Summary linear regression models

Marcos Angelini

January 12, 2016

### Introduction

In our first paper we got very low accuracy in predicted soil properties. They were: organic carbon and thickness of A horizon, total bases and base saturation at A horizon, clay ratio between B and A horizons and exchange sodium percentage at A and B horizons. Some reasons of the low accuracy were: high measurement error and low correlation between covariates and soil properties, particularly with those from A horizons.

For these reasons, we want to explore if there are potential soil properties that could be predicted with reasonable accuracy. In order to test this, this document shows the summary of stepwise linear regresion models of the main soil properties: **pH**, organic carbon (**OC**), **CEC**, total bases (**tb**), **clay**, silt between 2 and 20 μm (**silt20**) and very fine sand between 50 and75 µm (**sanf.mf**).

### Soil Properties

Soil dataset is organised by horizon. id.p is the profile number, id.p.h is the profile number and original horizon and another two columns contain Latitude (Y) and Longitude (X).

## Warning in `[<-.factor`(`\*tmp\*`, p$hor == "AB|BA", value =  
## structure(c(1L, : invalid factor level, NA generated

### Covariates

Covariates are:

* dem: altitude
* slope
* twi
* vertical distance to channel network (vdchn)
* maximum curvature (maxm)
* multi resolution index valley bottom (mrvbf)
* distance to Parana river (river)
* distance to water bodies (wdist)
* water mask (water, 0 land, > 0 inside-water body distance)
* env is a factor of 3 categories estimated from EVI images
* 14-years mean EVI (evim) and standard deviation (evisd)
* Land surface temperature mean (lstm) and standard deviation (lstsd)
* Seasonal mean EVI (XX1, XX2 and XX3)

# Soil dataset.  
knitr::kable(d[1:10,1:15])

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| id.p | id.p.h | hor | top | bottom | thick | tb | CEC | phw | phkcl | resist | OC | clay | silt20 | sand.mf |
| 3 | 3\_A1 | A | 12 | 25 | 13 | 16.2 | 18.4 | 6.9 | 5.7 | 5198 | 1.80 | 17.4 | 13.3 | 39.3 |
| 3 | 3\_Ap | A | 0 | 12 | 12 | 17.0 | 20.8 | 6.4 | 5.5 | 3751 | 2.80 | 17.2 | 13.6 | 40.7 |
| 3 | 3\_B2t | B | 35 | 65 | 30 | 16.8 | 18.3 | 7.5 | 5.9 | 2286 | 0.45 | 30.2 | 7.5 | 37.5 |
| 3 | 3\_B31 | BC | 65 | 100 | 35 | NA | 15.1 | 9.1 | 7.5 | 1039 | 0.23 | 17.8 | 10.0 | 42.5 |
| 3 | 3\_B32 | BC | 100 | 130 | 30 | NA | 13.4 | 9.5 | 7.8 | 1143 | 0.05 | 15.8 | 7.1 | 46.2 |
| 3 | 3\_C | C | 130 | 150 | 20 | NA | 13.4 | 9.3 | 7.5 | 1593 | 0.02 | 12.9 | 7.9 | 47.4 |
| 3 | 3\_A2 | E | 25 | 35 | 10 | 14.6 | 14.1 | 7.2 | 5.8 | 5891 | 0.82 | 14.6 | 12.7 | 45.2 |
| 326 | 326\_Ap | A | 0 | 18 | 18 | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| 326 | 326\_A1 | A | 18 | 32 | 14 | 14.8 | 19.9 | 5.6 | 5.3 | 3673 | 1.55 | 23.0 | 24.0 | 6.6 |
| 326 | 326\_B21 | B | 40 | 80 | 40 | 25.8 | 28.1 | 6.7 | 5.3 | 2112 | 0.52 | 42.1 | 21.3 | 4.2 |

summary(unique(d[,c(21:34,83:85)]), lm=T)

