna 0.011491 0.009887 1.162 0.2561

land\_coverShrubland 0.004978 0.007206 0.691 0.4961

land\_coverTemperate coniferous forest 0.009254 0.005959 1.553 0.1330

land\_coverTemperate deciduous forest -0.001977 0.005456 -0.362 0.7201

land\_coverTundra -0.015919 0.010720 -1.485 0.1501

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.008112 on 25 degrees of freedom

(4 observations deleted due to missingness)

Multiple R-squared: 0.4806, Adjusted R-squared: 0.3352

F-statistic: 3.305 on 7 and 25 DF, p-value: 0.01252

**Model with more covariates:**

Call:

lm(formula = beta\_soc ~ beta\_roots + land\_cover + map + mat +

clay, data = beta.all)

Residuals:

Min 1Q Median 3Q Max

-0.018323 -0.003149 0.000000 0.003189 0.017717

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.181e+00 1.043e-01 11.329 1.19e-10 \*\*\*

beta\_roots -2.254e-01 1.112e-01 -2.027 0.0549 .

land\_covergrassland 1.404e-02 6.001e-03 2.339 0.0288 \*

land\_coverSavanna 1.375e-02 1.022e-02 1.345 0.1924

land\_coverShrubland 6.321e-03 7.299e-03 0.866 0.3958

land\_coverTemperate coniferous forest 1.693e-02 7.966e-03 2.125 0.0451 \*

land\_coverTemperate deciduous forest 2.511e-03 6.428e-03 0.391 0.6998

land\_coverTundra -1.469e-02 1.055e-02 -1.393 0.1775

map -7.037e-06 4.879e-06 -1.442 0.1634

mat 5.672e-04 3.407e-04 1.665 0.1101

clay 1.954e-04 1.393e-04 1.402 0.1749

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.00796 on 22 degrees of freedom

(4 observations deleted due to missingness)

Multiple R-squared: 0.5599, Adjusted R-squared: 0.3598

F-statistic: 2.798 on 10 and 22 DF, p-value: 0.02118

> Anova(mod)

Anova Table (Type II tests)

Response: beta\_soc

Sum Sq Df F value Pr(>F)

beta\_roots 0.00026044 1 4.1103 0.05491 .

land\_cover 0.00139664 6 3.6737 0.01113 \*

map 0.00013177 1 2.0796 0.16337

mat 0.00017564 1 2.7720 0.11011

clay 0.00012455 1 1.9656 0.17487

Residuals 0.00139398 22

---

**Without land cover:**

Call:

lm(formula = beta\_soc ~ beta\_roots + map + mat + clay, data = beta.all)

Residuals:

Min 1Q Median 3Q Max

-0.022146 -0.006004 0.001774 0.007554 0.014923

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 9.671e-01 5.857e-02 16.513 <2e-16 \*\*\*

beta\_roots 1.065e-02 6.280e-02 0.170 0.866

map -3.454e-06 3.926e-06 -0.880 0.386

mat 9.375e-05 3.331e-04 0.281 0.780

clay 3.458e-06 1.649e-04 0.021 0.983

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0109 on 32 degrees of freedom

Multiple R-squared: 0.026, Adjusted R-squared: -0.09575

F-statistic: 0.2136 on 4 and 32 DF, p-value: 0.929

# **Discussion**

* If we can suggest new and equally simple mathematical functions that would describe the variability of profiles (for eg., something that describes dimorphic functions)
* Jessica: how about fleshing out the discussion with hypotheses of when and where we’d expect the root/soil relationships to fall apart.. And show some profile examples from NEON

Next steps 2/11/2020:

* Rereun everything with new tarball
* Try B analysis without O horizon
* Make sure ecosystem type and land cover is solid
* Can we cluster the B’s to figure out what drives the clusters?
* Add hydrological proxies- ET could explain beta? What about DOC transport?
* Discussing the disconnect:
  + Need something other than beta? Weak R 2 as a proxy for dimorhphic?
  + Vertical transport: pH (Will), PET (Kate), soil order (get hypotheses from Kate :))
  + What would priming do (Stuart)?
* QAQC issues: dead or alive? Variability in mm?
* Other covariates: bedrock, WTD, nutrient profiles? Dominant species/PFT? Ask Emily for beta N.

# AGU 2019 comments:

Hararuk asked about what then can predict the depth profiles?

KL: maybe its all DOC. surface litters signals are going to throw off curves, so try looking only below 10 or 20 cm. Also total c includes a lot of particulate unstabilized C so what about beta curves with heavy frac c (wont have data but call to action)

Verity Salmon: what if we did beta cumulative from the bottom instead of from the top [calc from 100 -> 0]

CourtneyCreamer: do you think the diff in root vs C depth dist could be a legacy thing (roots are now, C is accum over time). So 13c or 14c data might help? Similar, do you think it might have to do w hydrology (wetter areas would have more doc moving down leading to more even C depth dist). Anyway, u prob already thought of all this. Would love to hear how you're thinking of the legacy of C vs the now of roots sometime (I have a root - soil C dataset that doesn't match up w depth - veg change)

Billings

I have a lot of questions and ideas about trying to understand the discrepancy between rooting depth distributions and the depth dependence of soil organic carbon concentrations. I think that stable isotopes and, probably even more importantly, radiocarbon can help. I also think that understanding shrink swell tendencies of soils and that feature’s influence on vertical mechanisms of organic carbon distribution other than roots is probably important. The fact that Beto values for roots can go quite low with no corresponding decrease in SOC better values tells us something about SOC losses.

**SW:** It’s interesting that you mention stable isotopes. Others have also asked about this, but unfortunately it’s not something we measured as part of the ‘chemical properties’ of megapit soils. I am thinking of proposing that we measure it now on archived soil, since so many people seem interested. That said, even if I get this approved the data will not be forthcoming for some months, maybe even a year.

In terms of soil structure and other edaphic properties, we might be able to get at that with some of the other megapit data (particle size distributions, soil taxonomy), but maybe not as well as you’d like. And we certainly do not have radiocarbon measurements…

**SB:** A few thoughts have occurred to me since AGU when thinking of SOC and root depth distributions.

Sam are the NEON samples that contain carbonates acid treated prior to SOC (and perhaps eventually d13C) analyses? More broadly, do we know for sure that the SOC data from LTER sites in the database were arrived at after acid treatment if need be? It sometimes is neglected, and it’s a big deal.

Also, I’ve attached a [paper](https://www.nature.com/articles/srep45635) that highlights a process probably important for variation across sites in the relationship between depth declines in SOC vs roots: clay illuviation. Lousy title but valuable concept.

I have been working with some data from the Alps, and there we see almost linear declines with depth in roots but the SOC has a large beta.

**SW:** For NEON samples, yes we are sure that only OC is reported. In the megapit data, they (NRCS) arrive there by difference, in that they analyze total carbon, then IC, and get OC by difference. In our periodic sampling, the lab removes IC by acidification prior to EA-IRMS analyses for sites that contain carbonates, so the values we report are OC only. If we do go back and get the stable isotope data for the megaptit samples, sites will be acidified as needed prior to analysis.

For the non-NEON sites, I guess we can’t be sure but I know that there is an TC variable but also a TOC variable in the key-key so the hope is that people are cognizant and only report what they actually measured.

That’s super interesting about roots being linear but SOC having a more shallow distribution in your Alps work…if I’m getting this right, I think the NEON data shows the opposite? But maybe I’m confused…will have a look at the paper, I certainly appreciate the role of vertical translocation within the soil profile.