## dem river wdist maxc   
## Min. : 10.03 Min. : 1326 Min. : 0.00 Min. :-161.0267   
## 1st Qu.: 45.54 1st Qu.: 41725 1st Qu.: 94.87 1st Qu.: -7.9939   
## Median : 60.48 Median : 71406 Median : 226.36 Median : -3.8412   
## Mean : 59.66 Mean : 72389 Mean : 339.24 Mean : -6.5698   
## 3rd Qu.: 75.01 3rd Qu.:105982 3rd Qu.: 480.95 3rd Qu.: -0.4944   
## Max. :102.50 Max. :144937 Max. :1946.59 Max. : 61.1544   
## mrvbf slope twi vdchn   
## Min. :0.000 Min. :0.006406 Min. : 8.038 Min. :-2.7561   
## 1st Qu.:2.726 1st Qu.:0.069177 1st Qu.:11.209 1st Qu.: 0.5595   
## Median :3.872 Median :0.122954 Median :11.964 Median : 1.7094   
## Mean :3.829 Mean :0.192647 Mean :12.511 Mean : 2.8300   
## 3rd Qu.:4.968 3rd Qu.:0.233391 3rd Qu.:13.214 3rd Qu.: 3.9953   
## Max. :7.406 Max. :1.661317 Max. :23.138 Max. :18.7957   
## water env lstm lstsd   
## Min. :0.0000 Min. :1.000 Min. :295.8 Min. :6.874   
## 1st Qu.:0.0000 1st Qu.:1.000 1st Qu.:296.5 1st Qu.:7.584   
## Median :0.0000 Median :1.000 Median :296.8 Median :7.740   
## Mean :0.1669 Mean :1.575 Mean :296.8 Mean :7.722   
## 3rd Qu.:0.0000 3rd Qu.:2.000 3rd Qu.:297.1 3rd Qu.:7.879   
## Max. :2.0000 Max. :3.000 Max. :297.9 Max. :8.421   
## evim evisd XX1 XX2   
## Min. :2510 Min. : 469.6 Min. :76.14 Min. :1582   
## 1st Qu.:3297 1st Qu.:1224.7 1st Qu.:76.36 1st Qu.:2871   
## Median :3530 Median :1627.2 Median :76.52 Median :3249   
## Mean :3545 Mean :1535.4 Mean :76.56 Mean :3122   
## 3rd Qu.:3706 3rd Qu.:1862.1 3rd Qu.:76.71 3rd Qu.:3486   
## Max. :5014 Max. :2407.6 Max. :77.35 Max. :3955   
## XX3   
## Min. :1528   
## 1st Qu.:2053   
## Median :2196   
## Mean :2219   
## 3rd Qu.:2366   
## Max. :3073

[step(lm(sand.mf ~ dem \* river + wdist + maxc + mrvbf + slope + hor + twi + vdchn + water + lstm \* lstsd + evim \* evisd + env + X \* Y + XX1 \* XX2 \* XX3, p[p$hor == "A" | p$hor == "B" | p$hor == "C", ]),direction = "both")]

### pH

summary(lm(formula = phw ~ dem + river + mrvbf + hor + twi + vdchn +   
 water + evim + evisd + X + Y + XX1 + XX2 + XX3 + dem:river +   
 evim:evisd + X:Y + XX1:XX2 + XX1:XX3 + XX2:XX3, data = p[p$hor ==   
 "A" | p$hor == "B" | p$hor == "C", ]))

##   
## Call:  
## lm(formula = phw ~ dem + river + mrvbf + hor + twi + vdchn +   
## water + evim + evisd + X + Y + XX1 + XX2 + XX3 + dem:river +   
## evim:evisd + X:Y + XX1:XX2 + XX1:XX3 + XX2:XX3, data = p[p$hor ==   
## "A" | p$hor == "B" | p$hor == "C", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.64736 -0.50017 -0.07078 0.41024 3.13609   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.522e+02 5.932e+02 0.425 0.67072   
## dem 8.554e-03 3.749e-03 2.282 0.02264 \*   
## river 8.329e-05 1.709e-05 4.873 1.22e-06 \*\*\*  
## mrvbf -3.902e-02 1.410e-02 -2.768 0.00572 \*\*   
## horB 9.121e-01 4.396e-02 20.745 < 2e-16 \*\*\*  
## horC 1.638e+00 5.363e-02 30.545 < 2e-16 \*\*\*  
## twi 3.344e-02 1.219e-02 2.743 0.00616 \*\*   
## vdchn -1.971e-02 7.541e-03 -2.614 0.00904 \*\*   
## water -1.323e-01 6.107e-02 -2.166 0.03049 \*   
## evim -8.418e-04 3.270e-04 -2.575 0.01013 \*   
## evisd -1.354e-03 7.499e-04 -1.805 0.07124 .   
## X -1.696e-04 1.072e-04 -1.582 0.11388   
## Y -1.343e-04 9.391e-05 -1.430 0.15279   
## XX1 4.870e+00 1.199e+00 4.063 5.11e-05 \*\*\*  
## XX2 6.053e-02 1.326e-02 4.564 5.45e-06 \*\*\*  
## XX3 6.881e-02 3.605e-02 1.909 0.05645 .   
## dem:river -2.641e-07 3.598e-08 -7.341 3.52e-13 \*\*\*  
## evim:evisd 3.249e-07 2.000e-07 1.625 0.10443   
## X:Y 3.341e-11 1.766e-11 1.892 0.05874 .   
## XX1:XX2 -7.746e-04 1.723e-04 -4.496 7.46e-06 \*\*\*  
## XX1:XX3 -8.630e-04 4.663e-04 -1.851 0.06442 .   
## XX2:XX3 -7.167e-07 2.611e-07 -2.745 0.00613 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7486 on 1457 degrees of freedom  
## (298 observations deleted due to missingness)  
## Multiple R-squared: 0.549, Adjusted R-squared: 0.5425   
## F-statistic: 84.46 on 21 and 1457 DF, p-value: < 2.2e-16

### OC

summary(lm(formula = OC ~ dem + river + mrvbf + slope + hor + evim +   
 evisd + X + XX1 + XX2 + XX3 + dem:river + evim:evisd + XX2:XX3,   
 data = p[p$hor == "A" | p$hor == "B" | p$hor == "C", ]))

##   
## Call:  
## lm(formula = OC ~ dem + river + mrvbf + slope + hor + evim +   
## evisd + X + XX1 + XX2 + XX3 + dem:river + evim:evisd + XX2:XX3,   
## data = p[p$hor == "A" | p$hor == "B" | p$hor == "C", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.5076 -0.1637 -0.0196 0.1344 2.6921   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.060e+00 5.642e+00 0.365 0.715016   
## dem 5.662e-03 1.513e-03 3.743 0.000189 \*\*\*  
## river 3.675e-06 9.759e-07 3.766 0.000172 \*\*\*  
## mrvbf 2.303e-02 7.361e-03 3.128 0.001793 \*\*   
## slope 1.395e-01 6.955e-02 2.006 0.045046 \*   
## horB -1.409e+00 2.096e-02 -67.201 < 2e-16 \*\*\*  
## horC -1.764e+00 2.567e-02 -68.741 < 2e-16 \*\*\*  
## evim 2.337e-04 1.529e-04 1.529 0.126457   
## evisd 5.925e-04 3.520e-04 1.683 0.092519 .   
## X 1.929e-06 4.402e-07 4.382 1.26e-05 \*\*\*  
## XX1 -1.375e-01 5.684e-02 -2.419 0.015693 \*   
## XX2 -5.070e-04 2.560e-04 -1.980 0.047893 \*   
## XX3 -7.298e-04 3.641e-04 -2.004 0.045211 \*   
## dem:river -4.524e-08 1.466e-08 -3.085 0.002071 \*\*   
## evim:evisd -1.326e-07 9.356e-08 -1.417 0.156745   
## XX2:XX3 2.133e-07 1.115e-07 1.913 0.055896 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3556 on 1446 degrees of freedom  
## (315 observations deleted due to missingness)  
## Multiple R-squared: 0.8208, Adjusted R-squared: 0.819   
## F-statistic: 441.6 on 15 and 1446 DF, p-value: < 2.2e-16

### CEC

summary(lm(formula = CEC ~ dem + river + wdist + mrvbf + hor + lstm +   
 evim + X + Y + XX1 + XX3 + dem:river, data = p[p$hor == "A" |   
 p$hor == "B" | p$hor == "C", ]))

##   
## Call:  
## lm(formula = CEC ~ dem + river + wdist + mrvbf + hor + lstm +   
## evim + X + Y + XX1 + XX3 + dem:river, data = p[p$hor == "A" |   
## p$hor == "B" | p$hor == "C", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -17.9402 -3.1300 -0.3089 2.3086 27.1458   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.355e+03 7.830e+02 -3.008 0.00268 \*\*   
## dem 1.320e-01 1.822e-02 7.246 6.94e-13 \*\*\*  
## river 2.253e-04 9.799e-05 2.299 0.02166 \*   
## wdist -1.148e-03 3.978e-04 -2.886 0.00396 \*\*   
## mrvbf 2.024e-01 7.771e-02 2.605 0.00928 \*\*   
## horB 6.158e+00 2.730e-01 22.556 < 2e-16 \*\*\*  
## horC 5.465e-01 3.344e-01 1.634 0.10240   
## lstm 2.356e+00 5.858e-01 4.021 6.09e-05 \*\*\*  
## evim -1.902e-03 4.792e-04 -3.969 7.58e-05 \*\*\*  
## X 1.651e-04 5.768e-05 2.862 0.00427 \*\*   
## Y 1.409e-04 7.090e-05 1.987 0.04711 \*   
## XX1 -1.539e+00 5.828e-01 -2.642 0.00834 \*\*   
## XX3 1.515e-03 6.900e-04 2.195 0.02833 \*   
## dem:river -1.228e-06 1.940e-07 -6.330 3.25e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.644 on 1457 degrees of freedom  
## (306 observations deleted due to missingness)  
## Multiple R-squared: 0.4756, Adjusted R-squared: 0.4709   
## F-statistic: 101.7 on 13 and 1457 DF, p-value: < 2.2e-16

### tb

summary(lm(formula = tb ~ dem + river + wdist + maxc + mrvbf + slope +   
 hor + twi + lstm + evim + X + Y + XX1 + XX2 + XX3 + dem:river +   
 XX1:XX2 + XX1:XX3 + XX2:XX3 + XX1:XX2:XX3, data = p[p$hor ==   
 "A" | p$hor == "B" | p$hor == "C", ]))

##   
## Call:  
## lm(formula = tb ~ dem + river + wdist + maxc + mrvbf + slope +   
## hor + twi + lstm + evim + X + Y + XX1 + XX2 + XX3 + dem:river +   
## XX1:XX2 + XX1:XX3 + XX2:XX3 + XX1:XX2:XX3, data = p[p$hor ==   
## "A" | p$hor == "B" | p$hor == "C", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -17.5108 -2.8060 -0.4576 2.3503 22.6252   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -7.155e+03 2.519e+03 -2.840 0.00459 \*\*   
## dem 1.694e-01 2.091e-02 8.100 1.31e-15 \*\*\*  
## river 2.528e-04 9.736e-05 2.596 0.00954 \*\*   
## wdist -8.056e-04 3.970e-04 -2.029 0.04266 \*   
## maxc -1.377e-02 8.381e-03 -1.643 0.10068   
## mrvbf 1.678e-01 1.002e-01 1.675 0.09429 .   
## slope 1.851e+00 9.373e-01 1.975 0.04854 \*   
## horB 6.781e+00 2.573e-01 26.358 < 2e-16 \*\*\*  
## horC 2.636e+00 3.702e-01 7.121 1.82e-12 \*\*\*  
## twi 1.868e-01 7.929e-02 2.355 0.01867 \*   
## lstm 2.981e+00 6.493e-01 4.592 4.85e-06 \*\*\*  
## evim -1.464e-03 5.081e-04 -2.882 0.00402 \*\*   
## X 1.736e-04 5.734e-05 3.028 0.00252 \*\*   
## Y 1.381e-04 7.083e-05 1.949 0.05147 .   
## XX1 5.816e+01 3.091e+01 1.882 0.06010 .   
## XX2 1.514e+00 7.342e-01 2.062 0.03940 \*   
## XX3 1.988e+00 1.089e+00 1.825 0.06827 .   
## dem:river -1.481e-06 1.928e-07 -7.679 3.25e-14 \*\*\*  
## XX1:XX2 -1.977e-02 9.561e-03 -2.068 0.03890 \*   
## XX1:XX3 -2.591e-02 1.418e-02 -1.827 0.06790 .   
## XX2:XX3 -6.498e-04 3.385e-04 -1.920 0.05512 .   
## XX1:XX2:XX3 8.481e-06 4.409e-06 1.924 0.05462 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.192 on 1228 degrees of freedom  
## (527 observations deleted due to missingness)  
## Multiple R-squared: 0.5372, Adjusted R-squared: 0.5293   
## F-statistic: 67.89 on 21 and 1228 DF, p-value: < 2.2e-16

### Clay

summary(lm(formula = clay ~ dem + wdist + hor + twi + lstm + lstsd +   
 evim + evisd + X + Y + XX1 + XX2 + XX3 + lstm:lstsd + XX1:XX2 +   
 XX1:XX3 + XX2:XX3 + XX1:XX2:XX3, data = p[p$hor == "A" |   
 p$hor == "B" | p$hor == "C", ]))

##   
## Call:  
## lm(formula = clay ~ dem + wdist + hor + twi + lstm + lstsd +   
## evim + evisd + X + Y + XX1 + XX2 + XX3 + lstm:lstsd + XX1:XX2 +   
## XX1:XX3 + XX2:XX3 + XX1:XX2:XX3, data = p[p$hor == "A" |   
## p$hor == "B" | p$hor == "C", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -21.563 -4.069 -0.353 3.752 44.494   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.567e+03 6.080e+03 0.751 0.4526   
## dem 1.735e-01 2.525e-02 6.871 9.49e-12 \*\*\*  
## wdist -9.809e-04 5.738e-04 -1.709 0.0876 .   
## horB 1.289e+01 3.818e-01 33.761 < 2e-16 \*\*\*  
## horC -5.366e+00 4.746e-01 -11.305 < 2e-16 \*\*\*  
## twi 1.737e-01 9.841e-02 1.765 0.0777 .   
## lstm -3.551e+01 1.604e+01 -2.214 0.0270 \*   
## lstsd -1.431e+03 6.123e+02 -2.338 0.0195 \*   
## evim -1.304e-03 8.105e-04 -1.609 0.1078   
## evisd 2.016e-03 1.270e-03 1.587 0.1127   
## X 1.356e-04 9.580e-06 14.153 < 2e-16 \*\*\*  
## Y 6.784e-05 9.965e-06 6.808 1.46e-11 \*\*\*  
## XX1 6.281e+01 4.137e+01 1.518 0.1292   
## XX2 2.005e+00 9.978e-01 2.009 0.0447 \*   
## XX3 2.472e+00 1.451e+00 1.704 0.0886 .   
## lstm:lstsd 4.824e+00 2.063e+00 2.338 0.0195 \*   
## XX1:XX2 -2.617e-02 1.299e-02 -2.014 0.0442 \*   
## XX1:XX3 -3.221e-02 1.888e-02 -1.706 0.0883 .   
## XX2:XX3 -9.961e-04 4.592e-04 -2.169 0.0302 \*   
## XX1:XX2:XX3 1.300e-05 5.980e-06 2.173 0.0299 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.473 on 1426 degrees of freedom  
## (331 observations deleted due to missingness)  
## Multiple R-squared: 0.6611, Adjusted R-squared: 0.6566   
## F-statistic: 146.4 on 19 and 1426 DF, p-value: < 2.2e-16

### Silt

summary(lm(formula = silt20 ~ dem + river + hor + twi + vdchn + lstm +   
 lstsd + evim + evisd + X + Y + XX1 + XX2 + XX3 + dem:river +   
 lstm:lstsd + evim:evisd + X:Y + XX1:XX2 + XX1:XX3 + XX2:XX3 +   
 XX1:XX2:XX3, data = p[p$hor == "A" | p$hor == "B" | p$hor ==   
 "C", ]))

##   
## Call:  
## lm(formula = silt20 ~ dem + river + hor + twi + vdchn + lstm +   
## lstsd + evim + evisd + X + Y + XX1 + XX2 + XX3 + dem:river +   
## lstm:lstsd + evim:evisd + X:Y + XX1:XX2 + XX1:XX3 + XX2:XX3 +   
## XX1:XX2:XX3, data = p[p$hor == "A" | p$hor == "B" | p$hor ==   
## "C", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -18.0633 -3.2621 -0.2327 2.7372 24.4128   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -9.655e+03 5.521e+03 -1.749 0.080566 .   
## dem 2.627e-01 2.539e-02 10.347 < 2e-16 \*\*\*  
## river 4.465e-04 1.147e-04 3.892 0.000104 \*\*\*  
## horB -4.966e+00 2.873e-01 -17.287 < 2e-16 \*\*\*  
## horC -1.150e+00 3.565e-01 -3.226 0.001282 \*\*   
## twi 2.606e-01 7.794e-02 3.344 0.000847 \*\*\*  
## vdchn -2.183e-01 4.900e-02 -4.456 9.01e-06 \*\*\*  
## lstm -2.585e+01 1.278e+01 -2.023 0.043292 \*   
## lstsd -9.320e+02 4.886e+02 -1.907 0.056671 .   
## evim 4.054e-03 2.162e-03 1.875 0.060942 .   
## evisd 8.053e-03 4.966e-03 1.622 0.105078   
## X 1.925e-03 7.300e-04 2.637 0.008455 \*\*   
## Y 1.785e-03 6.398e-04 2.790 0.005337 \*\*   
## XX1 6.423e+01 3.130e+01 2.052 0.040359 \*   
## XX2 1.893e+00 7.559e-01 2.504 0.012397 \*   
## XX3 2.261e+00 1.096e+00 2.062 0.039396 \*   
## dem:river -3.093e-06 2.520e-07 -12.276 < 2e-16 \*\*\*  
## lstm:lstsd 3.139e+00 1.646e+00 1.907 0.056749 .   
## evim:evisd -2.129e-06 1.319e-06 -1.614 0.106830   
## X:Y -2.712e-10 1.197e-10 -2.266 0.023627 \*   
## XX1:XX2 -2.480e-02 9.844e-03 -2.519 0.011868 \*   
## XX1:XX3 -2.965e-02 1.427e-02 -2.077 0.037944 \*   
## XX2:XX3 -8.987e-04 3.475e-04 -2.586 0.009803 \*\*   
## XX1:XX2:XX3 1.178e-05 4.526e-06 2.601 0.009379 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.857 on 1419 degrees of freedom  
## (334 observations deleted due to missingness)  
## Multiple R-squared: 0.4847, Adjusted R-squared: 0.4764   
## F-statistic: 58.04 on 23 and 1419 DF, p-value: < 2.2e-16

### Sand

summary(lm(formula = sand.mf ~ dem + river + wdist + maxc + mrvbf + slope +   
 hor + vdchn + lstsd + evisd + X + Y + XX1 + XX2 + XX3 + dem:river +   
 X:Y + XX1:XX2 + XX1:XX3 + XX2:XX3 + XX1:XX2:XX3, data = p[p$hor ==   
 "A" | p$hor == "B" | p$hor == "C", ]))

##   
## Call:  
## lm(formula = sand.mf ~ dem + river + wdist + maxc + mrvbf + slope +   
## hor + vdchn + lstsd + evisd + X + Y + XX1 + XX2 + XX3 + dem:river +   
## X:Y + XX1:XX2 + XX1:XX3 + XX2:XX3 + XX1:XX2:XX3, data = p[p$hor ==   
## "A" | p$hor == "B" | p$hor == "C", ])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -21.2734 -2.7658 -0.0026 2.6845 26.1474   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.295e+04 4.789e+03 8.967 < 2e-16 \*\*\*  
## dem -3.167e-01 2.800e-02 -11.312 < 2e-16 \*\*\*  
## river -9.309e-04 1.183e-04 -7.869 7.05e-15 \*\*\*  
## wdist 1.752e-03 4.504e-04 3.889 0.000105 \*\*\*  
## maxc 2.250e-02 1.015e-02 2.218 0.026714 \*   
## mrvbf -3.707e-01 1.132e-01 -3.275 0.001084 \*\*   
## slope -1.993e+00 1.083e+00 -1.841 0.065895 .   
## horB -1.522e+00 3.043e-01 -5.004 6.32e-07 \*\*\*  
## horC 2.799e+00 3.781e-01 7.403 2.28e-13 \*\*\*  
## vdchn 8.390e-02 5.537e-02 1.515 0.129932   
## lstsd 4.322e+00 8.927e-01 4.841 1.43e-06 \*\*\*  
## evisd -1.327e-03 9.120e-04 -1.455 0.145923   
## X -4.962e-03 7.508e-04 -6.609 5.45e-11 \*\*\*  
## Y -4.567e-03 6.577e-04 -6.944 5.77e-12 \*\*\*  
## XX1 -1.506e+02 3.284e+01 -4.585 4.93e-06 \*\*\*  
## XX2 -3.979e+00 7.943e-01 -5.009 6.15e-07 \*\*\*  
## XX3 -4.964e+00 1.150e+00 -4.319 1.68e-05 \*\*\*  
## dem:river 4.749e-06 2.648e-07 17.934 < 2e-16 \*\*\*  
## X:Y 7.119e-10 1.238e-10 5.751 1.08e-08 \*\*\*  
## XX1:XX2 5.184e-02 1.034e-02 5.012 6.07e-07 \*\*\*  
## XX1:XX3 6.465e-02 1.496e-02 4.320 1.67e-05 \*\*\*  
## XX2:XX3 1.725e-03 3.645e-04 4.733 2.43e-06 \*\*\*  
## XX1:XX2:XX3 -2.247e-05 4.748e-06 -4.734 2.42e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.153 on 1423 degrees of freedom  
## (331 observations deleted due to missingness)  
## Multiple R-squared: 0.7325, Adjusted R-squared: 0.7283   
## F-statistic: 177.1 on 22 and 1423 DF, p-value: < 2.2e-16